

DAS GRUPPENDELPHI

EINE DISKURSIVE METHODE ZUR ERMITTLUNG VON EXPERT*INNENURTEILEN

Von der Fakultät Wirtschafts-und Sozialwissenschaften der Universität Stuttgart zur Erlangung der Würde eines Doktors der Wirtschafts-und Sozialwissenschaften (Dr. rer. pol.)
genehmigte Abhandlung

Vorgelegt von
Rainer Kuhn
aus Prüm

Erstgutachter:
Prof. Dr. Dr. h.c. Ortwin Renn

Zweitgutachter:
Prof. Dr. André Bächtiger

Tag der mündlichen Prüfung: 28. Januar 2020

Institut für Sozialwissenschaften der Universität Stuttgart
2020

Zusammenfassung

Das Gruppendelphi stellt eine innovative und diskursive Methode dar, die auf Grundlage des klassischen Delphiverfahrens weiterentwickelt wurde. Die Methode wird zur Erhebung von Expert*innenurteilen eingesetzt, um eine Ermittlung von Handlungsempfehlungen, eine Einschätzung zu (politischen) Leitbildern und Rahmenbedingungen sowie eine Bewertung von konkreten Maßnahmen zu erzielen. Diese Methode bildet den Untersuchungsgegenstand dieser Arbeit. Da sich ein Gruppendelphi im Vergleich zum klassischen Delphiverfahren durch die Integration von diskursiven Elementen auszeichnet, wird besonders die Frage danach verfolgt, ob die Qualitätsmerkmale handlungsleitender Diskurse von der Methode des Gruppendelphis eingelöst werden können. Um diesen methodischen Ansatz zu legitimieren, sind sowohl eine theoretische Grundlage als auch ein Bemessungskriterium für die Qualität der Ergebnisse dieser Methode notwendig, die mittels der Theorieklasse mentaler Modelle, dem Linsenmodell aus der Social Judgment Theory sowie der Theorie des kommunikativen Handelns geschaffen werden. Das daraus resultierende Testdesign besteht aus vier Dimensionen: (1) Evidenz, (2) Verständlichkeit, (3) normativen Angemessenheit und (4) grundlegende Funktionslogik. Dieses Testdesign dient der Überprüfung der Fragestellung anhand von fünf Fallbeispielen einer konkreten empirischen Anwendung der Methode innerhalb verschiedener Forschungsprojekte. So soll diese Arbeit nicht nur eine theoretische Fundierung des Gruppendelphis erzielen und eine Überprüfung der Funktionsweise der Methode ermöglichen, sondern eine Einordnung möglicher Einsatzfelder vornehmen und zu einer verbesserten praktischen Anwendung beitragen.

Abstract

A group Delphi is an innovative and discursive method, which was developed as a variation of the conventional Delphi technique. The method is designed to consolidate expert opinions and judgments, to achieve recommendations for action, assessments of (policy) statements and framework programs as well as recommendations of concrete actions. The main focus of this work lies on the group Delphi method. Compared to the conventional Delphi technique a group Delphi integrates discursive elements. This integration leads to the question whether the method of group Delphi provides a framework inside the quality criteria of action-guiding discourses. To legitimize the methodological approach of a group Delphi, there is a need for assessment criteria regarding the quality of process results, which are based on a theoretical approach developed in this work. This approach is based on the theory of mental models, the lens model in the social judgment theory as well as the theory of communicative action. In the following step the test design is divided in four dimensions: (1) evidence, (2) comprehensibility, (3) normative appropriateness, and (4) basic functional logic. This test design will be applied on five empirical cases of the group Delphi method, which were carried out in respective research projects. The aim is not only to deliver a theoretical framework for the method and to test its functionality but also to define future fields of application as well as to improve the application of the process.

Inhalt

1	Einleitung.....	5
2	Eponym: Die Mythe des Orakels von Delphi	8
3	Das klassische Delphiverfahren	10
3.1.	EINSATZFELDER EINES DELPHIS	10
3.2.	TYPISIERUNG DES DELPHIVERFAHRENS.....	11
3.3.	DEFINITION DES VERFAHRENS	13
3.4	MERKMALE UND ABLAUF DES DELPHIVERFAHRENS	14
4	Das Gruppendelphi.....	18
4.1	STRUKTUR EINES GRUPPENDELPHIS.....	19
4.1.1	<i>Ablauf eines Gruppendelphis</i>	<i>19</i>
4.1.2	<i>Variationen innerhalb des Gruppendelphis.....</i>	<i>21</i>
4.1.3	<i>Abbruchkriterium</i>	<i>25</i>
4.1.4	<i>Anzahl der Teilnehmenden.....</i>	<i>27</i>
4.1.5	<i>Aufgabe der Anonymität.....</i>	<i>28</i>
4.2	FRAGEBOGENKONSTRUKTION.....	29
4.2.1	<i>Einsetzbare Skalentypen eines Delphifragebogens.....</i>	<i>30</i>
4.2.2	<i>Integration weiterer methodischer Elemente</i>	<i>32</i>
4.2.3	<i>Offene Abfragen.....</i>	<i>34</i>
4.3	VORTEILE DER METHODE.....	35
4.4	BEGRIFF UND AUSWAHL DER EXPERT*INNEN	36
5	Wissenschaftstheoretische Grundlagen.....	39
5.1	THEORIEKLASSE DER MENTALEN MODELLE UND SOCIAL JUDGMENT THEORY	41
5.1.1	<i>Wissensbestände und Urteilsbildung</i>	<i>41</i>
5.1.2	<i>Überführung globaler Wissensbestände in rationale Urteile.....</i>	<i>45</i>
5.1.3	<i>Urteilsänderung und Annäherung der Urteile.....</i>	<i>47</i>
5.2	ERWEITERUNG DES THEORETISCHEN RAHMENS DURCH DIE THEORIE DES KOMMUNIKATIVEN HANDELNS...	51
5.2.1	<i>Begriffsbestimmung: Diskurs und Konsens</i>	<i>51</i>
5.2.2	<i>Rationalität und Wissen</i>	<i>55</i>
5.2.2.1	<i>Kognitiv-instrumentelle und kommunikative Rationalität</i>	<i>57</i>
5.2.2.2	<i>Normenregulierte Handlungen und expressive Selbstdarstellung</i>	<i>58</i>
5.2.3	<i>Argumentationspraxis: Geltungsansprüche und deren Einlösung im Diskurs</i>	<i>59</i>
5.2.4	<i>Implikationen der diskurstheoretischen Erweiterung</i>	<i>64</i>
5.2.5	<i>Einordnung des Gruppendelphis als Diskurstypus.....</i>	<i>65</i>
6	Forschungsdesign.....	68
6.1	FRAGESTELLUNG	68
6.2	OPERATIONALISIERUNG.....	70
6.2.1	<i>Dimension: Evidenz</i>	<i>71</i>
6.2.2	<i>Dimension: Verständlichkeit.....</i>	<i>76</i>
6.2.3	<i>Dimension: normative Angemessenheit</i>	<i>78</i>
6.2.4	<i>Dimension: grundlegende Funktionslogik eines Delphiansatzes.....</i>	<i>81</i>

7	Empirische Fälle: Darstellung und Diskussion der Ergebnisse	83
7.1	SAUBER+	86
7.1.1	<i>Fragestellung des Gruppendelphis</i>	86
7.1.2	<i>Merkmale des Gruppendelphis</i>	87
7.1.3	<i>Analyse des Gruppendelphis in SAUBER+</i>	88
7.1.3.1	Dimension: Evidenz	89
7.1.3.2	Dimension: Verständlichkeit	99
7.1.3.3	Dimension: normative Angemessenheit	107
7.1.3.4	Dimension: grundlegende Funktionslogik	110
7.1.4	<i>Zwischenfazit: Gruppendelphi in SAUBER+</i>	112
7.2	SEAWIND	115
7.2.1	<i>Fragestellung des Gruppendelphis</i>	115
7.2.2	<i>Merkmale des Gruppendelphis</i>	116
7.2.3	<i>Analyse des Gruppendelphis in SEAWIND</i>	117
7.2.3.1	Dimension: Evidenz	117
7.2.3.2	Dimension: Verständlichkeit	128
7.2.3.3	Dimension: normative Angemessenheit	135
7.2.3.4	Dimension: grundlegende Funktionslogik	137
7.2.4	<i>Zwischenfazit: Gruppendelphi in SEAWIND</i>	140
7.3	SMART METERING	142
7.3.1	<i>Fragestellung des Gruppendelphis</i>	143
7.3.2	<i>Merkmale des Gruppendelphis</i>	144
7.3.3	<i>Analyse des Gruppendelphis zu Smart Metering</i>	145
7.3.3.1	Dimension: Evidenz	145
7.3.3.2	Dimension: Verständlichkeit	158
7.3.3.3	Dimension: normative Angemessenheit	161
7.3.3.4	Dimension: grundlegende Funktionslogik	165
7.3.4	<i>Zwischenfazit: Gruppendelphi zu Smart Metern</i>	168
7.4	ECO-FAHRASSISTENTEN	170
7.4.1	<i>Fragestellung des Gruppendelphis</i>	171
7.4.2	<i>Merkmale des Gruppendelphis</i>	171
7.4.3	<i>Analyse des Gruppendelphis zu Eco-Fahrerassistenten</i>	173
7.4.3.1	Dimension: Evidenz	173
7.4.3.2	Dimension: Verständlichkeit	180
7.4.3.3	Dimension: normative Angemessenheit	184
7.4.3.4	Dimension: grundlegende Funktionslogik	190
7.4.4	<i>Zwischenfazit: Gruppendelphi zu Eco-Fahrerassistenten</i>	194
7.5	PROSUITE	196
7.5.1	<i>Fragestellung des Gruppendelphis</i>	197
7.5.2	<i>Merkmale des Gruppendelphis</i>	199
7.5.3	<i>Analyse des Gruppendelphis in PROSUITE</i>	200
7.5.3.1	Dimension: Evidenz	200
7.5.3.2	Dimension: Verständlichkeit	208
7.5.3.3	Dimension: normative Angemessenheit	215
7.5.3.4	Dimension: grundlegende Funktionslogik	217
7.5.4	<i>Zwischenfazit: Gruppendelphi in PROSUITE</i>	222
8	Diskussion der Ergebnisse	224
9	Fazit	241

10 Literatur	247
11 Anhang.....	256
11.1 FRAGEBOGEN DER VORBEFRAGUNG DES GRUPPENDELPHIS IN SAUBER+	256
11.2 FRAGEBOGEN DES GRUPPENDELPHIS IN SAUBER+	265
11.3 STATISTISCHE ERGEBNISSE DES GRUPPENDELPHIS IN SAUBER+	274
11.4 AUSGEWERTETE AUSSAGEN: NORMATIVE DIMENSION IN SAUBER+	279
11.5 FRAGEBOGEN DER VORBEFRAGUNG DES GRUPPENDELPHIS ZU SEAWIND	282
11.6 FRAGEBOGEN DES GRUPPENDELPHIS ZU SEAWIND	297
11.7 STATISTISCHE ERGEBNISSE DES GRUPPENDELPHIS IN SEAWIND	306
11.8 AUSGEWERTETE AUSSAGEN: NORMATIVE DIMENSION IN SEAWIND	317
11.9 FRAGEBOGEN DES GRUPPENDELPHIS ZU SMART METERING	318
11.10 STATISTISCHE ERGEBNISSE DES GRUPPENDELPHIS ZU SMART METERING	334
11.11 AUSGEWERTETE AUSSAGEN: NORMATIVE DIMENSION (SMART METER)	343
11.12 FRAGEBOGEN DES GRUPPENDELPHIS ZU ECO-FAHRASSISTENTEN.....	346
11.13 STATISTISCHE ERGEBNISSE DES GRUPPENDELPHIS ZU ECO-FAHRASSISTENTEN.....	361
11.14 AUSGEWERTETE AUSSAGEN: NORMATIVE DIMENSION (ECO-FAHRASSISTENTEN)	372
11.15 FRAGEBOGEN DES GRUPPENDELPHIS ZU PROSUITE	377
11.16 STATISTISCHE ERGEBNISSE DES GRUPPENDELPHIS IN PROSUITE	384
11.17 AUSGEWERTETE AUSSAGEN: NORMATIVE DIMENSION (PROSUITE).....	390
11.18 ERKLÄRUNG ZUR EIGENSTÄNDIGKEIT DIESER ARBEIT.....	391

„Anstelle des Sprechens ‚mit fremden Zungen‘, der Offenbarung des göttlichen Worts, des Flüsterns der Dämonen, tritt die (...) Sach- und Fachsprache“ (Hitzler 1994: 15).

1 Einleitung

Das Gruppendelphi stellt eine innovative und diskursive Methode dar, die zur Erhebung von Expert*innen- und Stakeholderurteilen eingesetzt wird. Entstanden als eine Modifikation des klassischen Delphiverfahrens bestehen die mit dem Einsatz der Methode verbundenen Ziele in der Entwicklung von wissensbasierten Interpretationen, der Ermittlung von Handlungsempfehlungen, der Bewertung (politischer) Leitbilder und Rahmenbedingungen sowie der Bewertung von konkreten Maßnahmen (vgl. Kuhn et al. 2014: 62; Niederberger/Renn 2018: 1).¹

Der Begriff *Delphi*, der sowohl in der Bezeichnung des klassischen Verfahrens sowie in seiner Modifikation enthalten ist, ist nicht zufällig ein Homonym zur antiken griechischen Pilgerstätte Delphi. Die Mythe des Orakels von Delphi wurde namensgebend im Jahre 1948 von der *RAND Corporation*² für das damals neu entwickelte Verfahren aufgegriffen. Geprägt wurde der Name *Delphi* durch den Philosophen Joseph Kaplan, einem Mitglied der RAND Corporation. Das Design des Verfahrens sollte dabei vorwiegend militärischen Zwecken dienlich sein, erstmals getestet wurde es allerdings mittels einer Prognose des möglichen Siegers eines Pferderennens³ (vgl. Helmer 1963; Woudenberg 1991; Bowles 1999). In seinem ursprünglichen Design ging es beim klassischen Delphiverfahren folglich besonders darum, Sachverhalte zu klären, bei denen unsicheres bzw. unvollständiges Wissen vorliegt. *„Die Delphimethode ist ein vergleichsweise stark strukturierter Gruppenkommunikationsprozess, in dessen Verlauf Sachverhalte, über die naturgemäß unsicheres und unvollständiges Wissen existiert, von Experten beurteilt werden“* (Häder/Häder 1995: 12).

In den späten 1980er Jahren wurde eine Modifikation des klassischen Verfahrens entwickelt, die den Untersuchungsgegenstand dieser Arbeit bildet. Kern dieser Modifikation namens Gruppendelphi ist die Durchführung eines Workshops, bei dem die rein schriftliche Befragung durch die Elemente der Kleingruppendiskussion sowie moderierten Plenarsitzungen ergänzt wird. Durch die Integration von diskursiven Elementen soll besonders ein Nachteil des klassischen Delphiverfahrens behoben werden, da die diskursiven Elemente die Möglichkeit bieten, Argumente für die jeweiligen Urteile auszutauschen und zu erfassen.

¹ Die Bezeichnung Methode wird in dieser Arbeit bewusst zur Kategorisierung des Gruppendelphis verwendet um eine semantische und konzeptionelle Trennung zwischen den Begriffen *Methode* und *Verfahren* einzuhalten. Da ein Gruppendelphi sich aus mehreren Elementen (u.a. Befragung, Kleingruppendiskussion und moderiertes Plenum) zusammensetzt, folgt der Autor der Begriffsbestimmung von Niederberger: *„Der Begriff ‚Methoden‘ meint hier empirische Verfahren, die systematisch und ergebnisorientiert geplant und durchgeführt werden. Diese Methoden können sich aus einzelnen Erhebungsinstrumenten oder der Kombination verschiedener Verfahren zusammensetzen“* (ebd. 2015a: 33).

² Ein Projekt der U.S. Air Force.

³ Ob es sich bei den Vorhersagen um ein Hunde- oder ein Pferderennen gehandelt hat, ist nicht eindeutig überliefert. So beschreibt Häder (2009: 15), dass sich die Angaben in der Literatur widersprechen. Dies ist vermutlich dem Umstand geschuldet, dass keine Publikation dieser ersten Anwendung eines Delphis oder über die Ergebnisse des besagten Rennens erfolgte (vgl. Helmer 1963: 5). Bis auf Seeger (vgl. 1979: 57) sprechen allerdings alle gesichteten Publikationen von einem Pferderennen, darunter auch frühe Publikationen der RAND Corporation, welche auch für die Konzeption und Durchführung des besagten Delphis zuständig war (vgl. u.a. Helmer 1963: 5; Woudenberg 1991: 132).

Um den Untersuchungsgegenstand einzugrenzen, werden in einem ersten Schritt sowohl das Eponym des Orakels (Kapitel 2) und ebenfalls die begriffliche Bestimmung des klassischen Ansatzes mit seinen spezifischen Merkmalen beschrieben (Kapitel 3). Da sich durch die Integration der diskursiven Elemente einige Strukturmerkmale des Gruppendelphis von denen des klassischen Verfahrens unterscheiden, werden sowohl die Merkmale der Methode des Gruppendelphis näher beschrieben (Kapitel 4.1), wie auch der Fragebogen als eines der zentralen Instrumente näher erläutert, denn die Fragebogenkonstruktion und dessen Anwendung spielen bei einem Delphi eine zentrale Rolle (Kapitel 4.2). Modifikationen von Methoden folgen dem Ziel einer Verbesserung des ursprünglichen Ansatzes, einer neuen Zielstellung und/oder einer effizienteren Anwendung. Auch das Gruppendelphi wurde folglich mit dem Ziel erdacht, „(...) *die positiven Aspekte beizubehalten und die negativen zu kompensieren*“ (Schulz/Renn 2009: 13). Welche Vorteile sich neben der Integration von Möglichkeitsräumen zur Argumentation versprochen werden, wird im Anschluss umfassend darlegt (Kapitel 4.3), bevor ein wesentlicher Punkt für die Konzeption und Durchführung eines Gruppendelphis näher beleuchtet wird. So kommt der Auswahl der Expert*innen eine zentrale Bedeutung zu, die bei einer Definition der Rolle *Expert*in* beginnt und mit deren personellem Einfluss auf die Ergebnisse des Prozesses endet (Kapitel 4.4).

Um einen methodischen Ansatz zu legitimieren, sind sowohl eine theoretische Grundlage als auch ein Bemessungskriterium für die Qualität der Ergebnisse dieser Methode notwendig. Da das klassische Delphi und das Gruppendelphi eine gemeinsame Funktionslogik teilen, kann eine gemeinsame theoretische Grundlage der Prozesse aufgrund der Theorieklasse mentaler Modelle und des Linsenmodells aus der Social Judgment Theory erfolgen. Dabei geht es nicht nur um eine theoretische Bestimmung *wie* die Teilnehmenden zu einem Urteil gelangen, sondern auch *wie* eine im Delhiprozess angestrebte Annäherung sowie überdies eine Verbesserung der Urteile erfolgen kann und welche Grundlage es für eine mögliche Änderung individueller Urteile gibt (Kapitel 5.1).

Da das Gruppendelphi diskursive Elemente enthält, ist eine Erweiterung des theoretischen Grundkonzeptes des klassischen Ansatzes notwendig. So werden die *psychologischen Erklärungsmodelle* um die *Theorie des kommunikativen Handelns* ergänzt, um insbesondere eine theoretische Fundierung der diskursiven Prozesse in den Kleingruppendiskussionen und moderierten Plenarsitzungen während eines Gruppendelphis vorzunehmen (Kapitel 5.2). Dabei geht es neben einer Begriffsbestimmung von *Diskurs* und *Konsens* insbesondere darum, *Geltungsansprüche* als Bemessungsgrundlage für rationale Argumentation und Wissen einzuführen und auf die Logik des Gruppendelphis zu übertragen. Diese Geltungsansprüche und die vorhergehenden theoretischen Konzeptionen bilden die Grundlage für das Untersuchungsdesign dieser Arbeit, das sich in die Dimensionen der *Evidenz*, *Verständlichkeit*, *normativen Angemessenheit* und *grundlegenden Funktionslogik* des Gruppendelphiansatzes aufspannt.

Die zentrale Fragestellung dieser Arbeit folgt dieser Bemessungsgrundlage und stellt die Frage: *Können die Qualitätsmerkmale handlungsleitender Diskurse von der Methode des Gruppendelphis eingelöst werden?* Dafür muss das Gruppendelphi einen Rahmen für die Einlösung der Geltungsansprüche innerhalb eines handlungsleitenden Diskurses bieten, der es zulässt, dass die Gültigkeit der einzelnen Argumente in Frage gestellt werden kann, eine semantische Klärung von Begriffen und Definitionen stattfinden kann, die Teilnehmenden die

Möglichkeit haben ihre normativen Einstellungen und Werturteile offen zu legen und letztlich eine Verständigung über jeweilige Relevanzsysteme erfolgen kann (Kapitel 6).

Die Dimensionen des Untersuchungsdesigns werden in dieser Arbeit auf fünf empirische Fälle von Gruppendelphis analytisch angewandt. Daneben werden sowohl die spezifische Fragestellung wie auch die konkreten Merkmale des jeweiligen Falles erläutert. Folgende fünf Fälle stehen aus vier Forschungsprojekten für eine tiefergehende Analyse zur Verfügung:

- *SAUBER+ - Innovative Konzepte und Technologien für die separate Behandlung von Abwasser aus Einrichtungen des Gesundheitswesens* (Kapitel 7.1)
- *SEAWIND – Sound exposure and risk assessment of wireless network devices* (Kapitel 7.2)
- *Soziale, öffentliche, ökonomische und politische Akzeptanz von Fördermaßnahmen zur Verbesserung der Effizienz im Bereich Elektrizitätsnutzung und von nachhaltiger Mobilität bei privaten Haushalten:*
 - *Teilprojekt Elektrizitätsnutzung: Smart Metering* (Kapitel 7.3)
 - *Teilprojekt nachhaltige Mobilität: Eco-Fahrassistenten* (Kapitel 7.4)
- *PROSUITE – Development and application of a standardized methodology for the prospective sustainability assessment of technologies* (Kapitel 7.5).

Nachdem für jeden Fall ein Zwischenfazit gezogen wurde, werden die Erkenntnisse aus der Analyse der empirischen Fälle vergleichend diskutiert und generelle Schlüsse abgeleitet (Kapitel 8), bevor letztlich im Fazit dieser Arbeit nicht nur die zentrale Frage beantwortet werden wird, ob *die Qualitätsmerkmale von handlungsleitenden Diskursen von der Methode des Gruppendelphis eingelöst werden können*, sondern überdies, welche weiteren Implikationen sich aus der durchgeführten Analyse ergeben (Kapitel 9). Letztlich soll diese Arbeit nicht nur eine theoretische Fundierung des Gruppendelphis und eine Überprüfung der Funktionsweise der Methode erbringen, sondern eine Einordnung möglicher Einsatzfelder vornehmen und zu einer verbesserten praktischen Anwendung beitragen.

Bevor aber die methodischen Besonderheiten des klassischen Delphiverfahrens und des Gruppendelphis fokussiert betrachtet werden, erfolgt eine kurze Beschreibung des Eponyms und dessen Ursprungs in der griechischen Mythologie, da diese einen ersten Anknüpfungspunkt an die Ziele des Delphiverfahrens erlauben.

2 Eponym: Die Mythe des Orakels von Delphi

Bereits im 8. Jahrhundert vor Christus stand in Delphi der Tempel des Orakels, in den sich Apollon auf seiner Suche nach einer Heimstätte niederließ. Der Mythe nach musste Apollon erst den Drachen Python töten, um das Heiligtum der Mutter- und Erdgöttin Gaia einzunehmen. Apollon gab sich danach den Beinamen Phythios, seine Priesterinnen wurden Phythia genannt. Diese Priesterinnen rekrutierten sich meist aus jungen Frauen aus dem Dorf Delphi. Wenn jemand auf der Suche nach einem Rat war, begab diese Person sich zu Phythia - also der zu diesem Zeitpunkt in der Rolle der Priesterin agierenden Frau - um eine Frage an sie zu richten und eine Opfergabe zu bringen. Nach einem Tag der Wartezeit bekam man die meist in Versen gesprochenen zweideutigen Ratschläge der Pythia mitgeteilt. Die Blütezeit des Orakels von Delphi lag im 6. und 5. Jahrhundert vor Christus. Das Orakel von Delphi hatte zu dieser Zeit im Vergleich zu anderen Orakelstätten einen besonderen Einfluss (vgl. Giebel 2001: 10; Häder 2009: 13). *„Die besondere Rolle von Delphi im Vergleich zu anderen Orakelstätten beruhte auf seiner Einflussnahme auf politische und gesellschaftliche Verhältnisse in Griechenland: Sein Rat war gefragt, wenn sich eine Stadt eine neue Verfassung gab, Kultur-reformen vornahm oder eine Neugründung plante (...)“* (Giebel 2001: 29).

Neben staatsrechtlichen Fragen wurde das Orakel auch zu Fragen der Kolonisation oder der Entscheidung über Krieg und Frieden konsultiert. Nach heutigem Stand der Forschung lässt sich dem Orakel eine bewusste politische Einflussnahme in der antiken Welt zuschreiben. So besteht die Hypothese, dass Pythia die Antwort auf wichtige Fragen bereits im Vorfeld festlegte und auf eine umfangreich geführte Dokumentation der Fragen und der Bittsteller zurückgreifen konnte. Die Priesterinnen verfügten zudem über ein hohes Bildungsniveau sowie ein weitreichendes Netzwerk, dessen Kontakte den Ansässigen im Tempel bereits frühzeitig Informationen zuspielten (vgl. Häder 2009: 14).

Das Eponym des Orakels von Delphi für moderne Delphibefragungen ist letztlich nicht unangebracht, denn heute *„(...) hat die empirische Sozialforschung Pythias Aufgabe übernommen. Sie entwickelte hierzu eine Methode, die nicht ohne Grund auf das antike Orakel anspielt. Denn wie im Apollontempel werden auch bei modernen Delphi-Studien Personen befragt, denen man einen besonderen Zugang zum Thema „Zukunft“ unterstellt“* (Brosi et al. 2003: 4). Auch Häder sieht gewisse Überschneidungen und Vergleichspunkte zwischen dem klassischen Delphiverfahren und dem namensgebenden Orakel:

- Die Vorhersagen des Orakels von Delphi waren nicht kostenlos, denn die Fragenstellenden mussten Opfergaben erbringen, um Antworten zu erhalten.
- Es musste eine gewisse Wartezeit eingeplant werden bis das Orakel die Antworten gab.
- Die Anfragen an das Orakel mussten schriftlich erfolgen.
- Dem Orakelspruch lagen Beratungen unter Expert*innen zugrunde.
- Die Antworten besaßen immer einen gewissen Interpretationsspielraum.
- Das Orakel lieferte Informationen, die Einfluss auf das politische Leben nehmen sollten (vgl. Häder 2009: 14f.).

Wenn es um Fragen der Zukunft geht, setzen Politik, Gesellschaft und Wissenschaft heute nicht mehr auf die nebulösen Worte eines Orakels. Allerdings liegen ähnlich dem antiken Orakelspruch auch moderne Entscheidungen über zukünftige Maßnahmen meist Beratungen und Prognosen von Expert*innen zugrunde. So braucht es hier strukturierende Methoden und Verfahren, die eine Hilfestellung bieten, um die wahrscheinlichsten Entwicklungen einzugrenzen. Das gilt besonders in Feldern, in denen sich die Modelle und Prognosen der Expert*innen widersprechen oder Themenfeldern, in denen die Generierung von Wissen noch am Anfang steht. Einen solchen Ansatz möchte das Verfahren des Delphis bieten, dessen Merkmale im folgenden Kapitel beschrieben werden.

3 Das klassische Delphiverfahren

In seinem ursprünglichen Design geht es bei einem Delphi darum, Erkenntnisse zu Sachverhalten zu erlangen, bei denen unsicheres oder unvollständiges Wissen zu Grunde liegt (vgl. Häder/Häder 1995: 12). Dieses unsichere und unvollständige Wissen beruht meist auf der Zielstellung, die den Rahmen bestimmt, in dem das Verfahren zur Anwendung kommt. So soll ein Delphi häufig Erkenntnisse zu zukünftigen Sachverhalten generieren. Besonders in der BRD finden Delphis folglich oftmals im Rahmen von Foresight-Studien statt (vgl. Cuhls et al. 1998; Cuhls/Möhrle 2005; von Oertzen et al. 2006; Cuhls/Kimpeler 2008). Gerade bei diesem Anwendungsfeld findet sich häufig ein Hinweis auf das Orakel von Delphi. Einer prognostischen Zielstellung diente auch der Einsatz des Delphiverfahrens durch die *RAND Corporation* in den 50er und 60er Jahren des letzten Jahrhunderts, die das Verfahren im Umfeld militärischer Fragestellungen entwickelte. So ging es bei einer frühen Anwendung beispielsweise darum, mögliche Ziele von atomaren Angriffen durch die Sowjetunion auf dem Territorium der USA vorherzusagen, um sich auf diese mögliche Bedrohung besser vorbereiten zu können (vgl. Bowles 1999: 32). Dabei sollten die eingebundenen Experten, die Sichtweise eines sowjetischen Strategen einnehmen. „*The experiment was designed to apply expert opinion to the selection, from the viewpoint of a Soviet strategic planner, of an optimal U.S. industrial target system and the estimation of the number of A-bombs required to reduce the munitions output by a prescribed amount*“ (Dalkey/Helmer 1962: 1).

3.1 Einsatzfelder eines Delphis

Neben dem Bericht von Dalkey und Helmer (vgl. ebd. 1962) benennen später weitere Berichte der *RAND Corporation* „*forecasting of technological events*“ (Dalkey 1969: 15) oder „*scientific and technological forecasting*“ (Sackman 1974: 3) als Einsatzfelder für Delphis. Ziel dieser Studien war es, eine langfristige Vorhersage der technologischen und wissenschaftlichen Entwicklungen in einer Zeitspanne von zehn bis fünfzig Jahren einzugrenzen und vorherzusagen. Diese Untersuchungen führten dazu, dass Delphis auch in der wissenschaftlichen Öffentlichkeit Bekanntheit erlangten (vgl. Häder 2009: 15). Seit 1971 finden in Japan regelmäßig Delphistudien zur Entwicklung von Wissenschaft und Technologie statt, die vom *National Institute of Science and Technology Policy (NISTEP)* durchgeführt werden. Die letzte Erhebung fand 2009 bereits zum neunten Mal statt und bestand immer aus mindestens zwei Befragungsrunden (vgl. NISTEP 2010: 3).

Der Einsatz des Delphiverfahrens beschränkt sich aber nicht nur auf Foresight-Studien und besonders in der sozialwissenschaftlichen Forschung werden dem Delphiverfahren eine Reihe weiterer Anwendungsmöglichkeiten zugesprochen (vgl. Häder/Häder 1995: 12). So werden Delphis eingesetzt um eine Ideenaggregation, eine Bestimmung diverser Sachverhalte bzw. eine Bewertung bestimmter Maßnahmen sowie auch eine Ermittlung von Expert*innenmeinungen und/oder eine Konsensbildung zu betreiben. So existieren damit vielfältige Vorstellun-

gen über die Form und den Einsatz sowie unzählige empirische Einsatzgebiete des Verfahrens. „*Besides technological event and trend forecasting, the Delphi process has been used to evaluate budgets, define policy options, expose hidden agendas, and assess the significance to past events*“ (Webler et al. 1991: 256).

Diese Einsatzgebiete basieren meist auf den jeweiligen Erwartungen der Forscher*innen an die Methode. So liegt für Häder der Gedanke nahe, „*(...) das Instrument Delphi-Befragungen nicht länger als einen kompakten Universalansatz zu betrachten, sondern es gezielter für bestimmte Ziele auszuarbeiten beziehungsweise zu modifizieren und dann entsprechend auf das jeweilige Ziel ausgerichtet einzusetzen*“ (Häder 2009:30). Ein solch gezielter Einsatz sollte dabei einem gewissen Ordnungsschema folgen, wie in der folgenden Typisierung dargestellt wird.

3.2 Typisierung des Delphiverfahrens

Nach Häder gewinnen besonders vier Typen des klassischen Delphiverfahrens ein methodisches Profil, wenn man als Kriterium die Verknüpfung des Einsatzes des Verfahrens mit einem jeweilig verfolgten Ziel ansetzt (vgl. Häder 2009: 24-30).

Als erste Klassifizierung werden *Delphi-Befragungen zur Ideenaggregation* typisiert (vgl. Häder 2009: 31). Bei diesem Typus geht es prioritär darum Ideen zu aggregieren. Hier werden vor allem Erfahrungen und Expertise der Teilnehmenden dazu genutzt, erste Problemlösungsvorschläge zu elaborieren. Anders als bei der ursprünglichen Variante des Verfahrens wird hier kein quantitativer, sondern ein qualitativer Ansatz verfolgt. Ein Delphi dieses Typs ist dann besonders erfolgreich, wenn viele Ideen generiert werden. Dabei ist es nicht wichtig, ob diesen konsensual zugestimmt wird. Die eingebundenen Teilnehmenden sollten in ihrer Zusammensetzung eine breite Expertise abbilden (vgl. Häder 2009: 31f).

Den zweiten Typus bilden *Delphi-Befragungen für eine möglichst exakte Vorhersage eines unsicheren Sachverhalts bzw. für dessen genaue(re) Bestimmung* (vgl. Häder 2009: 32). Wie beim ersten Einsatz des Delphiverfahrens durch die *RAND Corporation* (vgl. Dalkey/Helmer 1962, Helmer 1963) geht es bei diesem Typus darum einen unsicheren Sachverhalt zu schärfen und zu klären. Das bereits oben erwähnte „*forecasting of technological events*“ (Dalkey 1969: 15) spielt hier eine besondere Rolle. Bei diesem Typus ist die Zielgröße eine möglichst genaue Prognose der Zukunft. Nach Häder liegt hier „*(...) stets ein „wahrer Wert“ vor, der mit dem Verfahren ermittelt werden soll und an dem der Erfolg des Einsatzes schließlich gemessen werden kann*“ (Häder 2009: 32).

Beim dritten Typus von *Delphi-Befragungen zur Ermittlung und Qualifikation der Ansichten einer Expertengruppe über einen diffusen Sachverhalt* geht es um die Erhebung und Qualifizierung einer Einschätzung bzw. einer Meinung einer klar bestimmbar Expert*innengruppe (vgl. Häder 2009: 33). Angestrebt wird hier eine Verbesserung der Urteile der Expert*innen und die methodisch einwandfreie Abbildung der jeweiligen Ansichten. Vordergründiges Ziel ist

nicht eine bloße Vorhersage der Zukunft, sondern auch die Ermöglichung einer aktiven Gestaltung dieser.⁴ Hier zeigt sich der Unterschied zum Forecasting der Nachkriegszeit, in dem die möglichst exakte Determination der Zukunft im Detail vordergründig verfolgt wurde und die mit dem Delphiverfahren des vorhergehenden zweiten Typs weiterhin verfolgt wird (vgl. Cuhls 2000: 7, Häder 2009: 33).

Den vierten und letzten Typus bilden nach Häder *Delphibefragungen zur Konsensbildung unter den Teilnehmern* (vgl. Häder 2009: 35). Während eines Delphiverfahrens werden durch das Feedback Gruppenprozesse ausgelöst, die dazu dienen das Ergebnis des Delphis zu qualifizieren. Bei diesem Typus wird ein großes Gewicht daraufgelegt, dass das Verfahren in besonderem Maße zu hohen Konsensbewertungen führen kann. Dabei wird das Verfahren als effizient eingestuft, wenn ein Konsens zwischen ausgewählten Expert*innen zu verschiedenen Themen erzielt werden kann. Neben Expert*innen können bei diesem Typ ebenso Interessengruppen sowie Teilnehmende aus der allgemeinen Öffentlichkeit eingebunden werden (vgl. Häder 2009:34, Juwana et al. 2010: 1642).

Tabelle 1: Typen von Delphibefragungen

	Typus 1	Typus 2	Typus 3	Typus 4
Bezeichnung	Ideenaggregation	Bestimmung eines Sachverhaltes	Ermittlung von Expert*innenmeinungen	Konsens
Ziel	Sammlung von Ideen zur Lösung eines Problems	Verbesserung der Bestimmung eines Sachverhaltes (Vorhersagen)	Ermittlung und Qualifikation der Ansichten von Expert*innen	Hohes Maß an Übereinstimmung bei den Teilnehmer*innen
Methodische Ausrichtung	qualitativ	qualitativ und quantitativ	qualitativ und (vor allem) quantitativ	quantitativ
Auswahl der Teilnehmer*innen	Auswahl der Expert*innen erfolgt aufgrund der Expertise	Hypothesen zur Auffindung der Expert*innen nötig, keine formalisierbaren Regeln	Totalerhebung oder bewusste Auswahl der Expert*innen	Auswahl der Teilnehmer*innen kann aufgrund eines bestimmbareren Rahmens erfolgen

Quelle: Eigene Darstellung in Anlehnung an Häder (2009: 36)

Da die Typologie nach Häder (vgl. 2009: 30-37) auch eine zeitliche Dimension beinhaltet und die Entwicklung und hinzukommenden Einsatzfelder der Methode widerspiegelt, lässt sie eine gewisse Trennschärfe vermissen. Besonders bei Betrachtung der Typen 2 und 3 zeigt sich, dass die Anforderungen und Ziele, die mit dem Delphiverfahren verbunden werden, mit dem Einsatzfeld korrelieren. Das Beispiel des Einsatzes im Bereich Foresight zeigt, dass die Universalismethode zur Prognose einen Wandel zu einer Komponente beim Einsatz eines Methodenmixes vollzogen hat, der sich durch eine klare Zielgruppe und Zielbestimmung auszeichnet (vgl. Cuhls 2000: 14). Auch Häder sieht hier ein Spannungsfeld zwischen den verschiedenen Typen und spricht von einem Dilemma falls unterschiedliche Ziele mit dem Einsatz eines Delphis verfolgt werden, wenn also beispielsweise sowohl Konsens als auch neue Ideen am Ende des Verfahrens stehen sollen. Dieses Dilemma kann nur dann aufgelöst werden, wenn bei der

⁴ Nach Cuhls ist allen neueren Konzepten des Forecastings gemein, dass nicht bloß eine Methode zum Einsatz kommen sollte und kommunikative Prozesse angestoßen werden sollten, die unterschiedliche Akteure einbeziehen. Delphiverfahren sind bei Cuhls ein Teil eines sinnvollen Einsatzes von „Vorausschau-Verfahren“ (vgl. ebd. 2000:7-14).

Konzeption eines Delphiverfahrens eine differenzierte Betrachtung der Ziele erfolgt und Prioritäten festgelegt werden. Die oben dargestellte Typologie kann hier nach Häder einen Ansatzpunkt liefern (vgl. ebd. 2009: 37). Generell sieht allerdings auch Häder einen möglichen universellen Einsatz eines Delphiverfahrens: „*Delphi-Befragungen sind dazu in der Lage, gleichzeitig auch alle beziehungsweise mehrere Ziele zu verfolgen, also beispielsweise sowohl Konsens unter den Teilnehmern zu schaffen als auch einen unklaren Sachverhalt aufzuklären*“ (Häder 2009: 35). Dabei sollte aber einem generellen Grundsatz gefolgt werden: „*Perhaps the most important advice to be given to the potential users is to suit the method to the problem, not the problem to the method*“ (Lindstone 1985: 627).

3.3 Definition des Verfahrens

Unter dem Label *Delphi* sind bereits Verfahren unterschiedlicher Dauer, Reichweite, Intensität und verschiedener Formen angewandt worden. „*Die Adaptierbarkeit wird einerseits als Vorteil der Methode angesehen, die daraus resultierende Variationsbreite hat aber auch zur Unschärfe des Profils dieser Methodik geführt (...)*“ (Ammon 2009: 460). Der Umstand, dass Delphiverfahren mit ganz unterschiedlichen Zielen und Designs initiiert und durchgeführt werden, führt auch dazu, dass es bis heute keine einheitliche Definition von Delphiverfahren gibt (vgl. Häder 2009: 19; Schulz/Renn 2009: 11). Dies gilt weiterhin bezüglich einer fehlenden einheitlichen Bezeichnung. So wird beispielsweise von *Delphi-Befragungen* (vgl. Häder 2009), *Delphi-Verfahren* (vgl. Renn/Webler 1998: 78), der *Delphi-Methode* (vgl. Vorgrimler/Wübben 2003), der *Delphi-Erhebung* (vgl. Brosi et al. 2003), der *Delphi technique* (vgl. Bowles 1999), der *Delphi procedure* (vgl. Quade 1970: 16) oder aber vom *Delphi Assessment* (vgl. Sackman 1974) gesprochen. Neben der Vielfalt an Begriffen führen die verschiedenen Zielstellungen, die mit dem Delphiverfahren verbunden werden, zu ebenso vielfältigen Definitionen, die sich in der publizierten Literatur finden und damit eben eine Ausgangslage dafür bilden, warum bis dato keine einheitliche Definition besteht. Allerdings findet sich in der vorhandenen Literatur eine Vielzahl an Begriffsbestimmungen und Arbeitsdefinitionen, wie folgend dargestellt wird.

Die Bestimmung von spezifischen Sachverhalten zeigt sich bei Goodman, wenn sie davon spricht ein Delphi sei „*(...) a survey approach which would be very useful for initiating discussion on particular issue*“ (Goodman 1987: 733). In der Definition von Häder und Häder ist die Bestimmung eines unsicheren Sachverhaltes ein zentrales Charakteristikum: „*Die Delphimethode ist ein vergleichsweise stark strukturierter Gruppenkommunikationsprozess, in dessen Verlauf Sachverhalte, über die naturgemäß unsicheres und unvollständiges Wissen existiert, von Experten beurteilt werden*“ (Häder/Häder 1995: 12). Eine ähnliche inhaltliche Ausrichtung findet sich auch bei Brosi et al., hier wird auch der Komponente der Bestimmung einer zukünftigen Entwicklung mehr Gewicht verliehen: „*Charakteristisch für Delphi-Studien ist die Durchführung mehrerer Erhebungsrunden, um die Rückkoppelungen der Zwischenergebnisse an die in die Befragung eingebundenen Experten zu ermöglichen und durch Nachjustierung der Expertenurteile die Bandbreiten der wahrscheinlichsten Entwicklungen einzugrenzen*“ (Brosi et al. 2003: 4). Eine ebenso klare Verortung des Delphiverfahrens im Bereich der *Zukunftsforschung* nehmen von Oertzen et al. mit ihrer Definition vor: „*Die Grundidee einer Delphi-Befragung liegt darin, Thesen zur Zukunft zu formulieren und sie ausgewählten Befragten in zwei (oder mehr) Wellen zur Beurteilung (...) vorzulegen*“ (von Oertzen et al. 2006: 16). Neben der

Bestimmung eines Sachverhaltes werden oft auch gruppendynamische Prozesse und die Gewinnung eines konsensualen Urteils herausgestellt: „*Ganz allgemein kann ein Delphi als ein Verfahren verstanden werden, bei dem in einem iterativen Prozess Expertenurteile zu einer bestimmten Fragestellung ermittelt werden, mit dem Ziel Konsens und Dissens in den Urteilen zu erfassen und zu begründen*“ (Schulz/Renn 2009: 11). Quade spricht dem Delphiverfahren einen Nutzen bei der Verfolgung mehrerer Ziele zu und benennt neben Forecasting ebenso Werturteile als Gegenstandsbereich des Verfahrens: „*In view of the accelerating adoption of Delphi procedures by such a wide spectrum of public and private institutions, two uses are of immediate practical concern: for forecasting technological and social events and for value judgments*“ (Quade 1970: 16). Auch Webler et al. sprechen dem Delphi Problemlösungskompetenzen bei der Frage nach sozialen Werten zu. Das Verfahren sei besonders dann geeignet, wenn eine Lösung basierend auf einer Vermischung aus wissenschaftlicher Erkenntnis und sozialen Werten angestrebt wird (vgl. Webler et al. 1991: 256).

Die Definitionen lassen zwei zentrale Charakteristika erkennen. Zum einen geht es bei einem Delphi darum, inhaltliche Aspekte zu beleuchten, indem das Verfahren Lösungsansätze im Umgang mit Unsicherheiten bieten kann. Zum zweiten ermöglicht das Verfahren des Delphis eine verbesserte Erfassung von Gruppenmeinungen, indem es die Kommunikation innerhalb der Gruppe der Befragten steuert und strukturiert (vgl. Vorgrimler/Wübben 2003: 764; Häder 2009: 19). Diese beiden Charakteristika spielen in der Definition von Linstone und Turoff eine zentrale Rolle: „*Delphi may be characterized as a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem*“ (Linstone/Turoff 1975: 3).

3.4 Merkmale und Ablauf des Delphiverfahrens

Wie im vorherigen Absatz erkennbar, legen die Definitionen des Delphiverfahrens häufig gänzlich unterschiedlich Schwerpunkte, die häufig mit dem jeweiligen Erkenntnisinteresse bzw. der Zielstellung des Einsatzes des Verfahrens verknüpft sind. Dies führt zur dargestellten Vielfalt an Definitionen und dem Fehlen einer allgemein akzeptierten Definition (vgl. Vorgrimler/Wübben 2003: 764; Häder 2009: 13; Schulz/Renn 2009: 11). Gerade deshalb sollte der Einsatz des Delphiverfahrens an bestimmte Voraussetzungen gekoppelt werden, die verhindern sollen, dass das Design beliebig ausgelegt und verändert wird. Häder sieht die Gefahr, dass ein Delphi als *bedenkenlos veränderbare Universalmethode* verstanden werden könnte, die diffuse Erwartungen weckt (vgl. Häder 2009: 23).

Auf diesen Umstand rekurrierend ist es wichtig, die Grundmerkmale eines Delphiverfahrens näher zu betrachten, da sich hieraus die spezifischen Zielsetzungen des Einsatzes der Methode ergeben. „*Wenn auch in der Literatur einstimmig zu entnehmen ist, dass es keine allgemein anerkannte Definition der Delphi Methode gibt, kann doch davon ausgegangen werden, dass über die Grundidee der Delphi-Methode Konsens besteht*“ (Vorgrimler/Wübben 2003: 764). Dieser Konsens hinsichtlich der Grundidee zeigt sich besonders bei den spezifischen Merkmalen, die dem Verfahren zu eigen sind. Bereits in den frühen Veröffentlichungen der *RAND Corporation* werden drei Merkmale genannt, die das Verfahren in ihrer Verknüpfung kennzeichnen: *Anonymität*, *kontrolliertes Feedback* und *statistische Ergebnisdarstellung* (vgl. Dalkey 1969: 16; Rescher 1969: 1; Sackman 1974: 4). Bei einem Delphiverfahren wird dabei

ein quantitativ formalisierter Fragebogen eingesetzt, welcher die statistische Ergebnisdarstellung ermöglicht. Die Anonymität der Einzelantworten wird hierbei garantiert, da diese den Teilnehmenden in Form eines kontrollierten Feedbacks als aggregierte Gruppenantwort zurückgespielt wird. Der Prozess von Befragung, Auswertung und Feedback wird bei einem Delphiverfahren mindestens einmal wiederholt, es finden also mehrere, mindestens aber zwei Wellen statt (vgl. Niederberger/Kuhn 2013: 58). Ein entscheidendes Merkmal ist damit die Durchführung von mindestens zwei Befragungsrunden, so dass das Delphiverfahren als *mehrstufiges Befragungsverfahren* mit Rückkopplung verstanden werden kann (vgl. Vorgrimler/Wübben 2003: 764).

Zusammenfassend lassen sich fünf zentrale Charakteristika des Delphiverfahrens herausstellen:

- *Gruppenergebnisse*: Produkt der Befragung sind quantitativ statistische Gruppenergebnisse
- *Anonymität*: Die Anonymität wird den Teilnehmenden versichert und durch ein Feedback mittels aggregierter Gruppenantworten sichergestellt.
- *Iterativer Prozess*: Beim Delphiverfahren handelt es sich um einen strukturierten Kommunikationsprozess.
- *Kontrolliertes Feedback*: Ablauf und Aufbau eines Delphiverfahrens erlauben ein kontrolliertes Feedback.
- *Mehrstufiges Befragungsverfahren*: Ein Delphiverfahren besteht aus mindestens zwei Befragungsrunden (vgl. Rowe et al., 1991: 237; Vorgrimler/Wübben 2003: 764f.; Häder 2009: 24f.; von der Gracht 2012: 1526; Niederberger/Kuhn 2013: 58).

Der mehrstufige Aufbau eines Delphis sowie die Rückkopplung der Ergebnisse gehören hierbei zu den besonders charakteristischen Merkmalen eines Delphiverfahrens, da sie bei den vielseitigen und vielzähligen Varianten dieser Methode einen durchgängigen Bestandteil bilden (vgl. Rowe et al. 1991: 237f.; Webler et al. 1991: 255; Häder 2009: 25f.). So heben auch Brosi et al. diese Merkmale als Charakteristikum hervor: *„Charakteristisch für Delphi-Studien ist die Durchführung mehrerer Erhebungsrunden, um die Rückkoppelungen der Zwischenergebnisse an die in die Befragung eingebundenen Experten zu ermöglichen und durch Nachjustierung der Expertenurteile die Bandbreiten der wahrscheinlichsten Entwicklungen einzugrenzen“* (Brosi et al. 2003: 4). Dieser Grundgedanke bildet ein strukturelles Fundament des Delphiverfahrens und zeichnet sich auch im konkreten Ablauf ab. *„Die Grundidee von Delphi besteht darin, in mehreren Wellen Expertenmeinungen zur Problemlösung zu nutzen und in diesen ein anonymes Feedback einzuschalten“* (Häder/Häder 2000: 13).

Die Mehrstufigkeit des Verfahrens verfolgt das Ziel, den Teilnehmenden die Möglichkeit zu bieten ihre Einschätzungen aufgrund einer Rückmeldung der vorhergehenden Runde zu ändern. Die Rückkopplung der Ergebnisse der vorherigen Runde an die Teilnehmenden soll eine Verortung der eigenen Einschätzung im Vergleich zum gesamten Sample bereitstellen und dabei eine kritische Auseinandersetzung mit dem eigenen vorhergehenden Urteil anregen (vgl. Niederberger/Kuhn 2013: 58). Gerade durch den iterativen Prozess und das eingebundene Feedback, wird von einem *Mehr* an Informationen ausgegangen, als es beispielsweise bei einem traditionellen Survey der Fall wäre: *„In addition to the richness issues of traditional surveys, Delphi studies inherently provide richer data because of their multiple iterations and their response due to feedback“* (Okoli/Pawlowski 2004: 21).

Auch wenn es viele Varianten des Delphiverfahrens gibt, so kann von einem allgemeinen Ablaufmodell ausgegangen werden (vgl. Schulz/Renn 2009: 12). Die beschriebenen Merkmale wie auch die Funktionsweise eines Delphiverfahrens werden umso deutlicher, wenn man den Ablauf des Verfahrens näher betrachtet. Dieser Ablauf besteht aus vier Hauptkomponenten, die überdies die Charakteristika dieses Verfahrens widerspiegeln:

Befragung: Am Anfang des Verfahrens steht die Entwicklung eines standardisierten Fragebogens, der die relevanten Sachverhalte, die mit der Zielvorstellung des Delphis verbunden sind, enthalten sollte. Dieser kann beispielsweise eine Abfrage von zu erwartbaren Konsequenzen einer Maßnahme, eine Entscheidungsoption oder alternative Wege in die Zukunft enthalten (vgl. Schulz/Renn 2009: 12; Häder 2009: 125). Die Zielgruppe eines Delphis bilden Expert*innen die häufig mittels eines quantitativ formalisierten Fragebogens befragt werden. Da die Auswertung nur die aggregierten Antworten beinhaltet, kann den Teilnehmenden Anonymität ihrer Einzelantworten garantiert werden (vgl. Niederberger/Kuhn 2013: 58). Diese Anonymität ist ein wesentlicher Bestandteil des klassischen Delphiverfahrens. Dies erfolgt mit dem Grundgedanken, dass es so den Expert*innen leichter fällt ein gegebenes Urteil bzw. eine gegebene Einschätzung in späteren Runden zu revidieren und eine Meinungsführerschaft verhindert werden kann (vgl. Dalkey 1969: 16; Häder 2009: 148).

Auswertung: Die Auswertung, der von den Expert*innen beantworteten Fragebögen, besteht aus der Ermittlung der Durchschnittswerte sowie der Extremwerte und Varianzen. Diese bilden die statistischen Maßzahlen, die ein Feedback an die Teilnehmenden ermöglichen, ohne die zugesicherte Anonymität aufzugeben. Es findet lediglich eine Auswertung der aggregierten Gruppenantworten statt (vgl. Sackman 1974: 7; Schulz/Renn 2009: 12; Niederberger/Kuhn 2013: 58).

Feedback: Die Rückkopplung der Ergebnisse der vorhergehenden Befragungsrunde an die Teilnehmenden erfolgt mittels einer Übersicht über die aggregierten Gruppenantworten. So können die Expert*innen einen Vergleich ihrer eigenen Bewertung/Einschätzung der einzelnen Sachverhalte zu denen der anderen Teilnehmenden ziehen. Ziel dieses Feedbacks ist es, dass sie diesen Vergleich in die Urteilsbildung des mit dem Feedback übermittelten nächsten Fragebogens einbeziehen. Das Feedback bildet so ein Korrektiv der eigenen Einschätzungen, auch anhand möglicher neu gewonnener Erkenntnisse (vgl. Häder 2009: 24; Schulz/Renn 2009: 12). Dadurch, dass das Forscherteam das Feedback kontrolliert an die Teilnehmenden zurückspielen kann, können auch äußere Einflüsse gemindert werden: „*Controlled feedback in the Delphi process is designed to reduce the effect of noise*“ (Hsu/Sandford 2007: 2).

Wiederholung: Die Wiederholung von Befragung, Auswertung und Feedback kennzeichnet eines der wichtigsten Merkmale im Ablauf eines Delphiverfahrens. Diese Wiederholung findet idealiter solange statt, bis keine Änderungen der Urteile der Expert*innen mehr auftreten. Im Ideal sollen die Ergebnisse eines Delphis damit Bewertungen hervorbringen, die in der Expert*innengruppe einen klaren Konsens markieren oder aber einen Konsens über einen Dissens (vgl. Häder/Häder 2000: 13; Schulz/Renn 2009: 13f.).

So lassen sich aus den beschriebenen Charakteristika sowie des Ablaufs einige Kennzeichen herausstellen, die in der Vielzahl der Variationen des Delphiverfahrens eine Orientierung bieten:

- Expert*innen bilden die Zielgruppe
- Einsatz eines formalisierten Fragebogens
- Anonymität der Einzelantworten wird garantiert
- Ermittlung aggregierter Gruppenantworten
- Feedback mittels statistischer Ergebnisse
- Wiederholung der Befragung bis zu einem festgelegten Abbruchkriterium
- Konsens innerhalb der eingebundenen Expert*innengruppe wird angestrebt (vgl. Schulz/Renn 2009: 13).

Die Funktionsweise eines Delphiverfahrens lässt sich damit wie folgt zusammenfassen: *„Im Idealfall sortiert das Delphi-Verfahren die Bewertungen aus, die innerhalb der Expertengruppe konsensfähig sind oder einen Dissens begründen. Durch die Anonymisierung der Teilnehmer und den iterativen Prozeß der Befragung kann der jeweilige Kenntnisstand ohne Rücksicht auf den Prestigewert eines jeden Teilnehmers am Delphi-Prozeß dargestellt werden“* (Renn/Webler 1998: 78). Das Delphiverfahren hat sich bei der Beantwortung prognostischer Fragen, bei denen es darum geht, notwendiges Wissen zusammenzutragen und Ambivalenz und Unsicherheiten bezüglich einzelner Fragen zu bestimmen, als geeignet gezeigt. Gerade bei komplexen Themen und bei vorliegendem Dissens bezüglich eines Themenkomplexes sind dialogische Verfahren der kognitiven Urteilsbildung gefragt. Unter diesen hat sich das Delphiverfahren besonders bewährt (vgl. Renn/Webler 1998: 77).

4 Das Gruppendelphi

In den 1980er Jahren wurde mit dem Gruppendelphi eine Modifikation des klassischen Delphiverfahrens vorgenommen, bei der die schriftliche Befragung der Expert*innen durch einen moderierten Workshop mit den Elementen der Kleingruppendiskussionen sowie Plenarsitzungen ersetzt wird. Dies erfolgte mit dem Ziel „(...) *die positiven Aspekte beizubehalten und die negativen zu kompensieren*“ (Schulz/Renn 2009: 13). So ist die Hauptmodifikation zum klassischen Verfahren „(...) *die Nutzung von persönlich-mündlichen Workshops zur Erhebung und zur Diskussion der Daten (...)*“ (Häder 2009: 25). Durch die Integration von diskursiven Elementen soll ein Nachteil des klassischen Delphiverfahrens behoben werden. So bietet der Diskurs die Möglichkeit auch inhaltliche Begründungen für die Urteile der Expert*innen auszutauschen und zu erfassen. Niederberger beschreibt das Gruppendelphi wie folgt: „*Das Gruppendelphi ist eine Art Workshop, bei dem in rotierenden Kleingruppen ein standardisierter Fragebogen von Experten beantwortet wird. Das Ziel ist in der Regel Konsens bzw. Konsens über Dissens über einen Wissensdiskurs zu erhalten*“ (Niederberger 2015a: 41). So kann ein Gruppendelphi sowohl konsensuale Bewertungen befördern, wie auch deutlich getrennte Positionen ausmachen (Konsens über Dissens). Im Falle eines Konsenses über einen Dissens wissen die Expert*innen nach dem Gruppendelphi, warum eine Seite für eine Maßnahme ist und die andere dagegen. Darüber hinaus wurden die jeweiligen Begründungen im Gespräch auf Stärken und Schwächen überprüft. So kann ein so erzielter Konsens über einen Dissens ausschließen, dass das Ergebnis auf Scheinkonflikten oder Fehlurteilen beruht, sondern auf klaren Differenzen bei der Bewertung der Folgen potentiell verfügbarer Maßnahmen. Für beide Fälle – ob Konsens oder Konsens über Dissens - gilt im Rahmen der Modifikation zur klassischen Methode, dass am Ende des Prozesses ausführliche Begründungen für die Positionen vorliegen (vgl. Webler et al. 1991: 256; Renn/Webler 1998: 78f.; Renn 2007: 180; Schulz/Renn 2009: 13).

Ein Gruppendelphi soll anhand seiner Struktur dazu beitragen, „(...) *dass günstige Konstellationen für einen intensiven Austausch von inhaltlichen Begründungen und Argumenten bereitgestellt werden. So kann durch einen inhaltlichen Austausch der Argumente die Gefahr gemindert werden, dass sich eine falsche Einschätzung durchsetzt, die sich an der bloßen statistischen Mehrheit orientiert*“ (Kuhn et al. 2014: 66f.). Das Ziel der Methode des Gruppendelphis ist somit, den Rahmen für einen (interdisziplinären) offenen und konstruktiven Diskurs zwischen Expert*innen bereitzustellen und idealiter eine gemeinsame Beurteilung wissenschaftlicher Fragestellungen zu erhalten (vgl. Niederberger 2015b: 117).

Durch die physische Anwesenheit der Teilnehmenden kann bei einem Gruppendelphi im Gegensatz zu dem klassischen Verfahren keine Anonymität in der Befragungssituation aufrechterhalten werden. Das Element des Fragebogens bleibt dem Gruppendelphi aber erhalten. Dieser strukturiert die Diskussion zwischen den Teilnehmenden (vgl. Renn/Webler 1998: 78; Renn et al. 1999: 99; Niederberger/Kuhn 2013: 59).

4.1 Struktur eines Gruppendelphis

Die Struktur eines Gruppendelphis orientiert sich an der des klassischen Verfahrens. Als strukturierendes Element wird auch hier ein quantitativ formalisierter Fragebogen eingesetzt. Durch den Rahmen eines Workshops und die Integration diskursiver Elemente ergeben sich für das Gruppendelphi Unterschiede im Vergleich zum Ablauf des klassischen Delphiverfahrens. Diese Unterschiede sollen in den folgenden Unterkapiteln näher beschrieben werden.

4.1.1 Ablauf eines Gruppendelphis

Wie auch das klassische Delphiverfahren folgt ein Gruppendelphi einem idealtypischen Ablauf, der eine Orientierung für die Durchführung bietet und hilft, die zentralen Charakteristika bei der konkreten Anwendung des Verfahrens umzusetzen.

Beginn des Workshops: Zu Beginn des Workshops werden den Teilnehmenden sowohl die Struktur des Fragebogens sowie auch der Ablauf des Gruppendelphis erläutert, da aufgrund der relativen Neuheit der Methode davon ausgegangen werden sollte, dass diese nicht der gesamten Gruppe der Teilnehmenden bekannt ist. Darauf werden die Teilnehmenden je nach Gesamtanzahl in drei bis vier Kleingruppen aufgeteilt. Dies erfolgt meist nicht systematisch, sondern per Zufall. Die Teilnehmendenanzahl der Kleingruppe sollte anhand von Erfahrungswerten dabei zwischen drei und sechs Personen liegen (vgl. Renn/Webler 1998: 79; Schulz/Renn 2009: 17; Niederberger/Kuhn 2013: 59).

Kleingruppenphase: Jede Kleingruppe erhält die Aufgabe, den standardisierten Fragebogen gemeinsam zu beantworten. Eine Person aus dem Kreis der Kleingruppe füllt den Fragebogen aus, während alle Teilnehmenden sich auf ein gemeinsames Votum einigen sollen. Wenn bei einzelnen Items ein gemeinsames Urteil nicht möglich ist, da ein Dissens zwischen den jeweiligen Einschätzungen vorliegt, können Minderheitsvoten im Fragebogen vermerkt werden. Zwar wird ein Konsens angestrebt, abweichende Voten sind aber durchaus möglich. Um ungestört diskutieren zu können, sollten die Kleingruppen einen eigenen Raum bereitgestellt bekommen. Weiter sollte sich das Team der Forscher*innen darauf beschränken, die Kleingruppen nur in gewissen Abständen an die noch verbleibende Zeit zu erinnern und keinen Einfluss auf die Diskussion zu nehmen, um externe Effekte zu vermeiden (vgl. Schulz/Renn 2009: 17; Niederberger/Kuhn 2013: 59).

Auswertung: Nach der ersten Kleingruppenphase erfolgt durch das Team der Forscher*innen eine unmittelbare Auswertung der Ergebnisse. Während dessen haben die teilnehmenden Expert*innen eine Pause. Hierzu werden, ähnlich wie bei einem klassischen Delphi, die Mittelwerte, die Abweichungen sowie der daraus gebildete Variationskoeffizient berechnet und für eine Präsentation in der Plenarsitzung vorbereitet. Wie auch bei dem klassischen Verfahren, ist diese Auswertung die Grundlage der Rückkopplung der Ergebnisse der ersten Befragungsrunde und bildet so ein Korrektiv der jeweiligen Einschätzungen der Expert*innen (vgl. Schulz/Renn 2009: 12-18).

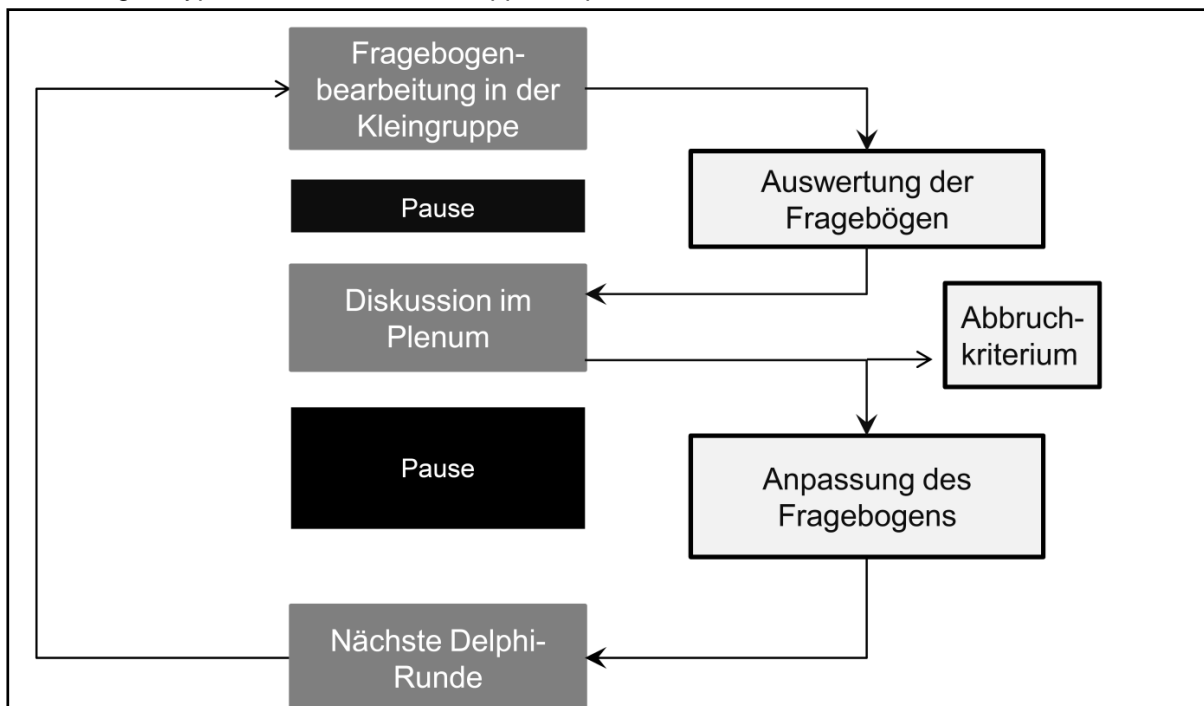
Plenarsitzung: In der Plenarsitzung soll die zur Verfügung stehende Zeit dazu genutzt werden, die Diskussion auf Themen zu lenken, bei denen eindeutig Dissense also die größten Diskrepanzen zwischen den Einschätzungen der Teilnehmenden vorliegen. Dabei müssen die Gruppen bzw. deren jeweilige Mitglieder, deren Einschätzungen am stärksten vom Mittelwert abweichen, ihre Standpunkte darlegen. „Ziel der Diskussion ist es herauszufinden, worin der Dissens begründet liegt und ob die Diskrepanzen durch Informationen und Argumente der anderen Experten aufzulösen sind“ (Renn/Webler 1998: 79). Auch kann es sein, dass bereits die Formulierung einer Frage für den Dissens verantwortlich ist, da diese beispielsweise semantisch nicht klar formuliert ist oder einen wichtigen Aspekt nicht beinhaltet.

In diesem Fall kann eine Anpassung des Fragebogens vorgenommen werden (vgl. Niederberger/Kuhn 2013: 63).

Anpassung des Fragebogens: Bereits während der Plenarsitzung werden bei Bedarf Formulierungen der Fragen angepasst. Es können aber auch neue Items formuliert werden, die in der zweiten Runde abgefragt werden. Um die Diskussion der Teilnehmenden in der zweiten Runde besser zu strukturieren, sollten im Fragebogen nur noch die Items enthalten sein, bei denen noch weiterer Diskussionsbedarf besteht. Es werden also lediglich die strittigen bzw. unklaren Fragen in den Fragebogen der folgenden Runde integriert (vgl. Schulz/Renn 2009: 18; Ruddat/Beninghaus 2009: 59; Niederberger/Kuhn 2013: 63).

Folgende Delphirunden: In der folgenden Kleingruppenphase sollten die Kleingruppen anders zusammengesetzt sein, als in der vorhergehenden. Dies kann durch eine systematische Permutation der Teilnehmenden sichergestellt werden. So werden auch die Repräsentant*innen der Extremgruppen auf andere Kleingruppen verteilt. Im Ideal wird die Abfolge von Kleingruppensitzung, Auswertung und Plenarsitzung solange wiederholt, bis keine signifikanten Verschiebungen der Urteile auftreten.⁵ Als Resultat erhält man am Ende eines Gruppendedelphis in der Regel eine eindeutige Verteilung der Beurteilung der einzelnen Items durch die Expert*innen. Entweder liegen die Einschätzungen der Gruppen nah um einen Mittelwert und es liegt damit eine konsensuale Bewertung des Sachverhaltes vor oder es bilden sich mehrgipflige Verteilungen, die auf getrennte Positionen bei der Bewertung hinweisen und einen Konsens über einen Dissens markieren sollten. Für beide Fälle liefert das Gruppendedelphi ausreichende Begründungen der Urteile der Expert*innen, die in der Plenarsitzung ausgetauscht und erfasst werden konnten (vgl. Renn/Webler 1998: 79; Schulz/Renn 2009: 18f.).

Abbildung 1: Typischer Ablauf eines Gruppendedelphis



Quelle: Kuhn et al. 2014: 66

Ähnlich wie beim klassischen Delphiverfahren können auch bei einem Gruppendedelphi verschiedene Varianten im Ablauf ausgemacht werden. „Der Ablauf des Gruppendedelphis ist nicht

⁵ Aufgrund von fehlender Zeit kann es auch dazu kommen, dass ein Gruppendedelphi früher zu Ende geht. Dies soll in Abschnitt 4.1.3 – Abbruchkriterium näher erläutert werden.

als strikte, unverrückbare Vorgabe zu verstehen, sondern vielmehr als Orientierung“ (Schulz/Renn 2009: 20). Allerdings bildet der in Abbildung 1 dargestellte typische Ablauf das Grundgerüst, auf dem die verschiedenen Variationen eines Gruppendelphis jeweils aufbauen. Diese und das ebenfalls in der Abbildung zu findende Abbruchkriterium sollen in den folgenden zwei Unterkapiteln näher beleuchtet werden.

4.1.2 Variationen innerhalb des Gruppendelphis

Zu einer der Variationen innerhalb eines Gruppendelphis gehört die zusätzliche Durchführung einer Befragungsrunde im Vorfeld des eigentlichen Workshops. Eine solche **Vorbefragung** folgt meist der Absicht einer Zeitersparnis sowie einen zusätzlichen thematischen Input für das eigentliche Gruppendelphi zu gewinnen. Dies empfiehlt sich besonders, wenn zu den abgefragten Sachverhalten des geplanten Gruppendelphis auf keine oder wenige theoretische Vorkenntnisse aufgebaut werden kann oder keine Erfahrungen bzw. Einschätzungen mittels anderer Methoden (beispielsweise Leitfadeninterviews oder Wertbaumanalysen) im Vorfeld erhoben werden konnten (vgl. Schulz/Wassermann 2010: 6; Niederberger/Kuhn 2013: 62).

Eine solche Vorbefragung ermöglicht den Teilnehmenden zudem „(...) *eine gute Vorbereitung auf den Workshop, lässt ihnen ausreichend Zeit, sich mit relevanten Fragestellungen inhaltlich auseinander zu setzen und zeigt auf der anderen Seite dem Forscherteam erste Einblicke in diskrepante Urteile*“ (Schulz/Renn 2009: 20). So erhalten die Teilnehmenden die Möglichkeit, sich im Vorfeld des Gruppendelphis nochmals vertieft mit dem abgefragten Thema auseinander zu setzen und den Wissenstand gegebenenfalls zu aktualisieren (vgl. Schulz/Wassermann 2010: 5f.). Ebenso soll die Vorbefragung Zeit einsparen, denn aufgrund dieser ersten Einblicke ist es für das Forscherteam möglich den Fragebogen auf die Punkte zu kürzen, bei denen die Diskrepanz zwischen den Urteilen der Expert*innen hoch ist. Die individuelle Vorbefragung bietet damit eine stärkere Fokussierung auf die diskrepanten Punkte und die Zielstellung diesen Fokus im eigentlichen Gruppendelphi weiter zu verfolgen (vgl. Schulz/Renn 2009: 20; Niederberger/Kuhn 2013: 62).

Die Durchführung einer Vorbefragung birgt allerdings auch Nachteile. So kann es passieren, dass die Befragten zwar die gleiche Antwort angeben, allerdings nicht auf einer gleichen kognitiven Grundlage sowie auf verschiedenen Relevanzsystemen basierend. Die Gefahr hier einen Scheinkonsens mittels schriftlich standardisierter Befragung zu produzieren, kann bei der Vorbefragung nicht gänzlich vermieden werden. Ebenfalls kann nicht ausgeschlossen werden, dass sich die Teilnehmenden auf ihre Einschätzung in der Vorbefragung festlegen und diese dann in Kleingruppe und Plenum unbedingt verteidigen wollen. Eine anonymisierte Auswertung und Präsentation der Ergebnisse dieser Befragung kann diesem Effekt allerdings entgegenwirken (vgl. Schulz/Renn 2009: 20f.).

Das Für und Wider einer Vorbefragung lässt sich auch in der praktischen Umsetzung von Gruppendelphis beobachten, in der häufig eine Vorbefragung mit dem Hintergrund der Zeitersparnis und Wissensgenerierung durchgeführt wird (vgl. Gallego Carrera 2009b: 80-83; Rudat/Beninghaus 2009: 58; Schulz et al. 2009: 47f.; Niederberger/Kuhn 2013: 62; Niederberger 2015b: 132), oder auf eine Vorbefragung verzichtet wird (vgl. Deuschle/Sonnberger 2009: 71; Kuhn et al. 2014: 63ff.). Die praktische Anwendung lässt weitere Argumente für beide Seiten erkennen. Dort wo ein Gruppendelphi als eine Methode im Rahmen eines Forschungsdesigns

neben weiteren Methoden zur Anwendung kommt, kann auf thematisches Vorwissen und/oder theoretische Vorkenntnisse zurückgegriffen werden. Hier wird auf eine Vorbefragung gerne verzichtet, da kein zusätzlicher Erkenntnisgewinn erwartet wird (vgl. Kuhn et al. 2014: 62ff.). Dies gilt ebenso für die Seite der Teilnehmenden, wenn diese über ausreichend Wissen zu aktuellen wissenschaftlichen Erkenntnissen und politischen Prozessen innerhalb des Themengebiets des Gruppendelphis verfügen (vgl. Schulz et al. 2009: 55). *„Die Erfahrungen aus diesem Delphi zeigen, dass eine vorab Versendung des Delphifragebogens nicht immer notwendig ist. Wenn in irgendeiner Form bereits eine Einbindung der Teilnehmer in den Planungs- und Entscheidungsprozess erfolgt ist, ist die Versendung eines Individualfragebogens nicht notwendig“* (Schulz et al. 2009: 54).

Da auch bei einem Gruppendelphi ein quantitativer Fragebogen eingesetzt wird, empfiehlt sich die Durchführung eines *Pretests*, um Verständlichkeit, Vollständigkeit und sachliche Richtigkeit zu überprüfen. Da bei einem Gruppendelphi Expert*innen eingebunden sind, ist hier besonders die sachliche Richtigkeit ein wichtiges Element. Denn Expert*innen erwarten Fragebögen in denen die Items nicht pauschal formuliert sind, disziplinären Sachverstand zeigen und wissenschaftliche Fundierung erkennen lassen (vgl. Schulz/Renn 2009: 113). Eine Vorbefragung macht auch dann Sinn, wenn ein explorativer Anspruch an diese gestellt wird. Hier können neben den quantitativen Items auch offene Antwortfelder (*Sonstiges*) integriert oder offene Fragen in den Fragebogen aufgenommen werden. So können für das Gruppendelphi am Workshoptag thematische Bereiche in den Fragebogen integriert werden, die nach Expert*inneneinschätzung im Hinblick auf die Forschungsfrage des Delphis relevant sind, im Vorfeld vom Forscher*innenteam aber nicht berücksichtigt wurden. *„Ob im Vorfeld des Workshops der Fragebogen mit der Bitte zum Ausfüllen verschickt wird, sollte von Fall zu Fall entschieden werden. Wenn sehr wenig Zeit für den Workshop eingeplant ist, kann die Versendung des Individualfragebogens helfen, die Diskussion stärker zu fokussieren und somit Zeit einzusparen“* (Schulz/Renn 2009: 21). Neben dem Argument der Zeiteinsparung - das im Forschungsalltag meist relevant ist - sollten auch die anderen Gründe des Für und Wider einer Vorbefragung abgewogen werden. In diese Abwägung sollten insbesondere Vorkenntnisse des Forscher*innenteams im thematischen Bereich des Delphis, wie auch die Einbindung in das methodische Konzept beachtet werden.

Bei der **Gruppenzusammensetzung** der Kleingruppen existieren zwei verschiedene Herangehensweisen. In der Regel erfolgt die Zusammensetzung der Kleingruppen zufällig. Diese kann beispielsweise durch Ziehen einer Nummer geschehen, anhand derer die Teilnehmenden anschließend auf die einzelnen Gruppen verteilt werden (vgl. Schulz/Renn 2009: 17; Niederberger/Kuhn 2013: 62). In wenigen Fällen erfolgt die Zusammensetzung der Kleingruppen hingegen durch bewusste Zuteilung durch das Forscherteam im Vorfeld. So soll zum einen verhindert werden, dass Expert*innen mit ähnlicher oder gleicher institutioneller Zugehörigkeit gemeinsam in einer Kleingruppe sitzen und diese im negativen Falle durch gemeinsame Argumentationsmuster dominieren könnten. Zum anderen können auch unterschiedliche Temperamente getrennt werden, um eine Verzerrung durch Dominanz einer Person zu verhindern (vgl. Deuschle/Sonnberger 2009: 71). Für eine Einschätzung des individuellen Temperamentes ist es allerdings notwendig die Teilnehmenden im Vorfeld des Gruppendelphis (persönlich) zu kennen, was aber in den seltensten Fällen für eine Einschätzung aller Expert*innen möglich

ist. Eine Trennung unter Beachtung der institutionellen Zugehörigkeiten der einzelnen Teilnehmenden ist allerdings in der Praxis häufig möglich und wird damit häufiger angewandt, allerdings selten dokumentiert.

Um negative Effekte wie eine Meinungsführerschaft über mehrere Runden hinweg zu vermeiden, empfiehlt sich generell eine Permutation des Teilnehmer*innenfeldes zu den verschiedenen Kleingruppenphasen (vgl. Niederberger/Kuhn 2013: 59). Auch kann durch eine Permutation der Teilnehmenden die Basis geschaffen werden, dass in der zweiten Runde die Repräsentant*innen der Extremgruppen aus der ersten (oder weiteren) Runden vertreten sind (vgl. Renn/Webler 1998: 79).

Da die Kleingruppen während der Beantwortung des Fragebogens ohne Moderation ungestört bleiben, ist es bei einem Gruppendelphi wichtig den Austausch der Begründungen für die jeweiligen Bewertungen während der Plenarsitzung zu dokumentieren. So finden sich Vorschläge die gesamte Diskussion mittels einer Videokamera aufzuzeichnen, um den Diskussionsverlauf in seiner Gänze zu dokumentieren (vgl. Renn/Webler 1998: 80). Ebenfalls gibt es Fälle von Gruppendelphis in denen Audioaufzeichnungen der Diskussion im Plenum angefertigt wurden (vgl. Deuschle/Sonnberger 2009: 70; Gallego Carrera 2009a: 103). Neben audio(-visuellen) Aufzeichnungen gibt es ebenso die Möglichkeit einer direkten Protokollierung der Diskussion (vgl. Schulz/Renn 2009: 115; Kuhn et al. 2014: 78). Es existieren allerdings auch Praxisbeispiele in denen ein Protokoll oder gar eine Aufzeichnung nicht erwähnt werden und/oder nicht stattgefunden haben. Da aber die Argumente für die jeweiligen Bewertungen der einzelnen Kleingruppen wichtige Anhaltspunkte für darauffolgende Bewertungen enthalten und die Diskussion nachvollziehbar machen können, ist von einem Verzicht einer **Dokumentation der Plenardiskussion** abzuraten.

Auch nach der Durchführung eines Gruppendelphis kann es bei Betrachtung einzelner praktischen Anwendungen zu Variationen kommen. Generell empfehlen Schulz und Renn in ihrer Beschreibung des idealtypischen Ablaufs eines Gruppendelphis die Erstellung eines belastbaren Protokolls durch das Forscher*innenteam, „(...) welches den Teilnehmern mit der Bitte um Freigabe oder ggfs. Korrektur zugeschickt wird. Dieses Protokoll sollte die zentralen statistischen Auswertungen, sprich die Angabe von Maßen der zentralen Tendenz und adäquate Streuungsmaße sowie ausführliche Darstellungen der Begründungen der Urteile beinhalten“ (ebd. 2009: 19). So sollen sich die zentralen Argumente und Interpretationen der Teilnehmenden in diesem Protokoll wiederfinden. Auf diesem Weg kann sichergestellt werden, dass die Ergebnisse des Gruppendelphis mitsamt den jeweiligen Begründungen und Argumente adäquat erhoben wurden und die Validität der erzielten Ergebnisse zusätzlich gesteigert wird. Ebenfalls kann so ein möglicher Scheinkonsens aufgedeckt werden. Falls einzelne Teilnehmende nicht die Zeit oder die Motivation für zusätzliche Argumentationen am Workshoptag aufbringen wollten, können sie diese als Rückmeldung zum Protokoll nachreichen und den Scheinkonsens damit aufdecken (vgl. Schulz/Renn 2009: 19; Schulz et al. 2009: 52; Niederberger/Kuhn 2013: 64). Die **Ausgestaltung des Feedbacks** variiert durchaus in der Praxis. So wurde beispielsweise bereits eine zusätzliche Erhebungsrunde mittels Individualfragebogen an das Ergebnisprotokoll gekoppelt, um die Einschätzungen der Expert*innen zusätzlich zu prüfen - ob also nach zwei Wochen keine signifikanten Änderungen in der Beurteilung durch die einzelnen Teilnehmenden vorgekommen sind (vgl. Niederberger/Kuhn 2013: 64). Auch möglich ist die Versendung einer Ergebniszusammenfassung mit Bitte um Zustimmung und falls gewünscht möglicher Korrekturanmerkungen durch die Teilnehmenden. Auch hier wird

den Expert*innen die Möglichkeit des Widerspruchs eingeräumt, „(...) so dass eventuelle Gruppeneffekte (jemand hat sich beispielsweise im Workshop der Gruppenmeinung gebeugt, ohne dies im Prozess bemerkbar zu machen) erkannt werden können“ (Kuhn et al. 2014: 78).

In der zweiten und den darauffolgenden Runden eines Gruppendelphis sollen nur noch die strittigen bzw. unklaren Fragen diskutiert werden, die aus der vorhergehenden Diskussion in der Plenarrunde übriggeblieben sind (vgl. Schulz/Renn 2009: 18). Hierfür wird in den meisten - besonders neueren - Anwendungen des Gruppendelphis eine **Modifikation des Fragebogens** für die nächste Delphirunde vorgenommen und ein angepasster auf die jeweiligen zu diskutierenden Fragen bzw. Items gekürzter Fragebogen an die Kleingruppen ausgehändigt (vgl. Deuschle/Sonnberger 2009: 75; Ruddat/Beninghaus 2009: 62; Niederberger/Kuhn 2013: 63).⁶ Diese Kürzungen sind auch beim klassischen Delphiverfahren ein oft angewandter Arbeitsschritt. So kann auch hier der Fragebogen „(...) schrittweise auf die von den Teilnehmern divergierend eingeschätzten Aspekte reduziert werden“ (Vorgrimmler/Wübben 2003: 766).

Es gibt allerdings auch Beispiele in denen eine Kürzung des Fragebogens, besonders aus Zeitgründen, nicht erfolgen konnte, wenngleich hier in der folgenden Kleingruppenphase ebenso ein Fokus auf die strittigen bzw. unklaren Fragen gelegt wurde (vgl. Webler et al. 1991: 257).⁷ Neben der Kürzung auf die strittigen Fragen kommt es in vielen Fällen auch vor, dass sich während der Plenarsitzung auf eine Umformulierung einzelner Items geeinigt wird. Diese Umformulierung wird mit den Teilnehmenden diskutiert, bis eine für alle befriedigende neue Formulierung in den neuen Fragebogen integriert werden kann. So kann während der Plenarsitzung bei ursprünglich unscharf formulierten Items eine präzise Formulierung gefunden werden, die von allen Expert*innen getragen wird (vgl. Schulz et al. 2009: 53; Ruddat/Beninghaus 2009: 62; Niederberger/Kuhn 2013: 63). Zusätzlich ist es möglich, während der Plenarsitzung die Modifikation mittels eines Beamers für alle sichtbar zu machen, so dass die Teilnehmenden Widerspruch einlegen können, wenn die Umformulierung nicht in ihrem Sinne ist (vgl. Deuschle/Sonnberger 2009: 75). In einzelnen Fällen wird der Fragebogen auch in seinem Aufbau grundlegend verändert, so dass beispielsweise Ratingskalen durch Rankingskalen ersetzt werden, um eine Priorisierung zu fokussieren (vgl. Ruddat/Beninghaus 2009: 62) oder es wird eine offene Frage in den Fragebogen integriert, um den Kontext der gesamten Einschätzungen zu definieren (vgl. Kuhn et al. 2014: 76).

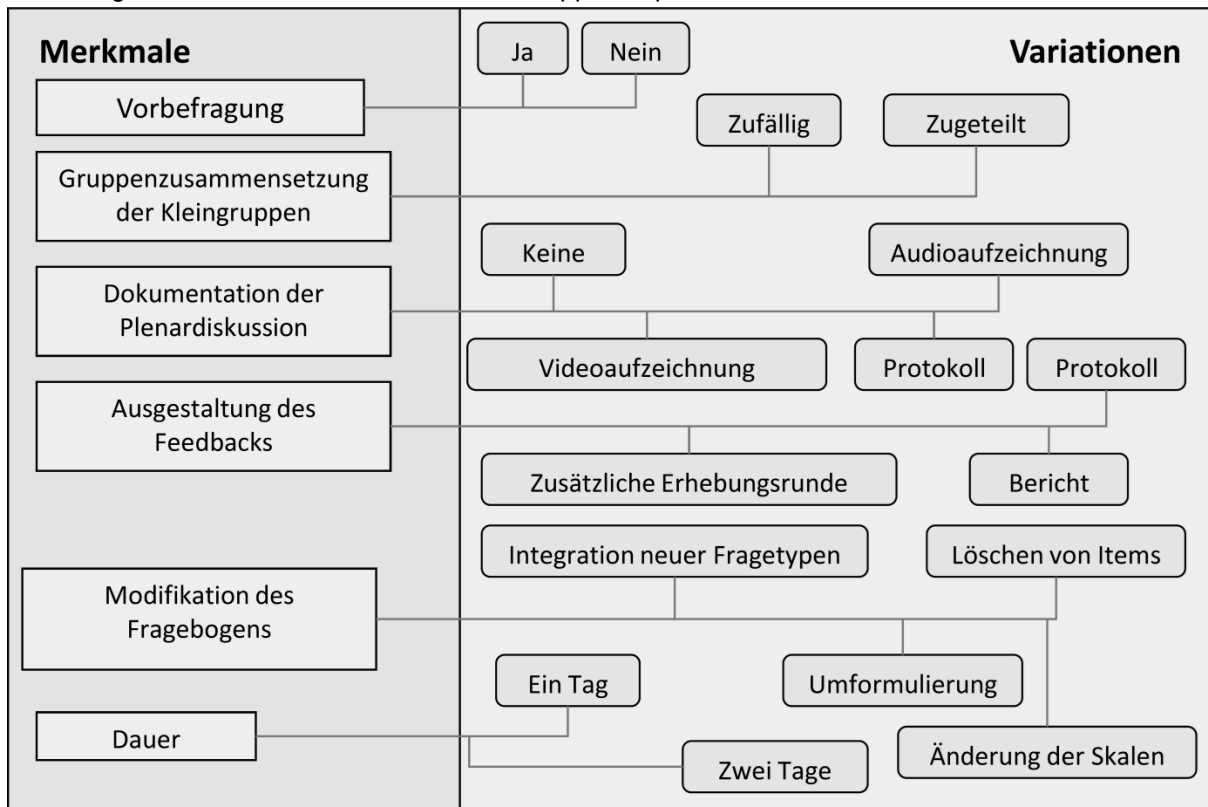
Die **Dauer des Gruppendelphis** - genauer die Dauer des eigentlichen Workshops - wird in der Literatur mit einer Zeitspanne von ein bis zwei Tagen angegeben (vgl. Renn/Webler 1998: 80; Schulz/Renn 2009: 20). Dieser Wert ist auch bei den bisher stattgefundenen praktischen Anwendungen der Methode zu beobachten, bei denen einige Beispiele an einem Tag durchgeführt wurden (vgl. Deuschle/Sonnberger 2009: 67; Ruddat/Beninghaus 2009: 58; Niederberger/Kuhn 2013: 62), sich aber ebenso auch Beispiele für einen zweitägigen Workshop finden (vgl. Gallego Carrera 2009a: 102; vgl. Gallego Carrera 2009b: 83; Schulz/Wassermann 2010: 6). Da das Teilnehmendenfeld bei einem Gruppendelphi aus Expert*innen bzw. aus Stakeholdern besteht, ist die Rekrutierung aufgrund von anderen Verpflichtungen der Teilnehmenden erschwert, so dass in der Forschungspraxis meist ein eintägiger Workshop Anwendung findet (vgl. Niederberger/Kuhn 2013: 62). Diese Erfahrungen wurden zum Beispiel auch

⁶ Der grundsätzliche Aufbau des Fragebogens wird in Kapitel 4.2 dieser Arbeit näher beschrieben.

⁷ Es ist davon auszugehen, dass mit technischen Neuerungen, wie stets verfügbare Notebooks und Drucker, auch die Möglichkeit des Druckes eines angepassten Fragebogens gestiegen ist. So kann der knappen Resource Zeit heute besser begegnet werden.

in einem Gruppendelphi des im 6. EU Forschungsrahmenprogramms geförderten Projektes *NEEDS*⁸ gemacht, bei dem es vielen der Teilnehmenden aufgrund ihrer sonstigen beruflichen Verpflichtungen nicht möglich war für zwei Tage an dem Gruppendelphi teilzunehmen. Aus diesem Grund empfiehlt Gallego Carrera, „(...) die zu investierende Zeit für Stakeholder möglichst gering zu halten und Workshops eher auf einen anstatt auf zwei Tage anzusetzen“ (ebd. 2009b: 92). Dies kann allerdings zur Abweichung von der Grundidee eines Gruppendelphis sowie auch der Grundidee des klassischen Delphiverfahrens führen, die davon ausgeht, dass ein Delphi solange durchgeführt wird, bis der Prozess *gesättigt* erscheint (vgl. Renn et al. 1999: 99) bzw. sich die Ergebnisse als *fruitful* erweisen (vgl. Häder 2009: 40). Folgt man dieser Grundidee, so ist die Durchführung eines Gruppendelphis als eintägiger Workshop „(...) ein Kompromiss, der jedoch nicht immer ermöglicht das ideale Abbruchkriterium zu erreichen, so, dass keine Verschiebung der Urteile der ExpertInnen mehr auftritt“ (Kuhn et al. 2014: 78). Wie genau ein solches Abbruchkriterium definiert werden kann und welche Implikationen sich daraus ergeben, wird im folgenden Kapitel 4.1.3 Abbruchkriterium näher beschrieben. Einen Überblick über die diskutierten Variationen beim Einsatz des Gruppendelphis gibt Abbildung 2.

Abbildung 2: Variationen beim Einsatz des Gruppendelphis



Quelle: Eigene Darstellung

4.1.3 Abbruchkriterium

Wie bereits in den vorhergehenden Kapiteln beschrieben wurde, beinhalten sowohl das klassische Delphiverfahren wie auch das Gruppendelphi als ein zentrales methodisches Element

⁸ *NEEDS* – *New Energy Externalities Development for Sustainability*. Informationen zum Projekt: www.needs-project.org

die Untergliederung in Delphirunden, die solange wiederholt werden bis der Prozess *gesättigt* erscheint (vgl. Renn et al. 1999: 99). Die Abfolge aus Kleingruppenphase, Feedback und Plinarsitzung soll solange fortgeführt werden, bis keine wesentlichen Verschiebungen der Einschätzungen bzw. Bewertungen der Teilnehmenden mehr auftreten (vgl. Webler et al. 1991: 255; Renn/Webler 1998: 79; Schulz/Renn 2009: 19). Das Ziel dieser Wiederholung besteht darin „(...), die kollektive Urteilssicherheit zu erhöhen und die Varianz der Antworten zu reduzieren. Die Schleife aus Befragung, Auswertung und Feedback wird idealiter solange wiederholt, bis die Experten keine Änderungen ihrer Urteile mehr vornehmen (...)“ (Niederberger/Kuhn 2013: 58). Für eine sinnvolle Anzahl der Delphirunden gibt es aber keinen festgeschriebenen Standard, allerdings wird die optimale Anzahl meist „(...) theoretisch an einem Abbruchkriterium festgemacht. Über die Definition eines Abbruchkriteriums existieren aber unterschiedliche Auffassungen“ (Vorgrimmler/Wübben 2003: 766). Ein theoretisch anvisiertes Abbruchkriterium ist die Erreichung eines Konsenses zwischen den eingebundenen Expert*innen zu allen abgefragten Sachverhalten - oder gleichfalls eines Konsenses über einen Dissens zu einzelnen Sachverhalten (vgl. Niederberger/Kuhn 2013: 59).

Da sowohl bei Gruppendelphis als auch beim klassischen Delphiverfahren meist quantitative Items abgefragt werden, lassen sich hier statistische Festlegungen eines Konsenses finden. So bestimmen Scheibe et al. beispielsweise einen Konsens „(...) when the interquartile range is no larger than two units on a ten units scale“ (ebd. 1975: 271). So findet sich auch bei Schulz et al. eine Bestimmung mittels statistischer Maßzahlen ebenfalls bei 10er skalierten Items. „Wenn z.B. die Standardabweichung im Max bei 1 lag, wurde Konsens angenommen. Auf der Skala entspricht dies einer maximalen Breite von drei Items“ (Schulz et al. 2009: 54). Oft finden sich aber keine theoretisch festgelegten Kriterien, die das Abbruchkriterium eines Delphis markieren. Allerdings zeigt sich die Festlegung auf statistische Maßzahlen gerade bei dem klassischen Delphiverfahren als ein gängiger Weg (vgl. Powell 2003: 4).

Auch Webler et al. definieren als Abbruchkriterium eines Gruppendelphis den Moment an dem sich die Ergebnisse stabilisiert haben und damit keine Verschiebungen der Urteile mehr stattfinden. Bei einem Gruppendelphi soll dieser Status aber klar erkennbar sein, da die möglichen Umformulierungen des Fragebogens, der iterative Prozess von Informationsaustausch und die Möglichkeit der Urteilsänderung einzelner Expert*innen das Ausmaß eines möglichen Konsenses über die Runden erkennbar machen – oder eben nicht (vgl. ebd. 1991: 255).

Neben den theoretischen Überlegungen haben in der Praxis allerdings meist zwei andere Faktoren einen Einfluss auf die Anzahl der Runden: Zeit und Kosten. Zudem bestätigen Erfahrungen nach Häder und Häder, dass die größten Veränderungen der Bewertungen zwischen der ersten und zweiten Runde eines klassischen Delphis auftreten und in den meisten Fällen ein für die Organisatoren des Delphis befriedigendes Ergebnis in der Mehrheit der Fälle nach circa drei Runden erzielt werden konnte (vgl. ebd. 2000: 17). Auch bei Gruppendelphis wird meist aufgrund von Zeitmangel auf die Durchführung von mehr als drei Runden verzichtet. In vielen Fällen zeigt sich allerdings auch hier, dass dies zu ebenfalls befriedigenden Ergebnissen führen kann (vgl. Gallego Carrera 2009b: 79-93; Ruddat/Beninghaus 2009: 57-63; Kuhn et al. 2014: 61-83). Es finden sich allerdings auch Fälle in denen die Zeit zu knapp war, um alle abgefragten Sachverhalte sättigend bewerten und diskutieren zu können. Dies ist meist bei eintägigen Workshops zu beobachten (vgl. Deuschle/Sonnberger 2009: 65-77; Schulz et al. 2009: 47-55). Eintägige Workshops sind, wie bereits in Kapitel 4.1.2 beschrieben, häufig ein

Kompromiss der nicht immer ermöglicht das theoretisch anvisierte Abbruchkriterium zu erreichen und dem Umstand entgegenkommt, das Expert*innen als Teilnehmende eines Gruppendelphis zeitlich nur begrenzt für Forschungsvorhaben zur Verfügung stehen können (vgl. Deuschle/Sonnberger 2009: 67; Kuhn et al. 2014: 78). Generell sollte sowohl für ein Gruppendelphi wie auch für ein klassisches Delphiverfahren das Ziel verfolgt werden, den Spagat zwischen den begrenzten Ressourcen Zeit und Geld und einem zufriedenstellenden Ergebnis - im Ideal ein Konsens - zu bewältigen. Oder wie es Häder und Häder formulieren: *“Als Optimum wird aber allgemein eine minimale Anzahl von Runden bei einem akzeptablen Maß an erzielter Genauigkeit angesehen“* (ebd. 2000: 17).

4.1.4 Anzahl der Teilnehmenden

Da die Teilnehmenden bei einem Gruppendelphi in einer Workshopsituation zusammenkommen, ist der Anzahl der Personen, die an diesem Verfahren teilnehmen können, eine Grenze nach oben gesetzt. So liegt die Begrenzung zum einen schlicht an der verfügbaren Raumgröße. Auch die Größe der Kleingruppen (siehe Kapitel 4.1.1) spielt eine Rolle, so dass pro Kleingruppe drei bis sechs Personen teilnehmen sollten, um eine Diskussion zur gemeinsamen Bewertung der jeweiligen Items zu ermöglichen. Da die Auswertung bei einem Gruppendelphi in einer kurzen Zeitspanne durchgeführt werden muss, sollten höchstens sechs Kleingruppen parallel stattfinden. Damit ergibt sich nach Summierung eine maximale Anzahl von 36 Teilnehmenden, die auch im Hinblick auf einen zielgerichteten Austausch von Argumenten in der Plenarsitzung nicht überschritten werden sollte. *„Bei Überschreiten einer gewissen Gruppengröße ist eine zielgerichtete Diskussion nicht mehr ohne weiteres sicherzustellen“* (Kuhn et al. 2014: 68). Auch bei der Mindestanzahl der teilnehmenden Personen an einem Gruppendelphi hilft der Rückgriff auf die Mindestgröße der Kleingruppendiskussion. In dieser sollten mindestens drei Personen miteinander diskutieren. Auch sollten mindestens zwei Kleingruppen stattfinden, um einen notwendigen Vergleich zwischen den Bewertungen der einzelnen Kleingruppen und damit eine Diskussionsgrundlage für die Plenarsitzung zu haben. Das absolute Minimum eines Gruppendelphis sollte damit nicht unter sechs teilnehmenden Personen liegen (vgl. Kuhn et al. 2014: 68). Es ist aber empfehlenswert von einem absoluten Minimum von neun Personen auszugehen, damit Diskussionen nicht schnell versiegen und der Austausch angeregt wird (vgl. Gallego Carrera 2009a: 108). Erfahrungen bisher durchgeführter Gruppendelphis zeigen, dass die Spanne von 9 bis 25 Teilnehmenden als durchaus praktikable Anzahl für die erfolgreiche Durchführung der Methode angesehen werden kann (vgl. Webler et al. 1991: 259; Deuschle/Sonnberger 2009: 69; Gallego Carrera 2009b: 83; Rudat/Beninghaus 2009: 58; Schulz et al. 2009: 48; Schulz/Wassermann 2010: 6; Hilpert et al. 2013: 10; Niederberger/Kuhn 2013: 62; Kuhn et al. 2014: 68). Auch bei einem klassischen Delphi wurden bereits Experimente zur Panelgröße durchgeführt. Häder berichtet von einem Experiment, in dem zwei Gruppen von einmal 16 und einmal 32 Personen die gleichen Fragen bearbeiteten. Die Ergebnisse zwischen beiden Gruppen stimmten zu 92,9% überein. Häder kommt zumindest für qualitativ ausgerichtete Befragungen zu dem Schluss, dass auf eine große Anzahl von Teilnehmenden verzichtet werden sollte (vgl. ebd. 2009: 98). Ebenso folgert Powell für das klassische Delphiverfahren, dass ein Delphi nicht nach Repräsentativität streben sollte und praktische Überlegungen bei der Personenanzahl eine der zentralen Rollen spielen sollten: *„Resources in terms of time and money are important and influential, yet an*

assessment of the magnitude of the problem and acceptability of answers are open to interpretation by researcher and commentator alike“ (Powell 2003: 378). So macht es vor diesem Hintergrund durchaus Sinn sich bei der Planung eines Gruppendelphis an den in der Praxis gewonnenen Erfahrungen zu orientieren und bei der Rekrutierung darauf zu achten einen Wert von 10 bis 36 Teilnehmenden zu erreichen (vgl. Webler et al. 1991: 257; Niederberger 2015b: 119).

4.1.5 Aufgabe der Anonymität

Einer der Hauptunterschiede zwischen einem klassischen Delphi und einem Gruppendelphi ist die Frage der Anonymität. Während bei einem klassischen Delphiverfahren sowohl eine anonymisierte Befragungssituation sowie auch ein anonymisiertes statistisches Feedback zu den zentralen Elementen des Verfahrens zu zählen sind (vgl. Häder/Häder 2000: 17; Häder 2009: 148), wird im Gegensatz dazu, die Anonymität bei einem Gruppendelphi als methodisches Element aufgegeben, da die Expert*innen sich zu einem gemeinsamen Workshop zusammenfinden (vgl. Webler et al. 1991: 254f.; Renn et al. 1999: 99) .

Der Grundgedanke von Anonymität als methodisches Element bei einem klassischen Delphi basiert auf der Annahme, dass dadurch ein zusätzliches Maß an Legitimität für den Prozess generiert werden kann. Durch eine anonyme Befragungssituation, wie auch ein anonymisiertes Feedback, soll der Einfluss von psychologischen und situativen Faktoren wie Überredung, Abneigung, Meinungsführerschaft und Mehrheitsdruck vermieden werden und es den Befragten damit erleichtern, ein einmal getroffenes Urteil in der nächsten Runde zu widerrufen (vgl. Cuhls/Möhrle 2005: 105; Ammon 2009: 459; Häder 2009: 148). Diese Vorteile sieht auch Goodman, wenn sie von den offensichtlichen Vorzügen der zugesicherten Anonymität im Delphiverfahren spricht: *„The obvious advantage of guaranteed anonymity is that it encourages opinions which are true and not influenced by peer pressure or other extrinsic factors“* (Goodman 1987: 730). Ebenfalls wird sich als weiterer Effekt durch die Gewährleistung von Anonymität eine Erhöhung des Rücklaufs für ein klassisches Delphiverfahren versprochen. So soll die Bereitschaft der Beteiligung, bei besonders unsicheren Sachverhalten eine Einschätzung abzugeben, erhöht werden (vgl. Häder/Häder 2000: 19f.; Vorgrimmler/Wübben 2003: 766).

Neben den positiven Effekten werden aber auch Nachteile durch die anonyme Befragungssituation beim klassischen Delphiverfahren gesehen. So besteht die Gefahr, dass die Teilnehmenden dazu verleitet werden nicht auseichend reflektierte oder übereilte Bewertungen abzugeben. Verstärkend kommt hinzu, dass die Teilnehmenden aufgrund der Anonymität ihres abgegebenen Urteils nicht für ihre jeweiligen Urteile verantwortlich gemacht werden können oder sich für diese rechtfertigen müssen (vgl. Goodman 1987: 730; vgl. Häder/Häder 2000: 17). Ebenfalls kann es dazu kommen, dass ein durch die Auseinandersetzung mit fremden Argumenten angetriebener Lernprozess, durch den anonymen Rahmen des Delphis nicht ausreichend gefördert wird (vgl. Bowles 1999: 32; Vorgrimmler/Wübben 2003: 767).

Bei einem Gruppendelphi kann die Anonymität aufgrund der Diskursivität des Verfahrens mit persönlichem Zusammentreffen der Expert*innen nicht zugesichert werden. Dieser Verlust eines zentralen methodischen Elementes wird teilweise kritisch betrachtet und als Nachteil interpretiert. *„Nachteil des offenen Diskussionsverfahrens beim Gruppen-Delphi ist es jedoch,*

daß die Anonymität der Teilnehmer nicht gewahrt bleibt“ (Renn/Webler 1998: 79). Eine mögliche Meinungsführerschaft tritt so als Gefahr zu Tage und kann mitunter zu einem unerwünschten Scheinkonsens führen. Diesem *bandwagon effect* kann allerdings begegnet werden, indem Personen mit einem ähnlichen Status als Teilnehmende für das Gruppendelphi rekrutiert werden (vgl. Webler et al. 1991: 257f.; Renn/Webler 1998: 79). Weiterhin kann einer Meinungsführerschaft vorgebeugt werden, indem die Struktur des Verfahrens genau eingehalten wird. In den Kleingruppen kann der Effekt bereits durch die vorgesehene Permutation der Teilnehmenden über die verschiedenen Runden gemindert werden (vgl. Kuhn et al. 2014: 67).⁹ Auch einem Effekt wie Mehrheitsdruck kann durch die Struktur des Gruppendelphis begegnet werden, indem es den Teilnehmenden möglich ist Minderheitsvoten abzugeben (vgl. Niederberger/Kuhn 2013: 59). In den Plenarsitzungen ist die Moderation für die Sicherstellungen einer ruhigen und sachlichen Diskussion zuständig. So können Effekte wie Meinungsführerschaft oder Mehrheitsdruck durch Versachlichung, Konfrontation und Konkretisierung auch gemindert werden (vgl. Schulz/Renn 2009:115; Niederberger/Renn 2018: 49).

Bezüglich der Autorität des Feedbacks hat die Aufgabe der Anonymität durchaus positive Aspekte. So wird das Feedback von den Expert*innen nur angenommen, wenn sie den Eindruck haben, dass die abgegebenen Bewertungen von in ihren Augen anderen, ebenfalls kompetenten Personen stammen. Eine Änderung der eigenen Bewertung auf Grundlage des Feedbacks wird begünstigt, wenn die anderen Teilnehmenden als glaubwürdig wahrgenommen und ihre Urteile nachvollziehbar begründet werden. Bei dem klassischen Delphiverfahren soll eine solche Glaubwürdigkeit unter anderem mittels Teilnehmer*innenlisten hergestellt werden, die den Expert*innen zugänglich gemacht werden (vgl. Häder 2009: 95). Begründungen für die jeweiligen Urteile finden hier in der Regel aber nicht statt. Die vielzähligen Argumente über das Für und Wider der Anonymität als methodisches Element eines Delphis zeigen, dass es durchaus auch legitim sein kann, das Element der Anonymität aufzugeben.

Da bei einem Gruppendelphi keine Anonymität zugesichert werden kann, sollte bei der Rekrutierung darauf geachtet werden, dass die potentiellen Teilnehmenden über einen ähnlichen Status verfügen. Ebenfalls sollte die vorgesehene Struktur eines Gruppendelphis eingehalten werden, um die angesprochenen Nachteile, die durch die Aufgabe der Anonymität auftreten können, zu mindern. Ein solch strukturierendes Element ist der Fragebogen, dessen Design im folgenden Kapitel näher beschrieben werden soll. Final soll aber noch eine unterstützende Aussage zur Aufgabe der Anonymität von Häder zitiert werden, der generell aber ein Fürsprecher dieser für das klassische Delphiverfahren ist (vgl. Häder 2009: 148ff.). *„Eigene Erfahrungen haben gezeigt, dass die meisten Experten (scheinbar) keinen Wert auf einen vertraulichen Umgang mit ihren Antworten legen“* (Häder 2009: 149).

4.2 Fragebogenkonstruktion

Der während eines Gruppendelphis eingesetzte Fragebogen erfüllt eine wichtige Funktion im Prozess. Neben den abgefragten Sachverhalten dient dieser auch der Strukturierung der Diskussion, sowohl in der Kleingruppe als auch im Plenum. Die im Fragebogen abgefragten Items werden in beiden Phasen des Prozesses einzeln diskutiert. Damit wird der Fokus auf einzelne

⁹ Siehe Kapitel 2.1.1.

Sachverhalte gerichtet und setzt der Diskussion gewisse Leitplanken. *Grundsatzdiskussionen* bzw. *Fensterreden* können damit weitgehend vermieden werden (vgl. Niederberger/Kuhn 2013: 59; Niederberger 2015a: 45). Denn das Ziel während des Workshops ist keine Ideenaggregation, die Förderung kreativer Prozesse oder die Bereitstellung einer politischen Plattform für die bereits angesprochenen Fensterreden, sondern ein Austausch von Wissen sowie das Finden einer im Idealfall von allen beteiligten Expert*innen gemeinsam getragenen Bewertung des jeweiligen Items im Fragebogen. Bei einem Gruppendelphi werden, wie im klassischen Delphiverfahren, vorwiegend quantitative Fragebögen eingesetzt. So wird ermöglicht, dass in der knappen Auswertungszeit während des Workshops Durchschnitte und Varianzen mathematisch bestimmt werden können (vgl. Schulz/Renn 2009: 23; Niederberger/Kuhn 2013: 59; Niederberger 2015a: 45). Die mittels Fragebogen erhobenen Daten dienen während eines Gruppendelphis besonders der Strukturierung einer Diskussion in der Plenarsitzung und sollen den Austausch von Argumenten und Begründungen auf die jeweiligen Einschätzungen strittiger Sachverhalte lenken: *„Die statistischen Werte dienen bei einem Gruppendelphi zur Strukturierung der Diskussion im Plenum und sind nicht wie bei einem klassischen Delphi auf die Erreichung einer hohen Fallzahl ausgelegt“* (Kuhn et al. 2014: 79). Diese Funktion des Fragebogens sollte bei der Konstruktion immer im Vordergrund stehen, ebenso wie der generelle Grundsatz für eine empirische Untersuchung: *„Denn bei jeder empirischen Untersuchung müssen Planung und Konstruktion des Fragebogens der inhaltlichen Schwerpunktsetzung der Studie angemessen sein“* (Schulz/Renn 2009:41).

4.2.1 Einsetzbare Skalentypen eines Delphifragebogens

Um Durchschnitte und Varianzen berechnen zu können, werden in der Regel vorwiegend *quasi-metrisch* skalierte Items verwendet. Die Zustimmung oder Ablehnung der eingebundenen Expert*innen könnte prinzipiell auch mittels dichotomer Abfrage erfolgen, allerdings wäre so keine differenzierte Bewertung möglich. *„Eine Aufstockung der Kategorien kann vorgenommen werden, um likert-skalierte oder kategoriale Variablen als quasi-metrisch skalierte Variable für eine methodische Auswertung nutzen zu können“* (Kuhn et al. 2014: 71). Als Minimum sollten mindestens fünf Kategorien verwendet werden, um von quasi-metrischen Items auszugehen und die Berechnung von Mittelwerten und der Standardabweichung¹⁰ vornehmen zu können. Allgemein gilt hier aber der Grundsatz: Je mehr Kategorien, desto besser. Denn bei genauer Betrachtung liefern **Ratingskalen** lediglich ordinale Daten (vgl. Urban/Mayerl 2008: 275). In Fragebögen eines Gruppendelphis werden aus diesem Grunde meist 10er-skalierte Items verwendet, weil hier davon ausgegangen werden kann, dass die Kategorien und die Abstände zwischen den Kategorien von den Befragten als gleich groß interpretiert werden. Die Auswahl einer 10er-Skalierung basiert auf der Überlegung, *„(...) dass Unterschiede in den Streuungen und Mittelwerten differenziert herausgearbeitet werden können, der kognitive Aufwand bei der Beantwortung aber begrenzt bleibt“* (Niederberger/Renn 2018: 70).

Die Vermeidung einer Mittelkategorie, wie bei ungerade skalierten Items, wird bei einem Gruppendelphi bewusst vermieden, da durch die fehlende Ausweichoption einer Mittelkategorie

¹⁰ Standardabweichung und Varianz sind als gleichwertige Streuungsmaße anzusehen. Für deskriptive Zwecke soll die Standardabweichung bevorzugt werden, da diese die gleiche Einheit besitzt, wie ursprüngliche Messwerte (vgl. Benninghaus 2002: 61). Dies kommt auch bei Gruppendelphis zur Anwendung.

eine Diskussion zwischen den Teilnehmenden sowie die Abgabe einer positiven oder negativen Tendenz angeregt werden soll. Ebenfalls wird für die Auswertung die Problematik der *Ambivalenz-Indifferenz* vermieden. Dadurch wäre mit Wahl der Mittelkategorie durch die Befragten bei der Auswertung nicht eindeutig bestimmbar, ob eine dezidierte Meinung vertreten wird oder ob die Bewertung dieses Items ambivalent ist. „Für ein Gruppendelphi erscheint eine solche Mittelkategorie weder geeignet noch notwendig (...)“ (Schulz/Renn 2009: 32), so dass immer eine gerade Anzahl an Antwortkategorien für die Items verwendet werden sollte. Die Sicherheit der Teilnehmenden kann zusätzlich entweder durch Abfrage der Urteilssicherheit im Fragebogen und/oder in der Diskussion während der Plenarsitzung näher bestimmt werden. Ebenso sollen im Rahmen eines Delphis Handlungsoptionen entweder als machbar oder eben als nicht-machbar bewertet werden, so dass eine Mittelkategorie hier wenig sinnvoll erscheint. (vgl. Mayer 2006: 82; Schulz/Renn 2009: 30-39; Kuhn et al. 2014: 71).

Neben den 10er-skalierten Items können mit Hilfe von Ratingskalen im Fragebogen eines Gruppendelphis beispielsweise Zeitintervalle oder prozentuale Wahrscheinlichkeiten abgefragt werden. Auch offene Fragen sind hier möglich, indem Fragen nach Häufigkeiten in den Fragebogen integriert werden können. Ratingskalen bieten den Vorteil eines vergleichsweise geringen kognitiven und zeitlichen Aufwandes für die Teilnehmenden während der Beantwortung, wie auch für das Forscherteam bei der Auswertung. Da Zeit ein wichtiger Faktor während eines Gruppendelphis ist, ist ein vorwiegender Einsatz dieser Fragetypen für die Fragebogenkonstruktion empfehlenswert. Zumal ein häufig beobachteter negativer Effekt bei Verwendung von Ratingskalen im Rahmen eines Gruppendelphis vermieden werden kann: Bei vielen hintereinander geschalteten Items kann ein Response Set bzw. Routineverhalten einsetzen. Diesem wird in vielen Fällen durch wechselnde Beurteilungsrichtungen zu begegnen versucht. Da in einem Gruppendelphi aber zu jedem Item eine Diskussion in der Kleingruppe erfolgt, ist für den hier einzusetzenden Fragebogen von diesem Effekt nicht auszugehen (vgl. Schulz/Renn 2009: 30-37; Kuhn et al. 2014: 71).

Um eine relative Einschätzung abzufragen, können im Fragebogen eines Gruppendelphis auch **Rankingskalen** zum Einsatz kommen. Diese haben den Vorteil, dass eine Priorisierung zwischen verschiedenen Alternativen durch die Befragten innerhalb einer Frage vorgenommen werden kann. Die Befragten werden in diesem Fall gebeten eine Rangordnung zwischen verschiedenen Antwortalternativen zu erstellen. Die Abfrage kann beispielsweise über die Vergabe von Schulnoten, über ein fiktives Budget, oder über eine Vergabe von Punkten erfolgen. Rankingskalen bieten den Vorteil einer geringeren notwendigen Abstraktionsebene als Ratingskalen, da beispielsweise Maßnahmen mittels der Verteilung eines fiktiven Budgets nah an realen Bedingungen bewertet werden können und eine differenzierte Aufteilung eben dieses Budgets durch die Teilnehmenden erfolgen kann (vgl. Kuhn et al. 2014 71f.; Niederberger/Renn 2018: 68-73).

Es empfiehlt sich allerdings nicht ausschließlich Rankingskalen im Gruppendelphi einzusetzen, da diese erfordern, „(...) dass die Alternativen mitbedacht werden müssen und somit steigen die kognitiven Anforderungen für den Befragten“ (Schulz/Renn 2009: 30). Auch die Entwicklung der Fragen mit eingesetzten Rankingskalen benötigt einen höheren Aufwand, da die möglichen Alternativen bereits im Vorfeld bestimmbar sein und mitbedacht werden müssen; Vorwissen ist hier also unbedingt erforderlich. Da es sich bei Rankingskalen um eine ipsative

Messung handelt - die Messwerte sich also gegenseitig bedingen - ist auch die statistische Auswertung aufwendiger (vgl. Schulz/Renn 2009: 33).

Betrachtet man beide Skalentypen zusammenfassend so bieten Rankingskalen den Vorteil, dass (1) die Alternativen bestimmbar sind, (2) die Fragen bei Beantwortung genau bedacht werden müssen und (3) eine Priorisierung im Rahmen der Frage möglich ist. Ratingskalen hingegen haben den Vorteil, dass sie (1) einen geringeren Zeitaufwand beim Ausfüllen und Auswerten benötigen, (2) einen geringeren kognitiven Aufwand für die Teilnehmenden voraussetzen und (3) absolute Interpretationen möglich sind. Beide Skalentypen bieten also Vorteile, die je nach Zeitbedarf und Ziel der Befragung eine Orientierung für die Konstruktion des Fragebogens bieten können (vgl. Schulz/Renn 2009: 33).

4.2.2 Integration weiterer methodischer Elemente

Bei standardisierten Befragungen sollen die Befragten nicht zu einer Antwortvorgabe gedrängt werden, falls diese bei der Abgabe einer Antwort bzw. Bewertung unsicher sind oder aber keine Angabe machen möchten. Diese Funktion erfüllen in quantitativen Fragebögen **Ausweichkategorien** wie „*Weiß nicht*“, „*Keine Angabe*“ oder „*Trifft nicht zu*“. Methodisch wird so das Dilemma gelöst, dass bei der Auswertung der Antworten nicht mehr festgestellt werden kann, wie die Befragten zu ihrem Urteil kamen und wie sicher sie sich in diesem Urteil sind. Ausweichkategorien erfüllen bei einer standardisierten quantitativen Befragung methodisch drei Funktionen:

1. Die Befragten werden bei bestehender Unsicherheit oder Ungewissheit nicht zu einer Antwort genötigt.
2. Die Angemessenheit und Relevanz der Frage kann anhand der Häufigkeit der Auswahl einer Ausweichkategorie durch die Befragten kritisch reflektiert werden.
3. Bei Einsatz von sowohl *Weiß nicht*- als auch *Keine Angabe*-Kategorien kann eine Differenzierung zwischen Nicht-Wissen und Antwortverweigerung bei der Auswertung durchgeführt werden (vgl. Schulz/Renn 2009: 38; Niederberger/Renn 2018: 77).

Obwohl bei einem Gruppendelphi Unsicherheit bei einigen Items eine große Rolle spielen kann, erscheint die Integration einer Ausweichkategorie im Fragebogen als nicht sinnvoll.¹¹ So besteht während der Kleingruppenphase die Möglichkeit Mehrheits- und Minderheitsvoten zu vergeben, die Expert*innen müssen also nicht auf ein *keine Angabe* aufgrund eines fehlenden Konsenses ausweichen. Vielmehr wird auf diesem Weg die Spanne der Uneinigkeit im Fragebogen abgebildet und erfasst. Weiter können kritische Fragen im Plenum aufgegriffen werden und der Grad der Uneinigkeit inhaltlich näher erläutert werden. Bei Themenkomplexen, die sich schon im Vorfeld des Gruppendelphis durch eine hohe Unsicherheit auszeichnen, kann zusätzlich eine Abfrage der Urteilssicherheit zu den jeweiligen Items in den Fragebogen integriert werden, die als funktionales Äquivalent zu herkömmlichen Ausweichkategorien gesehen werden kann (vgl. Schulz/Renn 2009: 39 Niederberger/Renn 2018: 77f.).

Wie bereits weiter oben in diesem Kapitel angesprochen, ermöglicht die Abfrage mittels quasi-metrischen Skalen den Teilnehmenden die Abgabe eines differenzierteren Urteils, als dies bei

¹¹ Der Umgang mit Unsicherheit ist eines der zentralen Ziele eines Delphis, wie bereits in Kapitel 3.1 angesprochen.

einer Abfrage von dichotomen Bewertungen der Fall wäre. Allerdings wachsen so auch die Herausforderung einer adäquaten Antwort und potentiell auch die Unsicherheit bezüglich des eigenen Urteils. *„Das heißt: umso differenzierter die Antwortvorgaben, desto schwerer ist die Abgabe eines exakten Urteils“* (Schulz/Renn 2009: 39). Besonders bei inhaltlich schwierigen Themen wächst hier die Relevanz von **Fragen zur Urteilssicherheit** (vgl. Schulz/Renn 2009: 39) bzw. *subjektiven Kompetenzfragen* (vgl. Häder 2009: 126), die eine Identifikation von unsicheren Ansichten der Expert*innen ermöglichen. So wäre es möglich, dass die Expert*innen bei einem Item zwar zu einer ähnlichen Einschätzung gelangen, dies könnte aber auf hoher Unsicherheit und jenseits ihrer fachlichen Expertise begründet liegen. Ein solcher Fall wäre ohne Integration von Fragen zur Urteilssicherheit bei der statistischen Auswertung nicht zu ermitteln. Da diese Frage aber aufgrund der ähnlichen Einschätzung konsensual bewertet wurde, wird sie nicht in der Plenarsitzung diskutiert. Die Unsicherheit würde also nicht aufgedeckt. Fragen zur Urteilssicherheit haben überdies eine psychologische Funktion, da sie den Befragten vermitteln, dass keine korrekten numerischen Antworten, sondern die am ehesten angenommenen Einschätzungen erwartet werden. Hohe Bewertungen gekoppelt mit hoher Sicherheit können letztlich auch höher gewichtet werden, so dass sich durch den Einsatz von Fragen ein klareres Bild zur Urteilssicherheit ergibt (vgl. Häder 2009: 128; Schulz/Renn 2009: 39f.; Niederberger/Renn 2018: 78ff.).

Da weitere Fragen allerdings auch zusätzliche Zeit für die Beantwortung benötigen, finden sich ebenso Praxisbeispiele bei denen - vorwiegend aus Gründen der Zeitersparnis - keine Urteilssicherheit abgefragt wurde (vgl. Hilpert et al. 2013; Niederberger/Kuhn 2013; Kuhn et al. 2014). Eine weitere praktische Erkenntnis zeigt, dass die Urteilssicherheit bei Ratingskalen relativ hoch ist. Gerade bei unmissverständlicher und eindeutiger Frageformulierung tendieren die Teilnehmenden zur Kennzeichnung sicherer Urteile. *„Möglicherweise zeigt sich hier ein Effekt der sozialen Erwünschtheit. Denn die Experten müssen innerhalb der Kleingruppendiskussion und später im Plenum Unsicherheiten eingestehen“* (Schulz et al. 2009: 53). Diese Beobachtung wurde während eines Delphis mit homogener Zusammensetzung des Feldes der Teilnehmenden gemacht. Bei einem heterogenen Feld der Teilnehmenden könnte aber erwartet werden, dass die Unsicherheiten zunehmen, so dass hier eine Abfrage der Urteilssicherheit wieder sinnvoll sein kann (vgl. Schulz et al. 2009: 53). Zudem ist zu vermuten, dass der Effekt der sozialen Erwünschtheit bezüglich hoher Urteilsicherheit in einem Teilnehmendenfeld mit heterogener Expertise nachlässt, da es in diesem Fall für die Teilnehmenden opportuner erscheint fehlende Expertise bei einzelnen Sachfragen offenzulegen.

Abschließend lässt sich für die Abfrage der Urteilssicherheit festhalten, dass sie für Ratingskalen weniger relevant scheint als bei einem Einsatz von Rankingskalen, da diese höhere kognitive Fähigkeiten von den Befragten verlangen. *„Je höher der kognitive Anspruch, desto wichtiger ist die Abfrage der subjektiven Sicherheit“* (Schulz/Renn 2009: 40). Ebenfalls kann eine Abfrage der Urteilssicherheit besonders in der ersten Delphirunde und in der möglichen Vorbefragung sinnvoll sein und in weiteren Runden aus dem Fragebogen ausgelassen werden (vgl. Schulz/Renn 2009: 40).

Expert*innenbegriffe und Fachtermini, die im Normalfall bei einer Befragung vermieden werden sollten, können allerdings im Fragebogen eines Delphi zum Einsatz kommen. Da das Teilnehmendenfeld eines Gruppendelphis aus Expert*innen besteht, sind diese Personen mit Fachtermini vertraut und deren Verwendung gewöhnt. So besteht hier überdies sogar eine Erwartungshaltung, dass diese Fachtermini auch im Gruppendelphi zum Einsatz kommen, um

eine trennscharfe und gezielte Diskussion der einzelnen abgefragten Sachverhalte führen zu können (vgl. Schulz/Renn 2009: 111; Kuhn et al. 2014: 72).

4.2.3 Offene Abfragen

Die Vorgehensweise des klassischen Delphiverfahrens einer standardisierten Befragung, in der alle Fragen und Antwortmöglichkeiten ausformuliert sind, wurde auch für das Gruppendelphi übernommen. Einer der entscheidenden Gründe hierfür ist, wie bereits in Kapitel 4.2.1 zu den Skalentypen angemerkt wurde, der Aspekt der Zeitersparnis. Allerdings findet dieser Fragetyp nicht nur Anwendung im Typus der ausschließlich qualitativ ausgelegten Delphibefragungen¹², in denen die Expert*innen um eine offene Stellungnahme gebeten werden, sondern auch bei vorwiegend quantitativ standardisierten Delphis. Dies geschieht vorwiegend bei Fragen, die sich im Vorfeld nur schwer als geschlossene Frage formulieren oder konstruieren lassen. Auch kann eine offene Frage am Ende des Fragebogens integriert werden, in der um weitere Aspekte und Bemerkungen gebeten wird (vgl. Häder 2009: 136f; Schulz/Renn 2009:27).

Dies kann ebenso für Gruppendelphis gelten, denn auch hier kann es dazu kommen, dass dem Forscherteam vor der Durchführung des Workshops nicht alle relevanten Aspekte bekannt sind, so dass neben der Abfrage von Konsens und Dissens auch die Generierung von zusätzlichem Wissen notwendig ist. *„Offene Fragen sind dann sinnvoll, wenn zusätzliche Optionen (etwa ein Maßnahmenkatalog) oder zusätzliche Bewertungsaspekte erwartet werden können, die den Konstrukteuren des Fragebogens nicht bekannt sind“* (Schulz/Renn 2009: 113). Wenn eine Vorbefragung stattfindet, können so bereits vor Durchführung des Workshops zusätzliche Elemente integriert und in den standardisierten Fragebogen transferiert werden. Falls keine Vorbefragung stattfindet, können offene Fragen auch im Fragebogen des Workshops integriert werden. Neben der offenen Frage am Ende des Fragebogens, kann bei geschlossenen Fragen ein offenes Item als Kategorie *Sonstiges* integriert werden. So kann ein zusätzlicher Input durch die Expert*innen für die nächsten Runden des Gruppendelphis sichergestellt werden. Der Einsatz offener Fragen und Kategorien sollte allerdings im Workshop-Fragebogen mit Bedacht und spärlich erfolgen, denn die Gefahr besteht, dass diese bei zu hoher Anzahl in den Pausen nicht ausgewertet und diese Informationen in den weiteren Runden nicht aufgegriffen werden können. Es sollte also sichergestellt sein, dass solche Ergebnisse in den Fragebogen der nächsten Runde integriert werden können (vgl. Webler et al. 1991: 262; Schulz/Renn 2009: 27; Kuhn et al. 2014: 76). So sollte bei der Verwendung von offenen Fragen im Fragebogen eines Gruppendelphis der Empfehlung von Schulz und Renn gefolgt werden, offene Fragen nur marginal einzusetzen (vgl. ebd. 2009: 27).

Abschließend sollte - wie bereits zu Beginn des Kapitels - angemerkt sein, dass die während eines Gruppendelphis erhobenen statistischen Ergebnisse besonders der Hervorhebung strittiger Sachverhalte sowie der Strukturierung der Diskussion im Plenum dienen und nicht auf die Erreichung einer hohen Fallzahl abzielen bzw. angewiesen sind. Somit eignet sich die statistische Aufarbeitung besonders für Fragestellungen zu denen wissenschaftliche Urteile und

¹² Typ 1: Delphi zur Ideenaggregation der in Kapitel 3.2 beschriebenen Typen von Delphibefragungen.

deren Argumente über bestimmte Sachverhalte abgegeben werden sollen. Damit ist das Gruppendelphi vielmehr ein Erkenntnisinstrument denn ein Erhebungsinstrument (vgl. Schulz/Renn 2009: 114; Kuhn et al. 2014: 79; Niederberger/Renn 2018: 87).

4.3 Vorteile der Methode

Eine Gefahr bei schriftlichen standardisierten Befragungen – wie beim klassische Delphiverfahren – ist, dass die Befragten die Fragen unterschiedlich interpretieren *„(...) und dementsprechend divergierende Relevanzsysteme ihren Urteilen zu Grunde legen. Dabei kann es vorkommen, dass zwar mehrere Teilnehmer die gleiche Antwort angeben, aber eben nicht auf der gleichen kognitiven Grundlage“* (Schulz/Renn 2009: 20). Solche Fehlinterpretationen sollen durch die diskursiven Elemente eines Gruppendelphis vermieden werden, denn diese Elemente bieten die Möglichkeit, auch inhaltliche Begründungen und Argumente für die jeweiligen Positionen der jeweiligen Teilnehmenden auszutauschen und sich auf eine gemeinsame Bewertungsgrundlage zu einigen (vgl. Webler et al. 1991: 258; Niederberger/Kuhn 2013: 59; Kuhn et al. 2014: 66f.). Solche Argumentationen können die gemeinsame Beurteilung befördern und fehlen beim klassischen Delphiverfahren gänzlich, falls das Feedback lediglich aus statistischen Mittelwerten besteht (vgl. Rowe et al. 1991: 244-248). Das Verfahren des Gruppendelphis kann einen Rahmen dafür bieten, *„(...) dass günstige Konstellationen für einen intensiven Austausch von inhaltlichen Begründungen und Argumenten bereitgestellt werden. So kann durch einen inhaltlichen Austausch der Argumente die Gefahr gemindert werden, dass sich eine falsche Einschätzung durchsetzt, die sich an der bloßen statistischen Mehrheit orientiert“* (Kuhn et al. 2014: 66f.). Durch iterative Gruppenprozesse können Expert*innen zudem zu einem intensiven Austausch von Gedanken und von Argumenten angeregt werden (vgl. Niederberger/Kuhn 2013: 61). Der Austausch von Begründungen wird somit ein zentrales Element des Feedbackprozesses, denn *„(...) recent reviews and research on the Delphi process suggest that to achieve improvements in accuracy from round to round, experts must be provided with rationales to accompany the feedback they receive about the responses from other group members, and that in the absence of these rationales, their responses will tend to converge only towards a majority position (...)“* (McBride et al. 2012: 907).

Der Austausch von Argumenten bietet auch für die Ergebnisinterpretation Vorteile: So kann bei der Auswertung der Ergebnisse nachvollzogen werden, warum manche Expert*innen ihre Urteile revidiert und andere dies eben nicht getan haben. Solche Bewertungs- und Begründungshorizonte können bei einem klassischen Delphi nicht berücksichtigt werden. Ein Gruppendelphi kann somit aufgrund einer Inklusion von inhaltlichen Argumenten und Begründungen mehr Transparenz und Klarheit in wissenschaftlichen Kontroversen herstellen. So können zusätzliche nützliche Informationen erfasst werden, die bei dem klassischen Delphiverfahren üblicherweise verloren gehen (vgl. Sackman 1974: 29; Niederberger 2015: 118f.).

Neben der Bereitstellung eines Rahmens für einen intensiven Austausch von inhaltlichen Begründungen und Argumenten für die jeweiligen Positionen ergeben sich durch die Modifikation noch weitere Vorteile:

- Während des Workshoptages eines Gruppendelphis ist eine direkte und unmittelbare Rückmeldung (Feedback) möglich, so dass keine äußeren Störfaktoren Einfluss auf die Bewertungen nehmen können.

- Die Diskussion über Sachverhalte bei denen die Urteile der Expert*innen auseinandergehen, gibt zeitgleich Auskunft darüber, welche Abweichungen zwischen den Bewertungen von den Befragten akzeptiert werden.
- Die interne Konsistenz der akzeptierten Sachverhalte wird durch die Diskussion zwischen den Teilnehmenden erneut geprüft.
- Der Austausch zwischen den Expert*innen kann zu einer semantischen Klärung von Begriffen und zur Präzisierung von Fragestellungen und Konzepten beitragen (vgl. Webler et al. 1991: 258; Schulz/Renn 2009: 114; Hilpert et al. 2013: 5; Kuhn et al. 2014: 67).

Ein letzter Vorteil des Gruppendelphis ist der Faktor Zeit. So liegen die Ergebnisse eines Gruppendelphis unmittelbar nach dem Workshop vor. Ebenfalls ist die zeitliche Einbindung der Expert*innen in den Prozess auf ein bis zwei Tage beschränkt, während sie im Rahmen des klassischen Delphiverfahrens über mehrere Monate hinweg zu bestimmten Zeiten eingebunden werden. So stellt eine Verzerrung durch systematische Ausfälle, wie im Paneldesgin eines klassischen Delphis, für das Gruppendelphi keine Gefahr dar. *„Experts are busy people. A questionnaire that requires substantial effort may not reach the desk of someone not directly involved in the matter. On the other hand, experts who have a stake in the particular subject are more willing to participate, but their opinions may be overly biased. The dropout rate of participants is typically high (...) In a circumstance where a particular expert viewpoint is weakly represented on the panel, a substantial dropout rate could sacrifice the validity of the process“* (Webler et al. 1991: 255). Damit ist für die Rekrutierung eines Gruppendelphis umso wichtiger, dass die Teilnehmenden die in der Fachwelt diskutierte Bandbreite an unterschiedlichen Auffassungen, Interpretationen und Begründungen vertreten (vgl. Renn/Webler 1998: 78f.; Schulz/Renn 2009: 117). Wie der Begriff einer Expert*in für ein Gruppendelphi zu definieren ist, soll im nächsten Kapitel erläutert werden.

4.4 Begriff und Auswahl der Expert*innen

„In einer stark ausdifferenzieren und arbeitsteilig organisierten Gesellschaft mit den verschiedenen gesellschaftlichen Subsystemen, dort ansässigen Organisationen, die sich wiederum in funktional sehr unterschiedliche Abteilungen und Arbeitsgruppen unterteilen, existiert kein einheitlicher Erfahrungsraum, den alle Gesellschaftsmitglieder teilen, sondern vielmehr sind vielfältige Realitätsausschnitte davon geprägt, dass dort exklusives Experten- bzw. „Sonderwissen“ (...) existiert“ (Wassermann 2015: 16). Bevor also eine Auswahl der Expert*innen für das Teilnehmendensample eines Gruppendelphis erfolgen kann, ist es wichtig zu bestimmen, was die Rolle als Expert*in in einem solche ausdifferenzierten System auszeichnet. Die Zuweisung der Rolle kann erfolgen durch:

1. Definition als Expert*innen durch die Profession;
2. Definition als Expert*innen durch die Klienten;
3. Definition als Expert*innen durch die Medien;
4. Selbstdarstellung als Expert*in (vgl. Kleimann 1996: 188ff.).

Bei einem Gruppendelphi erfolgt die Zuschreibung des Begriffs der Expertin bzw. des Experten durch das Forscherteam, die dieser Ordnung folgend die Klienten darstellen, da sie Sonderwissen von den Expert*innen abfragen möchten. *„Der Expertenbegriff ist daher relational, d.h.*

durch die Beziehung des Experten zur Profession und zu den Adressaten, definiert“ (Kleimann 1996: 189).

Die Auswahl des Samples für ein Gruppendelphi ist damit relational zu verstehen, da diese für einzelne Wissensgebiete und bestimmte Realitätsausschnitte über ein spezialisiertes und detailliertes Sonderwissen verfügen, das im Hinblick auf die jeweilige Fragestellung als relevant angesehen wird. Den Status Expert*in verleiht damit das Forscher*innenteam, das das Sample danach auswählen sollte, dass dieses entweder für die Implementierung oder die Kontrolle einer Problemlösung Verantwortung trägt oder über einen privilegierten Zugang zu Informationen verfügt (vgl. Kuhn et al. 2014: 67; Wassermann 2015: 16). Die Expert*innen können mit ihrem spezialisierten Sonderwissen somit dazu beitragen, beispielsweise wissenschaftlich-technische Entwicklungen zu beurteilen oder abzuschätzen, indem sie über komplexe Relevanzsysteme verfügen (vgl. Schenk/Döbler 1996: 78; Vorgrimler/Wübben 2003: 765; Martin et al. 2011: 30).

So sollte für die Rekrutierung der Teilnehmenden eines Gruppendelphis der generelle Grundsatz gelten, der u.a. auch für Expert*inneninterviews gilt: *„Expertin ist man nicht an sich, sondern im Hinblick auf ein bestimmtes Wissensgebiet“* (Przyborski/Wohlrab-Sahar 2008: 131). Als praktikables Auswahlkriterium hat sich sowohl für das klassische Delphi wie auch für das Gruppendelphi der Beruf oder das Tätigkeitsfeld der potentiellen Expert*innen erwiesen. So kann sichergestellt werden, dass die rekrutierten Personen in ihrem alltäglichen Umgang mit den für das Delphi relevanten Wissensgebieten vertraut sind. Aus diesem Grunde ist es wichtig im Vorfeld der Rekrutierung Klarheit über die Ziele des Delphis sowie über die Notwendigkeit der damit verknüpften Wissensgebiete zu erlangen. Nur so kann die Struktur der Expert*innen-gruppe mit der Aufgabenstellung des Delphis verknüpft werden und eine sinnvolle Rekrutierung erfolgen (vgl. Häder 2000: 3; Niederberger 2015b: 119).

Die Aufgaben, die Expert*innen bei einem Delphi erfüllen sollen, sind nicht trennscharf abgrenzbar, da die Teilnehmenden neben der Bereitstellung von Wissen oftmals ebenfalls bestimmte Erkenntnisse und Schlussfolgerung wie auch normative Vorschläge einbringen. So stellt die Bereitstellung von Wissen die Basis dar, *„(...) auf dessen Grundlage die Experten und Stakeholder bewerten, beraten oder vorausschauen“* (Niederberger 2015a: 39). So können sowohl Expert*innen sowie auch Stakeholder potentielle Teilnehmende eines Gruppendelphis sein. Hierbei sollen beide Gruppen aber trennscharf definiert werden. Eine Expertin bzw. ein Experte soll basierend auf der eigenen Kompetenz und Expertise objektiv argumentieren, bewerten und beraten. Dabei wird den Expert*innen ihre Objektivität dadurch zugeschrieben, dass sie keinerlei persönliche oder institutionelle Interessen vertreten. Stakeholdern wird ebenfalls inhaltliche Kompetenz bestimmte Realitätsausschnitte betreffend zugeschrieben, allerdings wird diesen im Gegensatz zur Objektivität ein subjektives Interesse innerhalb der Wissensgebiete zugesprochen und diese stehen mitunter für jeweils bestimmte politische Wünschbarkeiten. Aufgrund ihres Tätigkeitfeldes in Verbänden, Unternehmen und NGOs verfügen auch Stakeholder über Sonderwissen, das für Gruppendelphis relevant sein kann. Sie sollten bei bestimmten Sachfragen also ebenso zur Zielgruppe gezählt werden, nicht nur, wenn es darum geht die Wünschbarkeit einzelner Maßnahmen zu bestimmen (vgl. Niederberger/Wassermann 2015: 12f.; Niederberger 2015b: 119; Alcántara et al. 2016: 177).

Ein wichtiges Qualitätskriterium eines Gruppendelphis ist die Besetzung: Sind nur wenige Teilnehmende in den Prozess eingebunden, die im negativsten Fall aufgrund ihres institutionellen

Hintergrundes einen ähnlichen Standpunkt vertreten, dann ist nicht mit qualitativ hochwertigen und reliablen Ergebnissen zu rechnen (vgl. Deuschle/Sonnberger 2009: 68). Deshalb ist es aufgrund der begrenzten Anzahl an Teilnehmenden umso wichtiger bei einem Gruppendelphi die mögliche Bandbreite von Wissen, aber auch von möglichen Bewertungen für die jeweilige Fragestellung zu erfassen und die Teilnehmenden anhand dieser Ziele und Vorgaben zu rekrutieren: *„Unabhängig von der konkreten Zielstellung geht es in der Regel nicht darum, den einen Experten zu integrieren, sondern darum, eine möglichst umfassende Darstellung, idealerweise die gesamte Spannweite an Meinungen zu berücksichtigen“* (Niederberger/Wassermann 2015: 11). Falls davon ausgegangen werden kann, dass alle relevanten Argumente und Standpunkte durch die Anwesenheit vertreten werden, kann das Verfahren als legitim erachtet werden, und mit einem vernünftigen Ergebnis gerechnet werden (vgl. Abromeit 2002: 102).

„It is time for the oracles to move out and for the science to move in”

(Wechsler, zitiert in Häder/Häder 1995: 13).

5 Wissenschaftstheoretische Grundlagen

Oft wird als Argument gegen den Einsatz eines Delphiverfahrens dessen schwache theoretische Fundierung und Begründung angeführt. Eine solche theoretische Fundierung hat in dieser Form bisher kaum stattgefunden. Wenn eine Legitimation des Ansatzes auch theoretisch erfolgen soll, ist sowohl eine wissenschaftstheoretische Grundlage wie auch ein Bemessungskriterium für die Ergebnisse eines Delphis nötig (vgl. Häder 2009: 20-54). Diese Notwendigkeit besteht umso mehr für die Methode des Gruppendelphis, da sie als Modifikation des klassischen Verfahrens eine umso geringere theoretische Fundierung aufweisen kann, die durch die zusätzliche Integration diskursiver Elemente umfassender sein muss, als die des Delphiverfahrens.

Die theoretische Vorannahme, die dem klassischen Delphiverfahren zu Grunde liegt ist, dass in zwei Köpfen mindestens so viel Wissen vorhanden ist, wie in einem, aber - und das ist der ausschlaggebende Punkt – wahrscheinlich sogar mehr. Diese Argumentation basiert auf dem „1+n“ Argument (*n-heads rule*) von Dalkey der davon ausgeht, dass die Informationsqualität proportional mit der Anzahl der in die Befragung eingebundenen Expert*innen steigt. Das Wissen der vielen Köpfe wird sich mit der richtig eingesetzten Technik gegenüber den Fehlinformationen - die nach Dalkey ebenfalls in den Köpfen existieren – durchsetzen (vgl. ebd. 1969: 6). Cuhls und Möhrle sehen neben den positiven Synergieeffekten durch Einbeziehung möglichst vieler Personen auch die Möglichkeit einer größeren *Vorausschau-Genauigkeit* und nach Niederberger und Renn deuten bisherige Erfahrungen mit dem Gruppendelphi auf einen positiven Effekt dieses Prinzips hin (vgl. Cuhls/Möhrle 2005: 104; Häder 2009: 40; Niederberger/Renn 2018: 55).

Gerade beim Einsatz von Delphiverfahren, die eine Verbesserung der Bestimmung eines Sachverhaltes mittels Vorhersagen zum Ziel¹³ haben, wird von einer höheren Treffsicherheit der Prognose ausgegangen, wenn viele Expert*innen eingebunden werden. *„Da niemand genau wissen kann, wie die Zukunft werden wird, sollten möglichst viele Personen beteiligt sein. Denn es ist erwiesen, daß bei einer großen Anzahl von Antworten individuelle Schätzfehler ausgemittelt werden können und damit die Wahrscheinlichkeit einer ‚treffsicheren Prognose‘ steigt“* (Cuhls et al. 1998: 7). Dabei wird davon ausgegangen, dass je größer die Distanz des eigenen zu den anderen Urteilen zu verorten ist, bzw. je größer die eigene Urteilsunsicherheit ist, desto größer fällt auch die Änderung des eigenen Urteils aus. Diese Logik der möglichen Urteilsänderung der Expert*innen unterstützt die Annäherung an das im Kollektiv erzielte bessere Ergebnis (vgl. Häder 2009: 42). *„The improvements with iteration were hypothesized by Dalkey (...) to arise through the motion of the less knowledgeable panelists (known in RAND*

¹³ Typ 2 der in Kapitel 3.2 vorgestellten Typisierung.

jargon as “swingers”), accompanied by relative intransigence of the more accurate panelists (known as “holdouts”). It was suggested that the least accurate panelists will realize that this is the case and hence be drawn toward the median, while the most knowledgeable panelists will be more confident and so less drawn toward the mean or remain they are” (Rowe et al. 1991: 238).

Eng verknüpft mit dem *n-heads rule* ist die Frage nach einem Bemessungskriterium für ein erfolgreiches Delphiverfahren, denn die Regel geht davon aus, dass bei einer größeren Anzahl von eingebundenen Expert*innen eine größere Annäherung an den *wahren* Wert erfolgt. Dalkey führt hierfür - wie bereits in der Begründung von Rowe et al. (vgl. ebd. 1991: 238) erkennbar - eine statistisch basierte Begründung an: *„If the range of group answers includes the true, then, in general, the median is closer to the true answer than more than half of the group (...). In practical situations, the range of answers is very likely to include the true answer, in which case the stronger assertion is valid“ (Dalkey 1969: 7).*

Neben der bestehenden Kritik, dass sich allerdings auch minderwertige oder gar fehlerhafte Informationen unter einer Vielzahl von eingebundenen Personen durchsetzen können, besteht auch die Frage, ob es überhaupt Kriterien gibt, den hier angestrebten *wahren* Wert zu bestimmen.¹⁴ Häder äußert weitergehend Zweifel daran, ob die Annäherung oder Erreichung des wahren Wertes überhaupt das Ziel eines Delphiverfahrens darstellen sollte. Ist das Ziel eines Delphis einen Konsens bezüglich eines strittigen Themas oder aber eine gezielte Ideenaggregation¹⁵ zu erzielen, ist der wahre Wert im Vorfeld nicht definierbar und auch nicht von besonderem Interesse für den Ausgang des Delphiverfahrens selbst. Vielmehr geht es darum am Ende des Prozesses *fruchtbare* bzw. verwendbare Ergebnisse zu erzielen, mit denen die eingebundenen Expert*innen einverstanden sind und die sich für die ursprüngliche Fragestellung als *fruchtbar* bzw. hilfreich erweisen. *„Diese Gedanken fortgeführt macht es auch keinen Sinn, überhaupt die Frage nach der Wahrheit zu stellen, sondern es ist lediglich wichtig zu wissen, ob die bei einer Delphi-Befragung gewonnenen Informationen ‚fruitful‘ beziehungsweise produktiv sind“ (Hader 2009: 40).* Hilfreicher als die Beurteilung des wahren Wertes erscheint damit eine Beurteilung der Ergebnisse, mit den mit der Fragestellung des Delphis verknüpften Zielen. Hiermit steht eine andere Überlegung im Vordergrund, nämlich ob und wie diffuse bzw. unsichere Sachverhalte überhaupt von Expert*innen bewertbar sind und ob und wie diese Bewertungen im Verlauf eines Delphis verbessert werden können (vgl. Häder 2009: 40-44)?

Um die Aussage überprüfen zu können, dass die Ergebnisse eines Delphi im Hinblick auf bestimmte Fragestellungen fruchtbarer oder produktiver sein können als andere Befragungsformen, ist es notwendig zu klären, *wie* die Expert*innen innerhalb des Prozesses zu ihrem Urteil gelangen.

¹⁴ Bzw. *das Wahre* folgt man der wörtlichen Übersetzung von Dalkeys *„the true“* (vgl. ebd. 1969: 7).

¹⁵ Typ 1 (Ideenaggregation) und Typ 4 (Konsens) der in Kapitel 3.2 vorgestellten Typisierung.

5.1 Theorieklasse der mentalen Modelle und Social Judgment Theory

Bei einem Delphi geht es nicht darum eine bloße Einstellungsmessung vorzunehmen, sondern darum, eine wissensbasierte Beurteilung eines Sachverhaltes von Expert*innen zu erhalten. Ein solches Urteil der Expert*innen bezieht sich bei einem Delphi auf Sachverhalte, bei denen unvollständiges Wissen ein sicheres Urteil erschwert. *„Die bei Delphi-Studien durch die Experten grundsätzlich zu lösende Aufgabe besteht im weitesten Sinne im Beurteilen von Sachverhalten, für die - selbst beim kompetentesten Experten - nicht alle Informationen vorliegen (können), um ein sicheres Urteil zu fällen“* (Häder/Häder 1995: 14).

Bei bisherigen Untersuchungen des Delphiverfahrens wurde meist nicht erforscht, wie die Expert*innen zu ihrem abschließenden Urteil gelangen. Das Ziel der theoretischen Grundlage ist folglich nicht nur eine Basis dafür zu bereiten, warum es Teilnehmenden eines Delphis gelingt zu einem gemeinsamen Urteil bezüglich unsicherer Sachverhalte zu gelangen, sondern auch, ob es sich hierbei um qualitativ hochwertige bzw. fruchtbare Ergebnisse handelt. *„Wenn es möglich würde, theoretisch zu begründen, wie es den Experten gelingt, richtige beziehungsweise bessere Urteile über Sachverhalte zu fällen, über die ihnen zunächst nur unvollständiges Wissen vorliegt, so hätte dies sowohl für die Legitimation von Delphi-Befragungen generell als auch für die Entwicklung und Beurteilung konkreter Delphi-Designs wesentliche Implikationen“* (Häder 2009:43).¹⁶ Die Kognitionspsychologie bietet Erklärungsmodelle, die es ermöglichen die Mechanismen, die bei der Beurteilung ablaufen (auch unter suboptimalen Bedingungen), zu verstehen. Um die Abläufe dieser Beurteilungen transparent darzustellen wird die Urteilsfindung einem komplexen Informationsverarbeitungssystem zugeschrieben (vgl. Häder/Häder 1995: 13f.; Vorgrimler/Wübben 2003: 768). Diese *Theorieklasse der mentalen Modelle* dient dabei vorwiegend dazu, die Basis und den Begriff des Wissens zu bestimmen, auf dem die Expert*innen ihr Urteil bilden können. Wie aber das Urteil über einen spezifischen Sachverhalt gefällt werden kann und welchen Prozess diese Urteilsbildung durchläuft, kann durch eine Erweiterung der zentralen Annahmen der Theorieklasse der mentalen Modelle durch die zentralen Annahmen des *Linsenmodells* aus der *Social Judgment Theory* abgebildet werden (vgl. Hammond 1996: 86f.; Goldstein 2004: 41f.). Dabei ist darauf zu achten, dass es Unterschiede zwischen der ersten Runde eines Delphis (Urteilsbildung) und den folgenden Runden (Urteilsänderung oder Festigung der Urteile) gibt. In einem ersten Schritt sollte somit zuerst bestimmt werden, was unter Wissensbeständen und Urteilsbildung zu verstehen ist.

5.1.1 Wissensbestände und Urteilsbildung

Um die Wissensbestände, auf die Personen während der Urteilsbildung zurückgreifen, zu beschreiben, bilden die *Theorieklasse der mentalen Modelle* einen geeigneten Ansatz. Solche mentalen Modelle zielen auf die Fähigkeit, die Umwelt über die innere Repräsentation als aktive Konstruktion zu deuten, was die Umwelt für das einzelne Individuum verstehbar, erklärbar und - besonders in diesem Zusammenhang - vorhersagbar macht, so dass diese mentalen

¹⁶ Hervorhebungen im Original: Im Original sind die Hervorhebungen allerdings in kursiver Schrift wobei sie in dieser Zitation unterstrichen wurden.

Modelle auch von den jeweiligen Intentionen der Personen abhängig sind (vgl. Häder 2009: 45). „Bei solchen und ähnlichen Fragen wie: Was geschah vorher? Was geschieht danach? Wie funktioniert etwas? wird nach der Sichtweise moderner Kognitionspsychologie Wissen genutzt, das in Form sogenannter mentaler Modelle im Gedächtnis gespeichert ist und als Vorstellung wieder vor dem geistigen Auge sichtbar gemacht werden kann. Diese ‚inneren Bilder‘ können gemustert und abgelesen werden; mit ihrer Hilfe können Prozesse simuliert, Schlußfolgerungen gezogen, handlungswirksame Vorhersagen getroffen und Problemfragen beantwortet werden“ (Martschinke 2001: 71).

Mentale Modelle beruhen damit auf einer Interaktion von Wahrnehmung sowie Gedächtnis und sind direkt von den Intentionen der jeweiligen Informationsverarbeitenden abhängig. Sie sind zudem auch indirekt abhängig von dem zu schließenden Urteil bzw. der zu bewältigenden Aufgabe. Damit schließen mentale Modelle auch individuelles Alltagswissen bei der Lösung eines Problems mit ein. Unter Betrachtung des Delphiverfahrens sind diese mentalen Modelle besonders interessant, da sie Aussagen über eine Urteilsbildung im Status der Unsicherheit beinhalten. Diese Modelle können als Bindeglied zwischen theoretischem Wissen und praktischem Handlungswissen dienen, das stets auch für neue Erfahrungen adaptierbar sein kann. So können sowohl Schlussfolgerungen, als auch deren Veränderungen von den Teilnehmenden erfasst werden, die auf einem unvollständigen Informationsstand basieren (vgl. Häder 2009: 45; Schlemm 2006: 2).

In der Theorieklasse der mentalen Modelle ist für Beurteilungen bei Unsicherheit eine Unterscheidung zweier mentaler Modelle relevant. So kann erstens ein *lokales mentales Modell* vom Individuum abgerufen werden, bei dem die Person auf direkt abrufbares Wissen zurückgreifen kann und damit ihr Urteil als sicher wahrnimmt. Zweitens kann auf ein *probabilistisches Modell* zurückgegriffen werden, bei dem kein sicheres Wissen direkt für die Beurteilung des Sachverhaltes verfügbar ist bzw. direkt von der jeweiligen Person aufgegriffen werden kann. Da das Urteil eines Individuums bei einem *lokalen mentalen Modell* in seinem Gedächtnis auf direkt abrufbaren Informationen basiert, besteht für die jeweilige beurteilende Person in diesem Fall eine hohe subjektive Sicherheit das eigene Urteil betreffend (vgl. Gigerenzer et al. 1991: 506-511; Häder/Häder 1995: 15f. Häder 2009: 46). Bei einem Urteil mittels einer Konstruktion von probabilistischen mentalen Modellen wird alltägliches Erfahrungswissen bzw. Expert*innenwissen aus einem größeren Bewertungszusammenhang herangezogen: „A PMM solves the task by inductive inference, and it does so by putting the specific task into a larger context. A PMM connects the specific structure of the task with a probability structure of corresponding natural environment (stored in long term memory)“¹⁷ (Gigerenzer et al. 1991: 507).

Da in Delphiverfahren meist Urteile unter Unsicherheit gefällt werden, sind hier vorwiegend die probabilistischen mentalen Modelle entscheidend. Auch bei einem Delphi können die Expert*innen anhand der vorliegenden Informationen jeweilig ihre möglichen Informationen oder Referenzklassen bilden, die es den Befragten ermöglichen, sich aufgrund von Wahrscheinlichkeitshinweisen ein Urteil zu bilden (vgl. Häder 2009: 46).

¹⁷ Mit PMM ist ein *Probabilistic Mental Model* gemeint.

Die von den Expert*innen gebildeten hypothetischen Modelle bieten die Grundlage für ihre erste Einschätzung eines Sachverhaltes in der ersten Befragungsrunde eines Delphis.¹⁸ Zunächst beziehen sich die Expert*innen bei der Bildung dieser hypothetischen Modelle auf unterschiedliche Quellen, die in ihrer Wahrnehmung mit dem abgefragten Sachverhalt verknüpft sein könnten. Diese Quellen stammen vorwiegend aus dem beruflichen Umfeld, aber auch aus unterschiedlichen Medien sowie aus der spezifischen sozialen Situation der jeweiligen Expert*innen. Gerade die soziale Situation sorgt dafür, dass Selektivität und Variation in der Wahrnehmung des abgefragten Sachverhaltes zwischen den einzelnen Expert*innen bestehen. So kann es aufgrund der eigenen sozialen Umgebung schon bezüglich der Wahrnehmung des abgefragten Sachverhaltes zu Variationen hinsichtlich des Urteils kommen. Auch das jeweilige Fachwissen der Expert*innen kann je nach Profession bzw. fachspezifischen Hintergrund variieren, so dass die Expert*innen auf unterschiedliche empirische Erkenntnisse und spezifische Theorien zurückgreifen, mit deren Hilfe sie ihre hypothetischen Modelle zur Beurteilung bilden. *„Damit werden auch die von ihnen bei der Bildung probabilistischer Modelle benutzten Erkenntnishilfen, die sogenannten Cues, jeweils fachspezifischen Charakter tragen“* (Häder 2009: 47). Gerade bei einem interdisziplinär zusammengesetzten Expert*innensample können sich diese herangezogenen *Cues* bzw. *Probability Cues*¹⁹ zur Urteilsbildung zwischen den jeweiligen Einschätzungen deutlich unterscheiden und folglich zu unterschiedlichen Beurteilungen des Sachverhaltes führen. In die probabilistischen mentalen Modelle fließen - wie bereits oben erwähnt - zudem die Intentionen der Expert*innen ein, so dass auch diese zu unterschiedlichen Einschätzungen zwischen dem ersten Urteil führen können. So kann beispielsweise der eigene politische Standpunkt einer jeweiligen Expertin bzw. eines jeweiligen Experten auch bei der Beurteilung eines Sachverhaltes eine Rolle spielen (vgl. Häder/Häder 1995: 18f.; Häder 2009: 46f.). *„Selbst ein ähnlich strukturiertes Fachwissen wird vor dem Hintergrund unterschiedlicher politischer Standpunkte und unterschiedlicher wirtschaftlicher Interessen zu divergierenden Urteilen führen“* (Häder/Häder 1995: 19). Damit sollte eine Methode der Expert*inneneinbindung eben auch nicht versuchen *den einen Experten* einzubinden, sondern eine möglichst umfassende Spannbreite - idealiter die gesamte Spannbreite an Positionen und Meinungen - in die Abfrage integrieren (vgl. Niederberger/Wassermann 2015: 11).

Damit bildet die Theorieklasse der mentalen Modelle besonders über die probabilistischen mentalen Modelle ab, warum Expert*innen in ihren Urteilen auseinanderliegen können. Geht man davon aus, dass eine Person ein lokales mentales Modell abrufen kann, so wird dadurch auch die Funktionslogik eines Delphis begründet, in dem es über das Feedback und die weiteren Runden darum geht, dieses lokale mentale Modell zu identifizieren und als Bewertungsgrundlage der anderen eingebundenen Teilnehmenden ebenso in ihre mentalen Modelle zu integrieren. Auch kann so ein Austausch über das wahrscheinlichste bzw. fruchtbarste mentale Modell erfolgen, wenn aufgrund hoher Unsicherheit kein lokales mentales Modell vorhanden ist, wovon bei Delphibefragungen aufgrund der Zielbestimmungen – wie in diesem Kapitel bereits beschrieben - häufig auszugehen ist. Die Vielschichtigkeit der zu verarbeitenden Wissensbestände, die sich in den mentalen Modellen sammeln und überdies unter den eingebundenen Personen ausgetauscht werden müssen, benötigt allerdings nicht nur eine Rechtfertigung der Funktionslogik des Delphiansatzes, sondern sollte ebenfalls ein theoretisches Modell

¹⁸ Bei einem Gruppendelphi kann dies auch für die Vorbefragung gültig sein, falls eine solche stattfindet.

¹⁹ Mehr dazu in Gigerenzer et al. 1991: 507f.

beinhalten, dass den Weg rationaler Urteilsfindung aufzeigt und für die Logik des Delphis nutzbar macht. Nicht zuletzt auch um *fruchtbare* Ergebnisse bestimmbar zu machen.

Die Ausgangslage der Urteilsbildung knüpft damit an die bestehende Notwendigkeit der Bildung mentaler Modelle an. *„Wir können Objekte, Ereignisse und Zustände in der Welt nie „direkt“ erfahren. Die sensorische Information erlaubt nur eine Konstruktion einer mentalen Repräsentation von Objekten, Ereignissen und Zuständen. (...) Diese mentale Repräsentation ist Grundlage für schlußfolgernde Urteile“* (Jungermann et al. 2005: 164). Allerdings lassen sich Urteile näher eingrenzen, denn sie münden letztlich in Entscheidungen, die sich besonders dadurch kennzeichnen, dass sie im Allgemeinen auf konfliktbewusstes, abwägendes und zielorientiertes Handeln ausgelegt sind. Überdies werden Urteile dann notwendig, wenn sie entweder durch eine unsichere Informationslage oder einen unsicheren zu erwartenden Ausgang gekennzeichnet sind. Sie können sich aber auch aufgrund von Unsicherheit und Bedenken gegenüber den Präferenzen einer anderen Person begründen (vgl. Pitz/Sachs 1984: 140; Jungermann et al. 2005: 3).

Urteile werden häufig durch einen bestimmten Grad an Unsicherheit geprägt, bei dem gerade die möglichen Konsequenzen der Optionen unsicher sind, da die Konsequenzen auch von externen und damit für die Person nicht kontrollierbaren Ereignissen abhängig sind. Damit gibt es als Basis für ein Urteil unter Unsicherheit sowohl interne wie auch externe Ursachen. Dabei lassen sich besonders die externen Ursachen von Unsicherheit nicht gänzlich ausblenden, was letztlich einen natürlichen Prozess abbildet. *„Planning under uncertainty demands - and gets - human judgment (...)“* (Hammond 1996: 14). Dabei kann sich das menschliche Urteil an bestimmten festen Größen orientieren, die in globalen Wissensbeständen vorhanden sind, so wie etwa mathematische Regeln oder Naturgesetze. An diesen Referenzklassen orientieren sich Personen, um mit der bestehenden Unsicherheit umgehen zu können und, gekoppelt mit Intuition, zu einem individuellen Urteil zu gelangen (vgl. Adelman et al. 1975: 140; Hammond 1996: 14-20; Jungermann et al. 2005: 142).

Der Umgang mit Unsicherheiten hinsichtlich der Urteilsfindung wird in der Wissenschaft häufig mit dem Begriff der Wahrscheinlichkeit verknüpft. Eine Orientierung an Wahrscheinlichkeiten beinhaltet aber auch immer einen gewissen Grad an subjektiver Deutung. *„Nach der subjektivistischen Interpretation ist eine Wahrscheinlichkeitsangabe der Ausdruck eines rein subjektiven „Grades an Gewissheit“; insofern kann ein Wahrscheinlichkeitsurteil auch nicht richtig oder falsch sein. Unterschiedliche Personen können in bezug auf dasselbe Ereignis zu verschiedenen Urteilen kommen“* (Jungermann et al. 2005: 150). Um solche Wahrscheinlichkeiten bilden zu können, beziehen sich Personen auf Hinweisreize (*Cues*), womit der Wissensbestand letztlich auf den probabilistischen mentalen Modellen aus eben diesen *Probability Cues* gebildet wird. Die Qualität des Urteils hängt hier stark von der Validität der herangezogenen Hinweisreize ab (vgl. Goldstein 2004: 38; Jungermann et al. 2005: 177).

Um ein unverzerrtes wahrscheinliches Urteil zu erreichen, müssen sowohl die globale Umwelt sowie auch eine angemessene Erfassung und interne Präsentation des Problems erfolgen. *„But the task of achieving a good representation often relies on inside judgments - grouping units by similarity, determining relevance by causality, weighting alternatives by associate strength. And each of these factors can inhibit our ability to see the wider field. The outside view requires that we transcend the specific case, and yet it is the specific case, and all its*

properties, that make our representations of the world so compelling” (Lagnado/Sloman 2004: 173). Damit basieren Urteile auf einem Zusammenspiel aus internen sowie externen Hinweisreizen und Deutungen und besitzen Elemente die sowohl auf analytischen wie auch auf intuitiven Referenzklassen aufbauen (vgl. Adelman et al. 1975: 140). Ist eine Person aber mit der Bildung solcher Referenzklassen vertraut, sollte sie auch probabilistische Modelle erfolgreich verwenden können. Diese Annahme stützt sich auch auf die Erkenntnisse der *Social Judgment Theory*, in der der Zusammenhang zwischen der individuellen Nutzung von Hinweisreizen und die ökologische Validität in logischen Zusammenhang gesetzt werden und der Urteilsprozess von Individuen damit nachvollziehbar gemacht werden kann (vgl. Jungermann et al. 2005: 177). Denn letztlich bemessen sich fruchtbare Ergebnisse eines Delphis besonders daran, dass aus globalen Wissensbeständen unter Berücksichtigung der bestehenden Unsicherheiten *rationale Urteile* getroffen werden. Wie solche rationalen Urteile erreicht werden können, soll folgend beschrieben werden, denn „(...) *at least in certain circumstances, human judgment is rational and our judgments are rational*” (Hammond 1996: 7).

5.1.2 Überführung globaler Wissensbestände in rationale Urteile

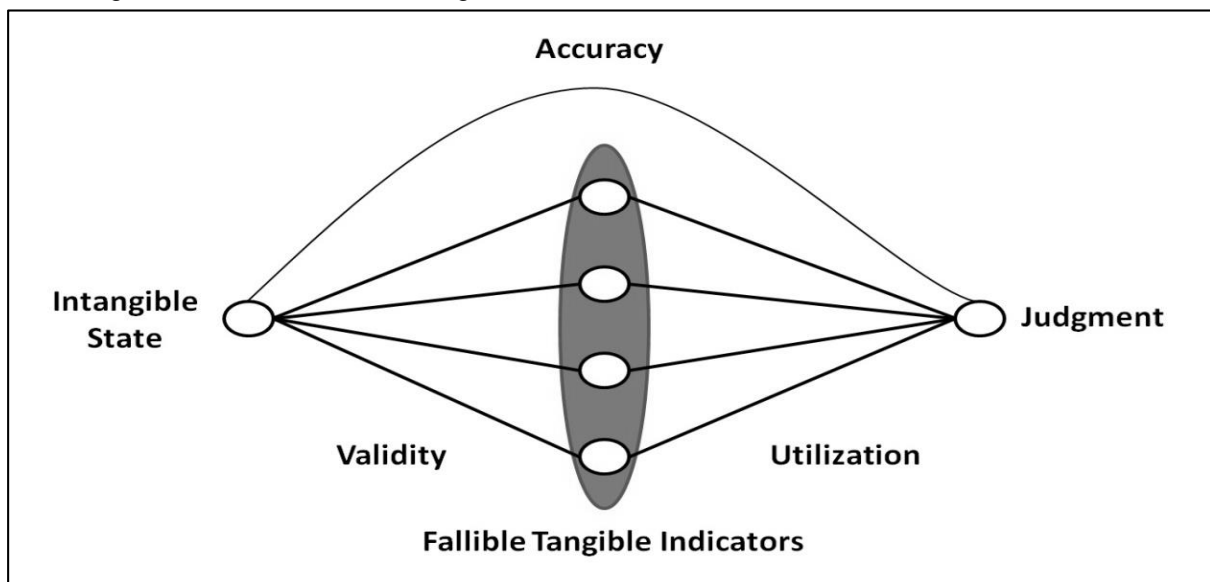
Der Umgang mit Unsicherheit basiert auf einer beständigen Rivalität zwischen Intuition und Analyse. Dabei liegt die lange bestehende Rivalität darin begründet, dass Intuition generell mit solchen Mythen wie Kreativität, Vorstellungskraft und bildlicher Repräsentation von Ideen verbunden wird, wohingegen die Analyse mit Logik, Mathematik und nachvollziehbaren Gedankengängen assoziiert wird (vgl. Hammond 1996: 60f.). Wird ein Urteil unter bestehender Unsicherheit notwendig, muss diese Rivalität überwunden werden und die Hinweisreize sinnvoll verknüpft werden, um die globalen Wissensbestände für eine rationale Entscheidung nutzbar zu machen. Die Unsicherheit bedingt, dass die getroffenen Urteile immer Gefahr drohen fehlerhaft (*fallible*) zu sein (vgl. Hammond 1996: 86), da die Abbildung der eigenen Person im Vergleich zu der globalen Umwelt immer zu einem gewissen Grad unzuverlässig, ambig oder zweideutig ist. „*At least, this was the argument for individual cues and means. The extent to which an organism could combine cues and select means to overcome their limitations when taken singly is, of course, the problem of achievement – the main subject of research*” (Goldstein 2004: 43). Der Fokus liegt damit auf den individuellen Hinweisreizen. Dabei spielt die Fehlerhaftigkeit dieser für eine analytische Betrachtung eine wichtige Rolle, denn diese *fehlerhaften Hinweisreize* stellen eine direkte Verbindung zur natürlichen Umwelt her (vgl. Hammond 1996: 86; Goldstein 2004: 42f.).

Um diesen Umstand in einem theoretischen Modell abzubilden hat Brunswik das *Linsenmodell (lens model)* entwickelt, das ebenso in direkter Anknüpfung in Hammonds *Social Judgment Theory* eine zentrale Rolle einnimmt (vgl. Pitz/Sachs 1984: 144). Dieses Linsenmodell ordnet das einzelne Individuum in eine natürliche Umwelt ein. Dabei werden im Modell sowohl die Variabilität der fehlerhaften Hinweisreize sowie auch ein Fokus auf Informationen innerhalb des individuellen kognitiven Systems gelegt, die wiederum Rückschlüsse auf externe Objekte zulassen. „*The lens model (...) represents an uncertain world represented by many fallible indicators and an organism that has the capacity to integrate them - without awareness - into a judgment that displays remarkable accuracy in visual perception and various degrees of accuracy in other circumstances. It is our lack of awareness of how we integrate information that makes it an intuitive process*” (Hammond 1996: 87).

Die verschiedenen Adaptionen des Linsenmodells können durchaus komplex sein und wurden in vielfacher Weise dargestellt (vgl. Wittmann 2002: 165; Goldstein 2004: 42; Jungermann 2005: 165). Diese komplexen Ausarbeitungen des Modells folgen dabei dem Ziel, die Korrelationen der fehlbaren Hinweisreize darstellbar und besonders auch in weiteren Schritten mittels mathematischer Formeln quantitativ abbildbar zu gestalten. Zu diesem Zweck wird ein Beziehungsgeflecht zwischen dem zu beurteilenden Kriterium über die fehlbaren Hinweisreize hin zu dem personellen Urteil konstruiert. Die Exaktheit des Urteils wird dabei mittels der Werte der Korrelation zwischen Urteil und Kriterium bemessen (vgl. Goldstein 2004: 41f.).

Um die Grundfrage, warum auch Expert*innen unterschiedliche Urteile hinsichtlich spezifischer Sachfragen fällen, zu betrachten, lohnt es sich die vereinfachte schematische Darstellung des Linsenmodells heranzuziehen, da in dieser alle zentralen Elemente der Funktionslogik bereits enthalten sind. Dieses in Abbildung 3 abgebildete Linsenmodell repräsentiert einen Organismus in einer Umwelt, die diesem Organismus eine Vielfalt an fehlbaren Hinweisreizen bietet. Zentrales Element dieses Modells ist der Fokus der Informationen in einem kognitiven System, das Schlussfolgerungen zu externen Objekten zulässt (vgl. Hammond 1996: 86f). Im Linsenmodell finden sich auf der linken Seite die globalen Wissensbestände der Umwelt, die durch nicht greifbare Zustände (*Intangible State*) gekennzeichnet sind. Der fehlende direkte Zugriff des Individuums auf diese Wissensbestände bedingt dabei den allgemeinen Zustand der Unsicherheit. Das Individuum findet sich auf der rechten Seite der schematischen Darstellung wieder, in dessen Feld sich letztlich das von der Person gebildete Urteil (*Judgment*) findet. In der Mitte des Linsenmodells finden sich die fehlbaren Hinweisreize (*Fallible Tangible Indicators*), die die Verknüpfung zwischen der Umwelt und dem individuellen Urteil bilden. Je exakter dieses Urteil ist (*Accuracy*), desto eher nähert sich das Individuum einem rationalen Urteil an. In der schematischen Darstellung wäre dies der oberste fehlbare Hinweisreiz der dem exakten Urteil am nächsten liegt.

Abbildung 3: Schematische Darstellung des Linsenmodells



Quelle: Eigene Darstellung nach Hammond 1996: 87.²⁰

²⁰ Um das namensgebende Element des Linsenmodells zu veranschaulichen, wurde in diesem Schaubild eine abgebildete Linse eingefügt, die sich in der schematischen Abbildung von Hammond nicht findet (vgl. ebd. 1996:

Die Beziehung zwischen dem nicht direkt greifbaren Zustand der Umwelt und den Hinweisreizen lässt sich über Validität (*Validity*) herstellen, während die Ableitung der fehlbaren Hinweisreize zu einem Urteil verwertet wird (*Utilization*). So illustriert das Modell, „(...) *how a person's judgment of an intangible object, event, or state is made on the basis of multiple fallible indicators (left side of the lens). The right side of the lens shows the utilization of several indicators*“ (Hammond 1996: 87). Das Linsenmodell stellt letztlich einen fokalen Zusammenhang zwischen den einzelnen Variablen her - die Beziehungen zwischen den globalen Wissensbeständen werden durch eine Linse zu einem subjektiven Urteil fokussiert (vgl. Hammond 1996: 86f.; Goldstein 2004: 42f.; Jungermann et al. 2005: 164).

Das Linsenmodell zeigt auf theoretischem Erklärungsweg zwei Perspektiven auf, die sowohl für das klassische Delphiverfahren als auch für das Gruppendelphi relevant sind. Zum einen zeigt es warum auch Expert*innen nicht immer einen direkten Zugriff auf rationale Urteile haben können und zum zweiten bietet es einen analytischen Weg um zu verstehen, wie Urteile gefällt werden (Urteilsbildung) und wie sie in Einklang gebracht werden können (Urteilsänderung). Dabei kommt es besonders darauf an, den relevanten Hinweisreiz zu identifizieren, der sich einem rationalen Urteil bei bestehender Unsicherheit so weit wie möglich annähert - oder idealiter diesen sogar trifft. Denn die „(...) *Wahrscheinlichkeit sich zu irren, hängt von der Validität des herangezogenen Hinweisreizes ab*“ (Jungermann et al. 2005: 177).

So haben die Teilnehmende die Möglichkeit, sich durch Bezug zu Referenzklassen einem rationalen Urteil anzunähern, da sie die Fähigkeit besitzen eine angemessene Bestimmung des Problems vorzunehmen und eine Perspektive auf die jeweiligen Beziehungen zwischen den relevanten Einheiten einnehmen zu können. Dieser Prozess besteht allerdings häufig in einer individuellen Urteilsfindung, bei der die verschiedenen Hinweisreize nach Ähnlichkeit gruppiert, Relevanzen anhand von Kausalbeziehungen hergestellt und Alternativen durch assoziative Gewichtung gegeneinander abgewogen werden (vgl. Lagnado/Sloman 2004: 173). Eine Verknüpfung von Analyse und Intuition kann zu einem rationalen Urteil und zu einem rationalen Ergebnis eines Urteilsprozesses führen. Welche Rolle dabei das Feedback und der Austausch zwischen den Teilnehmenden während eines Delphis spielen können soll im folgenden Abschnitt vertieft werden. In diesem Abschnitt wird der Fokus auf der Annäherung der Urteile liegen, welche eines der Ziele des Einsatzes eines Delphis bildet. Letztlich geht es darum der bestehenden Rivalität zwischen Analyse und Intuition zu begegnen, da beide bei Unsicherheit eine gewisse Notwendigkeit besitzen, um überhaupt ein Urteil fällen zu können. „*The rivalry will continue, as it should. Nevertheless, as long as the rivalry does continue, the study of human judgment must attempt to cope with the question of the competence of human judgment, whether analytical or intuitively based*“ (Hammond 1996: 93).

5.1.3 Urteilsänderung und Annäherung der Urteile

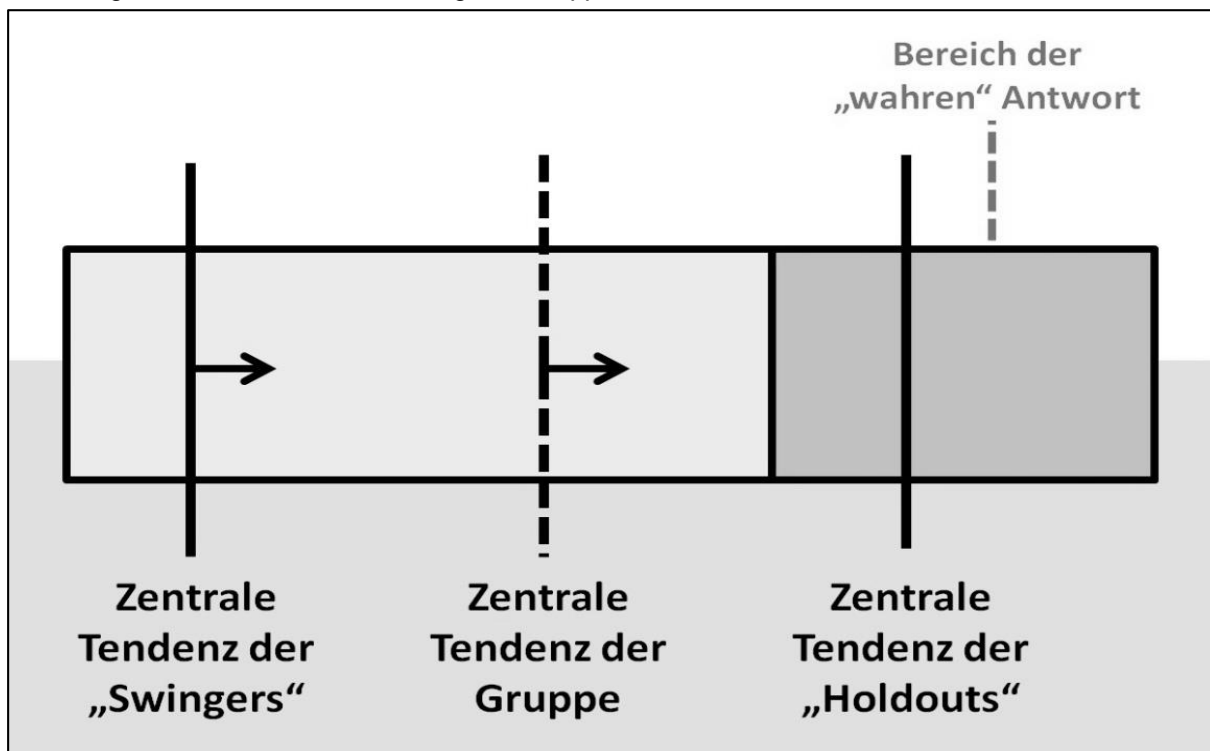
Das Ziel eines Delphis folgt dem ursprünglichen Gedanken des *n-head-rules* (vgl. Dalkey 1969: 6). Um dieser Logik zu entsprechen werden die Teilnehmenden durch das Feedback über die Urteile der anderen Teilnehmenden informiert. Ziel in dieser Konzeption ist es, durch die Rückmeldung und erneute Befragung eine Annäherung an den wahren Wert zu erreichen.

87). Hammond verwendet zudem anstatt des Begriffes der *cues* den Begriff der *indicators* (vgl. ebd. 1996: 86). Beide Termini werden in dieser Arbeit als *Hinweisreize* bezeichnet.

Dafür ist es bei Uneinigkeit notwendig, dass ein Teil der Teilnehmenden ihr ursprüngliches Urteil ändert. Wie bereits zu Beginn dieses Kapitels angesprochen, treten bei einem Delphi sowohl *Swingers* als auch *Holdouts* auf und die Gruppeneinschätzung nähert sich mit der Anzahl der Runden immer mehr einem gemeinsamen Wert an.

Wie in Abbildung 4 dargestellt, nähert sich damit die zentrale Tendenz der *Swingers* in Richtung der zentralen Tendenz der *Holdouts*, die sich innerhalb des angenommenen Bereichs einer wahren Antwort befinden. So bewegt sich letztlich auch die zentrale Tendenz der Gruppe in Richtung des wahren Wertes und erreicht diesen *wahren* Wert - oder nähert sich diesem zumindest an (vgl. Rowe et al. 1991: 238). Kritiker des Delphiverfahrens sehen hier allerdings lediglich eine Tendenz zum Gruppenmittel und keine sachgerechte Annäherung an ein wissenschaftsbasiertes Urteil. Andere Kritiker ziehen den Grad des Dogmatismus einer Person zur Erklärung heran, ob eine Person das Urteil ändert oder nicht. Damit würde sich im Endeffekt dogmatische Personen als *Holdouts* gegenüber den weniger dogmatischen Personen - den *Swingers* - durchsetzen (vgl. Vorgrimler/Wübben 2003: 769).

Abbildung 4: Theoretische Annäherung des Gruppenurteils



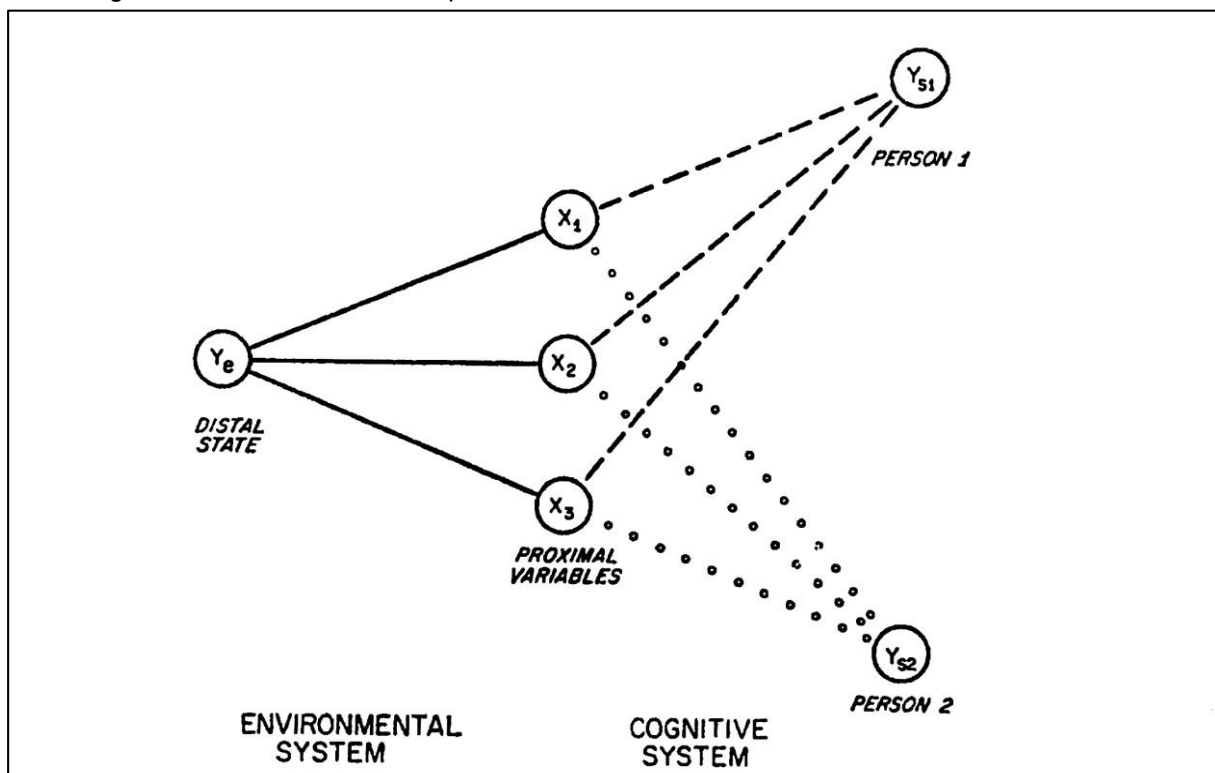
Quelle: Eigene Darstellung nach Rowe et al. 1991: 38

Das Feedback stellt allerdings eine zusätzliche Information dar, die an die Befragten herangebracht und von diesen interpretiert werden muss. So wird also nicht nur eine zusätzliche Suche nach neuen Informationen im Gedächtnis durch das Feedback angestoßen, sondern die Expert*innen müssen diese neuen Informationen im Vergleich zu ihrem eigenen vorhergegangenen Urteil prüfen. Es gibt damit vielfältigere Handlungsmotive aufgrund derer die Expert*innen eine Einschätzung teilen können oder diese ablehnen. Das Feedback bietet eine weitere Orientierungshilfe für das eigene Urteil, setzt aber nicht, wie in der Kritik geäußert, zwangsläufig zu einer blinden Orientierung am Gruppenmittel bzw. den *Holdouts* an. So orientieren sich die teilnehmenden Expert*innen nicht (nur) an Konsens und Gruppenurteil, sondern folglich auch an ihren eigenen auf den herangezogenen Hinweisreizen basierenden Einschätzungen und

gleichen das gelieferte Feedback mit diesen ab. Es kann also nicht davon ausgegangen werden, dass ab der zweiten Runde eine Bewegung der Urteile lediglich anhand der Rückmeldung hinsichtlich der Verweilenden erfolgt (vgl. Scheibe et al. 1975: 262-266; Häder/Häder 2000: 26f.)

Eine solche Begründung der Wirksamkeit des Feedbacks und der Wiederholung der Befragung ist allerdings letztlich rein deskriptiv und bietet noch kein theoretisches Erklärungsmodell für eine Urteilsänderung einzelner Teilnehmenden. Um eine mögliche Urteilsänderung theoretisch zu begründen, kann auch hier auf das Linsenmodell zurückgegriffen werden, dass die Validität eines rationalen Urteils in Verbindung mit dem Hinweisreiz setzt (vgl. Jungermann et al. 2005: 177). Hierbei können mittels der notwendigen Kombination von Analytik und Intention durch Heranziehen unterschiedlicher Hinweisreize Personen zu jeweilig anderen Urteilen eine spezifische Situation betreffend gelangen. „*Social judgment theory provides a theoretical framework for understanding such disagreement because it focuses primarily on the cognitive representations of relations between the variables (including their interdependencies) in the environment*“ (Adelman et al. 1975: 139f.). Wie ein solcher Widerspruch zwischen den Einschätzungen von zwei Personen entstehen kann, zeigt beispielhaft das in **Fehler! Verweisquelle konnte nicht gefunden werden.** dargestellte Schaubild auf.

Abbildung 5: Linsenmodell eines interpersonalen Konfliktes



Quelle: Adelman et al. 1975: S.139

Wie bereits in der schematischen Darstellung des Linsenmodells in Abbildung 3 im vorhergehenden Abschnitt, findet sich auch in diesem Schaubild die Umwelt mit ihren globalen Wissensbeständen auf der linken Seite. Der nicht greifbare Zustand (*Distal State*, Y_e) bildet die Zielgröße, der sich beide Personen mit ihrem jeweiligen Urteil (Y_{s1} , Y_{s2}) über die Hinweisreize (X_1 , X_2 , X_3) annähern möchten. Dem ursprünglichen Linsenmodell folgend findet im kognitiven

System auf der rechten Seite des Schaubilds eine Interpretation der verschiedenen Hinweisreize durch die Personen statt. Neben dieser Interpretation werden die Hinweisreize auch in Relation gesetzt und von der jeweiligen Person bewertet und gewichtet. Dies kann dazu führen, dass eine Person ihr Urteil (Y_{s1}) vorwiegend auf einen anderen Hinweisreiz (z.B. X_1) begründet als es die zweite Person bei ihrem Urteil (Y_{s2}) tut, die dieses auf Heranziehen von zwei anderen Hinweisreizen (X_2, X_3) bildet. Konflikte können dadurch besonders kognitive Quellen haben, „(...) *in that people can have shared goals but still differ in their assessments of the situation and the appropriate action*“ (Goldstein 2004: 48). Das Modell schließt allerdings ebenso wenig aus, dass beide Personen zu einem identischen Urteil gelangen können, indem sie die Hinweisreize gleich interpretieren und verwerten (vgl. Adelman et al. 1975: 139f.).

Die Darstellung in **Fehler! Verweisquelle konnte nicht gefunden werden.** zeigt damit nicht nur, wie Personen zu unterschiedlichen Urteilen gelangen können, sondern bietet darüber hinaus einen theoretischen Zugang dazu, dass eine Urteilsänderung möglich sein kann. So bleibt ein Urteil einer Person nicht statisch und der Wissensbestand, der sich über die in den mentalen Modellen zusammengesetzten Hinweisreizen, kann sich ändern. Dies kann über das Feedback eines Delphis erfolgen oder durch direkte Kommunikation zwischen den Personen, in denen neue Hinweisreize (z.B. Argumente) zu einem neuen Urteil und idealiter zu einem rationalen Urteil führen. Damit hängt ein Gruppenurteil nicht nur von den einzelnen individuellen Urteilen der jeweiligen Mitglieder dieser Gruppe ab, sondern überdies, wie gut die einzelnen Urteile miteinander abgeglichen werden. So kann ein gemeinsames Urteil erst kommuniziert, dann aggregiert und letztlich auch konsensual gebildet werden. Die Güte des Urteils hängt letztlich davon ab, ob die Teilnehmenden identische Hinweisreize verwenden und diese auf gleiche Weise gewichten. Dabei brauchen die Gruppenmitglieder eine gewisse Struktur und Hilfestellung um diesen Austausch zu gewährleisten (vgl. Wittmann 2002: 165; Baron 2004: 19; Sorkin et al. 2004: 465).

Die Darstellung des Linsenmodells veranschaulicht, dass globale Wissensbestände bei unsicheren Sachverhalten nicht immer über rein analytische Verfahren darstellbar sind. „*In solchen Situationen (...) dürfte Unsicherheit sich eher durch einen konstruktiven kognitiven Prozeß herstellen, in dessen Verlauf auf quantitative Informationen (beispielsweis Häufigkeiten) ebenso wie auf qualitative Informationen (beispielsweise Beschreibungen oder Argumente) zurückgegriffen wird. Das Interesse der Forschung verschiebt sich von probability-based reasoning zu knowledge-based reasoning*“ (Jungermann et al. 2005: 164). Eine Kritik am klassischen Delphiverfahren knüpft an das Fehlen von qualitativen Informationen während dieses Prozesses an. So wird kritisiert, dass die Überprüfung der individuellen mentalen Modelle, die durch die Auseinandersetzung mit fremden Argumenten hervorgerufen werden können, aufgrund der Anonymität des statistischen Feedbacks nicht ausreichend stattfinden kann. Es kann also nicht ausgeschlossen werden, dass die Expert*innen ein nicht ausreichend durchdachtes Urteil abgeben, da sie für dieses nicht persönlich verantwortlich gemacht werden können und es überdies ebenso nicht argumentativ begründen müssen. (vgl. Bowles 1999: 32; Vorgrimler/Wübben 2003: 767).

Diesen theoretischen Annahmen folgend, sollte der argumentative Austausch als diskursives Element eines Gruppendelphis eine Annäherung an rationale Expert*innenurteile unterstützen. Die Methode des Gruppendelphis benötigt damit allerdings neben der wissenschaftsthe-

oretischen Grundlage eines klassischen Delphis folglich eine erweiterte theoretische Legitimation der diskursiven Prozesse und deren vermuteten positiven Effekte. Die theoretische Legitimation muss folglich um eine weitere konzeptionelle theoretische Komponente ergänzt werden. Neben der Erklärung der Prozesse einer rationalen Urteilsfindung benötigt das Gruppendelphi auch eine diskurstheoretische Begründung der zusätzlichen Prozesse während des argumentativen Austauschs der Expert*innen und ob diese eine qualitative Verbesserung der Urteile bewirken können. Dem soll im folgenden Rechnung getragen werden.

5.2 Erweiterung des theoretischen Rahmens durch die Theorie des kommunikativen Handelns

Haben das klassische Delphiverfahren und das Gruppendelphi bisher eine genuine theoretische Grundlage, sollte diese für das Gruppendelphi durch zusätzliche Annahmen ergänzt werden, da der iterative Prozess im Falle des Gruppendelphis um diskursive Elemente erweitert wird (vgl. Niederberger/Kuhn 2013: 58; Kuhn et al. 2014: 66f.).

Diese Ergänzung basiert auf der bereits beschriebenen Annahme, dass das klassische Delphiverfahren sich als hilfreich erwiesen hat, um Expert*innen in einer Art demokratischem Gruppenprozess dazu zu befähigen, zu einem gemeinsamen *best guess* über zukünftige Entwicklungen zu gelangen (vgl. Webler et al. 1991: 254), wie auch eine Ideenaggregation zu betreiben, Expert*innenmeinungen zu ermitteln und nach Häder - auch als eigener Typ beschrieben - einen Konsens zu erreichen (vgl. Häder 2009: 36). Der Delphiansatz ist allerdings nicht beliebig auf alle Probleme und Fragestellungen anwendbar, sondern besonders auf solche, die eine Vermischung von wissenschaftlichen Bewertungen und normativen Wertungen notwendig machen: „*The Delphi process is not suited to all types of problems. it is best suited for those that involve a mixture of scientific evidence and social values*“ (Webler et al. 1991: 256). Gerade bei der Bewertung zukünftiger Entwicklungen, steht damit neben den objektiven Annahmen häufig auch die Frage nach dem *Was wollen wir?* im Raum. Sobald Annahmen und Bewertungsmodelle über zukünftige Ereignisse nebeneinander oder gar in konträre Richtungen führen, folgt aus der Frage *Was ist zu tun?* meist zwangsläufig ein Austausch über die jeweiligen zu Grunde liegenden Wertdimensionen (vgl. Rescher 1969: 2). Dies führt zur Theorie des kommunikativen Handelns von Habermas, in der die Verknüpfung von einem gemeinsamen Verständnis von Handlungen basierend auf Wissen über die möglichen Konsequenzen mit eben jenen Werten verknüpft und idealiter im Diskurs gelöst werden (vgl. Renn 2000a: 20).

5.2.1 Begriffsbestimmung: Diskurs und Konsens

Der Begriff des Diskurses bezeichnet keine rein beliebige Form der Kommunikation oder wie es Grunwald in einem Vortrag im Jahr 2008 treffend bemerkte: „*Diskurs ist kein Gerede*“.²¹

²¹ So Armin Grundwald in einem Vortrag zu den ethischen Anforderungen bei der Endlagerung nuklearer Abfälle am 31. Oktober 2008. Abrufbar: http://www.bmub.bund.de/fileadmin/bmu-import/files/pdfs/allgemein/application/pdf/endlagersymp_vortrag_grunwald.pdf.

Dies deutet bereits darauf hin, dass ein Diskurs einer klaren Konzeption folgt. Nach der Konzeption von Habermas ist ein Diskurs eine Form der Kommunikation, die durch Argumentationen gekennzeichnet ist und „(...) *in der problematisch gewordene Geltungsansprüche zum Thema gemacht und auf ihre Berechtigung hin untersucht werden*“ (Habermas 1973: 214). Ziel eines Diskurses ist dabei nicht, wie beispielsweise in einem Streitgespräch, die eigenen Interessen im Sinne der Nutzenmaximierung durchzusetzen, sondern vielmehr ein gemeinsam getragenes Ergebnis, basierend auf einem geteilten Einverständnis über Wertvorstellungen und moralische Standards, wie auch mögliches Wissen über die Konsequenzen zu erzielen (vgl. Renn 2000a: 20).

Der Ausgangspunkt des Konzeptes von Habermas ist eine Reflexion der Kommunikationssituation, die für eine Begründung von Aussagen notwendig ist. Gültig sind in diesem Zusammenhang nur diejenigen Aussagen, die sich gegenüber den anderen vorgebrachten Einwänden rechtfertigen – also begründen – lassen: „*Die Logik der Argumentation erfordert einen begrifflichen Rahmen, der dem Phänomen des eigentümlichen zwanglosen Zwang des besseren Argumentes Rechnung zu tragen erlaubt*“ (Habermas 1981: 52f.). Ein Diskurs bietet einen solchen Rahmen für die Prüfung der vorgebrachten Einwände. So zeichnen sich Diskurse dadurch aus, „(...) *daß die ihnen zugrundeliegenden Fragen prinzipiell durch für alle nachvollziehbare Gründe entschieden werden können*“ (Gottschalk 1999: 6). Besonders pluralistische Gesellschaften bilden die Eigenschaft aus, dass gemeinsame Grundlagen in Frage gestellt bzw. problematisiert werden, wodurch ein Diskurs über eben diese Grundlagen entsteht, also eine durch Argumentation gekennzeichnete Form der Kommunikation (vgl. Gottschalk 1999: 4f.; Keller et al. 2011: 12).

Aufgrund der, für die Entstehung einer Situation des Diskurses gesellschaftlichen Prozesse, muss die Argumentationstheorie ein umfassendes, nicht bloß auf Wahrheitsgeltung eingeschränktes Konzept für die Gültigkeit eines Diskurses anwenden. Das heißt im Umkehrschluss allerdings nicht, dass ein Diskurs darauf verzichten sollte ein wahrheitsanaloges Geltungskonzept anzustreben. Dieses Konzept muss aber die kritischen, raumzeitlichen und sozialen Beschränkungen beachten, die genau dieses Konzept zu einem späteren Zeitpunkt möglicherweise wieder anstreitbar machen können (vgl. Habermas 1981: 56). Weiter ergibt sich für die Gültigkeit eines Diskurses, dass nicht bloß ein *einziges* Konzept als wahr herausgestellt werden kann (vgl. Jäger 2008: 379f.). Dies führt zu einer Situation, in der im Diskurs nach der von allen akzeptierten bzw. idealiter nach der besten Lösung gesucht wird, was unter den oben genannten Beschränkungen als wahrheitsanaloges Geltungskonzept verstanden werden kann. So entsteht die Erwartung an den Diskurs, dass die darin getroffenen Aussagen verständlich formuliert und wahr sein sollten. Weiter sollen die Sprechenden bei ihren Äußerungen Wahrhaftigkeit an den Tag legen und das Geäußerte muss (soweit objektiv begründbar) richtig sein. Diese Eigenschaften an einen kommunikativen Austausch können von Diskursen bereitgestellt werden, denn Diskurse sind auf diesem Wege „(...) *die Fortsetzung des normalen kommunikativen Handelns mit anderen Mitteln, nämlich organisierte (Diskussions-)Prozesse argumentativer Auseinandersetzungen*“ (Keller 2011: 18).

Um zu einem qualitativ hochwertigen Ergebnis zu gelangen, muss ein Diskurs folglich mit besonderen Ansprüchen und unter notwendigen Voraussetzungen erfolgen: „*Von >Diskursen< will ich nur dann sprechen, wenn der Sinn des problematisierten Geltungsanspruches die Teil-*

nehmer konzeptuell zu der Unterstellung nötig, daß grundsätzlich ein rational motiviertes Einverständnis erzielt werden könnte, wobei >>grundsätzlich<< den idealisierenden Vorbehalt ausdrückt: wenn die Argumentation nur offen genug geführt und lange genug fortgesetzt werden könnte“ (Habermas 1981: 71). Wollen sich Teilnehmende nicht auf die Argumente der anderen einlassen, so kann ein Diskurs auch bei idealen Strukturen nicht zu einem konsensualen Ergebnis führen.

Neben diesem Vorbehalt spielt auch die Zeitdimension für den eigentlichen Prozess des Diskurses eine entscheidende Rolle. Nicht während jedes Diskurses ist genügend Zeit vorhanden, um alle notwendigen Argumentationen vorzubringen und gemeinsam voranzutreiben.²² Das jeweilige Gewicht einzelner Argumente kann überdies nur dann verstanden werden, wenn man die jeweiligen den Argumenten zu Grunde liegenden normativen Ansprüche und Ziele der Teilnehmenden - nach Habermas den *zugrundeliegende Sinn des Unternehmens* - versteht (Habermas 1981: 57). Auch diese Aussagen müssen begründet werden, damit die Teilnehmenden *konsensual* zu einem *wahrheitsanalogen Geltungskonzept* gelangen können. Gerade hier liegt trotz der vielfältigen Voraussetzungen, die an einen Diskurs gestellt werden, auch seine Stärke: *„Der Diskurs setzt gerade an dem Punkt an, an dem viele andere Entscheidungsverfahren versagen: Der Begründung von normativen und wertenden Aussagen und der Integration von pluralen Wertmustern in ein konsensfähiges Bewertungsraster (...)“* (Renn 2007: 175). Renn ist allerdings im Gegensatz zu Habermas der Ansicht, dass es für moralische Probleme auch mehrere gleichberechtigte Lösungen geben kann. Ein praktischer Diskurs kann damit nicht zum Ziel haben die einzige richtige normative Lösung eines moralischen Problems zu identifizieren. Dies gilt besonders beim Handeln unter Unsicherheit, wenn die Folgen und möglichen Nebenfolgen nur in Ansätzen oder gar gänzlich unbekannt sind (vgl. Renn 1999: 79).

Neben dem zu den kritischen, raumzeitlichen und sozialen Beschränkungen *transzendierenden Sinn von Gültigkeit* (vgl. Habermas 1981: 56) kommt eine weitere einschränkende Bedingung durch die Folgen und Nebenfolgen von Handlungsoptionen hinzu. Wie bereits in der Beschreibung des klassischen Delphiverfahrens (siehe Kapitel 3.1) beschrieben, sind die Folgen und Nebenfolgen von Handlungsoptionen in ihrer Gänze oft nicht bekannt oder gar abschätzbar. Gerade wenn eine zukunftsorientierte Perspektive für eine Bewertung von Nöten ist, führt dies oft zu unsicheren Folgen. Trotz dessen sind die Folgen aber nicht beliebig, *„(...) sondern ergeben sich als logische Konsequenz aus dem jeweiligen Wissen und der Anwendung methodischer Regeln innerhalb verschiedener wissenschaftlicher Disziplinen und Denkschulen. Für den Diskurs ist es entscheidend, die Spannweite wissenschaftlich legitimer Abschätzung so genau wie möglich zu bestimmen“* (Renn et al. 1999: 60).

So bleibt im Ideal am Ende eines diskursiven Prozesses als Ergebnis ein Konsens, der die Spannweite der Einschätzungen letztlich sehr konkret eingrenzt. Dies ist auch als Kernaufgabe des Gruppendelphis zu sehen, bei dem die divergierenden Urteile der Teilnehmenden aufgelöst und in einen Konsens überführt werden können (vgl. Niederberger/Renn 2018: 53). Ein solcher Konsens muss damit eine im Diskurs getroffene Entscheidung sein, die aus Sicht der einzelnen Teilnehmenden mindestens aus allen im Diskurs besprochenen Entscheidungen unter Beibehaltung des Status quo präferiert wird (vgl. Renn 1999: 80). Dies wäre letztlich ein *Argumentationskonsens*, der vorliegt *„(...) wenn der Konsens sich nicht nur auf das Ergebnis,*

²² Dies gilt wie angeführt auch in besonderem Maße für das Gruppendelphi (siehe Kapitel 4.1.3).

sondern auch auf die Art der Begründung erstreckt, die zu diesem Ergebnis führt“ (Giegel 1992a: 9). Diese Bestimmung eines Konsenses ist allerdings auch die anspruchsvollste Form eines Konsenses. Dieser unterscheidet sich nach Giegel von den weiteren Formen, einem *Ergebniskonsens*, der sich als Kompromiss darstellt, da zwar die Entscheidung von allen geteilt wird, allerdings nicht die Gründe für dieses Ergebnis. Und weiter, einem *Hintergrundkonsens*, der sich auf intersubjektiv geteilten Überzeugungen stützt, allerdings nicht zu einem geteilten Ergebnis gelangt, durch diese geteilten Überzeugungen allerdings erst einen Dissens ermöglicht. Giegel geht davon aus, dass in modernen pluralistischen Gesellschaften divergierende Parteien zwar einen Ergebniskonsens erzielen können, allerdings hinsichtlich ihrer grundlegenden Orientierungen ein Dissens bestehen bleibt. Dies führt zu einem *Konsensparadox der Moderne*. Diese Moderne zeichnet sich durch vielfältige Teilsysteme aus und weist damit zum einen eine gesteigerte Notwendigkeit der Konsensbildung auf, aber zum anderen ist aufgrund der Pluralität ein ebensolcher Konsens schwerer zu erzielen (Giegel 1992a: 8f.; Giegel 1992b: 59). Allerdings sollte die Konsensorientierung eines Diskurses nicht missverstanden werden: „*Diskurs bedeutet nicht: Einigung auf den kleinsten, meist trivialen Nenner. Es geht vielmehr um eine Konfliktaustragung, bei der die Argumente in aller Klarheit und, wenn notwendig, auch in aller Schärfe ausgetauscht und die unterschiedlichen Werte und Interessen dargelegt werden*“ (Renn 2007: 180). Somit müssen Diskurse nicht in einem harmonischen Konsens enden, sondern können auch zu einem Konsens über einen Dissens führen, denn das „(...) *Ergebnis eines Diskurses ist mehr Klarheit, nicht unbedingt Einigkeit*“ (Renn 2007: 180). Dieser Konsens über einen Dissens bildet somit einen Hintergrundkonsens, der die selbstverständlichen Gegebenheiten akzeptiert, „(...) *als welche sich die Grundstrukturen der gemeinsamen Welt vergesellschafteter Individuen darstellen*“ (Giegel 1992a: 9). Dem folgend definiert auch Miller einen Konsens über einen Dissens indem er diesen als rationalen Dissens einstuft, der sich durch die Identifikation des Strittigen auszeichnet, darüber allerdings voraussetzt, „(...) *daß es den Beteiligten aufgrund von kollektiv Geltendem gelingt, sich darüber zu verständigen, was bei ihrem Konflikt kollektiv gerade nicht gilt*“ (Miller 1992: 39). Somit beinhaltet auch ein im Diskurs erzielter Dissens immer intersubjektive Übereinstimmung, allerdings nicht in praktischer Hinsicht bezüglich des einen konsensualen Ergebnisses, es entstehen *vernünftige* Dissense, indem die Gründe für die jeweiligen Urteile transparent werden (vgl. Gottschalk 1999: 32). So können Konflikte letztlich drei Folgen haben: (1) Es kann keine Verständigung über die Grundlagen des Konfliktes und damit auch kein zufriedenstellendes Ergebnis erzielt werden, (2) weiter kann eine Identifikation des Strittigen erfolgen und damit ein rationaler Dissens erzielt und (3) letztlich ein Konsens erzielt werden (vgl. Miller 1992: 37).

Damit sind sowohl der Konsens als auch der Konsens über einen Dissens Zielgrößen eines diskursiven Verfahrens, denn jeder Diskurs sollte „(...) *Raum für vernünftige Dissense lassen und nicht aufgrund einer als Konsensversprechen falsch verstandenen Konsensorientierung einen einfachstufigen Konsens erzwingen*“ (Gottschalk 1999: 36). Durch diesen Möglichkeitsraum, kann sowohl dem Vorwurf eines Zirkelschlusses (vgl. Fach 1974: 222) begegnet werden, der den Diskurs letztlich als konsensuale Bestätigung durch die Konsenstheorie der Wahrheit beschreibt, als auch das Ziel verfolgt werden, die *Konsensformation* als *Prozess der rationalen Entscheidungsfindung* zu etablieren (vgl. Knorr-Cetina/Amann 1992: 212).

Letztlich muss am Ende eines diskursiven Prozesses nicht als zwingendes Ergebnis ein konsensuales Urteil stehen, sondern es kann auch eine im Diskurs getroffene Entscheidung für

zwei (oder mehr) Entscheidungsoptionen getroffen werden, „(...) *die in sich konsistent und begründungsfähig sind, ohne daß sich alle Teilnehmer auf die Superiorität der einen gegenüber einer anderen begründungsfähigen Option einigen können*“ (Renn 1999: 80). Dadurch wird das Ergebnis eines diskursiven Prozesses nicht beliebig. Es eröffnet vielmehr einen notwendigen Entscheidungsspielraum, der einen erzwungenen Konsens ausschließt und zeitgleich auf einem Bewertungshorizont, gestützt durch geteilte Grundannahmen und Argumente, verschiedene Urteile zulässt und gerade bei unsicheren Sachverhalten eine Kalibrierung der Urteile zulässt. Dies erscheint gerade bei hoher Unsicherheit der Sachverhalte eines Gruppendelphis als sinnvoll, widerspricht zeitgleich aber nicht dem Ziel eines Diskurses, wie auch der konkreten Methode des Gruppendelphis, denn „(...) *divergierende Expertenurteile aufzulösen, sie entweder in einen Konsens zu überführen oder die Hintergründe für die unterschiedlichen Urteile zu erfassen und offen zu legen, ist Kernaufgabe eines Gruppendelphis*“ (Niederberger/Renn 2018: 53). Die diskursiven Elemente eines Gruppendelphis in den Kleingruppen, wie auch im Plenum stellen sicher, dass auch abweichende Urteile über die dargestellten Begründungszusammenhänge und Argumentationen transparent und analysierbar werden und die Expert*innen den jeweiligen zugrundeliegenden Sinn des *Unternehmens* verstehen und bewerten können (vgl. Habermas 1981: 57; Niederberger/Kuhn 2013: 58; Kuhn et al. 2014: 67). Die Struktur des Gruppendelphis soll damit sicherstellen, dass ein Diskurs sein positives Potential entfalten kann. Nur so kann eine legitime – oder gar ideale - Sprechsituation erfolgen. Der Diskurs muss also am richtigen Ort, zur richtigen Zeit und an die richtigen Zuhörerenden gerichtet sein und im Idealfall auch die richtige Form aufweisen, um zum angestrebten Ergebnis eines Konsenses oder eines Konsenses über einen Dissens zu gelangen. So spielt überdies der jeweilige Sprachstil eine Rolle (vgl. Diaz-Bone 2002: 55f.) und nicht nur die Begriffsbestimmungen von Diskurs und Konsens verweisen auf die höheren Ansprüche, die an eine theoretische Fundierung des Gruppendelphis im Vergleich zum klassischen Delphi gestellt werden. Durch die methodisch diskursive Erweiterung werden die Begründungen für die Urteile der Expert*innen und deren Argumente integriert (siehe Kapitel 4.3), die auf verschiedenen Ebenen *rational* und damit *begründet* - *begründet* - sein müssen. Dieses diskurstheoretische Verständnis von Rationalität soll nachfolgend erläutert werden.

5.2.2 Rationalität und Wissen

Der Begriff der Rationalität steht in der Konzeption von Habermas in einer engen Verbindung zu dem Begriff des Wissens. Denn wenn etwas als *rational* angesehen wird, so wird in dieser Logik einer Bewertung eine Wissensbasis unterstellt, auf der diese Einschätzung beruht (vgl. Habermas 1981: 25).

Um sich dieser Konzeption des Wissens zu nähern, muss erst ein Rückgriff auf die Wissensbestände der involvierten Personen erfolgen. Der Begriff der Expert*innen wurde in dieser Arbeit bereits in Kapitel 4.4 näher definiert. Für eine erneute Betrachtung des Expert*innenbegriffs im Rahmen der Diskurstheorie soll an dieser Stelle nochmals darauf verwiesen werden, dass der Begriff als relationaler Begriff, basierend auf einer Form spezialisierten Sonderwissens, zu verstehen ist und über die bloße Form der Verfügbarkeit von Wissen hinausgeht (vgl. Kusche 2008: 54; Przyborski/Wohlrab-Sahr 2008: 131). Auch Habermas setzt ein weites Verständnis des Begriffes von Wissen an, „(...) *denn Rationalität hat weniger mit dem Haben von Erkenntnis als damit zu tun, wie sprach- und handlungsfähige Subjekte Wissen erwerben und*

verwenden. In sprachlichen Äußerungen wird Wissen explizit ausgedrückt, in zielgerichteten Handlungen drückt sich ein Können, ein implizites Wissen aus; auch dieses know-how kann grundsätzlich in die Form eines know-that überführt werden“ (Habermas 1981: 25). Neben dem Wissen, wie beispielsweise bestimmte Handlungen wirken könnten, zeichnet sich ein solches Sonderwissen dadurch aus, dass eine Person zusätzlich weiß, was diese Handlungen bewirken bzw., dass diese wirken können.

Eine auf ein solches Wissen basierende Handlung ist nach Habermas dann wirksam, wenn sie sich in Form einer Prognose als *wahr* herausstellt. Neben dem Begriff der *Wahrheit*, der sich auf die konkrete Existenz von Sachverhalten bezieht, führt Habermas weiterhin den Begriff der *Wirksamkeit* in sein Konzept mit ein. Wirksamkeit ist dabei als Eingriff zu verstehen, durch den Gegebenheiten bzw. Sachverhalte hervorgebracht werden können: „Mit seiner Behauptung nimmt A Bezug auf etwas, das in der objektiven Welt tatsächlich statthat, mit seiner Zwecktätigkeit nimmt B auf etwas Bezug, das in der objektiven Welt statthaben soll“ (Habermas 1981: 26).²³ Dieses Beispiel verdeutlicht einen grundlegenden Unterschied verschiedener Argumentationslogiken: Die von Habermas benannte *Person A* nutzt als Basis ihrer Einschätzung einen empirisch beobachtbaren Sachverhalt, während die zweite *Person B* auf normativer Grundlage argumentiert. Beide erheben allerdings mittels Argumentation gleichermaßen Ansprüche auf Richtigkeit ihrer Aussagen, entweder begründet durch angenommene Wahrheit oder im anderen Fall durch angenommene Wirksamkeit ihrer vorgebrachten Äußerungen. Sie erheben so „(...) mit ihren symbolischen Äußerungen Ansprüche, die kritisiert und verteidigt, d. h. begründet werden können. Die Rationalität ihrer Äußerungen bemißt sich an den internen Beziehungen zwischen dem Bedeutungsgehalt, den Gültigkeitsbedingungen und den Gründen, die nötigenfalls für ihre Gültigkeit, für die Wahrheit der Aussagen oder für die Wirksamkeit der Handlungsregel angeführt werden können“ (Habermas 1981: 26f.). Bis zu diesem Punkt der Konzeption ist eine Äußerung also dann rational, wenn sie kritisiert werden kann und die Person, die diese Aussage getätigt hat, diese auch (vernünftig) begründen kann (vgl. Habermas 1981: 26f.; Grunwald 1998: 208ff.; Keller et al. 2011: 12).

Die Rationalität einer Aussage nur daran zu bemessen, ob sie lediglich kritisierbar sein muss, hat nach Habermas allerdings zwei Schwächen: Eine solche Konzeption von Rationalität wäre...

1. ... zu abstrakt, da so wichtige Differenzierungen nicht in der Konzeption erfasst werden können.
2. ... zu eingrenzend, da mit rational nicht nur Aussagen bewertet werden die wahr oder wirksam sein können, sondern ein breites Spektrum von möglichen Argumentationen abdecken (vgl. Habermas 1981: 27).

Rationalität bezieht sich folglich auf unterschiedliche Formen des Austauschs von Argumenten, viele unterschiedliche Möglichkeiten gegenseitig Stellung zu beziehen und ein gemeinsames Ergebnis bzw. eine gemeinsam getragene Einschätzung auszuarbeiten (vgl. Habermas 1981: 27). Um sowohl dem breiten Spektrum von Rationalität gerecht zu werden, wie auch die einzelnen Differenzierungen beschreibbar und kategorisierbar zu machen, arbeitet Habermas Konzeption mit einer *engen* wie auch eine *breiten* Konzeption des Begriffs der Rationalität.

²³ Hervorhebung im Original.

5.2.2.1 Kognitiv-instrumentelle und kommunikative Rationalität

In der engen Betrachtung von Rationalität kann sich der Begriff in seiner Verwendung von Wissen entweder auf *instrumentelle Verfügung* oder auf *kommunikative Verständigung* beziehen. In beiden Fällen setzt die Verwendung des Wissens an der objektiven Welt an (vgl. Habermas 1981: 30). Dabei drückt sich *kognitiv-instrumentelle Rationalität* in zielgerichteten Handlungen aus. Eine instrumentelle Anwendung des Wissensbegriffs ist dann rational, wenn sich eine Person auf empirische Fakten stützt und ihre Handlung genau an diesen Fakten ausrichtet. So erfolgt eine zielgerichtete Handlung, die von einer „(...) *nicht-kommunikativen Verwendung propositionalen Wissens* (...)“ (Habermas 1981: 28) ausgeht. Personen vertrauen hierbei auf die beobachtbaren Fakten der objektiven Welt, die folgende Einschätzung bzw. Bewertung erfolgt damit rein subjektiv.

Der Begriff der *kommunikativen Rationalität* ist hingegen keine rein subjektive Bewertung, sondern stützt sich auf den kommunikativen Austausch zwischen an dieser Kommunikation beteiligten Personen. Hier geht Habermas davon aus, dass die verschiedenen „(...) *Teilnehmer ihre zunächst sehr subjektiven Auffassungen überwinden und sich dank der Gemeinsamkeit vernünftig motivierter Überzeugungen gleichzeitig der Einheit der objektiven Welt und der Intersubjektivität ihres Lebenszusammenhangs vergewissern*“ (Habermas 1981: 28). Die in eine Argumentation eintretenden Personen verlassen damit die an der objektiven Welt ausgerichtete Handlungsmotivation und unterziehen die von ihnen angenommenen Fakten einer Prüfung mittels argumentativen Austauschs mit anderen. Kommt es zu einer konsensualen Bewertung durch die beteiligten Personen, so kann von einer hohen kommunikativen Rationalität ausgegangen werden (vgl. Habermas 1981: 28). Oder vereinfacht ausgedrückt: Die eingebundenen Personen fühlen sich in ihrem gemeinsamen Urteil durch den Austausch untereinander bestärkt. Im Gegensatz zu einer rein instrumentellen Handlung stellt eine kommunikative Verständigung eine höhere Hürde für eine erfolgreiche Entscheidungsfindung dar. Denn im argumentativen Austausch muss es nicht zwangsläufig zu einer konsensualen Einschätzung kommen. Besonders dann, wenn die Gesprächspartner*innen verschiedene *pragmatische Rollen* einnehmen. Hierbei muss der „(...) *Sprecher, der eine Behauptung aufstellt, (...) über eine »Deckungsreserve« guter Gründe verfügen, um erforderlichenfalls seine Gesprächspartner von der Wahrheit der Aussage überzeugen und ein rational motiviertes Einverständnis herbeiführen zu können*“ (Habermas 1981: 29).

Letztlich unterscheiden sich beide Konzepte in ihrem Umfang der enthaltenen Autonomie. So ist eine Person, die aufgrund einer instrumentellen Verfügung zu ihrem Urteil gelangt, unabhängig von Beschränkungen ihrer Umwelt und kann so ohne Beeinflussung von außen zu ihrem, auf die objektive Datenlage fußenden, Urteil gelangen. Die Person hat damit ein hohes Maß an Selbstbehauptung. Kommunikative Verständigung hingegen erweitert den Spielraum für die Koordinierung von angestrebten Handlungen und kann überdies zu einer Einigung und damit Beilegung von Handlungskonflikten führen (vgl. Habermas 1981: 34). Nach Habermas sind letztlich sowohl instrumentelle Verfügungen wie auch kommunikative Handlungen Wege um Produkte zu generieren, die als rational angesehen werden können:

1. Effiziente Handlungen als kognitiv-instrumentelles Ergebnis.
2. Begründete Behauptungen als durch Kommunikation ausgearbeitetes Resultat (vgl. Habermas 1981: 34f.).

Oder anders ausgedrückt: „Im ersten Falle geht es um Handlungen, die darauf abzielen, die Welt zu verändern; im zweiten Falle um Handlungen, die darauf abzielen, andere zu überzeugen“ (Renn 1999: 70).

Aber wie sieht es aus, wenn eine Person ihr Urteil auf einen Wunsch bezieht? Auch diese normative Begründung kann durchaus als rational angesehen werden. Es gibt folglich noch andere Typen von rationalen Begründungen, die aber nicht mit Wahrheits- oder Erfolgsansprüchen verknüpft werden können. Diese finden sich bei Habermas in einer breiten Konzeption von Rationalität (vgl. Habermas 1981: 35).

5.2.2.2 Normenregulierte Handlungen und expressive Selbstdarstellung

Als rational werden im Rahmen von Kommunikation nicht bloß Behauptungen bezeichnet, die auf unumstößliche Tatsachen bzw. faktische Gegebenheiten gründen. Eine Äußerung kann auch mittels anderer Begründungen im kommunikativen Austausch als rational verstanden werden:

1. *Normenregulierte Handlung*: Die Begründung bezieht sich auf eine bestehende Norm.
2. *Expressive Selbstdarstellung*: Die Begründung bezieht sich auf ein Erlebnis.

Dieses Erlebnis kann dabei beispielsweise in einem Wunsch, in einem Gefühl oder in der Preisgabe eines Geheimnisses fortgeführt werden. Rational ist eine solche expressive Selbstdarstellung dann, wenn die Person, die diese äußert, praktische Konsequenzen daraus zieht und sich anhand dieser Konsequenzen konsistent verhält. Im Fall einer auf Normen begründenden Äußerung erhebt eine Person den Anspruch auf Richtigkeit ihrer Aussage und im Fall einer auf einem Erlebnis begründeten Äußerung, bezieht sich die Person auf die Wahrfähigkeit dieser Aussage (vgl. Habermas 1981: 34f.).

Wie bereits weiter oben erläutert, ist auch in diesem breiten Verständnis von Rationalität das notwendige Bemessungskriterium, dass die Aussage kritisiert werden kann. Sowohl normenregulierte Handlungen als auch expressive Selbstdarstellungen lassen sich zudem (vernünftig) begründen (vgl. Habermas 1981: 26f.). Damit ist auch „(...) für ihre Rationalität (...) die Möglichkeit der intersubjektiven Anerkennung eines kritisierbaren Geltungsanspruchs konstitutiv“ (Habermas 1981: 35).

Die Bedingungen der Kritisierbarkeit und Begründung für eine rationale Aussage gelten bei Habermas zudem für einen dritten Typus innerhalb des breiten Verständnisses: Evaluative Äußerungen, die weder rein expressiv noch normativ verbindlich sind. Habermas nennt hier als Beispiele die Vorliebe für eine herbstliche Landschaft, eine Ablehnung des Militärs oder auch eine bestehende Eifersucht auf Kolleg*innen. Solche Werturteile sind weder rein privat, noch sind sie von der Allgemeinheit intersubjektiv anerkannte Normen, werden aber von Angehörigen einer Kultur- und Sprachgemeinschaft ähnlich interpretiert. Wenn eine Person ihre Aussagen unter Verwendung solcher Wertstandards tätigt, die kulturell keine gemeinsam getragene Verwendung haben, dann verhalten sie sich nicht mehr rational, sondern *idio-synchronisch* (vgl. Habermas 1981: 35f.). Das kann innovativ sein, allerdings reicht für „(...) diese als

abweichend empfundenen Reaktionen (...) die rechtfertigende Kraft der herangezogenen kulturellen Werte nicht aus“ (Habermas 1981: 37).²⁴ Solche evaluativen Äußerungen zeigen anschaulich die Grenze der Aussagen, die als rational bezeichnet werden können. Auch hier gilt, dass sie dann als rational verstanden werden können, wenn sie begründet werden können und die sie äussernde Person, Argumente für ihre persönlichen Wünsche und Gefühle heranziehen kann. Wer sich aber „(...) in seinen Einstellungen und Bewertungen so privatistisch verhält, daß sie durch Appelle an Wertstandards nicht erklärt und plausibel gemacht werden können, der verhält sich nicht rational“ (Habermas 1981: 37).

Letztlich gilt für den Begriff der Rationalität sowohl im engen als auch im weiten Verständnis, dass vernünftige Gründe für die jeweilige Einschätzung vorliegen müssen. Diese müssen überdies auch im kommunikativen Austausch (besonders bei Kritik) geäußert und Argumente für die jeweilige Einschätzung herangezogen werden. All diese konstitutiven Sprechhandlungen werden so zu einer kommunikativen Praxis, „(...) die vor dem Hintergrund einer Lebenswelt auf die Erzielung, Erhaltung, und Erneuerung von Konsens angelegt ist, und zwar eines Konsenses, der auf der intersubjektiven Anerkennung kritisierbarer Geltungsansprüche beruht“ (Habermas 1981: 37). Ein fruchtbarer Diskurs benötigt somit gültige und gute Begründungen für alle während des Diskurses aufgestellten Behauptungen, was eine logische Kohärenz der Argumentation voraussetzt (vgl. Alcántara et al. 2016: 36).

Aber wie kann ein solcher Austausch von Begründungen erfolgen, der zu einer geteilten Einschätzung führt? Eine solche Theorie der Argumentation soll folgend weiter beschrieben werden.

5.2.3 Argumentationspraxis: Geltungsansprüche und deren Einlösung im Diskurs

Für Habermas stellt die Argumentationspraxis eine *Berufungsinstanz* dar, in der das kommunikative Handeln mit anderen Mitteln fortgesetzt werden kann, wenn (1) Alltagsroutinen einen bestehenden Dissens nicht mehr lösen können und (2) dieser Dissens nicht mit Gewalt gelöst werden soll. Ein so erzielter Konsens muss, wie oben beschrieben, basierend auf Gründen erzielt worden sein und Personen, die an dieser Kommunikation beteiligt sind, müssen ihre Äußerungen bei auftretender Kritik mit Argumenten verteidigen bzw. belegen. Argumente sind damit „(...) diejenigen Mittel, mit deren Hilfe die intersubjektive Anerkennung für den zunächst hypothetisch erhobenen Geltungsanspruch eines Proponenten herbeigeführt und damit Meinung in Wissen transformiert werden kann“ (Habermas 1981: 48). Ein Argument enthält so immer auch Gründe, die den Geltungsanspruch dieses Argumentes bei Kritik verfestigen können. Diese Geltungsansprüche müssen wie oben beschrieben nicht nur auf kognitive Aussagen bezogen werden, sondern können zudem expressive wie auch normative Äußerungen umfassen (vgl. Habermas 1981: 37f.; Renn 2000b: 125).

Kommunikatives Handeln ist unter diesen Bedingungen immer auch Argumentation, die für Habermas einen bestimmten Typ des Begriffes der Rede bildet: „Argumentation nennen wir den Typus von Rede, in dem die Teilnehmer strittige Geltungsansprüche thematisieren und

²⁴ Hervorhebung im Original.

versuchen, diese mit Argumenten einzulösen oder zu kritisieren“ (Habermas 1981: 38).²⁵ Dabei kann ein *symmetrischer Austausch* von Informationen und Argumenten nur dann den Ansprüchen von kommunikativer Rationalität entsprechen, wenn alle Teilnehmenden des Diskurses gleiche Rechte aber auch Pflichten besitzen und von strategischem Handeln absehen (vgl. Renn/Webler 1998: 48).

Somit müssen auch Eingeständnisse gemacht werden, und zwar sobald eine Person ihre Aussage nicht ausreichend begründen und eine andere Person bessere Gründe für ihre Einschätzung vorbringen kann. In diesem Moment wird deutlich, warum die Argumentationspraxis eine *Berufungsinstanz* darstellt, denn rationale Äußerungen sind dadurch, dass sie kritisiert und teilweise verworfen werden können, auch *verbesserungsfähig*. So stellt das Argumentieren einen Lernprozess dar und führt letztlich zu einem besseren Ergebnis als es ohne einen Diskurs der Fall gewesen wäre: „*Rationale Äußerungen sind aufgrund ihrer Kritisierbarkeit auch verbesserungsfähig: wir können fehlgeschlagene Versuche korrigieren, wenn es gelingt, die Fehler, die uns unterlaufen, zu identifizieren*“ (Habermas 1981:38f.).²⁶

Tabelle 2: Argumentationstypen in der Theorie des kommunikativen Handelns

		Bezugsgrößen	
		Problematische Äußerungen	Kontroverse Geltungsansprüche
Formen der Argumentation	Theoretischer Diskurs	Kognitiv-instrumentell	Wahrheit von Propositionen; Wirksamkeit teleologischer Handlungen
	Praktischer Diskurs	Moralisch-praktisch	Richtigkeit von Handlungsnormen
	Ästhetische Kritik	Evaluativ	Angemessenheit von Wertstandards
	Therapeutische Kritik	Expressiv	Wahrhaftigkeit von Expressionen
	Explikativer Diskurs	-	Verständlichkeit bzw. Wohlgeformtheit symbolischer Konstrukte

Quelle: Eigene Darstellung in Anlehnung an Habermas (1981: 45)

Der Austausch von Argumenten bei bestehendem Dissens kann also letztlich zu einem besseren Ergebnis führen. So verhält es sich damit wie in der Redewendung: *Aus Fehlern lernt man*. Allerdings sind in Habermas Konzeption mit den Fehlern explizite Fehler gemeint, die auf einer durch die Form der Argumentation ermöglichten Verhaltens basieren, das als *rational* angesehen werden muss. Dieses bereits beschriebene Verständnis von Rationalität²⁷ hilft, verschiedene Typen von Argumentation zu bilden und jeweilige Bezugsgrößen zu definieren, wie sie in Tabelle 2 dargestellt werden (vgl. Habermas 1981: 44).

²⁵ Hervorhebung im Original.

²⁶ Hervorhebung im Original.

²⁷ Siehe Kapitel 5.2.2.

Die Typologie und die jeweiligen Bezugsgrößen zeigen auf, dass legitime Aussagen im Diskurs vielfältige Formen annehmen können. So können beispielsweise Behauptungen, Gefühlsäußerungen, Appelle und Versprechungen geäußert und begründet werden. Diese stehen den anderen am Diskurs beteiligten Personen zur Kommentierung oder Kritik offen. *„Im Rahmen dieses Austauschs von Aussagen werden Geltungsansprüche angemeldet. Diese Ansprüche besagen, daß die Aussagen entweder hilfreich, wahr, wahrhaftig oder angemessen seien. Innerhalb des Diskurses muß dann die Gültigkeit dieser Ansprüche überprüft und eingelöst werden“* (Renn/Webler 1998: 50).

Mit der Verknüpfung der jeweiligen Bezugsgrößen bezieht sich der Begriff der Rationalität folglich auf ein System von Geltungsansprüchen, mit deren Hilfe die Teilnehmenden des Diskurses die Validität der getätigten Aussagen intersubjektiv bewerten können. Die Geltungsansprüche sind Hilfsmittel, um legitime von illegitimen Aussagen oder von bloßen Einwänden zu unterscheiden. Diese Unterscheidung von Aussagetypen und die Art ihrer Überprüfung mittels der Geltungskriterien sind wesentliche Merkmale für die Durchführung und Analyse von Diskursen und sollen deshalb folgend näher erläutert werden (vgl. Renn/Webler 1994: 35f.; Renn 1999: 83):

- **Verständlichkeit:** Um kommunikative Aussagen bewerten zu können, müssen solche Aussagen daraufhin überprüft werden, ob die Diskursteilnehmenden mit dem Geäußerten auch den gleichen Bedeutungsinhalt verbinden. Das heißt, die Aussagen müssen sowohl verständlich wie auch eindeutig sein. Dies gilt beispielsweise für im Diskurs geäußerte Begriffsdefinitionen, Klassifikationsschemata oder auch Übersetzungen, deren Bedeutungsinhalte von allen Teilnehmenden gleich verstanden und interpretiert werden sollten. Das Geltungskriterium dieses Typs von Aussagen kann dann in Frage gestellt werden, wenn Begriffe inkonsistent bzw. unklar genutzt werden oder der Aufbau von einzelnen Argumenten nicht logisch nachvollziehbar ist. Auch Übersetzungen können in ihrer Geltung für den bestehenden Diskurs angezweifelt werden, wenn keine Übereinstimmung zwischen Urtext und im Austausch verwendeten Text hergestellt werden kann. Es geht dabei hinsichtlich des Kriteriums der Verständlichkeit nicht darum, die Wahrheit der jeweiligen Aussagen anzuzweifeln. *„Definitionen können beispielsweise nicht danach beurteilt werden, ob sie wahr oder falsch sind, sondern nur ob sie im Rahmen der Diskussion zweckmäßig sind“* (Renn/Webler 1998: 50). Ein kommunikativer Aussagetyt ist letztlich dann für das Geltungskriterium der Verständlichkeit zulässig, wenn er dem Geltungsanspruch der Zweckmäßigkeit für den Diskurs genügen kann (vgl. Renn/Webler 1998: 50; Renn 1999: 72f.)
- **Evidenz (Wahrheit):** Unter dem Aussagetypus über kognitive Sachverhalte können die Aussagen auf systematischem Wissen, aber auch auf personellem Wissen bzw. anekdotischer Erfahrung oder sogar auf Intuition beruhen. Hier ist also - wie auch in Kapitel 5.1.1 im Rückgriff auf die mentalen Modelle beschrieben - nicht nur das als systematisches Expert*innenwissen angesehene Wissen von Bedeutung, sondern zusätzlich Erfahrungen und die Intuition darüber *Wie* etwas gelöst werden kann. *„Vielfach wird unterstellt, dass für kollektive Entscheidungen nur das systematische Expertenwissen von Bedeutung sei. Diese Forderung ist epistemologisch, aber auch aus praktischen Erwägungen wenig sinnvoll, denn es ist völlig irrelevant, woher das Wissen stammt“* (Renn 2007: 178). Worauf es allerdings ankommt ist, dass es gewisse Regeln gibt, nach denen die im Diskurs getätigten Aussagen auf ihre Zweckmäßigkeit bzw. ihre Richtigkeit hin überprüft werden können. Die

Verständigung über die Vorgehensweise im Diskurs ist hier der entscheidende Punkt: Es muss unter den Teilnehmenden festgelegt werden, nach welchen Beurteilungskriterien sie die Richtigkeit von solchen kognitiven Aussagen bemessen. Für einen Diskurs über eine naturwissenschaftliche Entscheidung (wie beispielsweise Grenzwerte) können dies Messdaten und methodologische Modelle sein. Plant man aber einen Spielplatz, können die Erfahrungen von Eltern und Kindern durchaus eher sinnvoll sein. Bei Vermutungen über zukünftige Folgen von angestrebten Maßnahmen braucht es letztlich Kriterien, die im Umgang mit unsicherem Wissen eine Überprüfung nach dem Kriterium der Gültigkeit ermöglichen. Ein kognitiver Aussagetyp ist somit dann unter dem Geltungskriterium der Evidenz zulässig, wenn er dem Geltungsanspruch der Wahrheit oder der Gültigkeit für den Diskurs genügen kann (vgl. Renn/Webler 1998: 51; Renn 1999: 73).

- **Wahrhaftigkeit:** Aussagen dieses Typus können in zwei Arten unterschieden werden: Zum einen können unter diesem Typus Äußerungen von Gefühlen, Ahnungen oder Befürchtungen gefasst werden, zum anderen Äußerungen über Verhaltensprognosen wie Versprechungen oder Vertrauensappelle. Die zweite Art von Äußerungen lässt sich dabei intersubjektiv überprüfen, indem beispielsweise ein im Diskurs von einer Person geäußertes Versprechen von den Diskursteilnehmenden nachträglich daraufhin geprüft wird, ob diese Person dieses auch einhält. Wenn eine Person aufgrund von vorhergehendem Verhalten bereits Vertrauen der anderen Personen genießt, kann dies auch im Diskurs dazu führen, dass ihre Äußerung einer Verhaltensabsicht von den anderen Teilnehmenden als wahrhaftig eingeschätzt wird. Bei der ersten Art von Äußerungen, wie den Gefühlen, ist die Wahrhaftigkeit nicht intersubjektiv gültig überprüfbar. Solche affektiven Äußerungen lassen sich auch nicht objektiv auf Fakten basierend überprüfen und werden deshalb für viele Diskursformen (wie beispielsweise Anhörungen) prinzipiell ausgeschlossen. Dies kann unter Umständen jedoch problematisch sein, da ein Ausschluss von affektiven Aussagen dazu führen kann, dass bestimmte Präferenzen im Diskurs nicht berücksichtigt oder nicht begründet werden können. Während Habermas davon ausgeht, dass für die Teilnehmenden die Erfahrungen aus der Lebenswelt genügen, um eine Beurteilung der Wahrhaftigkeit von affektiven Aussagen vorzunehmen, sind Renn und Webler in diesem Punkt skeptischer: *„Die Lebenswelt ist kein Deus ex machina, der uns lehrt, die Wahrhaftigkeit von Gefühlsäußerungen intersubjektiv gültig zu beurteilen“* (ebd. 1998: 51). Sie schlagen deshalb für den Diskurs die Notwendigkeit vor, affektive Aussagen verständlich in kognitive oder normative Aussagen zu übersetzen (vgl. Habermas 1984, S. 302ff.; Renn/Webler 1998: 51f.; Renn 2007: 178f.).
- **Normative Angemessenheit:** Unter diesem Aussagetypus können Äußerungen verstanden werden, die auf Werturteile, Präferenzen, Bewertungen von Optionen und moralischen Urteilen basieren. Solche normativen Äußerungen lassen sich als evaluative Aussagen und kollektiv verbindliche Verhaltensnormen aufgliedern. Dabei beziehen sich evaluative Äußerungen auf subjektive Präferenzen oder gruppenspezifische Vorlieben. Wie bereits in Kapitel 5.2.2.2 beschrieben, lassen sich für diese Art von Aussagen keine übergeordneten Überprüfungskriterien angeben, so dass die Teilnehmenden des Diskurses von guten oder schlechten Präferenzen ausgehen können, denn wenn *„(...) jemand A lieber mag als B und jemand anders B lieber mag als A, dann lassen sich zwar diese Präferenzen begründen, aber nicht auf ihre Wahrheit hin überprüfen. Aussagen über den Ist-Zustand lassen prinzipiell keine logisch schlüssigen Ableitungen über den Soll-Zustand zu (naturalistischer Fehlschluß)“* (Renn 1999: 74). Hier ließe sich zwar die Evidenz der persönlichen Beziehung von

A zu B durch weitere Kommunikation überprüfen, aber generell sind evaluative Aussagen nicht nach intersubjektiv gültigen Kriterien für eine Bewertung zugänglich. Solange die Präferenzen der einen nicht die Werte des anderen negativ beeinträchtigen, sollte und darf es keine Norm geben, die die jeweiligen Präferenzen beeinträchtigt. Eine mögliche Vielfalt von Lebensstilen oder auch individuelle Sinnorientierung sind folglich nicht Verhandlungsgegenstände von Diskursen. Dies gilt solange bis nicht eben diese Vielfalt selbst bedroht ist. In genau diesem Fall, in dem „(...) Präferenzen und Werte von Personen miteinander kollidieren, so daß die Ausübung dieser Präferenzen oder Werte unvermeidlich zu Konflikten führt, sind verbindliche Verhaltensnormen notwendig, die eine gerechte Aufteilung von präferenzgerechten Verhaltensweisen ermöglicht. Dazu benötigt man Kriterien, die von allen anerkannt werden“ (Renn 2007: 179). Nach der Konzeption von Habermas können solche übergeordneten Kriterien in einem konsensorientierten Diskurs intersubjektiv gültig bestimmt werden, der die für alle Betroffenen beste normative Lösung herausstellt (vgl. Habermas 1981: 37; Renn/Webler 1998: 52f.; Renn 1999: 74).

Tabelle 3: Typen von Aussagen und ihre Geltungskriterien

Geltungs-kriterium	Verständlich-keit	Evidenz	Wahrhaftigkeit	Normative Angemes-senheit
Aussagety-p	kommunikativ	kognitiv	expressiv	evaluativ/normativ
Geltungs-an-spruch	Zweckmäßigkeit	Wahrheit (Gültigkeit) ²⁸	Aufrichtigkeit	Gemeinwohl (evalua-tiv);
				Richtigkeit (normativ)
Über-prü-fungs-krite-rien	Erleichterung des Verständnisses, Kongruenz zwischen Übersetzung und Original, Autorisierung durch Verfasser*in	Für systematisches Wissen: Methodo-logische Regeln; Peer Review	Für affektive Aussagen: Übersetzung in kognitive oder normative Aussagen und Autorisierung der Übersetzung	Für evaluative Aussagen: Kohärenz und Verallgemeinerungs-fähigkeit
		Für anek-dotisches Wissen: Singuläre Überprüfbarkeit	Für Verhaltensprognosen: Verhalten in der Vergangenheit und Reputation	Für normative Aussagen: Universalisierungsgebot und Kompatibilität mit Gesetzen
Beispiel(e)	Definition	Tatsachen-behauptung	Versprechen, Gefühlsäußerung	Werte, Ziele, Interessen

Quelle: Eigene Darstellung in Anlehnung an Renn (1999: 72)

²⁸ Eigene Ergänzung.

Die Formen der Argumentation und die darin enthaltenen Äußerungen können folglich mit den beschriebenen vier Kriterien während eines Diskurses auf ihren Geltungsanspruch hin überprüft werden. Sowohl die Aussagetypen als auch die Kriterien sind Hilfsmittel um zulässige von unzulässigen Äußerungen während eines Diskurses zu trennen (vgl. Renn 2007: 177). Tabelle 3 bietet dabei einen zusammenfassenden Überblick über die jeweiligen Typen von Aussagen und wird ergänzt durch die zugehörigen Geltungskriterien.

5.2.4 Implikationen der diskurstheoretischen Erweiterung

An der Diskurstheorie kann hinsichtlich ihrer Reichweite sicherlich kritisiert werden, dass nicht a priori alle Überzeugungsänderungen via Argumentation erfolgen können. Einige Alternativen lassen sich nur direkt vorführen und nicht zwangsläufig sprachlich beschreiben. Allerdings bietet die Argumentationspraxis die Möglichkeit einer gegenseitigen Versicherung der Rationalität von Einstellungsänderungen, denn etwas Besseres „(...) *als unsere besten Gründe steht uns dafür nicht zur Verfügung*“ (Gottschalk 1999: 9).

Diskursive Verfahren erheben den Anspruch einen Rahmen für einen solchen Diskurs bereitzustellen, der einen Austausch über eben diese besten Gründe bietet und erheben damit ebenso den Anspruch gerechtere und kompetentere Lösungen zu liefern. In vielen praktischen Diskursen sind die Teilnehmenden am Ende aber enttäuscht oder verunsichert, da diese Verfahren mit großen Ansprüchen gestartet sind, dann aber mit trivialen Ergebnissen endeten. Die Umsetzung eines Verfahrens, das den Bedingungen der diskurstheoretischen Voraussetzungen gerecht wird, muss damit hohe Hürden überwinden. So darf sich nicht zuletzt beispielsweise eine *Diktatur des Sitzfleisches* durchsetzen und die Personen mit dem längsten Atem das Ergebnis dominieren oder gar diktieren, indem sie beharrlich bis zum Ende des Verfahrens auf ihrem Standpunkt beharren. Die Organisator*innen eines Diskursverfahrens stehen vor dem Dilemma den Austausch von Argumenten solange einen Rahmen zu bieten, bis er erschöpfend stattfinden konnte und zeitgleich dem Druck zu unterliegen zu einem Konsens zu gelangen (vgl. Weinrich 1973: 145; Gottschalk 1999: 9; Renn 1999: 66; Schweitzer 2007: 80).

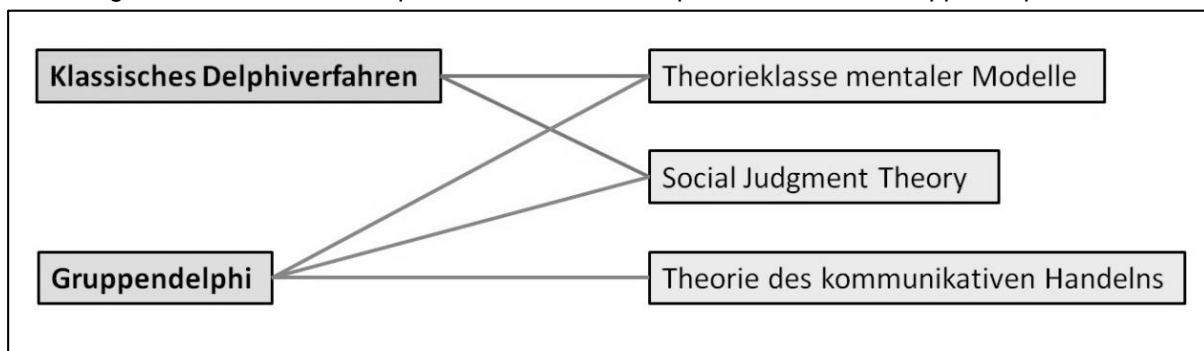
Warum aber bietet die Theorie des kommunikativen Handelns trotz dieser hohen Hürden einen geeigneten Rahmen eines theoretisch begründeten Diskurskonzeptes für eine praktische Umsetzung? Renn beantwortet diese Frage mit drei Begründungen: „*Zum Ersten hat Habermas die Grundelemente einer jeden Verständigung über Wissen, Bewertungen und Ziele identifiziert, theoretisch zugänglich gemacht und dadurch die Struktur von idealen Beteiligungsverfahren vorgezeichnet. Zum Zweiten verarbeitet Habermas Erkenntnisse aus der Entwicklungspsychologie, der Kleingruppenforschung, der politischen Soziologie und der Philosophie und integriert sie in einem logisch konsistenten Bezugsrahmen. Zum Dritten erscheinen die von Habermas verwandten Kategorien nach langjähriger praktischer Erfahrung mit Bürgerbeteiligung als intuitiv einsichtig und relevant*“ (Renn 2007: 175).

Neben einer intuitiv praktischen Begründung spricht somit ein interdisziplinär gebildeter, logisch konsistenter Bezugsrahmen dafür, auch einer praktischen Umsetzung eines Diskurses die Begründungszusammenhänge und Qualitätsanforderungen, die sich aus der Theorie des kommunikativen Handelns ableiten lassen, zu Grunde zu legen. Die Grundelemente eines idealen Diskursverfahrens bieten für den pragmatischen Vollzug eines realen Diskurses einen

brauchbaren Orientierungsrahmen. Dieser hilft den Organisator*innen und Beteiligten eines diskursiven Verfahrens gemeinsame Regeln der Beweisführung zu vereinbaren und die jeweiligen Geltungsansprüche zu prüfen. Das Diskurskonzept nach Habermas kann damit als Idealvorstellung begriffen werden, die allerdings für reale Verfahren und Prozesse in unterschiedlichen Formen handlungsrelevant sein kann (vgl. Renn 1999: 84; Renn 2007: 175; Keller et al. 2011: 12).

Ein diskursiver Austausch von Argumenten, bei bestehendem Dissens, kann somit letztlich zu einem besseren Ergebnis führen, als es ohne diese kontroversen Geltungsansprüche der Fall gewesen wäre: *„Jede explizite Überprüfung von kontroversen Geltungsansprüchen verlangt die anspruchsvolle Form einer Kommunikation, welche die Voraussetzungen der Argumentation erfüllt. Argumentationen ermöglichen ein Verhalten, das in einem besonderen Sinne als rational gilt, nämlich dem Lernen aus expliziten Fehlern“* (Habermas 1981: 44). Für die Verständigungsorientierung im Diskurs ist es notwendig, dass die Beteiligten die eigenen Bewertungsmaßstäbe transparent und nachvollziehbar vor anderen begründen können (vgl. Renn 1999: 83). Will man die, durch die Überprüfung der kontroversen Geltungsansprüche erzielten, *besseren* Ergebnis auch diskursanalytisch offenlegen, geht es nicht darum, den Aussagen genau eine *wahre* oder *objektive* Bedeutung zuzuweisen, sondern es macht eine gewisse methodische Systematik und entsprechende Klassifizierungen notwendig, ob sich die Ergebnisse nachweislich als *fruitful* erweisen (vgl. Häder 2009: 40; Keller 2011: 76f.).

Abbildung 6: Theoretische Konzeptionen klassisches Delphiverfahren und Gruppendelphi



Quelle: Eigene Darstellung

Um das methodische Design dieser Arbeit theoretisch zu begründen, ist es notwendig die genuinen theoretischen Annahmen, die sowohl für das klassische Delphiverfahren wie für das Gruppendelphi herangezogen werden können, für das Gruppendelphi weitergehend durch Annahmen aus der Diskurstheorie zu erweitern. Wie in Abbildung 6 dargestellt, widersprechen diese diskurstheoretischen Annahmen nicht den Annahmen, die auch für das klassische Delphi gelten, sondern bauen auf diesen auf, denn die Grundlogik der Iteration zwischen Ergebnis und erneuter Bewertung haben beide Ansätze gemein (vgl. Kuhn et al. 2014: 65f.).

5.2.5 Einordnung des Gruppendelphis als Diskurstypus

Final soll das Gruppendelphi in diesem Kapitel als Methode auf theoretischer Ebene verortet und als diskursives Verfahren einer entsprechenden Diskursart zugeordnet werden, um die Funktionsfähigkeit eines Gruppendelphis anhand der Theorie des kommunikativen Handelns

bewerten zu können. Eine solche Klassifizierung von Diskursarten kann anhand des Inputs, der von den Teilnehmenden erwartet wird, erfolgen (vgl. Niederberger 2015a: 37).

Um bei der Entscheidung über mögliche Maßnahmen die Wissensgrundlage dafür zu erarbeiten, welche zu erwartenden Folgen mit den jeweiligen Maßnahmen einhergehen, muss zunächst ein *Wissensdiskurs* geführt werden. Wird die Systemintegrität durch den Eindruck der Bedrohung seiner essentiellen Funktionen wahrgenommen, so ist es notwendig, Folgewissen auszuarbeiten und zu erlangen. Durch dieses Folgewissen kann die Priorität der Funktionserfüllung des Systems nachgewiesen und weiter auch (politisch) legitimiert werden. Dabei muss deutlich werden, welche Folge die jeweilige Maßnahme erwarten lässt. Genauso geht es um die Bestimmung von mit der Maßnahme verbundenen Unsicherheiten. Eine solche Abschätzung ad hoc zu treffen oder durch einzelne Expert*innen bewerten zu lassen, ist aufgrund der hohen Ambiguität, die sich durch die Komplexität des Systems ergibt, nur schwer – wenn nicht gar meistens unmöglich – sicher zu bewerkstelligen. Denn welche „(...) *Maßnahme aber welche einzelne Dimension im Einzelnen berührt und in welcher Weise, lässt sich vorab schwer abschätzen*“ (Renn 2007: 182). Eine solche Abschätzung kann nur im Rahmen eines Wissensdiskurses geschehen, in dem eine Analyse des aktuellen Wissensstandes auf Basis der im Diskurs festgelegten Bewertungskriterien erfolgen kann (vgl. Renn 2007: 182; Niederberger 2015a: 41). Im *Reflexionsdiskurs* wird sich der Frage nach der Behandlung von (zwangsweise) auftretenden Konflikten zwischen den angedachten Maßnahmen oder Maßnahmenpaketen angenommen. Ziel ist dabei die Abwägung zwischen sich teilweise ausschließenden oder aber mindestens überlappenden Zielgrößen und zusätzlich, deren Umsetzung. Solche potentiellen Konflikte müssen deutlich identifiziert werden. Zentral sind hier Fragen nach den Interessen und Werten, die mit den jeweiligen Maßnahmen verbunden sind und welche Ziele sie vorrangig befördern. An einem solchen Diskurs, der als Bewertungsdiskurs verstanden werden kann, sollten besonders organisierte Gruppen einer Gesellschaft eingebunden werden. „*Ziel dieses zweiten Diskurses ist es also, die mit einer Bewertung zwangsweise verbundenen Zielkonflikte zu verdeutlichen und die dadurch erforderlichen Prioritäten und Postprioritäten festzulegen*“ (Renn 2007: 183). Das Produkt eines Reflexionsdiskurses bildet eine Rangfolge von Maßnahmen, die idealiter ein abgestimmtes Maßnahmenpaket bilden. Mit diesem Maßnahmenpaket können implizit als relevant und handlungsbedürftig angesehene Dimensionen bestimmt werden (vgl. Renn 2007: 183). Ein *Gestaltungsdiskurs* bezieht sich letztlich auf die Auswahl und Umsetzung der in Frage kommenden Maßnahmen. Hier muss eine Entscheidung getroffen werden, die entweder von den von der Entscheidung selbst Betroffenen oder aber von Entscheidungsgremien, unter Einbeziehung der Präferenzen der Betroffenen, gestaltet werden müssen. Ziel eines solchen Diskurses ist die Legitimation der in den vorherigen Diskursarten behandelten Maßnahmen (vgl. Renn 2007: 182f.).

Zusammenfassend lässt sich damit eine Unterscheidung zwischen drei Diskursarten treffen:

1. Ein *Wissensdiskurs* über die zu erwartenden Folgen der potentiell wählbaren Maßnahmen.
2. Ein *Reflexionsdiskurs* als Abwägung von sich teilweise überlappenden oder sogar ausschließenden Zielgrößen.
3. Ein *Gestaltungsdiskurs* durch Auswahl und Umsetzung der in Frage kommenden Maßnahmen (vgl. Renn 2007: 182f.).

Gerade bei der Selektion und Begründung von Maßnahmenpaketen handelt es sich häufig um komplizierte Abwägungsprozesse, die für eine Umsetzung einer Gesamtstrategie zur Festlegung von Maßnahmenbündeln eine Integration von mehreren Diskursarten nötig machen. Das Gruppendelphi eignet sich als Methode einer wissensgestützten Argumentationsführung, mit dem Ziel „(...) *in der Regel, Konsens bzw. Konsens über Dissens über einen Wissensdiskurs zu erhalten*“ (Niederberger 2015a: 41). Folglich wäre ein Gruppendelphi idealtypisch als kognitiver Diskurs zu führen und nach dem Geltungsanspruch der Gültigkeit bzw. Wahrheit zu bewerten. Das exklusive Ziel wäre eine möglichst wirklichkeitsgetreue Abbildung und Erklärung (vgl. Renn 2007: 182f.; Renn 2000b: 127). Die einzelnen Aufgaben der Teilnehmenden eines Delphis sind realiter allerdings nicht so trennscharf, wie die Beschreibung der Art des Wissensdiskurses es vermuten lässt. Die Basis bildet zwar häufig die Bereitstellung von Wissen, auf dessen Grundlage die Teilnehmenden bewerten, beraten oder vorausschauen. Die Methode des Gruppendelphis geht in seiner praktischen Anwendung aber darüber hinaus: „*So werden bei einem Gruppendelphi die Experten oftmals gebeten, ihr Wissen zu teilen, aber auch Erkenntnisse oder normative Vorschläge zu bewerten (...)*“ (Niederberger 2015a: 39).

Folglich stellt Evidenz für die Prüfung der Funktionserfüllung eines Gruppendelphis ein wichtiges Geltungskriterium dar, die Geltungskriterien der Verständlichkeit, Wahrhaftigkeit und normativen Angemessenheit sollen aufgrund dessen, dass ein Gruppendelphi nicht als exklusiver Wissensdiskurs eingeordnet werden kann und sollte (vgl. Niederberger 2015a: 39), bei der tatsächlichen Durchführung eines Gruppendelphis für die weitere Analyse an diesem Punkt mit einbezogen werden. Wie die Geltungskriterien in das Forschungsdesign übernommen werden, soll im folgenden Kapitel näher beschrieben und ausgearbeitet werden.

6 Forschungsdesign

Bevor die Besonderheiten, die sich durch die Erweiterung des Delphiansatzes durch diskursive Elemente für die Forschungsfrage dieser Arbeit ergeben, in den Fokus genommen werden, soll nochmals kurz auf die Gemeinsamkeiten von klassischem Delphiverfahren und Gruppendelphi eingegangen und die in Kapitel 5 erläuterten wissenschaftstheoretischen Grundlagen Bezug genommen werden. Die vielzähligen Varianten eines Delphis haben gemeinsam, dass der Ablauf in einem mehrstufigen Aufbau erfolgt, der durch eine Rückkopplung der Zwischenergebnisse ergänzt wird (vgl. Kuhn et al. 2015: 318). Dies ist für die Grundannahme beider Variationen des Ansatzes nicht unerheblich, denn diese auf der Weiterführung des 1+n-Argumentes basierende Annahme (vgl. Dalkey 1969: 6), führt zu einem wichtigen Bemessungskriterium des Delphiansatzes. So ergibt es keinen Sinn, ein Delphi nach dem Wahrheitsgehalt seiner Ergebnisse zu bewerten, sondern danach, ob die gewonnenen Urteile rational sind. Das *Wie?* der Urteilsfindung rückt damit in den Fokus der Betrachtung, denn es liefert neben einer theoretischen Legitimation des Ansatzes zusätzlich eine Basis zur Bewertung konkreter Anwendungen des gesamten Delphiansatzes (vgl. Hader 2009: 40; Häder/Häder 2000: 23).

6.1 Fragestellung

Die theoretischen Überlegungen mithilfe des Linsenmodells in Kapitel 5.1 haben veranschaulicht, dass bei unsicheren Sachverhalten globale Wissensbestände nicht immer über rein analytische Verfahren in ein rationales Urteil überführt werden können. Neben quantitativen Informationen muss eine Person zur Urteilsbildung auch qualitative Informationen wie beispielsweise Argumente oder Beschreibungen nutzen (vgl. Jungermann et al. 2005: 164). Dies gilt insbesondere dann, wenn ein bereits gebildetes Urteil nochmals überprüft werden soll und neue Hinweisreize einbezogen werden sollen. Eine Kritik am klassischen Delphiverfahren knüpft an das Fehlen von qualitativen Informationen des Feedbacks während des Prozesses an. Das klassische Verfahren kann eine solche Auseinandersetzung aufgrund der Anonymität des Feedbacks nicht ausreichend gewährleisten. Ebenso kann nicht ausgeschlossen werden, dass die Teilnehmenden ein vorschnelles Urteil abgeben, da sie dies nicht argumentativ begründen müssen oder gar dafür verantwortlich gemacht werden können (vgl. Bowles 1999: 32; Vorgrimler/Wübben 2003: 767; McBride et al. 2012: 907).

Das Gruppendelphi versucht diesen Anspruch einzulösen, in dem es in einem Face-to-Face-Workshop den Austausch zwischen den Teilnehmenden über ihre jeweiligen Einschätzungen ermöglicht. Die Expert*innen können argumentativ ihre eigenen Urteile begründen und die Gefahr kann gemindert werden, dass ein übereiltes Urteil erfolgt. In einer solchen Situation ist es auch unwahrscheinlicher auf einem ursprünglichen (falschen) Urteil zu beharren, nur um dem Motiv zu folgen, sein Gesicht wahren zu wollen (vgl. McBride et al 2012: 907; Kuhn et al. 2014: 66f.).

Die Annahme, dass die diskursiven Elemente zu den beschriebenen Vorteilen eines Gruppendelphis gegenüber dem klassischen Verfahren führen, setzt allerdings voraus, dass der argumentative Austausch über die jeweiligen Urteile eine inhaltliche Verbesserung eben dieser

Urteile bewirkt und der diskursive Prozess während eines Gruppendelphis die konsensuale Annäherung der Expert*innenurteile unterstützt. Auch die Annahme, dass die Akzeptanz gegenüber neuen Informationen mit Einsatz dieser Methode durch vernünftige Begründungen befördert werde, kann nur gemacht werden, wenn die diskursiven Elemente einen solchen Austausch von Begründungen auch unterstützen. Hierfür müssen sich die Teilnehmenden über die jeweiligen der Bewertung zu Grunde liegenden individuellen Relevanzsysteme verständigen. Ebenso können normative Einstellungen nur durch den Austausch von Argumenten offengelegt werden und letztlich kann ein solcher Austausch zwischen den Expert*innen zur semantischen Klärung von Begriffen und zur Präzisierung von Fragestellungen und Konzepten beitragen. Das Ziel der Methode des Gruppendelphis ist somit, den Rahmen für einen (interdisziplinären) offenen und konstruktiven Diskurs zwischen Expert*innen bereitzustellen (vgl. Webler et al. 1991: 258; Schulz/Renn 2009: 20; Hilpert et al. 2013: 5; Kuhn et al. 2014: 67; Niederberger 2015b: 117).

Die Diskurstheorie steht dabei nicht im Widerspruch zu den wissenschaftstheoretischen Überlegungen des klassischen Delphiverfahrens, sondern stellt eine wissenschaftstheoretische Erweiterung der zusätzlichen Annahmen dar, die sich durch die Integration der diskursiven Elemente in das methodische Design des Gruppendelphis ergeben. Wenn die angenommenen Vorteile der Integration der diskursiven methodischen Elemente Bestand haben sollen, dann müssen die aufgestellten Anforderungen an einen Diskurs durch das methodische Design des Gruppendelphis eingelöst werden.

Da es sich - wie im vorgehenden Kapitel beschrieben - bei einem Gruppendelphi hinsichtlich der Art des Diskurses vorwiegend um einen Wissensdiskurs handelt, kann das Gruppendelphi hauptsächlich als kognitiver Diskurstypus begriffen werden. Ein solcher Diskurstyp eignet sich besonders dazu, komplexe Probleme aufzulösen und Lösungsstrategien zu entwickeln (vgl. Renn 2007: 182f.). Dies ist meist eines der primären Ziele des Einsatzes der betrachteten Methode. Durch die praktischen Erfahrungen mit der Durchführung von Gruppendelphis ist aber davon auszugehen, dass die Expert*innen ebenfalls ihre persönlichen Erfahrungen und normative Vorstellungen in die Bewertung und Argumentation der jeweiligen Urteile einbeziehen (vgl. Niederberger 2015a: 39). Dafür spricht zudem, dass sich die in einem Gruppendelphi behandelten Sachverhalte nicht immer nur als komplex, sondern aufgrund ihrer zukünftigen Ausrichtung ebenso durch hohe Unsicherheit - oder weiter gar durch hohe Ambiguität - auszeichnen. Solche Sachverhalte lassen sich nur schwer im strikten kognitiven Diskurs lösen, sondern benötigen zusätzlich Argumente, die in ihrer Lösungsperspektive auch auf das Gemeinwohl ausgerichtet sind. Sachverhalte mit hoher Ambiguität sollten somit neben der Ausrichtung auf das Gemeinwohl ebenfalls eine Beachtung der normativen Richtigkeit von Lösungsansätzen zulassen (vgl. Renn 2008: 280). Deshalb ist es für die Betrachtung der Methode des Gruppendelphis wichtig, auch normative Aussagen in die Auswertungen der einzelnen Fälle praktischer Anwendungen des Gruppendelphis mit einzubeziehen.²⁹ Überdies sollte im Diskurs zwischen den Expert*innen³⁰ darauf geachtet werden, dass die Verständlichkeit von kommunikativen Aussagen als Kriterium nicht vernachlässigt wird. Falls kein gemeinsames Verständnis besteht, sollte sich auf ein gemeinsames geeinigt werden. Im Rückgriff auf

²⁹ In einer Veröffentlichung der Rand Corporation hat Quade bereits normative Bewertungen als bewertbare Sachverhalte eines Delphis miteingeschlossen (vgl. ebd. 1970: 16).

³⁰ Besonders wenn diese einen unterschiedlichen disziplinären Hintergrund haben.

Kapitel 5.2 sollen allerdings affektive Äußerungen für den Diskurs während eines Gruppendelphis ausgeschlossen werden, denn deren expressive Rationalität lässt sich in diesem Rahmen weder interpretieren, noch im Diskurs intersubjektiv von den anderen Teilnehmenden als gültig beurteilen. Außerdem sind sie für die zu behandelnden Sachverhalte nicht als zielführend einzustufen. Wie bereits angemerkt wurde, sollten affektive Aussagen somit generell in kognitive oder normative Aussagen übersetzt werden (vgl. Renn 2007: 178).

Dies führt den argumentativen Bogen zurück zur Theorie des kommunikativen Handelns, in der die Verknüpfung von einem auf Wissen, über die möglichen Konsequenzen basierenden gemeinsamen Verständnis von Handlungen, mit eben jenen Werten verknüpft werden, die im Diskurs gelöst werden sollten (vgl. Renn 2000a: 20). Ein Gruppendelphi sollte dem optimistischen Anspruch der Modifikation Rechnung tragen und mittels seines methodischen Aufbaus einen intensiven Austausch von Begründung und Argumenten befördern (vgl. Kuhn et al. 2014: 66). Wenn die angenommenen Vorteile der diskursiven Elemente Bestand haben sollen, dann müssen durch das Gruppendelphi folglich auch die Anforderungen an einen Diskurs eingelöst werden, um diese Vorteile auch für die Durchführung der Methode konstatieren zu können. Die **zentrale Fragestellung für diese Arbeit** lautet folglich:

- **Können die Qualitätsmerkmale von handlungsleitenden Diskursen von der Methode des Gruppendelphis eingelöst werden?**

Um dieser Frage gerecht zu werden, muss die Logik der Argumentation aus der Theorie des kommunikativen Handelns auch für das Gruppendelphi nachweisbar sein, und dem *zwanglosen Zwang des besseren Argumentes* eine entsprechende Rahmung bieten. Die Form der Argumentation muss somit während eines Gruppendelphis in ihrer Funktionsweise als ein Lernprozess sowie einer Berufungsinstanz stattfinden, um auch strittige Geltungsansprüche zu thematisieren, mittels inhaltlicher Begründungen und durch Argumente zu kritisieren oder zu bestätigen (vgl. Habermas 1981: 37-53). Für die Fragestellung werden damit die Typen von Aussagen und ihre Geltungsansprüche innerhalb eines Diskurses zu einem zentralen Analyseschema. Mittels der Geltungskriterien lässt sich überprüfen, ob die Merkmale kommunikativer Rationalität anhand der vorliegenden Anwendungsbeispiele durch die Methode des Gruppendelphis eingelöst werden können, und am Ende des Prozesses für die einzelnen diskutierten Einschätzungen der Sachverhalte ein Konsens bestehen kann, „(...) *der auf der intersubjektiven Anerkennung kritisierbarer Geltungsansprüche beruht*“ (Habermas 1981: 37).

6.2 Operationalisierung

Für die empirischen Fallbeispiele der methodischen Anwendungen eines Gruppendelphis sollten sowohl die Geltungsansprüche sowie deren Einlösung mittels kommunikativer, kognitiver und normativer/evaluativer Diskurse offengelegt werden. Falls das Gruppendelphi einen Rahmen für die Einlösung dieser Geltungsansprüche innerhalb eines handlungsleitenden Diskurses bietet, können die Ansprüche einer inhaltlichen Verbesserung der Urteile, einer konsensualen Annäherung der Urteile und letztlich einer höheren Akzeptanz der Integration neuer Informationen erzielt werden. Was muss passieren, damit diese Ansprüche erreicht werden?

1. Die Gültigkeit der einzelnen Argumente kann in Frage gestellt und mittels eines diskursiven Austauschs zwischen den teilnehmenden Expert*innen überprüft werden.
2. Einer semantischen Klärung von Begriffen und Definitionen muss (gerade im interdisziplinären Kontext) ein Rahmen geboten werden.
3. Es muss ein diskursiver Rahmen existieren, in dem die Teilnehmenden ihre normativen Einstellungen und Wertvorstellungen offenlegen können.
4. Es muss für die Teilnehmenden möglich sein, eine Verständigung über die, den Beurteilungen zugrundeliegenden Relevanzsysteme, zu erzielen.

Für das Untersuchungsdesign erfolgt ein Rückgriff auf die in Kapitel 5.2.3 dargestellten Geltungsansprüche und deren Einlösung im Diskurs, denn sowohl die Aussagetypen als auch deren jeweilige Kriterien bilden Hilfsmittel, um zulässige von unzulässigen Äußerungen während eines Diskurses zu trennen (vgl. Renn 2007: 177). Tabelle 4 fasst diese für das Gruppendelphi relevanten Geltungskriterien zusammen, die für die weitere Analyse als zentrale Kategorien des Untersuchungsdesigns angewandt werden und die für eine Auswertung der empirischen Fallbeispiele in dieser Arbeit in den folgenden Unterkapiteln operationalisiert werden.

Tabelle 4: Typen von Aussagen und ihre Geltungsansprüche

Geltungskriterium	Evidenz	Verständlichkeit	Normative Angemessenheit
Aussagentyp	kognitiv	kommunikativ	evaluativ/normativ
Geltungsanspruch	Wahrheit (Gültigkeit) ³¹	Zweckmäßigkeit	Gemeinwohl (evaluativ)
			Richtigkeit (normativ)

Quelle: Eigene Darstellung in Anlehnung an Renn (1999: 72).³²

6.2.1 Dimension: Evidenz

Da es sich bei einem Gruppendelphi vorwiegend um den Typus eines Wissensdiskurses handelt, wird die Dimension der Evidenz als zentrale Dimension im Testdesign verortet. Dieses Geltungskriterium folgt grundsätzlich dem Geltungsanspruch der Wahrheit. Im Hinblick auf die, im Gruppendelphi diskutierten Sachverhalte allerdings vielmehr dem Geltungsanspruch der Gültigkeit, denn die angestrebten Ergebnisse sollen als *rational* bemessen werden, und dabei ein wahrheitsanaloges Geltungskonzept anstreben, dass vorhandene raumzeitliche und soziale Beschränkungen beachtet (vgl. Habermas 1981: 56; Renn/Webler 1998: 51; Häder 2009: 40; Niederberger 2015a: 41).

Grundlegend für die Einschätzungen und Bewertungen der Expert*innen ist unter dieser Dimension vorwiegend systematisches Wissen. Überdies können die Expert*innen argumentativ ihre Urteile bezüglich bestimmter Sachverhalte auf personelles bzw. anekdotisches Wissen - wie auch gar auf ihre eigene Intuition stützen (vgl. Renn 1999: 73). Die Frage des *Wie ein*

³¹ Eigene Ergänzung.

³² Das Geltungskriterium der Evidenz wurde in dieser Tabelle dem Geltungskriterium der Verständlichkeit vorangestellt, da diese Reihenfolge dem Aufbau des Untersuchungsdesigns entspricht.

Sachverhalt beurteilt bzw. in Zukunft gelöst werden kann, steht für diese Dimension im primären Fokus und stellt zusätzlich eine Bewertung konkreter Ansätze innerhalb des gesamten Delphi-Ansatzes dar (vgl. Häder 2009: 40). Das Ziel des Diskurses ist in diesem Fall nicht mehr Einigkeit, sondern mehr Klarheit über Sachverhalte zu gewinnen, in einem Rahmen in dem problematisch (gewordene) Geltungsansprüche zum Thema gemacht, auf ihre Berechtigung hin untersucht werden und in ihrer Ausrichtung der instrumentellen Rationalität folgen (vgl. Habermas 1973: 214; Renn 2007: 180). Die empirischen Fälle müssen in dieser Dimension damit nicht dem Anspruch genügen, dass nicht nur konsensuale Ergebnisse für das Gruppendelphi unter dem Geltungskriterium der Evidenz angenommen werden können. Vielmehr veranschaulichen auch Ergebnisse, die sich vom ursprünglichen Mittelwert der ersten Erhebung hin zu einem anderen Wert bewegen, dass der Austausch zu einer Veränderung der kollektiven Einschätzung führt, wie auch eindeutige Konsense über Dissens.

Die Bewertungen der abgefragten Sachverhalte durch die Teilnehmenden müssen insofern *rational* sein, als dass sie es ermöglichen Handlungsoptionen auszuarbeiten, die trotz Unsicherheit³³ von allen Teilnehmenden am Ende gemeinsam getragen werden. Aufgrund des Austauschs der Begründungen und Argumente für die jeweiligen subjektiven Bewertungen in den diskursiven Phasen eines Gruppendelphis, sollten sich die im Prozess erhobenen statistischen Bewertungen der teilnehmenden Expert*innen immer mehr annähern (Konsens) bzw. klarer an zwei (oder mehr Werten) konzentrieren (Konsens über Dissens). Diese Annäherung der Urteile folgt in seinem logischen Ablauf der von Rowe et al. beschriebenen Tendenz der Gruppe hin zur Einschätzung der *Holdouts* durch Urteilsänderung der *Swingers* (vgl. ebd. 1991: 38). Im Falle des Gruppendelphis werden die *Swingers* vom *zwanglosen Zwang des besseren Argumentes* der *Holdouts* von einem anderen Urteil überzeugt. Damit folgen sie nicht nur (wie in der Grundannahme von Rowe et al.) der statistischen Mehrheit, sondern sie nehmen die, in den diskursiven Phasen dargestellten Begründungszusammenhänge und die transparent gemachten Argumentationen an (vgl. Rowe et al. 1991: 38; McBride et al 2012: 907; Kuhn et al. 2014: 67).

Der am häufigsten eingesetzte Fragetypus bei Gruppendelphis sind 10er-skalierte Ratingskalen. Diese Skalierung ermöglicht es, die Bewertung der Teilnehmenden anhand von statistischen Maßzahlen zügig aufzubereiten und die Diskussion im Plenum anhand dieser Berechnungen zu strukturieren. Diese deskriptiven Werte bieten überdies die Möglichkeit die Entwicklungen der jeweiligen Urteile der Expert*innen auch nach einem Gruppendelphi mittels der einzelnen Items nachvollziehen zu können. Aufgrund der quasi-metrischen Skalierung³⁴ lässt sich, neben dem arithmetischen Mittel sowie der Standardabweichung, „(...) *dem üblichen Streuungsmaß für (mindestens) intervallskalierte Messwerte (...)*“ (Diekmann 2009: 679), ebenfalls das relative Streuungsmaß des Variationskoeffizienten errechnen, der als Quotient aus Standardabweichung und arithmetischem Mittel gebildet wird. Der Variationskoeffizient hat dabei den Vorteil, dass er ein Verhältnis zum arithmetischen Mittel der Häufigkeitsverteilung abbildet und für sich genommen interpretiert werden kann (vgl. Benninghaus 2002: 62; Diekmann 2009: 678ff.). Eben dieser Variationskoeffizient ermöglicht eine Einschätzung darüber, ob die einzelnen Bewertungen der Expert*innen als konsensual eingestuft werden können. Überdies wirkt sich zusätzlich die Bewertung des einzelnen Items auf der zehnstufigen Skala

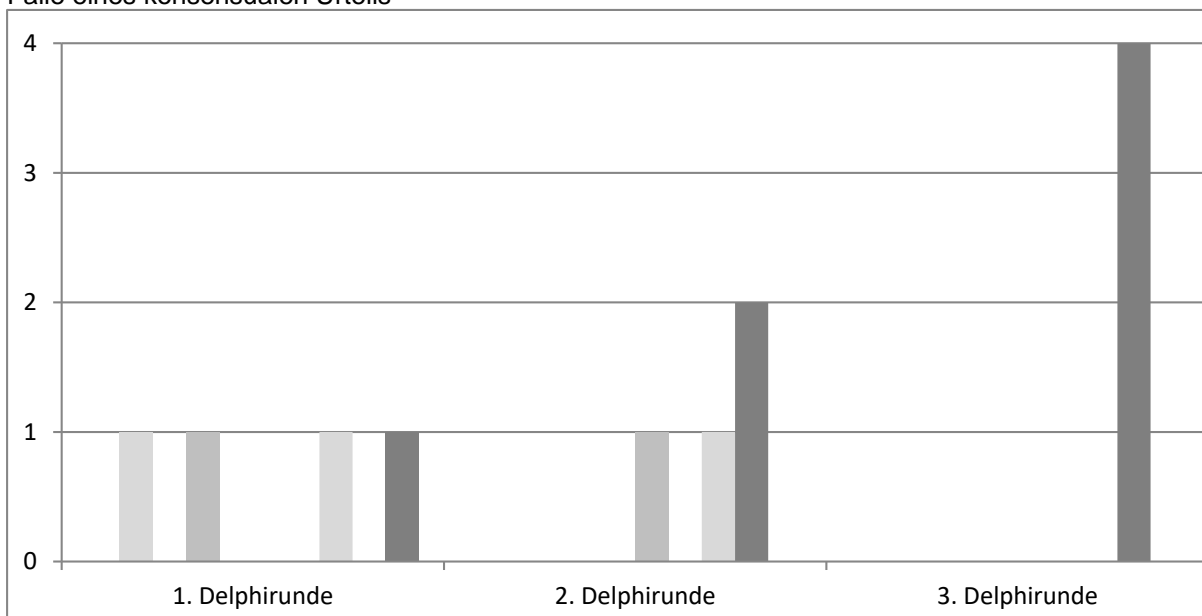
³³ Siehe Kapitel 3.4.

³⁴ Siehe Kapitel 4.2.1.

auf den Wert des Variationskoeffizienten aus, so dass auch Fragen bzw. Items mit hoher Ablehnung gezielt im Plenum diskutiert und die Gründe für die Ablehnung erörtert werden können. Ab einem Variationskoeffizienten von $\geq 0,5$ wird bei einer praktischen Durchführung eines Gruppendelphis von einem Dissens ausgegangen, Items die diesen Wert annehmen oder über diesem Quotienten liegen, werden damit vornehmlich diskutiert. Diese Schwelle basiert auf Erfahrungen durch vergangene Anwendungen der Methode des Gruppendelphis, sollte als empirischer Richtwert damit allerdings nicht exklusiv ausgelegt werden (vgl. Kuhn et al. 2014: 72f.).

Sollen die Qualitätsmerkmale von handlungsleitenden Diskursen eingelöst werden, so sollten sich über den Verlauf eines Gruppendelphis die Werte in Richtung eines Konsenses bewegen. Für die Häufigkeitsverteilung der einzelnen Items bedeutet dies, dass der Wert der Standardabweichung über die Delphirunden hinweg sinken sollte und sich im Ideal alle Einschätzungen der Kleingruppen an oder idealiter auf einem Wert der 10er Skala bündeln. Kurz, je mehr Delphirunden stattfinden, desto mehr konzentrieren sich die Einschätzungen der Expert*innen an einem Wert auf der Skala. Abbildung 7 veranschaulicht beispielhaft eine solche angenommene Entwicklung der Antwortverteilung der einzelnen Kleingruppen über die Runden des Gruppendelphis hinweg.

Abbildung 7: Ideale Annäherung der Bewertungen der Gruppen im Verlauf eines Gruppendelphis im Falle eines konsensualen Urteils

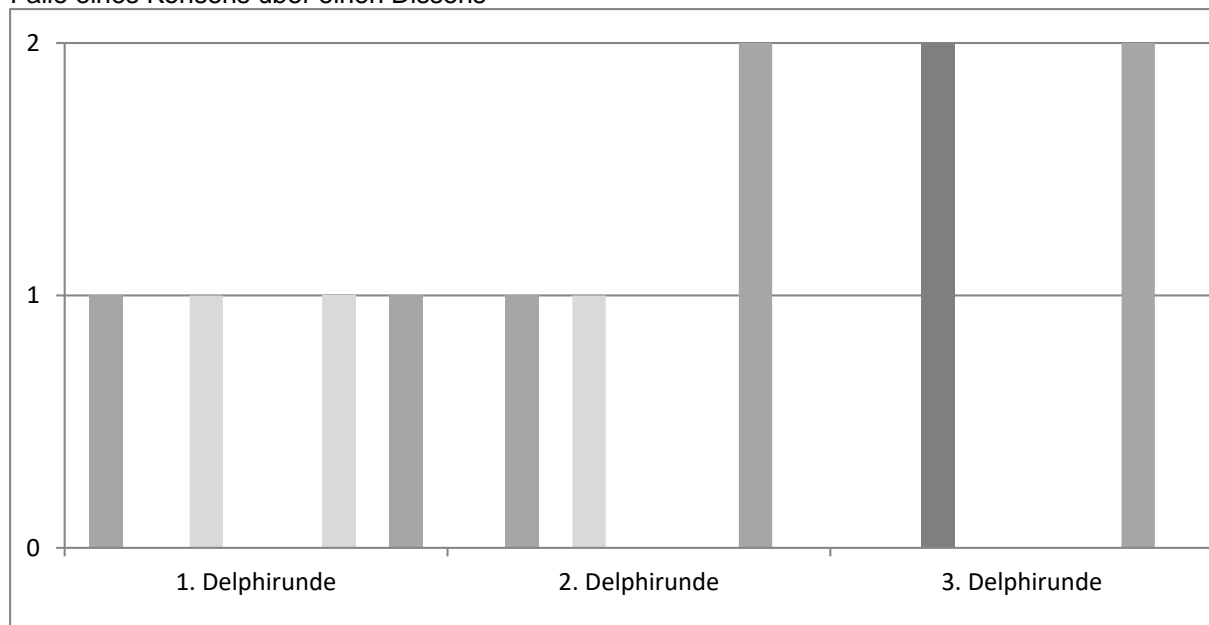


Die Balken zeigen die Verteilung der Antworten der einzelnen Kleingruppen auf der 10er-Skala über die Erhebungszeitpunkte hinweg. N=4. Quelle: Eigene Darstellung

Sollte ein Austausch über systematisches Wissen wie personelles bzw. anekdotisches Wissen und Intuition unter den Expert*innen stattfinden, ist davon auszugehen, dass eine solche Entwicklung in den Antwortmustern angezeigt wird. Der *zwanglose Zwang des besseren Argumentes* würde sich damit auch in den statistischen Werten widerspiegeln. Folglich sollte auch der Wert des Variationskoeffizienten bei den im Plenum diskutierten Items über die Delphirunden hinweg unter den ursprünglichen Schwellenwert von $\geq 0,5$ sinken.

Ebenfalls ist ein Konsens über einen Dissens unter den angenommenen Wirkmechanismen zulässig. So veranschaulicht dieser unter anderem zusätzlichen Forschungsbedarf, der beispielsweise auf den gegebenen Unsicherheiten basieren kann.³⁵ Je mehr Delphirunden stattfinden, desto mehr sollten sich die Einschätzungen der Expert*innen in diesem Falle an zwei Werten³⁶ konzentrieren. Bei einem Konsens über einen Dissens ist von höheren Werten der Standardabweichung auszugehen. Dadurch erhöht sich folglich auch der Wert des Variationskoeffizienten, da die Standardabweichung dessen Dividenden bildet. Die Werte können also bei einem Konsens über einen Dissens, anders als bei einem konsensualen Urteil, auch über einem Variationskoeffizienten von $\geq 0,5$ liegen. Darum ist es in diesen Fällen umso wichtiger die Häufigkeitsverteilungen der einzelnen Items zu betrachten, da auch das arithmetische Mittel nicht die jeweiligen Bewertungen der einzelnen Expert*innengruppen adäquat widerspiegelt. Abbildung 8 veranschaulicht eine idealtypische Verteilung der Häufigkeiten im Falle eines Konsenses über einen Dissens und dessen diskursive Entwicklung über die Runden des Gruppendelphis hinweg.

Abbildung 8: Ideale Annäherung der Bewertungen der Gruppen im Verlauf eines Gruppendelphis im Falle eines Konsens über einen Dissens



Die Balken zeigen die Verteilung der Antworten der einzelnen Kleingruppen auf der 10er-Skala über die Erhebungszeitpunkte hinweg. N=4. Quelle: Eigene Darstellung

Die in dieser Abbildung dargestellte Verteilung der Urteile zeigt, dass auch ein Konsens über einen Dissens nicht beliebig ist und nicht jeder Dissens als Konsens über einen Dissens bezeichnet werden kann. Allerdings kann hier die verwendete deskriptive Statistik keine Kenngrößen bereitstellen, an denen ein solcher erkennbar wäre. So ist auch unter dieser Dimension wichtig, die einzelnen Begründungszusammenhänge für die jeweiligen Urteile der Teilnehmenden zu betrachten, um beispielsweise Scheinkonsense ausschließen zu können. Ebenso müssen die einzelnen Häufigkeiten und deren Verteilung beachtet werden. Dies gilt ebenso

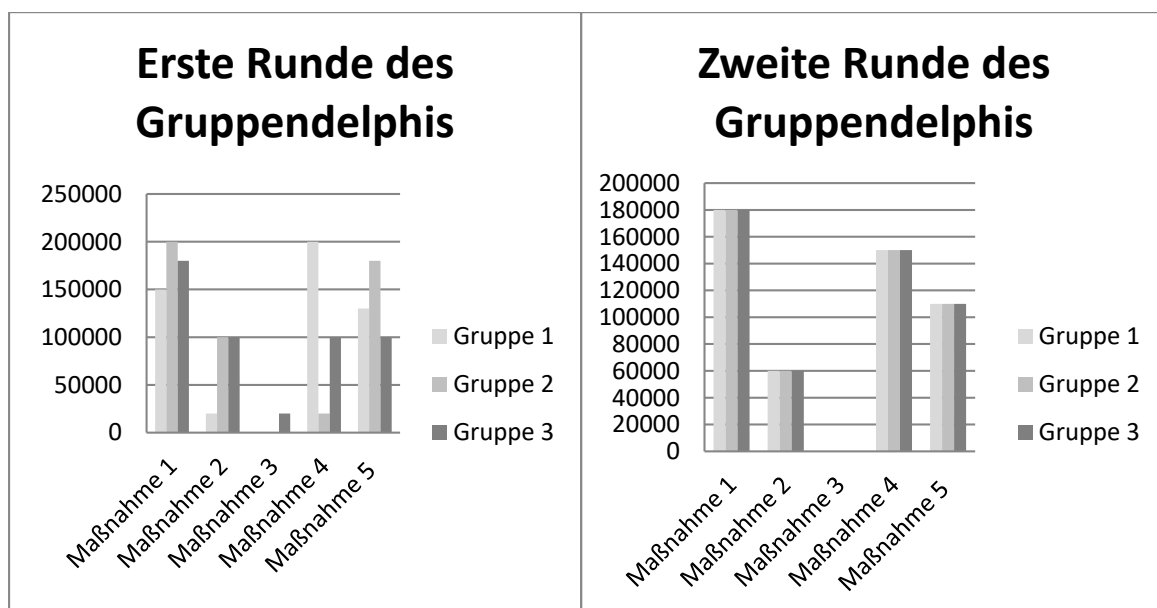
³⁵ Siehe Kapitel 5.2.5.

³⁶ Es ist durchaus möglich, dass ein Konsens über einen Dissens sich auch an mehr Werten ausdrückt, allerdings ist die Fallzahl mit zwei bis vier Kleingruppen zu gering, um dies für die betrachteten Beispiele als eindeutiges Antwortmuster zu erwarten.

für die anderen Fragetypen, die in einem Fragebogen eines Gruppendelphis eingesetzt werden können.

Beim Einsatz von Budgetfragen findet während der Pause eines Gruppendelphis häufig eine grafische Aufbereitung der von den Kleingruppen verteilten Budgets, beispielsweise mittels in der Pause erstellter Balkendiagramme, statt. Dies bietet den Vorteil, dass den Teilnehmenden des Plenums anschaulich gezeigt werden kann, wo die Unterschiede und Gemeinsamkeiten in den Budgetverteilungen zwischen ihren Kleingruppen liegen. Da bei einer solchen Rankingskala ein vorgegebenes Budget von den teilnehmenden Expert*innen auf die verschiedenen Maßnahmen aufgeteilt wird, sollten die Maßnahmen anders als bei den 10er-skalierten Ratingskalen nicht nur einzelnen für sich betrachtet werden, sondern im Kontext der Gesamtverteilung. Im Ideal bildet sich auch hier im Laufe des Gruppendelphis zwischen den einzelnen Gruppen eine konsensuale Verteilung des Budgets heraus. Abbildung 9 stellt eine solche idealtypische Entwicklung über zwei Runden des Gruppendelphis dar. Budgetfragen bieten aufgrund der fiktiven monetären Werte den Vorteil, dass sie für die Teilnehmenden eine realitätsnahe Bewertung der Maßnahmen darstellen.³⁷

Abbildung 9: Darstellung einer idealtypischen Häufigkeitsverteilung bei Budgetfragen über zwei Runden des Gruppendelphis



Anmerkungen: Die Abbildung stellt die Verschiebungen der Häufigkeiten bezüglich der gewählten Budgetausgaben für einzelne Maßnahmen dar. Quelle: Eigene Darstellung

So ist der Vergleich der einzelnen Budgets auch für die Teilnehmenden leicht nachvollziehbar und begründbar, was für die Auswertung neben einem Vergleich der einzelnen Häufigkeitsverteilungen hinweg, auch einen Konsens über einen Dissens nicht vorwiegend an unterschiedlichen Verteilungen des Budgets, sondern besonders an den im Plenum geäußerten Argumenten der einzelnen Gruppen erkenntlich macht.

Bei der Auswahl gezielter Maßnahmen aus einer Liste von, einen Sachverhalt potentiell adressierenden Maßnahmen sollten sich die Expert*innen am Ende des Gruppendelphis auf identische Maßnahmen geeinigt haben. Eine solche Abfrage von Maßnahmen kann zudem

³⁷ Siehe Kapitel 4.2.1.

als Rankingfrage aufgebaut sein, in der beispielsweise eine Rangfolge der priorisierten Maßnahmen durch die Teilnehmenden gebildet werden soll. Am Ende des Workshoptages sollte hier eine von den Expert*innen konsensual getragene Rangfolge der Maßnahmen ausgearbeitet sein.

Auch kann eine prozentuelle Abfrage von Maßnahmen bezüglich einzelner Zielvorstellungen eingesetzt werden, in der mittels Prozentangaben die Wirksamkeit einzelner Maßnahmen bezüglich verschiedener Zielvorstellungen von den teilnehmenden Expert*innen bewertet wird. Da dieser Fragetyp als Ratingskala konzipiert ist, sollten die Maßnahmenbewertungen getrennt betrachtet werden. Methodisches Ziel des Gruppendelphis ist idealiter eine konsensuale Prozentverteilung für die einzelnen Maßnahmen zwischen den Urteilen der Expert*innen herzustellen.

Die Betrachtung der verschiedenen Fragetypen lässt die Vermutung zu, dass sowohl die Entwicklung hin zu einem Konsens, wie auch zu einem Konsens über einen Dissens auf diesem Wege erkennbar ist. Ebenso die Frage, ob die Gültigkeit der einzelnen Argumente in Frage gestellt und mittels eines diskursiven Austauschs zwischen den teilnehmenden Expert*innen überprüft werden kann. Die Teilnehmenden müssen den Rahmen des Gruppendelphis anerkennen und eine Verständigung darüber erzielen, nach welchen Kriterien sie die Richtigkeit von kognitiven Aussagen beurteilen (vgl. Renn 1999: 73).

6.2.2 Dimension: Verständlichkeit

Wie bereits in der Herleitung der Fragestellung beschrieben, zeigen Erfahrungen mit Gruppendelphis, dass es während der diskursiven methodischen Elemente nicht nur um einen exklusiven Austausch von Wissen geht, sondern auch Formulierungen für die Bewertung einzelner Sachverhalte wichtig und entscheidend sein können (vgl. Niederberger 2015b: 134). Die Dimension der Verständlichkeit dient der Überprüfung, ob die Methode des Gruppendelphis dem Geltungskriterium der Verständlichkeit Rechnung tragen und somit dem Geltungsanspruch der Zweckmäßigkeit für den Diskurs entsprechen kann. Um die während eines Gruppendelphis von den Teilnehmenden getätigten kommunikativen Aussagen bewerten zu können, muss veranschaulicht werden, dass die teilnehmenden Expert*innen mit dem *Geäußerten* einen gleichen Bedeutungsinhalt verbinden. Sollte das nicht der Fall sein, so muss das Gruppendelphi einen Raum dafür bieten, eben einen solchen gemeinsamen Bedeutungsinhalt für die strittigen Begriffsdefinitionen, Klassifikationsschemata und Formulierungen von Aussagen auszuarbeiten und zu erzielen. Ein solcher Fall kann eintreten, wenn Begriffe und Aussagen inkonsistent oder unklar genutzt werden oder der Aufbau von einzelnen Argumenten nicht logisch nachvollziehbar ist. Es geht dabei in der Dimension der Verständlichkeit nicht darum, die Wahrheit der jeweiligen Aussagen anzuzweifeln, sondern ein gemeinsames, dem Zweck entsprechendes, Verständnis der strittigen Äußerungen zu erreichen (vgl. Renn/Webler 1998: 50; Renn 1999: 72f.).

Da bei der Konstruktion eines Delphifragebogens beispielsweise Fachtermini³⁸ Verwendung finden, kann es vorkommen, dass Expert*innen verschiedener Disziplinen zwar mit dem Begriff vertraut sind, allerdings andere Konzepte und Deutungsinhalte mit eben diesem verbinden. Auch können semantische Missverständnisse anfangs zu unsicheren Urteilen führen, deren Begründung den Teilnehmenden ohne die Klärung eines gemeinsamen Verständnisses und einer zweckmäßigen Verwendung der Formulierung nicht möglich ist. *„Unsichere Antworten sind häufig ein Resultat semantischer Missverständnisse, nicht zwangsläufig ein Zeichen für Unsicherheit“* (Schulz et al. 2009: 55). Ebenso kann die konkrete Formulierung einer der im Fragebogen zu bewertenden Aussagen für einen Dissens verantwortlich sein, da eine Aussage beispielsweise semantisch nicht klar formuliert ist oder einen wichtigen Aspekt nicht beinhaltet (vgl. Niederberger/Kuhn 2013: 63).

„Gerade bei Handlungsempfehlungen, die später an die Politik bzw. den Auftraggeber der Studie weiter gegeben werden sollen, spielt die Wortwahl eine wichtige Rolle“ (Niederberger 2015b: 134). Besonders in solchen Fällen zeigt sich, dass die teilnehmenden Expert*innen bei Formulierungen einzelner Aussagen sehr genau deren Zweckmäßigkeit prüfen und auf eine unmissverständliche Semantik Wert legen. In bisher durchgeführten praktischen Anwendungen finden sich häufig Beispiele bei denen besonders die im Fragebogen abgefragten Aussagen im Plenum neu formuliert wurden oder ein gemeinsamer Kontext für die Gesamtbewertung erörtert wurde (vgl. Deuschle/Sonnberger 2009: 75; Hilpert et al. 2013: 24; Kuhn et al. 2014:76f.; Kuhn et al. 2015: 326f.)

Wie ist nun erkennbar, ob einer semantischen Klärung von Begriffen und Definitionen während eines Gruppendelphis ein entsprechender Rahmen geboten wird und ob ein solcher Diskurs unter den Teilnehmenden überhaupt stattfinden kann?

Während eines Gruppendelphis besteht die Möglichkeit den Fragebogen zwischen der Plenarphase und der nächsten Runde in den Kleingruppen anzupassen. In dieser Arbeitsphase wird der Fragebogen nicht nur auf die weiterhin strittigen Items gekürzt, sondern die im Plenum beschlossenen Umformulierungen einzelner Fragen, Items bzw. Aussagen über Maßnahmen werden in diesen angepassten Fragebogen integriert.³⁹ Für die Betrachtung der Fallbeispiele ist ein Vergleich zwischen den in den jeweiligen Gruppendelphis eingesetzten Fragebögen somit der erste Schritt der Analyse. Wenn Umformulierungen stattgefunden haben, so sind diese bei einem Vergleich der Fragebögen der verschiedenen Runden erkennbar und auch die Plenardiskussion, in der die Umformulierung stattgefunden hat, da diese folglich direkt vor der Erstellung des neuen Fragebogens geschehen sein muss. Ob die Umformulierung den Geltungsanspruch der Zweckmäßigkeit erfüllt, kann über zwei weitere Analyseschritte überprüft werden. Zum einen sollte die Umformulierung zu einer eindeutigeren Bewertung führen, das heißt konkret, zu einem Konsens, falls alle Expert*innen nach der semantischen Klärung auch den Geltungsanspruch der Gültigkeit für das jeweilige Item bzw. die jeweilige Aussage gleichermaßen teilen. Allerdings kann eine Umformulierung ebenso zu einem Konsens über einen Dissens führen, da die semantische Klärung das Erkennen dieses - auch als kognitiver Aussagetypus - ermöglicht hat. Nimmt man an, dass ein Gruppendelphi auch dem Geltungs-

³⁸ Siehe Kapitel 4.2.3.

³⁹ Siehe Kapitel 2.1.1.

kriterium der Verständlichkeit Rechnung trägt, sollte eine Umformulierung somit zu einem klareren Ergebnis in der nächsten Runde des Gruppendelphis führen und sich auf die identische Weise anhand der statistischen Werte und Antwortmuster prüfen lassen, wie bereits unter der Dimension der Evidenz beschrieben.⁴⁰ Zusätzliche Hinweise können zudem die Protokolle der jeweiligen Gruppendelphis liefern, in denen Umformulierungen und deren Begründung in der Regel schriftlich festgehalten werden und die damit für eine inhaltliche Auswertung der einzelnen Fallbeispiele zur Verfügung stehen.

6.2.3 Dimension: normative Angemessenheit

Die Frage des *Wie ein Sachverhalt beurteilt bzw. in Zukunft gelöst werden kann*, beinhaltet häufig nicht nur Fragen danach, was beispielsweise die beste technische Lösung für einen Sachverhalt sein kann, sondern auch, was eine *wünschenswerte* Lösung für die Zukunft sein könnte.⁴¹ So wird häufig das Bedenken geäußert, dass ein Urteil das lediglich auf wissenschaftlichem und technischem Wissen begründet wird und folglich wertegeleitetes Denken und normative Vorstellungen gänzlich ausschließt, die Qualität der Entscheidung negativ beeinflusst. Es muss also eine Verknüpfung von einem auf Wissen basierenden gemeinsamen Verständnis von Handlungen und deren möglichen Konsequenzen mit den relevanten Werten stattfinden (vgl. Renn 2000a: 20; Failing et al. 2004: 1). *„(S)cience can help get the facts right but ultimately whether a suite of impacts is considered on balance, either „good“ or „bad“ is a question of values – a proper topic for dialogue, but one about which science is moot (...)“* (Gregory 2016: 12). Das Delphiverfahren wird als besonders geeignet erachtet Probleme zu bearbeiten, die eine Mischung aus wissenschaftlicher Beweisführung und sozialen Werten aufweisen, da es sich eignet die Ursprünge und Hintergründe solcher Wertkonflikte offen zu legen (vgl. Rescher 1969: 12; Webler et al. 1991: 256). Damit spielen Werte, die das Gemeinwohl berühren, wie auch subjektive Einstellungen der Teilnehmenden über die normative Richtigkeit eine Rolle für die Bewertungen der einzelnen Sachverhalte. Niederberger ist allerdings der Meinung, dass persönliche Einstellungen und Werte der Teilnehmenden besonders bei Uneinigkeit bezüglich einzelner Grundsatzfragen nicht in die Diskussion während eines Gruppendelphis mit einfließen dürfen (vgl. ebd. 2015b: 132). Die Rolle der Expert*innen während eines Gruppendelphis sollte ebenfalls nicht so ausgelegt werden, die Rolle von Entscheidungsträger*innen einzunehmen und eine finale Wertung und normative Entscheidung herbeizuführen. Allerdings sollten unter Anbetracht einer umfassenden Bewertung eben jene normativen Betrachtungen unter dem Gesichtspunkt der umfassenden Bewertung in die angenommenen Konsequenzen und wahrscheinlichen Implikationen einzelner Maßnahmen mit einfließen können (vgl. Failing et al. 2004: 16).

Das Gruppendelphi muss folglich einen diskursiven Rahmen ermöglichen, in dem die Teilnehmenden ihre normativen Einstellungen und Wertvorstellungen offenlegen können. Wenn die betrachteten Fälle weitergehend dem Geltungskriterium der normativen Angemessenheit entsprechen sollen, so müssen die, während der diskursiven Phasen geäußerten, evaluativen Aussagen den Kriterien der Kohärenz und Verallgemeinerungsfähigkeit sowie auch die nor-

⁴⁰ Siehe Kapitel 6.2.1.

⁴¹ Siehe Kapitel 5.2.

mativen Aussagen dem Gebot der Universalisierung und der Kompatibilität mit Gesetzen entsprechen. Eine solche Argumentation würde Ergebnisse befördern, die den Geltungsansprüchen des Gemeinwohls und der normativen Richtigkeit folgen, wenn die normative Angemessenheit während eines Gruppendelphis bei einzelnen Sachverhalten zum Thema erhoben wird (vgl. Renn 1999: 72).

Der Austausch über die normative Angemessenheit oder die Gemeinwohlorientierung einzelner Maßnahmen sind - wie bereits beschrieben - nicht das primäre Forschungsinteresse eines Gruppendelphis, sondern im Hinblick auf einzelne von den Teilnehmenden geäußerten Begründungen ihrer Urteile sowie Argumente für die Einschätzung eines einzelnen Sachverhaltes relevant.⁴² Anders als die Dimension der Verständlichkeit lässt sich die hier betrachtete Dimension der normativen Angemessenheit nicht mittels eines Vergleiches der Fragebögen innerhalb eines Falles überprüfen, da solche Argumente nicht als semantische Umformulierung in den jeweiligen Fragebogen der folgenden Runde des Gruppendelphis aufgenommen werden. Der Fokus muss vielmehr auf die von den Teilnehmenden geäußerten Begründungen während der Plenarphase gelegt werden. Wie bereits in Kapitel 4.1.1 beschrieben, stehen für eine Auswertung dieser Äußerungen in der Regel Protokollierungen der Diskussionen bzw. Tonaufnahmen oder audiovisuelle Aufzeichnungen zur Verfügung (vgl. Renn/Webler 1998: 80; Deuschle/Sonnberger 2009: 70; Gallego Carrera 2009a: 103; Schulz/Renn 2009: 115; Kuhn et al. 2014: 78).

So wird in diesem Fall eine rein qualitative Analyse der Aufzeichnungen und Protokolle aus den einzelnen Fallbeispielen gewählt, um zu überprüfen, ob das Gruppendelphi einen Rahmen bietet, in dem die Teilnehmenden ihre normativen Einstellungen und Wertvorstellungen offenlegen können. Eine offene Erhebung soll eben ein solches *hidden curriculum* vermeiden. Die qualitative Analyse dieser Erhebungen bietet nach Kuckartz et al. weiterhin Vorzüge gegenüber starren Bewertungskonzepten einer Skala, weil sie eine differenziertere Betrachtung der Ergebnisse ermöglichen, da die erhobenen Daten wesentlich komplexer sein können. Eine solche Betrachtungsweise kann auch mögliche Widersprüche in den Bewertungen und Äußerungen der Teilnehmenden aufdecken. Auch die Hintergründe, die hinter Ansichten und Urteilen stehen, können mittels einer qualitativen Auswertung offenkundig werden. Somit werden Ansichten und Urteile für die Analyse besser nachvollziehbar und können in den jeweiligen, der Äußerung zu Grunde liegenden Kontext gesetzt werden (vgl. Kuckartz et al. 2008: 66ff.).

Nach dem Forschungsverständnis von Kuckartz et al. ist es besonders die Forschungsfrage, der eine zentrale Bedeutung im Prozess zukommt, womit die Autoren einem pragmatischen Methodenverständnis folgen (vgl. ebd. 2008: 13). Dieses Verständnis bietet die nötige Kreativität für den Forschungsprozess, und ist unabdingbar für eine weitere Entwicklung bezüglich der sozialwissenschaftlichen Potentiale, insbesondere im Rahmen der Diskursforschung. Es bedarf damit letztlich einem gewissen Maß an Flexibilität hinsichtlich der Forschungsfrage (vgl. Mayring 2002: 65; Keller 2011: 119). Das heißt allerdings nicht, dass die qualitative Analyse unter dem Deckmantel der *Kreativität* und *Flexibilität* beliebig sein darf. Vielmehr muss hier besonders darauf geachtet werden, dass eine Offenlegung und Nachvollziehbarkeit der einzelnen Interpretationsschritte aufrechterhalten wird. Diesem Anspruch an die empirische Wissenschaft muss gefolgt werden, wenn die Analyse mehr sein will als *bloße Reportage*. Das

⁴² Siehe Kapitel 5.2.5.

heißt auch für qualitatives Forschen, dass es keineswegs verschwommen sein darf und hinsichtlich Offenlegung und Systematisierung der Vorgehensweise den quantitativen Analysen ein Äquivalent darstellt (vgl. Mayring 2002: 65; Keller 2011: 76f.). „Die zur Beantwortung der Forschungsfrage eingesetzten Methoden haben sich als gegenstandsangemessen zu legitimieren“ (Kuckartz et al. 2008: 13).⁴³

Um dem Gebot zu folgen die eingesetzte Methode *gegenstandsangemessen* (vgl. Kuckartz et al. 2008: 13) bzw. *gegenstandsadäquat* (vgl. Mayring 2002: 65) anzuwenden, soll die Vorgehensweise zur Überprüfung dieser Dimension hinsichtlich der angestrebten Analyseschritte bereits vor der Anwendung an den einzelnen Fallbeispielen skizziert werden. Die Auswertung der Protokolle und Zusammenfassungen der einzelnen Fälle erfolgt dabei mittels einer qualitativen Inhaltsanalyse. Die Form dieser Analyse strebt an, Texte systematisch mittels eines Kategoriensystems schrittweise zu verdichten. Ziel ist es, das Material so zu reduzieren, dass trotzdem noch ein Abbild des Originalmaterials erhalten bleibt (vgl. Mayring 2002: 115). Da die Ergebnisse eines Gruppendelphis besonders auf inhaltlicher Ebene von Interesse sind und die Plenarphasen ergebnisorientiert gestaltet werden, zielt die Analyse auf den Gruppenoutput, der in der Dokumentation festgehalten wurde und „(...) beinhaltet die Identifikation zentraler Themen des Gesprächs sowie eine Beschreibung und Erklärung der verschiedenen Meinungen“ (Schulz 2012: 17).

Im Falle dieser Dimension wird damit besonderes Augenmerk auf die von den Teilnehmenden geäußerten normativen und wertgeleiteten Argumente gelegt, die diese im Verlaufe der Plenardiskussionen äußern. Neben den Protokollen stehen meist auch die Zusammenfassungen der Ergebnisse, die an die Teilnehmenden nach dem Gruppendelphi gesandt wurden, für eine Analyse zur Verfügung. Diese werden ergänzend analysiert. Da bei dieser Dimension überprüft werden soll, ob die Teilnehmenden auch Argumente vorbringen, die dem Geltungskriterium der normativen Angemessenheit entsprechen, muss das jeweils vorliegende Material auf unterschiedliche Arten an Äußerungen mittels folgender Fragestellungen überprüft werden:

- Bezieht sich die Äußerung direkt auf den Geltungsanspruch nach Gemeinwohl oder normativer Richtigkeit?
- Beziehen sich anknüpfende Erwiderungen auf diesen Geltungsanspruch?
- Wird einer Person von einer anderen Person eine wertgeleitete Argumentation zugesprochen bzw. unterstellt?
- Werden diese Argumente als legitim erachtet?

Dabei werden vier Kategorien gewählt, welche die von den Teilnehmenden getätigte *Sollensaussagen* für die weitere Analyse strukturieren und mögliche Unterkategorien der jeweiligen Fälle ermöglichen. So sind unter den Geltungskriterien der Kohärenz und Verallgemeinerungsfähigkeit sowie dem Universalisierungsgebot und der Kompatibilität mit gesetzten Äußerungen als Kategorien möglich, die in **(1) Werturteile**, **(2) Präferenzen**, **(3) (individuelle) Bewertungen von Optionen** und **(4) moralische Urteile** unterteilt werden können.⁴⁴ Diese Kategorien bilden eine Verknüpfung von normativen Aussagen mit ihrer jeweiligen theoretischen Bezugs-

⁴³ Hervorhebung im Original.

⁴⁴ vgl. Kapitel 5.2.3.

größen von Rationalität und dem daraus abgeleiteten System von Geltungsansprüchen, welche eine Einordnung sowie eine Bewertung der einzelnen von Teilnehmenden getätigten Aussagen ermöglichen.

Bevor die bisher in den drei Dimensionen beschriebene Analyse Anwendung findet, soll eine weitere Dimension und deren Testdesign beschrieben werden, das einer Überprüfung der grundlegenden Funktionslogik eines Delphiansatzes und weiteren Randbedingungen Rechnung tragen soll.

6.2.4 Dimension: grundlegende Funktionslogik eines Delphiansatzes

Neben den Implikationen, die sich durch die diskurstheoretische Erweiterung des Gruppendelphi ergeben, sollte auch die Funktionsweise von Wiederholung und Rückkopplung, die Gruppendelphi und klassischem Delphiverfahren genuin ist, als grundlegende Funktionsweise an den empirischen Fallbeispielen überprüft werden. *„The inherent tendency of the Delphi process towards agreement is a strength in that it can guide group opinion toward a final decision (...) allowing for disagreement to be noted and included in the results. It does mean, however, the final judgments must always take into account the distribution of responses and that a group’s agreement may not be as significant as it first appears“* (Goodman 1987: 730). Da das Gruppendelphi eine Modifikation des klassischen Delphiverfahrens darstellt, wird in dieser Dimension die Einlösung der grundlegenden Funktionslogik des klassischen Delphiverfahrens ebenso für die jeweiligen empirischen Fallbeispiele überprüft. Dafür werden die beschriebenen Funktionsweisen auch für das Gruppendelphi angewandt, insofern diese nicht bereits durch die drei vorhergehenden Hauptdimensionen abgedeckt sind.⁴⁵ Auch sollten die Urteile der Teilnehmenden ebenso nach dem Gruppendelphi stabil bleiben, um einen Scheinkonsens ausschließen zu können. Dieser würde vorliegen, wenn die Teilnehmenden den Workshoptag nicht dazu genutzt hätten, um ihre Argumente einzubringen, obwohl sie die von der Gruppe getroffene Beurteilung nicht mittragen. Dies könnte ein Indikator für eine fehlende Motivation einzelner Teilnehmer*innen oder nicht ausreichender Zeit für einen erschöpfenden Diskurs sein.⁴⁶ Aufgrund der fehlenden Anonymität kann bei einem Gruppendelphi auch eine mögliche Meinungsführerschaft nicht gänzlich ausgeschlossen werden, die ebenso zu einem unerwünschten Scheinkonsens führen könnte.⁴⁷ Die Protokolle der empirischen Fallbeispiele bieten die Möglichkeit einen Scheinkonsens für die jeweiligen Gruppendelphis auszuschließen. So verfolgt ein nach dem Gruppendelphi an die Teilnehmenden versandtes Protokoll die Aufgabe der Prüfung, ob die Ergebnisse des Gruppendelphis mitsamt den jeweiligen Begründungen und Argumenten adäquat erhoben wurden. Auf diesem Wege kann die Validität des Gesamtprozesses zusätzlich gesteigert sowie ein möglicher Scheinkonsens aufgedeckt werden. Dem wäre der Fall, wenn die Teilnehmenden, den im Protokoll berichteten Argumenten und Ergebnissen des Gruppendelphis im Nachgang widersprechen würden. Um auch hier die grundlegende Funktionslogik eines Gruppendelphis zu überprüfen und damit solche potentiellen Scheinkonsense konkret auszuschließen, wird bei den jeweiligen Fällen geprüft werden,

⁴⁵ Zu den Funktionsweisen vgl. Kapitel 3.

⁴⁶ vgl. Kapitel 4.1.2.

⁴⁷ vgl. Kapitel 4.1.5.

ob es Widersprüche und Änderungsbedarfe der Teilnehmenden am Protokoll gab und ob diese inhaltlich auf einen Scheinkonsens hinweisen.

Letztlich muss sich ein Delphi nicht nur daran bemessen lassen, ob es zu einer möglichst hohen Zahl an konsensualen Einschätzungen geführt hat, sondern dass die Ergebnisse im Sinne einer rationalen Urteilsbildung eines sachgerechten und wertadäquaten Resultates hinsichtlich der Fragestellung des jeweiligen Falles geführt haben.⁴⁸ Aus diesem Grund soll in dieser Dimension nochmals das Forschungsinteresse der jeweiligen Fallbeispiele mit der Verwertbarkeit der Ergebnisse diskutiert werden.

Aus methodischer Perspektive ist es zudem ratsam, auch einzelne Elemente des Gruppendelphis im Rahmen der Fälle zu reflektieren. Dabei werden die methodischen Elemente der einzelnen Fälle (wie beispielsweise die Ausgestaltung des Fragebogens oder der Ablauf des Tages im Kontrast zur geplanten Agenda)⁴⁹ reflektiert. Falls sich daraus Implikationen auf die Verwertbarkeit der Forschungsergebnisse ergeben, so kann diese Betrachtung wertvolle Hinweise für die generelle Ausgestaltung zukünftiger Gruppendelphis geben.

Tabelle 5 bietet abschließend einen Überblick über die einzelnen im Forschungsdesign verwendeten Dimensionen. Diese in diesem Kapitel beschriebenen Dimensionen werden folgend auf die einzelnen Fälle angewandt.

Tabelle 5: Dimensionen des Forschungsdesigns

1. Evidenz	2. Verständlichkeit	3. Normative Angemessenheit
Austausch über Gültigkeit der einzelnen Argumente. Verständigung über, die, den Beurteilungen zugrundeliegenden Relevanzsysteme.	Semantischen Klärung von Begriffen und Definitionen	Möglichkeit der Offenlegung von normativen Einstellungen und Wertvorstellungen
4. Grundlegende Funktionslogik des Delphiansatzes		
Betrachtung der Ergebnisse (Forschungsinteresse) Generelle Funktionsweise Reflektion der methodischen Ausgestaltung		

Quelle: Eigene Darstellung

⁴⁸ Zu den theoretischen Hintergründen solcher Bemessungsgrundlagen siehe Kapitel 5.

⁴⁹ Die Merkmale eines Gruppendelphis wurden bereits in Kapitel 4 dargelegt.

7 Empirische Fälle: Darstellung und Diskussion der Ergebnisse

In diesem Kapitel wird das in Tabelle 5 dargestellte Testdesign auf verschiedene Fallbeispiele aus der Praxis angewandt. Ein ähnliches Vorgehen haben auch Schulz und Renn gewählt, die in einer Veröffentlichung Praxisbeispiele im Hinblick auf Schwächen und Stärken der Methode diskutiert haben. Der Fokus lag hier allerdings auf der praktischen Durchführung. Damit sollte die Diskussion der Fallbeispiele „*vor allem praktische Hilfestellung bei der Durchführung eines Gruppendelphis, insbesondere bei der Entwicklung eines adäquaten Fragebogens, liefern*“ (Schulz/Renn 2009: 8). Deuschle und Sonnberger verfolgten bei der Darstellung ihres Fallbeispiels bewusst eine Kontrastierung von *Ideal-Delphi* und der *Delphi-Praxis*, die sich beispielsweise aus Gegebenheiten der Rekrutierung und des Zeitmanagements ergaben (vgl. ebd. 2009: 76). In der Literatur zum klassischen Delphiverfahren sowie auch zum Gruppendelphi wird davon ausgegangen, dass es bisweilen keine allgemein anerkannte Definition des Delphiverfahrens gibt (vgl. Vorgrimler/Wübben 2003: 764; Häder 2009: 13; Schulz/Renn 2009: 11).⁵⁰ Ob es ein – wie von Deuschle und Sonnberger angesprochenes *Ideal-Delphi* also gibt, lässt sich anhand der Abwesenheit einer allgemein anerkannten Definition zumindest in Frage stellen. Für das klassische Delphiverfahren bilden die vier Typen von Häder allerdings eine Orientierung bei der Konzeption und Analyse innerhalb des klassischen Ansatzes (vgl. ebd. 2009: 36). Und auch wenn es viele Varianten des Delphiansatzes gibt, so kann doch von einem gemeinsamen Ablaufmodell ausgegangen werden. Die verschiedenen Ansätze bieten gemeinsame Kennzeichen, die der Vielzahl der Variationen genuin sind, so etwa beispielsweise, der Einsatz eines formalisierten Fragebogens, die Ermittlung aggregierter Gruppenantworten, das Feedback mittels statistischer Ergebnisse oder die Wiederholung der Befragung bis zu einem festgelegten Abbruchkriterium (vgl. Schulz/Renn 2009: 12f.).⁵¹

Bei der Vielzahl der Variationen ist es damit durchaus möglich, die praktische Anwendung eines Delphis auch als solche Methode - oder eben eine Anwendung einer anderen Methode unter dem Titel „Delphi“ als falsche Bezeichnung - zu identifizieren. Neben den Kennzeichen bieten die im Kapitel 4.1.2 beschriebenen Merkmale eines Gruppendelphis ein deskriptives Analyseschema, um trotz der Variationen innerhalb der Methode des Gruppendelphis, die konkreten empirischen Anwendungen als ein solches Gruppendelphi zu kategorisieren. Diese Merkmale helfen überdies eine Vergleichbarkeit der einzelnen Fälle sicherzustellen und sollen für jedes der folgenden Fallbeispiele angewandt werden:

- *Vorbefragung*: Wurde im konkreten Fall eine Vorbefragung durchgeführt?
- *Zusammensetzung*: Wie viele und welche Personen haben am Gruppendelphi teilgenommen?
- *Gruppenzusammensetzung*: Wurden die Kleingruppen zufällig oder zugeteilt zusammengesetzt?
- *Dokumentation der Plenardiskussion*: Wie wurde die Plenardiskussion dokumentiert?

⁵⁰ Siehe hierzu auch Kapitel 3.3.

⁵¹ Siehe hierzu auch Kapitel 3.

- *Fragebogen*: Welche Fragetypen und Skalentypen kamen im Fragebogen des jeweiligen Fallbeispiels zu Anwendung?
- *Dauer*: Wie lange hat der Workshop gedauert?
- *Anzahl der Runden*: Wie viele Runden haben im jeweiligen Gruppendelphi stattgefunden?
- *Ausgestaltung des Feedbacks*: Wie wurden die finalen Ergebnisse des Workshoptages an die Teilnehmenden rückgespiegelt?

Die konkrete Ausgestaltung der einzelnen Fälle ist immer von deren zentralem Forschungsinteresse bestimmt, so dass die Forschungsfrage des jeweiligen Gruppendelphis auch Einfluss auf die Merkmale dieses Falles Einfluss ausübt. So soll in diesem Kapitel neben einer allgemeinen Beschreibung der einzelnen Fälle auch immer die zentrale Fragestellung dargelegt werden, bevor die einzelnen Merkmale und die im vorhergehenden Kapitel beschriebenen Dimensionen Anwendung finden. Letztlich werden in diesem Kapitel die Ergebnisse jedes einzelnen Falles im Hinblick auf die Fragestellung dieser Arbeit diskutiert.

Bevor die einzelnen Fälle anhand dieses Schemas analysiert werden, sollen die beiden Institute, die jeweils für die Konzeption, Durchführung und Ergebnissicherung der einzelnen Fälle zuständig waren, in Tabelle 6 vorgestellt werden.

Tabelle 6: Fallbeispiele - Durchführende Institute

DIALOGIK gemeinnützige Gesellschaft für Kommunikations- und Kooperationsforschung mbH
<p>Wissenschaftlicher Direktor und Geschäftsführer: Prof. Dr. Dr. h.c. Ortwin Renn Geschäftsführerin: Agnes Lampke Stellvertretende Wissenschaftliche Direktorin und Prokuristin: Dr. Marion Dreyer</p> <p>Die Aufgabe von DIALOGIK ist es, „(...) Kommunikations- und Kooperationsformen im Spannungsfeld von Politik, Wirtschaft und Zivilgesellschaft systematisch und anwendungsorientiert zu erforschen. Hierfür setzt DIALOGIK gezielt diskursive Untersuchungsmethoden ein. In Gesellschaften, die durch eine Pluralität von Wissen und Werten gekennzeichnet sind, gewinnen Kommunikation und Kooperation zwischen und innerhalb von gesellschaftlichen Akteursgruppen immer mehr an Bedeutung. Dies trifft im Besonderen auf die Politikfelder Technik, Umwelt, Gesundheit und Verbraucherschutz zu. Hier führen kontroverse Ansichten über den angemessenen Umgang mit Chancen und Risiken häufig zu fruchtlosen Verhandlungen und Blockaden zwischen den relevanten gesellschaftlichen Kräften. Vor allem für diese risikosensiblen Politikfelder untersuchen wir, wie innovative Formen der Kommunikation und neuartige Partizipations- und Kooperationsverfahren Entscheidungsprozesse und die Umsetzung von Entscheidungen in die Praxis von Politik, Wirtschaft und Gesellschaft verbessern können. Unser Ziel dabei ist es, neue Wege zu erforschen, wie der Umgang mit wertvollen gesellschaftlichen Ressourcen – natürlichen, ökonomischen, kulturellen, gesellschaftlichen – verbessert werden kann.</p> <p>Wichtige Felder unserer Forschungsarbeit sind:</p> <ul style="list-style-type: none"> • Neue Ansätze der Risikokommunikation als Instrument eines effektiven und fairen Risiko- und Chancenmanagements • Verbesserte Kommunikationsinstrumente für die öffentliche Gesundheitsfürsorge • Innovative Formen der Partizipation und Konfliktlösung • Praktische Mediation in Technik- und Umweltkonflikten • Verbesserte Kommunikationsprogramme für Brückenbildung und wechselseitigen Austausch zwischen Wissenschaft und Öffentlichkeit • Integration von Technikfolgenabschätzung und Vorausschau (Foresight) -Aktivitäten in Governance-Prozesse

*Wir analysieren die Formen, Instrumente, Wirkungen und Potentiale von Kommunikation und Kooperation auf lokaler, regionaler, nationaler und transnationaler Ebene. DIALOGIK stellt sich den hier skizzierten wissenschaftlichen Herausforderungen mit dem Ziel, in einem stetigen Prozess des Suchens und Lernens einen **konzeptionellen, theoretischen** wie auch **handlungsrelevanten** Beitrag für eine **neue Kommunikationskultur** in dem gemeinsamen Streben nach einer **nachhaltigen** und zukunftsgerechten Entwicklung zu leisten“ (Homepage Dialogik)⁵².*

ZIRIUS - Zentrum für Interdisziplinäre Risiko- und Innovationsforschung der Universität Stuttgart (ehemals ZIRN)

Wissenschaftliche Direktorin: Prof. Dr. Cordula Kropp
Wissenschaftlicher Direktor: Prof. Dr. André Bächtiger
Stellvertretender wissenschaftlicher Direktor: Prof. Dr. Burkhard Pedell
Wissenschaftliche Geschäftsführerin: Sandra Wassermann
Verwaltungsleitung: Agnes Lampke

„Das Zentrum für interdisziplinäre Risiko- und Innovationsforschung (ZIRIUS) wurde 2003 als Interdisziplinärer Forschungsschwerpunkt Risiko und Nachhaltige Technikentwicklung (ZIRN) am Internationalen Zentrum für Kultur- und Technikforschung (IZKT) der Universität Stuttgart gegründet. Gründer und langjähriger Leiter war bis Anfang 2016 Ortwin Renn. ZIRN ging in fachlichen und sachlichen Ausstattungsmerkmalen auf die 2003 geschlossene Akademie für Technikfolgenabschätzung in Baden-Württemberg zurück. Seit August 2013 besteht ZIRIUS als Stuttgart Research Center an der Universität Stuttgart.

Im Juni 2017 übernahmen Prof. André Bächtiger und Prof. Cordula Kropp als gewähltes Direktorenteam die Leitung von ZIRIUS.

Im Februar 2018 wurde apl. Prof. Dr. Kathrin Braun als Forschungskordinatorin eingestellt.

ZIRIUS untersucht die Bedingungen und Chancen sozio-technischen Wandels und demokratischer Innovation auf dem Weg in eine gerechte und nachhaltige Gesellschaft.

Sozio-technischen Wandel verstehen wir als komplexes Ineinandergreifen von wissenschaftlich-technischen und gesellschaftlichen Wandlungsprozessen. Technologieentwicklung ist dabei selber als sozialer Prozess zu verstehen, der in soziale, ökonomische und kulturelle Kontexte eingebettet ist.

Unsere Forschung befasst sich mit den Fragen und Herausforderungen des sozio-technischen Wandels, seinen Gestaltungschancen und Risiken und den Chancen effektiver Bürgerbeteiligung.

Wir untersuchen das Spannungsfeld zwischen wissenschaftlich-technischen Innovationspotenzialen und der Notwendigkeit ihrer verantwortlichen Gestaltung. Schwerpunkte bilden Fragen der Technikfolgenabschätzung, sozio-technischer Transformation, der partizipativen Risikobeurteilung und der Entwicklung innovativer Beteiligungsformate.“ (Homepage ZIRIUS).

Als Autor dieser Arbeit war ich an allen der als Fälle analysierten Gruppendelphis in der Organisation, Planung und/oder Moderation des Workshops involviert. Die Reihenfolge der dargestellten Fälle erfolgt strukturiert nach Zeitpunkt der Durchführung, beginnend bei der jüngsten Durchführung hin zu älteren Fällen. Den Anfang bildet dabei die Analyse des Gruppendelphis im Rahmen des Forschungsprojektes Sauber+.

⁵² Hervorhebungen im Original.

7.1 SAUBER+

Projekt: SAUBER+ - Innovative Konzepte und Technologien für die separate Behandlung von Abwasser aus Einrichtungen des Gesundheitswesens

Institut: DIALOGIK gemeinnützige Gesellschaft für Kommunikations- und Kooperationsforschung mbH

Das hier betrachtete Fallbeispiel eines Gruppendelphis fand im vom *Bundesministerium für Bildung und Forschung (BMBF)* unter der Fördermaßnahme *Risikomanagement von neuen Schadstoffen und Krankheitserregern im Wasserkreislauf (RiSKWa)* geförderten Vorhabens SAUBER+ statt, das vom *Institut für Siedlungswasserwirtschaft der RWTH Aachen* unter Leitung von Prof. Dr.-Ing Pinnekamp koordiniert wurde. Im Rahmen dieses Projektes wurden die Emissionen von Arzneimittelrückständen aus Einrichtungen des Gesundheitswesens in die aquatische Umwelt qualitativ wie auch quantitativ bewertet und Vorschläge zu deren Minderung erarbeitet. Für diesen Zweck wurden Daten zur Beurteilung des Risikopotentials, das aus antibiotikaresistenten Krankheitserregern und Emissionen von Arzneimittelrückständen aus Gesundheitseinrichtungen resultieren kann, erhoben und transdisziplinär bewertet. Neben Technologien und Strategien zur Minderung dieser Emissionen wurden in diesem Forschungsvorhaben zusätzliche Bildungsmaßnahmen zur Sensibilisierung ausgearbeitet wie auch Stakeholderdialoge durchgeführt (vgl. Pinnekamp et al. 2015: 357).

Das Gruppendelphi war Teil einer Reihe von Stakeholderdialogen, die sich insgesamt in sechs Workshops aufteilte, von denen das betrachtete Gruppendelphi Teil des vierten Workshops dieser Reihe war. Für den Dialogprozess wurden Personen rekrutiert, die mit dem Konsum bzw. der Produktion von Arzneimitteln oder dem Weg der Arzneimittel durch die aquatische Umwelt zu tun hatten. Die Teilnehmenden konnten damit aus verschiedenen Bereichen rekrutiert werden, so dass sich die Gruppe der Stakeholder letztlich aus Vertreter*innen der Abwasserwirtschaft, der Apothekerschaft, der Ärzteschaft, Einrichtungen des Gesundheitswesens, dem Gewässerschutz, Ministerien und Fachbehörden, der pharmazeutischen Industrie, der Umweltforschung, Umweltschutzverbänden, dem Verbraucherschutz und aus der Wasserversorgung zusammensetzte. Neben dem Gruppendelphi wurde in diesem Stakeholderdialogprozess ebenso eine partizipative Szenarioentwicklung umgesetzt (vgl. Kerber et al. 2015a: 104ff.). Diese ausgearbeiteten Szenarien sollten zukünftige mögliche Entwicklungspfade erfassen und abbilden, weshalb Maßnahmen und Maßnahmenbündel über den Zeithorizont beschrieben und in Dynamik und Interaktion weiterentwickelt wurden. Die in diesem Szenarioprozess dargestellten Maßnahmen wurden anschließend mithilfe des Gruppendelphis auf ihre Umsetzbarkeit und mögliche Wirkungen von den Stakeholdern bewertet (vgl. Kerber et al. 2015b: 314).

7.1.1 Fragestellung des Gruppendelphis

Ziel des Einsatzes der Methode des Gruppendelphis im Projekt SAUBER+ war es, „(...) eine Bewertung verschiedener Maßnahmen zur Verringerung der Belastung der aquatischen Umwelt mit Rückständen von Humanarzneimitteln (...) durch die im Forschungsvorhaben mit dem Themenfeld verbundenen Stakeholder vornehmen zu lassen. Über die Maßnahmenbewertung

*hinaus sollte der Handlungsbedarf hinsichtlich Humanarzneimitteln im Vergleich zum Handlungsbedarf bei anderen Kontaminanten (die zu einer Belastung von Oberflächengewässern und von Grundwasser führen können) in Relation gesetzt werden“ (Kuhn et al. 2015: 317). Konkret bestand die Aufgabe der Stakeholder darin, den Handlungsbedarf in Bezug auf Humanarzneimittel, Diagnostika, Krankheitserregern sowie multiresistenten Keimen im Wasserkreislauf in Relation zum Handlungsbedarf bei anderen Kontaminanten, die zu einer Belastung von Gewässern führen, zu bewerten. Zudem sollte durch die Teilnehmenden ebenfalls eine Bewertung des relativen Handlungsbedarfes für verschiedene Eintragsquellen von Arzneimittelrückständen in die aquatische Umwelt erfolgen. Letztlich wurden überdies konkrete Maßnahmen in den Handlungsfeldern *Umgang mit Arzneimitteln*, *Arzneimittelentwicklung* und *Emissionsmanagement* durch die Stakeholder bewertet (vgl. Kuhn et al. 2015: 319). Die Methode des Gruppendelphis schien hier für einen Einsatz besonders geeignet, da Arzneimittelrückstände im Wasserkreislauf als relativ junges Risikothema im Gewässerschutz eingestuft werden können. Auf der diskursiven Ausrichtung der Methode lag hier besonders ein Fokus, da vorzüglich die argumentativen Grundlagen der verschiedenen Einschätzungen der teilnehmenden Stakeholder im Forschungsinteresse der Projektbeteiligten standen (vgl. Dreyer/Kuhn 2014: 54; Kuhn et al. 2015: 317).*

7.1.2 Merkmale des Gruppendelphis

Im Vorfeld des Gruppendelphis wurde eine Vorbefragung durchgeführt, die an 19 Personen gesandt wurde. Diese 19 Teilnehmenden waren Stakeholder, die an verschiedenen Formaten der Workshopreihe teilgenommen hatten. Von diesen 19 Personen beantworteten 14 Personen den Fragebogen, wobei zwei dieser Personen einen gemeinsamen Fragebogen in gegenseitiger Absprache ausfüllten.⁵³ Mit der Vorbefragung waren von Seiten des Forschungsteams drei Ziele verbunden:

1. Die Stakeholder sollten sich auf diesem Wege bereits mit der Struktur und den Inhalten des Fragebogens vertraut machen.
2. Das Team wollte Impulse und Anmerkungen für eine gegebenenfalls notwendige Weiterentwicklung des Fragebogens einholen.
3. Es sollten potentiell konsensuale Bewertungen von Maßnahmen ermittelt werden, die für den Workshopfragebogen somit entfallen könnten (vgl. Kuhn et al. 2015: 319).

Am eigentlichen Workshop nahmen zehn Personen teil, deren institutionelle Hintergründe in der Wasserwirtschaft, im Umweltschutz, der Pharmaindustrie, dem Apothekerverband, von Behörden, der Umweltmedizin, dem Verbraucherschutz und der Wasserforschung lagen (vgl. Kuhn et al. 2015: 320). Um keine stark homogenen Kleingruppen hinsichtlich der einzelnen institutionellen Bereiche in der ersten Runde des Gruppendelphis zu bilden, wurden die Teilnehmenden für diese Runde vom Team zugeteilt. Zur zweiten Runde erfolgte dann eine permutierte Gruppenzusammensetzung.

Der Fragebogen der Vorbefragung unterteilte sich in sieben Abschnitte. Unter vieren dieser Abschnitte wurden insgesamt 34 Ratingfragen mittels 10er-skaliertem Items abgefragt, wovon

⁵³ Dies erfolgte nicht in Absprache mit dem Forschungsteam, sondern aus eigener Motivation der Teilnehmenden, da beide Personen als Stakeholder aus dem Bereich der Pharmaindustrie am Prozess teilnahmen.

eine Itembatterie durch eine offene Abfrage der potentiellen Nennung „*Sonstige*“ ergänzt wurde. Neben diesem Itemtyp wurden zusätzlich eine offene Frage sowie zwei Rankingfragen mittels Verteilung eines fiktiven Budgets unter den restlichen Abschnitten des Fragebogens abgefragt.⁵⁴ Der Workshopfragebogen unterteilte sich ebenfalls in sieben Abschnitte. Insgesamt wurden in diesem Fragebogen 46 10er-skalierte Items abgefragt, wovon eine Itembatterie abermals durch eine offene Abfrage der potentiellen Nennung „*Sonstige*“ ergänzt wurde. Zwölf dieser 10er-skalierten Items wurden für die jeweiligen Aussagen in eine Bewertung der *Wirksamkeit der Maßnahme* und der *Umsetzbarkeit der Maßnahme* aufgeteilt. Neben diesen Fragetypen wurden zwei Rankingfragen mittels Verteilung eines fiktiven Budgets abgefragt.⁵⁵

Das Gruppendelphi im Projekt Sauber+ fand am 27. November 2013 im Ökohaus in Frankfurt am Main statt und war Teil des zweitägigen Stakeholder Workshops zu *Handlungsstrategien - Welche Schwerpunkte bei Einrichtungen und Maßnahmen?* Das Gruppendelphi fand am zweiten Tag des Gesamtworkshops statt und gliederte sich im Detail wie folgt auf:

9.00 Uhr	Eintreffen / Begrüßungskaffee
9.15 Uhr	Einführung in das Gruppen-Delphi und in die Ziele des Tages
9.30 Uhr	Bearbeitung des ersten Fragebogens in Kleingruppen
11.00 Uhr	Kaffeepause und Auswertung
11.30 Uhr	Plenarsitzung : Diskussion der Ergebnisse
13.00 Uhr	Mittagspause
14.00 Uhr	Bearbeitung des zweiten Fragebogens in Kleingruppen
14.45 Uhr	Kaffeepause und Auswertung
15.00 Uhr	Plenarsitzung : Diskussion der Ergebnisse
15.45 Uhr	Fazit und Feedback zur Methode
16.15 Uhr	Maßnahmen-Blitzlicht
16.30 Uhr	Feedback-Runde und Abschied
16.45 Uhr	Ende

Die beiden durchgeführten Plenarsitzungen wurden von einem Teammitglied protokolliert, so dass im Nachgang zum Workshop mit diesen schriftlichen Aufzeichnungen und auch den statistischen Ergebnissen ein 18 Seiten umfassendes Protokoll an die Teilnehmenden mit der Bitte um Feedback versendet werden konnte.

7.1.3 Analyse des Gruppendelphis in SAUBER+

Nachdem die Fragestellung und die Merkmale des Gruppendelphis in SAUBER+ beschrieben wurden, werden folgend die vier Dimensionen des Testdesigns für diesen Fall angewandt: Können für diesen Fall die Ansprüche einer inhaltlichen Verbesserung der Urteile, einer konsensualen Annäherung der Urteile sowie letztlich einer höheren Akzeptanz der Integration neuer Informationen erzielt werden? Dem Aufbau des Testdesigns folgend beginnt die Analyse mit der Dimension der Evidenz.

⁵⁴ Der Fragebogen der Vorbefragung befindet sich im Anhang 11.1.

⁵⁵ Der im Gruppendelphi eingesetzte Fragebogen befindet sich im Anhang 11.2.

7.1.3.1 Dimension: Evidenz

Wie bereits in Kapitel 6.2.1 beschrieben, wird die Dimension der Evidenz als zentrale Dimension im Testdesign verortet. Der Geltungsanspruch, der an die im Gruppendelphi erzielten Ergebnisse unter diesem Bemessungskriterium gestellt wird, folgt dem Geltungsanspruch der Gültigkeit.

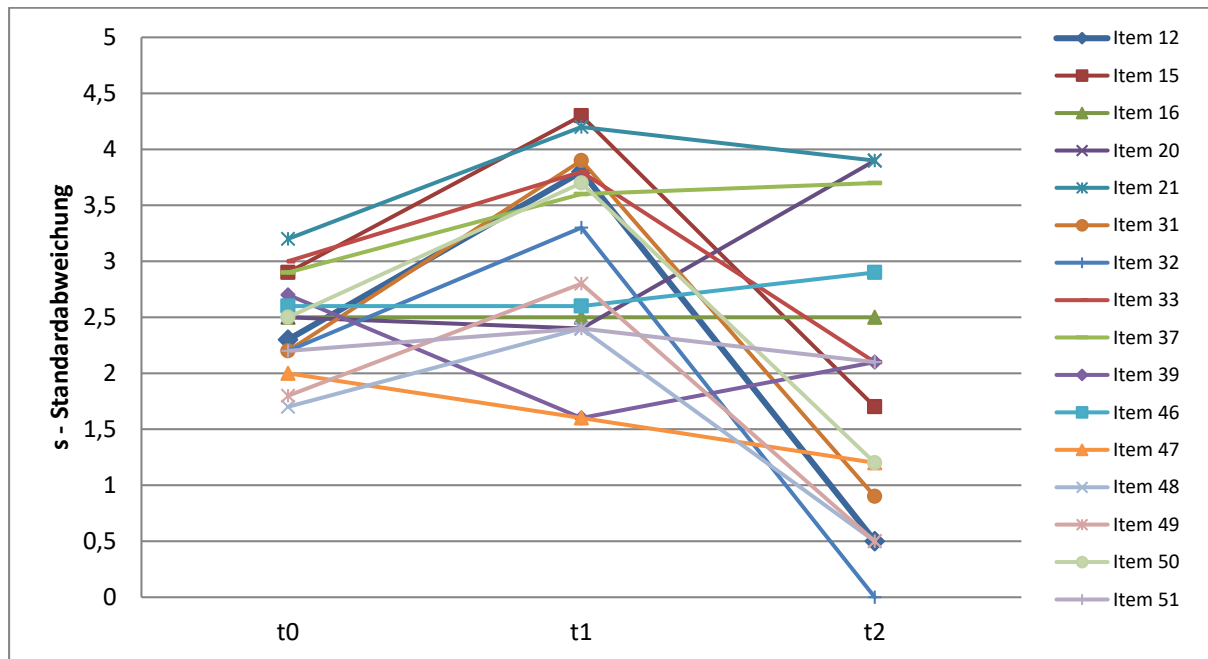
Den Beginn der Analyse bilden die Items, die mittels 10er-skalierten Ratingskalen abgefragt wurden.⁵⁶ Bei den in diesem Projekt eingesetzten Fragebögen bilden diese Items den am häufigsten eingesetzten Typus - mit 34 Items im Fragebogen der Vorbefragung und 46 Items im am Workshop eingesetzten Fragebogen. Die Items wurden mittels ausformulierter Aussagen gebildet und abgefragt. Die Aufgabe der Teilnehmenden bestand darin anzugeben, inwieweit sie den Aussagen über Maßnahmen aus den verschiedenen Abschnitten der drei Handlungsfelder *Arzneimittelentwicklung*, *Umgang mit Arzneimitteln* und *Emissionsmanagement* zustimmen. Die Stakeholder sollten bei diesen Fragen einen Wert von 1 angeben, wenn sie die jeweilige Aussage *ganz und gar ablehnen* und einen Wert von 10 wählen, wenn sie der jeweiligen Aussage *voll und ganz zustimmen*. Weiterhin wurde die jeweilige *Dringlichkeit des Handlungsbedarfs nach Eintragsquellen* mittels 10er-skalierten Items abgefragt. Bei dieser Frage wurden die einzelnen Eintragsquellen jeweils als einzelnes Item bewertet. Die Stakeholder wurden hier gebeten, einen Wert von 1 anzugeben, wenn sie der Ansicht waren, dass bei der jeweiligen Eintragsquelle *kein Handlungsbedarf* besteht. Wenn sie der Ansicht waren, dass bei der jeweiligen Eintragsquelle *hoher Handlungsbedarf* besteht, wurden sie gebeten einen Wert von 10 anzugeben.

Wenn die Qualitätsmerkmale eines handlungsleitenden Diskurses für diesen Fall eingelöst werden, so sollten sich über den Verlauf eines Gruppendelphis die Werte in Richtung eines Konsenses bewegen. Für die im ersten Analyseschritt betrachteten Werte der Standardabweichung der jeweiligen Items bedeutet dies, dass der jeweilige Wert der Standardabweichung über die Delphirunden hinweg sinken und sich die Einschätzungen der Teilnehmenden idealiter an einem Wert auf der Skala bündeln. Im Fall des Gruppendelphis in SAUBER+ wurden insgesamt 16 Items über die Delphirunden hinweg abgefragt, die somit für einen Vergleich der Entwicklung der Standardabweichung über die Runden als Daten zur Verfügung stehen. Der Verlauf des Wertes der Standardabweichung über die Vorbefragung (t0) sowie über die beiden Runden (t1; t2) des Workshoptages finden sich in Abbildung 10. In dieser Abbildung wird deutlich, dass die Standardabweichung einzelner Items verschiedene Entwicklungen nehmen kann. So steigt sie beispielsweise bei *Item 21* von $s=3,2$ zur ersten Delphirunde auf einen Wert von $s=4,2$ an, um dann zur zweiten Delphirunde leicht auf den Wert von $s=3,9$ zu sinken. Bei *Item 16* bleibt der Wert der Standardabweichung konstant bei $s=2,5$ über alle drei Abfragen hinweg. Solche Entwicklungen werden in der folgenden Analyse näher betrachtet werden, da hier zum einen die Häufigkeiten sowie auch die in der Plenarsitzung von den Stakeholdern geäußerten Argumente mögliche Erklärungsmodelle bereitstellen können.⁵⁷

⁵⁶ Die Antworten der einzelnen Kleingruppen sowie die statistischen Ergebnisse finden sich in Anhang 11.3.

⁵⁷ Die Entwicklung einzelner Items wird nachfolgend differenziert beschrieben.

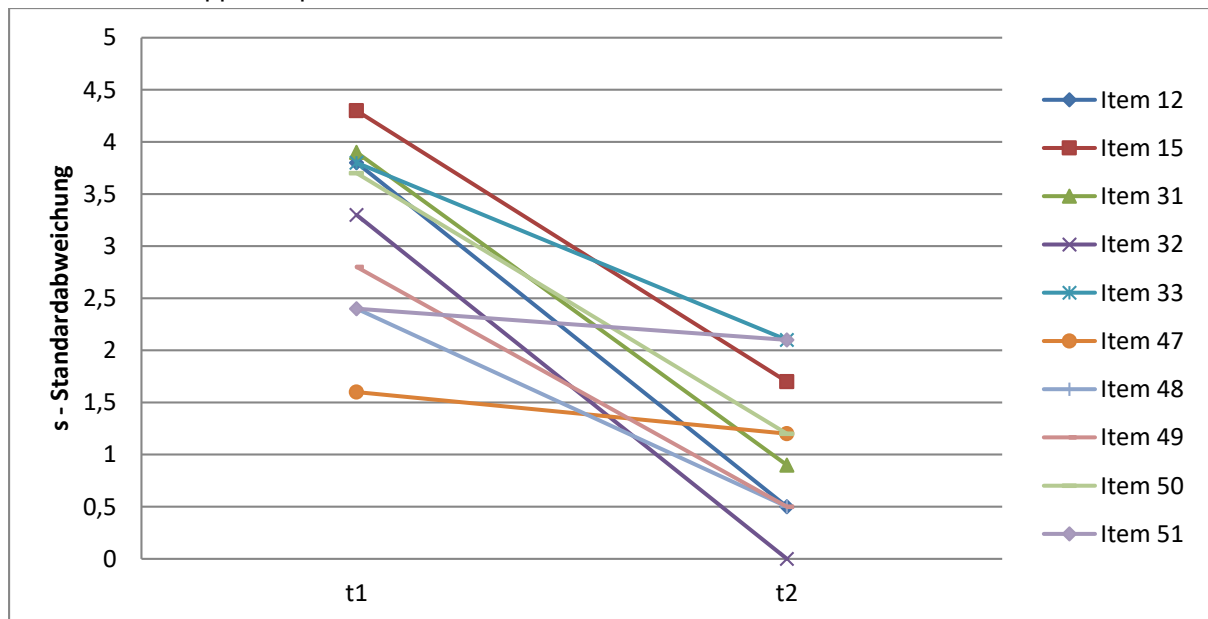
Abbildung 10: Vergleich der Werte der Standardabweichung über die Runden des Gruppendelphis in SAUBER+



Anmerkungen: $tn_0=13$; Item 12, 32, 33 $tn_1=4$, davon jeweils ein MV, alle übrigen Items $tn_1=3$; Item 20, 21 $tn_2=4$, davon jeweils ein MV, alle übrigen Items $tn_2=3$. Quelle: Eigene Darstellung

Verfolgt man die der Dimension der Evidenz zugrunde liegende Annahme, dass sich über den Verlauf eines Gruppendelphis die Werte in Richtung eines Konsenses bewegen, so wird ein solcher Verlauf durch die Werte der Standardabweichung von den in Abbildung 11 aufgezeigten zehn Items bestätigt. Dies gilt mit besonderem Fokus auf die Entwicklung zwischen der ersten Runde (t1) und der zweiten Runde (t2) des Workshoptages, denn hier finden die diskursiven Elemente eines Gruppendelphis ihre Anwendung.

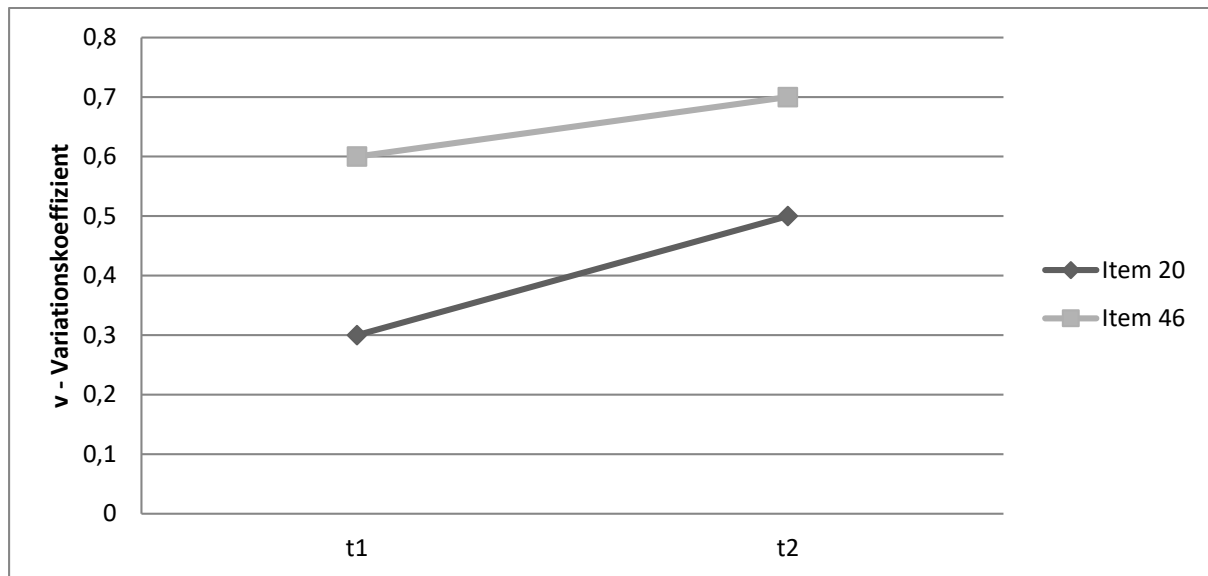
Abbildung 11: Vergleich der Werte der Standardabweichung mit konsensualer Entwicklung über die Runden des Gruppendelphis in SAUBER+



Anmerkungen: Item 12, 32, 33 $tn_1=4$, davon jeweils ein MV, alle übrigen Items $tn_1=3$; $tn_2=3$. Quelle: Eigene Darstellung

Diese zehn Items zeigen damit die idealiter angenommene Entwicklung hin zu einem konsensualen Ergebnis zwischen den Einschätzungen der Teilnehmenden. Aber wie sieht es mit der Entwicklung hin zu einem Konsens über einen Dissens aus? Dieser ist mittels der Standardabweichung nicht klar identifizierbar, allerdings nimmt die Entwicklung des Wertes der Standardabweichung einen gegenläufigen Verlauf als bei einem Konsens – der Wert steigt. Diese Entwicklung bildet zudem der Variationskoeffizient ab. Einen solchen Verlauf zeigen im Fall des Gruppendelphis in SAUBER+ die in Abbildung 12 dargestellten *Items 20* und *36*.

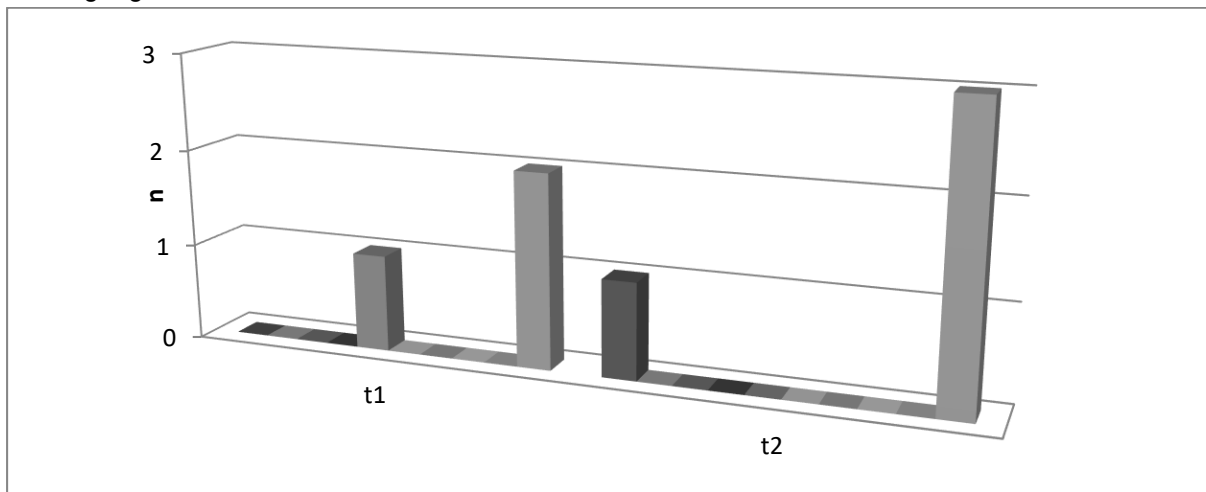
Abbildung 12: Steigende Werte des Variationskoeffizienten über die Runden des Gruppendelphis in SAUBER+



Anmerkungen: Item 20: $tn_1=3$, $tn_2=4$, davon ein Minderheitsvotum; Item 46: $tn_1=3$; $tn_2=3$. Quelle: Eigene Darstellung

Um zu erkennen, ob es sich bei beiden Items um einen Konsens über einen Dissens handelt, sollten für beide Items die Verteilung der jeweiligen Urteile der einzelnen Kleingruppen näher betrachtet werden, denn das arithmetische Mittel gibt in einem solchen Fall nicht die jeweiligen Bewertungen der einzelnen Kleingruppen adäquat wieder. Wie bereits erwähnt, ist ein Konsens über einen Dissens nicht beliebig und nicht jeder Dissens sollte als ein solcher ausgelegt werden. Aus diesem Grund muss für die Überprüfung der Annahme eines Konsenses über einen Dissens ein Fokus auf die Verteilung der Antworten der Teilnehmenden gelegt werden. Abbildung 13 stellt die Verteilung der einzelnen Beurteilungen der jeweiligen Kleingruppen hinsichtlich *Item 20* zur ersten Runde (t_1) und zur zweiten Runde (t_2) des Gruppendelphis dar. Betrachtet man diese Entwicklung, so zeigt sich, dass die Urteile der Stakeholder von der ersten zur zweiten Runde weiter auseinanderliegen. In der ersten Runde wählte eine Gruppe den Wert 5 auf der Skala, wobei die anderen beiden Gruppen den Wert 10 wählten. Zur zweiten Delphirunde bleibt eine unterschiedliche Beurteilung der Aussage, ob bundesweit ein einheitlicher und verbindlicher Standard für die Entsorgung unverbrauchter Arzneimittel über die Apotheken geschaffen werden sollte, bestehen. Allerdings votieren in der zweiten Runde zwei Gruppen dafür, dass sie der *Aussage voll und ganz zustimmen*. Die Möglichkeit der Abgabe von Minderheitsvoten wurde bereits dargestellt und ein solches Minderheitsvotum ist in diesem Fall zu beobachten, denn auch in der anderen Gruppe wählte eine Person den Wert von 10 per Minderheitsvotum. Die anderen Personen dieser Kleingruppe *lehnten* mit ihrem Urteil die Aussage zur zweiten Delphirunde *ganz und gar ab*.

Abbildung 13: Antwortverteilung bei Item 20 „Bundesweit einheitlicher und verbindlicher Standard: Entsorgung“



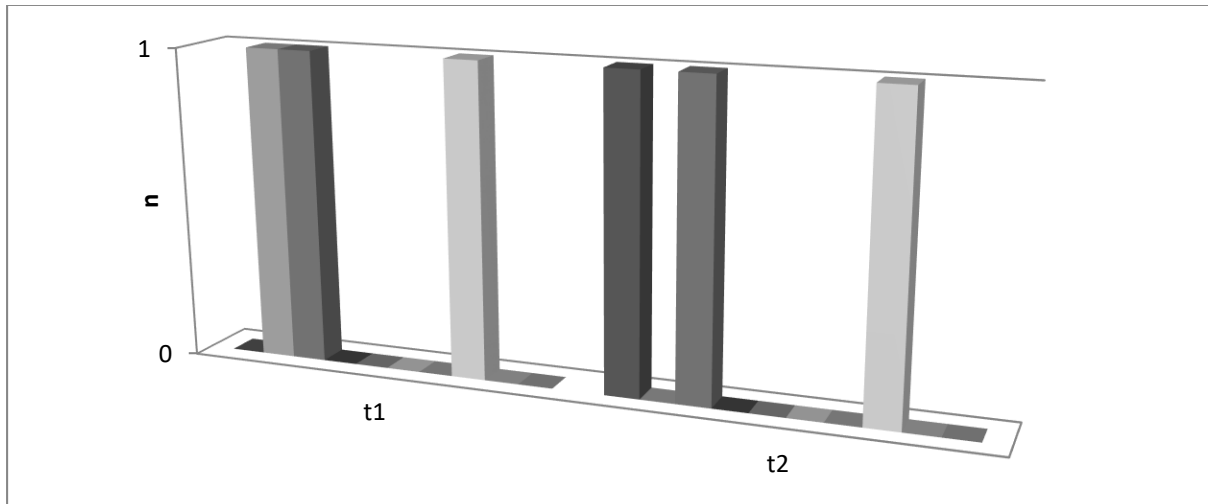
Anmerkungen: tn1=3; tn2=4, davon ein Minderheitsvotum.

Quelle: Eigene Darstellung

Der Diskurs während der ersten Plenarsitzung machte diese Ablehnung bereits deutlich, allerdings wurde sich in der ersten Kleingruppenphase in einer Gruppe auf einen Kompromiss geeinigt und die Aussage mit einem Skalenwert von 5 bewertet. Die Stakeholder äußerten dies in der ersten Plenarphase und stellten dar, dass sie sich in der Kleingruppe nicht einig waren und sich deshalb auf einen Wert *in der Mitte* der 10er-Skala verständigten, dieser aber nicht die Einschätzungen der Einzelpersonen widerspiegeln. In der zweiten gemischten Kleingruppe entschieden sich die Teilnehmenden dann dazu, ihre mehrheitliche Ablehnung mit einem Wert von 1 im Fragebogen anzugeben. Um die Zustimmung der anderen Person abzubilden wurde das Minderheitsvotum mit einem Wert von 10 angegeben. So wurde ein Konsens über einen Dissens in der zweiten Plenarsitzung von den Stakeholdern bestätigt, der sich im Besonderen auf die Strukturen und Rahmenbedingungen begründete, die sich aus Sicht des ablehnenden Teils der Teilnehmenden negativ auf eine erfolgreiche Umsetzung auswirken würden. Der zustimmende Teil der Teilnehmenden sah in diesen kein Hindernis der Aussage zuzustimmen, da sie diese nicht als hemmend einschätzten.

Bei *Item 46*, das den *Handlungsbedarf bei Arzneimittelproduktionsstätten* abfragte, stellt sich die Entwicklung der einzelnen Antworten der Kleingruppen hin zu einem Konsens über einen Dissens als genuine Entwicklung nicht so deutlich dar. Wie in *Abbildung 14* ersichtlich wird, besteht hier bereits zur ersten Runde (t1) eine Verteilung der Beurteilungen der *Dringlichkeit des Handlungsbedarfs bei Arzneimittelproduktionsstätten* durch die einzelnen Kleingruppen, die einen Konsens über einen Dissens markiert. Dieser bleibt auch zur zweiten Runde (t2) trotz veränderter Gruppenzusammensetzung bestehen, so dass zwei Gruppen (Gruppe 1, Gruppe 2) hier *keinen bzw. wenig Handlungsbedarf* sehen und die dritte Gruppe (Gruppe 3) einen *höheren Handlungsbedarf* schätzt. Um zu beurteilen, warum trotz veränderter Gruppenzusammensetzung für dieses Item ein ähnliches Ergebnis auch zur zweiten Runde zu Stande kommt, sollen im Folgenden die Diskussionsinhalte beider Plenarphasen in die Analyse mit aufgenommen werden.

Abbildung 14: Antwortverteilung bei Item 46 „Handlungsbedarf Arzneimittelproduktionsstätten“



Anmerkungen: tn1=3; tn2=3

Quelle: Eigene Darstellung

Die Frage, die der Bewertung der jeweiligen Eintragsquelle in diesem Fragebogen vorausging, lautete: *Wie beurteilen Sie die Dringlichkeit des Handlungsbedarfs bei verschiedenen Eintragsquellen? „Handlungsbedarf“ kann sich auf alle der oben angeführten Maßnahmen einschließlich Forschungsförderung und weitere mögliche Maßnahmen beziehen.*

Der Austausch im ersten Plenum deckte auf, dass während der ersten Kleingruppenphase teilweise das Missverständnis bestand, es ginge bei diesem Fragenblock nur um technische Maßnahmen. Ebenso herrschte in dieser Runde Unklarheit darüber, ob sich die Aussagen auf Deutschland bezogen oder ob die internationale Situation in den Blick genommen werden sollte. Der einführende Hinweis dieser Frage wurde daraufhin ergänzt und ging in die zweite Runde mit einer zusätzlichen Erläuterung der geographischen Reichweite und des Fokus ein: *Deutschland, kein Technikfokus.*⁵⁸

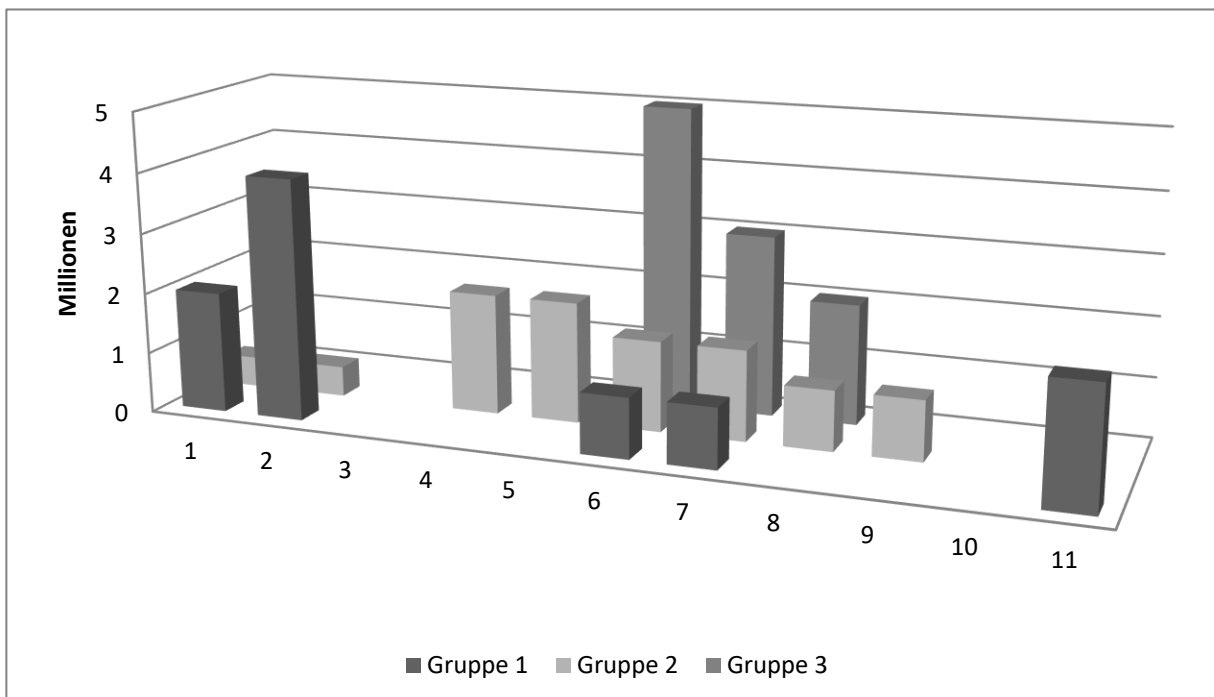
Während sich die Präzisierung bei den anderen Items dieser Frage hin zu einem konsensualeren Ergebnis auswirkte, war dies wie in Abbildung 14 hinsichtlich der Beurteilung des *Handlungsbedarfes bei Arzneimittelproduktionsstätten* nicht der Fall. Die Kleingruppe, die *keinen Handlungsbedarf* sah, führte für ihr Urteil als Argument an, dass in der BRD keine Wirkstoffproduktion stattfindet, sondern lediglich eine Produktion von Fertigarzneimitteln. Da besonders die Wirkstoffproduktion eine Produktionsstätte zur relevanten Eintragsquelle mache, wäre auf internationaler Ebene durchaus ein Handlungsbedarf gegeben, aber nicht bei der Produktion von Fertigarzneimitteln in Deutschland. Die Gruppe, die einen *hohen Handlungsbedarf* bei Arzneimittelproduktionsstätten sah, führte als Argument an, dass auch bei Fertigarzneimittelproduktion Unfälle und Unvorsichtigkeiten auftreten könnten. Inwieweit Unfälle durch beispielsweise ein Monitoring vermieden werden könnten, blieb dabei als offene Frage stehen und konnte im Rahmen des Workshops nicht besprochen werden. Die angebrachten Argumente der Teilnehmenden während der Plenumsphase zeigten, dass alle Stakeholder einen

⁵⁸ Diese Präzisierung wirkte sich auf die anderen abgefragten Eintragsquellen (Item 47, 48, 49, 50, 51) zu einer konsensualeren Bewertung aus (siehe Abbildung 11) und soll im Kapitel 7.1.3.2 näher erläutert werden.

hohen Handlungsbedarf hinsichtlich der *Wirkstoffproduktionsstätten* sahen, allerdings ein Konsens über einen Dissens darüber bestand, inwiefern dies auch für die Produktion von *Fertigarzneimitteln* der Fall sei.

Neben den 10er-skalierten Items wurden bei diesem Gruppendelphi in SAUBER+ auch zwei Budgetfragen abgefragt, bei denen die abgefragten Inhalte nicht nur einzeln für sich betrachtet werden sollten, sondern im Kontext der Gesamtverteilung. Im Ideal bildet sich diesbezüglich über das Gruppendelphi zwischen den Kleingruppen eine konsensuale Verteilung des Budgets heraus. Bei der ersten Budgetfrage des Fragebogens wurden die Teilnehmenden gebeten sich vorzustellen, dass die Bundesregierungen ihnen ein fiktives Budget von 10 Millionen Euro zur Verfügung stellt. Mit diesem Geld sollen neue Maßnahmen finanziert werden um *Gewässerbelastungen* zu verringern. Die Aufgabe der Stakeholder war, dieses Budget auf Maßnahmen zur Verringerung der Gewässerbelastung auf zehn Kontaminanten zu verteilen: *Krankheitserreger, Multiresistente Keime, Nährstoffe, Pestizide, Schwermetalle, Humanarzneimittel, Tierarzneimittel, Diagnostika, Organische Lösungsmittel* und *Streusalz*. Neben diesen gab es zusätzlich die Möglichkeit eine weitere Kontaminante unter der offenen Kategorie *Sonstige* anzugeben. Von dieser Möglichkeit machte eine Gruppe (Gruppe 1) während der ersten Kleingruppenphase gebrauch und verteilte 2 Millionen Euro des Gesamtbudgets auf *Nanopartikel* (Item 11).

Abbildung 15 - Verteilung des Budgets auf Kontaminanten - Runde 1 (t1)

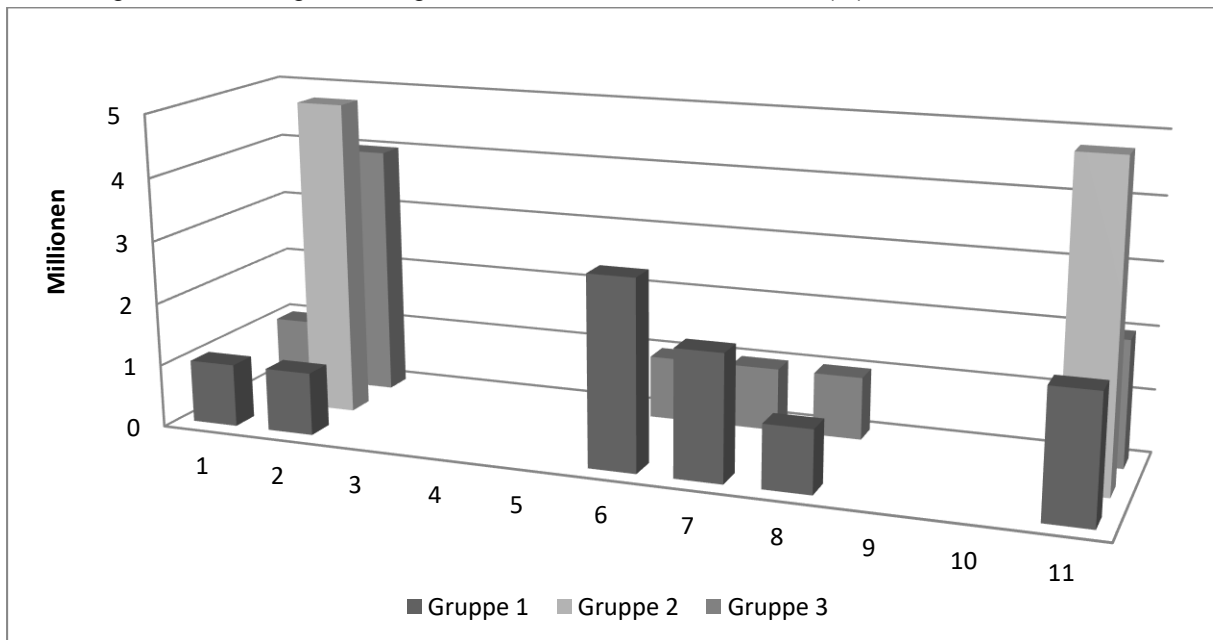


Anmerkungen: Item 1 = Krankheitserreger; Item 2 = Multiresistente Keime; Item 3 = Nährstoffe; Item 4 = Pestizide; Item 5 = Schwermetalle; Item 6 = Humanarzneimittel; Item 7 = Tierarzneimittel; Item 8 = Diagnostika; Item 9 = Organische Lösungsmittel; 1 Item 0 = Streusalz; Item 11 = Nanopartikel. Gesamtbudget = 10 Millionen Euro. Quelle: Eigene Darstellung

Wie in Abbildung 15 dargestellt, fiel die Zuteilung des Budgets über die Gruppen hinweg in der ersten Runde (t1) sehr unterschiedlich aus. Eine Gruppe (Gruppe 2) verteilte das Gesamtbudget von 10 Millionen Euro auf insgesamt 8 Kontaminanten, während eine Gruppe (Gruppe

3) das Gesamtbudget lediglich auf drei Kontaminanten verteilte, darunter auf *Humanarzneimittel (Item 6)* mit dem höchsten Betrag von 5 Millionen Euro. Gruppe 1 verteilte das Gesamtbudget auf fünf Kontaminanten unter diesen auch auf *Nanopartikel*, die als *Item 11* in die Auswertung mit aufgenommen wurden und so auch in den Fragebogen der zweiten Runde (t2) integriert wurden. Auf diese Ergänzung der abgefragten Kontaminanten wurde sich in der ersten Plenarsitzung geeinigt, da alle Teilnehmenden dem Urteil folgten, dass *Nanopartikel* als relevante Kontaminante in die Bewertung mit einfließen sollten. Während der Diskussion im ersten Plenum stellte sich überdies hinaus, dass die Stakeholder bei der Verteilung des Gesamtbudgets in den Kleingruppen unterschiedlichen Grundsätzen gefolgt waren. Während besonders die zweite Kleingruppe versucht hatte so viele Kontaminanten wie möglich mit einem Geldbetrag für Maßnahmen zu versehen, hatte besonders die dritte Gruppe darauf geachtet, besondere Schwerpunkte zu bilden und die aus ihrer Sicht wichtigsten Kontaminanten klar zu markieren. Während die drei Kleingruppen somit in der ersten Runde der Verteilung des Budgets unterschiedlichen Präferenzen folgten, einigte man sich im Rahmen der Plenarsitzung für diese Frage zur zweiten Runde auf den Leitsatz: *Nicht kleckern, sondern klotzen*. So sollte keinem Gießkannenprinzip gefolgt werden, sondern eine Schwerpunktbildung bei der Mittelverteilung gelegt werden. Dieser Schwerpunktbildung folgte in der zweiten Runde (t2) besonders die neu zusammengesetzte Gruppe 2, die das Gesamtbudget gleichermaßen für weitergehende Maßnahmen auf *multiresistente Keime (Item 2)* und *Nanopartikel (Item 11)* verteilte. Auf diese beiden Kontaminanten verteilten auch die beiden anderen Kleingruppen einen Teil des Gesamtbudgets, überdies aber auch über weitere vier Items, wie in Abbildung 16 ersichtlich wird.

Abbildung 16 - Verteilung des Budgets auf Kontaminanten - Runde 2 (t2)



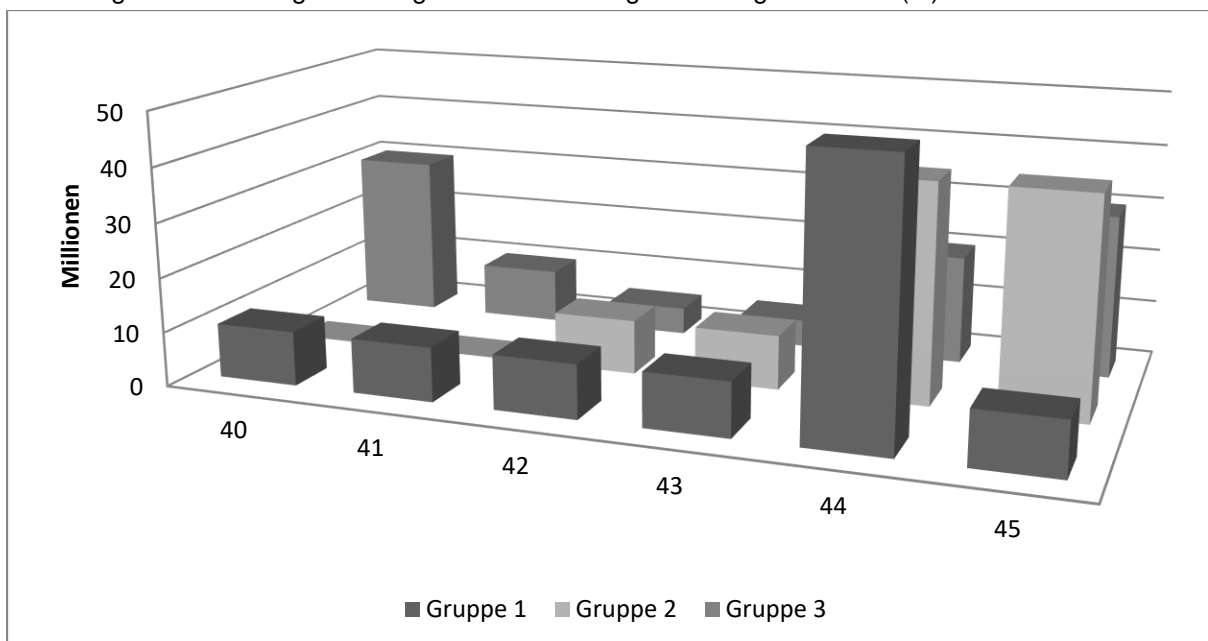
Anmerkungen: Item 1 = Krankheitserreger; Item 2 = Multiresistente Keime; Item 3 = Nährstoffe; Item 4 = Pestizide; Item 5 = Schwermetalle; Item 6 = Humanarzneimittel; Item 7 = Tierarzneimittel; Item 8 = Diagnostika; Item 9 = Organische Lösungsmittel; 1 Item 0 = Streusalz; Item 11 = Nanopartikel. Gesamtbudget = 10 Millionen Euro. Quelle: Eigene Darstellung

Damit stellt sich die Verteilung des Gesamtbudgets in der zweiten Runde (t2) im Vergleich zur ersten Runde (t1) durchaus kongruenter dar. Diese Einschätzung kann durch die Argumenta-

tion im Plenum zusätzlich bestärkt werden. Gruppe 2 erläuterte ihre Verteilung mit dem Verweis darauf, dass sie einen großen Handlungsbedarf bei den zwei von ihnen zugeteilten Maßnahmen sahen und diesen durch die Zuteilung von jeweils 5 Millionen Euro besonders zum Ausdruck bringen wollten. Die anderen Gruppen teilten dieses Urteil, das besonders bei *multiresistenten Keimen (Item 2)* und *Nanopartikeln (Item 11)* ein großer Bedarf an weitergehenden Maßnahmen und überdies der insgesamt höchste Handlungsbedarf bestände. Gruppe 2 teilte ebenso die Einschätzung der beiden anderen Kleingruppen, dass ebenso weitergehende Maßnahmen bei den von diesen Gruppen mit einem Budget bedachten Kontaminanten angemessen seien. Das Urteil der Stakeholder im Plenum war damit ebenso konsensual, dass bei *Krankheitserregern (Item 1)*, *Humanarzneimitteln (Item 6)*, *Tierarzneimitteln (Item 7)* und *Diagnostika (Item 8)* zusätzlicher Handlungsbedarf angemessen sei, der über bestehende Regulierungen und Maßnahmen hinausgehe.

Bei der zweiten Budgetfrage im Abschnitt zur *Forschungsförderung in Bezug auf Arzneimittelrückstände im Gewässerkreislauf* stand den Stakeholdern ein fiktives Budget von 100 Millionen Euro zur Verfügung das die Bundesregierung für Forschungsförderprogramme zur Verfügung stellen würde. Mit diesen Programmen solle die Forschung im Bereich *Arzneimittelrückstände im Gewässerkreislauf* vorangetrieben werden. Die Aufgabe der Teilnehmenden war dieses Geld auf die im Fragebogen aufgeführten Forschungsfelder zu verteilen, eine Kategorie für sonstige Forschungsfelder gab es bei dieser Budgetfrage nicht. Abbildung 17 zeigt die Verteilung des Budgets nach der ersten Kleingruppenphase (t1).

Abbildung 17 - Verteilung des Budgets auf Forschungsförderung - Runde 1 (t1)

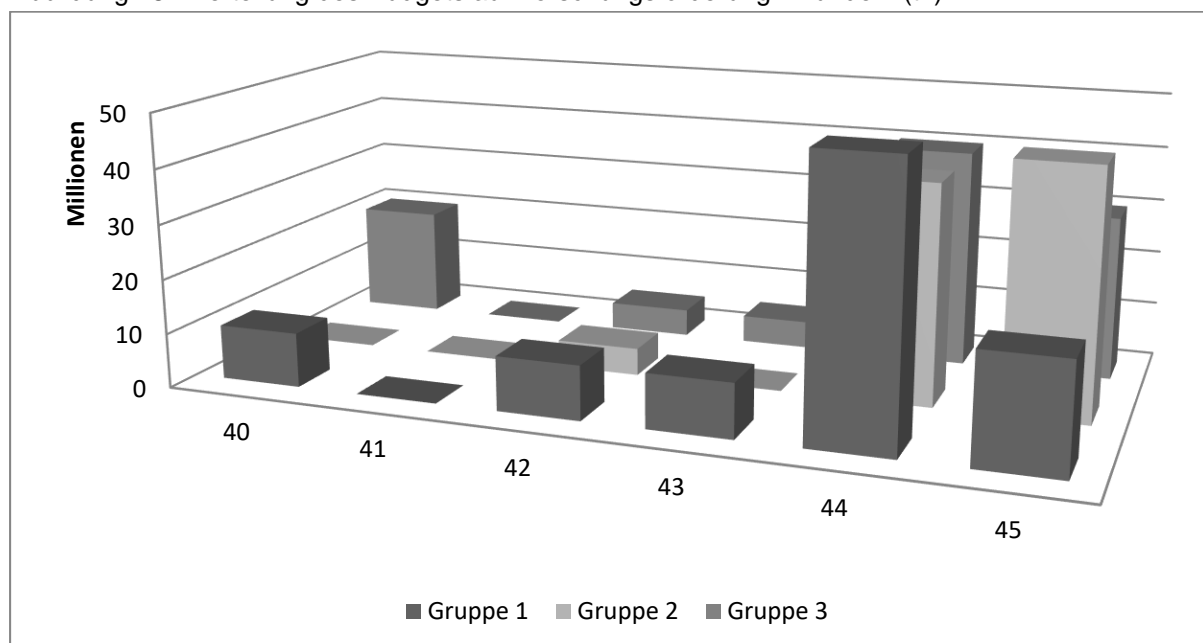


Anmerkungen: Item 40 - Forschung und Entwicklung im Bereich umweltverträgliche Arzneimittelwirkstoffe; Item 41 - Forschung- und Entwicklung im Bereich Verlustminderung pharmakologisch aktiver Substanzen, z.B. durch Optimierung der Formulierung und damit teilweise verbundene Änderung der Darreichungsform; Item 42 - Forschung zum konsequenten Befolgen ärztlicher Therapien durch die Patienten (Gute Compliance), Item 43 - Forschung und Entwicklung im Bereich nicht-medikamentöser Behandlungsformen, Item 44 - Forschung zur Entstehung, Verbreitung und Weitergabe von Antibiotika-Resistenzen, multiresistenten Keimen und Antibiotikaresistenzigenschaften, Item 45 - Forschung zu Abbau- und Transformationsprodukten von Arzneimittelwirkstoffen und ihrer Metaboliten. Gesamtbudget = 100 Millionen Euro. Quelle: Eigene Darstellung

Die grafische Darstellung zeigt auf, dass die Aufteilung des Gesamtbudgets durch die einzelnen Kleingruppen sehr unterschiedlich vorgenommen wurde. Während Gruppe 1 der *Forschung zur Entstehung, Verbreitung und Weitergabe von Antibiotika-Resistenzen, multiresistenten Keimen und Antibiotikaresistenzigenschaften* mit 50 Millionen Euro die Hälfte des Gesamtbudgets zugewiesen hat, hat Gruppe 3 diesem Forschungsfeld mit 20 Millionen Euro ein geringeres Budget zugewiesen. Auch zeigt sich bei der *Forschung und Entwicklung im Bereich umweltverträglicher Arzneimittelwirkstoffe* ein großer Unterschied zwischen den verschiedenen Kleingruppen. So teilte Gruppe 2 diesem Forschungsfeld kein Budget zu, Gruppe 3 hingegen einen hohen Betrag mit 30 Millionen Euro. Da die Zeit der ersten Plenarphase nicht gänzlich ausreichte, konnte diese Frage im Vorfeld der zweiten Kleingruppenphase nicht im Detail besprochen werden, so dass sich die Teilnehmenden für die zweite Runde auch bei dieser Frage zügig auf den Grundsatz *nicht kleckern, sondern klotzen* für die Verteilung des Gesamtbudgets einigten. Dieser wurde durch den Leitsatz ergänzt, dass das Geld vor allem dort hinfließen sollte, wo es notwendig sei Forschungsfelder aufzubauen. So wurde die Verteilung in der zweiten Kleingruppenphase mit dieser Perspektive angestrebt.

Die Verteilung des Gesamtbudgets nach der zweiten Kleingruppenphase (t2), wie in Abbildung 18 dargestellt, wirkt im Vergleich zur Verteilung der ersten Runde (t1) nicht eindeutiger.

Abbildung 18 - Verteilung des Budgets auf Forschungsförderung - Runde 2 (t2)



Anmerkungen: Item 40 - Forschung und Entwicklung im Bereich umweltverträgliche Arzneimittelwirkstoffe; Item 41 - Forschung- und Entwicklung im Bereich Verlustminderung pharmakologisch aktiver Substanzen, z.B. durch Optimierung der Formulierung und damit teilweise verbundene Änderung der Darreichungsform; Item 42 - Forschung zum konsequenten Befolgen ärztlicher Therapien durch die Patienten (Gute Compliance), Item 43 - Forschung und Entwicklung im Bereich nicht-medikamentöser Behandlungsformen, Item 44 - Forschung zur Entstehung, Verbreitung und Weitergabe von Antibiotika-Resistenzen, multiresistenten Keimen und Antibiotikaresistenzigenschaften, Item 45 - Forschung zu Abbau- und Transformationsprodukten von Arzneimittelwirkstoffen und ihrer Metaboliten. Gesamtbudget = 100 Millionen Euro.
Quelle: Eigene Darstellung

Einzig dem Forschungsfeld *Forschung zur Entstehung, Verbreitung und Weitergabe von Antibiotika-Resistenzen, multiresistenten Keimen und Antibiotikaresistenzigenschaften* (Item 44) wurde von den drei Kleingruppen ein ähnlich hohes Budget zugeteilt. Gruppe 2 und Gruppe 3

gaben diesem Forschungsfeld in der zweiten Runde jeweils 40 Millionen Euro, Gruppe 1 teilte diesem Forschungsfeld 50 Millionen Euro zu. In der zweiten Plenarphase konnte allerdings im Vergleich zur ersten Plenarphase ein längerer Austausch der Argumente für die jeweiligen Budgetverteilungen zwischen den Teilnehmenden stattfinden, so dass die verschiedenen Verteilungen auch inhaltlich näher betrachtet werden können. Wie in Abbildung 18 bereits erkennbar, wurde nicht nur Item 44 ein größerer Betrag zugeteilt, sondern auch dem Forschungsfeld zu Abbau- und Transformationsprodukten von Arzneimittelwirkstoffen und ihrer Metaboliten (Item 45).

Diese hohen Budgets wurden den Feldern aus zweierlei Gründen zugeteilt: Zum einen hielten die Stakeholder den Auf- und Ausbau dieser beiden Forschungsfelder für besonders wichtig, zum anderen wurden die Forschungskosten in diesen Bereichen im Vergleich zu den anderen Forschungsbereichen als besonders hoch eingeschätzt. Die Zuteilung erfolgte damit nicht nur unter der Prämisse der Wichtigkeit des Forschungsfeldes, sondern auch unter Beachtung der möglichen Kosten. Die Teilnehmenden führten beispielsweise als Vergleichsgröße die *Forschung zum konsequenten Befolgen ärztlicher Therapien durch die Patienten (Item 42)* an, das von allen Gruppen als ebenso wichtig eingestuft wurde, allerdings keine Labore oder Großgeräte und damit ein geringeres Budget benötige. Dieses Forschungsfeld *koste schlicht weniger, sei allerdings nicht minder wichtig*. Einen Ausbau der *Forschung- und Entwicklung im Bereich Verlustminderung pharmakologisch aktiver Substanzen (Item 41)* hielten die Teilnehmenden dabei konsensual für nicht erstrebenswert, da die pharmazeutische Industrie hier bereits Forschungsanstrengungen unternehmen würde. Beim Forschungsfeld zum *konsequenten Befolgen ärztlicher Therapien durch die Patienten (Item 43)* sahen die Teilnehmenden ebenfalls keine Notwendigkeit für größere Fördersummen. Gruppe 2, die hier kein Budget zugeteilt hatte, schätzte die Bedeutung dieses Feldes als niedrig ein, denn *gute Compliance* sei zwar nicht gänzlich irrelevant, allerdings kein primäres Forschungsfeld, was den Sachverhalt des Gruppendelphis beträfe. Dem widersprachen die beiden anderen Gruppen, die Forschung zu guter Compliance als wichtig erachteten und das Budget nur in geringer Summe verteilten, da sie wie bei *Item 42* den *notwendigen finanziellen Aufwand* für dieses Forschungsfeld als eher gering einstufen. Hier bestand folglich ein Konsens über einen Dissens hinsichtlich der Notwendigkeit dieses Forschungsfeld zu fördern. Insgesamt sahen die Stakeholder geschlossen für vier der abgefragten Forschungsbereiche einen hohen Bedarf an Forschungsförderung, stellten aber unterschiedliche Kostenbedarfe in der Forschungsdurchführung in Rechnung und befürworteten daher auch deutliche Unterschiede in den Fördersummen bei *Item 44* und *Item 45* den Bedarf an einem gezielten und finanzkräftigen Anstoß der Forschung.

Beide Budgetfragen lassen erkennen, dass der Vergleich der einzelnen Budgets beim Gruppendelphi in SAUBER+ für diesen Kreis der Teilnehmenden leicht nachvollziehbar und begründbar war. So war es bei beiden Fragen möglich einen Konsens zu erzielen oder einen Konsens über einen Dissens herauszuarbeiten. Insbesondere die *Forschungsförderung in Bezug auf Arzneimittelrückstände im Gewässerkreislauf* und die darin abgefragten Forschungsfelder zeigen überdies, dass die Abfrage in diesem konkreten Fall zu unterschiedlichen Interpretationen geführt hat, *wie* das Budget verteilt werden soll. Während einzelne Stakeholder aufgrund von Erfahrungen mit Budgetierungen der einzelnen Forschungsfelder die tatsächlich zu veranschlagenden Kosten mit Fokus auf den realen Finanzbedarf einzelner Forschungsfelder in die Verteilung des Budgets mit einbezogen haben, haben andere Teilnehmende das

Budget mit Fokus darauf verteilt, für *wie wichtig* sie die einzelnen Forschungsfelder erachten. Da diese Frage nicht erschöpfend im ersten Plenum diskutiert werden konnte, konnte dieser unterschiedliche Bewertungshorizont erst zur zweiten Plenarsitzung aufgedeckt und besprochen werden. So konnte durch den Austausch der Argumente letztlich klar herausgestellt werden, wo eine Forschungsförderung aus Sicht der Teilnehmenden aus- bzw. aufgebaut werden sollte (*Item 40, 42, 44, 45*), wo dies nicht geschehen sollte (*Item 41*), wie die Kosten dem Finanzierungsbedarf entsprechen könnten und um letztlich einen Konsens über Dissens bezüglich eines Forschungsfeldes (*Item 43*) zu markieren.

Neben den Items die in der zweiten Runde des Gruppendelphis abgefragt wurden, existieren überdies auch Items, deren enthaltene Aussagen in der ersten Runde von den Gruppen unterschiedlich beurteilt wurden, aber nicht mehr in den Fragebogen der zweiten Runde abgefragt wurden. Grund hierfür ist ein in der Plenarsitzung ausgearbeiteter Konsens zwischen den Teilnehmenden, diese Aussage aufgrund von vorgebrachten Argumenten geschlossen abzulehnen.

Ein Beispiel für eine solche Aussage, stellt das *Item 17* dar: *Das Angebot an Arzneimitteln in variablen Packungsgrößen sollte erweitert und die Packungsgröße an die jeweilige Therapie besser angepasst werden.* Der Variationskoeffizient betrug bei diesem Item nach der ersten Kleingruppenphase einen Wert von $v=0,6$. Während die Gruppen 2 und 3 die Aussage ablehnten (3), stimmte Gruppe 1 dieser Aussage *voll und ganz* zu (10). Die ablehnenden Gruppen begründeten in der Plenarsitzung ihre Ablehnung damit, dass es bereits ausreichend differenzierte Packungsgrößen mit den Kennzeichnungen *N1, N2* und *N3* gäbe. Durch das *Arzneimittelmarktneuordnungsgesetz* (2011) wurden die Kennzeichnungen *N1, N2* und *N3* zusätzlich auf die Dauer der typischen Behandlungszyklen abgestimmt. Eine weitere Differenzierung stiftete überdies nur Verwirrung und wäre ebenso ein kaufmännisches Problem. Die zustimmende Gruppe ließ sich von diesen Argumenten überzeugen und teilte nun die ablehnende Bewertung. Auf Grund des erreichten Konsenses wurde diese Maßnahme in der zweiten Runde des Delphis nicht nochmals abgefragt, da die Maßnahme geschlossen abgelehnt wurde.

Die evidenzbasierte Betrachtung der empirischen Daten zeigt, dass die vermuteten Entwicklungen wie dargestellt anhand der statistischen Daten und - ebenso gleichwertig - aufgrund der wissensbasierten Argumentationen während der Plenarsitzung nachgewiesen werden können. Als nächste Dimension soll dem Forschungsdesign folgend eine tiefergehende Betrachtung der Dimension der Verständlichkeit erfolgen.

7.1.3.2 Dimension: Verständlichkeit

Die Dimension der Verständlichkeit dient der Überprüfung, ob die Methode des Gruppendelphis dem Geltungsanspruch der Zweckmäßigkeit für den Diskurs entsprechen kann, die Teilnehmenden folglich mit dem *Geäußerten* einen gleichen Bedeutungsinhalt verbinden. Das Gruppendelphi muss einen Raum dafür bieten, eben einen solchen gemeinsamen Bedeutungsinhalt für die strittigen Begriffsdefinitionen, Klassifikationsschemata und Formulierungen von Aussagen zu erzielen. So können interdisziplinär unterschiedlich konzeptuierte Fachtermini, semantische Missverständnisse oder durch das Forscherteam nicht klar formulierte Aussagen im Fragebogen zu unterschiedlichen Interpretationen durch die Teilnehmenden führen.

Da der Fragebogen im Falle des Gruppendelphis in SAUBER+, wie in der methodischen Konzeption bei semantischen Missverständnissen angedacht, zwischen der ersten Runde und der zweiten Runde angepasst wurde, kann der erste Schritt der Analyse auf diese Items fokussiert erfolgen. Bei SAUBER+ wurden vier Aussagen zwischen der ersten und zweiten Runde des Gruppendelphis umformuliert. Dies erfolgte während des ersten Plenums in Abstimmung mit den teilnehmenden Stakeholdern:

- **Item 12:** Die Formulierung dieser Aussage im Fragebogen zur ersten Delphirunde (t1) lautete: *Bei der Zulassung von Humanarzneimitteln sollte das Umweltrisiko in der Nutzen-Risiko-Analyse des Arzneimittels berücksichtigt werden.* Der Variationskoeffizient wies nach der ersten Runde einen hohen Wert von $v=0,8$ auf. Während die Aussage von Gruppe 1 (9) und Gruppe 3 (8) mit hoher Zustimmung beurteilt wurde, lehnten Gruppe 2 (1) sowie ein Minderheitsvotum (1) in Gruppe 1 diese Aussage *ganz und gar* ab. Die Gründe der jeweiligen Bewertungen der Teilnehmenden wurden in der Plenarsitzung ausgetauscht und es wurde offenkundig, dass die Aussage bezüglich der Maßnahme geschärft werden müsse, da das Umweltrisiko bei der Zulassung zum Zeitpunkt des Delphis bereits berücksichtigt wurde. So müsse zwischen einer *generellen Berücksichtigung* des Umweltrisikos und einem Umweltrisiko als *Versagungsgrund* unterschieden werden. Dies war einigen Stakeholdern bei der Bewertung in der ersten Runde (t1) nicht bewusst und so wurde sich darauf geeinigt, die Aussage im Fragebogen für die zweite Runde (t2) um den *Versagungsgrund* zu ergänzen: *Bei der Zulassung von Humanarzneimitteln sollte das Umweltrisiko in der Nutzen- Risiko- Analyse des Arzneimittels mit Versagungsgrund berücksichtigt werden.*
- **Item 33:** Die Formulierung dieser Aussage im Fragebogen zur ersten Delphirunde (t1) lautete: *Krankenhäuser und Spezialkliniken sollten ihr Abwasser generell vorbehandeln und erst nach dieser Vorbehandlung in das kommunale Kanalnetz einleiten.* Der Variationskoeffizient wies nach der ersten Runde einen hohen Wert von $v=0,7$ auf. Während die Aussage von Gruppe 3 mit hoher Zustimmung (9) bewertet wurde, lehnten Gruppe 1 und Gruppe 2 diese Aussage mit Bewertungen von 3 und 1 ab. In Gruppe 1 gab es zudem ein Minderheitsvotum, das dieser Aussage *voll und ganz* zustimmte (10). Diese Aussage wurde während der Plenarsitzung mit den Stakeholdern vertiefend diskutiert. Die Gruppen mit hoher Ablehnung stellten heraus, dass spezielle Abwasserströme unterschiedlich behandelt werden sollten. Es sei immer sinnvoll, dort zu behandeln, wo wirklich hohe Frachten an problematischen Stoffen anfielen. Vor einer Behandlung des Abwassers vor Ort sollte geprüft werden, ob eine solche Vorbehandlung vor der Einspeisung auch notwendig und sinnvoll sei. Dem stimmten auch die Stakeholder aus der dritten Gruppe mit der hohen Bewertung (10) sowie der Stakeholder mit dem hohen Minderheitsvotum (10) zu. Diese ließen sich von dem Argument der ablehnenden Gruppen überzeugen, dass eine *generelle* Vorbehandlung bei Krankenhäusern und Spezialkliniken keine diskutabile Option sei. Auch ein Fokus auf Krankenhäuser und Spezialkliniken sei nicht zielführend und so wurden Gesundheitseinrichtungen in ihrer Gesamtheit in diese Aussage aufgenommen. Die von allen Stakeholdern konsensual angenommene und neu formulierte Aussage lautete für die zweite Runde des Gruppendelphis wie folgt: *Gesundheitseinrichtungen sollten bedarfsorientiert ihr Abwasser vor Einleitung in das kommunale Kanalnetz vorbehandeln.*
- **Item 37:** Die Formulierung dieser Aussage im Fragebogen zur ersten Delphirunde lautete: *Die Liste prioritärer Stoffe der EU-Wasserrahmenrichtlinie sollte um multiresistente Keime erweitert werden.* Diese Aussage wurde von Gruppe 1 und Gruppe 2 mit Werten von 2 und

1 stark abgelehnt, während Gruppe 3 dieser Aussage zustimmte (9), der Variationskoeffizient wies einen Wert von $v=0,9$ (t1) auf. Die ablehnenden Gruppen begründeten ihre Ablehnung damit, dass die *EU Wasserrahmenrichtlinie (WRR)* besonders den Schutz des Ökosystems als Ziel hätte und somit automatisch eine Fokussierung auf Schadstoffe erfolge. Es bedürfe damit anderer Regulierungsmöglichkeiten für multiresistente Keime, denn eine Aufnahme dieser Keime würde die Systematik der WRR sprengen. Zudem bestehe bei multiresistenten Keimen noch zu viel Ungewissheit und damit verbunden ein bestehend großer Forschungsbedarf. Die dritte Gruppe, die der abgefragten Aussage eine hohe Zustimmung zukommen gelassen hatte (9) begründete diese Zustimmungen damit, dass eben wegen dieser hohen Ungewissheit eine Überwachung des Vorkommens multiresistenter Keime und weitere Forschung notwendig wären. Alle Teilnehmenden waren sich darin einig, dass Maßnahmen nur dann ergriffen werden sollten, wenn die Ergebnisse von Forschung und Überwachung (Monitoring) dies anzeigten. Die teilnehmenden Stakeholder einigten sich konsensual darauf, die Aussage in diesem Sinne umzuformulieren: *Multiresistente Keime sollten überwacht und gegebenenfalls (nach Forschungsstand) Maßnahmen ergriffen werden.*

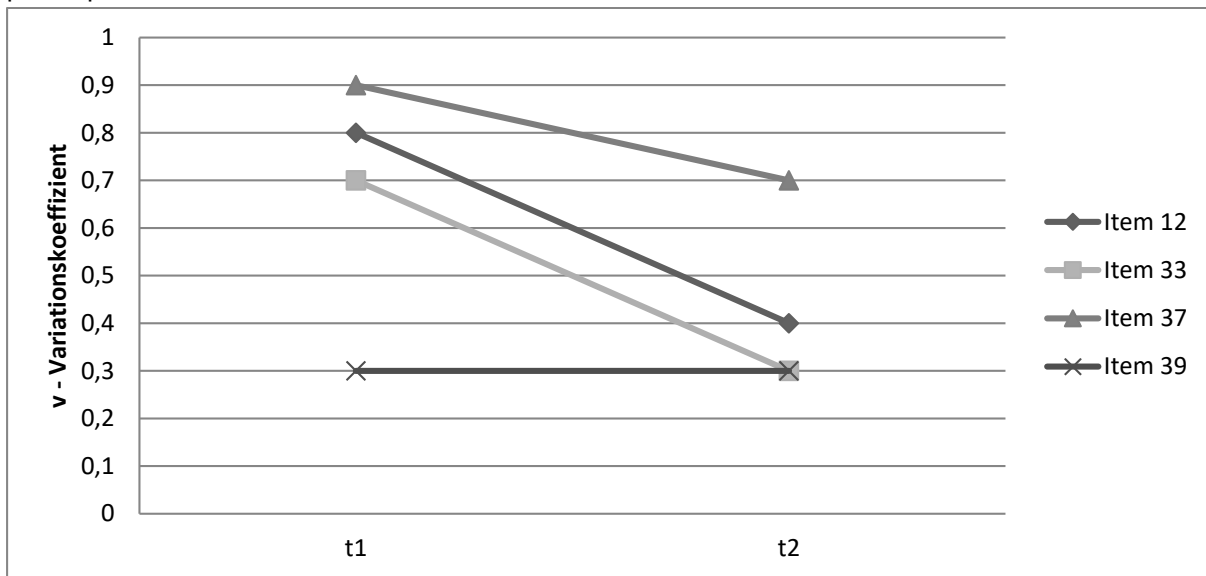
- **Item 39:** Die Formulierung dieser Aussage im Fragebogen zur ersten Delphirunde lautete: *Die novellierte deutsche Klärschlammverordnung sollte ein generelles Verbot des Aufbringens von Klärschlamm auf Lebensmittelböden (Anbauflächen für Nahrungs- und Futtermittel) umfassen. Für Energieböden (nachwachsende Rohstoffe) sollten andere Bestimmungen gelten.* Trotz eines niedrigen Wertes des Variationskoeffizienten von $v=0,3$ wurde diese Maßnahme in der Plenarsitzung (t1) diskutiert. Dies lag an der Verteilung der Urteile begründet, da Gruppe 1 der Aussage mit einer Bewertung von 7 eher zustimmte, während Gruppe 2 mit einem Wert von 5 und Gruppe 3 mit einem Wert von 3 die Aussage eher ablehnten, so dass auch hier ein Austausch über die Gründe für die unterschiedlichen Bewertungen aus Sicht der Moderation sinnvoll erschien. Besonders wurde bei dieser Aussage diskutiert, ob man Energieböden und Lebensmittelböden gesondert betrachten sollte. So argumentierten Teilnehmende der Gruppe 1, dass eine Unterscheidung von beiden Bodenarten in diesem Kontext deshalb keinen Sinn ergebe, da die Aufbringung von Klärschlamm immer eine Grundwassergefährdung beinhalten würde. Das Urteil der Gruppe begründete sich dabei besonders auf die generelle Gefährdung des Grundwassers, die mit dem Düngen mittels Klärschlamm verbunden sei. Gruppe 3 hingegen legte ihrer Beurteilung das Potential des Klärschlammes zu Grunde, da Alternativen bestünden bzw. weiterentwickelt werden sollten, bei denen die im Klärschlamm enthaltenen Nährstoffe gewonnen und als Dünger genutzt werden könnten. Die Teilnehmenden konnten sich für die zweite Runde auf eine Umformulierung der Aussage einigen, die beide Standpunkte pointierter abbildete: *Die novellierte deutsche Klärschlammverordnung sollte ein generelles Verbot des Aufbringens von Klärschlamm **auf landwirtschaftlich genutzten Böden** umfassen.*⁵⁹

Betrachtet man die Werte des Variationskoeffizienten der vier Aussagen, die zwischen den beiden Runden des Gruppendelphis umformuliert wurden, so ergibt sich, wie in Abbildung 19 veranschaulicht, kein gänzlich einheitliches Bild hin zu einer konsensualeren Bewertung der Items. Bei den *Items 12, 33 und 37* deutet der jeweilige Wert des Variationskoeffizienten darauf

⁵⁹ Hervorhebung im Original

hin, dass die Teilnehmenden zur zweiten Runde eine konsensualeren Beurteilung der jeweiligen Aussagen gefunden haben. Allerdings liegt der Variationskoeffizienten bei *Item 37* mit einem Wert von $v=0,7$ noch immer deutlich über dem Erfahrungswert von $0,5$, der auf ein konsensuales Urteil hindeuten würde. Bei *Item 39* ist der ohnehin schon niedrigere Wert von $v=0,3$ in der ersten Runde zur zweiten Runde bestehen geblieben.

Abbildung 19: Werte des Variationskoeffizienten der umformulierten Items über die Runden des Gruppendelphis in SAUBER+



Anmerkungen: Item 12, 13: $tn_1=4$, Item 37, 39: $tn_1=3$; $tn_2=3$.

Quelle: Eigene Darstellung

Bei *Item 37* bleiben die Bewertungen der Stakeholder trotz neu zusammengesetzter Gruppen sehr unterschiedlich. So lehnte Gruppe 2 die Aussage stark ab (1), während Gruppe 3 der Aussage voll und ganz zustimmte (10) und Gruppe 1 eine eher ablehnende Beurteilung (5) abgab. Die ablehnende Haltung wurde damit begründet, dass momentan zu wenig Wissen vorhanden sei, um zu beurteilen, ob die Maßnahme sinnvoll sei. Die Befürworter*innen der Maßnahme sprachen sich für eine systematische Erhebung über *Sondermessprogramme* an einer begrenzten Zahl von Kläranlagen aus, um so ein grobes Bild über die aktuelle Situation zu generieren. Der Dissens, ob eine Überwachung multiresistenter Keime grundsätzlich notwendig sei, blieb damit bestehen. Es gab allerdings Einigkeit darüber, dass es keine flächendeckende Überwachung der Keime geben sollte. Bei *Item 39* konnte in der zweiten Plenarrunde aufgrund von zeitlichen Beschränkungen kein Austausch über die jeweiligen Gründe für die einzelnen Bewertungen erfolgen. Diese Diskussion fiel anhand des niedrigen Variationskoeffizienten mit einem Wert von $v=0,3$ im Vergleich zu anderen Items mit höheren Werten für dieses Item aus. Allerdings lohnt sich bei diesem Item der Einbezug des arithmetischen Mittels, das von $\bar{x}=5,0$ (t_1) auf $\bar{x}=7,3$ (t_2) angestiegen ist. Über die Gruppen hinweg wurde diese Aussage also von den Stakeholdern nach der Umformulierung insgesamt positiver beurteilt.⁶⁰ *Item 12* wurde in der zweiten Runde von den teilnehmenden Stakeholdern geschlossen abgelehnt $\bar{x}=1,3$, während *Item 33* mit $\bar{x}=7,3$ (t_2) höhere Zustimmungswerte erfuhr.⁶¹

⁶⁰ Gruppe 1 = 10; Gruppe 2 = 5; Gruppe 3 = 7

⁶¹ Die Argumente für die Ablehnung von Item 12 werden in Kapitel 7.1.3.3 näher beschrieben

Neben den umformulierten Aussagen wurden bei diesem Gruppendelphi im Fragebogen der zweiten Runde (t2) optische Veränderungen (bzw. Hervorhebungen) bei einzelnen Aussagen vorgenommen. Dies erfolgte bei drei Items:⁶²

- **Item 15:** *Es sollte eine auf Umweltverträglichkeit bezogene Produktkennzeichnung von Medikamenten als Entwicklungsanreiz für eine „grüne Produktpolitik“ der Arzneimittelhersteller eingeführt werden.* Die Hervorhebung erfolgte aufgrund der Diskussion im ersten Plenum. So wurde das *Item 15* von Gruppe 1 (9) und Gruppe 3 (10) als sehr relevant bewertet, während die zweite Gruppe (1) sowie ein Minderheitsvotum in Gruppe 1 (1) die abgefragte Aussage stark ablehnten, der Variationskoeffizient hatte einen Wert von $v=0,8$. Die ablehnende Gruppe begründete die hohe Ablehnung damit, dass die Produktkennzeichnung aus ihrer Sicht kein effektiver Entwicklungsanreiz für ein Pharmaunternehmen sei, da hier die ökonomischen Gewinne gegen Null tendierten. Die Arzneimittelentwicklung und die Wirkstoffentwicklung würden in diesem frühen Entwicklungsstadium völlig anderen Prioritäten folgen, wie beispielsweise neuen Krankheitsbildern. Die grafische Veränderung der Aussage erfolgte weiterhin darauf, dass es unterschiedliche Bewertungsgrundlagen der Stakeholder in den verschiedenen Kleingruppen gab. So verstanden die zustimmenden Teilnehmenden die Produktkennzeichnung generell als eine gute Maßnahme, wobei die ablehnenden Gruppen explizit eine Produktkennzeichnung als potentiellen *Entwicklungsanreiz* bewerteten. Da dies aus Sicht der Stakeholder den Kern der Aussage bildete, wurde die Textformatierung zur zweiten Runde vorgenommen.
- **Item 21:** *Ein bundesweit einheitlicher und verbindlicher Standard für die Entsorgung unverbraucher Arzneimittel über die Apotheken, bei dem die nicht frei verkäuflichen Medikamente einem Pfandsystem unterliegen, sollte erprobt werden (der Arzneimittelnutzer erhält bei Rückgabe des angebrochenen Medikaments oder der Verpackung einen monetären Wert zurück, den er beim Kauf des Arzneimittels mit entrichtet hat).* *Item 21* wurde von der ersten sowie der zweiten Kleingruppe in der ersten Runde *voll und ganz* abgelehnt (1/1) während die dritte Kleingruppe diesem absolut zustimmte (10), der Variationskoeffizient hatte einen ungewöhnlich hohen Wert von $v=1,1$. Die zustimmende Gruppe wies umgehend während der ersten Plenarsitzung auf ein in der Planung bestehendes Pilotprojekt hin, das den Rücklauf von Medikamenten und deren Verpackungen mittels eines Bonussystems honorieren sollte. Ein solches System müsse dann allerdings beispielsweise über die Krankenkassen organisiert werden, da aufgrund von gesetzlichen Bestimmungen kein Geld von den Apotheken an die Kundschaft ausgezahlt werden dürfe. Die erste ablehnende Gruppe begründete ihre starke Ablehnung mit dem immensen bürokratischen Aufwand, der durch ein solches System entstehe. Zudem erwecke ein solches Pfandsystem den falschen Eindruck, dass Medikamente wie Wertstoffe zu behandeln seien, was sie de facto nicht wären. Die zweite Gruppe begründete ihre starke Ablehnung mit der Befürchtung, dass ein solches Pfandsystem für die Kund*innen einen Anreiz schaffen könnte, die Medikamenteneinnahme vorzeitig abzubrechen, um für die restlichen Medikamente einen Geldbetrag zu erhalten. Der Austausch im Plenum deckte hier allerdings ein Missverständnis auf, da diese Gruppe in ihr Urteil zur Aussage nicht mit einbezogen hatte, dass die Maßnahme auch für

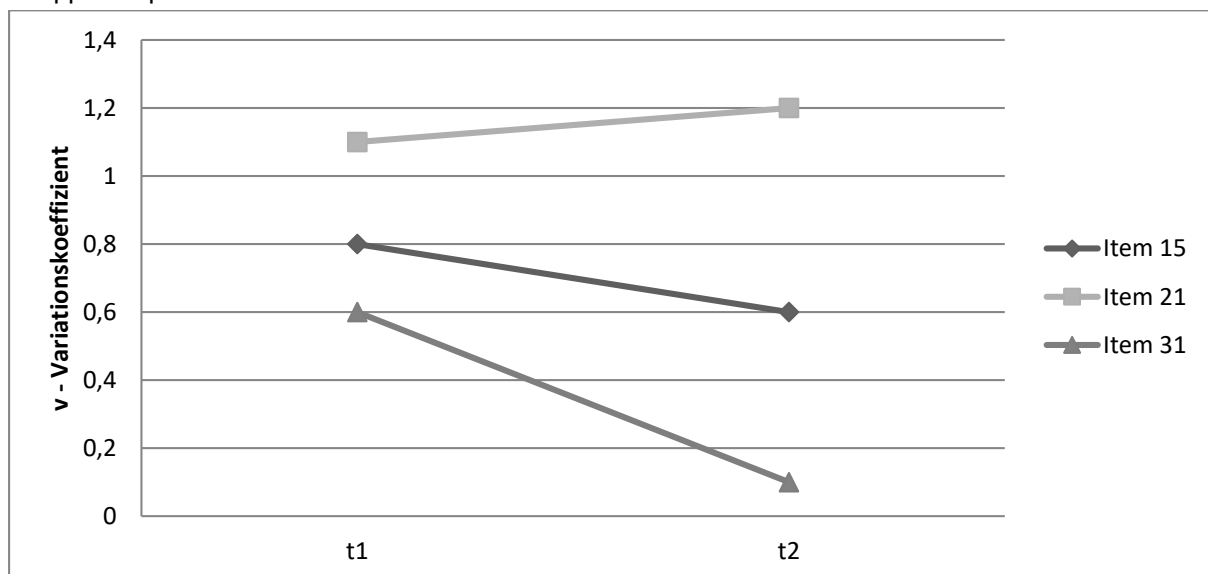
⁶² Bei *Item 15* und *Item 21* wurden einzelne Teile der Aussage mittels fetter und unterstrichener Schrift zusätzlich hervorgehoben, bei *Item 31* erfolgte dies mittels Unterstrich.

leere Verpackungen ein solches Pfand vorsah. Dieser Teil der Aussage wurde im Fragebogen für die zweite Runde als Erinnerungsstütze optisch hervorgehoben.

- **Item 31:** Bundesweit sollten Apotheker und Ärzte zu „Antibiotikabeauftragten“ weitergebildet werden, die im Anschluss in Krankenhäusern als Berater tätig sind (Bsp. Niedersachsen). Die in diesem Item abgefragte Aussage erhielt in der ersten Runde von der ersten Gruppe (10) und der zweiten Gruppe (8) hohe Zustimmungswerte, wurde allerdings von der dritten Gruppe stark abgelehnt (1), so dass der Wert des Variationskoeffizienten $v=0,7$ betrug. Die ablehnende Gruppe begründete ihre Ablehnung mit dem Argument, dass das Wissen über Antibiotika bereits zum Zeitpunkt des Delphis ein Element der Grundausbildung jeder Ärzt*in sowie jeder Apotheker*in sei. Dem wurde von den Teilnehmenden mit zustimmender Bewertung entgegnet, dass ein solches Grundlagenwissen zwar vorhanden sei, aber häufig nicht danach gehandelt würde, da ein bloßes Grundlagenwissen hinsichtlich des Sachverhaltes nicht ausreichend sei. So wurde auf ein System in den *Niederlanden* hingewiesen, wo es bereits entsprechendes Fachpersonal gäbe. In *Nordrhein-Westfalen* würde dem Beispiel der Niederlande gefolgt und in diesem Bundesland sei ein spürbarer Rückgang von Antibiotika zu verzeichnen. Daher könnten *Beauftragte* besonders in Krankenhäusern darauf achten, dass die Sensibilität für einen angemessenen Umgang mit Antibiotika erhöht würde. Aufgrund der Verortung der Relevanz in die Einrichtung des Krankenhauses, wurde dies im Fragebogen (t2) mittels Unterstrichs hervorgehoben.

In Abbildung 20 wird ersichtlich, dass die Entwicklung der Werte des Variationskoeffizienten hinsichtlich der grafisch markierten Items unterschiedlich verläuft. Während sich der Wert des *Items 31* von 0,6 (t1) zu einem konsensualen Ergebnis von $v=0,1$ (t2) entwickelt, zeigen *Item 21* mit einem Anstieg auf den sehr hohen Wert von $v=1,2$ (t2) und *Item 15* zeigt mit einem Wert von $v=0,6$ (t2) weiterhin eine unterschiedliche Beurteilung der Aussagen durch die verschiedenen Kleingruppen an.

Abbildung 20: Werte des Variationskoeffizienten der grafisch bearbeiteten Items über die Runden des Gruppendelphis in SAUBER+



Anmerkungen: tn1=3; Item 15, 31: tn2=3, Item 21: tn2=4.

Quelle: Eigene Darstellung

Auch wenn der Variationskoeffizient mit einem Wert von $v=0,6$ bei *Item 15* hoch ist, so waren sich die Kleingruppen in der Ablehnung dieser Aussage in der zweiten Runde einig.⁶³ Eine konsensualere Bewertung bestätigt dabei auch der Blick auf die Standardabweichung, die mit einem Wert von $s=4,3$ (t_1) auf einen Wert von $s=2,7$ (t_2) gesunken ist. Der hohe Wert des Variationskoeffizienten lässt sich durch das niedrige arithmetische Mittel von $\bar{x}=2,7$ als Divisor des Variationskoeffizienten ableiten. Inhaltlich waren sich die Teilnehmenden darüber einig, dass eine auf Umweltverträglichkeit bezogene Produktkennzeichnung von Medikamenten kein effektiver Entwicklungsanreiz für Pharmakonzerne sei. Da die Entwicklung einer neuen Substanz mit circa einer Milliarde Euro bis zur Zulassung zu veranschlagen sei, wurde der Zusatznutzen über die Nachfrage nach gekennzeichneten umweltverträglicheren Medikamenten als vernachlässigbar eingestuft. In dieser Einschätzung waren sich die Teilnehmenden einig. Die höhere Bewertung der abgefragten Maßnahme durch Gruppe 1, wurde von dieser damit begründet, dass diese Logik insbesondere auf *große Pharmakonzerne* zuträfe. Für Start-up-Unternehmen und kleinere Unternehmen könnte eine solche *grüne Produktkennzeichnung* aber sehr wohl ein gutes Marketinginstrument sein und sei deshalb aus dieser Perspektive nicht generell abzulehnen. Auf diese Beurteilung konnten sich die Gruppen in der zweiten Plenarphase konsensual einigen.

Bei *Item 21* lehnten alle drei Gruppen die Aussage zum Pfandsystem ganz und gar ab (1). Allerdings gab es in Gruppe 2 ein Minderheitsvotum, dass der Aussage voll und ganz zustimmte (10). Diese Zustimmung des einzelnen Stakeholders erfolgte mit der Begründung, dass ein Bonussystem sinnvoll sei, um eine Rückgabe unverbrauchter Arzneimittel zu garantieren. Dabei seien verschiedene Bonussysteme denkbar: über Geld, geldwerte Boni oder verminderte Zuzahlungen. Die restlichen Teilnehmenden betonten daraufhin erneut ihre Ablehnung eines pfandgestützten Systems, denn ein solches System sei neben zusätzlichen bürokratischen Hürden überdies nicht verbraucherfreundlich. Hier gab es letztlich einen Dissens über die Wirksamkeit eines pfandgestützten Systems. Die Teilnehmenden konnten sich allerdings als *Minimalkonsens* darauf einigen, dass ein bundesweit einheitlicher Entsorgungsstandard notwendig sei.

Der abgefragten Aussage in *Item 31* wurde zur zweiten Runde von den Gruppen 2 und 3 voll und ganz zugestimmt (10) und erfuhr ebenso von Gruppe 1 (8) eine hohe Zustimmung, so dass die Maßnahme, dass bundesweit Apotheker*innen und Ärzt*innen zu *Antibiotikabeauftragten* weitergebildet werden sollten, die im Anschluss in Krankenhäusern als Berater*innen tätig sein könnten, konsensual zugestimmt wurde. Die Aussage wurde somit im zweiten Plenum nicht weiter diskutiert.

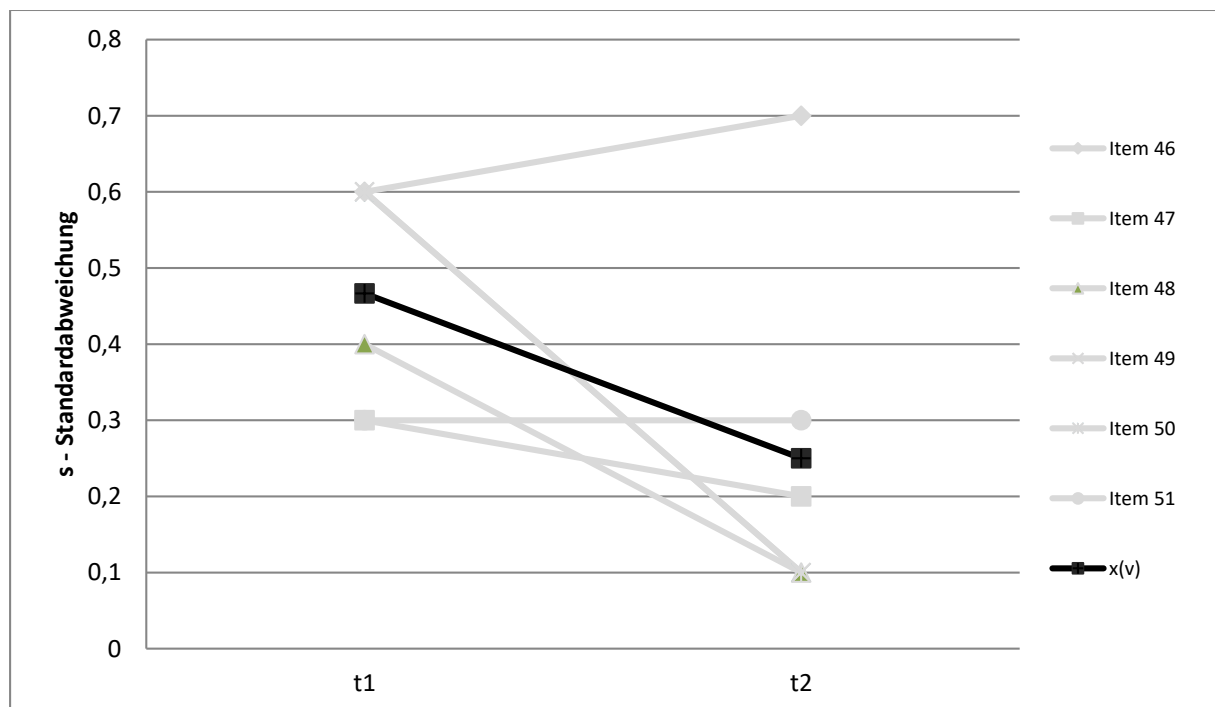
Wie bereits im vorherigen Kapitel 7.1.3.1 beschrieben, wurde beim sechsten Frageabschnitt über die *Dringlichkeit des Handlungsbedarfes nach Eintragsquellen* zur zweiten Runde eine Ergänzung des Einführungstextes im Fragebogen vorgenommen, da der Austausch zwischen den Teilnehmenden im ersten Plenum aufdeckte, dass teilweise das Missverständnis bestand, es ginge bei diesem Fragenblock nur um technische Maßnahmen. Überdies bestand in der ersten Runde Uneinigkeit darüber, ob die Maßnahmen sich auf die *BRD* bezogen oder die internationale Perspektive beurteilt werden sollte. Der einführende Hinweis dieser Frage wurde

⁶³ Gruppe 1 = 5; Gruppe 2 = 2; Gruppe 3 = 1.

daraufhin um *Deutschland, kein Technikfokus* ergänzt. Aus Zeitmangel konnte nicht jede Eintragsquelle hinsichtlich des von den Stakeholdern eingeschätzten Handlungsbedarfs diskutiert werden, sondern es erfolgte ein Maßnahmen-übergreifender Austausch.

Betrachtet man die in Abbildung 21 veranschaulichte Entwicklung der Werte des Variationskoeffizienten über die beiden Runden des Gruppendelphis, so zeigt sich am Mittelwert der Rundenwerte des Variationskoeffizienten ($\bar{x}(v)$), dass die Summe sich hin zu einer konsensualeren Beurteilung der *einzelnen Handlungsbedarfe bei verschiedenen Eintragsquellen* entwickelt hat. Einzig bei *Item 46* ist ein Anstieg des Wertes des Variationskoeffizienten zu verzeichnen. Die unter der Dimension der Evidenz⁶⁴ beschriebenen Argumentationen der Stakeholder zeigten für diese Eintragsquelle, dass alle Stakeholder einen hohen Handlungsbedarf hinsichtlich der *Wirkstoffproduktionsstätten* sahen, allerdings ein Konsens über einen Dissens darüber bestand, inwiefern dies auch für die Produktion von Fertigarzneimitteln der Fall sei.

Abbildung 21: Werte des Variationskoeffizienten der Items des Abschnitts IV und dem Mittelwert des Variationskoeffizienten in SAUBER+



Anmerkungen: tn1=3; tn2=3

Quelle: Eigene Darstellung

Während der zweiten Runde konnten überdies drei *Eintragsquellen* identifiziert werden, bei denen die Teilnehmenden geschlossen einen hohen Handlungsbedarf sahen und bei denen der Variationskoeffizient jeweils einen Wert von $v=0,1$ (t2) betrug:

- **Item 48:** *Spezielle Einrichtungen des Gesundheitswesens z.B. Ärzthäuser, psychiatrische Kliniken, Sportkliniken etc.*
- **Item 49:** *Pflegeeinrichtungen und Betreuungsheime, Seniorenresidenzen, Hospize etc.*
- **Item 50:** *Privathaushalte.*

⁶⁴ siehe Kapitel 7.1.3.1.

So konnten letztlich vor gleichem Bewertungshorizont drei Eintragsquellen für weitergehende Maßnahmen identifiziert werden.

Auch in der Dimension zur Verständlichkeit findet sich ein Beispiel für eine Aussage, die trotz unterschiedlicher Bewertung in der ersten Kleingruppenphase nicht mehr in der zweiten Runde abgefragt wurde, da ein Konsens zwischen den Teilnehmenden während des Plenum erzielt werden konnte. Das *Item 36* wurde im Fragebogen der ersten Runde wie folgt abgefragt: *Alle Wasserwerke, die ihr Trinkwasser aus Uferfiltrat, Flusswasser oder angereichertem Grundwasser gewinnen, sollten erweiterte Reinigungstechniken wie beispielsweise die Technik der Aktivkohlefiltration einsetzen.* Diese Aussage wurde von Gruppe 1 und Gruppe 2 jeweils abgelehnt (3), während Gruppe 3 dieser Aussage *voll und ganz* zustimmte (10), der Variationskoeffizient betrug $v=0,6$ (t1). Die ablehnende Haltung wurde damit begründet, dass es als nicht sinnvoll angesehen wurde, einen reinen *end-of-pipe-Ansatz* zu verfolgen und somit erst am Punkt der Trinkwassergewinnung anzusetzen. Es müsse konsequent von der Quelle her gedacht werden, dann würden Trinkwasserwerke nur in *Ausnahmefällen* die Aktivkohlefiltration hinzuschalten, da die Technik überdies zu zusätzlichem Energieverbrauch führe. Hier sei ein *bedarfsorientierter* Ansatz zu wählen, eine generelle Anforderung sei der falsche Ansatz. Die Gruppe der Stakeholder einigte sich darauf, die Maßnahmenoption wie folgt umzuformulieren: *Wasserwerke, die ihr Trinkwasser aus Uferfiltrat, Flusswasser oder angereichertem Grundwasser gewinnen, sollten bedarfsorientiert erweiterte Reinigungstechniken wie beispielsweise die Technik der Aktivkohlefiltration einsetzen.* Ein zusätzlicher Einsatz erweiterter Reinigungstechniken durch Wasserwerke war nicht die Handlungsoption, der die Stakeholder die höchste Priorität zusprachen. Da diese Einschätzung mit der vorgenommenen Umformulierung bezüglich dieses Items aufgrund des Austausches im Plenum konsensual zustimmend war, wurde die spezifizierte Maßnahme in der zweiten Runde des Gruppendelphis nicht mehr abgefragt.

7.1.3.3 Dimension: normative Angemessenheit

Der Fokus der Analyse hinsichtlich der normativen Angemessenheit und der damit verbundenen Möglichkeit der Offenlegung von normativen Einstellungen und Wertvorstellungen der Teilnehmenden, richtet sich auf die von den Teilnehmenden geäußerten Begründungen während der Plenarphasen. Für die qualitative Inhaltsanalyse werden vier Kategorien angewandt, die die von den Teilnehmenden getätigten *Sollensaussagen* strukturieren: (1) Werturteile, (2) Präferenzen, (3) (individuelle) Bewertungen von Optionen und (4) moralische Urteile. Für diese Analyse stehen beim Gruppendelphi in SAUBER+ zum einen ein Protokoll der ersten und der zweiten Plenarsitzung sowie das Ergebnisprotokoll des Gruppendelphis zur Verfügung.

Nach Auswertung der Dokumente⁶⁵ finden sich bezüglich der Aussage von *Item 12*⁶⁶ in den drei Protokollen Argumente, die die normative Dimension adressieren. So lautete die Aussage: *Bei der Zulassung von Humanarzneimitteln sollte das Umweltrisiko in der Nutzen-Risiko-Analyse des Arzneimittels [mit Versagungsgrund] berücksichtigt werden.* Hier wurde in der ersten Plenarsitzung, die bereits im vorhergehenden Kapitel beschriebene Umformulierung vorgenommen und die Ergänzung *mit Versagungsgrund* eingefügt. Die Aufnahme der Aussage in

⁶⁵ Eine Aufstellung der in die Auswertung aufgenommenen Aussagen findet sich in Anhang 11.4.

⁶⁶ Das Item wurde zusätzlich im letzten Abschnitt des Fragebogens abermals abgefragt. Hier wurde die identische Formulierung in ihrer Beurteilung in Wirksamkeit (53) und Umsetzbarkeit (54) unterteilt.

den Fragebogen der zweiten Runde erfolgte zudem, da ein Stakeholder die Relevanz dieser Aussage mit der Präferenz betonte, dass sich bezüglich des abgefragten Sachverhaltes etwas ändern solle. Nach der zweiten Kleingruppenphase wurde die Aussage geschlossen abgelehnt und die Begründung für diese teilweise veränderte Einschätzung der Teilnehmenden in der Plenarsitzung ausgetauscht. Die Ablehnung einer Kleingruppe stützte sich vorwiegend auf das Werturteil, dass *das Wohl des Patienten immer den Umweltaspekten vorgelagert* sein müsse, damit der Leitsatz gelten kann, dass bei *der Arzneimittelzulassung Gesundheitsaspekte Priorität vor Umweltaspekten haben müssen*. Ein Vorschlag zur Berücksichtigung dieses Leitsatzes und möglicher zeitversetzter Integration des Umweltaspektes war, *neue umweltschädliche Arzneimittel zwar zunächst zuzulassen sie aber dann wieder vom Markt zu nehmen, wenn umweltverträglichere Alternativen zur Verfügung stünden*. Gegen dieses Argument wurde ein moralisches Urteil entgegengestellt, dass die ethische Verantwortung gegenüber den Patient*innen betonte, wie beispielsweise Personen mit dem Krankheitsbild der Epilepsie. Bei diesen Personen stelle die Einstellung auf ein bestimmtes Medikament einen langwierigen Prozess dar, so *dass eine Umstellung auf ein alternatives Medikament nicht immer problemlos erfolgen könne*. Der Versagungsgrund könne damit nicht das Bemessungskriterium der Umwelt sein, wenn ein Medikament sinnvoll für eine Patient*innengruppe sei. Dieses moralische Prinzip wurde von den Stakeholdern geschlossen angenommen und befürwortet, so dass die abgefragte Maßnahme folglich geschlossen abgelehnt wurde. Die in der Diskussion zu diesem Item geäußerten Argumente adressierten sowohl Aspekte des Gemeinwohls, als auch der normativen Richtigkeit. Der Lösungsvorschlag adressierte direkt das geäußerte Werturteil. Auch die auf moralischen Prinzipien beruhende Erwiderung wurde von den Diskursteilnehmenden als legitim erachtet, so dass die Aussage nicht nur geschlossen, sondern aus den identischen Gründen geschlossen abgelehnt wurde.

Der Austausch in den Plenarsitzungen über die Aussage des *Items 15* wurde bereits im vorhergehenden Kapitel dargestellt, soll an dieser Stelle allerdings nochmals im Hinblick auf die normative Angemessenheit der Argumente betrachtet werden. Die Aussage im Fragebogen (t1) lautete: *Es sollte eine auf Umweltverträglichkeit bezogene Produktkennzeichnung von Medikamenten als Entwicklungsanreiz für eine ‚grüne Produktpolitik‘ der Arzneimittelhersteller eingeführt werden*. Das Item wurde letztlich von den Stakeholdern bezogen auf *große Pharmakonzerne* konsensual abgelehnt. In Gruppe 3 kam es allerdings in der zweiten Runde zu einer höheren Bewertung (5) dieser Aussage als in den anderen Kleingruppen. Diese höhere Bewertung stellte sich in der folgenden Plenarsitzung als ein Kompromiss zwischen den Teilnehmenden dieser Kleingruppe dar, da eine Person die ablehnende Einschätzung für Pharmakonzerne teilte, allerdings im Gegensatz zu den restlichen Stakeholdern als individuelle Bewertung eine *grüne Produktkennzeichnung* für kleinere Unternehmen und Start-Ups als gutes Marketinginstrument einstufte. Diese Person führte für diese Position zusätzlich an, dass das Auffinden neuer Wirksamkeiten einen *erstrebenswerten Wert* darstelle und die Produktion gefördert werden solle. Diesem Argument der normativen Richtigkeit des Auffindens neuer Wirksamkeiten widersprachen die Teilnehmenden nicht, legten ihrer Einschätzung allerdings das Argument zu Grunde, dass die Arzneimittelentwicklung und die Wirkstoffentwicklung in diesem frühen Entwicklungsstadium völlig anderen Prioritäten folgen würden. Dabei bezogen sich die beiden argumentierenden Stakeholder auf ihre Erfahrungen aus der Alltagspraxis in Pharmakonzerne. Die jeweiligen Argumente wurden zwar gegenseitig als legitim erachtet,

allerdings betonten beide Positionen ihren jeweiligen Geltungsanspruch, so dass ein Konsens über einen Dissens bei der Beurteilung dieses Sachverhaltes bestehen blieb.

Auch bei der in *Item 21* abgefragten Aussage fand sich am Ende des Gruppendelphis eine Person, die ein Minderheitenvotum vertrat und bei dieser Einschätzung blieb: *Ein bundesweit einheitlicher und verbindlicher Standard für die Entsorgung unverbrauchter Arzneimittel über die Apotheken, bei dem die nicht frei verkäuflichen Medikamente einem Pfandsystem unterliegen, sollte erprobt werden (der Arzneimittelnutzer erhält bei Rückgabe des angebrochenen Medikaments oder der Verpackung einen monetären Wert zurück, den er beim Kauf des Arzneimittels mit entrichtet hat).* Während der ersten Plenarsitzung wurde bereits ein moralisches Urteil vorgetragen, das eine Gefahr für das *Patientenwohl* in der Aussage erkannte. So wurde die Ablehnung mit der Befürchtung begründet, dass ein solches System für die Patient*innen *einen Anreiz schaffen könnte, die Medikamenteneinnahme vorzeitig abzubrechen.* Hier lag allerdings ein Verständnisproblem in der Interpretation des Frageinhalts vor, so dass sich die Teilnehmenden der Kleingruppe von dem Argument überzeugen ließen, dass es auch für die leere Verpackung Pfand zurückgabe und dies auch so in der Frage enthalten sei. Während der Diskussion im zweiten Plenum wurde von einem Teilnehmenden, als individuelle Bewertung der Option eines pfandgestützten Rückgabesystems, als *verbraucherunfreundlich* bewertet. Diese ablehnende Haltung basierte auf der Annahme, dass sich die Verbraucher*innen nach einer *Kosten/Aufwand-Nutzen-Kalkulation* gegen eine Rückgabe und gegen das System entscheiden würden. Die Person, die die Aussage befürwortete argumentierte mittels Werturteil, dass Patient*innen bewusstgemacht werden müsse, *dass Medikamente ein besonderes Produkt* seien. Ein geeignetes didaktisches Mittel dies zu vermitteln wäre deshalb eine Rückgabe über die Apotheken: *Kommt aus der Apotheke raus, muss in die Apotheke rein.* Das Argument, dass Medikamente *ein besonderes Produkt* seien, teilten die restlichen Stakeholder als legitime Einschätzung, allerdings teilten sie nicht den didaktischen Weg, dies an die Verbraucher*innen über ein an Apotheken gebundenes Pfandsystem zu vermitteln. Es bleibt letztlich ein Konsens über einen Dissens zwischen *Verbraucherfreundlichkeit* und *Bewusstmachung des besonderen Gutes* bestehen.

Bei *Item 39* konnten sich beide Gruppen letztlich, wie im vorhergehenden Kapitel 7.1.3.2 beschrieben, auf eine Umformulierung der Aussage einigen: *Die novellierte deutsche Klärschlammverordnung sollte ein generelles Verbot des Aufbringens von Klärschlamm auf landwirtschaftlich genutzten Böden umfassen.* Der bei dieser Aussage vorherrschende Konsens über einen Dissens blieb allerdings auch nach dem zweiten Plenum bestehen. In dieser zweiten Plenarsitzung wurde deutlich, dass die jeweiligen Positionen auf verschiedenen Werturteilen gründeten. Während ein Teil der Teilnehmenden ein generelles Verbot des Düngens mittels Klärschlamm für sinnvoll erachtete, da der *Schutz des Grundwassers* als oberstes Gut vorzuziehen sei, sah der andere Teil der Teilnehmenden das Potential einer technischen Nutzung der Nährstoffe im Klärschlamm. Diese Option wurde aus dieser Sicht als technisch machbar und gefahrlos eingeschätzt.

Ein weiteres in der zweiten Plenarsitzung als legitim erachtetes Argument, findet sich im Werturteil, dass es bei der Frage der Entsorgung vorrangig um Umweltschutz und nicht um Gesundheitsschutz gehe und dies exakt bei einer Aufklärungskampagne kommuniziert werden

müsse. Dies bezog sich auf den Inhalt, der in *Item 59* adressierten Aussage, dass eine *bundesweite Aufklärungskampagne zur sachgerechten Entsorgung unverbrauchter Arzneimittel durchgeführt werden sollte*.

Die Stakeholder konnten im Rahmen des Gruppendelphis in mehreren Fällen individuelle Präferenzen zum Ausdruck bringen, wie die Urteile zu einzelnen Maßnahmen gefällt werden sollten. So wurde während der ersten Plenarsitzung für die Verteilung der Budgets in Abschnitt I und V des Fragebogens die Präferenz der Teilnehmenden geäußert, keine *Gießkannen-Taktik* zu verfolgen. So sollte unter dem Leitsatz *nicht Kleckern, sondern Klotzen* eine klare Verteilung der Budgets vorgenommen werden, um die aus Sicht der Teilnehmenden sinnvollen Maßnahmen klar zu markieren. Dieser Argumentation folgten die Teilnehmenden zu zweiten Runde geschlossen.

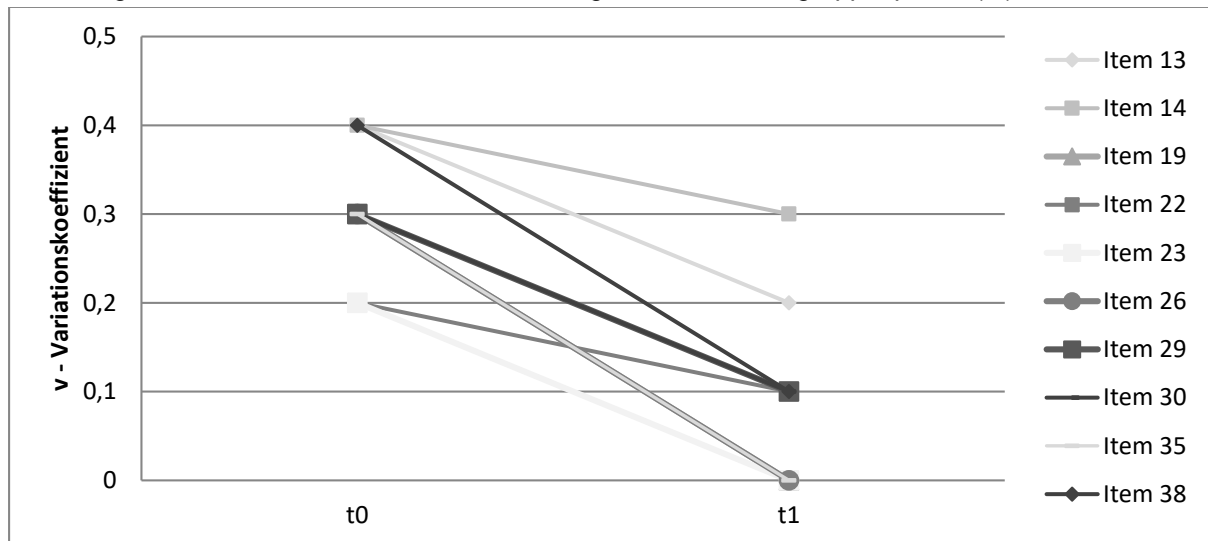
So finden sich in dieser Dimension Beispiele von Aussagen, die entweder den Geltungsanspruch des Gemeinwohls, als auch den Geltungsanspruch an normative Richtigkeit in Anspruch nehmen. Teilweise wurden diese Argumente als legitim erachtet und von den Teilnehmenden in die weitere Urteilsfindung mit einbezogen, teilweise führten diese auch zu einer unterschiedlichen Bewertung des Sachverhaltes, der allerdings auf Basis der verschiedenen Geltungsansprüche als Konsens über einen Dissens markiert werden konnte.

7.1.3.4 Dimension: grundlegende Funktionslogik

Das Gruppendelphi in SAUBER+ muss sich ebenso daran bemessen lassen, ob Maßnahmen identifiziert werden konnten, die aus Sicht der Teilnehmenden zur Verringerung der Belastung von Gewässern mit Humanarzneimitteln beitragen können, also zu einer Lösung der grundlegenden Fragestellung des Gruppendelphis beitragen. Dafür standen den Urteilen der Teilnehmenden eine Reihe von prominent diskutierten und weiteren Maßnahmen zur Verringerung der Gewässerbelastung zur Verfügung. Das Gruppendelphi sollte dabei Bewertungen durch die Stakeholder hervorbringen, die einen klaren Konsens oder einen Konsens über einen Dissens markieren. Die jeweiligen Urteile sollten auch nach dem Gruppendelphi stabil bleiben.

Betrachtet man die bereits in Kapitel 7.1.3.1 diskutierte Entwicklung der statistischen Kenngrößen, so zeigt sich bei einer Vielzahl der Items ein Anstieg des Wertes des Variationskoeffizienten zwischen der Vorbefragung (t0) und der ersten Kleingruppenbewertung (t1) am eigentlichen Workshoptag. Diese Items sind zur besseren Nachvollziehbarkeit in Abbildung 22 rot dargestellt. Auch wenn diese einzelnen Items bereits in der vorhergehenden Analyse näher besprochen wurden, so soll an dieser Stelle nochmals auf diese atypische Entwicklung eingegangen werden, da diese statistischen Kennzahlen der Grundannahme eines Delphis widersprechen, dass sich die Expert*innen in ihrem Urteil über die Runden hinweg immer weiter annähern, besonders, wenn man den Umstand betrachtet, dass die Teilnehmenden den Fragebogen der Vorbefragung individuell ausgefüllt haben.

Abbildung 23: Items mit konsensualer Entwicklung zur ersten Kleingruppenphase (t1) in SAUBER+



Anmerkungen: tn0=13; tn1=3;

Quelle: Eigene Darstellung

Diese Items wurden im Plenum (t1) und der weiteren Runde des Delphis (t2), wie in der methodischen Konzeption vorgesehen, nicht weiter diskutiert und abgefragt (t2), so dass sie in der Analyse der vorherigen Dimensionen nicht tiefergehend analysiert wurden.

Nach dem Gruppendelphi bereitete das Forscherteam die Ergebnisse in einem 18 Seiten umfassenden Protokoll auf, das sowohl die statistischen Ergebnisse der jeweiligen Runden enthielt, wie auch die Entwicklung der einzelnen Urteile über die Runden des Workshops (t1/t2) hinweg. Neben diesen wurden auch die geäußerten Argumente der Stakeholder während der Plenarsitzung in dem Protokoll aufgenommen. Das Protokoll wurde allen Teilnehmenden zwei Monate nach dem Workshop mit der Bitte um Feedback zugesandt. Es gab von Seiten der Teilnehmenden weder Widersprüche noch inhaltliche Korrekturvorschläge, was dafür spricht, dass während des Workshops kein Scheinkonsens vorlag.

Letztlich muss sich auch das Gruppendelphi in SAUBER+ daran messen lassen, ob es ermöglicht hat *fruchtbare* Ergebnisse auszuarbeiten, die von allen Teilnehmenden konsensual getragen werden. Dies ist in diesem Fallbeispiel gelungen, denn es konnten in jedem der drei adressierten Handlungsfelder (*Arzneimittelentwicklung, Umgang mit Arzneimitteln* und *Emissionsmanagement*) Maßnahmen identifiziert werden, die konsensual hohe Zustimmungswerte erhielten, dies sowohl nach der ersten Runde des Gruppendelphis als auch nach Austausch im Plenum in der darauffolgenden Runde des Gruppendelphis. Feedback und Wiederholung haben damit dazu geführt, dass neue Informationen *aufgenommen* und von den Teilnehmenden *angenommen* wurden und so eindeutigere Bewertungen ausgearbeitet werden konnten.

7.1.4 Zwischenfazit: Gruppendelphi in SAUBER+

Das Gruppendelphi in SAUBER+ hat sein mit dem Einsatz der Methode verbundenes Ziel letztlich erreicht. So wurden von relevanten Stakeholdern Bewertungen zu *Maßnahmen zur Verringerung der Gewässerbelastung mit Humanarzneimittelwirkstoffen* erhoben und gemeinsam im diskursiven Austausch strittige Bewertungen zu einem Konsens oder einem Konsens über einen Dissens herangeführt. Auch konnte der *Handlungsbedarf bei Humanarzneimitteln*

in Vergleich zum Handlungsbedarf bei anderen Kontaminanten, die zu einer Belastung von Grundwasser und von Oberflächengewässern führen, in Relation gesetzt sowie der aktuelle Forschungsbedarf im Bereich Arzneimittelrückstände im Gewässerkreislauf erhoben werden.

Die Entwicklung zu einem Konsens in den Bewertungen der jeweiligen Maßnahmen zeigt ferner auf, dass die Qualitätsmerkmale eines handlungsleitenden Diskurses eingelöst werden konnten. So konnte zwischen der ersten Runde des Gruppendelphis zur zweiten Runde bei zehn Items⁶⁷ anhand der statistischen Maßzahlen eine konsensuale Entwicklung veranschaulicht werden. Auch konnte beispielsweise im ersten Abschnitt zum *Handlungsbedarf bei verschiedenen Kontaminanten* zur zweiten Runde eine konsensuale Bewertung erzielt werden, indem die Teilnehmenden im ersten Plenum des Gruppendelphis die Möglichkeit hatten, eine Verständigung über die, den Urteilen zugrundeliegenden, Relevanzsysteme zu erzielen. Dass die Gültigkeit der einzelnen Argumente in Frage gestellt und mittels eines diskursiven Austauschs zwischen den Teilnehmenden überprüft werden konnten, zeigt sich bereits in den dargestellten Inhalten der Diskussionen während den Plenarsitzungen. So konnte beispielsweise im Austausch über die in *Item 12* diskutierte Maßnahme die Präzisierung zwischen *generellen Berücksichtigung des Umweltrisikos* und einem *Umweltrisiko als Versagungsgrund* eingeführt werden, die einem Teil der Stakeholder bis zu diesem Zeitpunkt nicht bewusst war. Nach dem Wissensaustausch zu dieser Bewertungsgrundlage und zusätzlicher Integration des Zusatzes *mit Versagungsgrund* in der Frageformulierung führten zu einer geschlossenen Ablehnung durch die Gruppe der Stakeholder und damit zu einem konsensualen Ergebnis. Dies verdeutlicht überdies, dass einer Präzisierung der Formulierung von Maßnahmen bei diesem Gruppendelphi ein Rahmen geboten wurde.

Wie wichtig auch die semantische Klärung ist, zeigt sich in diesem Fallbeispiel nicht nur im Hinblick auf Begriffe und Definitionen, sondern auch in der Interpretation der im Fragebogen formulierten Aussagen. So kam es, wie in Abschnitt 7.1.3.2 beschrieben, beim sechsten Frageabschnitt über die *Dringlichkeit des Handlungsbedarfes nach Eintragsquellen* in der ersten Runde dazu, dass die verschiedenen Gruppen ihren Urteilen verschiedene geografische Reichweiten zu Grunde legten. Da das Gesamtprojekt SAUBER+ sich aufgrund der Förderung durch das BMBF auf bundesdeutsche Maßnahmen fokussierte, übertrug das Team diese Reichweite ohne weitere Reflexion auch auf den Fragebogen, ohne diesen Fokus an die Teilnehmenden zu vermitteln. Anders war es bei der Interpretation der Maßnahmen als technische Maßnahmen, die in einer Kleingruppe stattfand. Hier wurde der im Fragebogen eingefügte Hinweis überlesen: „*Handlungsbedarf*“ kann sich auf alle der oben angeführten Maßnahmen einschließlich Forschungsförderung und weitere mögliche Maßnahmen beziehen. So konnte sich mit dem Austausch im ersten Plenum auf einen gemeinsamen Bewertungshorizont verständigt werden, der durch die in den Fragebogen eingefügte Ergänzung *Deutschland, kein Technikfokus* zusätzlich als Gedankenstütze festgehalten wurde.

Wie die Dimension der Evidenz ebenso eine Umformulierung der Frage bedingen kann, zeigt sich beispielsweise bei *Item 33*, in dem in der ursprünglichen Formulierung eine *generelle* Vorbehandlung der Abwässer bei Krankenhäusern und Spezialkliniken vorgeschlagen wurde. Die Stakeholder begründeten die Ablehnung mit dem Argument, dass es aus ökonomischer Sicht nicht sinnvoll sei, eine solche Vorbehandlung generell durchzuführen. Nur da, wo hohe

⁶⁷ Items 12, 15, 31, 32, 33, 47, 48, 49, 50, 51.

Frachten an problematischen Stoffen anfielen, mache solch ein Vorgehen Sinn - also bei Bedarf. Dies führte dazu, dass *generell* durch *bedarfsorientiert* ausgetauscht wurde.

In diesem Fallbeispiel zeigt sich die Annahme bestätigt, dass während eines Gruppendelphis ein diskursiver Rahmen besteht, in dem die Teilnehmenden ihre normativen Einstellungen und Wertvorstellungen offenlegen können. So wurden von den Stakeholdern Argumente angeführt, die beispielsweise die Wertvorstellung vertraten, dass das Wohl der Patient*innen immer dem Umweltschutz vorgezogen werden sollte. Dieses normative Argument wurde dabei nicht verdeckt, von den anderen Teilnehmenden als legitim aufgenommen und überdies in die weitere Urteilsfindung mit einbezogen.

Mit Blick auf die Merkmale eines Gruppendelphis zeigt sich in diesem Fall das bereits in Kapitel 4.1.3 geschilderte Zeitproblem, so dass das im Ideal angenommene Abbruchkriterium bei diesem Fallbeispiel nicht eingehalten werden konnte und der Faktor Zeit das natürliche Ende des Workshops bestimmte. So wäre es sinnvoll gewesen, beim fünften Abschnitt über die *Forschungsförderung in Bezug auf Arzneimittelrückstände im Gewässerkreislauf* eine weitere Kleingruppenphase durchzuführen, da die Ergebnisse insbesondere bei dieser Budgetfrage nach der zweiten Plenardiskussion weiter geschärft hätten werden können. Im Fragebogen des Gruppendelphis gab es überdies einen siebten Frageblock, in dem die zwei jeweils in den Handlungsfeldern bestbewerteten Maßnahmen⁶⁸ der Vorbefragung mit der Differenzierung auf ihre *Wirksamkeit* und *Umsetzbarkeit* ebenfalls 10er-skaliert abgefragt wurden. Diese Ergebnisse konnten aber sowohl in der ersten Plenarsitzung sowie in der zweiten Plenarsitzung nicht tiefergehend diskutiert werden. Die Zeit reichte hier nicht für einen Austausch der Argumente zu den jeweiligen Urteilen, so dass diese Ergebnisse nicht analysiert werden konnten. Diese Ergebnisse erfüllen damit nicht den, mit der Funktionsweise eines Gruppendelphis erhobenen, qualitativen Anspruch für eine Analyse, die auch die inhaltlichen Begründungen miteinschließt. So konnten einzelne Inhalte und Argumente, die in der zweiten Plenarsitzung zu den im Wortlaut identischen Aussagen unter der Differenzierung in *Wirksamkeit* und *Umsetzbarkeit* in die Gesamtanalyse der ursprünglichen Items, teilweise integriert werden.

Am Ende des Gruppendelphis wurde im Plenum von den Teilnehmenden abgefragt, wie sie die Methode wahrgenommen haben und wie sie diese beurteilen. Die Einschätzungen der Stakeholder waren dabei nicht geschlossen. Einige stellten heraus, dass sie ihre Urteile sehr pragmatisch gefällt hätten, da das Wissen bei einzelnen Sachverhalten doch eher als oberflächlich einzustufen sei. Andere fanden den Aspekt des Pragmatismus jedoch auch von Vorteil, da man von anderen Teilnehmenden viel Input erhalte. Gerade dadurch bestünde die schnellere Bereitschaft seine eigene Meinung in den Kleingruppendiskussionen umzustellen und auf diesem Wege wäre eine Konsensbildung in der Diskussion auch möglich. Ein Stakeholder war allerdings der Meinung, dass eine Runde ausreichend gewesen und eine längere Diskussionsphase mit umfassendem Protokoll besser gewesen wäre. Ebenso wurde die Frage geäußert, ob möglicherweise andere Urteile getroffen worden wären, wenn die Teilnehmenden einen anderen inhaltlichen Hintergrund gehabt hätten.

⁶⁸ Aus dem Handlungsfeld *Arzneimittelforschung*: Item 12, Item 13; Aus dem Handlungsfeld *Umgang mit Arzneimitteln*: Item 20, Item 23; Aus dem Handlungsfeld *Emissionsmanagement*: Item 34, Item 35.

Das Feedback der Teilnehmenden zeigt somit nochmals drei Merkmale auf, die sich besonders auf die Methode des Gruppendelphis beziehen:

1. Die Möglichkeit das eigene Urteil aufgrund zusätzlicher Argumente anderer Teilnehmenden zu ändern, wird als gegeben angesehen.
2. Ein Gruppendelphi bietet die Möglichkeit Sachverhalte, bei denen zusätzlicher Forschungsbedarf besteht, auszuarbeiten. Allerdings wird dies von den Teilnehmenden nicht geschlossen als befriedigende Erkenntnis dieses Prozesses wahrgenommen.
3. Bei der Rekrutierung und Zusammensetzung muss besonders darauf geachtet werden, im Hinblick auf die abgefragten Sachverhalte alle möglichen Positionen und Disziplinen einzubinden.

7.2 SEAWIND

Projekt: SEAWIND – *Sound exposure and risk assessment of wireless network devices*

Institut: DIALOGIK gemeinnützige Gesellschaft für Kommunikations- und Kooperationsforschung mbH

Dieses Gruppendelphi fand im Rahmen des 7. EU-Forschungsrahmenprogramms geförderten Vorhabens SEAWIND statt, das von der *Foundation for Research on Information Technologies in Society (IT²IS)* unter Leitung von Prof. Niels Kuster koordiniert wurde. Ziel des Projektes war die interdisziplinäre Erforschung der möglichen gesundheitsschädigenden Auswirkungen, die bei der Nutzung drahtloser Kommunikationstechnologien, wie beispielsweise UMTS, LTE oder WiFi, bestehen könnten. So wurden mittels Einsatzes von Technologien und Messverfahren die Strahlenexpositionen ermittelt, die bei der Verwendung von solchen Kommunikationstechnologien entstehen. Ebenfalls wurden die durch eine solche externe Exposition im Körper entstehenden induzierten Felder bestimmt, um eine Beurteilung der biologischen Relevanz vornehmen zu können. Erweiternd, wurden computerbasierte Modelle berechnet, da die induzierten Felder stark von Anatomie und Körperhaltung abhängen können. Um einen Umgang mit den gesellschaftlich wahrgenommenen Risiken und Unsicherheiten zu erforschen, wurde neben Fokusgruppen, hinsichtlich der sozialwissenschaftlichen Fragestellungen das hier untersuchte Gruppendelphi als Methode eingesetzt (vgl. Hilpert et al. 2013: 1).

7.2.1 Fragestellung des Gruppendelphis

Das mit dem Einsatz des Gruppendelphis verbundene Ziel war es, Kommunikationsstrategien zu entwickeln, die es möglich machen, die Klassifizierung von Strahlungsrisiken basierend auf der 2b-Klassifizierung der *International Agency of Research on Cancer (IARC)*⁶⁹ an die Bevölkerung zu vermitteln. „Das Gruppendelphi wurde in diesem Zusammenhang durchgeführt, um Bewertungen zentraler Elemente in der Kommunikation von Unsicherheiten an Nicht-Experten zu generieren“ (Hilpert et al. 2013: 9).

⁶⁹ International Agency of Research on Cancer (IARC) (2012): List of Classifications. Online: <http://monographs.iarc.fr/ENG/Classification/index.php>, zugegriffen am: 19.06.2017.

7.2.2 Merkmale des Gruppendelphis

Das Gruppendelphi fand am 24. September 2014 im GENO-Haus in Stuttgart statt. Da in diesem Projekt hinsichtlich der abgefragten Sachverhalte nicht auf vorhergehende Analysen zurückgegriffen werden konnte, entwickelte das Team einen ersten Fragebogen für eine schriftliche Vorbefragung, um die ersten Einschätzungen der Expert*innen abzufragen, auf deren Ergebnissen der Fragebogen für das Gruppendelphi erstellt wurde. Um eine hohe Fallzahl zu generieren und damit so viele Bewertungen wie möglich zu erhalten, wurde die Vorbefragung nicht nur den bereits rekrutierten Expert*innen zugesandt, sondern auch jeder Person, die bereits ihr Interesse an dem Thema bzw. dem Gruppendelphi bekundet hatte (vgl. Hilpert et al. 2013: 10).

Die Teilnehmenden wurden anhand ihres Tätigkeitsfeldes rekrutiert, deren professionelle Handlungsfelder eine Schnittmenge mit der Kommunikation von Unsicherheit bezüglich Strahlungsquellen aufwiesen, so dass Personen mit Wissen über elektromagnetische Verträglichkeit, Technikfolgenabschätzung, Strahlenschutz, Risikoforschung, elektromagnetische Felder und Mikrowellen, öffentlichen Gesundheit, Informationstechnologie und Gesellschaft, medizinische Strahlenbiologie und mobile Kommunikation eingeladen wurden. Von den angeschriebenen 130 Expert*innen nahmen letztlich 14 Personen am Workshop teil, zugesagt hatten 18 Personen (vgl. Hilpert et al. 2013: 9). Die Teilnehmenden wurden in drei Kleingruppen per Zufallsverteilung mit vier bis fünf Teilnehmenden pro Kleingruppe aufgeteilt. In der zweiten Runde erfolgte eine permutierte Gruppenzusammensetzung.

Der Fragebogen der Vorbefragung im SEAWIND-Gruppendelphi untergliederte sich in sechs Abschnitte. In diesen Abschnitten wurden in Summe 54 Ratingfragen mittels 10er-skaliertes Items abgefragt. Diese Itembatterien wurden jeweils um eine offene Antwortkategorie zur potentiellen Nennung *Sonstige* ergänzt. Neben diesem Fragetypus wurden zusätzlich fünf offene Fragen sowie zwei Rankingfragen zur Auswahl der wichtigsten Aspekte abgefragt.⁷⁰ Der im Workshop eingesetzte Fragebogen gliederte sich in fünf Abschnitte. Insgesamt wurden 47 Ratingfragen mittels 10er-skaliertes Items abgefragt, deren Itembatterien, wie auch in der Vorbefragung durch eine offene Antwortkategorie zur potentiellen Nennung *Sonstiger* ergänzt wurden. Ebenso fanden sich auch in diesem Fragebogen zwei Rankingfragen zur Auswahl der wichtigsten Aspekte. Die Fragen wurden dabei in fünf Abschnitte untergliedert: *Kommunikationsakteure* (I), *Kommunikationsinhalte* (II), *Kommunikation von Messdaten* (III), *Konkrete Maßnahmen* (IV) und *Kommunikation von Untersuchungsergebnissen* (V).⁷¹

⁷⁰ Der Fragebogen der Vorbefragung befindet sich im Anhang 11.5.

⁷¹ Der im Gruppendelphi eingesetzte Fragebogen befindet sich im Anhang 11.6.

Während des eintägigen Workshops konnten wie geplant zwei Delphirunden stattfinden. Der geplante Ablauf gliederte sich im Detail wie folgt auf:

09:30 Uhr	Welcome Kaffee
10:00 Uhr	Begrüßung und Vorstellung
10:15 Uhr	Einführung in das Gruppendelphi
10:30 Uhr	Bearbeitung des ersten Fragebogens
12:00 Uhr	Mittagspause
13:00 Uhr	Plenarsitzung: Diskussion der Ergebnisse
14:30 Uhr	Bearbeitung des zweiten Fragebogens
15:45 Uhr	Kaffeepause
16:15 Uhr	Plenarsitzung: Diskussion der Ergebnisse
17:30 Uhr	Fazit und Verabschiedung

Während der Plenarphase dokumentierten zwei Personen aus dem Forschungsteam die jeweiligen Äußerungen der Teilnehmenden. Diese Aufzeichnungen wurden gemeinsam mit den statistischen Ergebnissen als Ergebnisprotokoll aufbereitet. Dieses sechsseitige Ergebnisprotokoll wurde den Teilnehmenden mit der Bitte um Mitteilung etwaiger Korrekturwünsche zugesandt.

7.2.3 Analyse des Gruppendelphis in SEAWIND

Nachdem die Merkmale und die zentrale Fragestellung des Gruppendelphis im Forschungsprojekt SEAWIND dargestellt wurden, findet auch für dieses Beispiel eine Fallanalyse statt, die dem Aufbau des Testdesigns folgend mit der Dimension der Evidenz beginnt.

7.2.3.1 Dimension: Evidenz

Sollten die Qualitätsmerkmale eines handlungsleitenden Diskurses im Falle des Delphis in SEAWIND eingelöst werden, so sollten sich die Werte der jeweiligen Items über den Verlauf des Gruppendelphis in Richtung eines Konsenses bewegen oder einen Konsens über einen Dissens markieren. Den Beginn der Analyse dieser Dimension bilden die Items, die mittels 10er-skalierten Ratingskalen abgefragt wurden.⁷² Mit 47 abgefragten 10er-skalierten Items im Fragebogen der ersten Runde bildet dieser Typ der Ratingskalen auch in SEAWIND den am häufigsten eingesetzten Fragetypus. Im Gegensatz zu SAUBER+ wurden diese Items nochmals komplett in der zweiten Runde des Gruppendelphis abgefragt, so dass im Falle von SEAWIND mehr Items für die Analyse zur Verfügung stehen.

Im Fragebogen des SEAWIND-Gruppendelphis wurden die 10er-skalierten Items nicht mittels formulierter Aussagen abgefragt, sondern je nach Frage unterschiedlich konstruiert. So standen in Frage 1.1 im ersten Abschnitt des Fragebogens *Kommunikationsakteure* zur Bewertung, die nach Meinung der Teilnehmenden die Öffentlichkeit über die Einstufung neuer Risiken unter die *Kategorie 2b des Krebsrisikos durch elektromagnetische Felder (EMF)* im Bewertungskonzept der *International Agency for Research on Cancer (IARC)* informieren sollten. Ein Votum von 1 auf der Skala bedeutet dabei, *dass diese Akteursgruppe die Öffentlichkeit*

⁷² Die Antworten der einzelnen Kleingruppen sowie die statistischen Ergebnisse finden sich in Anhang 11.7.

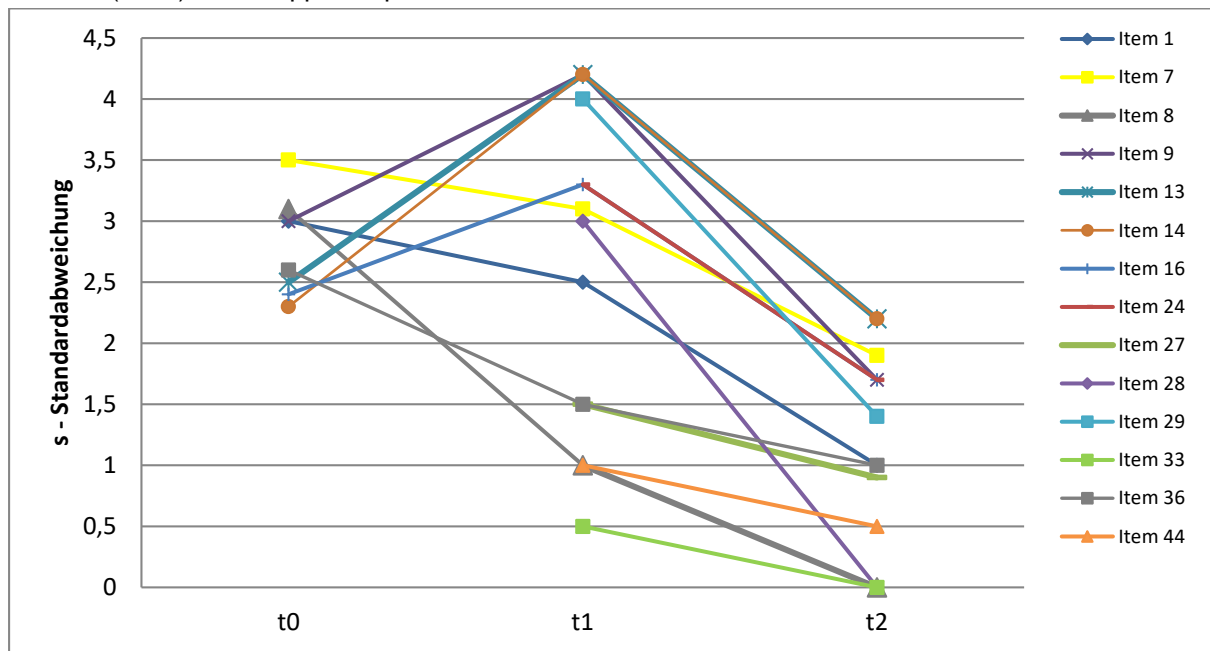
auf keinen Fall über diese Einstufung und ihre Implikationen informieren solle. Eine 10 auf der Skala hingegen, dass diese Akteursgruppe dies auf jeden Fall tun solle. Die Kommunikationsakteure wurden bei dieser Frage aufgelistet, so dass beispielsweise *Journalisten, Wissenschaftler* oder *nationale Organisationen, wie z.B. die Bundeszentrale für gesundheitliche Aufklärung* zur Bewertung standen. In Frage 2.1 des zweiten Abschnitts zu *Kommunikationsinhalten* wurden unterschiedliche Inhalte abgefragt, die im Falle, dass *elektromagnetische Felder „möglicherweise krebserregend“ (Risikogruppe 2b)* sind, kommuniziert werden könnten. Die Wertung mit einer 1 bedeutete dabei, dass der jeweilige Kommunikationsinhalt *ganz und gar nicht* kommuniziert werden solle und die Bewertung mit einer 10 bedeutete, dass dieser Inhalt *unbedingt* kommuniziert werden müsse. Abgefragt wurden hier beispielsweise die *Datengrundlage zur Festlegung der Risikogruppe* und die *Unabhängigkeit der fachlich bewertenden Institution*. In Frage 3.1 wurden im dritten Abschnitt zur *Kommunikation von Messdaten* die *Wege der Übermittlung* solcher Daten ebenfalls mittels einfacher Nennung dieser Wege abgefragt, wobei die *generelle Effektivität der dargestellten Kommunikationsweise* mittels Einschätzung als *ineffektiv* durch eine Wertung von 1 bis zu *sehr effektiv* durch eine Wertung von 10 bewertet wurden. So wurden die Teilnehmenden gebeten, beispielsweise *fachliche Beiträge auf Konferenzen und Workshops* aber auch *Videos und Podcasts auf Youtube und ähnlichen Portalen* durch ihr Votum auf der 10er-Skala zu beurteilen. In Frage 3.2 wurden diese Übermittlungswege dann nochmals im Hinblick auf ihre *Glaubwürdigkeit* abgefragt. Ein Wert von 1 bedeutete dabei, dass die Expert*innen diesen Kommunikationsweg als *überhaupt nicht glaubwürdig* einschätzten, ein Wert von 10, dass sie diesen Kommunikationsweg als *sehr glaubwürdig* bewerteten. Unter dem vierten Abschnitt zu *konkreten Maßnahmen* wurden in Frage 4 *Maßnahmen zur Steigerung der Aufmerksamkeit* abgefragt, die auf das *Problem der potentiellen Strahlung von elektromagnetischen Feldern* aufmerksam machen können. Ein Wert von 1 bedeutete dabei, dass die Teilnehmenden diese Maßnahme als *ganz und gar nicht* geeignet ansahen, ein Wert von 10 hingegen, dass sie diese Maßnahme als *voll und ganz geeignet* einschätzten. Als Maßnahme wurden beispielsweise *mobile Dienstleister* abgefragt, die *Messungen vor Ort zur Überprüfung der Strahlenintensität durchführen*. Somit wurden letztlich bei diesem Gruppendelphi keine formulierten Aussagen abgefragt, sondern entweder Akteure oder Maßnahmen per Nennung zur Beurteilung durch die Teilnehmenden in den Fragebogen integriert.

Verfolgt man im Falle von SEAWIND die der Dimension der Evidenz zugrunde liegende Annahme, dass sich über den Verlauf eines Gruppendelphis die Werte in Richtung eines Konsenses bewegen, so wird ein solcher Verlauf durch die Werte der Standardabweichung von den in Abbildung 24 aufgezeigten 14 Items bestätigt. Dies gilt mit Fokus auf die Entwicklung zwischen der ersten Runde (t1) und der zweiten Runde (t2) des Workshoptages, denn hier finden die diskursiven Elemente eines Gruppendelphis ihre Anwendung. Diese dargestellten Items verteilen sich dabei quer über die oben beschriebenen Fragen, so dass sich immer mindestens ein Item aus dem jeweiligen Frageblock in der Abbildung wiederfindet.

Betrachtet man die Entwicklung zwischen der Vorbefragung (t0) und der Beantwortung des ersten Workshopfragebogens (t1) so ist hier eine gänzlich unterschiedliche Entwicklung des Wertes der Standardabweichung (s) zu beobachten. Während beispielsweise bei den abgefragten *internationalen Organisationen (Item 1)* die Standardabweichung bereits von s=3 auf s=2,5 zwischen diesen beiden Befragungszeitpunkten sinkt, steigt der Wert des Items über die

Methode zur Datenbewertung (Item 14) von $s=2,3$ auf $s=4,2$ an. Trotz dieser unterschiedlichen Entwicklungen zwischen Vorbefragung und erster Kleingruppenphase sinkt der Wert der Standardabweichung bei den in Abbildung 24 dargestellten Items zwischen der ersten Runde (t_1) und der zweiten Runde (t_2). Dies gilt sowohl für das *Item 35* bei dem der Wert von $s=1,0$ auf $s=0,0$ sinkt, als auch bei *Item 14* bei dem der Wert von $s=4,2$ auf $s=2,2$ sinkt. Den größten Abfall des Wertes des Streuungsmaßes zeigt allerdings das Item zu *Videos und Podcasts auf Youtube und ähnlichen Portalen (Item 28)* auf. Bei diesem sinkt der Wert von $s=3,0$ in der ersten Runde des Delphis auf einen Wert von $s=0$ zur zweiten Runde.

Abbildung 24: Vergleich der Werte der Standardabweichung mit konsensualer Entwicklung über die Runden (t_1 ; t_2) des Gruppendelphis in SEAWIND



Anmerkungen: $tn_0=30$; $tn_1=3$; $tn_2=3$

Quelle: Eigene Darstellung

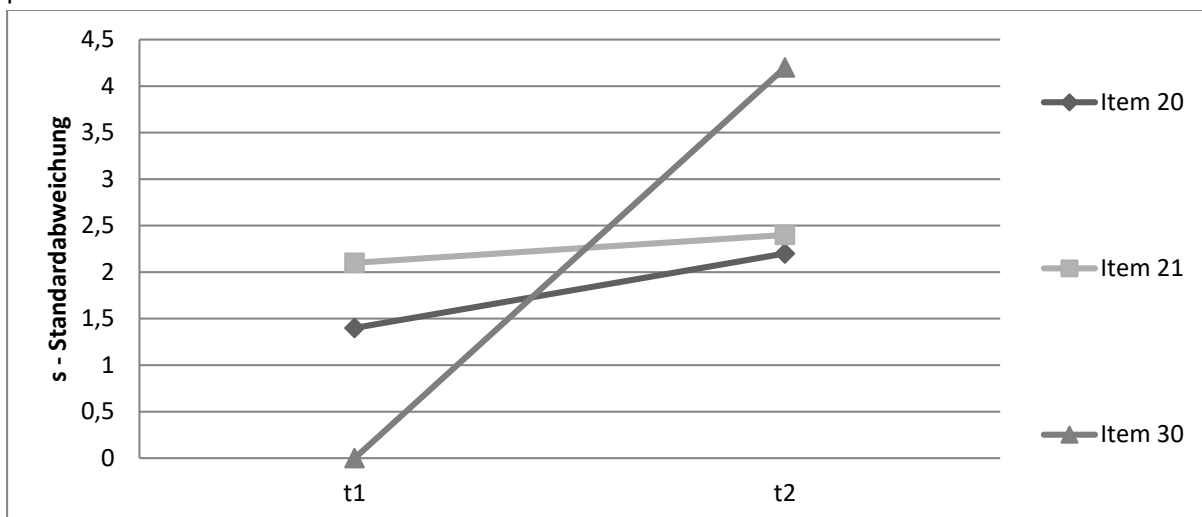
Diese 14 Items zeigen folglich die angenommene Entwicklung hin zu einem konsensualen Ergebnis zwischen den Einschätzungen der Teilnehmenden anhand der Werte der Standardabweichung auf. Das Beispiel SEAWIND zeigt allerdings, dass ein exklusiver Blick auf die Standardabweichung - und folglich auch den Variationskoeffizienten - die Gefahr einer Fehlinterpretation birgt. So muss ebenfalls die Anzahl der Antworten (n im Vergleich zu N) beim jedem Item in die Analyse mit einbezogen werden. Dies gilt nicht nur für mögliche Minderheitsvoten, sondern zeigt sich im Falle von SEAWIND besonders auch im Hinblick auf Enthaltungen.

So finden sich in Abbildung 24 zwei Items, die anhand des Wertes der Standardabweichung als Entwicklung zu einem konsensualen Ergebnis interpretiert werden könnten. Allerdings hat sich bei *Item 36* eine Kleingruppe (Gruppe 1) im Fragebogen der zweiten Runde (t_2) enthalten, so dass nur Urteile von zwei Kleingruppen ($n=2$) im Wert der Standardabweichung von $s=1,0$ abgebildet werden. Da sich hier eine Gruppe enthalten hat, kann damit nicht von einer konsensualen Bewertung aller Teilnehmenden für dieses Item ausgegangen werden. Noch eindeutiger kann dies ebenso für *Item 8* angenommen werden, denn hier hat nur eine Kleingruppe (Gruppe 1) ein Votum abgegeben. Der Wert der Standardabweichung ist bei einer Fallzahl von

$n=1$ folglich zwingend $s=0$, ebenfalls aber nicht als konsensuales Ergebnis zwischen Teilnehmenden des Gruppendelphis zu werten, da sich hier gar zwei Gruppen ihres Urteils hinsichtlich des Items enthalten haben.⁷³ Beide Items (*Item 8*, *Item 36*) wurden aufgrund dieser Enthaltungen in Abbildung 24 bereits bewusst grau markiert, da sie nicht als konsensuale Entwicklung gewertet werden können. So bleiben letztlich zwölf Items, die mittels des Wertes der Standardabweichung die Konsensannahme abbilden.

Bei Betrachtung der einzelnen Gruppenantworten fällt generell auf, dass sich in der ersten Runde eine Kleingruppe auffällig oft enthalten hat. So hat Gruppe 2 bei insgesamt 28 Items kein Votum abgegeben. Dies erschwert die Interpretation der Standardabweichung hinsichtlich eines steigenden Wertes zwischen der ersten und der zweiten Runde am Workshoptag. So kann bei diesen Items bei einer exklusiven Betrachtung der Standardabweichung nicht von einem steigenden Dissens zwischen den Runden ausgegangen werden, wenn in der zweiten Runde eine höhere Fallzahl ($n=3$) vorliegt und folglich keine Enthaltung stattgefunden hat. Ähnlich wie bei Betrachtung der *Items 8* und *36* würde hier ein Fehlschluss vorliegen, nur das dieser im Falle eines steigenden Wertes einer anderen Logik folgt - problematisch sind damit nicht nur eine Enthaltung zum Zeitpunkt t_2 , sondern überdies auch zum Zeitpunkt t_1 .

Abbildung 25: Items mit steigendem Wert der Standardabweichung über die Runden des Gruppendelphis in SEAWIND



Anmerkungen: $tn_0=30$; $tn_1=3$; $tn_2=3$

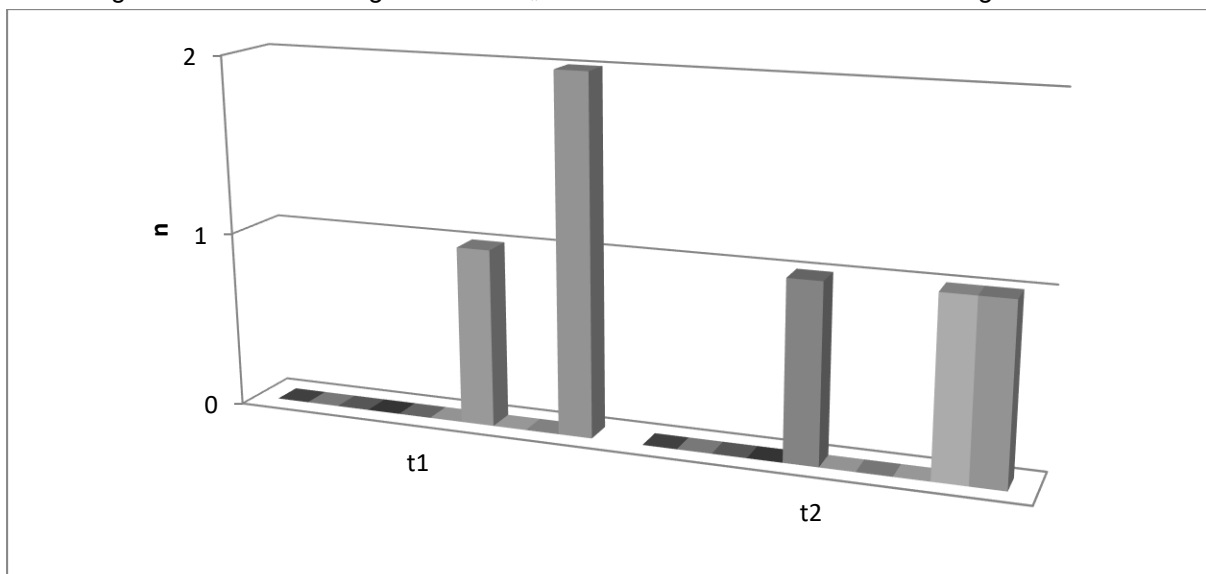
Quelle: Eigene Darstellung

Unter diesen Rahmenbedingungen stehen aufgrund der häufigen Enthaltungen von Gruppe 2 in der ersten Runde, insgesamt die in Abbildung 25 abgebildeten drei Items für eine weitere Auswertung mit Fokus auf einen Konsens über einen Dissens zur Verfügung. Da das arithmetische Mittel (aufgrund des vorliegenden Dissens) nicht die jeweiligen Bewertungen der einzelnen Kleingruppen adäquat wiedergibt, sollten für diese drei Items auch in diesem Fallbeispiel die Verteilung der jeweiligen Urteile der einzelnen Kleingruppen in die Auswertung integriert werden.

⁷³ Auch in der ersten Runde (t_1) des Gruppendelphis kam es bei den *Items 1,8,27,28,29,33,35,36* durch eine Kleingruppe (Gruppe 2) zu Enthaltungen. Da dies allerdings im Gegensatz zur zweiten Runde (t_2) nicht der Überprüfungshypothese der Konsensannahme widerspricht, konnten diese Items in der Analyse integriert werden. Dies ist allerdings ebenso nicht gänzlich unproblematisch, wie in der weiteren Analyse aufgezeigt werden wird.

Das erste Item, dessen Verteilung der einzelnen Urteile der jeweiligen Kleingruppen folgende betrachtet werden soll, bietet der Kommunikationsinhalt des *angestrebten Risikomanagements* (Item 20). Dieses Item war Teil der abgefragten Itembatterie in Frage 2.1, die auf der Einschätzung der *Agency for Research on Cancer (IARC)* basierte, dass auf „möglicherweise krebserregend“ (Risikogruppe 2b) gebildet wurde. Dies kann mittels unterschiedlicher Inhalte kommuniziert werden. Die Teilnehmenden sollten beurteilen, welche Inhalte bei solch einem Risiko mit hoher Unsicherheit an die allgemeine Bevölkerung kommuniziert werden sollten. Abbildung 26 stellt die Verteilung der einzelnen Urteile in den beiden Runden des Gruppendelphis dar.

Abbildung 26: Antwortverteilung bei Item 20 „Kommunikationsinhalte - Risikomanagement“



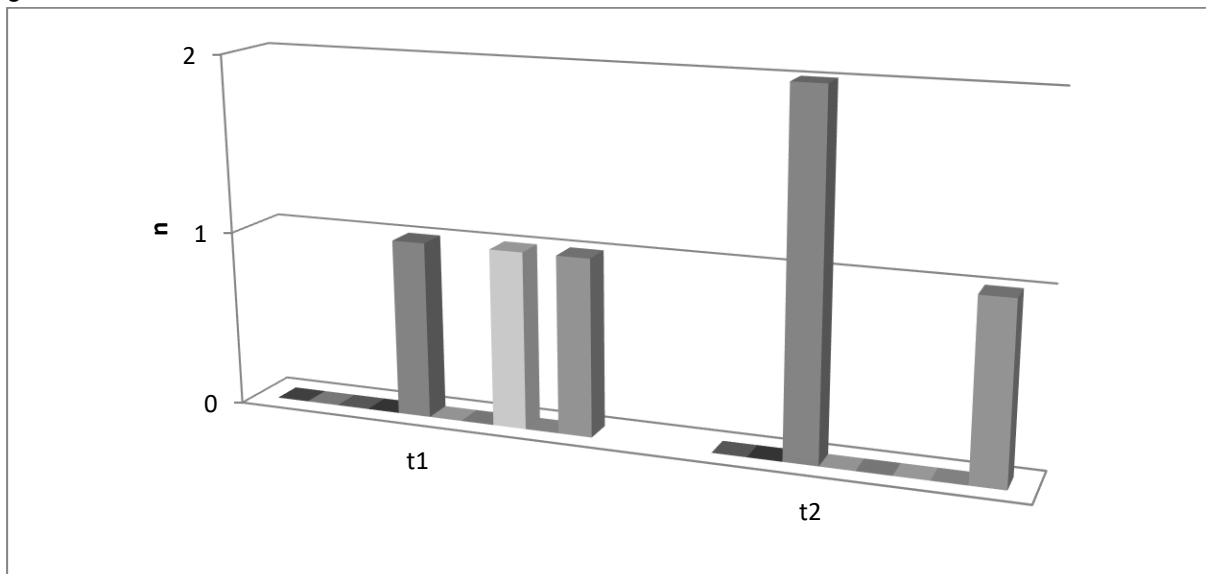
Anmerkungen: tn1=3; tn2=3

Quelle: Eigene Darstellung

In der ersten Runde (t1) werteten zwei Kleingruppen (Gruppe 1, 3), dass das *angestrebte Risikomanagement* unbedingt kommuniziert werden sollte (10), während eine Kleingruppe (Gruppe 2) diesen Kommunikationsinhalt zwar positiv (7), allerdings nicht so eindeutig, wie die beiden anderen Gruppen bewertete. Zur zweiten Runde des Gruppendelphis (t2) stellt sich die Verteilung der Antworten anders dar. So bewerteten zwei der neu zusammengesetzten Kleingruppen (Gruppe 2 und 3) das *angestrebte Risikomanagement* als einen effektiven *Kommunikationsinhalt*, der unbedingt kommuniziert werden müsse (9/10), während die erste Kleingruppe diesen Kommunikationsinhalt eher nicht (5) als Kommunikationsinhalt, der kommuniziert werden sollte, einstufte. Der zur zweiten Runde ersichtliche Konsens über einen Dissens gründete besonders darauf, dass die Expert*innen unterschiedliche Urteile fällten, inwieweit das angestrebte Risikomanagement als Information für die Öffentlichkeit relevant sei. Im ersten Plenum wurde im Einverständnis aller Teilnehmenden festgelegt, dass es sich bei der abgefragten *Bevölkerung* um die *allgemeine Öffentlichkeit* als Zielgruppe handelte.⁷⁴ Hinsichtlich der Komplexität, die mit dem Risikomanagement verbunden ist, begründete Gruppe 1 die Ablehnung des Kommunikationsinhaltes, während die beiden anderen Gruppen diesen Kommunikationsinhalt durchaus als verständlich (und damit relevant) für die allgemeine Öffentlichkeit beurteilten.

⁷⁴ Diese Präzisierung wird in Kapitel 7.2.3.2 näher beleuchtet.

Abbildung 27: Antwortverteilung bei Item 21 „Kommunikationsinhalte - Forschungsergebnisse mit vergleichenden Risiken“



Anmerkungen: tn1=3; tn2=3

Quelle: Eigene Darstellung

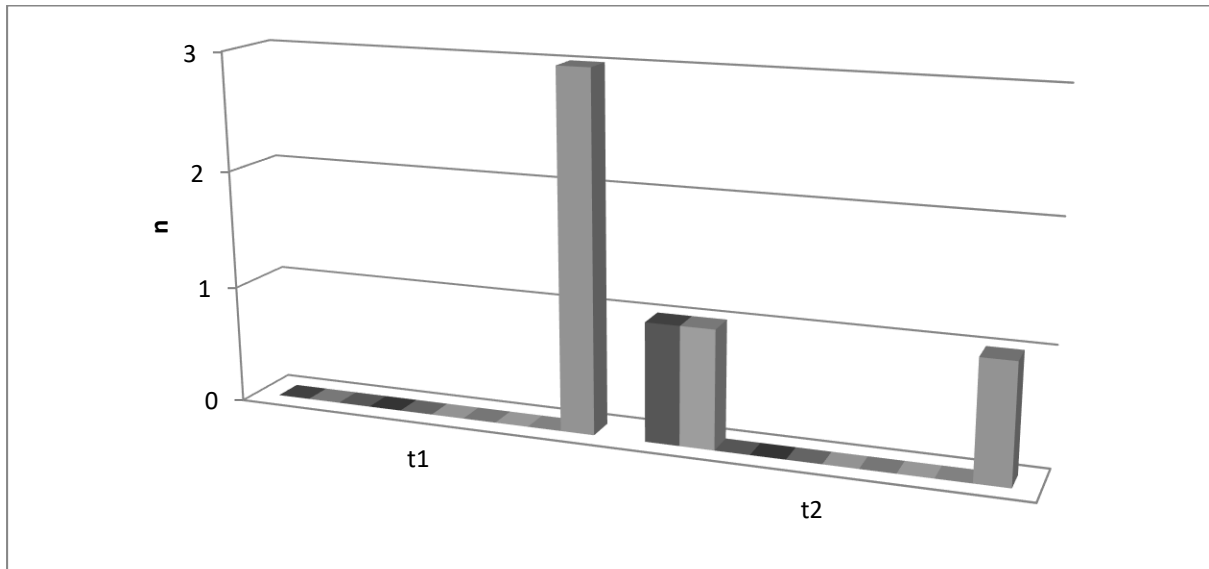
Der Effekt dieser näheren Eingrenzung der Zielgruppe spiegelt sich auch bei *Item 21* wider, das ebenso Teil der Itembatterie in *Frage 2.1* war. Dieses Item lautete: *Zur besseren Vergleichbarkeit sollten Forschungsergebnisse mit vergleichbaren Risiken dargestellt werden (etwa andere Risiken in der Kategorie 2b).*⁷⁵ Wie in Abbildung 27 dargestellt, bildet die Verteilung der Antworten der einzelnen Kleingruppen in der zweiten Runde (t2) einen Konsens über einen Dissens ab, der mittels Einbezug der vorgetragenen Argumente der einzelnen Gruppen bestätigt werden kann. So kamen zwei der Kleingruppen (Gruppe 1 und 2) in der zweiten Runde zu dem Urteil, dass ein solcher Vergleich für die allgemeine Öffentlichkeit nicht hilfreich und somit eher nicht als notwendiger Kommunikationsinhalt zu beurteilen wäre (5), während eine Kleingruppe (Gruppe 3) einen solchen Vergleich als sehr hilfreich für eine Kommunikation mit der allgemeinen Öffentlichkeit einstufte, und diese Vergleichbarkeit als einen Inhalt bewerteten, der unbedingt kommuniziert werden solle (10). Mit Blick auf die Verteilung der Antworten in der ersten Runde stellt sich die Verteilung der einzelnen Antworten in Runde 2 folglich eindeutig als Konsens über einen Dissens dar.

Bei Betrachtung von Abbildung 28 drängt sich die Frage auf, warum das *Item 30* überhaupt nochmals abgefragt wurde. So stellt sich die Beurteilung dieses Items durch die Teilnehmenden in Runde 1 (t1) mit einem Variationskoeffizienten von $v=0$ als eindeutig konsensual dar. Dies ist in der zweiten Runde (t2) nicht mehr der Fall. Hier zeigen die einzelnen Antworten der Kleingruppen einen Dissens zwischen der jeweiligen Bewertung auf. Während Gruppe 1 und Gruppe 2 wissenschaftliche Publikationen nicht als geeignete Kommunikationsweise bestimmter Messdaten bewerten (1/2), sieht Gruppe 3 solche Publikationen als sehr effektiven Weg an (10). Grund für die erneute Abfrage ist eine Diskussion im ersten Plenum, in der von den Teilnehmenden geschlossen entschieden wurde, den einleitenden Fragetext zu *Frage 3.1* hinsichtlich der Übermittlungswege zur Kommunikation von Messdaten durch eine Ergänzung durch wichtige *wissenschaftlichen Studienergebnisse – die Handlungsempfehlungen und*

⁷⁵ Hervorhebung im Original. Siehe SEAWIND-Fragebogen im Anhang 11.6.

Schlussfolgerungen begründen für das gesamte Itemset in dieser Frage zu übernehmen (Items 23 – 33).

Abbildung 28: Antwortverteilung bei Item 30 „Kommunikation von Messdaten - Wissenschaftliche Publikationen (Effektivität)“



Anmerkungen: tn1=3; tn2=4

Quelle: Eigene Darstellung

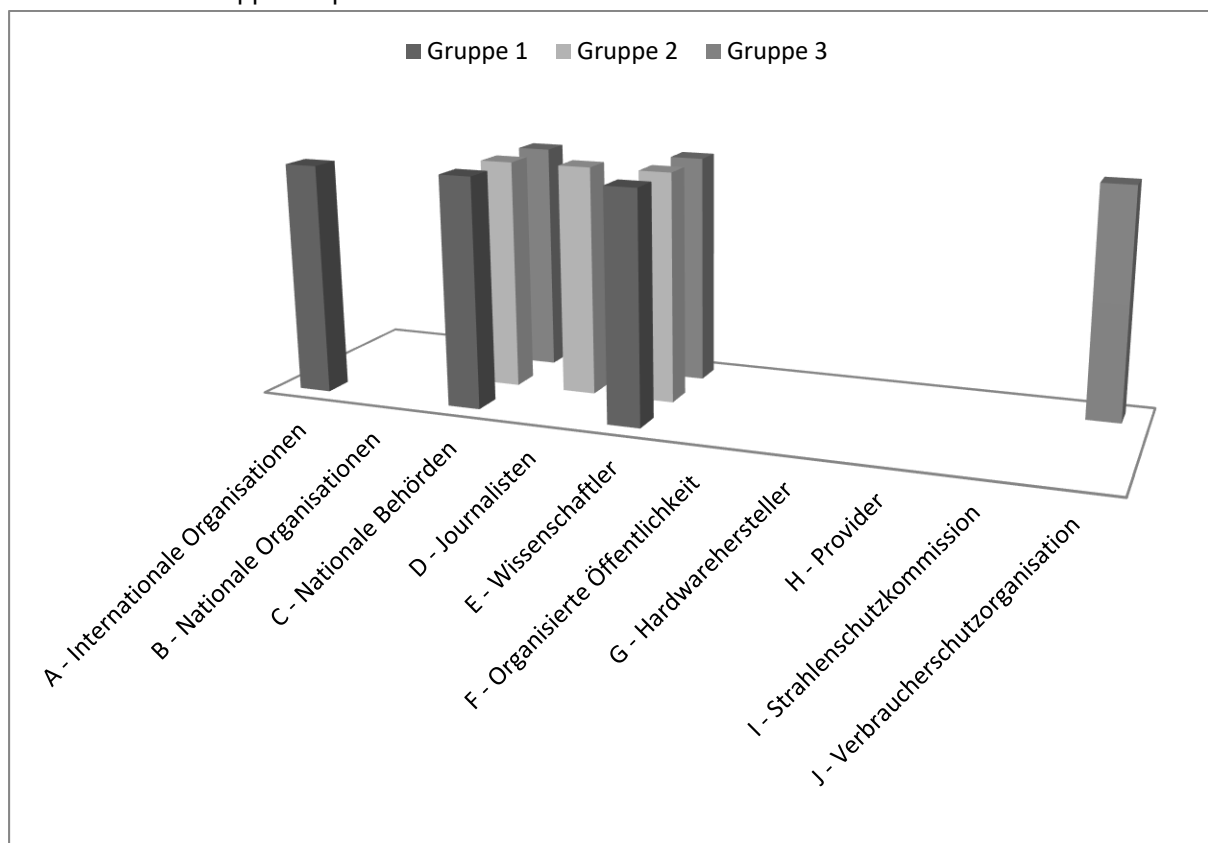
Unter diesem veränderten Kontext wurden folglich alle Items dieser *Frage 3.1* zur erneuten Bewertung in die zweite Kleingruppenphase des Delphis gegeben. Der Dissens begründete sich letztlich darauf, dass die zwei ablehnenden Gruppen (Gruppe 1 und 2) wissenschaftliche Publikationen nicht als effektive *zielgruppenspezifische Kommunikation* an die allgemeine Öffentlichkeit ansahen, während die dritte Gruppe dies sehr wohl tat, da solche Quellen als seriös einzustufen seien und nur in entsprechender, verständlicher Form publiziert werden müssten. Die Gruppen konnten sich zur zweiten Plenarsitzung hinsichtlich dieses Items auf einen Konsens über einen Dissens verständigen und stufen damit letztlich die Effektivität von wissenschaftlichen Publikationen zur Kommunikation von Messdaten sehr unterschiedlich ein.

Die Analyse der 10er-skalierten Items leidet im Falle des SEAWIND-Gruppindelphis unter der sehr hohen Anzahl von Enthaltungen durch die Kleingruppen. So treten in der ersten Kleingruppenphase insgesamt 28 Enthaltungen auf. Diese stammen geschlossen von Gruppe 2, die sich über die einzelnen Abschnitte des Fragebogens übermäßig oft enthalten hat. In der zweiten Runde sinkt die Zahl der Enthaltungen auf insgesamt 13 Enthaltungen und verteilt sich über die drei gebildeten Kleingruppen hinweg: Gruppe 1, 8 Enthaltungen; Gruppe 2, 4 Enthaltungen und Gruppe 3, eine Enthaltung.

Da in der Einführung des Delphis nicht explizit darauf hingewiesen wurde, dass für die Expert*innen die Möglichkeit besteht Minderheitsvoten abzugeben, lässt sich die hohe Zahl der Enthaltungen von Gruppe 2 in der ersten Runde damit erklären, dass sich die Personen dieser Kleingruppe bei Uneinigkeit darauf verständigten kein Urteil für das abgefragte Item abzugeben. Dies wurde so von den Teilnehmenden dieser Kleingruppe im ersten Plenum geäußert. Warum aber auch in der zweiten Runde eine große Zahl an Enthaltungen vorliegt, lässt sich damit allerdings nicht begründen. In Folge dessen bleiben letztlich für die Analyse in der Dimension der Evidenz die 17 weiter oben dargestellten Items übrig.

Im Fragebogen des SEAWIND-Gruppendelphis wurden neben den 10-er skalierten Items überdies zwei Rankingfragen integriert, in denen die Teilnehmenden gebeten wurden, die *drei wichtigsten* Aspekte auszuwählen. Die erste *Frage 1.2* war Teil des ersten Abschnitts des Fragebogens, der thematisch das Feld der *Kommunikationsakteure* abdeckte. Diese Frage beschäftigte sich dabei mit der möglichen Notwendigkeit einer Zusammenarbeit verschiedener Akteure, die bei der *Kommunikation von Risiken mit hohen Unsicherheiten* sinnvoll sein könnte. Die Expert*innen wurden gebeten, die drei wichtigsten Akteursgruppen auszuwählen, die *ihrer Meinung nach zusammenarbeiten sollten, um die Risiken elektromagnetischer Felder an die Öffentlichkeit zu kommunizieren*.

Abbildung 29: Antwortverteilung bei Frage 1.2 „Zusammenarbeit von Akteuren“ in der ersten Runde des SEAWIND-Gruppendelphis.



Anmerkungen: tn1=3

Quelle: Eigene Darstellung

Wie in Abbildung 29 dargestellt, wählten in der ersten Runde des Gruppendelphis alle drei Kleingruppen bereits geschlossen zwei Akteursgruppen, deren Zusammenarbeit hinsichtlich dieses Sachverhaltes notwendig sei:

1. *Nationale Behörden, wie z.B. das Bundesministerium für Gesundheit, das Bundesamt für Strahlenschutz oder das Bundesamt für Risikobewertung.*
2. *Wissenschaftler*

In der Plenardiskussion waren sich alle Teilnehmenden einig, dass die zuständigen nationalen Behörden der zentrale Akteur seien, um die Öffentlichkeit zu informieren. Damit sei eine Zusammenarbeit bei einer Kommunikation über potentielle Risiken elektromagnetischer Felder mit dieser Akteursgruppe unverzichtbar. Um die Qualität der Einschätzung zu erhöhen, sollte nach Ansicht der Teilnehmenden eine Zusammenarbeit mit der Wissenschaft erfolgen. Gruppe

2 brachte ein, dass die Qualität der Botschaft auch die Zusammenarbeit mit der Akteursgruppen der Journalist*innen sehr relevant sei, da diese für die Übermittlung der wissenschaftlichen Analysen an die Öffentlichkeit die relevanteste Akteursgruppe seien. Da auch die anderen Teilnehmenden eine Zusammenarbeit zwischen Wissenschaftler*innen und Journalist*innen als essentiell für die Qualität der Botschaft ansahen, wurde eine neue Auswahlkategorie in den Fragebogen der zweiten Delphirunde aufgenommen: *Tandem Journalist und Wissenschaftler*.

Ebenfalls wurden *Verbraucherschutzorganisationen* in den Fragebogen der zweiten Runde integriert. Dieser Vorschlag wurde von Gruppe 3 im Fragebogen der ersten Kleingruppenphase unter der offenen Kategorie *Sonstige* eingebracht. Die Teilnehmenden argumentierten, dass solche Organisationen aufgrund ihrer Glaubwürdigkeit gegenüber der Öffentlichkeit als relevante Akteursgruppe in die Zusammenarbeit der anderen Akteure unbedingt einbezogen werden müssten. Da das Projekt SEAWIND von Europäischer Ebene gefördert wurde, wählte die erste Kleingruppe *internationale Organisationen wie z.B. die Europäische Kommission* aus. Die anderen beiden Kleingruppen sahen diese Akteursgruppe nicht als besonders relevant für eine Zusammenarbeit an, da zwar auch ein internationaler Austausch über die möglichen Risiken elektromagnetischer Felder geführt werden, die Information der Öffentlichkeit aber national angepasst erfolgen müsse. Dieses Argument wurde von Gruppe 1 als legitim erachtet, so dass sie der Argumentation folgte und die internationalen Organisationen ebenso als keine der relevanten Akteursgruppen einstufte.

Die Diskussion im Plenum (t1) führte zu zwei weiteren Ergänzungen im Fragebogen der zweiten Delphirunde, da die nationalen Behörden nach Ansicht eines Experten, wie konsensual im Teilnehmendenfeld angenommen, eine relevante Akteursgruppe für eine Zusammenarbeit seien, diese aber weiter spezifiziert werden müssten. So schlug dieser Teilnehmer vor, sowohl das *Umweltministerium* sowie auch das *Bundesamt für Strahlenschutz* als konkrete Akteure in den Fragebogen zu integrieren. Diese waren aus Sicht des Teilnehmenden die zentralen behördlichen Akteure, wenn es darum ginge über Risiken elektromagnetischer Felder zu informieren. Dieser Argumentation widersprachen die übrigen Expert*innen nicht, so dass beide Akteure als jeweils eigene Auswahlkategorie in den Fragebogen aufgenommen wurden. Da weiterhin ein Konsens darüber bestand, dass sowohl nationale Behörden im Allgemeinen, wie auch die Wissenschaft relevante Akteure seien, wurden diese im Fragebogen der zweiten Runde als konsensual markiert. So konnten die gemischten Kleingruppen zur zweiten Delphirunde jeweils nur eine (weitere) Akteursgruppe auswählen, die für die Zusammenarbeit als wichtigste Akteursgruppe eingestuft wurde.

An diese Vorgabe hielt sich in der zweiten Runde des Gruppendelphis lediglich die erste neu zusammengesetzte Kleingruppe (Gruppe 1), die die Akteursgruppe der Journalist*innen wählte, allerdings mit dem schriftlichen Hinweis versah, dass diese *für Inhalte, nicht für allgemeine Strategien* zuständig sein sollten. Gruppe 2 hingegen strich den vermerkten Konsens im Fragebogen und wählte sowohl die *Verbraucherschutzorganisationen* sowie das *Tandem Journalist und Wissenschaftler*. Gruppe 3 wählte ohne Streichung des vermerkten Konsenses sowohl das *Umweltministerium* als auch das *Bundesamt für Strahlenschutz*. Aufgrund der Differenz bezüglich des Antwortverhaltens zwischen den jeweiligen Kleingruppen, wurde zum zweiten Plenum auf eine graphische Aufarbeitung der einzelnen Antworten verzichtet, da im Besonderen die beiden Gruppen 2 und 3 durch ihre zusätzliche Auswahl überrepräsentiert

gewesen wären. Allerdings fand eine Diskussion der jeweiligen Priorisierungen im Plenum statt. Da aufgrund der Aufnahme des Tandems zwischen Journalist*innen und Wissenschaft und der zusätzlichen Abfrage der einzelnen Akteursgruppen *Wissenschaft* und *Journalisten* im Fragebogen der zweiten Runde eine fehlende Trennschärfe zwischen den einzelnen Auswahlitems verursacht wurde, konnte dieser Dissens in der Diskussion schnell aufgelöst werden. Alle Gruppen sahen sowohl die Journalist*innen wie auch die Wissenschaft neben den Behörden als relevante Akteure mit den ihnen jeweiligen zugeschriebenen Rollen an. Dazu kam als konkrete Behörden das von Gruppe 3 ausgewählte *Bundesamt für Strahlenschutz* sowie das *Umweltministerium* hinzu. Neben diesen stimmten alle Teilnehmenden darin überein, dass die Verbraucherschutzorganisationen ebenfalls als eine wichtige Akteursgruppe zu werten seien. So ergab sich statt der ursprünglich angedachten drei Akteursgruppen ein konsensuales Ergebnis für eine Zusammenarbeit von vier Akteursgruppen:

1. *Nationale Behörden*, insbesondere das Umweltministerium wie auch das Bundesamt für Strahlenschutz
2. *Wissenschaftler*
3. *Journalisten*
4. *Verbraucherschutzorganisationen*

Die zweite Rankingfrage mit Auswahloption von drei Items wurde im zweiten Abschnitt des Fragebogens abgefragt, der das Thema der *Kommunikationsinhalte* behandelte. Konkret wurden in Frage 2.2 die *Ziele der Kommunikation* abgefragt, in der die Teilnehmenden gebeten wurden, die Ziele zu benennen, die ihrer *Meinung nach durch eine Kommunikation zu Risiken* mit hoher Unsicherheit erreicht werden sollten. Die Expert*innen wurden gebeten, die drei Ziele anzugeben, die ihrer Meinung nach *durch eine entsprechende Risikokommunikation auf jeden Fall erreicht werden sollten*.

Tabelle 7: Antwortverteilung bei Frage 2.2 „Ziele der Kommunikation“ in der ersten Runde des SEAWIND-Gruppendelphis.⁷⁶

Kommunikationsziele	Gruppe 1	Gruppe 2	Gruppe 3
Das Vertrauen in Risikobewertungsverfahren steigern.			
Das Vertrauen in die kommunizierenden Institutionen steigern.		x	
Einseitige Risiko-Interpretationen vermeiden.		X	
Unsicherheitsquellen in der Risikobewertung identifizieren.	X	x	
Den Grad der Evidenz der Risikoanalysen charakterisieren und bewerten.			
Das Management durch zuständige Behörden im Umgang mit dem Risiko darlegen.			
Schaffung von Transparenz in der Risikobewertung und im Risikomanagement.	X	X	X
Hervorhebung der Kompetenzen der risikobewertenden Institutionen.			
Nachvollziehbarkeit der Datengrundlage zur Risikobewertung für jedermann.			
Den Umgang mit (unvermeidbaren) Risiken verdeutlichen.	X		
Nennung Sonstige: Handlungsempfehlungen im Umgang mit konkreten Risiken			X

Anmerkungen: Da Gruppe 2 eine ihrer Auswahlmöglichkeiten auf zwei Kommunikationsinhalte aufgeteilt hat, ist eine ganze Auswahl als X und besagte halbe Stimmen als x gekennzeichnet.

⁷⁶ Da Gruppe 2 eine ihrer drei Auswahlmöglichkeiten auf zwei Items aufgeteilt hat, wurde zur Verdeutlichung der Antworten nicht die Form einer Abbildung, sondern die einer Tabelle gewählt.

Tabelle 7 bietet einen Überblick über die Auswahl der einzelnen Kleingruppen in der ersten Runde des Gruppendelphis. Wie in der Tabelle anhand der Markierung erkennbar, wählten die Kleingruppen zur ersten Runde bereits die *Schaffung von Transparenz in der Risikobewertung und im Risikomanagement* konsensual als eines der wichtigsten Kommunikationsziele. Auch im ersten Plenum waren sich die Expert*innen geschlossen einig, dass dies als wichtigstes Ziel mittels entsprechender Risikokommunikation auf jeden Fall erreicht werden solle. Dieser Konsens wurde im Fragebogen der zweiten Runde markiert und die Teilnehmenden wurden - ähnlich wie in *Frage 1.2* - gebeten, in der zweiten Kleingruppenphase nur zwei weitere Ziele zu wählen. Gruppe 3 nutzte in der ersten Runde die offene Kategorie um ein weiteres Kommunikationsziel vorzuschlagen: *Handlungsempfehlungen im Umgang mit konkreten Risiken*. Die Teilnehmenden argumentierten im ersten Plenum, dass gerade Handlungsempfehlungen als wichtiges Ziel der Kommunikation von Risiken verstanden werden sollen. Dieser Argumentation folgten die Teilnehmenden der beiden anderen Kleingruppen, so dass konsensual entschieden wurde, dieses Kommunikationsziel in den Fragebogen der zweiten Runde aufzunehmen.

Überdies erläuterte Gruppe 2 ihre Motive für die Teilung einer ihrer Auswahlmöglichkeiten. So war sich die Kleingruppe uneinig, ob die Steigerung des Vertrauens in die kommunizierenden Institutionen als eines der wichtigsten Ziele eingestuft werden könne. Ein Teil der Kleingruppe kam zu dem Urteil, dass dies eines der wichtigsten Ziele sei. Ein anderer Teil der Gruppe sah dieses Ziel als positives Begleitziel, falls die anderen Ziele erreicht würden. Somit müsse der Schwerpunkt besonders auf die Transparenz gelegt werden. Dieser Argumentation konnte die gesamte Gruppe der Expert*innen letztlich zustimmen. Die zweite gesplittete Auswahl bezog sich darauf, dass nicht nur *Unsicherheitsquellen in der Risikobewertung* identifiziert werden müssten, sondern ein Teil der Gruppe 2 die Meinung vertrat, dass insgesamt jegliche *Unsicherheit in der Risikobewertung* identifiziert werden müsse. Auch dieser Argumentation konnte das Teilnehmendenfeld geschlossen folgen, so dass konsensual beschlossen wurde, das Item für die zweite Runde von *Unsicherheitsquellen in Unsicherheit* zu ändern.

In der zweiten Kleingruppenphase wählten die neu zusammengesetzten Kleingruppen neben der Transparenz zwei weitere, der im ersten Plenum besprochenen Ziele aus, so dass letztlich, wie bei der Konzeption der Frage vorgesehen, nach Urteil der Teilnehmenden die drei wichtigsten Ziele identifiziert wurden:

1. *Schaffung von Transparenz in der Risikobewertung und im Risikomanagement.*
2. *Handlungsempfehlungen im Umgang mit konkreten Risiken.*
3. *Unsicherheit in der Risikobewertung identifiziert.*

Während dieser evidenzbasierten Betrachtung des Fallbeispiels in SEAWIND ist bereits aufgefallen, dass bei diesem Gruppendelphi zwischen dem Fragebogen der ersten und der zweiten Delphirunde bereits vermehrte und vielfältige Umformulierungen und Ergänzungen vorgenommen wurden. Welche Effekte diese Modifikationen im Bezug auf die Verständlichkeit hatten und ob dies einen positiven Einfluss auf eine gesteigerte konsensuale Einschätzung der Expert*innen hatte, soll im folgenden Unterkapitel näher betrachtet werden.

7.2.3.2 Dimension: Verständlichkeit

Auch im Falle des SEAWIND-Gruppendelphis kann eine Analyse hinsichtlich der Zweckmäßigkeit des Diskurses durchgeführt werden, da der Fragebogen zwischen der ersten Runde und der zweiten Runde angepasst wurde. So kann im ersten Schritt überprüft werden, ob eine Anpassung aufgrund von interdisziplinär unterschiedlich verstandenen Fachtermini, semantischen Missverständnissen oder durch das Forscherteam nicht klar formulierter Fragen und Items zu unterschiedlichen Interpretationen zwischen den Teilnehmenden führten und überdies, ob die Anpassungen dazu führten, dass die Expert*innen nach der Anpassung des Fragebogens letztlich mit dem *Geäußerten* in der zweiten Runde einen gleichen Bedeutungsinhalt verstanden.

In SEAWIND wurden die Items, anders als in der methodischen Konzeption im Ideal vorgesehen, nicht als formulierte Aussagen abgefragt, sondern durch bloße Nennung der abgefragten Sachverhalte dem Urteil der Teilnehmenden gestellt. Trotz dessen - oder gerade deswegen⁷⁷ - wurden in diesem Fall besonders in den einleitenden Fragetexten zu den einzelnen Itembatterien im Laufe dieses Gruppendelphis Umformulierungen zwischen dem ersten Plenum und der zweiten Kleingruppenphase vorgenommen. So stehen insgesamt fünf Fragen für eine tiefergehende Betrachtung in der Dimension der Verständlichkeit zur Verfügung.

Die erste Umformulierung betrifft *Frage 1.1* die *Inhalte einer Information* der Öffentlichkeit hinsichtlich des Krebsrisikos durch elektromagnetische Felder thematisierte. Der Fragetext lautete: *Im Jahr 2011 hat die International Agency for Research on Cancer (IARC) ihre Einschätzung des Krebsrisikos durch elektromagnetische Felder (EMF) bekanntgegeben. Die Einschätzung lautet: „möglicherweise krebserregend“ (Risikogruppe 2b).⁷⁸ Welche der nachfolgend aufgelisteten Akteursgruppen sollten Ihrer Meinung nach die Öffentlichkeit über diese Einstufung der Risiken unter die Kategorie 2b informieren? Bitte bewerten Sie diesbezüglich nachstehend die Akteursgruppen. Die Ziffer 1 (-) bedeutet hierbei, dass diese Akteursgruppe die Öffentlichkeit auf keinen Fall über diese Einstufung und ihre Implikationen informieren sollte und die Zahl 10 (+) bedeutet, dass diese Gruppe auf jeden Fall die Öffentlichkeit über diese Einstufung und ihre Implikationen informieren sollte.*

Während der ersten Plenarphase merkten die Expert*innen bereits vor der Diskussion einzelner Items zu dieser Itembatterie an, dass die Formulierung der Frage grundsätzlich hinterfragt werden solle. So argumentierte ein Teilnehmer aus Gruppe 2, dass verschiedene Aspekte die Bewertung der einzelnen Items beeinflussen, da die Formulierung *wer sollte* nicht auf alle der abgefragten Akteure zuträfe. Dies begründet auch die hohe Zahl an Enthaltungen dieser Gruppe bei insgesamt 9 der 11 abgefragten Items in der ersten Kleingruppenphase. Die *Sollensaussage* träfe prinzipiell nur auf Politiker auf nationaler Ebene und nationale Behörden zu, da diese laut Definition ihrer Rolle, je nach Kontext der Dringlichkeit, teilweise der Informationspflicht folgen müssten. Die anderen beiden Gruppen (1, 3) konnten dieses Argument nachvollziehen, hatten dies aber zur ersten Runde nicht in ihre Bewertung einfließen lassen. Beide Gruppen interpretierten die Frage in der ersten Kleingruppenphase so, dass es darum ginge zu beurteilen, welche Gruppen besonders geeignet wären, eine solche informierende Rolle

⁷⁷ Dies soll im Zwischenfazit 7.2.4 nochmals tiefer diskutiert werden.

⁷⁸ IARC Risikogruppen: (1) krebserregend, (2a) wahrscheinlich krebserregend, (2b) möglicherweise krebserregend, (3) nicht klassifizierbar für Menschen, (4) wahrscheinlich nicht krebserregend.

einzunehmen. Gruppe 2 stimmte zu, dass ein solcher Bewertungsrahmen sinnvoller wäre. Da dies aber nicht aus der Frageformulierung eindeutig hervorging, einigte man sich geschlossen auf eine Umformulierung des Fragetextes im Fragebogen der zweiten Kleingruppenphase (t2): *Welche der nachfolgend aufgelisteten Akteursgruppen ist Ihrer Meinung nach am besten geeignet, die Öffentlichkeit über diese Einstufung der Risiken unter der Kategorie 2b informieren?*⁷⁹ Richtet man den Blick auf die beiden Items, bei denen ein Vergleich über die Runden hinweg möglich ist, weil Gruppe 2 sich in der ersten Runde bei diesen nicht enthalten hatte, so lässt sich kein konsensuales Ergebnis erkennen.

Tabelle 8: Werte des Variationskoeffizienten bei Frage 1.1 „Inhalte der Information“ in SEAWIND

Item #	Frageinhalt	v(t1)	v(t2)
7	Hardwarehersteller	0,8	0,8
9	Politiker auf nationaler Ebene	1,1	0,6

Wie in Tabelle 8 aufgeführt, hat sich der Wert des Variationskoeffizienten bei den *Hardwareherstellern* (Item 7) nicht in Richtung eines konsensualeren Ergebnisses verändert. Bei Item 9 über *Politiker auf nationaler Ebene* ist der sehr hohe Wert des Variationskoeffizienten von $v=1,1$ (t1) auf einen Wert von $v=0,6$ (t2) gesunken, liegt aber weiterhin über dem Erfahrungswert für einen Konsens von $v=0,5$. Beide Items werden zur zweiten Runde (t2) von den Teilnehmenden allerdings geschlossen abgelehnt.⁸⁰ Auch wenn sich damit hinsichtlich der statistischen Maßzahlen kein sichtbarer Effekt der Umformulierung des Fragetextes erkennen lässt, so kann anhand der Fallzahlen der Antworten bei den abgefragten Items dieser *Frage 1.1* durchaus ein Effekt beobachtet werden. Während in der ersten Kleingruppenphase (t1) neun der elf Items nur eine Fallzahl von $n=2$ Antworten aufweisen, ist dies in Runde 2 nicht mehr der Fall. Die neu zusammengesetzten Gruppen 2 und 3 enthielten sich lediglich bei *Item 8*, so dass alle anderen Items eine Fallzahl von $n=3$ vorweisen können.

Auch bei der Abfrage von *Inhalten der Kommunikation* (Frage 2.1) kam es zu einer Änderung der Frageformulierung im Fragebogen für die zweite Runde (t2). Die Frage im ersten Workshopfragebogen lautete: *Die Einschätzung der Agency for Research on Cancer (IARC), dass elektromagnetische Felder „möglicherweise krebserregend“ (Risikogruppe 2b) sind, kann mittels unterschiedlicher Inhalte kommuniziert werden. Was denken Sie: Welche Inhalte sollten bei solch einem Risiko mit hoher Unsicherheit an die allgemeine Bevölkerung kommuniziert werden?*

Auch bei dieser Frage wurde bereits vor der tiefergehenden Diskussion der Items auf eine, aus Sicht der Teilnehmenden, missverständlich Formulierung im Fragetext hingewiesen. So könne der Begriff *allgemeine Bevölkerung* im Prinzip jeden miteinschließen, was aber aufgrund der zielgruppenspezifischen Abhängigkeit der einzelnen Items problematisch sei. So wurde gewünscht, dass für die Beurteilung der einzelnen Items ein Rahmen definiert würde, an welche Gruppe sich die Inhalte der Kommunikation richten sollen. Klarer als der Begriff *allgemeine Bevölkerung* sei hier schlicht der Begriff *Öffentlichkeit*. Auch wenn nicht alle Teil-

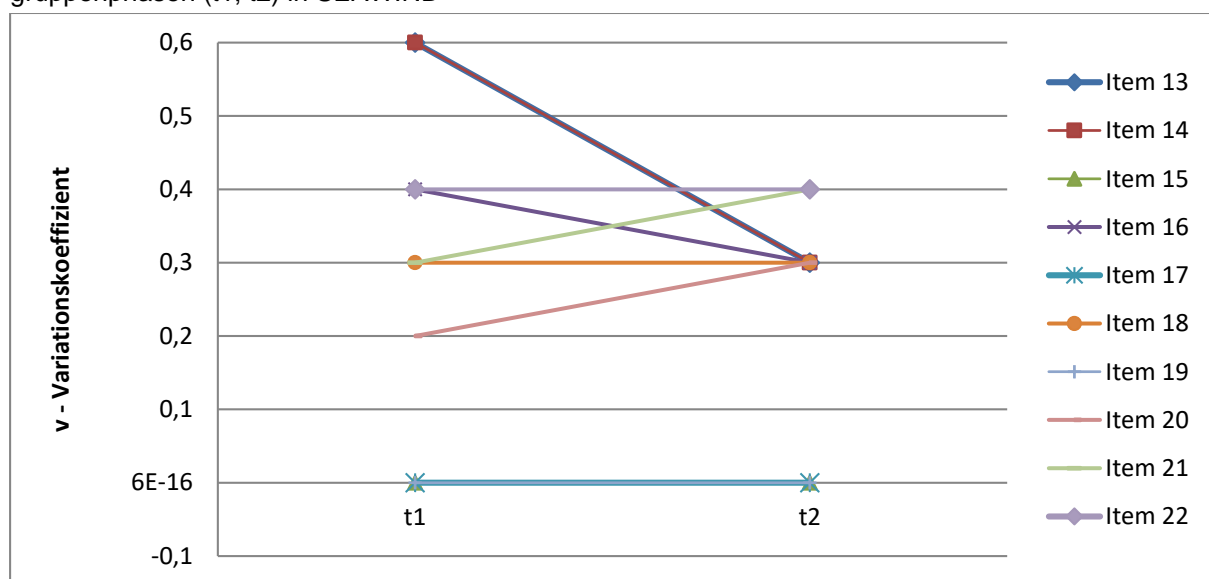
⁷⁹ Hervorhebung im Original des Fragebogens der zweiten Runde.

⁸⁰ *Item 7* – Hardwarehersteller (1/1/5) und *Item 9* – Politiker auf nationaler Ebene (1/2/5).

nehmenden die von einzelnen Expert*innen thematisierte Problematik mit dem Begriff der *allgemeinen Bevölkerung* nachvollziehen konnten, so konnten sie sich darauf einigen, dass alle unter dem Begriff der *Öffentlichkeit* einen identischen Bewertungsrahmen für die Items teilen konnten. So wurde der Fragetext wie folgt angepasst: *Was denken Sie: Welche Inhalte sollten bei solch einem Risiko mit hoher Unsicherheit an die **Öffentlichkeit** kommuniziert werden?*⁸¹

Im Gegensatz zu *Frage 1.1* gab es bei *Frage 2.1* keine Enthaltungen hinsichtlich einzelner Items durch einzelne Kleingruppen, so dass diese Itembattery in Abbildung 30 gesamt dargestellt werden kann. Mit Blick auf diese Abbildung wird erkennbar, dass alle abgefragten Items zur zweiten Runde unter einen Wert des Variationskoeffizienten von $v=0,5$ fallen, bzw. auf ihrem Wert verweilen.

Abbildung 30: Entwicklung des Wertes des Variationskoeffizienten bei Frage 2.1 zwischen den Kleingruppenphasen (t1; t2) in SEAWIND



Anmerkungen: $tn1=3$; $tn2=3$;

Quelle: Eigene Darstellung

Ebenso fällt auf, dass die bereits in Runde 1 mit einem Wert des Variationskoeffizienten von $v=0$ beurteilten *Items 15, 17 und 19* auch zur zweiten Runde stabil auf diesem konsensualen Wert bleiben. Die Urteile der Expert*innen bewegen sich unter Beachtung des Wertes des Variationskoeffizient lediglich bei *Item 20* und *Item 21* zum Zeitpunkt t2 weniger konsensual als zu t1, wobei sich bei beiden eindeutig ein Konsens über ein Dissens zur zweiten Runde herausbildet.⁸²

Überdies fand bei dieser *Frage 2.1* eine Umformulierung eines einzelnen Items statt. So lautete *Item 13* im Fragebogen der ersten Kleingruppenphase: *Die Datengrundlage zur Festlegung der Risikogruppe.*⁸³ In der ersten Kleingruppenphase beurteilten Gruppe 1 und Gruppe 3 dieses Item als Kommunikationsinhalt, der *unbedingt kommuniziert werden sollte*. Gruppe 2 hingegen beurteilte dieses *Item 13* als Kommunikationsinhalt der *ganz und gar nicht kommuniziert werden sollte*. Basierend auf diesem Dissens ($v=0,6$) zwischen den Urteilen zu diesem Inhalt wurde dieses *Item 13* folglich im ersten Plenum tiefgehend diskutiert. Gruppe 2 begründete

⁸¹ Hervorhebung im Original des Fragebogens der zweiten Runde.

⁸² *Item 20*: Runde 1 (10/7/10), Runde 2 (5/9/10); *Item 21*: Runde 1 (8/5/10), Runde 2 (5/5/10).

⁸³ Hervorhebung im Original des Fragebogens der ersten Runde.

ihre Ablehnung dieses Inhaltes anhand der Einschätzung, dass die Datengrundlage für die Öffentlichkeit nicht besonders von Interesse sei. Als zweites Argument führte die zweite Kleingruppe weiterhin an, dass die Datengrundlage auch nicht für die Zielgruppe verständlich sei, besonders wenn es sich um Rohdaten handle. Diese Rohdaten müssten aufbereitet sein und mit den entsprechenden Handlungsempfehlungen sinnvoll verknüpft werden, bevor man diese kommuniziere. Damit müsse einzig die Datengrundlage, auf der die Einschätzung *möglicherweise krebserregend* für *elektromagnetische Felder* basiere, kommuniziert werden, um die Seriosität der Einschätzung zu vermitteln. Überdies führte ein Teilnehmer an, dass die Veröffentlichung von Rohdaten auch die Gefahr eines Missbrauchs beherberge.

Dem Argument, dass Rohdaten als Kommunikationsinhalt für die Öffentlichkeit weder verständlich noch geeignet seien, konnten sich die beiden anderen Gruppen anschließen. Sie teilten überdies das Argument, dass die aufbereitete Datengrundlage grundsätzlich ein wesentlicher Inhalt der Kommunikation sein solle, um die wissenschaftliche Basis der Einschätzung transparent zu machen. Für die Wissenschaft sei überdies die Datengrundlage sehr relevant. Dem konnte auch Gruppe 2 folgen und so einigte sich das Plenum darauf, dass das Item die *Datengrundlage* enthalten solle, aber eben *keine Rohdaten*. Dies wurde mittels Ergänzung der Itemformulierung im Fragebogen der zweiten Kleingruppenphase kenntlich gemacht, so dass das *Item 13* wie folgt abgefragt wurde: Die *Datengrundlage (keine Rohdaten)* zur Festlegung der Risikogruppe.⁸⁴

Die Einigkeit in der ersten Plenarsitzung spiegelt sich in den statistischen Ergebnissen der zweiten Delphirunde nicht gänzlich wieder. So bewertet Gruppe 3 das Item mit einem Wert von 10, während Gruppe 2 mit einem Wert von 6 nur bedingt zustimmt und Gruppe 1 mit einem Wert von 5 dieses Item sogar eher ablehnt. Der Wert des Variationskoeffizienten sank zwar von $v=0,6$ (t_1) auf $v=0,3$ (t_2), allerdings kann bei *Item 13* nicht von einem konsensualen Ergebnis zwischen den Kleingruppen zur zweiten Runde ausgegangen werden. Das Item wurde zur zweiten Plenarsitzung nicht tiefergehend besprochen, so dass die Gründe für die bestehende Uneinigkeit hinsichtlich des jeweiligen Urteils, trotz Umformulierung des *Items 13* sowie Umformulierung der *Frage 2.1*, nicht näher bestimmbar bleiben. Um dieses Item erfüllend bewerten zu lassen, wäre eine dritte Runde notwendig gewesen. So hätte sowohl eine erneute Diskussion im zweiten Plenum sowie auch in einer dritten Kleingruppenphase zu einem klareren Ergebnis führen können - oder zumindest die Gründe für die bestehende Uneinigkeit aufgedeckt. Aufgrund des knappen Zeitplans des eintägigen Workshops war dies nicht möglich.

Die nächste Umformulierung im Fragetext fand bei *Frage 3.1* im dritten Abschnitt zur *Kommunikation von Messdaten* statt, die sich thematisch mit den möglichen *Wegen der Übermittlung von Messdaten und deren Implikationen* befasste. Der Fragetext im Fragebogen der ersten Runde lautete: *Grundlage für die Beurteilung eines Risikos sind Messdaten. Die Kommunikation von Messdaten und deren Implikationen kann auf verschiedene Art und Weise erfolgen. Bitte bewerten Sie die generelle Effektivität der dargestellten Kommunikationsweise, wobei Effektivität hier ein Beurteilungskriterium darstellen soll, mit dem sich beschreiben lässt, in welchem Ausmaß eine Kommunikationsweise dafür geeignet ist, bestimmte Messdaten zu*

⁸⁴ Hervorhebung im Original des Fragebogens der zweiten Runde.

kommunizieren. 1 (-) bedeutet, dass Sie diesen Weg als ineffektiv bewerten, 10 (+) bedeutet, dass Sie diesen Weg als sehr effektiv bewerten.

Bei dieser Frage fand die bereits in Kapitel 7.2.3.1 dargestellte Diskussion im ersten Plenum hinsichtlich der Begriffe *Messdaten* und *Implikationen* statt. Diese Begriffe waren den Teilnehmenden nicht präzise genug, so dass der Fragetext stattdessen durch *wichtige wissenschaftliche Studienergebnissen - die Handlungsempfehlungen und Schlussfolgerungen begründen* ersetzt wurde. Da die zweite Gruppe in der ersten Kleingruppenphase lediglich bei drei der elf abgefragten Items dieses Itemsets ein Urteil abgegeben hatte, kann nur bei drei *Übermittlungswegen* ein Vergleich zwischen den beiden Runden t1 und t2 erfolgen.

Tabelle 9: Werte des Variationskoeffizienten bei Frage 3.1 „Wege der Übermittlung“ in SEAWIND

Item #	Frageinhalt	v (t1)	v (t2)
24	Bürger-Gesprächsgruppen (Effektivität)	0,6	0,1
30	Wissenschaftl. Publikationen (Effektivität)	0,0	0,6
31	Bürgerforen und Bürgerparlamente (Effektivität)	0,6	0,7

Wie in Tabelle 9 ersichtlich wird, ist die Entwicklung des Wertes des Variationskoeffizienten zwischen den beiden Delphirunden bei den jeweiligen Items sehr unterschiedlich. Während der in der ersten Runde bestehende Dissens bei den *Bürger-Gesprächsgruppen (Item 24)* von einem Wert des Variationskoeffizienten von $v=0,6$ zur zweiten Delphirunde auf einen konsensualen Wert von $v=0,1$ sinkt, steigt dieser bei wissenschaftlichen Publikationen von einem konsensualen Ergebnis von $v=0$ auf einen Wert von $v=0,6$. Dieser Dissens begründete sich - wie bereits beschrieben - darauf, dass zwei ablehnenden Gruppen wissenschaftliche Publikationen nicht als effektive zielgruppenspezifische Kommunikation an die allgemeine Öffentlichkeit ansahen, während eine Gruppe dies sehr wohl tat. So basierte der Dissens letztlich nicht auf der vorgenommenen Ergänzung des einleitenden Fragetextes, sondern auf der Einigung der Teilnehmenden aus dem Austausch zur Zielgruppe aller Maßnahmen, die sie geschlossen als die *allgemeine Öffentlichkeit* festlegten. Die Effektivität von *Bürgerforen und Bürgerparlamenten (Item 31)* wurde in der ersten Runde mit $v=0,6$ von den Gruppen geschlossen abgelehnt (5/2/1). Diese Verteilung änderte sich zur zweiten Runde des Gruppendelphis, in der die dritte Gruppe diesen Weg der Übermittlung von *wichtigen wissenschaftlichen Studienergebnissen* nun als sehr effektiv (10) bewertete, während die erste und zweite Gruppe diesen Übermittlungsweg weiterhin als ineffektiv beurteilten (1/4), der Wert des Variationskoeffizienten stieg folglich auf $v=0,7$. Die ablehnende Haltung gründete darauf, dass solche *Bürgerforen und Bürgerparlamente* von den ablehnenden Gruppen als sehr teuer und aufwendig eingestuft wurden, so dass der Aufwand im Vergleich zu anderen Übermittlungswegen als zu hoch zu bewerten sei, um diesen Weg als *effektiv* zu beurteilen. Die zustimmende dritte Gruppe sah hier trotz hohem Aufwand ebenso auch eine höhere Glaubwürdigkeit, so dass sie diesen Weg als effektiv einstufte. Hier blieb ein Konsens über einen Dissens bestehen.

Neben der Umformulierung des einleitenden Fragetextes, wurden in dieser Itembatterie ebenso einzelne Items in ihrer Formulierung geändert. So wurden nach der ersten Plenardiskussion sowohl das *Item 23* zur Effektivität des Übermittlungsweges der *Pressekonferenzen* geändert, wie auch das *Item 24* der Effektivität der *Bürger-Gesprächsgruppen*. Hier wurde bei beiden Items der Kommunikationsinhalt der *Messdaten* (t1) durch *Studienergebnisse* (t2) als

zu kommunizierende Inhalte ersetzt. Wie bereits in Bezug auf Tabelle 9 dargestellt, wurde das *Item 24* zur zweiten Runde des Delphis konsensual als effektiver Weg der Übermittlung beurteilt. Wurde sich in der ersten Runde des Gruppendelphis bei *Item 23* von Gruppe 2 enthalten, so stufen in der zweiten Runde des Gruppendelphis alle Gruppen, die in diesem Item abgefragten *Pressekonferenzen* geschlossen als sehr effektiv ein (10), so dass hier der Wert des Variationskoeffizienten zum Zeitpunkt t2 bei $v=0$ lag.

Inwiefern die Umformulierung des einleitenden Fragetextes oder vielmehr die Ersetzung des Begriffes der *Messdaten* durch den Begriff der *Studienergebnisse* bei den *Bürger-Gesprächsgruppen (Item 24)* jeweilig Einfluss auf das konsensuale Ergebnis bei der Beurteilung der Effektivität in der zweiten Runde hatten, lässt sich durch die Synchronität der Umformulierungen zur Abfrage (t2) indes nicht bestimmen.

Bei *Frage 3.1* findet sich zudem ein Beispiel dafür, dass sich auch im zweiten Plenum geschlossen auf eine Präzisierung eines Items geeinigt werden konnte, und so im diskursiven Austausch ein konsensuales Ergebnis ohne erneute Abfrage mittels Fragebogen erreicht werden konnte. So wurde das *Item 29* zur Bewertung der Effektivität von *TV und Radio* als Übermittlungswege in der zweiten Runde bereits von den Teilnehmenden als effektiv eingestuft. Während aber die Gruppen 1 und 3 diesen Weg als sehr effektiv beurteilten (10), bewertete die zweite Gruppe diesen Weg lediglich als eher effektiv (7). Diese Einschätzung wurde von dieser Gruppe damit begründet, dass bei TV und Radio die Gefahr einer unseriösen Übermittlung bestände. Daher könne man diesen Weg nicht pauschal als effektiv einstufen. Die beiden anderen Gruppen begründeten ihr Urteil damit, dass sie Seriosität in ihrem Urteil implizit als gegeben vorausgesetzt hätten. So wurde von Seiten der Teilnehmenden vorgeschlagen, das *Item 29* um den Zusatz *seriös* zu ergänzen. Gruppe 2 folgte daraufhin dem Urteil der anderen Gruppen, diesen Übermittlungsweg als *sehr effektiv* einzustufen.

Die in *Frage 3.1* erfolgten Umformulierungen des Fragetextes wirkten sich ebenso auf *Frage 3.2* zur *Glaubwürdigkeit der Übermittlungswege* aus, da auch hier die Formulierung der *Übermittlung von Messdaten durch wichtige wissenschaftliche Studienergebnisse - die Handlungsempfehlungen und Schlussfolgerungen begründen* ersetzt wurde. Hier finden sich allerdings keine über beide Runden des Workshops vergleichbaren Items, da sich Gruppe 2 in der ersten Kleingruppenphase geschlossen bei der Beurteilung dieser Itematterie enthalten hatte. Dies gilt damit ebenso für *Item 35* und *Item 36*, bei denen wie bereits bei *Item 23* und *Item 24* der Begriff der *Messdaten* durch den *Begriff der Studienergebnisse* ersetzt wurden.

Letztlich wurde ebenso bei *Frage 4* zu *Maßnahmen zur Steigerung der Aufmerksamkeit* in der ersten Plenarsitzung eine Umformulierung von den Teilnehmenden vorgeschlagen. Der ursprüngliche Fragetext lautete: *Im Folgenden finden Sie eine Reihe von Maßnahmen, die auf das Problem der potentiellen Strahlung von elektromagnetischen Feldern aufmerksam machen sollen. Bitte tragen Sie hierfür einen Wert ein. Geben Sie den Wert 1 (-) an, wenn Sie der Ansicht sind, dass sich diese Maßnahme ganz und gar nicht eignet und geben Sie die Ziffer 10 (+) an, wenn sich die Maßnahme voll und ganz eignet.*

Da in der ersten Kleingruppenphase keines der abgefragten vier Items konsensual bewertet wurde, äußerten die Teilnehmenden in der Diskussion im ersten Plenum zu einem frühen Zeitpunkt das Bedürfnis, auch diesen Fragetext zu konkretisieren, da keine Aufmerksamkeit auf

die potentielle Strahlung erzeugt werden müsse, sondern vielmehr auf die potentiellen gesundheitlichen Gefährdungen durch Strahlung hingewiesen werden solle. Diese Umformulierung wurde für den Fragebogen der zweiten Delphirunde übernommen, so dass das Itemset zum Zeitpunkt t2 wie folgt zur Beurteilung der Expert*innen einleitend abgefragt wurde: *Im Folgenden finden Sie eine Reihe von Maßnahmen, die auf das Problem der potentiellen gesundheitlichen Gefährdung durch Strahlung von elektromagnetischen Feldern aufmerksam machen sollen.*

Da sich in der zweiten Kleingruppenphase die erste der Kleingruppen einer Beurteilung des Items 48 zu Labels enthielt, stehen für einen Vergleich die drei in Tabelle 10 aufgeführten Items zu Verfügung. Die Werte des Variationskoeffizienten der Items 47 und 49 deuten darauf hin, dass keine Annäherung der Urteile der Kleingruppe zur zweiten Runde stattgefunden hat. Dies liegt allerdings besonders in den ablehnenden Urteilen der Teilnehmenden begründet, die zu niedrigen Werte des arithmetischen Mittels bei beiden Items führen. Während beispielsweise bei Item 49 zu Messungen in der ersten Runde durch Gruppe 2 noch eine positive Beurteilung stattfand (7), lehnen die Gruppen in der zweiten Runde solche Messungen als geeignete Maßnahme geschlossen ab (1/3/1). Letztlich besteht auch hier ein konsensuales Urteil zur zweiten Runde.

Tabelle 10: Werte des Variationskoeffizienten bei Frage 4 „Maßnahmen zur Steigerung der Aufmerksamkeit“ in SEAWIND

Item #	Frageinhalt	v (t1)	v(t2)
47	Apps	0,6	0,6
49	Messungen	0,7	0,6
50	Messgeräte	0,5	0,0

Dies ist ebenso beim Item 47 zu den Apps der Fall, in der die Kleingruppen zur zweiten Delphirunde diese Maßnahme ebenso geschlossen ablehnen (5/2/1), wenn auch Gruppe 1 die Maßnahme lediglich als eher nicht geeignet abgelehnt. Beim Item 50 zu den Messgeräten ist das Urteil der drei Kleingruppen zur zweiten Runde konsensual, denn sie sind sich in dem Urteil einig, dass sich diese Maßnahme *ganz und gar nicht eignet* (1/1/1). Letztlich waren sich die Teilnehmenden in der zweiten Plenarsitzung darin einig, dass sich keine der Maßnahmen zur Steigerung der Aufmerksamkeit eignet, wenn es darum ginge, die *potentielle gesundheitliche Gefährdung durch Strahlung* zu vermitteln.

Bei den Auswahlfragen fand, wie in Kapitel 7.2.3.1 dargestellt, bei Frage 2.2 beim Item zur Identifikation von *Unsicherheitsquellen in der Risikobewertung* eine Umformulierung statt. So teilten die Teilnehmenden, das von Gruppe 2 im ersten Plenum vorgebrachte Argument, dass jegliche *Unsicherheit in der Risikobewertung* identifiziert werden müsse, so dass das Item im Fragebogen der zweiten Runde lautete: *Unsicherheit in der Risikobewertung identifizieren*. Diese Umformulierung führte letztlich dazu, dass dieses Ziel nach Urteil der Teilnehmenden zum Ende des Gruppendelphis zu einem der drei wichtigsten Kommunikationsziele gezählt wurde.

Letztlich fällt bei diesem Fallbeispiel die ungewöhnlich hohe Zahl an Umformulierungen an den Frage- sowie Itemformulierungen auf. So wurden nicht nur vereinzelt Items, sondern eine

hohe Zahl an einleitenden Fragetexten zwischen den beiden Runden umformuliert. Dies begründet sich weniger in semantischen Missverständnissen, sondern an, aus Sicht der Teilnehmenden, nicht klar formulierten Fragen, Akteuren und Maßnahmen.

7.2.3.3 Dimension: normative Angemessenheit

In der Dimension der normativen Angemessenheit liegt der Fokus auf den von den Teilnehmenden geäußerten Begründungen während der Plenarphasen und ob diese Phasen die Möglichkeit der Offenlegung von normativen Einstellungen und Wertvorstellungen bieten. Auch im Falle des SEAWIND-Gruppendelphis werden hierfür die vier Kategorien angewandt, die die geäußerten Aussagen in (1) Werturteile, (2) Präferenzen, (3) (individuelle) Bewertungen von Optionen und (4) moralische Urteile einordnen. Für diese Analyse stehen beim Gruppendelphi in SEAWIND zum einen das Ergebnisprotokoll des Gruppendelphis⁸⁵ sowie die Veröffentlichung von Hilpert et al. (vgl. ebd. 2013: 11-32) zur Verfügung.

Im Fallbeispiel des SEAWIND-Gruppendelphis, finden sich in der Dokumentation der Plenarphasen wenige Beispiele, in denen die Expert*innen für ihre Begründungen Argumente der normativen Dimension heranzogen.⁸⁶ Als erstes Item, bei dem im ersten Plenum normative Argumente von den Expert*innen artikuliert wurden, ist das *Item 13: Die Datengrundlage zur Festlegung der Risikogruppe*. Die Diskussion zu diesem Item wurde bereits im vorhergehenden Abschnitt zur Verständlichkeit näher erläutert, soll hier allerdings nochmals mit spezifischem Blick auf die normative Dimension beleuchtet werden. Das *Item 13* führte in der ersten Kleingruppenphase zu einer sehr unterschiedlichen Bewertung durch die Teilnehmenden in der die Gruppe 1 und Gruppe 3 dieses Item als Kommunikationsinhalt einstufen, der *unbedingt kommuniziert werden sollte*; Gruppe 2 hingegen beurteilte die Datengrundlage zur *Festlegung der Risikogruppe* als Kommunikationsinhalt, der *ganz und gar nicht kommuniziert werden sollte*. Während einige Expert*innen darauf hinwiesen, dass die Daten zu Missbrauch führen könnten, da dies bereits an anderer Stelle im Falle von Rohdaten geschehen sei, argumentierte ein Experte mit einer anderen Bewertung dieser Option, nicht die Rohdaten zu veröffentlichen, sondern zu erklären, wie die Bewertungen zustande gekommen wären. Diese Bewertung der Option gründete er auf dem Werturteil, dass der gesamte Bewertungsprozess transparent zu gestalten sei. Dem Argument der Transparenz stimmte eine Mehrzahl der Expertinnen und Experten zu, auch wenn die Bewertungen darüber auseinandergingen, dass für einen transparenten Bewertungsprozess auch eine Veröffentlichung der Rohdaten unumgänglich sei oder wie von dem Experten vorgeschlagen, nur lediglich zu vermitteln sei, wie die Bewertung zustande gekommen wäre. Neben diesen beiden Positionen, die dem Argument der Transparenz hohes Gewicht einräumten, folgten einzelne Expert*innen dem Argument der Gefahr eines möglichen Missbrauchs der Rohdaten. Hier wurde von einem Experten auf *unzählige Fälle* verwiesen, bei denen dies bereits geschehen sei und evidenzbasiert auf das normative Ziel der Transparenz reagiert. Alle Positionen bezogen sich in ihren Erwiderungen zwar auf die jeweils anderen Argumente, allerdings ließen sich die gegensätzlichen Positionen nicht auflösen, so dass letztlich ein Konsens über einen Dissens darüber bestand, dass die

⁸⁵ Das Protokoll findet sich auf der beigelegten Daten CD.

⁸⁶ Eine Verdichtung der kategorisierten Inhalte finden sich im Anhang 11.8.

Gefahr des Missbrauchs dem Argument der Transparenz entgegenstünde. Die Position, die die Expert*innen einnahmen, hing letztlich von der eigenen individuellen Bewertung ab.

Im ersten Plenum gab es eine Diskussion der Expert*innen darüber, ob die Einschätzung der *International Agency for Research on Cancer (IARC)* - wie im einleitenden Text der *Frage 1.1* vorgesehen - überhaupt kommuniziert werden sollte. In diesem ersten Plenum wurde entschieden, dies zu diesem Zeitpunkt nicht weiter zu diskutieren, sondern die Frage als neues *Item 80* der *Frage 1.1* vorrauszustellen: *Ist es Ihrer Meinung nach notwendig, die Einschätzung der IARC in die breite Öffentlichkeit zu kommunizieren?* Dieses Item wurde in der zweiten Kleingruppenphase in den Fragebogen integriert. Während sich dabei Gruppe 1 enthielt, sah es Gruppe 2 als nicht notwendig an (2) die Einschätzung der IARC zu kommunizieren, während Gruppe 3 dies als unbedingt notwendig ansah (10). Dieses Item wurde zu Beginn des zweiten Plenums näher diskutiert. Die zustimmende Gruppe 3 begründete ihre Zustimmung mit der Relevanz der IARC als Autorität auf diesem Gebiet und der großen Relevanz der Nachricht. Auch wenn die Inhalte auf die jeweiligen Rezipienten abgestimmt werden müssten, sei es aus diesen Gründen notwendig, die Einschätzung bereits früh zu kommunizieren. Die ablehnende Gruppe 2, die sich gegen eine aktive Kommunikation aussprach, begründete dies mit dem moralischen Urteil, dass eine solche Kommunikation der Einschätzung potentieller Gefahren nur die jeweiligen Zielgruppen verunsichern würde. Überdies führte diese Gruppe der ablehnenden Expert*innen die individuelle Bewertung an, dass man den meisten Menschen die Relevanz der Einstufung nicht vermitteln könne. Auch wenn die Teilnehmenden beide Positionen als legitim anerkannten, konnte dieser Dissens nicht aufgelöst werden. Ebenso die sich in der zweiten Kleingruppenphase enthaltende Gruppe 1 begründete ihre Enthaltung auf diesen zwei gegensätzlichen Positionen und fasste diese in zwei Leitsätzen zusammen:

- *Right to know*: Lediglich eine reaktive Kommunikation der Einschätzung der IARC.
- *Need to know*: Unbedingt eine aktive Kommunikation der Einschätzung der IARC.

Auch wenn beide Argumentationsstränge von den jeweils anderen Positionen als legitimer Geltungsanspruch anerkannt wurden, fanden keine anknüpfenden Erwiderungen statt. Die Positionen wurden von den Teilnehmenden als so diametral wahrgenommen, dass ein Teilnehmer letztlich für eine Abstimmung plädierte, die von den anderen Expert*innen angenommen wurde: *Dabei sprachen sich 10 Teilnehmer für eine reaktive Kommunikation (right to know) und 4 Teilnehmer für eine aktive Kommunikation (need to know) aus.*

Die Diskussion dieser verschiedenen Einschätzungen und Präferenzen setzte sich allerdings auch nach der Abstimmung im Plenum bei der offenen *Frage 5* zur Kommunikation bei hoher Unsicherheit fort: *Wie sollen Ergebnisse von Strahlungsuntersuchungen kommuniziert werden, die mit hoher Unsicherheit behaftet sind und bei welchen bisher keine verlässlichen Aussagen über gesundheitliche Konsequenzen möglich sind?* Aufgrund von Zeitmangel, konnte keine der jeweiligen Kleingruppen in beiden Kleingruppenphasen schriftlich antworten. Allerdings wurde diese Frage am Ende der zweiten Plenarphase mit den Teilnehmenden diskutiert. So wurde auch hier das Werturteil als Argument angeführt, dass es entscheidend sei Vertrauen und Transparenz zu schaffen und einen offenen Diskurs über die Chancen und Risiken zu führen. Damit wäre auch eine aktive Kommunikation unter Unsicherheit zu wählen - *need to know*. Ein Experte der Gegenposition argumentierte allerdings mit der individuellen Einschätzung, dass aber die Einstufung der IARC keinen Mehrwert darstelle und Risiken vielmehr

vorher geprüft werden sollten, ob die Meldung darüber nicht gar kontraproduktiv wirken könnte - *right to know*. Die Präferenz des Teilnehmers bestand darin, erst die Frage nach dem Mehrwert hinter einer Meldung zu klären. Beide bestehenden Positionen konnten auch in der Diskussion um *Frage 5* nicht aufgelöst werden. War es allerdings bei dem diskursiven Austausch um die Bewertungen bei *Item 80* noch so, dass sich anknüpfende Erwidierungen auf die jeweiligen Argumente bezogen, so kann dies für den Austausch bei *Frage 5* nicht mehr beobachtet werden. Vielmehr wurden die jeweiligen Positionen mehrfach argumentativ bekräftigt ohne Bezug zur anderen Position herzustellen.

7.2.3.4 Dimension: grundlegende Funktionslogik

Das Gruppendelphi in SEAWIND bietet anhand der Kategorie *Sonstige* bei den jeweiligen Fragen der ersten Runde die Möglichkeit, auch Items in der Analyse zu betrachten, die nicht dem typischen Ablauf eines Gruppendelphis folgen. So zeichnen sich diese Items durch drei Aspekte aus, die bei der Abfrage in einem Gruppendelphi idealiter nicht zu beobachten sind:

1. Die Formulierung der Items wurde von den Teilnehmenden im Fragebogen der ersten Runde durchgeführt.⁸⁷
2. Die Items wurden erst im Plenum (t1) diskutiert, bevor sie in der Kleingruppenphase (t2) abgefragt wurden.
3. Es fand keine erneute Abfrage (auch bei Dissens) statt.

Neben dem bereits im vorhergehenden Abschnitt zur normativen Angemessenheit⁸⁸ besprochenem *Item 80*, wurden im Gruppendelphi sechs weitere Items durch die Teilnehmenden vorgeschlagen und erst zur zweiten Kleingruppenphase abgefragt:

- *Item 81: Berufsständische Organisationen im Gesundheitsbereich (Enthaltung)*
- *Item 82: Fachgesellschaften*
- *Item 83: „gestufte“ Information der zuständigen Behörden in Print und Internet*
- *Item 84: Netzbetreiber*
- *Item 85: Fachbezogene Vereine/Organisationen*
- *Item 87: Eigene Websites*

Bei der ersten Betrachtung der Urteile der jeweiligen Kleingruppen, fällt bei diesen Items auf, dass es eine Vielzahl von Enthaltungen gibt. So wurde sich bei den *Items 80, 81* und *83* jeweils von einer Gruppe enthalten, während sich bei *Item 87* sogar zwei Gruppen enthalten haben.⁸⁹ Da die Enthaltungen bei bloßer Betrachtung des Variationskoeffizienten zu einer falschen Auslegung eines konsensualen Ergebnisses zwischen allen drei Gruppen führen könnten, sind diese in Abbildung 31 schwarz dargestellt. Wie anhand der jeweiligen Werte des Variationskoeffizienten erkennbar, wurden auch die von allen drei Gruppen bewerteten Items (*82, 84, 85*) nicht durchweg konsensual beurteilt. Diese drei Items wurden geschlossen unter der Ka-

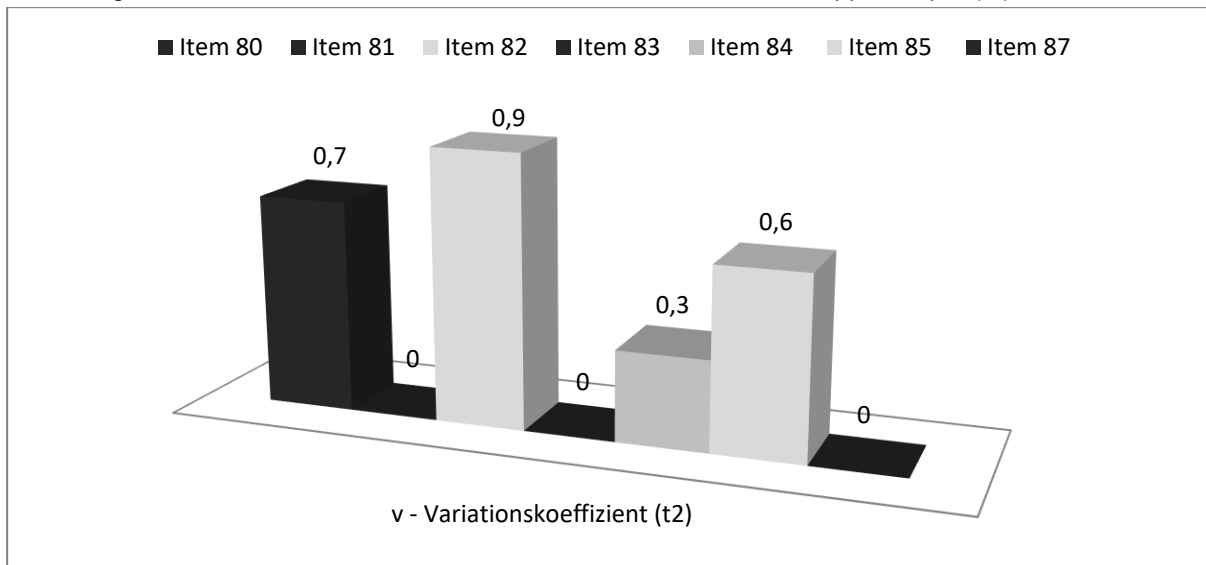
⁸⁷ Dadurch, dass im SEAWIND-Gruppendelphi als Items keine Aussagen formuliert wurden, sondern direkte Abfragen durch Nennung von Akteuren und Maßnahmen stattfand, konnten die Teilnehmenden ihre eigenen Items auf einfachem Wege durch bloße Nennung formulieren.

⁸⁸ siehe Kapitel 7.2.3.3.

⁸⁹ Gruppe 1 enthielt sich bei den *Items 80, 83* und *87*, während sich Gruppe 2 bei den *Items 81* und *87* enthalten hat.

tegorie *Sonstige* in der ersten Runde des Gruppendelphis als weitere Akteursgruppen vorgeschlagen, die eine Rolle bei der Information der Öffentlichkeit bezüglich der *Einschätzung des Krebsrisikos durch elektromagnetische Felder* übernehmen sollen.

Abbildung 31: Variationskoeffizient bei neuen Items im SEAWIND-Gruppendelphi (t2)



Anmerkungen: tn2=3

Quelle: Eigene Darstellung

Von diesen drei Items wurde lediglich das *Item 84* mit einem Wert des Variationskoeffizienten von $v=0,3$ durch die drei Gruppen konsensual bewertet, wenn auch durch Ablehnung dieser abgefragten Akteursgruppe (*Netzbetreiber*) als relevanten Akteur (1/2/2). Sowohl bei *Item 82* zu *Fachgesellschaften* als auch bei *Item 85* zu *Fachbezogene Vereine und Organisationen* lagen die Einschätzungen der Kleingruppen auseinander. Bei *Item 82* bewertete Gruppe 1 dieses Item ablehnend mit einem Wert von 1, wie auch Gruppe 2 ein ablehnendes Urteil mit einem Wert von 2 fällte, während hingegen Gruppe 3 dieses Item mit einem Wert von 10 als sehr relevant bewertete. Der Wert des Variationskoeffizienten für dieses Item lag folglich mit $v=0,9$ sehr hoch. Bei den *fachbezogenen Vereinen und Organisationen (Item 85)* verteilen sich die einzelnen Bewertungen weitflächig auf der 10er-Skala: So bewertete die erste Gruppe das Item mit voller Zustimmung hinsichtlich der Relevanz dieser Akteursgruppe (10), während das Urteil von Gruppe 2 im unteren Bereich der Skala bei einem Wert von 2 lag und letztlich Gruppe 3 die leichte Ablehnung dieser Akteursgruppe kennzeichnete (5). Letztlich führen die unter der Kategorie *Sonstige* durch die Teilnehmenden vorgeschlagenen Items nicht zu zufriedenstellenden (bzw. *fruchtbaren*) Ergebnissen. So werden zwar weitere Sachverhalte - oder konkreter Akteursgruppen - vorgeschlagen, die hinsichtlich der Fragestellung einer *Kommunikation des Krebsrisikos* bedacht werden sollten, allerdings können diese aufgrund der unterschiedlichen Einschätzungen nicht als relevante Akteure in die Ergebnisse mit einfließen, auf die sich konsensual verständigt werden konnte. Ebenso besteht auch bei den von den Teilnehmenden vorgeschlagenen Items in Runde 1 das Problem einer hohen Zahl an Enthaltungen einzelner Gruppen in der zweiten Kleingruppenphase.

Im Falle des SEAWIND-Gruppendelphis tritt diese hohe Zahl der Enthaltungen nicht nur in dieser Kategorie auf, sondern zeigt sich - wie bereits erläutert - über alle Abschnitte des Fragebogens hinweg. So enthielt sich die zweite Kleingruppe in der ersten Kleingruppenphase bei in der Summe 28 Items und auch in der zweiten Kleingruppenphase ist die Summe von 13

Enthaltungen relativ hoch, auch wenn diese sich auf drei Kleingruppen verteilt: Gruppe 1, 8 Enthaltungen; Gruppe 2, 4 Enthaltungen; Gruppe 3, eine Enthaltung. Dies stellt sich, wie besonders die Unterkapitel der Evidenz⁹⁰ und der Verständlichkeit⁹¹ bereits aufgezeigt haben, besonders als Problem für die Interpretation der Ergebnisse dar, denn um einen Konsens oder einen Dissens über einen Dissens anhand der Ergebnisse der Befragungsrunden herauszustellen, braucht es eine geschlossene Beurteilungen aller drei Kleingruppen. Überdies sollte dieses Urteil sowohl geschlossen zur ersten Runde sowie auch zur zweiten Runde erfolgen. Da die Kleingruppen zur zweiten Runde neu zusammengesetzt werden, kann sich so die *Kultur der Enthaltung* aus einer Gruppe (t1) in die folgende Kleingruppenphase (t2) auf die anderen Kleingruppen übertragen, wie es bei diesem Fall zu vermuten ist.⁹²

Eine Begründung für die vielen Enthaltungen in der ersten Kleingruppenphase wurde von Gruppe 2 in der ersten Plenarsitzung geäußert: So wurde bei bestehender Uneinigkeit zu einzelnen Items in der Gruppe bei diesen kein Urteil im Fragebogen abgegeben. Bei Dissens innerhalb der Gruppe folgte eine Enthaltung. Dies liegt überdies darin begründet, dass in der Einführung des Gruppendelphis am Workshoptag nicht darauf hingewiesen wurde, dass bei Uneinigkeit bei der Beurteilung einzelner Items auch die Möglichkeit besteht, Minderheitsvoten im Fragebogen zu vermerken. Auch wenn von der Moderation in der ersten Plenarsitzung mit Vermerk auf die zweite Kleingruppenphase auf die mögliche Verwendung von Minderheitsvoten hingewiesen wurde, findet sich ebenso in der zweiten Runde mit 13 Enthaltungen eine ungewöhnlich hohe Zahl dieser wieder. Neben einem Gewöhnungseffekt (*Kultur der Enthaltung*), der von den Teilnehmenden in Gruppe 2 (t1) zur zweiten Runde in alle Gruppen getragen wurde, kann aber auch die knappe Zeit einen Erklärungsansatz bieten. Denn die knappe Zeitspanne, um sich auf ein gemeinsames Urteil zu einigen, machte die Wahl einer zeitsparenden Enthaltung bei einzelnen Items für die Teilnehmenden attraktiv.

Darauf, dass die Zeit während der Kleingruppenphasen zu knapp bemessen war, deutet auch die fehlende Beantwortung der letzten Frage im Fragebogen durch alle Gruppen zu allen Phasen des Gruppendelphis hin. Diese *Frage 5 zur Kommunikation bei hoher Unsicherheit* bildete den Abschluss des Fragebogens und war als offene Frage konzipiert: *Wie sollen Ergebnisse von Strahlungsuntersuchungen kommuniziert werden, die mit hoher Unsicherheit behaftet sind und bei welchen bisher keine verlässlichen Aussagen über gesundheitliche Konsequenzen möglich sind?* Die Frage lieferte letztlich keine Ergebnisse, da sie aufgrund fehlender Beantwortung auch nicht gänzlich in den beiden Plenarsitzungen besprochen werden konnte.

Das Gruppendelphi in SEAWIND lieferte zwar für viele der Fragestellungen des Gesamtprojektes zufriedenstellende Ergebnisse, wie beispielsweise in *Frage 2.1 zu Kommunikationsinhalten* oder *Frage 1.1* mit der konsensualen Benennung der relevanten *Akteursgruppen*, die für eine Information der Öffentlichkeit relevant wären. Allerdings bleiben durch die hohe Zahl an Enthaltungen und der fehlenden Beantwortung von *Frage 5* auch einige Bewertungen offen,

⁹⁰ siehe Kapitel 7.2.3.1.

⁹¹ siehe Kapitel 7.2.3.2.

⁹² Da die Kleingruppenphase (wie konzeptionell vorgesehen) in geschlossener Runde der Teilnehmenden stattfindet, kann dies nicht anhand von Aufzeichnungen belegt, sondern bloß konstatiert werden.

anders wie es im Falle des Gruppendelphis in SAUBER+ der Fall war. Die generelle Funktionslogik eines Gruppendelphis kann im Falle von SEAWIND als nur bedingt erfüllt angesehen werden.

7.2.4 Zwischenfazit: Gruppendelphi in SEAWIND

Wie bereits in den vorhergehenden Unterkapiteln beschrieben, weist der Fall des Gruppendelphis in SEAWIND einige problematische Bereiche auf, die von dem im Ideal angenommenen Ablauf abweichen. Auf die hohe Anzahl der Enthaltungen durch die Kleingruppen wurde bereits eingegangen, wie auch die zu knappe Zeit bzw. der zu umfangreiche Fragebogen thematisiert wurden.

Dieser Fall weist überdies auch inhaltlich einige problematische Bereiche auf, die für die Konzeption eines Fragebogens für ein Gruppendelphi wichtige Hinweise liefern. Zum einen wurde in der ersten Frage die *Einschätzung des Krebsrisikos durch elektromagnetische Felder* durch die *International Agency for Research on Cancer (IARC)* als *zu kommunizierender Inhalt* vorgegeben. Dies führte zwar in der Vorbefragung zu keinen Kommentaren durch die teilnehmenden Expert*innen, wurde aber von einigen Teilnehmenden in der ersten Runde des Gruppendelphis in Frage gestellt. Dies führt zum einen zu der Erkenntnis, dass nicht alle problematischen Formulierungen oder weiter auch Inhalte des Fragebogens durch die Vorbefragung aufgedeckt werden konnten. Des Weiteren ist hier aber auch ein Effekt auf die inhaltliche Rahmung einzelner Fragen und die des gesamten Fragebogens erkennbar. Das Forscherteam reagierte auf die Kritik an der IARC-Einschätzung in Absprache mit den Teilnehmenden im ersten Plenum darauf, dass eine Filterfrage zu Beginn des zweiten Fragebogens integriert wurde, die das Urteil der Expert*innen dazu abfragte, ob eine *aktive Kommunikation der Einschätzung der IARC in die breite Öffentlichkeit sinnvoll sei (Item 80)*. Dies konnte dazu beitragen, dass die weiteren Fragen des Fragebogens unter einem von den Teilnehmenden individuell getroffenen Bewertungshorizont bewertet werden konnten. So wurden beispielsweise die Items in *Frage 1.1* und *Frage 2.1* zur zweiten Runde konsensual bewertet und führten zu brauchbaren Ergebnissen, sie bezogen sich thematisch allerdings beide in der Frageformulierung auf die jeweilige Einschätzung der IARC. Da das *Item 80* in der zweiten Delphirunde mit einem Variationskoeffizienten von $v=0,7$ (bei $n=2$) nicht konsensual bewertet wurde, hatten die Teilnehmenden durchaus verschiedene Bewertungshorizonte für die folgenden Fragen (mal mit Zustimmung des Kommunikationsinhaltes, mal ohne, mal mit Uneinigkeit innerhalb der Kleingruppe). Der bestehende Dissens in der Beurteilung der Einschätzung der IARC als sinnvollem Kommunikationsinhalt für die Öffentlichkeit, konnte auch im Austausch im zweiten Plenum nicht aufgelöst werden. Letztlich bestanden einzelne Teilnehmende darauf eine Abstimmung durchzuführen, um *klare Zahlen* des Zuspruchs bzw. der Ablehnung zu erhalten. Dabei sprachen sich 10 Teilnehmende für eine reaktive Kommunikation (*right to know*) und 4 Teilnehmende für eine aktive Kommunikation (*need to know*) der Einschätzung von elektromagnetischen Feldern als *möglicherweise krebserregend* aus. Der Wunsch nach einer Abstimmung, die klare Mehrheitsverhältnisse geschaffen hat, spiegelt gewissermaßen die Einstellung einzelner Expert*innen zu diesem Workshop wider. So wurde der *Konsensgedanke* nicht vom gesamten Teilnehmendenfeld angenommen. Dies erschwerte den Austausch von Argumenten für die jeweiligen Positionen, denn diese wurden von einigen Teilnehmenden entweder als

wahr oder *falsch* bewertet, sogar, wenn sich andere evidenzbasierte Argumentationen auf andere wissenschaftliche Studien beriefen.

Wertgeleitete Argumentationsmuster wurden im Falle von SEAWIND kaum vorgetragen und wenn dies doch der Fall war, so wurden sie (häufig) nicht als legitim erachtet. Dies zeigt sich daran, dass die anknüpfenden Argumente wieder mit Bezug zu wissenschaftlichen Daten oder Studien dem Kriterium der Evidenz folgten und der Anspruch nach *Wahrheit* dem der *normativen Richtigkeit* vorgezogen wurde. So wurde beispielsweise eine *effektive Kommunikation von Risiken* von einem Teil der Expert*innen als eine dem *Gemeinwohl* zuträgliche Kommunikation gleichgesetzt.

Überdies verlagerte sich der Diskurs aufgrund folgender evidenzbasierter oder nach Verständnis zielender Argumente zügig wieder in eine andere Richtung. Ob dies auch durch die Formulierung der Items als direkte Abfrage der Sachverhalte und nicht, wie idealiter vorgesehen, anhand von *Sollensaussagen* befördert wurde, ist nach Betrachtung dieses einzelnen Falles (noch) nicht abschließend zu konstatieren. Nach Analyse der Diskursinhalte der beiden Plenarsitzungen fällt allerdings auf, dass die eingesetzten Items weniger (besonders normativen) Spielraum für einen Diskurs boten, als es beispielsweise im Falle des Gruppendelphis in SAUBER+ der Fall war. Befördert wurde diese Tendenz überdies durch die starke Dominanz einzelner teilnehmenden Personen, die nicht nur der Einschätzung der *International Agency for Research on Cancer* kritisch gegenüberstanden, sondern auf andere *sinnvollere* Studien verwiesen und sich zwar auf die Argumente anderer Teilnehmenden bezogen, diese aber meist mit der Erwidern entgegneten, diese seien *falsch*.

In diesem Fall findet sich zudem ein Beispiel dafür, dass auch der Fragebogen zu folgenden Kleingruppenphasen mit Bedacht angepasst werden muss. Dies stellt sich als eine große Herausforderung dar, da dieser während der Plenarsitzungen in Absprache mit den Teilnehmenden angepasst, erweitert, an anderen Stellen gekürzt und bei Bedarf umformuliert werden sollte. So besteht die Gefahr, dass aufgrund der vielen kontinuierlichen Änderungen während des laufenden Austauschs im Plenum Fehler bei der Konstruktion des Fragebogens unterlaufen. Ein solcher Fehler lässt sich nach der Auswertung der beiden Runden (t1/t2) und der Protokolle der Plenarsitzungen (t1/t2) bei der Auswahlfrage *Frage 1.2* über die *Zusammenarbeit von Akteuren* identifizieren. Hier wurden in der Diskussion im Plenum bereits in der ersten Runde zwei Akteursgruppen konsensual als wichtige Akteure identifiziert, die aus Sicht der Teilnehmenden zusammenarbeiten sollten: Nationale Behörden und Verbraucherschutzorganisationen. Überdies waren sich die Teilnehmenden nach einem Austausch im ersten Plenum darüber einig, dass ein *Tandem* aus Wissenschaftler*innen und Journalist*innen ebenso als Gruppe zusammenarbeiten sollte. Dieses Tandem wurde auch als solches Tandem in den Fragebogen der zweiten Kleingruppenphase in die Auswahlmöglichkeiten dieser Frage aufgenommen: *Tandem aus Wissenschaftlern und Journalisten*. Da im Fragebogen ebenso sowohl *Wissenschaftler* als auch *Journalisten* jeweils als einzelne Items abgefragt wurden, bestand im Fragebogen zur zweiten Runde eine Überschneidung zwischen einzelnen Auswahlitems und damit letztlich eine fehlende Trennschärfe. So wählten die Kleingruppen, wie in Kapitel 7.2.3.1 bereits dargestellt, teilweise die einzeln abgefragten Items oder aber das Tandem. Letztlich führte hier die fehlende Trennschärfe zu einem Dissens bei der Betrachtung der Verteilung der Auswahlkategorien, der letztlich aber inhaltlich keiner war. Denn nach Verständni-

gung im zweiten Plenum wurden sowohl die Wissenschaftler*innen als auch die Journalist*innen von den Teilnehmenden konsensual als die Akteursgruppen eingestuft, die neben den zwei weiteren Akteursgruppen zusammenarbeiten sollten. Dieses Ergebnis hätte, abgesehen von der hervorgerufenen Trennschärfeproblematik im zweiten Delphifragebogen, überdies auch nach dem ersten Plenum bereits als Konsens entschieden werden können, da bereits zu diesem Zeitpunkt alle Teilnehmenden diese Gruppen als relevante Gruppen für eine Zusammenarbeit einstufen. Betrachtet man den Umstand der knappen Ressource Zeit, besonders bei diesem Gruppendelphi, bei dem die offene *Frage 5* am Ende des Fragebogens zu keiner der Runden bearbeitet werden konnte, so hätte diese *Frage 1.2* auch nach der ersten Plenarsitzung bereits aus dem Fragebogen für die zweite Gruppendelphirunde entfernt werden können. Eine Abfrage durch die Moderation im ersten Plenum, ob sich alle Teilnehmenden auf diese vier Akteursgruppen verständigen können, hätte letztlich zu einem identischen Ergebnis geführt. Eine Validierung des Konsenses über eine weitere Runde während eines Gruppendelphis ist in der Idealkonzeption nicht vorgesehen, und dieser, wenn auch in der Konstruktion des Fragebogens fehlerbehaftete Fall zeigt auf, dass dies aus Gründen der Zeitersparnis nicht das prioritäre Ziel während eines Gruppendelphis sein sollte.

Für die Validierung der Ergebnisse und der Aufdeckung eines Scheinkonsenses bestand die Möglichkeit der Versendung eines Ergebnisprotokolls mit der Bitte um Rückmeldung. Im Falle von SEAWIND wurde den Teilnehmenden ein sechs seitiges Ergebnisprotokoll zugesandt. Die Teilnehmenden hatten keine Korrekturwünsche oder Ergänzungen, so dass die Ergebnisse des Gruppendelphis auf diesem Wege nochmals rückgesichert werden konnten.

7.3 SMART METERING

Projekt: *Soziale, öffentliche, ökonomische und politische Akzeptanz von Fördermaßnahmen zur Verbesserung der Effizienz im Bereich Elektrizitätsnutzung und von nachhaltiger Mobilität bei privaten Haushalten – Teilprojekt Elektrizitätsnutzung: Smart Metering*⁹³

Institut: *ZIRN – Interdisziplinärer Forschungsschwerpunkt Risiko und Nachhaltige Technologieentwicklung am Internationalen Zentrum für Kultur- und Technikforschung der Universität Stuttgart*

Die Zielsetzung der Studie zu Smart Metern war es, im Rahmen der *nationalen Klimaschutzinitiative* geeignete Empfehlungen für das *Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU)*⁹⁴ auszuarbeiten. Das Projekt verband dabei umweltpsychologische Fragestellungen mit Fragestellungen nach politischen Rahmenbedingungen, um im Bereich des Verbraucher*innenverhaltens einen Beitrag zur Weiterentwicklung der nationalen Klimaschutzinitiative zu leisten (vgl. Mack/Tampe-Mai 2012: 69). Die Fragestellung des Teilprojektes war es, „(...) *durch welche Kombination von Maßnahmen die Verbreitung von Smart Metern sehr gut gefördert und deren Wirksamkeit im Hinblick auf die Erhöhung der Energieeffizienz in Haushalten optimiert werden kann. In der Studie wurde zum einen untersucht, wie Smart Meter*

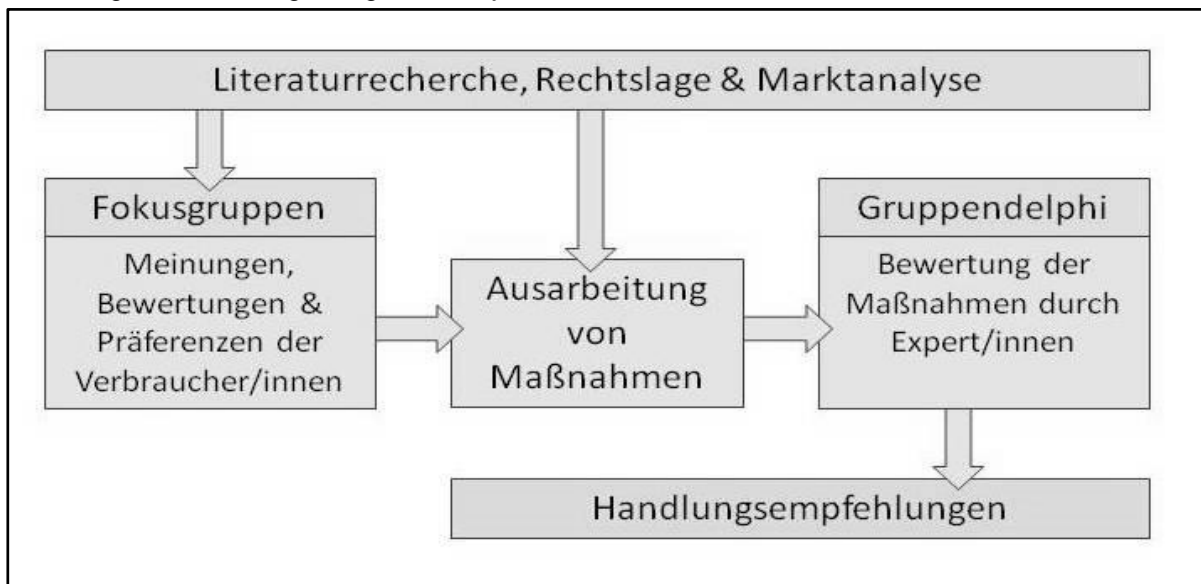
⁹³ „Smart Meter sind digitale elektronische Stromzähler, die Feedback zum Stromverbrauch eines Haushalts über Home Displays, Webportale oder Smartphone-APPs geben können“ (Renn et al. 2012a: 7).

⁹⁴ Das Bundesministerium wurde zwischenzeitlich umbenannt. Es wird derzeit als Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit bezeichnet. Quelle: <https://www.bmu.de/> (Zugegriffen am 28.03.2019)

ergänzende Feedbacksysteme ausgestaltet werden müssen, um VerbraucherInnen optimal beim Stromsparen zu unterstützen und Verhaltensänderungen hin zu mehr Energieeffizienz zu fördern. Zum anderen wurde erfasst, welche Rahmenbedingungen und begleitenden Maßnahmen zur Verbreitung und Förderung der tatsächlichen Nutzung solcher Systeme bereitgestellt werden sollten“ (Kuhn et al. 2014: 63).

Ausschlaggebend für die Wahl der Themenstellung zu *Smart Metering* war die Aktualität zum Zeitpunkt der Teilprojektdurchführung, da die Bundesregierung mit dem *Gesetz zur Öffnung des Messwesens bei Strom und Gas für den Wettbewerb* im August 2008 die EDL-Richtlinie⁹⁵ des Europäischen Parlaments und des Rates über *Endenergieeffizienz und Energiedienstleistungen* umsetzte. Um sowohl die Transparenz des Stromverbrauchs zu erhöhen, als auch eine effiziente Stromnutzung der Verbraucher*innen zu fördern, sollten mit dem Gesetz die Einführung intelligenter Zähler und technische Innovationen am Zähl- und Messwesen gefördert werden (vgl. Renn et al. 2012a: 4).

Abbildung 32: Forschungsdesign des Projektes zu Smart Metern



Quelle: Eigene Darstellung nach Kuhn et al. 2014: 63

Das Forschungsdesign des Teilprojektes gliederte sich dabei in verschiedene methodische Elemente auf, welches neben einer Literaturrecherche, einer Prüfung der Rechtslage sowie einer Marktanalyse, auch die Durchführung von Fokusgruppen und eines Gruppendelphis beinhaltete. Abbildung 32 zeigt den Ablauf des Forschungsdesigns als auch die Ausarbeitung der im Gruppendelphi abgefragten Maßnahmen basierend auf den vorhergehenden Schritten.

7.3.1 Fragestellung des Gruppendelphis

Während in den Fokusgruppen prioritär Meinungen, Bewertungen und Präferenzen potentieller Nutzer*innen ermittelt wurden, verfolgte das Gruppendelphi das Ziel, Maßnahmenbewertungen der konkreten Ausgestaltung solcher Smart-Metering-Systeme durch Expert*innen zu erheben. Dies erfolgte insbesondere im Hinblick auf die politische, technische sowie finanzielle

⁹⁵ Richtlinie 2006/32/EG.

Machbarkeit einzelner Maßnahmen. Wie sollten also Feedbacksysteme ausgestaltet sein, um eine effiziente Stromnutzung bei Verbraucher*innen mittels Smart Meter zu fördern und welche Kommunikationsmaßnahmen zur Einführung und Verbreitung von Smart Metern sollten verfolgt werden? Das Forschungsteam verfolgte mit dem Gruppendelphi das Ziel, am Ende einen Maßnahmenkatalog sowie konkrete politische Handlungsempfehlungen auf Basis der Expert*innenurteile zu formulieren (vgl. Kuhn et al. 2014: 64-70).

7.3.2 Merkmale des Gruppendelphis

Im Rahmen des Teilprojektes *Elektrizitätsnutzung* wurde im Vorfeld des Gruppendelphis keine Vorbefragung der rekrutierten Expert*innen durchgeführt, da auf die zentralen Ergebnisse der Literaturanalyse, der Marktanalyse sowie ergänzender Expert*inneninterviews und der Fokusgruppen aufgebaut werden konnte. Diese Ergebnisse wurden in Form von Maßnahmen für den im Gruppendelphi eingesetzten Fragebogen aufbereitet (vgl. Renn et al. 2012a: 16).

Im Hinblick auf die vorhergehenden Analysen, wurden für das Gruppendelphi Teilnehmende rekrutiert, die Wissen und Erfahrungen im Bereich der Akzeptanz und Effektivität von Smart Metern und ein breites Spektrum an Meinungen und entscheidungsrelevantes Wissen hinsichtlich dieses Sachverhaltes abdecken konnten. Insgesamt nahmen 16 Expert*innen an dem Delphi teil, so dass sich das Feld der Teilnehmenden aus Vertreter*innen der Energiewirtschaft, Forschung und Wissenschaft, Politik sowie Anbietern von Smart Meter Lösungen, Dienstleistern und Umweltorganisationen zusammensetzte (vgl. Renn et al. 2012a: 16f).

Die Teilnehmenden diskutierten in vier Kleingruppen mit jeweils vier Personen. Zur zweiten Runde folgten mittels Permutation neu zugeteilte Kleingruppen. Die Plenardiskussion wurde mittels handschriftlicher Notizen durch das Forscher*innenteam notiert und nach Durchführung des Gruppendelphis wurden die statistischen Ergebnisse mit diesen Anmerkungen für die weitere Auswertung ergänzt. Der eingesetzte Fragebogen unterteilte sich in acht Abschnitte, unter denen insgesamt 14 Fragen abgefragt wurden. Insgesamt wurden davon 38 Items als 10er skalierte Ratingfragen abgefragt. In sieben Fragen wurden die Itembatterien um eine offene Antwortkategorie zur potentiellen Nennung *Sonstige* ergänzt. In Abschnitt 4 und 5 fanden sich Fragen mit Items zur Auswahl, die ebenfalls um die Kategorie *Sonstige* ergänzt wurden. Zusätzlich wurden zwei Items in Abschnitt 2 mittels prozentualer Abfrage abgefragt. Neben dieser Frage wurden zwei Rankingfragen mit prozentualer Abfrage in Abschnitt 3 integriert, um eine Budgetverteilung in Prozenten auf drei mögliche Kostenträger zu verteilen. Am Ende des Fragebogens wurden zwei offenen Fragen abgefragt.⁹⁶

Das Gruppendelphi wurde unter dem Titel *Stromsparen im Haushalt fördern durch Smart Meter* am 24. Mai 2011 im Hotel Maritim in Stuttgart durchgeführt. Die Agenda des eintägigen Workshops gliederte sich im Detail wie folgt auf:

10:00 Uhr	Beginn der Veranstaltung
10:00 -10:10 Uhr	Begrüßung
10:10 -10:30 Uhr	Vorstellungsrunde
10:30 -11:10 Uhr	Ergebnisse der Fokusgruppen

⁹⁶ Der im Gruppendelphi eingesetzte Fragebogen befindet sich im Anhang 11.9.

11:10 -11:15 Uhr	Einführung in die Methode des Gruppendelphi
11:15 -12:45 Uhr	Ausfüllen des Fragebogens in Kleingruppen I
12:45 -13:45 Uhr	Mittagspause / Auswertung der Fragebögen
13:45 -15:15 Uhr	Diskussion im Plenum I
15:15 -16:30 Uhr	Ausfüllen des Fragebogens in Kleingruppen II
16:30 -17:15 Uhr	Kaffeepause / Auswertung der Fragebögen
17:15 -18:30 Uhr	Diskussion im Plenum II
18:30 Uhr	Ende der Veranstaltung

Bei diesem Gruppendelphi wurden wie im Programm vorgesehen zwei Runden durchgeführt. Die Ergebnisse wurden nach dem Workshop mittels eines Ergebnisprotokolls und der Bitte um Korrekturen oder Ergänzungen an die Teilnehmenden versandt.

7.3.3 Analyse des Gruppendelphis zu Smart Metering

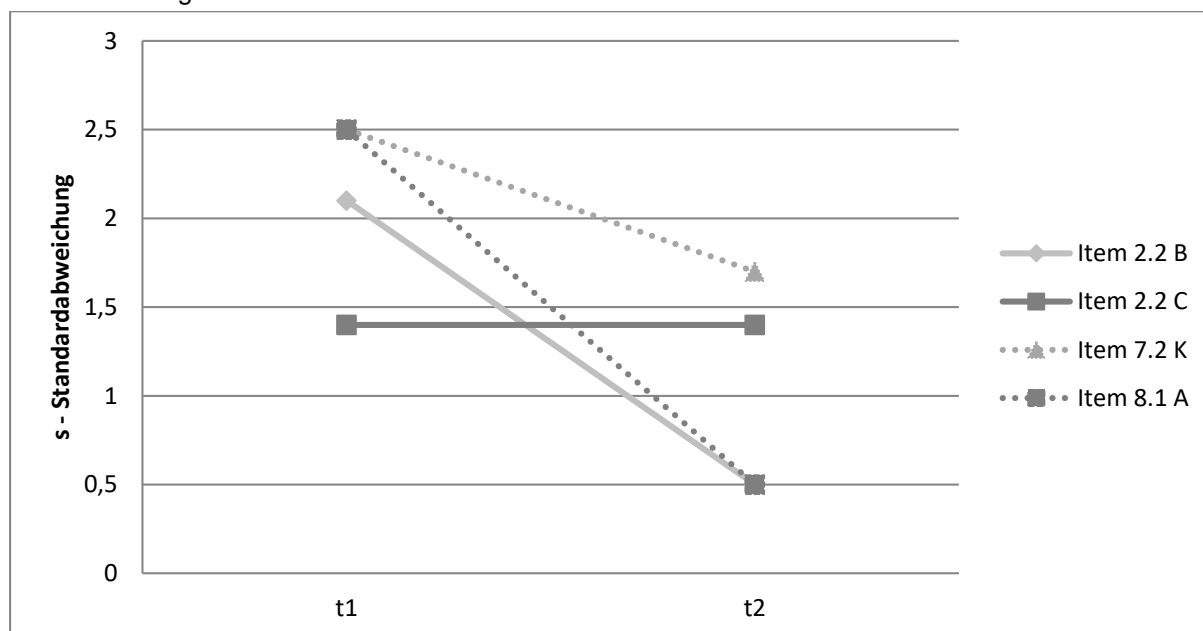
Nachdem nun auch im Falle des Gruppendelphis zur Akzeptanz und Effektivität von Smart Metern der thematische Rahmen der zentralen Fragestellungen, die Merkmale des Falles sowie der Umfang und die Inhalte des Fragebogens erläutert wurden, finden die vier Dimensionen des Testdesigns ebenfalls Anwendung.

7.3.3.1 Dimension: Evidenz

Dem Geltungsanspruch der Gültigkeit folgenden, sollen ebenso bei diesem Fallbeispiel die Qualitätsmerkmale eines handlungsleitenden Diskurses eingelöst werden, indem sich die Einschätzungen der Teilnehmenden in die Richtung eines konsensualen Ergebnisses oder zu einem Konsens über einen Dissens entwickeln. Auch bei diesem Gruppendelphi beginnt die Analyse in der Dimension der Evidenz mit den 10er-skalierten Items. Hier bildet das Gruppendelphi zu Smart Metering allerdings eine Ausnahme. So wurden zwar im ersten Fragebogen mit insgesamt 38 abgefragten 10er-skalierten Items eine beachtliche Anzahl dieses Itemtypus abgefragt, das Gruppendelphi zu Smart Metering stellt hinsichtlich der nochmals abgefragten 10er-skalierten Items allerdings einen starken Kontrast, besonders zum vorhergehenden Fallbeispiel dar. Während im Falle von SEAWIND alle dieser Items aus dem ersten Fragebogen auch im Fragebogen der zweiten Kleingruppenphase nochmals abgefragt wurden, waren es über beide Zeitpunkte im Smart Meter Gruppendelphi lediglich vier 10er-skalierte Items, die in beiden Fragebögen abgefragt wurden. So stehen in der Dimension der Evidenz für eine Analyse der Entwicklung der statistischen Kennwerte lediglich die in Abbildung 33 dargestellten vier Items zur Verfügung, die den Verlauf über beide Runden des Workshoptages abbilden können.

Bei Betrachtung von Abbildung 33 fällt allerdings auf, dass die *Items 7.2 K* und *8.1 A* gepunktet dargestellt sind. Bei beiden Items gab es in der ersten Kleingruppenphase (t1) eine Enthaltung der dritten Kleingruppe. Anders wie im Falle von SEAWIND ist diese Enthaltung allerdings nicht auf Uneinigkeit der Teilnehmenden dieser Kleingruppe begründet, sondern darauf, dass diese Kleingruppe im bestehenden Zeitrahmen der ersten Kleingruppenphase nicht den ganzen Fragebogen ausfüllen konnte. So konnte die dritte Kleingruppe im Gruppendelphi zu Smart Metering in der ersten Runde keine Bewertung für die *Frage 7.2* sowie für die *Fragen 8.1, 8.2* und *8.3* abgeben.

Abbildung 33: Vergleich der Werte der Standardabweichung über die Runden des Gruppendelphis zu Smart Metering



Anmerkungen: $tn_1=4$; $tn_2=4$

Quelle: Eigene Darstellung

Beachtet man beide Items, die folglich für eine Betrachtung der Entwicklung über beide Runden (t_1/t_2) durch Vorliegen der Urteile aller vier Kleingruppen ($n=4$) zur Verfügung stehen, so zeigt sich bereits bei Betrachtung des Wertes der Standardabweichung eine unterschiedliche Entwicklung. Während bei *Item 2.2 B* der Wert von $s=2,1$ (t_1) auf $s=0,5$ (t_2) absinkt, bleibt der Wert beim *Item 2.2 C* mit $s=1,4$ über die beiden Runden des Gruppendelphis (t_1/t_2) identisch. Die Werte des Variationskoeffizienten unterstreichen diese Entwicklung ebenfalls, denn auch hier sinkt der Wert des Variationskoeffizienten bei *Item 2.2 B* von $v=0,3$ (t_1) auf $v=0,1$ (t_2), während der Wert bei *Item 2.2 C* mit $v=0,2$ (t_1/t_2) keine Entwicklung über die beiden Runden zeigt.

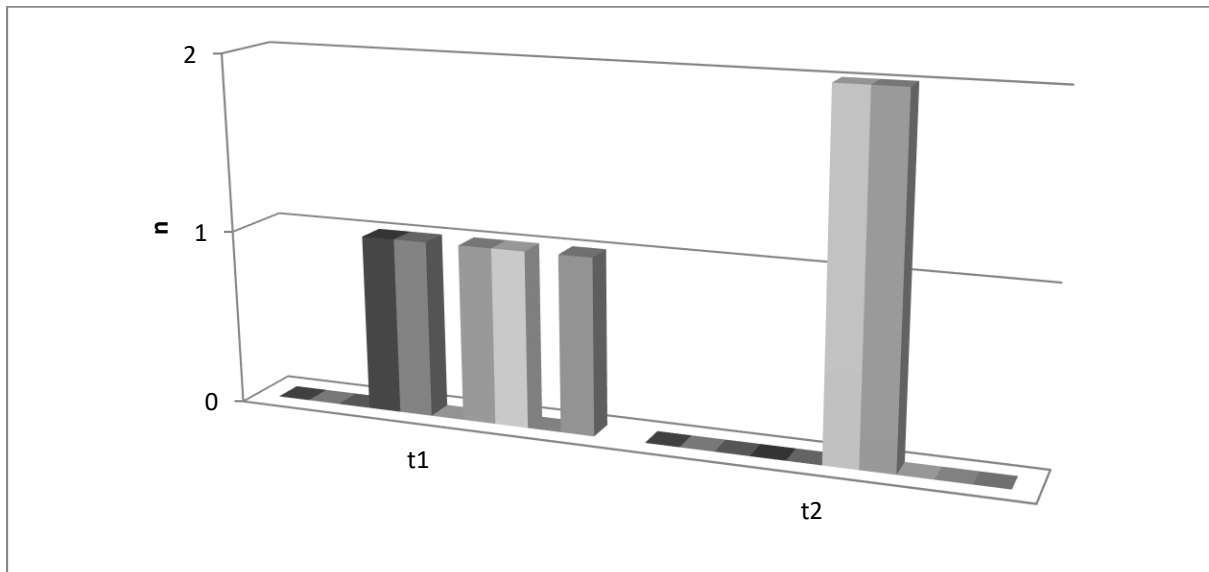
Beide Items zählten zur Itembatterie der *Frage 2.2* unter dem zweiten Abschnitt des Fragebogens, der das *Stromsparerpotential von Smart Metern* und deren *Produktvarianten* behandelte. Der Fragetext der *Frage 2.2* im Fragebogen der ersten Runde lautete: *Die Umsetzung von Smart Metering kann in verschiedenen Varianten erfolgen. Für wie erfolgversprechend bewerten Sie die einzelnen Varianten bezogen auf die Effektivität im Hinblick auf das Stromsparen?*⁹⁷ Wenn die Teilnehmenden die *Effektivität der Variante im Hinblick auf das Stromsparen* als niedrig bewerteten, so wurden sie gebeten den Wert von 1 auf der 10er-Skala zu wählen. Wenn sie die Effektivität der Variante allerdings als hoch bewerteten, so wurden sie gebeten einen Wert von 10 anzugeben.

Betrachtet man die Verteilung der Urteile zu *Item 2.2 B* in Abbildung 34, so wird die Entwicklung der statistischen Kennzahlen hin zu einem konsensualen Ergebnis hinsichtlich dieses

⁹⁷ siehe Fragebogen des Gruppendelphis im Anhang 11.9.

Items weiter bestätigt. In diesem Item wurde die Variante von *Smart Meter mit Datenübertragungsfunktion und monatlicher informativer Abrechnung* abgefragt.

Abbildung 34: Antwortverteilung bei Item 2.2 B



Anmerkungen: tn1=5 davon ein MV (10); tn2=4

Quelle: Eigene Darstellung

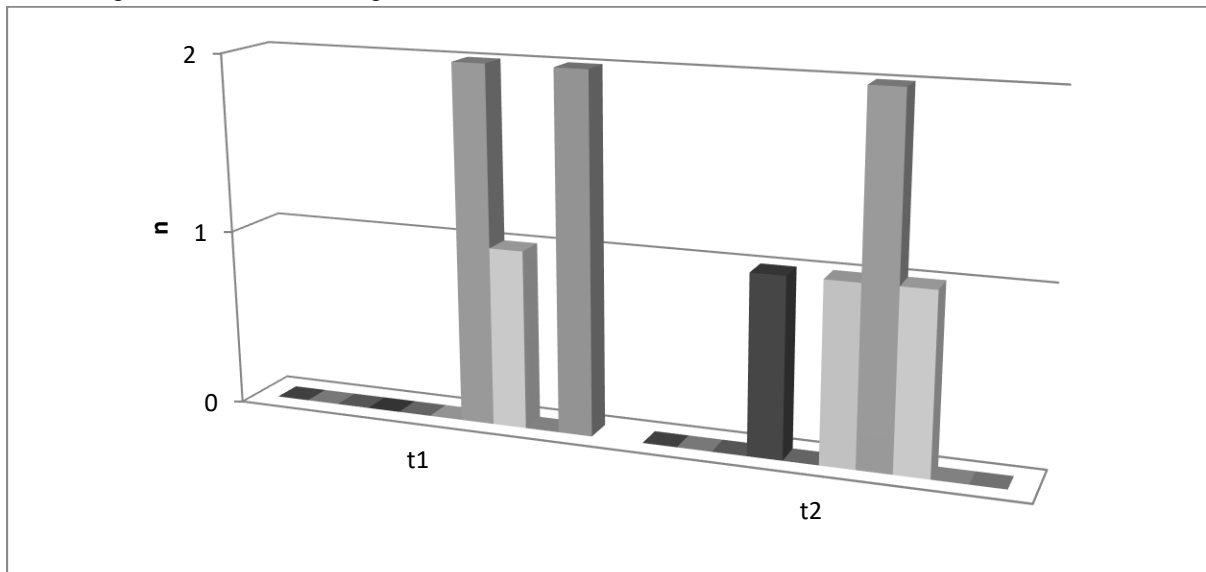
Bei Betrachtung der einzelnen Urteile der Kleingruppen zeigt sich auch, warum diese Frage trotz eines Variationskoeffizienten von unter 0,5 ($v=0,3$) im ersten Plenum des Gruppendelphis diskutiert und letztlich in den Fragebogen der zweiten Kleingruppenphase belassen wurde. So finden sich zwar drei Urteile, die diese Variante als effektiv bewerteten (Gruppe 2=7, Gruppe 3=8 und ein Minderheitsvotum in Gruppe 3=10). Es finden sich allerdings auch gegenteilige Urteile, die diese Form als eher nicht effektiv bewertet (Gruppe 1=4 und Gruppe 4=5). Da sowohl negative als auch sehr positive Urteile hinsichtlich der Effektivität dieser Variante abgegeben wurden, wurde dieses Item in der ersten Plenarphase diskutiert.

Gruppe 1 argumentierte hinsichtlich ihres eher ablehnenden Urteils, dass die Maßnahme nur retrospektiv Informationen zu Verbrauch und Kosten liefere. Dies wäre mit keinem unmittelbaren Handlungsanreiz für die Verbraucher*innen verbunden und deshalb kamen sie zu dem Urteil, dass eine solche Variante eines Smart Meters nicht als effektiv eingestuft werden könne. Die Gruppen, die die Maßnahme als effektiv eingestuft hatten, bezogen sich in ihrem Urteil darauf, dass eine monatliche informative Abrechnung eine wichtige Basis der Verbraucher*innen für Handlungsentscheidungen sei, da sie stetig unmittelbares Feedback zu Verbrauch und Kosten liefern könne. Die Person, die das Minderheitsvotum in Form eines Wertes von 10 abgegeben hatte, führte überdies das Argument an, dass eine regelmäßige monatliche Erinnerung zu einem Langzeiteffekt führe, während die Informationsbereitstellung direkt am Gerät eher nicht langfristig wirksam sei, da die Verbraucher*innen das Gerät nur am Anfang häufig in Augenschein nehmen würden. Zudem wurde argumentiert, dass der aufgeführte Geldbetrag gerade bei einer solchen Abrechnung einen unmittelbaren Handlungsanreiz setzen könne. Das Item wurde in der zweiten Runde des Gruppendelphis konsensual als eher effektiv eingestuft, so bewerteten die zur zweiten Runde neu zusammengesetzten Gruppen 1 und 3 dieses Item mit 6 und die Gruppen 2 und 4 das Item mit 7 auf der 10er-Skala. Wie dieses

konsensuale Ergebnis nach der Diskussion in der ersten Plenarsitzung und den durchaus unterschiedlichen Bewertungen in dieser Runde zustande kommen konnte, kann nach der Betrachtung des *Items 2.2 C* genauer erklärt werden.

Mit Blick auf die in Abbildung 35 dargestellte Verteilung der jeweiligen Urteile der Kleingruppen hinsichtlich des *Items 2.2 C*, zeigt sich, dass sowohl die Werte der Standardabweichung sowie auch die Werte des Variationskoeffizienten zwar über beide Runden des Gruppendelphis (t1/t2) keine Veränderungen aufzeigen, die Antworten sich allerdings anders verteilen.

Abbildung 35: Antwortverteilung bei Item 2.2 C



Anmerkungen: tn1=5 davon ein MV (10); tn2=5 davon ein MV (10)

Quelle: Eigene Darstellung

Mittels dieses *Items* wurde die Variante *Smart Meter mit Datenübertragungsfunktion und persönlichen Feedbackinstrumenten über Internetportal, PC-Software, Home-Display oder Smartphone-App* abgefragt. Ergänzend dazu wurden auch die, das System beinhaltende Informationen, im Rahmen des *Items* aufgezählt:

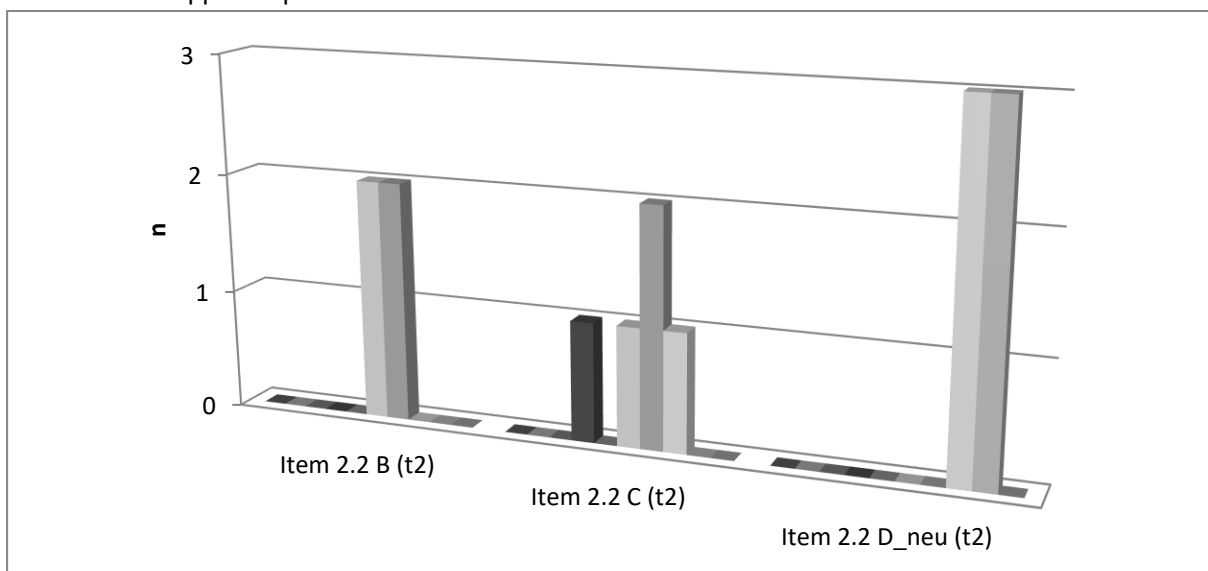
- *Lastgänge zum Gesamtstromverbrauch*
- *Gerätespezifisches Verbrauchsfeedback*
- *Bewertung des Verbrauchs durch Vergleich des Stromverbrauchs mit strukturähnlichen Haushalten bzw. den energieeffizientesten Geräten am Markt*
- *Kombination von Verbrauchsfeedback mit bereichsspezifischen Handlungsempfehlungen zu energieeffizienten Geräten und Nutzungsweisen*
- *Strategien zum Setzen von eigenen Stromsparzielen und zur Erfolgskontrolle durch den Smart Meter*
- *Erinnerungshilfen*

Betrachtet man die Verteilung der Bewertungen, so gilt für alle Urteile, dass sie diese Ausgestaltung eines Smart Meters in der ersten Kleingruppenphase durchweg als effektiv bewerteten: Gruppe 1 und Gruppe 2 mit einer Wertung von 7, Gruppe 3 mit einer Bewertung von 8 sowie Gruppe 4 und ein Minderheitsvotum in Gruppe 3 mit einer Wertung von 10. Diese Verteilung begründet, neben dem niedrigen Wert des Variationskoeffizienten von $v=0,2$ (t1), dass dieses *Item* im ersten Plenum nicht näher diskutiert wurde. Folglich ergibt sich die Frage, warum dieses *Item 2.2 C* im zweiten Fragebogen des Gruppendelphis nochmals abgefragt wurde.

Hier bietet der Austausch zwischen den Teilnehmenden im ersten Plenum einen Erklärungsansatz: Wie bereits in der Beschreibung der vorgebrachten Argumente zu *Item 2.2 B* ersichtlich wird, wurden die verschiedenen Ausgestaltungen von Smart Metering Systemen von den Teilnehmenden nicht als jeweilig - wie vorgesehen - einzeln zu bewertende Varianten angesehen, sondern vielmehr vergleichend bewertet. Dies zeigt sich beispielsweise an der Argumentation, dass die Informationsbereitstellung direkt am Gerät eher nicht langfristig wirksam sei. Da die *einfache Variante eines Smart Meters ohne Datenübertragungsfunktion (Item 2.2 A)* bereits in der ersten Kleingruppenphase mit einem arithmetischen Mittelwert von $\bar{x}=1,8$ und $s=0,4$ von allen vier Kleingruppen geschlossen als nicht effektiv bewertet wurde, blieben die zwei Arten der Ausgestaltung (*Item 2.2 B* und *Item 2.2 C*) als Optionen für die Expert*innen im ersten Plenum bestehen, die gegeneinander abgewogen werden mussten. Dieser Aspekt war den Teilnehmenden wichtig, da beide Varianten gerade im Kontrast ihrer jeweiligen Stärken und Schwächen gesehen werden müssten. Im Hinblick darauf, dass beide Varianten von Smart Metering jeweils positive Aspekte beinhalten würden, wurde aus dem Kreis der Expert*innen vorgeschlagen, auch beide Varianten gemeinsam als Kombination abzufragen. Diesem Vorschlag wurde durch die Moderation gefolgt, indem ein weiteres Item in den Fragebogen der zweiten Kleingruppenphase aufgenommen wurde: *Kopplung von monatlicher Abrechnung mit Datenübertragungsfunktion sowie persönlichen Feedbacklösungen (Item 2.2 D_neu)*.

Die in Abbildung 36 dargestellte Verteilung der drei Items der zweiten Kleingruppenphase spiegelt den Austausch im ersten Plenum wider. So wurde die Variante der *Kopplung von monatlicher Abrechnung mit Datenübertragungsfunktion sowie persönlichen Feedbacklösungen (Item 2.2 D_neu)* von allen vier Kleingruppen in der zweiten Runde konsensual als effektive Gestaltungsart eingestuft.

Abbildung 36: Vergleich der Antwortverteilung der Items 2.2 B, 2.2 C und 2.2 D_neu in der zweiten Runde des Gruppendelphis



Anmerkungen: Item 2.2 B (t2): n=4; Item 2.2 C (t2): n=5; Item 2.2 D (t2): n=6. Quelle: Eigene Darstellung

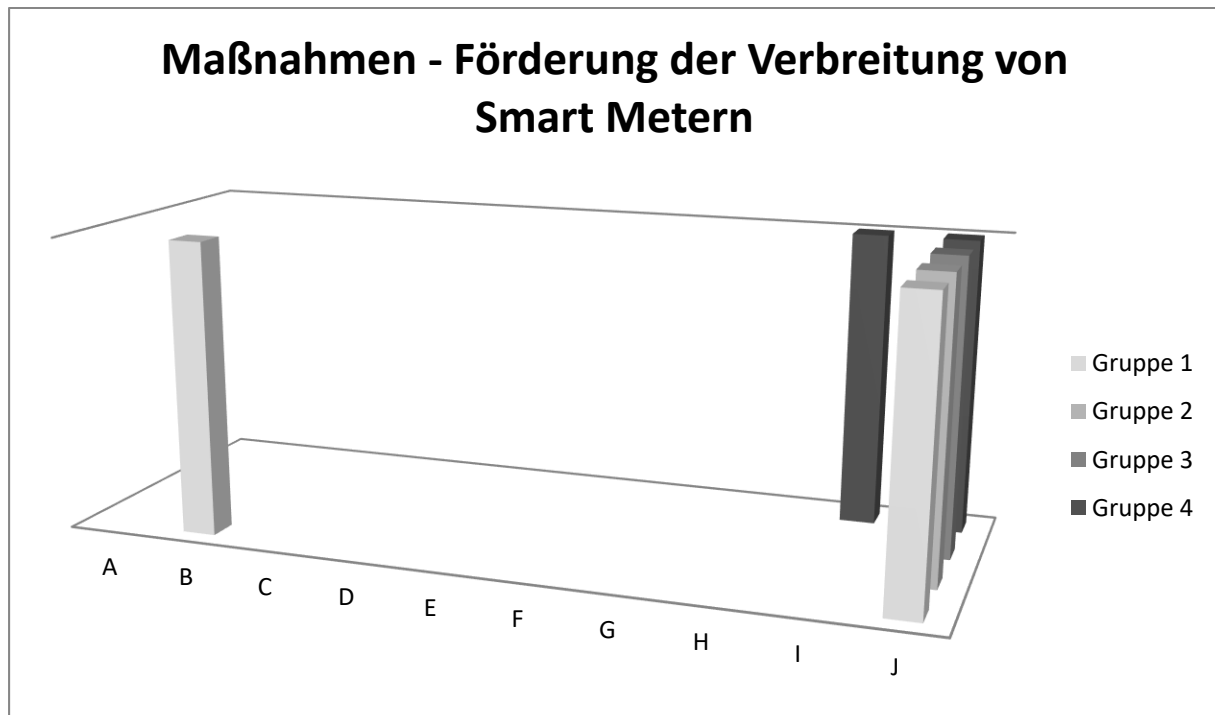
Hinzu kommen zwei Minderheitsvoten aus Gruppe 1 und 4, die die Effektivität dieser Lösung zwar hoch (8), aber etwas geringer einstufen wie ihre jeweiligen Gruppen (9). Die beiden Teilnehmenden begründeten die geringfügig niedrigere Bewertung im zweiten Plenum damit,

dass sie bereits die Lösung in *Item 2.2 B* als effektive Variante beurteilten, allerdings eine Kombination nicht schaden würde. Einer dieser beiden Teilnehmenden fand ein *monatliches Feedback* wesentlich effektiver, als die in *Item 2.2 C* aufgeführten *Feedbacklösungen*, so dass dieser Experte in der Kleingruppe 4 dies mit dem Minderheitsvotum von 4 beim *Item 2.2 C* zusätzlich markierte. Unter Berücksichtigung der vorgebrachten Argumente, ergibt sich unter Beachtung der Verteilung der Antworten in Runde 2 ein für die Ergebnisverwertung eindeutiges Bild. So haben die Teilnehmenden beide Varianten der Ausgestaltung von Smart Metern (*2.2 B* und *2.2 C*) in der zweiten Runde jeweils für die einzelne Variante im Hinblick auf ihre Effektivität bewertet, wie es in der ursprünglichen Konzeption des Gruppendelphis auch vorgesehen war. Die Variante von *Smart Metern mit Datenübertragungsfunktion und monatlicher informativer Abrechnung (Item 2.2 B)* beurteilen die Teilnehmenden konsensual zwar als effektiv, mit einem arithmetischen Mittel von $\bar{x}=6,5$ allerdings auch nicht als sehr effektive Art der Ausgestaltung. Bei der Ausgestaltung eines *Smart Meters mit Datenübertragungsfunktion und persönlichen Feedbackinstrumenten über Internetportal, PC-Software, Home-Display oder Smartphone-App (Item 2.2 C)* liegt die Beurteilung durch die Teilnehmenden mit $\bar{x}=6,4$ auf einem ähnlichen Niveau. Allerdings besteht hier durch das Minderheitsvotum eines Experten durch den Skalenwert von 4 eine Bewertung, die diese Variante als ineffektiv einstuft. Im zweiten Plenum bestand hier Konsens über einen Dissens zwischen dem Experten und den übrigen Teilnehmenden, die dieser Variante ein gewisses Maß an Effektivität zusprachen. Letztlich konnte aber für die Itematterie der *Umsetzung von Smart Metering in verschiedenen Varianten und deren Erfolg auf die Effektivität im Hinblick auf das Stromsparen (Frage 2.2)* ein konsensuales Ergebnis für eine Ausgestaltung von Smart Metern erzielt werden, da der neuen Variante *Kopplung von monatlicher Abrechnung mit Datenübertragungsfunktion sowie persönlichen Feedbacklösungen (Item 2.2 D_neu)* mit einem arithmetischen Mittel von $\bar{x}=8,5$ (t_2) konsensual eine hohe Effektivität zugesprochen wurde. Diese Bewertung wurde von den Teilnehmenden mit dem zusätzlich im zweiten Plenum geäußerten Argument gestützt, dass *je umfassender eine Variante sei, desto besser sei sie auch*. Auch die anderen Varianten können zu Erfolg führen, jedoch sollte das Maximale angestrebt werden, da eine Kombination keinesfalls schädlich sein könne. Dem wurde konsensual zugestimmt.

Einen weiteren Vergleich zwischen beiden Runden des Gruppendelphis, bietet *Frage 4.1*, die als Auswahlfrage innerhalb der Kategorie der Ratingskalen konzipiert wurde und in der die Teilnehmenden drei aus insgesamt neun Maßnahmen auswählen konnten. Ergänzt wurde die Auswahl durch ein zusätzliches Item, indem sonstige Maßnahmen vorgeschlagen werden konnten. Diese Frage befand sich im vierten Abschnitt des Fragebogens zur *Förderung der Verbreitung von Smart Metern durch ein Markteinführungsprogramm und empfohlene Tarifgestaltungen (keine verpflichtende Einführung für alle Haushalte)*. Der Fragetext der *Frage 4.1* lautete dabei wie folgt: *Smart Meter haben über die gesetzlich geforderten Basislösungen in Neubauten und bei grundlegenden Sanierungen hinaus bislang lediglich eine sehr eingeschränkte Verbreitung gefunden. Wie sich in den Fokusgruppen gezeigt hat und weitere Studien belegen, sind ein wesentlicher Grund dafür die Kostenstrukturen (Einmalkosten für Geräte und Einbau, im Vergleich zu den Standardtarifen höhere monatliche Grundpreise, höhere kWh-Preise außerhalb der Schwachlastzeiten). Eingesparte Kosten durch Verbrauchsreduktionen werden durch diese Mehrkosten oftmals aufgezehrt. Welche Maßnahmen halten Sie für am besten geeignet, um die Verbreitung von Smart Metern (kombiniert mit monatlicher infor-*

mativer Abrechnung und Feedbackinstrumenten) im Rahmen eines Markteinführungsprogramms zu fördern? Begleitende kommunikative Maßnahmen werden hier vorausgesetzt, die detaillierte Diskussion darüber erfolgt in der nächsten Frage. Die drei wichtigsten, erfolgversprechendsten Maßnahmen im Hinblick auf Verbreitung bitte ankreuzen.⁹⁸ Die bei dieser Frage zur Auswahl stehenden Maßnahmen basierten auf den vorhergehenden methodischen Anwendungen und Analysen des wissenschaftlichen Konzeptes des Teilprojektes, bestehend aus Literaturrecherche, Marktanalyse und Fokusgruppen.

Abbildung 37: Auswahl der Maßnahmen in Frage 4.1 zur ersten Runde (t1) des Gruppendelphis zu Smart Metern



Legende: A: Subvention der einmaligen Einbaukosten für EVUs unter der Bedingung, dass eine vorgegebene Mindestanzahl SM pro Jahr eingebaut wird; B: Subvention der einmaligen Einbaukosten für Verbraucher (analog zur „Abwrackprämie“ im Kfz-Bereich); C: Subvention des monatlichen Grundpreises für Verbraucher; D: Steuerliche Begünstigungen des Stromverbrauchs für die SM nutzenden Endverbraucher, d.h. der Strompreis sinkt für diese Verbraucher; E: Steuerliche Begünstigungen für die SM anbietenden Unternehmen unter der Bedingung, dass eine vorgegebene Mindestanzahl SM pro Jahr eingebaut wird; F: Ergänzendes Bonus- oder Prämienangebot, z.B. kostenlose Nutzung eines Elektromobils für einen Tag; G: Leasingmodelle für die Anschaffung der Geräte; Gesamtkosten über die Laufzeit dürfen nicht höher sein als die Kosten bei Einmalzahlung; H: An SM gekoppelte Tarife mit einem Kontingent an Inklusiv-KWh, wobei sich die Inklusivstrommenge dabei an einem sehr stromsparenden Durchschnittsverbraucher orientiert; I: Progressive Tarife gekoppelt an den SM: deutlicher Preisanstieg bei überdurchschnittlichem Verbrauch, dafür aber geringere Preise im Bereich unterhalb des Durchschnittsverbrauchs; J: Sonstiges.
Quelle: Eigene Darstellung

Trotz dieser Vorarbeiten bei der Konzeption der Maßnahmen wird in Abbildung 37 schnell deutlich, dass die Teilnehmenden die Auswahl, der vom Forscherteam formulierten Maßnahmen, nicht favorisierten. So wurde die *Maßnahme B* einer *Subvention der einmaligen Einbaukosten für Verbraucher (analog zur „Abwrackprämie“ im Kfz-Bereich)* nur von Gruppe 1 als sinnvolle Maßnahme ausgewählt. Dies gilt ebenso für die *Maßnahme I* über *progressive Tarife gekoppelt an den SM: deutlicher Preisanstieg bei überdurchschnittlichem Verbrauch, dafür*

⁹⁸ siehe Fragebogen des Gruppendelphis zu Smart Metern im Anhang 11.9.

aber geringere Preise im Bereich unterhalb des Durchschnittsverbrauchs, die lediglich von Gruppe 4 ausgewählt wurde. Obwohl für alle vier Gruppen insgesamt die Möglichkeit bestand drei dieser neun vorgeschlagenen Maßnahmen auszuwählen, blieb es bei der Auswahl von in Summe zwei der abgefragten Maßnahmen, Gruppe 2 und 3 verzichteten dabei gänzlich auf eine Auswahl einer der vorgeschlagenen Maßnahmen. In Abbildung 37 wird mit Blick auf die Kategorie *J* aber ebenso deutlich, dass sich die Kleingruppen nicht gänzlich bei dieser Frage in ihrem Urteil enthalten haben. Hier hatten die Teilnehmenden in der ersten Kleingruppenphase die Möglichkeit eigene Maßnahmenvorschläge offen zu formulieren. Diese Möglichkeit nutzen alle vier Kleingruppen, so dass nach der ersten Kleingruppenphase insgesamt vier weitere Maßnahmen als Vorschläge für *erfolgsversprechende Maßnahmen im Hinblick auf eine Verbreitung* von Smart Metern zu Auswahl standen:

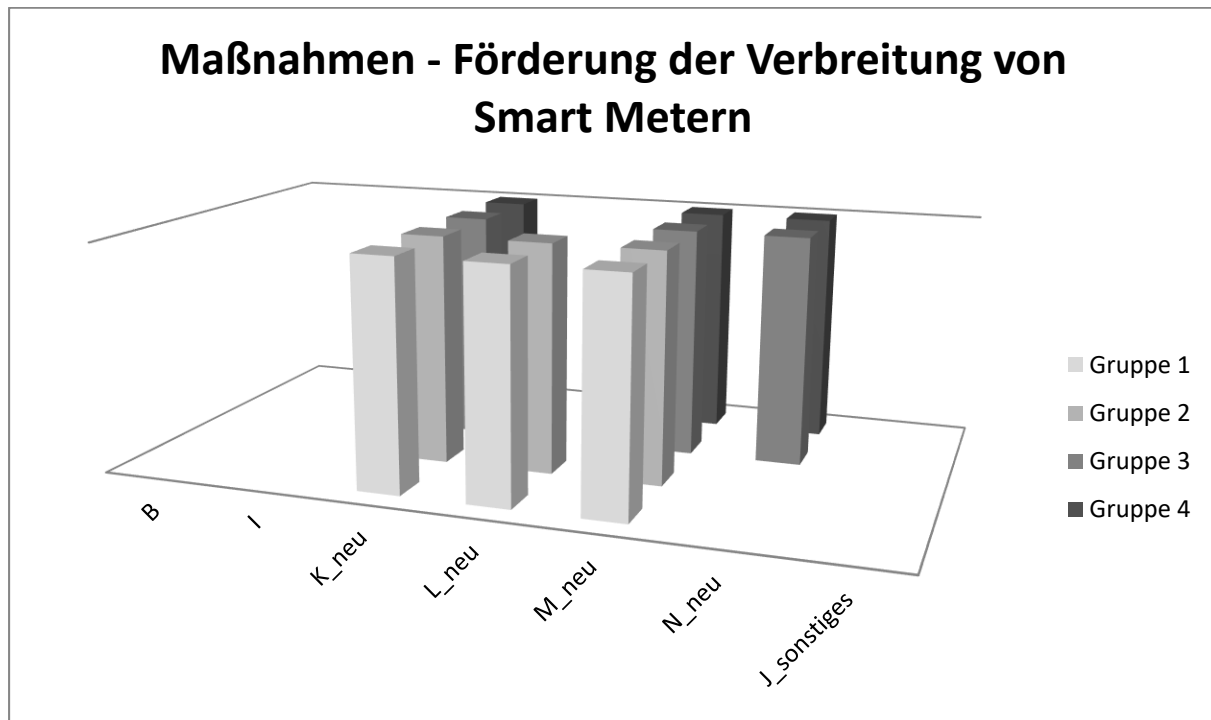
- *Tatsächliche Lastspezifische Preise, Preissignale vom Smart Meter (K_neu)*
- *Klare Zielvorgaben in Kombination mit gesetzlichen Rahmen (L_neu)*
- *Kostenneutralität / Umlage über Netzentgelte (M_neu)*
- *Möglicherweise staatliche Finanzierung, Abwicklung mit Mindestquote über EVUs, Anreiz in großen Chargen, keine Abschreckung der Kunden durch hohe Anschaffungspreise (N_neu)*

Da diese Maßnahmen der einzelnen Kleingruppen den jeweils anderen Kleingruppen zum Zeitpunkt des ersten Plenums unbekannt waren, reagierte die Moderation auf die große Anzahl der neuen Maßnahmen, indem sie jede der Kleingruppen bat, ihre vorgeschlagene Maßnahme kurz zu erläutern. Dies erfolgte mit dem Ziel in der zweiten Kleingruppenphase bereits einen gemeinsamen Bewertungshorizont zu schaffen. Nach dieser Erläuterung wurden die von den Teilnehmenden vorgeschlagenen Maßnahmen in den Fragebogen der zweiten Runde in die Auswahlfrage 4.1 aufgenommen. Überdies wurden die Gruppen 1 und 4 gebeten, ihre Auswahl der bereits im ersten Fragebogen vorhandenen Maßnahmen zu erläutern. Gruppe 1 gab an, dass sie *Maßnahme B* gewählt habe, da sie als Kleingruppe diese Auswahl im Hinblick auf eine Wahl *noch am besten* fanden, sie aber im Vergleich zu den neuen Maßnahmen nicht (mehr) favorisierten. Um dies auch in der zweiten Runde bestätigen zu lassen, wurde die *Maßnahme B* im Fragebogen belassen. Eine identische Begründung für die Auswahl der *Maßnahme I* hatte auch Gruppe 4, so dass auch diese Maßnahme nochmals zur Bestätigung dieser Aussage im Fragebogen der zweiten Runde belassen wurde. Die Teilnehmenden einigten sich darauf, alle anderen im ersten Fragebogen bereits abgefragten Maßnahmen (*A, C-H*) als konsensual abgelehnt zu betrachten, so dass diese nicht mehr in der zweiten Runde (t2) zur Auswahl gestellt wurden.

Bei Betrachtung der in Abbildung 38 dargestellten Auswahl bestätigt sich bereits die Aussage der Kleingruppen zu den *Maßnahmen B* und *I*, dass diese Maßnahmen nur als beste Alternative in einer nicht idealen Auswahl in der ersten Kleingruppenphase angesehen wurden. Diese wurden in der zweiten Kleingruppenphase von keiner Gruppe ausgewählt. Die Ergebnisse der zweiten Kleingruppenphase zeigen überdies, dass sich alle Teilnehmenden konsensual auf zwei Maßnahmen einigen konnten, die sie als wichtige und erfolgsversprechende Maßnahmen zur Förderung der Verbreitung von Smart Metern bewerteten:

- *Tatsächliche lastspezifische Preise, Preissignale vom Smart Meter (4.1 K_neu)*
- *Kostenneutralität für die Haushalte/ Umlage über Netzendgeld (4.1 M_neu)*

Abbildung 38: Auswahl der Maßnahmen in Frage 4.1 zur zweiten Runde (t2) des Gruppendelphis zu Smart Metern



Legende: B: Subvention der einmaligen Einbaukosten für Verbraucher (analog zur „Abwrackprämie“ im Kfz-Bereich); I: Progressive Tarife gekoppelt mit Smart Meter: Deutlicher Preisanstieg bei überdurchschnittlichem Verbrauch, dafür aber geringe Preise im Bereich unterhalb des Durchschnittsverbrauchs; K_neu: Tatsächliche Lastspezifische Preise, Preissignale vom Smart Meter; L_neu: Klare Zielvorgaben in Kombination mit gesetzlichen Rahmen; M_neu: Kostenneutralität / Umlage über Netzentgelte; N_neu: Möglicherweise staatliche Finanzierung, Abwicklung mit Mindestquote über EVUs, Anreiz in großen Chargen, keine Abschreckung der Kunden durch hohe Anschaffungspreise; J: Sonstiges. Quelle: Eigene Darstellung.

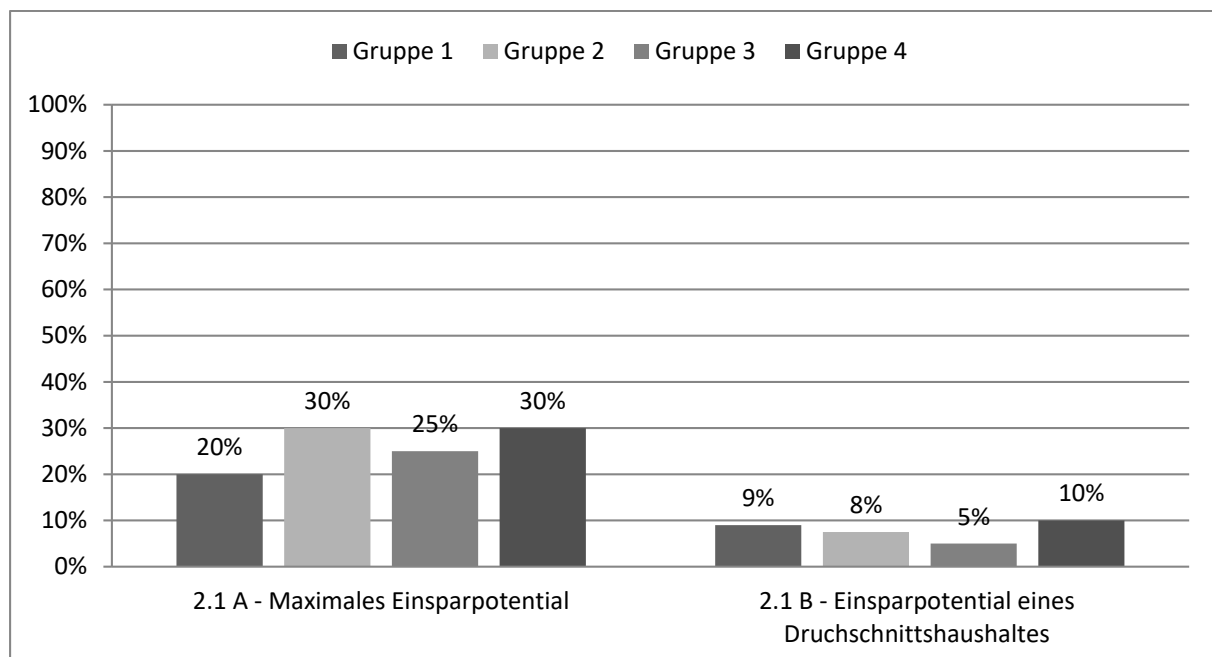
Neben diesen beiden Maßnahmen führten die Teilnehmenden im zweiten Plenum das Argument an, dass aufgrund einer hohen Komplexität des Energiemarktes auch weitere Maßnahmen ergänzt werden sollten, die über Tarifstrukturen hinaus zum Stromsparen anregen. So sahen im Austausch auch die Gruppen 3 und 4, die von Gruppe 1 und 2 ausgewählte Maßnahme *klare Zielvorgaben in Kombination mit gesetzlichen Rahmen (L_neu)* als wichtig und erfolgversprechend an. Dies galt auch für das umfassende Maßnahmenbündel von *möglicher staatlicher Finanzierung, Abwicklung mit Mindestquote über EVUs, Anreiz in großen Chargen, keine Abschreckung der Kunden durch hohe Anschaffungspreise (N_neu)*, das von Gruppe 3 und 4 in der zweiten Kleingruppenphase ausgewählt wurde und dies auch die Gruppen 1 und 2 im zweiten Plenum als wichtig und erfolgversprechend ansahen. So ergab sich letztlich eine klare und konsensual getragene Auswahl von zwei Maßnahmen, die nach Bewertung der Teilnehmenden durch weitere Maßnahmen flankiert werden sollten.

Der zweite Abschnitt bezüglich des *Stromsparpotentials von Smart Metern* im Fragebogen der ersten Delphirunde enthielt auch eine Ratingfrage mittels prozentualer Abfrage. In dieser wurden die teilnehmenden Expert*innen gebeten, sowohl das *maximale Einsparpotential bei der Nutzung von Smart Metern (Item 2.1 A)*⁹⁹ als auch das für einen *Durchschnittshaushalt zu*

⁹⁹ *Wie viel Prozent Einsparung halten Sie bei der Nutzung von Smart Metern für maximal möglich?* (Hervorhebung im Original, siehe Anhang 11.9).

erwartende Einsparpotential (Items 2.1 B)¹⁰⁰ mittels prozentualer Angabe zu bewerten. In Abbildung 39 finden sich die jeweiligen prozentualen Bewertungen der Einsparpotentiale für beiden Items des Fragebogens der ersten Kleingruppenphase. Sowohl bei *Item 2.1 A* als auch bei *Item 2.1 B* zeigt sich hier bereits, dass das mögliche Einsparpotential von Smart Metern durch die verschiedenen Kleingruppen ähnlich eingeschätzt wird. So wird das *maximale Einsparpotential (Item 2.1 A)* mit einem Wert von $\bar{x}=26,5\%$ eingeschätzt. Gruppe 1 liegt mit einer Einschätzung von 20% möglichen Einsparpotentials zwar am weitesten von diesem Mittelwert entfernt, jedoch unterscheidet sich die Einschätzung dieser Gruppe von den Einschätzungen der anderen drei Kleingruppen nicht signifikant.

Abbildung 39: Prozentuale Angaben der Kleingruppen zu Frage 2.1 in Runde 1 des Gruppendelphis zu Smart Metern



Anmerkungen: tn1=4; tn2=4

Quelle: Eigene Darstellung

Ein ähnliches Bild zeigt sich auch bei der Einschätzung des *Einsparpotentials eines Durchschnittshaushaltes (Item 2.1 B)*, bei der das arithmetische Mittel im Vergleich zu *Item 2.1 A* einen deutlich niedrigeren Wert von $\bar{x}=8\%$ hat. Hier liegt Gruppe 3 mit einer Einschätzung des Einsparpotentials von 5% am weitesten von den Einschätzungen der anderen Kleingruppen entfernt, allerdings unterscheidet sich auch hier die Einschätzung nicht signifikant von den anderen Kleingruppen. Aufgrund dieser ähnlichen Bewertungen des Einsparpotentials von Smart Metern wurde diese *Frage 2.1* im ersten Plenum nur kurz von der Moderation angesprochen. Die Teilnehmenden konnten dabei die jeweils anderen Einschätzungen der Kleingruppen mittragen und führten überdies als Argument an, dass sowohl das *maximale Einsparpotential* wie auch das *Einsparpotential in einem Durchschnittshaushalt* stark von einer konkreten Ausgestaltung der Smart Meter abhingen, so dass für diese Frage sowohl bei *Item 2.1*

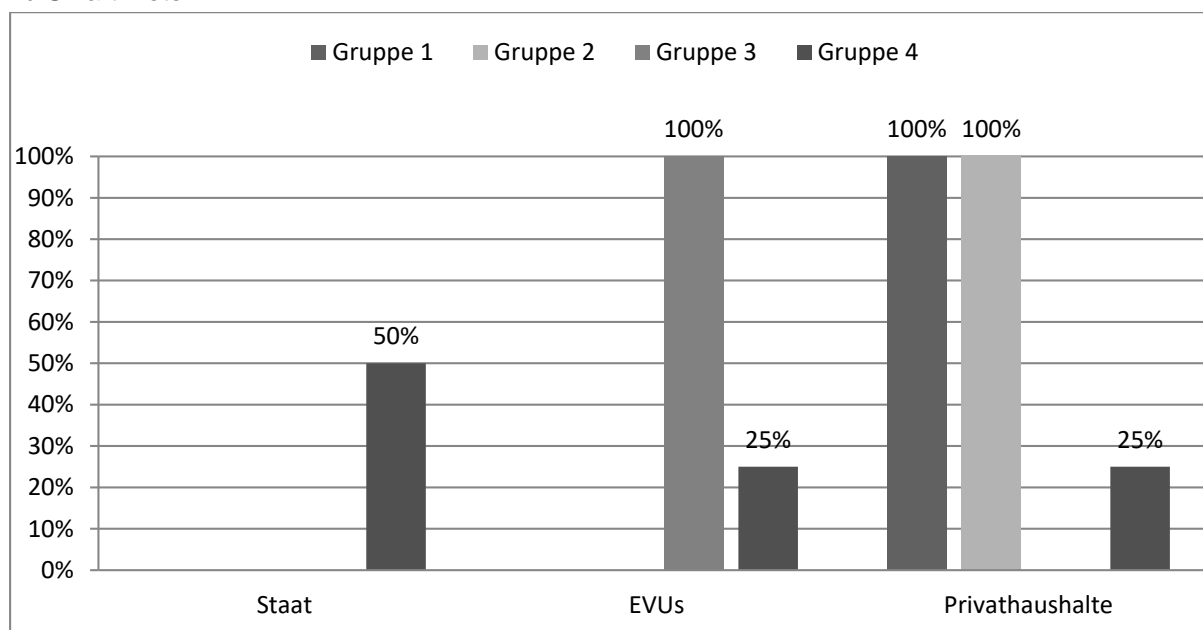
¹⁰⁰ Wie viel Prozent Einsparung ist in einem Durchschnittshaushalt zu erwarten? (Hervorhebung im Original, siehe Anhang 11.9).

A ein Range von 20-30% wie auch bei *Item 2.1 B* ein Range von 5-10% von allen Teilnehmenden je nach Variante¹⁰¹ konsensual getragen wurde. So konnte ein Konsens bereits nach dem ersten Plenum erreicht werden.

Eine weitere Ratingfrage mittels prozentualer Abfrage enthielt *Frage 3.1* des Fragebogens, die Teil des dritten Abschnitts war, der eine *verpflichtende Einführung von Smart Metern für alle Haushalte* zum Inhalt hatte. Abgefragt wurde ein weitergehender Ausbau von Smart Metern als dieser zum Zeitpunkt des Delphis vorgesehen war: *Smart Meter mit Datenübertragungsfunktion müssen verpflichtend von jedem Energieversorger bzw. Netzbetreiber bei allen Kunden zu dem Zeitpunkt eingebaut werden, wenn der alte Zähler geeicht werden müsste. Mechanische Zähler müssen je nach Typ alle 12 oder 16 Jahre geeicht werden.*

Das *Item 3.1 B* fragte danach, wie die Teilnehmenden die Kosten einer solchen Maßnahme prozentual verteilen würden. Dabei konnten die Expert*innen *die Kosten für einen verpflichtenden Einbau auf drei Akteure* verteilen: Staat, Energieversorgungsunternehmen (EVU)¹⁰² und Privathaushalte. Abbildung 40 zeigt dabei ein divergierendes Bild hinsichtlich der Verteilung bei *Frage 3.1*.

Abbildung 40: Prozentuale Verteilung der Kleingruppen in Frage 3.1 B in Runde 1 des Gruppendelphis zu Smart Metern



Anmerkungen: tn1=4; tn2=4

Quelle: Eigene Darstellung

Während Gruppe 1 und 2 die gesamten 100% auf die *Privathaushalte* und damit die Verbraucher*innen verteilten, verteilte Gruppe 3 die gesamte Menge der Kosten auf die *Energieversorgungsunternehmen (EVUs)*. Eine Aufteilung der Kosten über die drei abgefragten Akteursgruppen hinweg, nahm lediglich die vierte Kleingruppe vor, die 50% der Kosten an den *Staat*,

¹⁰¹ Diese Varianten wurden in Frage 2.2 genauer mittels 10er-skalierten Items abgefragt. Siehe weiter oben in diesem Kapitel.

¹⁰² Mit dem zusätzlichen Hinweis: *Es wird gesetzlich festgelegt, dass dieser Anteil durch die EVUs getragen wird* (siehe Anhang 11.9).

und jeweils 25% der Kosten an die *EVUs* und *Privathaushalte* verteilte. Diese signifikant unterschiedlichen Bewertungen ließen nach der ersten Kleingruppenphase einen Dissens zwischen den Kleingruppen vermuten, so dass die einzelnen Bewertungen der jeweiligen Kleingruppen im ersten Plenum ausgetauscht wurden. Dabei erläuterte zuerst Gruppe 3 die Verteilung der gesamten Kosten auf die EVUs damit, dass die Kosten erst mal bei diesen anfallen würden, wenn eine verpflichtende Einführung von Smart Metern gesetzlich vorgeschrieben sei, da diese die neuen Systeme anschaffen und einbauen müssten.

Dieses Argument teilten ebenso die Gruppen 1 und 2, nur, dass sie ihrer Bewertung den Umstand zugrunde legten, dass diese Kosten von den EVUs letztlich an die Privathaushalte umverteilt würden, und damit immer die Verbraucher*innen zahlen würden, denn auch Subventionen von staatlicher Seite würden letztlich ebenso über Steuerabgaben von den Verbraucher*innen finanziert. Dieser Argumentation konnten alle vier Gruppen zustimmen. Die Verteilung auf alle drei Akteure, die Gruppe 4 in der ersten Kleingruppenphase vorgenommen hatte, basierte auf einem anderen Verständnis des Frageinhaltes. Dabei erläuterte Gruppe 4 ihr Verständnis der Frage und die daraus resultierende Verteilung. Da diese Gruppe die Verteilung als subjektiv sinnvoll erachtete, die einer persönlichen Einschätzung darüber folge, wie die Kosten verteilt sein sollten:¹⁰³ *Wie würden Sie die zusätzlich entstehenden Kosten verteilen?*¹⁰⁴ Auch dieser Perspektive konnten alle Teilnehmenden zustimmen, so dass sich letztlich zwei gemeinsame Blickwinkel ergaben:

- Letztlich werden sowohl die anfallenden Kosten über die Abrechnung vom EVU als auch von Staat über Steuern auf die Privathaushalte verteilt (Evidenz)
- Alle Akteure sollten sich an den Kosten beteiligen (normative Angemessenheit)¹⁰⁵

Diese zwei Perspektiven und der damit verbundene Konsens zwischen den Kleingruppen nach der Diskussion im ersten Plenum konnte so auch für die zweite Frage in diesem Design übernommen werden. Auch wenn sich die in Abbildung 41 dargestellten Verteilungen der jeweiligen Kleingruppen zu *Frage 3.2 D* in der ersten Runde teilweise von denen in *Frage 3.1 B* unterschieden. Auch Frage 3.2 hatte einen *verpflichtenden Einbau von Smart Metern* zum Frageinhalt, nur, dass dieser im Gegensatz zu *Frage 3.1* mit der zusätzlichen Komponente *der Ausgestaltung von Smart Metern mit Datenübertragungsfunktion* ergänzt wurde, also zusätzlich eine *Kombination mit persönlichen Feedbackinstrumenten über Internetportal, PC-Software, Home-Display oder Smartphone-App als zusätzliche Maßnahme* integriert wurde.

Die Gruppen 1 und 4 wählten auch bei dieser *Frage 3.2 D* eine identische Verteilung, wie bei *Frage 3.1 B* und äußerten dafür die jeweilig gleiche Begründung, wie in Abbildung 41 ersichtlich wird. Gruppe 3 nutzte als Kontrast zu *Frage 3.1 B* bei *Frage 3.2 D* die Gelegenheit, hier ihre als subjektiv angesehene faire Bewertung anzugeben, wie es Gruppe 4 bereits bei der vorhergehenden Frage mit gleicher Argumentation getan hat. Gruppe 2 sah allerdings durch

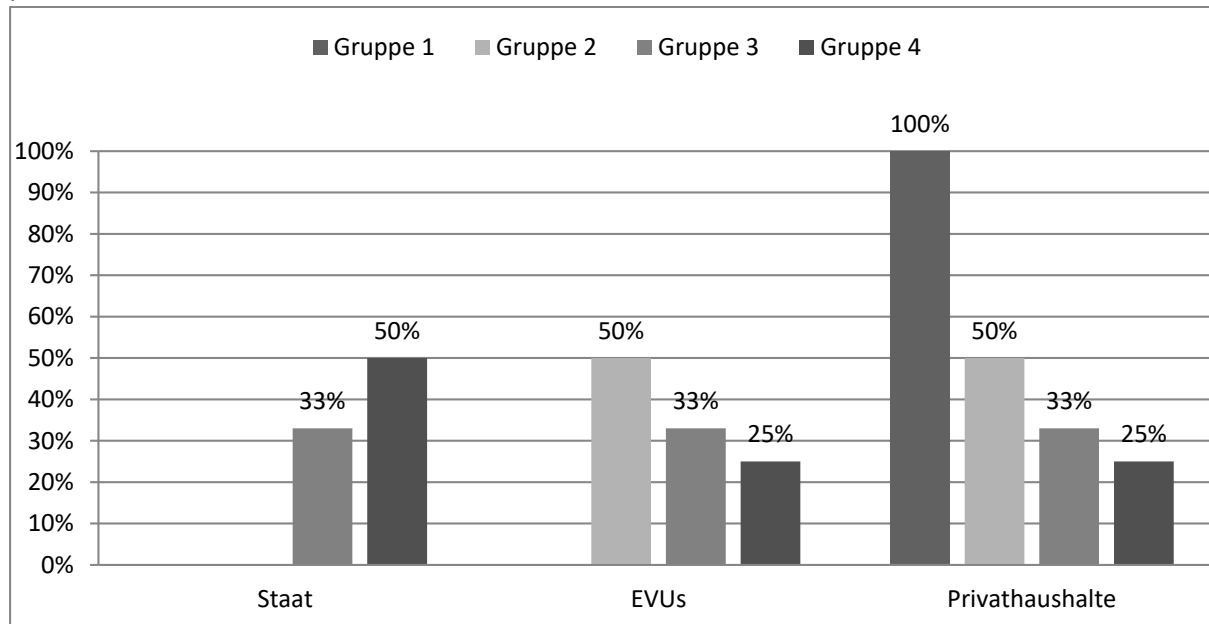
¹⁰³ Mehr dazu in Kapitel 7.3.3.2.

¹⁰⁴ Siehe Frage 3.1 B im Fragebogen (Anhang 11.9).

¹⁰⁵ Mehr dazu in Kapitel 7.3.3.3.

die zusätzliche Integration von Datenübertragungsfunktion und Feedbackinstrumenten bei dieser Form der Ausgestaltung eine Chance für einzelne Verbraucher*innengruppen Kosten einzusparen, die mit entsprechender gesetzlicher Vorgabe vom EVU getragen werden sollten.¹⁰⁶

Abbildung 41: Prozentuale Verteilung der Kleingruppen in Frage 3.2 D in Runde 1 des Gruppendelphis zu Smart Metern



Anmerkungen: tn1=4; tn2=4

Quelle: Eigene Darstellung

Die Teilnehmenden ergänzten beide *Frage 3.1 B* und *Frage 3.2 D* in der folgenden Diskussion im ersten Plenum um Argumente, die einen breiten Kontext mit einbezogen. Denn das Ziel der Einführung von Smart Metern sei, mit dem Hintergrund der Integration von Erneuerbaren Energien, einen Einklang von Einspeisung und Verbrauch herzustellen. Damit stellte sich nach Ansicht der Teilnehmenden nicht die Frage danach, ob Smart Meter eingeführt werden, sondern *wann* dies geschehe. Gerade diese zeitliche Dimension sei beispielsweise deshalb relevant, wann eine punktuelle Finanzierung durch den Staat wichtig sei. Eine solche staatliche Finanzierung könne als zeitlich begrenzter Anschlag durchaus sinnvoll sein. Wenn überdies bestimmt sei, vor welchem Hintergrund eine Einführung von Smart Metern stattfinde, könne die Kostenfrage behandelt werden. Ein volkswirtschaftlich sinnvoller Einsatz sei damit bei einer entsprechenden gesamtwirtschaftlichen Regulierung möglich.

Die Expert*innen artikulierten für die zweite Runde des Gruppendelphis den Wunsch, die Einführung von Smart Metern in einen solchen weiteren Zusammenhang einzuordnen und damit die Rahmenbedingungen für die im Gruppendelphi abgefragten Maßnahmen bestimmen zu können. Der Vorschlag eines Teilnehmenden in der zweiten Kleingruppenphase *Präambeln* zu formulieren, die die Einführung von Smart Metern in diese Zusammenhänge einordnen, wurde von den Teilnehmenden geschlossen als sinnvoll erachtet. Dieser Forderung wurde von Forscher*innenteam entsprochen, so dass in der zweiten Runde zu Beginn des Fragebogens eine offene Frage eingefügt wurde, in der die Teilnehmenden die Möglichkeit hatten, in ihrer jeweiligen Kleingruppe eine solche Präambel zu formulieren. Zwar sind offene Fragetypen bei

¹⁰⁶ Hier ergänzten sich in der Argumentation Evidenz und normative Angemessenheit, dies wird in Kapitel 7.3.3.3 nochmals thematisiert werden.

einem Gruppendelphi ein gängiger Fragetypus, allerdings stellen diese Präambeln, die von den Teilnehmenden als Abfrage der Rahmenbedingungen des Delphis vorgeschlagen wurden, in dieser Form eine besondere Abfrage dar. Da diese Präambeln zwar einer Verständigung über den Beurteilungen zugrundeliegenden Relevanzsysteme dienlich sind, was auch eine Diskussion in der Dimension der Evidenz rechtfertigen würde, folgen diese umso mehr dem Geltungsanspruch der Zweckmäßigkeit und sollen deshalb im folgenden Kapitel weiter untersucht werden.

7.3.3.2 Dimension: Verständlichkeit

Diese Dimension dient der Untersuchung, ob die Methode des Gruppendelphis dem Geltungsanspruch der Zweckmäßigkeit für den Diskurs entsprechen kann. Das heißt, ob die Teilnehmenden folglich mit dem *Geäußerten* letztlich einen gleichen Bedeutungsinhalt verbinden? Das Gruppendelphi muss, wie bereits in den vorhergehenden Fällen erwähnt, einen Raum dafür bieten, eben einen solchen gemeinsamen Bedeutungsinhalt für die strittigen Begriffsdefinitionen, Klassifikationsschemata und Formulierungen von Aussagen zu erzielen. Anders als bei den zwei vorhergehenden Fallbeispielen finden sich beim Gruppendelphi zu Smart Metern keine Umformulierungen einzelner Aussagen oder Frageinhalte, die auf interdisziplinär unterschiedlich verwendete Fachtermini, semantische Missverständnisse oder vom wissenschaftlichen Team nicht klar formulierte Aussagen hindeuten. So wurden in diesem Fallbeispiel im Vergleich von ersten Fragebogen und Fragebogen der zweiten Runde keine Umformulierungen bzw. Ergänzungen von Aussage- und Frageformulierungen vorgenommen. Wenn die Teilnehmenden die Auswahl der, vom Forscherteam vorgeschlagenen Items nicht teilen konnten, so kam es, wie im vorhergehenden Kapitel 7.3.3.1 zu gänzlich neuen Vorschlägen für Items, wie im Falle der Auswahlfrage 4.1 von Maßnahmen zur *Förderung der Verbreitung von Smart Metern*. Diese Möglichkeit wurde durch die Kategorie *Sonstige* bereits im Fragebogen der ersten Runde geboten.

Mit *Frage 3.3* wurde allerdings auch zur zweiten Runde des Gruppendelphis eine gesamte Frage in den Fragebogen der zweiten Runde integriert: *Es gibt einen volkswirtschaftlich sinnvollen Einsatz von SM, der gesamtgesellschaftlich geregelt werden muss, Einbezug von erneuerbaren Energien, energiewirtschaftlicher Workflow und Anreiz zur Einsparung. Nennen Sie relevante Akteure, auf die Kosten verteilt werden können.* Diese Frage wurde unter den Teilnehmenden konsensual für die zweite Runde abgestimmt und formuliert. Die Notwendigkeit dieser Integration ergab sich aus dem Austausch im ersten Plenum zu den *Fragen 3.1 B* und *3.2 D*, in denen jeweils die EVUs als ein zu bewertendes Item abgefragt wurden, auf die ein Teil der Kosten prozentual verteilt werden konnte. Alle Teilnehmenden in den Kleingruppen sahen bereits in der ersten Kleingruppenphase den Begriff der *EVUs* als problematisch an, wie sie im folgenden Plenum geschlossen äußerten. Denn hier können unterschiedliche Akteure unter einem Begriff verstanden werden, was eine Bewertung vor identischem Bewertungshorizont erschwere. Nach der Liberalisierung des Strommarktes und des Messwesens müsse zwischen verschiedenen Akteuren unterschieden werden, so dass für die zweite Runde eine Abfrage dieser möglichen Akteure entschieden wurde. Wie viele Akteure von den Teilnehmenden als relevante Akteure benannt wurden, zeigen die Antworten aus der zweiten Runde zu der neuen *Frage 3.3: Endverbraucher, Messdienstleister, Messstellenbetreiber,*

Netzbetreiber (ÜNB/VNB), Vertrieb, Erzeuger (regenerativ/konventionell), Bund, Land, Kommune, EU und Energiehandel. Bei der Abfrage in der zweiten Runde zeigte sich, dass auch der Begriff des *Staates*, der ebenfalls als Item zur Auswahl stand, für die Abfrage einer Kostenverteilung nicht differenziert genug war. So wurde hier von den Teilnehmenden zwischen *Bund, Land* und *Kommune* unterschieden und auf supranationaler Ebene auch die *EU* aufgeführt, die jeweils für spezifische Kostenübernahmen in Frage kämen. Der Austausch im zweiten Plenum brachte hinsichtlich des Begriffs der *Energieversorgungsunternehmen* den Konsens hervor, dass dieser Terminus im liberalisierten Markt nicht mehr zutreffend sei und vorwiegend zwischen *Stromlieferant, Netzstellenbetreiber* und *Messstellenbetreiber* unterschieden werden müsse. Da aus zeitlichen Gründen keine dritte Runde des Gruppendelphis stattfinden konnte, war eine erneute Abfrage der Kostenverteilung unter den ausdifferenzierten Akteuren nicht mehr möglich, was vor dem ab dem zweiten Plenum geteilten Bewertungshorizont zu spezifischeren Ergebnissen hätte führen können, als es in den ursprünglich konzipierten *Fragen 3.1 B* und *3.2 D* der Fall war.¹⁰⁷

Da die Integration der Präambeln besonders mit dem Ziel erfolgte, den Maßnahmen einen gemeinsamen Bedeutungshorizont voranzustellen, soll diese im Fragebogen (t2) eingesetzte offene Frage in dieser Dimension nun näher betrachtet werden. In Tabelle 11 finden sich die von den Teilnehmenden im Fragebogen der zweiten Delphirunde in den jeweiligen Kleingruppen formulierten Präambeln. Diese wurden im Anschluss der Kleingruppenphase im zweiten Plenum von den jeweiligen Gruppen vorgetragen und anschließend zur Diskussion gestellt. Die von Gruppe 1 vorgetragene und erläuterte Präambel *Smart Metering in einem Gesamtkontext aus Smart Grid und Smart Home zu sehen* erhielt dabei von allen Teilnehmenden volle Zustimmung. Besonders *eine Integration und ein Ausbau von erneuerbaren Energien, Energieeffizienzsteigerung und Energieeinsparung* wurden als wichtige Elemente des *Gesamtkontextes* gesehen. Bei der von Gruppe 2 vorgestellten Präambel gab es hingegen Diskussionsbedarf hinsichtlich des Verlagerungspotentials von Smart Metern auf die *Lastgänge des Stromnetzes*. Hier wurde als möglicher leistungsbremsender Faktor die *Berufstätigkeit* von Verbraucher*innen als Gegenargument angeführt. Dieses Argument erkannte die zweite Gruppe als Hemmnis an, erwiderte allerdings, dass auch in der *Alltagsführung von Berufstätigen* ein auf Lastgänge angepasster Verbrauch durch den Einsatz von *Smart Metering* und *Smart Grid* erfolgen könne. So konnten sich die Teilnehmenden letztlich diskursiv darauf einigen, dass *Berufstätigkeit* eine einschränkende, aber nicht gänzlich hemmende Komponente darstelle.

Bei einer personenbezogenen Profilbildung, wie sie in den Erläuterungen der Präambel von Gruppe 2 vorgestellt wurde, konnten die Teilnehmenden keinen Konsens darüber erzielen, ob *Pseudonymisierung* aus Gründen des *Datenschutzes* zielgruppenspezifische Angebote ausschließen würde. Insgesamt wurde eine *Profilbildung* bei Smart Metern und Smart Home allerdings als weniger kritisch eingeschätzt, als dies bei Handynutzung oder Internetzugang der Fall sei. So konnten letztlich auch alle Teilnehmenden dem von Gruppe 2 verfassten schriftlichen Inhalt dieser Präambel zustimmen.

¹⁰⁷ siehe Kapitel 7.3.3.1

Tabelle 11: Von den Kleingruppen in der zweiten Runde (t2) des Gruppendelphis zu Smart Metern formulierte Präambeln

Gruppe	Präambel
1	<p><i>Smart Metering ist in einem Gesamtkontext aus Smart Grid und Smart Home zu sehen. Insbesondere sind dabei die Integration und Ausbau von erneuerbaren Energien, Energieeffizienzerhöhung und Energieeinsparung in dem Gesamtkontext zu berücksichtigen.</i></p> <p><i>Der Staat hat Interesse an CO2-Anspargung, der Kunde hat Interesse, Kosten zu sparen und die Energieversorger ein Interesse optimal Erlöse zu gewinnen. Nur wenn sich alle diese Interessen gemeinsam realisieren, ist der Einsatz von Smart Metering flächendeckend erfolgreich.</i></p>
2	<p><i>Politischer Nutzen: Unterstützung der Umwelt und Klimaschutzziele durch die langfristige Integration von erneuerbare Energien mit ihrer volatilen Verfügbarkeit in die Netzinfrastruktur und Lastflüsse. → Zielerreichung durch zeitliche Lastverteilung</i></p> <p><i>Volkswirtschaftlicher Nutzen und Nutzen beim regulierten natürlichen Monopol (Netzbetreiber / "Messtellenbetreiber"): Reduzierung von Netzinfrastrukturinvestitionen durch verbesserte Auslastung der heutigen Netze und Reduzierung der anstehenden Neuinvestitionen. → Zielerreichung durch zeitliche und räumliche Lastverteilung (dezentrale Einspeisung) („smart Grid“)</i></p> <p><i>Nutzen für Lieferanten und Kunden: Stromkosteneinsparung durch Berücksichtigung des tatsächlichen Lastverlauf beim Einkauf und setzen von Preissignalen und Preisanreize für den Kunden. → Zielerreichung durch Bilanzierbarkeit des tatsächlich gemessenen Lastprofils Anstelle von Standardlastprofile</i></p> <p><i>Kundennutzen und politischer Nutzen: Energieeinsparung durch transparente Darstellung des Energieverbrauches und damit verbunden Verhaltensänderung beim Kunden. → Zielerreichung durch transparente Darstellung des Energieverbrauches</i></p>
3	<p><i>Es ist ein durchgängiges Konzept von der politischen Zielstellung bis zum Nutzen für den Endkunden durch aktive Einbindung aller Akteure zu schaffen. Es müssen Anreize durchgängig geschaffen werden für Endkunden, Lieferant/Vertrieb, MSB/MDL, Netzbetreiber, Erzeuger / dezentraler Erzeuger (Anlagenbetreiber).</i></p> <p><i>Es müssen dynamische Transfers durchgängig abbildbar sein:</i></p> <ul style="list-style-type: none"> • <i>Smart Grid i.w.S. (e-Mobility, Smart Home) – Fähigkeit</i> • <i>Es müssen weitere Sparten (Wärme, Wasser etc.) berücksichtigt werden</i>
4	<p><i>Kommunikation zu SM immer im Zusammenhang mit SG und Integration EE. Mehr Transparenz, mehr Ehrlichkeit, was kann ein SM wirklich für wen tun. Keine reine Kosteneinsparkommunikation. SM in Zusammenhang mit einer Gesamtstrategie stellen: Konsumer → Prosumer</i></p>

Bei den jeweiligen Präambeln von Gruppe 3 und Gruppe 4 gab es keinen weiteren Diskussionsbedarf, so dass die Teilnehmenden den formulierten Inhalten dieser beiden Präambeln letztlich ebenfalls geschlossen zustimmten. Die Teilnehmenden teilten somit folglich konsensual die Bewertung, dass Smart Meter für den generellen Kontext und ebenso den Kontext des Gruppendelphis, aus einer engen Betrachtung als einzelnes Instrument zum Stromsparen herauszulösen seien. Vielmehr sollten Smart Meter als Element in einem breiten Kontext mit den übergeordneten Themen der Integration der Erneuerbaren Energien in die Netze, Einführung von Smart Grids und Smart-Home-Konzepten, Elektromobilität, CO2-Einsparung und einer Steigerung der Energieeffizienz verstanden werden. Diese breitere Betrachtung sei insbesondere für die Kommunikationsstrategie gegenüber den Verbraucher*innen von zentraler Bedeutung. Auch wenn sich mit Smart Metern für Verbraucher*innen keine Kosten sparen lassen sollten, so ist die Einführung vor dem genannten Hintergrund dennoch unabdingbar.

Durch die Einigung, dass alle Präambeln nebeneinanderstehend den Kontext des Gruppendelphis im Hinblick auf die bewerteten Maßnahmen sowie den generellen Bewertungshorizont

vorgaben, konnte ein konsensual getragener Bedeutungsinhalt - wenn auch ex-post während des zweiten Plenums - für das Gruppendelphi durch die Integration und den Austausch über die Präambeln erreicht werden. Inwieweit dabei auch Gründe der normativen Angemessenheit eine Rolle spielten, soll im folgenden Abschnitt näher untersucht werden.

7.3.3.3 Dimension: normative Angemessenheit

Für die Analyse, der von den Teilnehmenden in den beiden Plenarsitzungen des Gruppendelphis zu Smart Metern getätigten Aussagen, stehen sowohl ein Protokoll beider Plenarsitzungen und das Ergebnisprotokoll, das nach dem Gruppendelphi an die Teilnehmenden versandt wurde, zur Verfügung. Überdies können im Falle dieses Gruppendelphis, die bereits im vorhergehenden Abschnitt besprochenen, von den Teilnehmenden in der zweiten Runde des Gruppendelphis formulierten Präambeln in die Auswertung mit einfließen. Ziel dieses Abschnitts ist es zu überprüfen, ob es den Teilnehmenden in diesem Fallbeispiel möglich war, neben evidenzbasierten Äußerungen ebenso normative Einstellungen und Wertvorstellungen als Argumente anzubringen. Und überdies, ob diese als legitime Aussagen von den Teilnehmenden im Diskurs aufgenommen wurden.

Im Falle des Gruppendelphis zu Smart Metern finden sich, in allen für eine Inhaltsanalyse zur Verfügung stehenden Dokumenten Aussagen, die unter der Dimension der normativen Angemessenheit betrachtet werden können. Im ersten Plenum findet sich bei *Frage 4.1* ein erster Austausch von Argumenten, die auf einer normativen Begründung aufbauen, in der die *Förderung der Verbreitung von Smart Metern durch ein Markteinführungsprogramm und empfohlene Tarifgestaltungen* behandelt wurde. Hier wurden die Teilnehmenden gebeten, die Maßnahmen auszuwählen, die nach ihrer Ansicht *die Verbreitung von Smart Metern (kombiniert mit monatlicher informativer Abrechnung und Feedbackinstrumenten) im Rahmen eines Markteinführungsprogramms (zu) fördern*. In Abschnitt 7.3.3.1 wurde bereits dargestellt, dass es bei der Abfrage während der ersten Kleingruppenphase keinen Konsens bei den im Fragebogen vorgeschlagenen Maßnahmen gab und jede der vier Kleingruppen jeweils eine eigene formulierte Maßnahme vorschlug. Letztlich gelang es aufgrund dieser vorgeschlagenen Maßnahmen ein konsensuales Ergebnis zur zweiten Runde des Gruppendelphis zu erzielen. Im ersten Plenum wurden bereits die Begründungen der einzelnen Kleingruppen für die *Nicht-Auswahl*, der in der Frage aufgelisteten Maßnahmen sowie den Vorschlag eigener Maßnahmenformulierungen ausgetauscht. Dabei äußerten Teilnehmende auch individuelle Bewertungen von Optionen, die sie transparent als subjektive Einschätzungen offenlegten. So äußerte ein Experte, dass er die potentielle *Stromeinsparung* von Smart Metern als *nice to have* bewerte, diese allerdings an dritter Stelle hinter *Lastverlagerung* und *Kostensparpotential* einzuordnen sei. Dieser Meinungsäußerung schloss sich ein Teil der Expert*innen im weiteren Verlauf des Plenums an. Die Teilnehmenden der dritten Kleingruppe stellten das Einsparpotential hinsichtlich der Kosten in der weiteren Argumentation in den Vordergrund und äußerten überdies die Präferenz, dass *Smart-Meter-Lösungen* für die Kund*innen *kostenneutral sein sollten* und nicht von diesen getragen werden müssten, wie es zum Zeitpunkt des Delphis der Fall war. Ein direkter Einwand aus dem Plenum bezog sich auf diese Äußerung und brachte das Argument an, dass die Kund*innen *letztlich immer zahlen würden*, ob nun *individuell* (Rechnung) oder *sozialisiert* (Steuern). Der darauffolgende Diskurs im Plenum fokussierte sich daran anschließend noch mehr auf den Aspekt einer erfolgreichen Verbreitung von Smart Metern.

So sei ein wichtiger Schritt *die Transaktionskosten zu verringern*. Die von den Teilnehmenden geäußerten Argumente, die sich auf den Aspekt der normativen Angemessenheit bezogen, wurden bei dieser Frage im ersten Plenum von den Teilnehmenden als legitime Begründungen akzeptiert und in den folgenden Argumenten erfolgte ein direkter Bezug zu diesen.

Auch *Frage 3.1 B* und *Frage 3.2 D* wurden bereits in Abschnitt 7.3.3.1 hinsichtlich der Evidenz diskutiert. Der Dissens zwischen den Bewertungen lag hier allerdings besonders in der Perspektive begründet, hinsichtlich derer die einzelnen Kleingruppen das Budget verteilten. Während sich drei Kleingruppen für ihre Urteile in der ersten Kleingruppenphase auf Evidenz stützten, wählte die vierte Gruppe eine Verteilung der Kosten vor dem Hintergrund der normativen Angemessenheit.

Der Hintergrund der Bewertung dieser Kleingruppe war damit nicht, wie die Verteilung der Kosten letztlich unter den gegebenen Bedingungen ausfallen würde, sondern eine Einschätzung, wie die *Kosten verteilt sein sollten*. Auf diese beiden Perspektiven konnten sich die Teilnehmenden in der folgenden Diskussion verständigen und sahen beide als gleichberechtigt relevant an. Zudem wurde die zeitliche Dimension nochmals herausgestellt, denn die geteilte Einschätzung lautete, dass in der *Einführungsphase ein zeitlich begrenzter Anschub durch den Staat durchaus sinnvoll* wäre. Auch wenn beide Fragen (*3.1 B* und *3.2 D*) nicht in der zweiten Runde des Gruppendelphis abgefragt wurden, so wurden hinsichtlich der Verteilung des Budgets aufgrund der detaillierten Abfrage weiterer involvierter Akteursgruppen in *Frage 3.3* ebenfalls die Budgetverteilung mit einem weiteren Argument thematisiert. Da durch die Auflistung weiterer Akteure eine Aufteilung der Kosten auf diese zu komplex gewesen wäre, konnte man sich konsensual auf das wertgeleitete Argument einigen, dass die Kostenverteilung sich danach richten sollte, *wer profitiert und dann eher entsprechende Kosten (mit-)tragen kann*.

Frage 5.1 thematisierte *Tarifstrukturen, die das Stromsparen fördern können*. Bei dieser Frage wurden die Teilnehmenden aufgefordert, *die zwei wichtigsten, erfolgsversprechendsten Tarifstrukturen im Hinblick auf erfolgreiches Energiesparen* auszuwählen. Die an die Kleingruppenphase (t1) anschließende Diskussion im Plenum fokussierte sich zu Beginn auf die erste Maßnahme A: *Progressive Tarife gekoppelt an den SM: deutlicher Preisanstieg bei überdurchschnittlichem Verbrauch, dafür aber geringere Preise im Bereich unterhalb des Durchschnittsverbrauchs*. Diese Maßnahme wurde von drei Kleingruppen als erfolgsversprechend bewertet, während Gruppe 1 diese Maßnahme nicht wählte. Diese Gruppe begründete ihre Nicht-Auswahl der Maßnahme mit einem Hinterfragen, ob *eine Unterscheidung zwischen „gutem Strom“ (geringer Verbrauch) und „bösem Strom“ (darüber hinausgehender Verbrauch) sinnvoll sei*. So seien *Sondersituationen*, wie beispielsweise *Nachwuchs* oder *Besuch*, aber auch die *Hausinfrastruktur* als Faktoren zu berücksichtigen, bei denen progressive Tarife zu Ungerechtigkeiten führen könnten. Überdies wurde hinterfragt, wie ein in der Maßnahme enthaltener *überdurchschnittlicher Verbrauch* ermittelt werden könne, bzw. was als durchschnittlicher Verbrauch festgelegt werden könne. Berücksichtige man die verschiedenen Lebensumstände der Verbraucher*innen, *seien Konstrukte, wie Durchschnittsverbräuche fragwürdig* und könnten damit zu Ungerechtigkeiten führen. Diese Argumente wurden von den übrigen Gruppen als legitim erachtet. Die folgende Diskussion stellte die übergeordnete Frage, ob es überhaupt das *primäre Ziel* der Einführung von Smart Metern sei, *Strom zu sparen*. Die Teilnehmenden waren

sich darin einig, dass mit der Einführung von Smart Metering nicht das Stromsparen das prioritäre Ziel sein sollte, sondern an erster Stelle *die Integration der erneuerbaren Energien* stünde. Damit seien auf Tarife ausgelegte Maßnahmen letztlich zweitrangig und sollten mit Hinblick auf dieses Ziel zurückgestellt werden.

An dieses Argument wurde ebenfalls in der Plenardiskussion (t1) zu *Frage 7.2* angeknüpft, bei der *Inhalte*, die *bei kommunikativen Maßnahmen zu Smart Metern thematisiert werden* sollten, abgefragt wurden. Nachdem ein Teilnehmer hier erst das Argument anführte, dass für die Verbraucher*innen als primärer Kommunikationsinhalt und wichtigster Aspekt die persönliche Kostenersparnis stünde, änderte sich die anfängliche Zustimmung einer Vielzahl der Teilnehmenden, nachdem das Argument vorgetragen wurde, dass *aufgrund der komplexen Materie eine ebenso komplexe Informationsvermittlung notwendig* sei. So wurde folglich die persönliche Präferenz geäußert, dass die Spannbreite aller Maßnahmen erhalten werden müsse, um der Komplexität und der großen Menge an möglichen Zielgruppen gerecht zu werden. Das an *Frage 5.1* anschließende Argument wurde auch bei dieser Frage konsensual geteilt, das die dieser Frage folgende Vision in den Mittelpunkt stellte: *Die Energiewende, für die Smart Meter zwingend notwendig* seien. Diese Argumentation findet sich letztlich auch in den Präambeln wieder, wie beispielsweise im von Gruppe 1 (t2) formulierten Text: *Smart Metering ist in einem Gesamtkontext aus Smart Grid und Smart Home zu sehen. Insbesondere sind dabei die Integration und Ausbau von erneuerbaren Energien, Energieeffizienzerhöhung und Energieeinsparung in dem Gesamtkontext zu berücksichtigen*. Oder auch in der von Gruppe 4 formulierten Präambel, in der eine *Kommunikation zu Smart Metern* immer in den *Zusammenhang von Smart Grid und den Erneuerbaren Energien* gesetzt werden sollte.

Frage 6.1 hatte eine *Einsparquote* zum Inhalt, die eine *gesetzlich festgelegte Verpflichtung der EVUs zur Einsparung vorgegebener Energiemengen in einem festgelegten Zeitraum* beinhalten sollte. Bei dieser Frage gab es in der ersten Runde des Gruppendelphis einen großen Dissens über alle drei Items dieser Itematterie hinweg. Während Gruppe 1 alle drei abgefragten Maßnahmen voll und ganz ablehnte (1), stimmte Gruppe 2 allen Maßnahmen voll und ganz zu (10). Die beiden anderen Kleingruppen bewegten sich dabei im Mittelfeld der 10er-Skala zwischen Skalenwerten von 2 bis 6.¹⁰⁸ Aufgrund der hoch unterschiedlichen Bewertungen bei allen drei Items wurde diese Frage als Ganze im ersten Plenum diskutiert. Gruppe 1 nannte als Begründung für ihr ablehnendes Urteil, dass die EVUs nicht für die Einsparungen der Kund*innen verantwortlich gemacht werden sollten und der Staat *hier nicht über Bande spielen* sollte. Es sei die *Rolle des Staates Instrumente zu entwickeln, die direkt bei den Verbraucher*innen wirken* können. Für die Stromanbieter ergebe sich ein Glaubwürdigkeitsproblem, wenn diese *zum Stromsparen animieren* würden. Gruppe 2 konnte diese Argumentation zwar nachvollziehen, sah aber eine mögliche Verpflichtung der Anbieter durch den Staat als mögliches Instrument, wenn die *Netzbetreiber* über ein Regulierungskonto Ausgleichsmöglichkeiten für Mindereinnahmen erhalten würden. Auch diese Argumente wurden von den Teilnehmenden als legitim erachtet und führten letztlich zurück zur Diskussion, um die Komplexität des Strommarktes, verbunden mit einer Vielzahl von Akteuren, bei denen auch zwischen Netzbetreibern (monopolistisch) und Anbietern (liberalisiert) unterschieden werden müsse. Diese in

¹⁰⁸ Alle Antworten sowie die statistischen Werte der *Items 6.1 A, 6.1 B und 6.1 C* finden sich im Anhang 11.10 und sollen mit Fokus auf die normative Dimension adressierende Argumente an dieser Stelle nicht weiter vertieft werden.

Frage 3.3 erst zur zweiten Runde des Gruppendelphis abgefragte Vielzahl an Akteursgruppen verhinderte eine abschließende Bewertung dieser *Frage 6.1*. So wurde im ersten Plenum konsensual entschieden, diese Frage, aufgrund *der Komplexität im liberalisierten Markt*, nicht mehr weiter in diesem Gruppendelphi zu behandeln.

Am Ende des zweiten Plenums wurde auf Nachfrage der Moderation nochmals *Frage 7.1* behandelt, die *Szenarien der kommunikativen Vermittlung von Smart Metern* zum Inhalt hatte. Eines dieser Szenarien beinhaltete *Kommunikationsangebote an Bürger/innen durch Kommunen*. Dieses Szenario wurde zur ersten Runde des Gruppendelphis mit einem arithmetischen Mittel von $\bar{x}=4,6$ und einer Standardabweichung von $s=1,2$ von den Gruppen geschlossen im Mittelfeld der 10er-Skala bewertet.¹⁰⁹ Da dieses Item im ersten Plenum aufgrund fehlender zur Verfügung stehender Zeit nicht mehr diskutiert werden konnte, allerdings in der zweiten Runde des Gruppendelphis nicht mehr abgefragt wurde, fand eine Diskussion im zweiten Plenum statt. Hierbei führten die Teilnehmenden verschiedene Argumente an, warum Kommunen nicht ideale Akteure seien, um eine gänzlich effektive kommunikative Vermittlung von Smart Metern an die Bürger*innen zu übernehmen. Neben dem Argument, dass Kommunen zur Neutralität verpflichtet seien und damit keine Werbung für Anbieter von Smart Metering machen könnten, wurden auch gleichberechtigt wertende Argumente vorgetragen, wie dass *Kommunen selbst häufig schlampig mit Energie umgingen*. Im Sinne einer *Vorbildwirkung* sollen sie selber mit der *Installation von Smart Metern vorangehen*. Auch wurde als Argument vorgebracht, dass staatliche Mittel nicht für die Erstellung von Informationsmaterial verwendet werden sollten, wenn andere Akteure von der Einführung der Smart Meter profitieren würden. Letztlich gab es auch Argumente, die eine nicht gänzlich schlechte Bewertung der Kommunen als Kommunikationsakteur unterstützten, da diese für eine neutrale Informationsvermittlung wichtig seien. Diese Argumente begründeten letztlich die Bewertung der Kommunen als zum Teil geeignet und wurden von den Teilnehmenden als verschiedene Faktoren dieser Bewertung gleichberechtigt akzeptiert.

So finden sich im Falle des Gruppendelphis zu Smart Metern in den Protokollen der beiden Plenarsitzungen Äußerungen der Teilnehmenden, die als Argumentation den Geltungsanspruch nach Gemeinwohl und normativer Richtigkeit direkt adressieren. Solche Argumente wurden bei diesem Gruppendelphi gegenüber Argumenten, die den Geltungsanspruch der Gültigkeit adressierten, von den Teilnehmenden immer als gleichberechtigt behandelt. Dies zeigt sich besonders bei der Budgetverteilung, bei der die Teilnehmenden sowohl die Perspektive einnehmen konnten, wie das Budget zwischen den Akteuren verteilt werden *sollte* und der Perspektive, wie das Budget nach ihrer Einschätzung unter den gegebenen Rahmenbedingungen verteilt werden *wird*. So konnte der Dissens in der Einschätzung letztlich durch Einnahme beider Perspektiven in einen Konsens aufgelöst werden. Da die Dimension der normativen Angemessenheit adressierende Argumente in diesem Gruppendelphi als legitim angesehen wurden, kam es überdies nicht zur fehlenden Anerkennung eines Argumentes mittels der Äußerung einer Unterstellung einer wertgeleiteten Argumentation.

¹⁰⁹ Auch bei diesem Item 7.1 F finden sich die Antworten der einzelnen Kleingruppen sowie die statistischen Werte im Anhang 11.10.

7.3.3.4 Dimension: grundlegende Funktionslogik

Das Gruppendelphi zu Smart Metern erfolgte mit dem Ziel zu zwei zentralen Fragestellungen lösungsorientierte Maßnahmen zu identifizieren und Erkenntnisse zu generieren:

- Wie lässt sich die Verbreitung von Smart Metern erhöhen?
- Welche ergänzenden Maßnahmen sind sinnvoll, um das Stromsparen mit Hilfe von Smart Metern zu fördern?

Auch dieser Fall muss sich im Sinne der generellen Funktionslogik eines Gruppendelphis daran bemessen lassen, ob Erkenntnisse zu handlungsleitenden Empfehlungen ausgearbeitet werden und letztlich die Fragestellungen mit fruchtbaren Antworten gefüttert werden konnten.

Dabei konnten durch den Einsatz des Gruppendelphis im Hinblick auf eine zügige Einführung von Smart Metern hilfreich Erkenntnisse gewonnen werden. So bewerteten die Teilnehmenden beispielsweise *eine Einführung von Smart Metern kombiniert mit monatlicher informativer Abrechnung und persönlichen Feedbackinstrumenten im Hinblick auf das Stromsparen* als sehr bedeutsam an. Zur Förderung der Verbreitung von Smart Metern empfahlen die Teilnehmenden Maßnahmen innerhalb eines schlüssigen Konzeptes *mit entsprechenden Gesetzen und Verordnungen, die Schaffung von mehr Klarheit und mehr Transparenz bezüglich der Einführungsgründe und der damit in Zusammenhang stehenden Kosten*. Um eine Effizienzsteigerung im Hinblick auf das Stromsparen zu erreichen, wählten die Teilnehmenden Maßnahmen, die die *Modernisierung des Eichrechts*, eine *Aktualisierung des Datenschutzes* und die *Schaffung von Anreizen für alle Beteiligten* betrafen. Ebenso wurden Maßnahmen identifiziert, die zur Förderung der Verbreitung von Smart Metern aus Sicht der Teilnehmenden als erfolgversprechend bewertet wurden, wie Smart Meter mit Preissignalen zu *tatsächlichen lastspezifischen Preisen* oder *Kostenneutralität mit einer Umlage über das Netzentgelt*. Auch einen *verpflichtenden Einbau von Smart Metern mit Datenübertragungsfunktion durch persönliche Feedbackinstrumente* bewerteten die Teilnehmenden als erfolgversprechende Maßnahme. So konnte das Gruppendelphi die angestrebten Ziele zufriedenstellend erreichen und es ist überdies gelungen, über die Integration, der von den Teilnehmenden vorgeschlagenen Präambeln einen Rahmen zu definieren, in dem sich die Maßnahmen bewegen müssen. Dies sowohl im Sinne einer effektiven und erfolgreichen Umsetzung sowie anhand der normativen Zielrichtung einer sinnvollen Förderung von erneuerbaren Energien im Rahmen der angestrebten Energiewende.

Das Fallbeispiel zeigt allerdings ebenso auf, wie wichtig eine thematische Fokussierung im Hinblick auf die für einen Workshop zur Verfügung stehende Zeit ist. So wurden am Ende beider Fragebögen (t1, t2) zwei offene Fragen gestellt, die von keiner der Kleingruppen in der zur Verfügung stehenden Zeitspanne beantwortet werden konnten. Diese offenen Fragen waren Teil des achten Abschnitts des Fragebogens, der direkte *Handlungsempfehlungen für die Bundesregierung* beinhaltete. So sollten die Expert*innen unter *Frage 8.2* beantworten, welche Handlungsempfehlungen sie der Bundesregierung geben würden, *um die Verbreitung von Smart Metern zu fördern* und ergänzend dazu in *Frage 8.3* welche Handlungsempfehlungen sie der Bundesregierung geben würden, *um Smart Meter im Hinblick auf das Stromsparen effizienter zu machen*? Letztlich zeigt sich hier auch ein Problem bei der Konzeption des Fragebogens, denn solche Handlungsempfehlungen wurden bereits durch die abgefragten Maßnahmen in den vorhergehenden Abschnitten des Fragebogens thematisch abgedeckt. Auch

wenn hier nochmals Platz für eigene Ideen und Vorschläge der Teilnehmenden bereitgestellt werden sollte, so brachten die Teilnehmenden solche Ideen unter den jeweiligen Frageblöcken unter, besonders da ihnen auch konzeptionell durch die Kategorie *Sonstige* dazu Raum gegeben wurde. Folglich wurde so durch die späte Abfrage der offenen Fragen im Fragebogen eine Redundanz hervorgerufen, die auch problematisch gewesen wäre, wenn die Teilnehmenden in den Kleingruppen mehr Zeit zur Verfügung gehabt hätten. Letztlich kann man aus Forscher*innenperspektive sogar froh sein, dass diese nicht beantwortet wurden, da dies schlimmstenfalls gar zu Widersprüchen geführt hätte.

Abbildung 42: Ausschnitt aus Frage 7.1 des Gruppendelphis zu Smart Metern (t1)

Abschnitt VII: Kommunikative Maßnahmen

Frage 7.1:

Neben einem Markteinführungsprogramm zur besseren Verbreitung der Smart Meter und Maßnahmen, die Anreize zum Stromsparen über an SM gekoppelte Tarife setzen, sind begleitende kommunikative Maßnahmen von großer Bedeutung, um das Stromsparen mittels Smart Metern im Haushalt zu fördern. Insbesondere solche kommunikativen Maßnahmen können vom BMU im Rahmen der Nationalen Klimaschutzinitiative gefördert werden.

Im Folgenden finden Sie einige Szenarien, die bezüglich der kommunikativen Vermittlung von Smart Metern denkbar wären. Vorausgesetzt wird hier, dass veränderte Rahmenbedingungen herrschen, in denen die Verbraucher über die Nutzung der Smart Meter beim Stromsparen auch Kosten sparen. Bitte bewerten Sie diese Szenarien bezüglich ihres Erfolgs hinsichtlich der Verbreitung von Smart Metern. Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass dieses Szenario überhaupt nicht erfolgversprechend ist, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass dieses Szenario sehr erfolgversprechend ist.

Für welche Zielgruppen ist das jeweilige Szenario besonders gut geeignet? Bitte tragen Sie in die Spalte der Zielgruppen die Zahl der Zielgruppe(n) ein, für die Sie das Szenario als erfolgversprechend bewerten.

Zielgruppen:

1. 18-30 jährige Paare / Singles ohne Kinder im Haushalt
2. 31-50 jährige Paare / Alleinerziehende mit Kindern im Haushalt
3. 51-65 jährige Paare / Singles ohne Kinder im Haushalt
4. Rentner
5. Hartz IV-Empfänger

Bitte beachten Sie auch die Szenarien auf der folgenden Seite.

Szenario	Szenarienbeschreibung	-	2	3	4	5	6	7	8	9	+	Für welche Zielgruppen (1-5) ist diese Maßnahme besonders gut geeignet?
7A	Klassische Medien: Fernsehen, Rundfunk, Zeitung, Flyer Problem- und handlungsorientierte Information sowie Vorbilder nutzen bei Kommunikationsangeboten: <ul style="list-style-type: none"> • Bekannte Persönlichkeiten aus Politik und Gesellschaft werben für SM; • Videoclips mit Modellfamilien, WGs usw., überzeugte Nutzer kommen zu Wort Akteure: BMU, DENA, Energieagenturen, EVUs											

Quelle: Fragebogen des Gruppendelphis zu Smart Metern (siehe Anhang 11.9)

Neben diesen beiden Fragen aus dem achten Abschnitt des Fragebogens, stellt auch die im siebten Abschnitt abgefragte *Frage 7.1* in der Nachbetrachtung des Gruppendelphis keinen idealen Fragetypus dar. In Abbildung 42 ist ein Ausschnitt aus dieser Frage dargestellt, der die erste Seite der zweiseitigen Frage abbildet. Neben dem, in der Abbildung erkennbaren, *Szenario 7A* wurden im Fragebogen noch fünf weitere Szenarien (*7B – 7F*) sowie ein offenes Item *7G* abgefragt, in dem die Teilnehmenden die Möglichkeit hatten, bei Bedarf noch ein weiteres Szenario zu bilden. Dabei soll die Abbildung nur einen optischen Eindruck der Komplexität dieser Frage vermitteln. Einen Einblick, wie umfassend die Szenarien formuliert waren, gibt Tabelle 12.

Tabelle 12: Unter Frage 7.1 zu bewertende Szenarien im Gruppendelphi zu Smart Metern

Szenario	Szenarienbeschreibung ¹¹⁰
7A	<p>Klassische Medien: Fernsehen, Rundfunk, Zeitung, Flyer</p> <p>Problem- und handlungsorientierte Information sowie Vorbilder nutzen bei Kommunikationsangeboten:</p> <ul style="list-style-type: none"> • Bekannte Persönlichkeiten aus Politik und Gesellschaft werben für SM; • Videoclips mit Modellfamilien, WGs usw., überzeugte Nutzer kommen zu Wort <p>Akteure: BMU, DENA, Energieagenturen, EVUs</p>
7B	<p>Soziale Netzwerke im Internet</p> <p>Einrichtung einer „Fan-Plattform“(Facebook o.a.) / eines Blogs, Bereitstellung von Informationen, Kommunikation zwischen den „Fans“</p> <p>Akteure: BMU, Kommunikationsagenturen, Multiplikatoren, engagierte Privatpersonen</p>
7C	<p>Persönliche Beratung / Einführung durch Servicepersonal beim Einbau und im Vorfeld des Kaufs</p> <p>Akteure: Kundenservice von EVUs, Stadtwerke, Netz- und Messstellenbetreiber, Verbraucherzentralen</p>
7D	<p>Soziale Settings für Kommunikationsangebote nutzen: Schulen, Kindergärten, Universitäten</p> <p>Infoveranstaltungen, Vorführung, praktische Nutzung, Beratungsangebote, schriftliche Materialien</p> <p>Akteure: BMU, Schulen, Kindergärten, Kommunen, Universitäten</p>
7E	<p>Soziale Settings für Kommunikationsangebote nutzen: Wohnungsbaugesellschaften und Wohnheime</p> <p>Begleitende Informationskampagne zur Einführung: Demonstration, Beratung, schriftliche Materialien, Wettbewerbe</p> <p>Akteure: BMU, DENA, Wohnungswirtschaft</p>
7F	<p>Kommunikationsangebote an Bürger durch Kommunen</p> <ul style="list-style-type: none"> • Kommunale Kampagnen mit Plakaten, Flyern u.a. • Meldestellen: Kommunale Beratung zu SM beim Wohnungswechsel: schriftlich und persönlich • Informationsbroschüren für Eigentümer beim Wohnungskauf und für Mieter <p>Akteure: BMU, Kommunen, Stadtwerke</p>
7G	<p>Halten Sie andere Szenarien für wichtig?</p> <p>Wenn ja, welche?</p>

Neben den bereits komplexen Szenarien, die neben konkreten Maßnahmen überdies auch die Akteure beinhalteten, die die jeweiligen kommunikativen Maßnahmen durchführen sollten, wurden im einleitenden Fragetext auch potentielle Zielgruppen für die Maßnahmen benannt, die auf den im Forschungsdesign vorhergehenden Methoden ausgearbeitet worden waren:

1. 18-30 jährige Paare / Singles ohne Kinder im Haushalt
2. 31-50 jährige Paare / Alleinerziehende mit Kindern im Haushalt
3. 51-65 jährige Paare / Singles ohne Kinder im Haushalt
4. Rentner
5. Hartz IV-Empfänger

¹¹⁰ Die Szenarien bilden einen Teil der Frage 7.1 des Gruppendelphis zu Smart Metern (siehe Anhang 11.9).

So sollten die Teilnehmenden des Gruppendelphis in ihrer jeweiligen Kleingruppe sowohl die Szenarien bewerten, wie auch die jeweiligen primären Zielgruppen für jedes Szenario bestimmen. Die Bewertung der einzelnen Szenarien erfolgte dabei anhand einer 10er-Skala, bei der die Teilnehmenden gebeten wurden einen Wert von 1 anzugeben, wenn Sie der Ansicht seien, dass dieses Szenario *überhaupt nicht erfolgversprechend* ist und einen Wert von 10, wenn dieses Szenario *sehr erfolgversprechend* sei. Nach diesem Schritt sollten die Teilnehmenden für jedes Item die primäre Zielgruppe bzw. Zielgruppen per Nennung der jeweiligen Zielgruppennummer für dieses Szenario angeben. Neben der komplexen Bewertung sowie der ausführlichen Ausgestaltung der Items und aufgrund der verschiedenen Bewertungsaufgaben hatte der einleitende Fragetext ebenfalls einen großen Umfang. Diese Vermischung verschiedener Bewertungsoptionen sowie verschiedener Fragetypen gebündelt als eine Frage während des Gruppendelphis abzufragen war zu komplex. So wurden die Szenarien von drei der Gruppen (1, 2 und 4) mittels der 10er-skalierten Abfrage bewertet, bei der Zuschreibung der jeweiligen Zielgruppe gaben nur zwei der Gruppen (1 und 4) ein Votum ab.

Der Austausch im Plenum (t1) zeigte auf, dass die Szenarien von den Teilnehmenden größtenteils als komplementär angesehen wurden und dass je nach Zielgruppe ein Szenario etwas wirksamer sei, aber letztlich mehrere (bis alle) Szenarien nebeneinander zu einem Erfolg führen könnten. Die zusätzlichen Benennungen von den für die Szenarien notwendigen Akteuren im Text des jeweiligen Szenarios führten zu weiterem Diskussionsbedarf der Teilnehmenden über die Eignung und Glaubwürdigkeit der einzelnen kommunizierenden Akteursgruppen. Die mit dem Einsatz dieser *Frage 7.1* erreichten Ergebnisse spiegeln sich im Zeitaufwand, der für die Bearbeitung dieser Frage in den Kleingruppen notwendig war, nicht wider.¹¹¹ Letztlich hätte ein Ranking der abgefragten Szenarien die Teilnehmenden auf effizientem Wege dazu leiten können, eine Abstufung zwischen diesen Szenarien vorzunehmen. Dadurch, dass bei jedem einzelnen Szenario ein Rating vorgenommen werden konnte, gibt es zwar Items, die besser oder schlechter abschneiden, der Fokus der Diskussion im Plenum (t1) konnte so aber nicht auf eine Abstufung geleitet werden, zumal die zusätzliche Integration der Zielgruppenabfrage die Items zusätzlich relativierte, da das jeweilige Szenario für *jene oder diese* Zielgruppe besonders bedeutend sein konnte oder nicht.

Letztlich sollte im Hinblick auf die Funktionsweise eines Gruppendelphis darauf geachtet werden, die Fragen effizient zu gestalten und den Fokus auf die zu bewertenden Sachverhalte zu beschränken und nicht wie in *Frage 7.1* geschehen, mehrere Sachverhalte und Fragetypen zu vermischen. Wie dieser Einsatz einer solch komplexen Frage für das generelle Fazit des Fallbeispiels zu bewerten ist, soll im folgenden Abschnitt weiter besprochen werden.

7.3.4 Zwischenfazit: Gruppendelphi zu Smart Metern

Das Gruppendelphi zu Smart Metern zeigt besonders auf, dass ein Fragebogen nicht zu umfassend sein darf. Dies war nicht alleinig der Abfrage der komplexen *Frage 7.1* geschuldet, sondern der Fragebogen war mit insgesamt 14 Fragen aufgeteilt auf acht Abschnitte generell zu umfassend. Da die im Gruppendelphi abgefragten Maßnahmen auf vorhergehenden Analysen mittels Literatur-, Rechtslagen- und Marktanalysen sowie Fokusgruppen mit Verbraucher*innen des Forschungsdesigns basierten, wurde im Falle des Gruppendelphis zu Smart

¹¹¹ Die Ergebnisse der Frage 7.1 finden sich im Anhang 11.10.

Metern keine Vorbefragung durchgeführt. Diese hätte durch Abbrüche oder potentielle Anmerkungen der Befragten bereits Hinweise auf einen zu zeitaufwendigen Fragebogen liefern können und überdies bereits konsensual bewertete Maßnahmen und Einschätzungen markieren können, die nicht mehr beim Gruppendelphi hätten abgefragt werden müssen. Dies hätte eine Fokussierung auf die Maßnahmen ermöglicht, bei denen Dissens zwischen den Bewertungen der Expert*innen und Experten vor dem Gruppendelphi im Vorfeld des Workshops bestanden hätte. Dass viele Items bereits konsensual von den Teilnehmenden befürwortet oder abgelehnt wurden, zeigen besonders, die in diesem Fallbeispiel abgefragten 10-er skalierten Ratingfragen. Von diesen in der ersten Kleingruppe abgefragten 38 Items wurden in den Fragebogen der zweiten Runde lediglich vier Items nochmals abgefragt. Dies deutet bereits darauf hin, dass mit der Durchführung einer Vorbefragung bereits einige Items für den eigentlichen Workshoptag hätten aussortiert werden können, da kein Austausch von Argumenten mehr notwendig gewesen wäre.

Dass bereits in der ersten Runde so viele Maßnahmen identifiziert werden konnten, die in der zweiten Runde des Gruppendelphis nicht nochmals abgefragt wurden, deutet an, dass die Aussagen in den einzelnen Fragen bereits klar formuliert wurden. Dadurch, dass die Aussagen zu den Maßnahmen nicht nur auf einer Literaturrecherche, sondern bereits in einem diskursiven Format, wie den Fokusgruppen angewandt wurden, gab es, wie in Kapitel 7.3.3.2 dargelegt wurde, kaum semantische Missverständnisse. Einzig der Begriff der Energieversorgungsunternehmen (*EVUs*) wurde in den Fokusgruppen von den Verbraucher*innen als ausreichende Akteursbezeichnung angenommen, wohingegen dieser Begriff für die im Gruppendelphi involvierten Expert*innen alle relevanten Akteursgruppen nicht spezifisch genug erfasste, so dass mit der Integration von *Frage 3.3* im zweiten Fragebogen auf diese notwendige Unterscheidung reagiert werden konnte. Die unterschiedlichen Bewertungshorizonte von Expert*innen und Verbraucher*innen zeigten sich auch bei *Frage 4.1*, bei denen die Teilnehmenden des Gruppendelphis die bereits in den Fokusgruppen diskutierten Maßnahmen ablehnten und eigene Maßnahmen unter der Kategorie *Sonstige* ausformulierten (t1), die folgend in den Fragebogen der zweiten Runde des Gruppendelphis integriert werden konnten.¹¹²

Das Fallbeispiel zeigt überdies, dass aufgrund des Austausches im Plenum die Möglichkeit besteht Fragen zurückzustellen, da der Gesamtkontext der abgefragten Sachverhalte einen Einfluss auf die einzelnen Maßnahmen hat, bzw. dazu führt, dass diese zunächst zurückgestellt werden müssen. So führte der Austausch zu *Frage 5.1* neben der normativen Einschätzung, dass progressive Tarife zu sozialen Ungerechtigkeiten führen können und daher abzulehnen sind, auch übergeordneter Frage, ob es überhaupt das primäre Ziel der Einführung von Smart Meter sei, Strom zu sparen.¹¹³ Die Teilnehmenden waren sich einig, dass das vordergründige Ziel eine Integration der *Erneuerbaren Energien* sei und damit auf Tarife ausgelegte Maßnahmen letztlich im Hinblick auf dieses Ziel zurückgestellt werden sollten. Aufgrund dieses von den Teilnehmenden geteilten Kontextes wurde diese *Frage 5.1* im Fall des Gruppendelphis zu Smart Metern für die zweite Runde des Gruppendelphis verworfen und die, allen Maßnahmen im Fragebogen vorgestellte Frage nach einer Präambel gemeinsam beschlossen. Durch die Integration der Präambeln konnten der Kontext des Gruppendelphis im Hinblick auf die bewerteten Maßnahmen sowie ein genereller Bewertungshorizont als Rahmung für die

¹¹² siehe Kapitel 7.3.3.1.

¹¹³ siehe Kapitel 7.3.3.3

im Gruppendelphi abgefragten Sachverhalte hergestellt werden. Wenn dieser also auch am Anfang des Gruppendelphis nicht eindeutig war, so konnte er ex post in der zweiten Runde des Gruppendelphis geschaffen werden. Dass die Teilnehmenden die Notwendigkeit sahen sich auf eine normative Ausrichtung der Rahmung des Gruppendelphis zu verständigen und damit prioritäre Ziele zu definieren, stellt eine Besonderheit dieses Fallbeispiels dar.

Letztlich bleibt zusammenzufassen, dass nicht nur die Anzahl der Fragen, sondern auch der Einsatz vieler verschiedener Fragetypen auf die sich die Teilnehmenden während der Kleingruppenphase einstellen mussten, bei einer Zeitspanne von 1,5 Stunden für die erste Kleingruppenphase und mit 1,25 Stunden für die zweite Kleingruppenphase zu hoch waren. Dies hat nicht nur vom Forscherteam, sondern auch von den Teilnehmenden für dieses Gruppendelphi viel Disziplin und Energie abverlangt. Ob dies auch beim in diesem Projekt durchgeführten zweiten Gruppendelphi zu Eco-Fahrassistenten der Fall war, soll auch in der Betrachtung des Falles im nächsten Kapitel untersucht werden.

7.4 ECO-FAHRASSISTENTEN

Projekt: *Soziale, öffentliche, ökonomische und politische Akzeptanz von Fördermaßnahmen zur Verbesserung der Effizienz im Bereich Elektrizitätsnutzung und von nachhaltiger Mobilität bei privaten Haushalten – Teilprojekt nachhaltige Mobilität: Eco-Fahrassistenten*

Institut: *ZIRN – Interdisziplinärer Forschungsschwerpunkt Risiko und Nachhaltige Technologieentwicklung am Internationalen Zentrum für Kultur- und Technikforschung der Universität Stuttgart*

Zielsetzung des Teilprojektes zur nachhaltigen Mobilität war es - wie auch in Kapitel 7.3 zum Teilprojekt zu Smart Metering beschrieben - einen Beitrag zur strategischen Weiterentwicklung der Klimaschutzinitiative des *BMU* zu leisten. Dabei sollte die Akzeptanz und Bewertung klimapolitischer Maßnahmen mittels Analyse unterschiedlicher Verbraucher*innengruppen untersucht und unter Einbindung relevanter Stakeholder und wissenschaftlicher Expert*innen weiterentwickelt werden. In diesem zweiten *Teilprojekt zur nachhaltigen Mobilität* wurde der Forschungsfokus auf *Eco-Fahrerassistenzsysteme* und das Praktizieren eines spritsparenden Fahrstils¹¹⁴ gelegt, da der Individualverkehr ein relevantes energiepolitisches Handlungsfeld markiert. Eco-Fahrerassistenzsysteme haben hinsichtlich eines spritsparenden Fahrstils im Gegensatz zu rein technischen Lösungen ein größeres Potential, da sie die Fahrenden beim Spritsparen einbeziehen - was allerdings auch eine aktive Umstellung des individuellen Fahrstils erfordert. Wie dies erfolgreich geschehen kann, war eine der zentralen Forschungsfragen des Teilprojektes (vgl. Renn et al. 2012b: 4f.).

¹¹⁴ „Der spritsparende Fahrstil ist durch Ausführen einiger wichtiger Praktiken gekennzeichnet. So spielen unter anderem das Schalten und Fahren bei niedrigen Drehzahlen und mäßiges Beschleunigen, vorausschauendes Fahren mit Ausrollen-lassen, gleichmäßig Fahren und Abstände einhalten und gelassene Fahrweise ebenso eine Rolle wie das Abstellen des Motors bei Haltezeiten über 20 Sekunden, das Vermeiden von hohen Geschwindigkeiten und Kurzstrecken, der sparsame Gebrauch von Klima- und sonstigen Anlagen und die regelmäßige technische Wartung des Fahrzeuges“ (Renn et al. 2012b: 4).

Die Funktionsweise und der Grad der Involviertheit der Fahrerenden ist dabei bei verschiedenen Formen von Eco-Fahrerassistenzsysteme unterschiedlich: „Manche Systeme geben Hinweise zum Spritsparen, ohne den Fahrer weiter zu involvieren. Andere Systeme geben während oder nach der Fahrt visuelle, auditive oder taktile Feedbacks zum jeweiligen Fahrstil, mit der konkreteren Absicht beim Fahrer eine Verhaltensänderung herbeizuführen“ (Renn et al. 2012b: 4). So ergab sich für das Teilprojekt als weitere zentrale Forschungsfrage die Frage danach, wie Eco-Fahrerassistenzsysteme gestaltet sein und welche Informationen sie bereitstellen müssen, um die Fahrerenden zu einem spritsparenden Fahrstil zu motivieren. Um eine breitenwirksame Förderung dieser Assistenzsysteme zu erreichen, schloss sich ebenso die Frage nach der Gestaltung von begleitenden Maßnahmen und Rahmenbedingungen einer solchen Förderung an (vgl. Renn et al. 2012b: 4f.).

Wie im Teilprojekt zu Smart Metering folgte das Forschungsdesign im Teilprojekt zur nachhaltigen Mobilität der gleichen methodischen Abfolge. So wurde im ersten Schritt eine Literaturrecherche durchgeführt, die um eine ausführliche Analyse des Marktes von Eco-Fahrerassistenzsystemen sowie der gesetzlichen Grundlagen bereichert wurde. Diese bildeten die inhaltliche und konzeptionelle Basis für insgesamt sechs Fokusgruppen und des folgenden Gruppendelphis aus dem Handlungsempfehlungen entwickelt werden sollten (Renn et al. 2012b: 8).

7.4.1 Fragestellung des Gruppendelphis

Der Einsatz des Gruppendelphis im Teilprojekt nachhaltige Mobilität diente im Ablauf des Forschungsdesigns besonders dazu, unter Einbeziehung von Expert*innen Bewertungen hinsichtlich drei zentraler Fragestellungen auszuarbeiten. Die erste Frage hatte die technische Ausgestaltung von Eco-Fahrerassistenten im Fokus, bei der es besonders darum ging zu erarbeiten, welche Informationen den Fahrerenden vermittelt werden sollten. Diese Informationen sollen dazu beitragen die Fahrerenden zu einem spritsparenden Fahrstil zu motivieren und einen solchen Fahrstil durch die Bereitstellung der Informationen überdies zu unterstützen. Die zweite zentrale Fragestellung beschäftigte sich mit der Verbreitung von Eco-Fahrerassistenten. Ausgehend hierfür war die Annahme, dass eine Verbreitung der Assistenzsysteme angestrebt werden sollte. Die Ergebnisse des Gruppendelphis sollten hierfür Maßnahmen bereitstellen, die eine Verbreitung von Eco-Fahrerassistenten befördern können. Neben Maßnahmen zur Verbreitung beschäftigte sich die dritte zentrale Fragestellung mit weiteren begleitenden Maßnahmen und den übergeordneten Rahmenbedingungen, um ebenso einen spritsparenden Fahrstil wie auch eine aktive Nutzung von Eco-Fahrerassistenten zu erhöhen (vgl. Renn et al. 2012b: 18).

7.4.2 Merkmale des Gruppendelphis

Im Vorfeld dieses Gruppendelphis fand keine Vorbefragung der teilnehmenden Expert*innen statt, da aufgrund des Forschungsdesigns bereits auf die Ergebnisse der Literaturrecherche, wie auch der Fokusgruppen, sowohl für die Konstruktion des Fragebogens des Gruppendelphis als auch auf die Formulierungen der einzelnen Maßnahmen zurückgegriffen werden konnte (vgl. Renn et al. 2012b: 18).

Das Gruppendelphi fand am 12. Juli 2011 im Hotel Le Meridien in Stuttgart statt. Insgesamt nahmen 14 Expert*innen teil. Die Rekrutierung dieser Teilnehmenden erfolgte aufgrund ihrer Expertise, indem sie sich in ihrem beruflichen Feld mit den Themenfeldern von Fahrerassistenzsystemen, persuasive Interfaces und Aspekten eines spritsparenden Fahrstils beschäftigten. So konnten fünf Vertreter der Automobilindustrie¹¹⁵, sieben Wissenschaftler*innen¹¹⁶ und zwei Vertreter von Verbänden¹¹⁷ für dieses Gruppendelphi gewonnen werden. Weitere drei Personen, die bereits zugesagt hatten, mussten ihre Teilnahme absagen, darunter jeweils eine Vertreterin wie ein Vertreter aus Verbänden und ein Journalist, dessen Profession im Rahmen der Automobilpresse lag (vgl. Renn et al. 2012b: 18). Ein Teilnehmer aus der Automobilindustrie, musste sich nach der ersten Runde des Gruppendelphis aufgrund eines konfligierenden Termins von der Veranstaltung verabschieden.

Die Teilnehmenden wurden in vier Kleingruppen von jeweils drei bis vier Personen aufgeteilt. Die Aufteilung bei diesem Gruppendelphi erfolgte nicht zufällig, da das Forscher*innenteam vermeiden wollte, dass eine Kleingruppe nur mit Personen aus der Automobilindustrie oder nur mit Personen aus der Wissenschaft besetzt sein würde, was aufgrund der Verteilungen mit einer reinen Zufallsauswahl nicht gänzlich ausgeschlossen werden konnte. Dies wurde durch die Zuteilung der Personen mittels jeweiliger nummerierter Zettel, die neben der Agenda für die Teilnehmenden mit ihren Namensschildern auf den Plätzen bereitlagen, bewusst vermieden. Bei der Zuteilung der Gruppen wurde ebenso darauf geachtet, dass mit der neuen Besetzung zur zweiten Runde ebenfalls eine institutionell homogen besetzte Gruppe vermieden werden konnte.

Der in diesem Gruppendelphi eingesetzte Fragebogen unterteilte sich in sechs Abschnitte unter denen insgesamt 14 Fragen abgefragt wurden. Die Fragen gliederten sich dabei in insgesamt 48 Items auf, die als 10er skalierte Ratingfragen konstruiert waren. Überdies fünf offene Abfragen, zwei Fragen mittels prozentualer Abfrage, und eine Rankingskala mit Auswahl der wichtigsten Maßnahme. Die 10er skalierte Ratingfragensets, wurden jeweils um eine offene Antwortkategorie zur potentiellen Nennung durch *Sonstige* ergänzt.¹¹⁸

Das Gruppendelphi am 12. Juli 2011 begann um 10:00 Uhr und endete um 18:35 Uhr. In dieser Zeitspanne konnten zwei Runden durchgeführt werden. Der Workshop begann nach einer Begrüßung der Teilnehmenden mit einer Information zu den Ergebnissen der Fokusgruppen. Die Agenda gliederte sich im Detail wie folgt auf:

10:00 Uhr	Begrüßung der Teilnehmer, Vorstellungsrunde
10:30 -11:10 Uhr	Ergebnisse der Fokusgruppen
11:10 -11:15 Uhr	Einführung zur Methode
11:15 -12:45 Uhr	Ausfüllen des Fragebogens in Kleingruppen – Runde 1
12:45 -13:45 Uhr	Mittagspause / Auswertung der Fragebögen
13:45 -15:15 Uhr	Diskussion im Plenum I
15:15 -16:30 Uhr	Ausfüllen des Fragebogens in Kleingruppen Runde 2

¹¹⁵ Adam Opel AG, BMW AG, Mercedes Benz cars -Daimler AG, Volkswagen AG.

¹¹⁶ Ifeu Institut für Energie- und Umweltforschung Heidelberg GmbH, DFKI GmbH Saarbrücken, Helmholtz-Zentrum für Umweltforschung – UFZ Leipzig, Universität Stuttgart, Universität Magdeburg, Technische Universität Dresden, Technische Universität Chemnitz.

¹¹⁷ ACE Berlin, Deutscher Verkehrssicherheitsrat e.V. Bonn.

¹¹⁸ Der im Gruppendelphi eingesetzte Fragebogen befindet sich im Anhang 11.12.

16:30 -17:15 Uhr	Kaffeepause / Auswertung der Fragebögen
17:15 -18:30 Uhr	Diskussion im Plenum II
18:35 Uhr	Ende der Veranstaltung

Die Plenardiskussion wurde bei diesem Gruppendelphi mittels Protokoll festgehalten, das mittels einer Transkription von Tonaufnahmen des ersten und zweiten Plenums erstellt wurde. Dieses Protokoll beinhaltet die jeweiligen Redebeiträge der Teilnehmenden sowie die Äußerungen der Moderation, so dass die einzelnen Argumente zu den diskutierten Sachverhalten detailliert nachvollzogen werden können. Die Ergebnisse wurden im Rahmen dieses Gruppendelphis in Form eines 10-seitigen Ergebnisprotokolls an die Teilnehmenden mit Bitte um Korrekturen und Ergänzungen zugesandt (vgl. Renn et al. 2012b: 122-132).

7.4.3 Analyse des Gruppendelphis zu Eco-Fahrassistenten

Nachdem die Fragestellung und die Merkmale dieses Fallbeispiels erläutert wurden, wird die Analyse, wie auch in den vorhergehenden Fällen, anhand der vier Dimensionen des Testdesigns erfolgen, beginnend mit der Dimension der Evidenz.

7.4.3.1 Dimension: Evidenz

Die Dimension der Evidenz folgt dem Geltungsanspruch der Gültigkeit, der an die im Gruppendelphi erzielten Ergebnisse unter diesem Bemessungskriterium gestellt wird. Werden die Qualitätsmerkmale eines handlungsleitenden Diskurses eingelöst, so müssen sich die Werte der einzelnen Items über den Verlauf auch dieses Falles idealerweise in Richtung Konsenses bewegen oder einen Konsens über einen Dissens markieren. Die Analyse beginnt auch im Falle des Gruppendelphis zu Eco-Fahrassistentensystemen mit den 10er-skalierten Ratingskalen. Auch bei den in diesem Teilprojekt eingesetzten Fragebögen bilden diese Items mit einer Zahl von 48 Items den am häufigsten eingesetzten Typus im Fragebogen der ersten Kleingruppenphase.

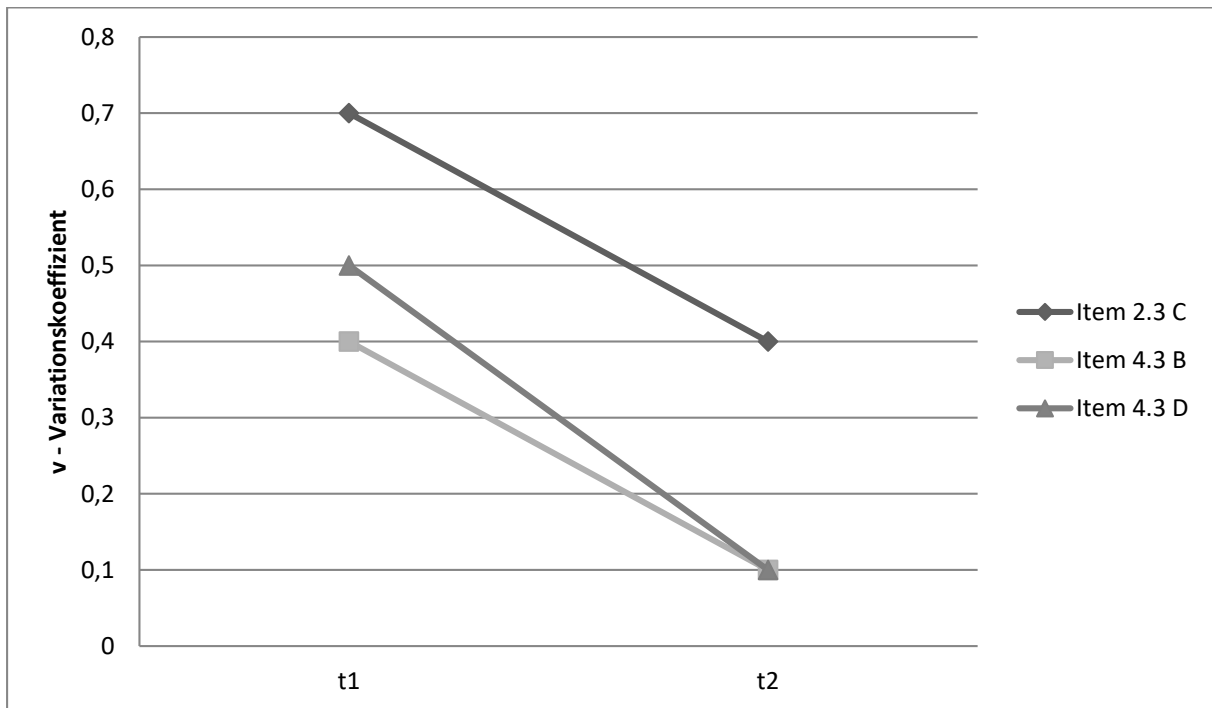
Aufgrund der im Testdesign nicht vorgesehenen Vorbefragung stehen auch in diesem Fall für den Vergleich die statistischen Werte der ersten Runde (t1) und der zweiten Runde (t2) dieses Gruppendelphis zur Verfügung.¹¹⁹ In diesem Gruppendelphi wurden jeweils vier Kleingruppen für die erste und zweite Runde des Gruppendelphis gebildet, so dass die Urteile von diesen insgesamt jeweils vier Gruppen für einen Vergleich bereitstehen. Dieser Fall ähnelt dem Gruppendelphi zu Smart Metern nicht nur darin, dass beide Fälle in einem identischen Ablauf eines Forschungsprozesses integriert waren und keine Vorbefragung stattfand, sondern auch darin, dass bei den 10er-skalierten Items nur eine sehr geringe Fallzahl, der in der ersten Runde abfragten Items auch im zweiten Fragebogen nochmals abgefragt wurden. Im Falle dieses Gruppendelphis, stehen unter Beachtung des Ausschlusskriteriums von Enthaltungen, insgesamt drei Items für einen Vergleich zur Verfügung.

Diese drei Items finden sich in Abbildung 43 in der ein Vergleich der Werte des Variationskoeffizienten über die beiden Runden des Gruppendelphis zu Eco-Fahrassistenten abgebildet ist. Richtet man den Blick auf die abgebildeten Items, so stellt das *Item 4.3 D* einen im Ideal

¹¹⁹ Eine Aufstellung der gesamten statistischen Ergebnisse beider Runden findet sich im Anhang 11.13.

angenommen sinkenden Wert des Variationskoeffizienten über die Runden dar. Während der Wert des Variationskoeffizienten noch in der ersten Runde des Gruppendelphis mit einem Wert von $v=0,5$ einen Dissens markiert, zeigt er in der zweiten Runde des Gruppendelphis mit einem Wert von $v=0,1$ ein konsensuales Ergebnis an. Eine ähnliche Entwicklung zeigt sich auch beim Item 4.3 B, bei dem der Wert des Variationskoeffizienten von $v=0,4$ (t1) ebenfalls auf einen Wert von $v=0,1$ zur zweiten Runden des Gruppendelphis sinkt.¹²⁰

Abbildung 43: Vergleich der Werte des Variationskoeffizienten über die Runden des Gruppendelphis zu Eco-Fahrassistenten



Anmerkungen: $tn_1=4$; $tn_2=4$

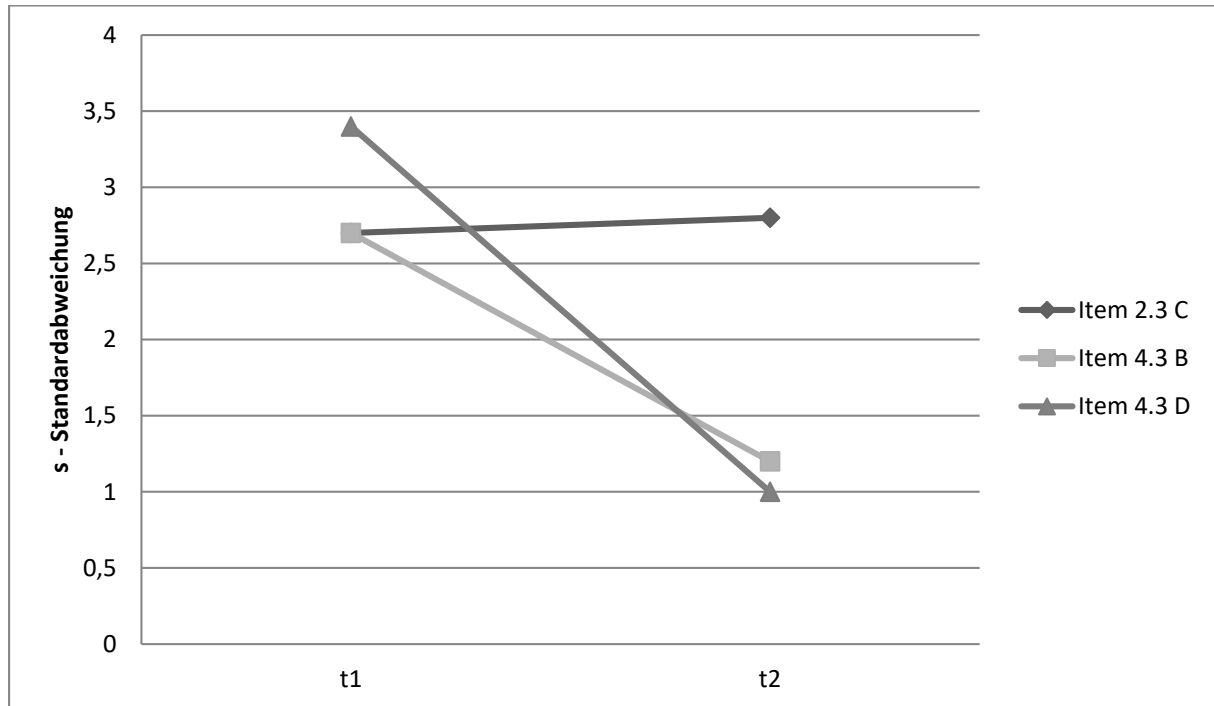
Quelle: Eigene Darstellung

Auch bei den Werten des *Items 2.3 C* zeigt sich ein deutliches Absinken des Wertes des Variationskoeffizienten über die beiden Runden des Gruppendelphis. Allerdings sinkt hier der Variationskoeffizient von $v=0,7$ (t1) lediglich auf einen Wert von $v=0,4$ (t2). Ob dies an einer konsensualen Ablehnungen dieses Items und damit einem niedrigen Wert des arithmetischen Mittels als Divisor begründet ist oder die Urteile der einzelnen Kleingruppen aufgrund eines höheren Wertes der Standardabweichung als Dividend weiter auseinanderlagen, lässt sich anhand einer exklusiven Betrachtung des Variationskoeffizienten nicht beurteilen. Deshalb soll im Folgenden die Entwicklung der Werte der Standardabweichung in die Auswertung mit einbezogen werden. Ein Vergleich dieser Werte über die Runden des Gruppendelphis zu Eco-Fahrassistenten findet sich in Abbildung 44. Hinsichtlich der *Items 4.3 B* und *4.3 D* wird die konsensuale Entwicklung zur zweiten Runde des Gruppendelphis auch unter Beachtung der Werte der Standardabweichung bestätigt. So sinkt dieser Wert beim *Item 4.3 B* von $s=2,7$ (t1) auf $s=1,2$ (t2). Auch beim *Item 4.3 D* sinkt der Wert der Standardabweichungen zwischen den beiden Runden und zwar von $s=3,4$ (t1) auf $s=1,0$ (t2). Der Wert der Standardabweichung hinsichtlich des *Items 2.3 C* deutet bereits an, dass die Einschätzungen der Teilnehmenden

¹²⁰ Warum dieses Item trotz eines Variationskoeffizienten von $v=0,4$ nochmals abgefragt wurde, wird in diesem Abschnitt bei der Betrachtung der einzelnen Antwortverteilungen näher erläutert.

sich über beide Runden des Gruppendelphis nicht angenähert haben. So steigt dieser von einem Wert von $s=2,7$ (t_1) auf einen Wert von $s=2,8$ (t_2) an.

Abbildung 44: Vergleich der Werte der Standardabweichung über die Runden des Gruppendelphis zu Eco-Fahrassistenten



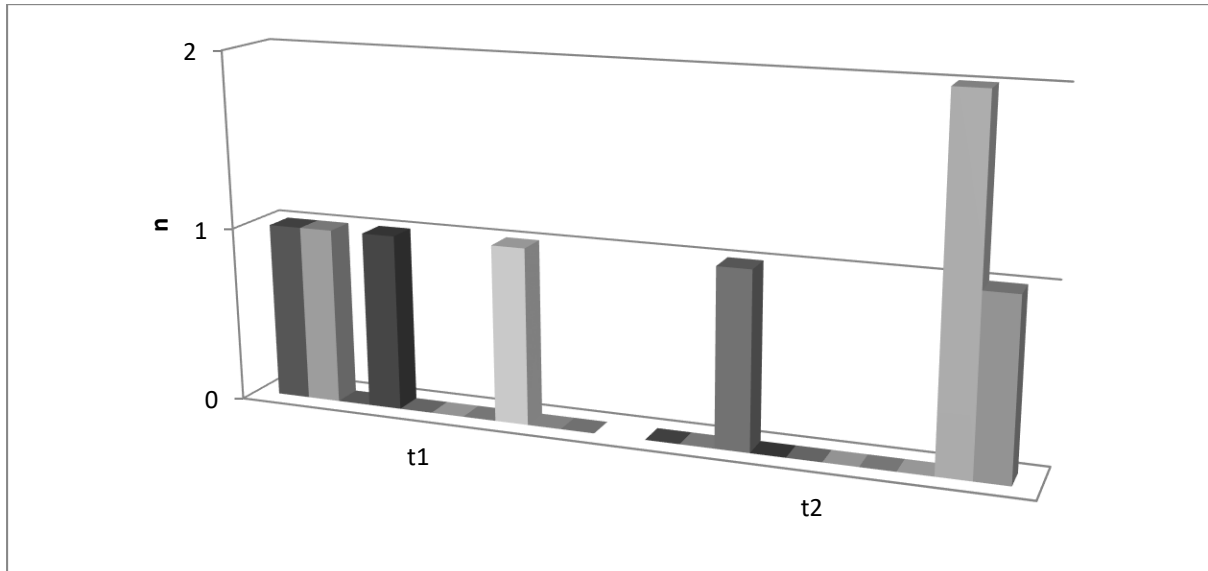
Anmerkungen: $tn_1=4$; $tn_2=4$

Quelle: Eigene Darstellung

Hier lohnt sich eine nähere Betrachtung der Antwortverteilungen, um zu beurteilen, ob sich bei diesem *Item 2.3 C* in der zweiten Runde im Vergleich zur ersten Runde ein klarer Konsens über einen Dissens herausgestellt hat oder ob die jeweiligen Gruppenantworten über die 10er-Skala verteilt liegen und diesem Item kein eindeutiges Urteil zugewiesen werden kann. Dieses Item war im zweiten Abschnitt des Fragebogens eingeordnet, in dem die *Gestaltung von Eco-Fahrassistenten* behandelt wurde und es wurde unter der *Frage 2.3* abgefragt, in der *Gestaltungsmerkmale eines Eco-Fahrassistenten, die einen spritsparenden Fahrstil unterstützen und / oder die Motivation, spritsparend zu fahren, positiv beeinflussen* zur Bewertung standen. Das Item 2.3 C lautete dabei wie folgt: *Basisinformationen zu spritsparenden Fahrstilaspekten (Beschleunigen, Ausrollen, gleichmäßig Fahren, effiziente Geschwindigkeit) werden während der Fahrt permanent rückgemeldet und sind nicht abschaltbar*. Die Teilnehmenden wurden gebeten, dieses Item mit einem Wert von 1 auf der Skala zu bewerten, wenn sie der Meinung waren, dass diese *Gestaltungsmerkmale überhaupt nicht wichtig sind* und mit einem Wert von 10 zu bewerten, wenn sie der Ansicht waren, dass diese *Gestaltungsmerkmale sehr wichtig sind*. Bezieht man die in Abbildung 45 dargestellte Verteilung der Antworten mit in die Analyse ein, so zeigt sich für die erste Runde, dass drei Gruppen dieses Gestaltungsmerkmal als unwichtig beurteilten, während Gruppe 4 dieses als wichtig wertete. Diese Frage wurde im ersten Plenum diskutiert und Gruppe 4 führte für ihr Urteil das Argument an, dass solche *Basisinformationen zu spritsparenden Fahrstilaspekten* immer als Grundvoraussetzung für einen Eco-Fahrassistenten wichtig seien. Gruppe 1, die das Item mit einem Wert von 1 auf der 10er-Skala als *überhaupt nicht wichtig* beurteilt hatte, stimmte dieser Aussage zwar zu, begründete ihr ableh-

nendes Urteil aber darauf, dass in der Formulierung der Aussage dieses Items eine *permanente* und *nicht-abschaltbare* Information enthalten war. Ein solches System würde von den meisten Fahrerenden als penetrant wahrgenommen, da ständig zu viele Detailinformationen angezeigt würden. Im Plenum konnte sich darauf geeinigt werden, dass die Formulierung¹²¹ von *permanent* in *dezent* geändert und die Ergänzung *nicht abschaltbar* entfernt wurde.

Abbildung 45: Antwortverteilung bei Item 2.3 C



Anmerkungen: tn1=4; tn2=4

Quelle: Eigene Darstellung

Die Verteilung der einzelnen Gruppenantworten zur zweiten Runde des Gruppendelphis (t2) zeigt eine eindeutigere Verteilung hinsichtlich eines Konsenses über einen Dissens. Während drei der Kleingruppen diese Gestaltungsform als sehr wichtig (9/9/10) beurteilen, stuft Gruppe 1 dieses Gestaltungsmerkmal als unwichtig ein (3). Diese Gruppe begründete diese Ablehnung damit, dass sie in dieser Art der Gestaltung zwar wie alle Gruppen eine hohe Akzeptanz der Fahrerenden sahen, durch die Möglichkeit des Abschaltens aller Funktionen aber der Grundgedanke des Systems, die Fahrerenden durchgängig zu informieren, nicht mehr gegeben sei. Dieser Konsens über einen Dissens hinsichtlich der Gestaltbarkeit blieb am Ende des zweiten Plenums bestehen.

Wenn auch bei diesem Gruppendelphi lediglich drei 10er-skalierte Items für die Überprüfung des Geltungsanspruchs der Gültigkeit über die beiden Runden (t1, t2) zur Verfügung stehen, so bietet der Fall eine Besonderheit, die ebenfalls für einen solchen Vergleich genutzt werden sollte und im idealtypischen Ablauf eines Gruppendelphis nicht vorgesehen ist. So wurde beim Gruppendelphi zu Eco-Fahrassistenten der Skalentypus von der ersten Runde des Gruppendelphis, in der Frage 5.1 als 10er-skalierte Ratingskala abgefragt wurde, zur zweiten Runde des Gruppendelphis zu einer Rankingskala umgestaltet. Frage 5.1 stellte im Fragebogen der ersten Runde eine Reihe von Argumenten zur Bewertung, die *besonders geeignet* seien, *einen spritsparenden Fahrstil zu praktizieren und zum Kauf und zur tatsächlichen Nutzung von Eco-Fahrassistenten zu motivieren*. In dieser Runde wurden die Teilnehmenden gebeten die zugrundeliegenden Themen der Argumente mit einer 1 zu bewerten, wenn sie die *Themen für überhaupt nicht wichtig* erachteten und mit einer 10, wenn sie *die Themen für sehr wichtig*

¹²¹ Diese Umformulierung wird im Kapitel 7.4.3.2 (Verständlichkeit) vertiefend dargestellt.

erachteten. In der ersten Runde des Gruppendelphis gab es hinsichtlich eines Themas bereits einen Konsens bei den Bewertungen der Kleingruppen. So wurde das *Item 5.1 D - Persönliche (finanzielle) Vorteile durch Sprit- und Kosteneinsparung* von den Teilnehmenden mit einem Wert des arithmetischen Mittels von $\bar{x} = 9,0$ mit einer Standardabweichung von $s = 0,7$ und einem Variationskoeffizienten von $v = 0,1$ bewertet. Ebenfalls durchweg als geeignet, wurde auch das *Item 5.1 G - „Grünes“ Image für Firmen und öffentliche Einrichtungen bei Ausrüstung der Dienstwagenflotte* mit einem Wert des arithmetischen Mittels von $\bar{x} = 8,0$ bewertet. Der Wert der Standardabweichung lag hier aber mit $s = 1,2$ etwas höher, wenn auch der Variationskoeffizienten von $v = 0,2$ auf ein konsensuales Ergebnis hindeutet.

So finden sich aber auch Items, die von den Teilnehmenden im ersten Fragebogen nicht konsensual bewertet wurden. So das *Item 5.1 C - Vorsorge für nachfolgende Generationen, globale Gerechtigkeit* das beispielsweise von Gruppe 1 mit einem Wert von 7 als eher wichtig beurteilt wurde, während Gruppe 4 dieses Item mit einem Wert von 1 als unwichtig beurteilte. Da die anderen beiden Kleingruppen dieses Item ebenfalls als eher unwichtig einstufen (4/2), drückt der Wert des arithmetischen Mittels trotz der Zustimmung einer Gruppe mit einem Wert von $\bar{x} = 3,5$ (t1) eher Ablehnung aus. Der bestehende Dissens zeigt sich bei der Betrachtung des hohen Wertes der Standardabweichung von $s = 2,3$ (t1) sowie des hohen Wertes des Variationskoeffizienten mit $v = 0,7$ (t1). Auch beim *Item 5.1 B - Ressourcenschonung durch Energieeinsparung* gehen die Urteile der Kleingruppen in der ersten Runde auseinander. So liegt der Wert des arithmetischen Mittels mit $\bar{x} = 5,8$ zwar höher als beim *Item 5.1 C*, allerdings beurteilte hier eine Kleingruppe (Gruppe 4) das Item mit einem Wert von 2 als unwichtig, während die anderen Kleingruppen das Items als eher wichtig beurteilten (8/6/7). Wenn auch der Wert des Variationskoeffizienten mit $v = 0,4$ noch unter dem im Ideal angenommenen Grenzwert eines konsensualen Ergebnisses liegt, zeigt auch der hohe Wert der Standardabweichung mit $s = 2,3$ bereits den Dissens zwischen den Einschätzungen der Kleingruppen an.

Aufgrund dieses hohen Wertes wurde dieses *Item 5.1 B* auch als erstes Item der *Frage 5.1* im ersten Plenum des Gruppendelphis diskutiert. Die ablehnende vierte Gruppe führte als Argument an, dass dieses Thema zwar als Argument für einzelne gesellschaftliche Gruppen relevant sei, aber nicht als zielgruppenübergreifendes Argument funktionieren könne. Dem stimmten die Teilnehmenden der anderen Kleingruppen zu, allerdings sahen sie dieses Thema für einen Großteil der Fahrerenden doch als wichtiges Argument an. Ebenfalls stimmten die Teilnehmenden, dem Argument einer Person zu, dass ökonomische Themen oft in ihrer Wirkung überschätzt werden. Im ersten Plenum stellte sich bereits bei der Diskussion dieses Items heraus, dass die Teilnehmenden in den jeweiligen Kleingruppen die per Items abgefragten Themen nicht, wie durch die 10er-Skalierung angedacht, einzelstehend bewertet hatten, sondern diese in Relation zu den anderen abgefragten Items gesetzt hatten. Überdies hatten drei Kleingruppen die abgefragten Themen ohne genauere Beachtung der potentiellen Zielgruppen, also wie angedacht, zielgruppenübergreifend bewertet, während die erste Kleingruppe die Items immer auch unter dem Blick auf bestimmte Zielgruppen (wie zum Beispiel umweltbewusste Fahrerende) beurteilt hatte. Da alle Kleingruppen die abgefragten Items in Relation zueinander bewertet hatten, wurde vorgeschlagen eine Rangliste der drei am wichtigsten zu erachteten Themen zu Argumentationen für Maßnahmen zur Förderung eines spritsparenden Fahrstils und Nutzung von Eco-Fahrassistenten zu erstellen. So wurde konsensual beschlossen die Items der *Frage 5.1* im zweiten Fragebogen als Rankingfrage abzufragen und mit dem

Zusatz zu versehen, als Bewertungshorizont die allgemeine Öffentlichkeit als Zielgruppe zu definieren. Um einen - zumindest optischen - Vergleich zwischen beiden Delphirunden (t1, t2) und den beiden eingesetzten Fragetypen mit unterschiedlichen Skalenniveaus zu ermöglichen, finden sich in Tabelle 13 sowohl die Werte des arithmetischen Mittels der jeweiligen Items in der ersten Runde sowie die Reihenfolgen der permutierten Kleingruppen in der zweiten Runde des Gruppendelphis bzw. die Auswahl von Gruppe 4 in der zweiten Delphirunde, da diese Kleingruppe statt eines Rankings lediglich eine Auswahl getroffen hat.

Tabelle 13: Vergleich zwischen den Ergebnissen der Ratingskalen (t1) und der Rankingskala (t2) zu Frage 5.1 des Gruppendelphis zu Eco-Fahrassistenzsystemen

Kommunikative Maßnahmen – Themen für Argumente:	Rating (t1)	Ranking (t2)			
	\bar{x}	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4
Klimarelevanz und CO2-Einsparung (5.1 A)	4,8	3	2		
Ressourcenschonung durch Energieeinsparung (5.1 B)	5,8	3	2	3	
Vorsorge für nachfolgende Generationen, globale Gerechtigkeit (5.1 C)	3,5				
Persönliche (finanzielle) Vorteile durch Sprit- und Kosteneinsparung (5.1 D)	9,0	1	1	1	X
Zusammenhang zwischen Verkehrssicherheit und spritsparendem Fahrstil (5.1 E)	4,8				X
Vorbildfunktion für Kinder/Andere (5.1 F)	4,3				
„Grünes“ Image (5.1 G)	8,0	2		2	X
Neuberechnung der KFZ-Steuer (5.1 H)	5,2				
Wirksamkeit von Eco-Fahrassistenten (5.1 I)	2,0				

Anmerkungen: In der zweiten Spalte der Tabelle finden sich die Werte des arithmetischen Mittels zur ersten Runde des Gruppendelphis. Die Spalten 3 bis 6 stellen die jeweilige Rangfolge der einzelnen Kleingruppen dar. Da Gruppe 1 zwei Themen auf Platz 3 platzierte, sind hier insgesamt vier Themen in dieser Spalte enthalten. Gruppe 2 wählte zwei Themen gleichberechtigt auf Platz 2, so dass hier kein dritter Platz vergeben wurde. Gruppe 4 behandelte die Frage wie eine Auswahlfrage, so dass hier nur die Auswahl von drei Themen per „X“ gekennzeichnet wurde; $t_1=4$; $t_2=4$.

Auch wenn ein direkter Vergleich unterschiedlicher Skalenniveaus nicht möglich ist, besonders unter Beachtung, dass Gruppe 4 die Frage in der zweiten Runde als Auswahlfrage interpretiert hat, so lassen sich doch durch die Gegenüberstellung Rückschlüsse auf die Entwicklung der Urteile der Teilnehmenden des Gruppendelphis ableiten. So wurde das *Item 5.1 D* zum Thema *persönlicher (finanzieller) Vorteile durch Sprit- und Kosteneinsparung* auch in der zweiten Runde des Gruppendelphis von allen vier Gruppen als wichtiges Thema erachtet. Die Gruppen 1, 2 und 3 platzierten dieses Thema auf den ersten Rang, während Gruppe 4 es ebenfalls auswählte, wenn auch hier keine Rangfolge gebildet wurde.

Auch das *Item 5.1 G* mit Thema *„Grünes“ Image*, das in der ersten Kleingruppenphase als 10er-skaliertes Item mit einem arithmetischen Mittel von $\bar{x}=8,0$ als wichtig eingestuft wurde, wurde auch in der zweiten Kleingruppenphase von den Gruppen 1 und 3 auf den zweiten Rang platziert und ebenso von Gruppe 4 ausgewählt und damit von drei Gruppen als wichtig erachtet. Gruppe 2 hat dieses Item allerdings nicht in das Ranking mit aufgenommen, sondern stattdessen mit den *Items 5.1 A* und *5.1 B* zwei Items gleichermaßen auf den zweiten Rang gesetzt und damit ebenfalls zwei Themen ausgewählt, die Nachhaltigkeit adressieren. Diese beiden

Items wurden auch von Gruppe 1 in die Rangfolge aufgenommen, hier allerdings beide gleichberechtigt auf dem dritten Rang. Ab diesem Punkt zeigen sich spätestens deutlich die Probleme, die bei einem Vergleich unterschiedlicher Skalentypen zwangsläufig auftauchen. Verfolgte die Abfrage über die 10-er skalierten Items in der ersten Runde noch eine alleinstehende Bewertung der jeweiligen Items (wenn dies auch nach Aussagen der Teilnehmenden im ersten Plenum nicht gänzlich geschehen ist), so sind die Bewertungen der Items in der zweiten Runde in jeweiliger Beziehung zueinander und in drei Kleingruppen (1-3) überdies in hierarchischer Reihenfolge erfolgt, was einen direkten Vergleich verhindert.

Aus diesen Ergebnissen lässt sich letztlich nur das *Item 5.1 D - Persönliche (finanzielle) Vorteile durch Sprit- und Kosteneinsparung* als konsensuales Ergebnis markieren, auf das sich die Teilnehmenden in beiden Gruppendelphirunden einigen konnten. Weitere Erkenntnisse dazu liefert auch der Austausch im zweiten Plenum, bei dem die Teilnehmenden nochmals geschlossen betonten, dass die *persönlichen finanziellen Vorteile* das geeignetste Argument seien, um die Nutzung von Eco-Fahrassistenzsystemen zu befördern. Dieses Ergebnis stand allerdings bereits in der ersten Runde des Gruppendelphis anhand der 10er-skalierten Abfrage fest und wurde in der folgende Runde durch den Einsatz eines anderen Skalentypus lediglich nochmals bestätigt.

Während der zweiten Plenarsitzung begründete Gruppe 2, warum sie das *Item 5.1G* in der zweiten Runde nicht mehr in ihr Ranking mit aufgenommen hatte. So kam die Gruppe während der Kleingruppenphase zu der Einschätzung, dass ein solches „*Grünes*“ *Image* weniger für die Fahrerenden der Kraftfahrzeuge wichtig sei, sondern vielmehr für die Fahrzeughersteller. Die anderen Kleingruppen stimmten zu, dass für die Hersteller ein „*Grünes*“ *Image* für den Verkauf von Fahrzeugen ein wichtiges Argument sei, dies aber nicht ausschließen könne, dass dies ebenso ein wichtiges Thema für die Fahrerenden sei. Gruppe 2 folgte diesem Argument allerdings nicht und verwies auf die beiden *Items 5.1 A - Klimarelevanz und CO2-Einsparung* und *Item 5.2 B - Ressourcenschonung durch Energieeinsparung*, die hinsichtlich einer nachhaltigen Argumentation wichtigere Themen für die Verbraucher*innen seien. Dieser Dissens blieb bestehen, die Teilnehmenden konnten sich im zweiten Plenum allerdings konsensual darauf verständigen, dass neben ökonomischen Themen auch auf Nachhaltigkeit ausgerichtete ökologische Themen wichtige Argumente seien. Letztlich bleibt trotz eines konsensualen Ergebnisses hinsichtlich des *Items 5.1 D* der Wechsel zwischen den Skalentypen als kritisch anzusehen, da die Teilnehmenden dadurch gezwungen waren in den jeweiligen Delphirunden unterschiedliche Bewertungshorizonte einzunehmen, was letztlich auch einen direkten Vergleich der Daten verhindert. Ein Indiz, dass ein solches *Umschalten* von Skalentypen auch für die Teilnehmenden nicht unproblematisch ist, zeigt die „falsche“ Interpretation der Rankingskala durch Auswahl der Items anstatt von Platzierungen von Gruppe 2 (t2).

Wie bereits weiter oben in diesem Abschnitt dargestellt, ermöglichen ebenfalls die *Items 4.3 B* und *4.3 D* einen Vergleich über beide Runden des Gruppendelphis. Da bei beiden Items eine Umformulierung zwischen erster Runde und zweiter Runde des Gruppendelphis stattgefunden hat, sollen diese im folgenden Kapitel der Dimension der Verständlichkeit tiefer betrachtet werden.

7.4.3.2 Dimension: Verständlichkeit

Dem Geltungsanspruch der Zweckmäßigkeit folgend, wird auch im Fall des Gruppendelphis zu Eco-Fahrassistenten die Frage danach untersucht, ob die Teilnehmenden über den Prozess des Gruppendelphis mit dem *Geäußerten* letztlich einen gleichen Bedeutungsinhalt verbinden.

Die Entwicklung der statistischen Werte hin zu einem konsensualen Ergebnis der *Items 4.3 B* und *4.3 D* über die beiden Runden des Gruppendelphis wurde bereits im vorhergehenden Abschnitt dargestellt.¹²² Beide Items wurden unter dem vierten Abschnitt des Fragebogens abgefragt, der eine Förderung *der Verbreitung von Eco-Fahrassistenten durch regulative Maßnahmen* thematisierte. Überdies waren beide Teile der Itematterie von *Frage 4.3* zugeordnet, die sich mit dem verpflichtenden *Einbau von Eco-Fahrassistenzsystemen in Neufahrzeuge* beschäftigte: *Bei der Verbreitung von Eco-Fahrassistenzsystemen stellt sich die Frage, ob es einen verpflichtenden Einbau in Neuwagen seitens der Hersteller geben soll. Dadurch soll die Verbreitung von Eco-Fahrassistenzsystemen sichergestellt werden. Außerdem könnte damit auch eine Standardisierung von Eco-Fahrassistenzsystemen erreicht werden.*¹²³ Die Teilnehmenden wurden bei dieser Frage gebeten, einen Wert von 1 auf der 10er-Skala zu wählen, wenn sie der Ansicht waren, *dass die Aussage überhaupt nicht zutrifft* und einen Wert von 10 anzugeben, wenn sie der Ansicht waren, *dass sie sehr zutrifft*. Die konkreten Formulierungen der Aussagen zu den Items im Fragebogen lauteten:

- **4.3 B:** *Diese Maßnahme ist politisch durchsetzbar*
- **4.3 D:** *Für Eco-Fahrassistenzsysteme sollte ein Standard entwickelt werden*

In der Diskussion im ersten Plenum argumentierte Gruppe 1, dass sie das *Item 4.3 B* ablehnend beurteilt hätte (3), da im Frageinhalt der *Frage 4.3* die mögliche Erreichung einer *Standardisierung von Eco-Fahrassistenzsystemen* thematisiert wurde. Eine solche Standardisierung wurde von der ersten Gruppe als zu große Hürde wahrgenommen, um eine politische Durchsetzbarkeit, wie in *Item 4.3 B* thematisiert konstatieren zu können. So hätten die verschiedenen Automobilhersteller ihre jeweiligen Patente auf ihre Ausführungen der Eco-Fahrassistenzsysteme, die eine generelle Standardisierung erschweren oder gar verhindern könnten. Auch wäre es begrifflich schwer zu definieren, was hinter dem Begriff eines Standards in der konkreten technischen Ausführung zu verstehen sei. Die Teilnehmenden der übrigen Kleingruppen konnten diese Problematik nachvollziehen und hatten die Frage einer konkreten Ausgestaltung eines Standards in ihrer Beantwortung in den jeweiligen Kleingruppen ausgeklammert, da eine solche Definition aufgrund ihrer hohen technischen Komplexität nicht in dem Rahmen des Gruppendelphis definiert werden könne. Da in *Item 4.3 D* ein solcher Standard überdies in der Itemformulierung thematisiert wurde, indem konkret die Entwicklung eines Standards für Eco-Fahrassistenten abgefragt wurde, konnte sich das Plenum auch bei diesem Item darauf verständigen, dass die Problematik einer möglichen Definition auch hier zuträfe. Die Teilnehmenden waren sich ebenfalls darin einig, dass der zeitlich begrenzte Rahmen des

¹²² So verminderte sich beim Item 4.3 B der Wert der Standardabweichung von $s=2,7$ (t1) auf $s=1,2$ (t2) und der Wert des Variationskoeffizienten von $v=0,4$ (t1) auf einen Wert von $v=0,1$ (t2). Eine ähnliche Entwicklung zeigten auch die Werte beim Item 4.3 D auf, bei dem der Wert der Standardabweichung von $s=3,4$ (t1) auf $s=1,0$ (t2) und der Wert des Variationskoeffizienten von $v=0,5$ (t1) auf einen Wert von $v=0,1$ (t2) gesunken sind.

¹²³ Die Frage findet sich im Fragebogen des Gruppendelphis zu Eco-Fahrassistenten im Anhang 11.12.

Gruppendelphis nicht ausreiche, eine Definition eines solchen Standards mit verbundener hoher technischer Komplexität zu formulieren. So wurde sich darauf geeinigt, als Arbeitsdefinition für die zweite Runde von einem *Minimalstandard* auszugehen, um damit die konkrete technische Ausgestaltung bewusst auszuklammern und einen gemeinsamen Bewertungshorizont zwischen den Teilnehmenden zu schaffen, der in der ersten Runde aufgrund der unterschiedlichen Gewichtungen der Problematik durch die jeweiligen Gruppen nicht gegeben war. Der Fragetext der *Frage 4.3* wurde im Fragebogen der zweiten Runde wie folgt geändert: *Bei der Verbreitung von Eco-Fahrassistenzsystemen stellt sich die Frage, ob es einen verpflichtenden Einbau in Neuwagen seitens der Hersteller geben soll. Dadurch soll die Verbreitung von Eco-Fahrassistenzsystemen sichergestellt werden. Außerdem könnte damit auch ein Minimalstandard (egal wie dieser im Detail aussieht) für Eco-Fahrassistenzsysteme erreicht werden.*¹²⁴

Letztlich beförderte der durch die Umformulierung von *Frage 4.3* geschaffene gemeinsam geteilte Bewertungshorizont der Teilnehmenden eine konsensuale Beurteilung der *Items 4.3 B* und *4.3 D*, wie sie auch anhand der in Abbildung 46 dargestellten Werte des Variationskoeffizienten nochmals nachvollzogen werden kann. Weil mit der Umformulierung der Frage allerdings nicht nur diese beiden Items betroffen waren, finden sich in der Abbildung ebenso die beiden *Items 4.3 A* und *4.3 C*. Die konkrete Formulierung der Items im Fragebogen lautete:

- **4.3 A:** *Eco-Fahrassistenzsysteme sollten verpflichtend in Neufahrzeuge eingebaut werden, um deren Verbreitung optimal zu gewährleisten.*
- **4.3 C:** *Diese Maßnahme ist erfolgversprechender als die Förderung der Verbreitung durch die Anrechnung der Eco-Fahrassistenten als Öko-Innovation im Rahmen der oben genannten EU Verordnung 443/2009.*¹²⁵

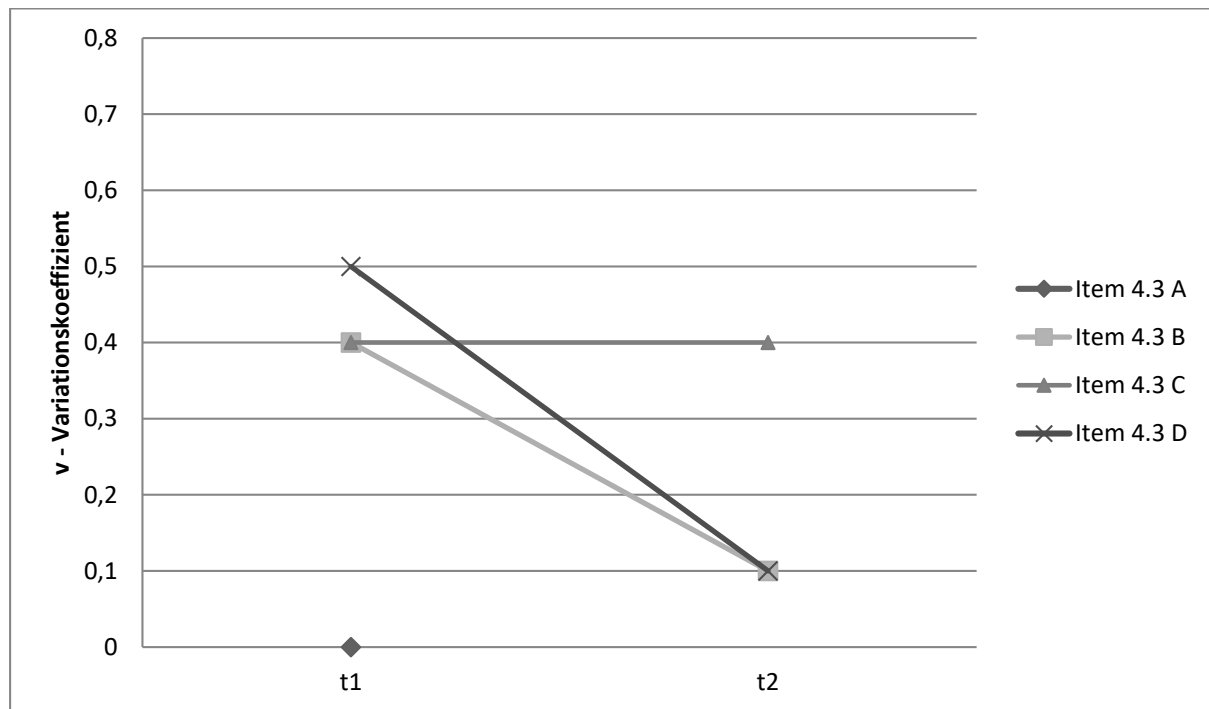
Beachtet man den Inhalt des *Items 4.3 A* so fällt auf, dass dieses Item thematisch im ersten Teil der Frageformulierung von *Frage 4.3* verankert ist und keinen direkten inhaltlichen Bezug zu einer *Standardisierung von Eco-Fahrassistenzsystemen* hat. Dieses Item ist damit das einzige Item dieser Itematterie, das nicht von der Problematik einer fehlenden Definition dieses Standards betroffen ist. Betrachtet man nun, die in Abbildung 46 dargestellten Werte des Variationskoeffizienten über die beiden Delphirunden (t_1/t_2), so fällt bereits auf, dass das *Item 4.3 A* nicht in der zweiten Runde des Gruppendelphis abgefragt wurde. Dies liegt zum einen in der konsensualen Zustimmung aller vier Kleingruppen in der ersten Runde des Gruppendelphis begründet. So wurde das *Item 4.3 A* mit einem Wert des Variationskoeffizienten von $v=0$ und einer Standardabweichung von $s=0,8$ bereits in der ersten Runde geschlossen als zutreffende Aussage bewertet und zweitens, anders als bei den restlichen drei Items (*4.3 B*, *4.3 C*, *4.3 D*), keine inhaltliche Abhängigkeit zu dem in der ersten Runde bestehenden unterschiedlichen Bewertungshorizont hinsichtlich eines Standards bestand. *Item 4.3 C*, das die Aussage, die in *Frage 4.3* beschriebene Maßnahme, *als erfolgversprechender als die Förderung der Verbreitung durch die Anrechnung der Eco-Fahrassistenten als Öko-Innovation im Rahmen der EU Verordnung 443/2009* herausstellte, wurde über beide Runden abgefragt. Denn anstatt der generellen Aussage unter *Item 4.3 A* stellt dieses *Item 4.3 C* einen Vergleich

¹²⁴ Die Kennzeichnung durch Unterstrich wurde in dieser Arbeit vorgenommen, um den geänderten Abschnitt des Fragetextes zu verdeutlichen.

¹²⁵ Verordnung (EG) Nr. 443/2009 des Europäischen Parlaments und des Rates vom 23. April 2009 zur Festsetzung von Emissionsnormen für neue Personenkraftwagen im Rahmen des Gesamtkonzepts der Gemeinschaft zur Verringerung der CO₂-Emissionen von Personenkraftwagen und leichten Nutzfahrzeugen.

an, der die Ausgestaltung der Eco-Fahrassistenten in Bezug auf eine bestehende Verordnung setzt.

Abbildung 46: Vergleich der Werte des Variationskoeffizienten bei Frage 4.3 des Gruppendelphis zu Eco-Fahrassistenten



Anmerkungen: $tn_1=4$; $tn_2=4$, bei Item 4.3 C (t2) allerdings eine Enthaltung durch Gruppe 1 sowie ein Minderheitsvotum in Gruppe 4.
Quelle: Eigene Darstellung

Damit stand das *Item 4.3 C* thematisch im direkten Bezug zum Minimalstandard. Der in der ersten Runde bestehende Dissens hinsichtlich der Bewertung dieses Items konnte mit Blick auf Abbildung 46 allerdings auch zur zweiten Runde des Gruppendelphis nicht aufgelöst werden. So blieb der Wert des Variationskoeffizienten von $v=0,4$ (t1) auch zur zweiten Runde mit $v=0,4$ bestehen, ähnlich auch die Werte der Standardabweichung, bei der $s=2,9$ (t1) nur leicht auf $s=2,7$ (t2) sinkt. Diese Werte sollten aber im Falle des *Items 4.3 C* mit Vorsicht behandelt werden, da die Fallzahl zu beiden Runden zwar $n=4$ beträgt, die Fallzahl $n=4$ (t2) in der zweiten Runde allerdings auf einem Minderheitsvotum in Gruppe 4 und einer Enthaltung durch Gruppe 1 basiert. Damit sind die Werte letztlich nicht vergleichbar und können in exklusiver Betrachtung keine Aussage hinsichtlich eines Konsenses stützen.¹²⁶ Da aber die Argumente der Teilnehmenden zu *Item 4.3 C* im zweiten Plenum Bezug auf die Umformulierung der *Frage 4.3* nehmen, wurde dieses Item mit Fokus auf die Dimension der Verständlichkeit in diesem Abschnitt aufgenommen. Die Mehrheit der Teilnehmenden war sich im zweiten Plenum darin einig, dass die in *Frage 4.3* dargestellte Maßnahme erfolversprechender wäre als die *Förderung der Verbreitung durch die Anrechnung der Eco-Fahrassistenten als Öko-Innovation im Rahmen der EU Verordnung 443/2009*. Die Person, die das Minderheitsvotum von einem Skalenwert von 3 in Gruppe 4 abgegeben hatte (t2), argumentierte hingegen, dass beides zum Zeitpunkt des Delphis nicht vergleichbar sei, da noch kein Standard definiert wäre. Ohne ge-

¹²⁶ Aus diesem Grund wurde das Item auch nicht in Kapitel 7.4.3.1 hinsichtlich der Dimension der Evidenz aufgenommen.

naue Kenntnisse über einen technischen Standard zu haben, könne man beide Lösungsansätze zur Förderung von Eco-Fahrassistenzsystemen nicht vergleichen. Der in den Fragetext integrierte *Minimalstandard* als gemeinsamer Bewertungshorizont reiche zwar aus, um die beiden anderen *Items 4.3 B* und *4.3 D* zu beurteilen, sei allerdings mit einem Vergleich einer konkreten bestehenden Verordnung zu vage. Diesem Argument folgte ebenso Gruppe 1, deren Teilnehmende schilderten, dass die Ausgestaltung eines Standards für einen solchen Vergleich unabdingbar sei und diese Gruppe sich deshalb einem Urteil verwehrt hätte. Der im zweiten Plenum bestehende Konsens über einen Dissens konnte im Verlauf der Diskussion nicht in einen Konsens überführt werden. Generell einig waren sich die Teilnehmenden hingegen insgesamt, dass eine Form der Standardisierung von Eco-Fahrassistenzsystemen zukünftig erreicht werden müsse. Mit Blick auf die Effekte der getätigten Umformulierung der *Frage 4.3* lässt sich überdies festhalten, dass die Schaffung eines gemeinsamen Bewertungshorizontes und semantischen Verständnisses des Begriffs eines *Minimalstandards* im Falle der beiden *Items 4.3 B* und *4.3 D* zu einem konsensualen Urteil der Teilnehmenden (t2) führte und bei *Item 4.3 C* eindeutig zukünftigen Forschungs- bzw. Regulierungsbedarf hinsichtlich eines Standards als Bezugsrahmen für einen Vergleich herausstellte.

Das *Item 2.3 C* wurde im vorhergehenden Abschnitt 7.4.3.1 bereits näher im Hinblick auf die Entwicklung zum eindeutigen Konsens über einen Dissens untersucht.¹²⁷ Dieses Item ist überdies auch für eine genauere Analyse unter dem Geltungskriterium der Verständlichkeit interessant, da bei diesem Item die Formulierung der Aussage über die beiden Delphirunden geändert wurde. Das *Item 2.3 C* lautete im ersten Fragebogen des Gruppendelphis: *Basisinformationen zu spritsparenden Fahrstilaspekten (Beschleunigen, Ausrollen, gleichmäßig Fahren, effiziente Geschwindigkeit) werden während der Fahrt permanent rückgemeldet und sind nicht abschaltbar.* Mit Verweis auf die Wahrnehmung der Fahrenden als störend und einen damit verbundenen Akzeptanzverlust von permanenten Fahrassistenten, wurde sich darauf geeinigt die Aussage des *Items 2.3 C* durch eine Umformulierung anzupassen, so dass es im zweiten Fragebogen des Gruppendelphis wie folgt abgefragt wurde: *Basisinformationen zu spritsparenden Fahrstilaspekten (Beschleunigen, Ausrollen, gleichmäßig Fahren, effiziente Geschwindigkeit) werden während der Fahrt dezent rückgemeldet.*¹²⁸

Das Item wurde zur zweiten Runde zwar nicht gedreht abgefragt, es wurde aber eine Abschwächung der Aussage hinsichtlich der technischen Ausgestaltung vorgenommen, die auch den Bewertungshorizont der Teilnehmenden veränderte. Dies wird auch durch den Austausch der Begründungen für die jeweiligen Urteile der Kleingruppen im zweiten Plenum inhaltlich unterstützt. Während die ablehnende erste Kleingruppe (3) ihr Urteil besonders auf das Argument stützte, dass durch die potentielle Abschaltbarkeit des Systems nun der Grundgedanke der durchgängigen Information nicht mehr gegeben sei, führten hingegen die zustimmenden Kleingruppen (9/9/10) die höhere Akzeptanz der Nutzer*innen durch ein dezentes System als Argument an. So konnte der Konsens über einen Dissens, auf den sich im zweiten Plenum

¹²⁷ So stieg der Wert der Standardabweichung bei diesem Item von $s=2,7$ (t1) minimal auf einen Wert von $s=2,8$ (t2) an und der Wert des Variationskoeffizienten verringerte sich aufgrund der höheren Zustimmungswerte der Gruppen 2,3 und 4 in der zweiten Runde von $v=0,7$ (t1) auf $v=0,4$ (t2).

¹²⁸ Die Kennzeichnung durch Unterstrich wurde in dieser Arbeit vorgenommen, um den geänderten Abschnitt des Fragetextes zu verdeutlichen.

geeinigt werden konnte, letztlich durch den veränderten Bewertungshorizont an zwei Prioritätensetzungen der Teilnehmenden festgemacht werden: Akzeptanz versus ständige Information.¹²⁹

In wieweit solche Prioritäten der Teilnehmenden auch als Werturteile, Präferenzen, (individuelle) Bewertungen von Optionen und moralische Urteile während dieses Gruppendelphis als legitime Argumentationen angenommen wurden, soll im folgenden Kapitel näher untersucht werden.

7.4.3.3 Dimension: normative Angemessenheit

Die Analyse in der Dimension der normativen Angemessenheit zielt auf die von den Teilnehmenden geäußerten Begründungen während des Plenums ab und ob es den Teilnehmenden möglich war, ihre normativen Einstellungen und Wertvorstellungen offenzulegen.

Die Plenarsitzungen dieses Gruppendelphis wurden umfassend protokolliert. Sowohl dieses Protokoll, das die einzelnen Redebeiträge der Teilnehmenden enthält, wie auch ein Ergebnisprotokoll, das an die Teilnehmenden nach dem Gruppendelphi versandt wurde, stehen für die Auswertung zur Verfügung. Gerade das erste Dokument bietet für die Analyse des Geltungsanspruchs nach Gemeinwohl und normativer Richtigkeit im Falle des Gruppendelphis zu Eco-Fahrassistenten durch die umfassende Dokumentation der einzelnen Redebeiträge mit jeweiliger Zuordnung der Person, die diese Äußerung getätigt hat, eine umfassende Grundlage für die Analyse. Spezifisch bei diesem Fallbeispiel ist, dass auch eine Zuordnung zu den jeweiligen Personen erfolgen kann, die ein Argument im Plenum geäußert haben und überdies, welche Personen, wie darauf reagiert haben.¹³⁰ So kann in diesem Fall auch nachvollzogen werden, ob einzelne Teilnehmende normative oder wertgeleitete Argumente nicht als legitim erachten oder diese übergehen bzw. nicht mit (eigenen) Argumenten erwidern.

Die konsensuale Entwicklung des *Items 2.3 C* wurde bereits in den beiden vorhergehenden Abschnitten dargestellt. Es finden sich in der Dokumentation hier auch Argumente, die für die Dimension der normativen Angemessenheit von Interesse sind. Überdies finden sich ebenso unter der *Frage 2.3* zwei weitere Items (*2.3 H*, *2.3 M*) bei deren Diskussion in den Plenarsitzungen der Geltungsanspruch nach normativer Richtigkeit oder Gemeinwohl adressiert wurde. Diese Frage wurde im zweiten Abschnitt zur *Gestaltung von Eco-Fahrassistenten* abgefragt und lautete: *Gestaltungsmerkmale eines Eco-Fahrassistenten, die einen spritsparenden Fahrstil unterstützen und / oder die Motivation, spritsparend zu fahren, positiv beeinflussen.* Da das *Item 2.3 C* als erstes Item in der Reihe der zu besprechenden Items abgefragt wurde, soll auch die Darstellung dieses Items als erstes erfolgen.

¹²⁹ Unter Frage 2.3 wurde in der zweiten Runde überdies ein neues Item 3.2 T integriert, das durch die Teilnehmenden konsensual als wirksam beurteilt wurde und einen Lösungsansatz zu diesem Konsens über einen Dissens bieten kann. Dieses Item 3.2 T wird in Abschnitt 7.4.3.4 tiefergehend behandelt.

¹³⁰ Wenn dies auch nur in anonymisierter Form erfolgen kann, da den Teilnehmenden im Sinne eines geschützten Raums für den Diskurs Anonymität hinsichtlich der einzelnen Redebeiträge zugesichert wurde. Die 14 Teilnehmenden wurden nach chronologischer Reihenfolge der Vorstellungsrunde zu Beginn dieses Gruppendelphis in diesem Protokoll mit der Bezeichnung TN1 bis TN 14 durchnummeriert, so dass für die Auswertung keine Bezüge durch die anonymisierte Form verloren gehen.

Basisinformationen zu spritsparenden Fahrstilaspekten (Beschleunigen, Ausrollen, gleichmäßig Fahren, effiziente Geschwindigkeit) werden während der Fahrt permanent rückgemeldet und sind nicht abschaltbar (Item 2.3.C). Die bereits bestehende inhaltliche Diskrepanz hinsichtlich dieses Gestaltungsmerkmals von Eco-Fahrassistenten wurde bereits in den beiden vorhergehenden Kapiteln erläutert. So dürfen Eco-Fahrassistenzsysteme nach Ansicht der teilnehmenden Expert*innen nicht aufdringlich sein, um die Akzeptanz der Nutzer*innen sicher zu stellen, sie dürfen allerdings auch nicht zu dezent (bzw. ganz abschaltbar) sein, so dass die Fahrer*innen die Informationen gänzlich ignorieren können. Dieses Dilemma spiegelte sich bereits in der Umformulierung des Items *permanent rückgemeldet* und *nicht abschaltbar* (t1) zu *dezent rückgemeldet* (t2) wider. Diese Umformulierung ist eine Folge des Austauschs zwischen den Teilnehmenden im ersten Plenum, bei dem einzelne Expert*innen bereits deutlich persönliche Präferenzen zum Ausdruck brachten. Eine Person¹³¹ stellte dabei direkt die Frage nach dem *Wollen wir das?* So äußerte TN10 während der ersten Plenarsitzung eine Frage, die direkt die Präferenzen der anderen Expert*innen adressierte: *„Wollen wir was kontinuierlich angezeigtes Dezentes haben“* (TN10, t1)? Daraufhin erwiderte TN12 im direkten Anschluss die eigene Präferenz, dass wenn *„(...) es so weit abzuschalten ist, dass man es nicht mehr wahrnimmt, dann sage ich, nein, das braucht man dann auch nicht, das kann man vergessen. Die Drehzahlanzeige ist auch immer da. Es muss im Blickpunkt des Fahrers sein“* (TN12, t1). Die Frage nach den persönlichen Präferenzen durch TN12 wurde dabei im folgenden Austausch als legitim erachtet. Die Moderation stellte an diese Argumente anknüpfend eine Nachfrage nach der möglichen technischen Ausgestaltung eines dezenten Systems, bis sich letztlich konsensual auf die neue Formulierung des *Items 2.3 C* verständigt werden konnte. In der zweiten Plenarsitzung äußerten die Teilnehmenden keine Präferenzen, allerdings individuelle Bewertungen der Optionen, wie: *„Akzeptanz käme nur, wenn es so dezent ist, dass man es übersehen kann, dann bringt es aber nichts, man kann es ignorieren, dann bringt es aber nichts mehr“* (TN3, t2). Dieser individuellen Bewertung wurde durch eine Person (TN12, t2) wieder aufgegriffen, indem diese Person inhaltlich anknüpfend argumentierte, dass wenn ein System in einer solchen Ausgestaltung abschaltbar wäre, dann bräuchte man es auch nicht. Letztlich wurden in beiden Runden sowohl Präferenzen, als auch individuelle Bewertungen von Optionen nicht von anderen Teilnehmenden als illegitime Argumente im Diskurs angezweifelt, sondern überdies wurden an diese auch in weiteren Argumentationen angeknüpft. Letztlich wurde das eingehend erwähnte Dilemma zwischen Akzeptanz und Wirksamkeit ebenso im Ergebnisprotokoll des Gruppendelphis als ausstehende Entscheidung für eine Präferenz aufgenommen. Da auch hier keine Korrekturwünsche der Teilnehmenden nach dem Gruppendelphi geäußert wurden, bestätigt dies die Annahme, dass solche individuellen Einschätzungen während dieses Gruppendelphis von allen Teilnehmenden als legitim erachtet wurden.

Integration des Eco-Fahrassistenten in ein Multifunktionsgerät, z.B. Navi, das hohe Akzeptanz erfährt (Item 2.3 H). Der Austausch zu diesem Item im ersten Plenum wurde hauptsächlich durch Evidenz adressierende Argumente der technischen Ausgestaltung und Verknüpfung

¹³¹ Anstatt der Bezeichnung *Expertin/Experte* oder *Teilnehmerin/Teilnehmer* wird in diesem Kapitel bei Redebeiträgen einzelner Teilnehmerinnen und Teilnehmer bewusst ausschließlich die Bezeichnung als *Person* bzw. im konkreten Fall als *TN#* gewählt, da eine geschlechterspezifische Bezeichnung durch die Anonymisierung nicht (mehr) möglich ist und einzelne weibliche Redebeiträge aufgrund der wenigen weiblichen Teilnehmerinnen nicht gänzlich dem Gebot der Anonymität entsprechen würden.

von externen Navigationssystemen geprägt. Allerdings äußerten auch hier einzelne Teilnehmende während des Diskurses individuelle Bewertungen zu dieser konkreten Option eines Nachrüstgerätes durch Dritthersteller. So führte eine Person hinsichtlich der Nachrüstgeräte an, dass ein solches Navigationsgerät, „(...) *nicht hundertprozentig ist, besser, als es nicht zu haben*“ (TN10, t1). Eine Person (TN2, t1) stimmte dieser Aussage zwar zu, da die Daten, die mithilfe eines Navigationsgerätes ermittelt werden könnten, *extrem wichtig* seien, allerdings wäre es überdies ebenso *extrem wichtig*, das Navigationssystem als Datenquelle für den Eco-Fahrassistenten im Fahrzeug zu nutzen. Die Probleme der technischen Verknüpfung zwischen einem externen Navigationsgerät und den Fahrzeugdaten bestimmten darauf die folgende inhaltliche Diskussion. Dabei sahen die Teilnehmenden zwei mögliche technische Lösungen: Zum einen könne das externe Navigationsgerät Daten an den im Fahrzeug verbauten Eco-Fahrassistenten liefern oder aber das Navigationsgerät Daten vom Fahrzeug erhalten und die Funktion eines Eco-Fahrassistenten übernehmen. Für beide Optionen sahen die Teilnehmenden keine bestehende technische Lösung. Die geäußerte individuelle Bewertung von TN10 wurde dabei nicht in Frage gestellt oder als illegitimes Argument abgelehnt, sondern führte letztlich zu einem Konsens darüber, dass externe Navigationssysteme zwar nicht auf die Fahrzeugdaten zugreifen könnten, allerdings eine Einbindung von Eco-Fahrinformationen in ein solches Gerät konsensual befürwortet wurde. So konnte ein bestehender Dissens im ersten Plenum aufgelöst werden, denn das *Item 2.3 H* wurde in der ersten Kleingruppenphase von Gruppe 1, 2 und 3 positiv bewertet (10/9/8), während Gruppe 4 dieses Item ablehnte (2).¹³² Aufgrund des im ersten Plenum getroffenen Konsenses, dass die Nutzung *besser als nichts* sei, wurde dieses Item zur zweiten Runde nicht mehr abgefragt. Überdies wurde Forschungsbedarf hinsichtlich der technischen Ausgestaltung einer Verknüpfung von externem Navigationsgerät und Fahrzeug konstatiert.

Vergleich mit anderen Fahrern (desselben Fahrzeugs oder Fahrzeugtyps) (Item 2.3 M). Bei diesem Item wurde der Vergleich zwischen verschiedenen Personen als mögliches Gestaltungsmerkmal eines Eco-Fahrassistenten abgefragt. Hier waren sich die Teilnehmenden in der ersten Kleingruppenphase nicht einig: So beurteilten sowohl Gruppe 1 als auch Gruppe 3 sowie ein Minderheitsvotum in Gruppe 2 dieses Gestaltungsmerkmal als wichtig (8), während Gruppe 2 dieses Gestaltungsmerkmal als eher unwichtig (5) und Gruppe 4 es als gar nicht wichtig (1) beurteilten.¹³³ Im anschließenden Plenum führten die Teilnehmenden für ihre Positionen häufig Werturteile als Argumente an. So wurde von einer Person thematisiert, dass durch dieses Item eine soziale Norm angesprochen würde, die durch den Vergleich mit anderen entstehen würde. Diese Person führte weiter aus, „(...) *vielleicht wäre eine individuelle Norm besser in dem Sinne wie war ich vor einem Monat, das erleichtert vielleicht eher ein Lernen*“ (TN8, t1). Auf dieses Argument bezog sich direkt eine Person, die diesem Werturteil widersprach: „*Der soziale Vergleich hat durchaus auch in anderen Ländern einen hohen Wert „der Beste zu sein*““ (TN10, t1). Daraufhin entgegnete wiederum eine weitere Person (TN7, t1), dass es ebenso Untersuchungen gäbe, die zeigen würden, dass auch die sparsamen Fahrer*innen sich dem allgemeinen Mittel angleichen. Diese Personengruppe sich also verschlechtern würde.

¹³² Item 2.3 H: v=0,4 (t1) und s=3,1 (t1).

¹³³ Item 2.3 M: v=0,5 (t1) und s=2,8 (t1).

Neben dieser zu Studienergebnissen rekurrierenden Argumentation, betonte diese Person (TN7, t1), dass dies aus normativer Sicht unbedingt ausgeschlossen werden solle, da so sparsame Fahrende wieder mehr verbrauchen würden, was nicht das Ziel sein könne. Nur die Nutzer*innen, die unter dem Durchschnitt lägen, hätten somit eine Motivation besser zu werden. Eine weitere Person führte als wertbasiertes Argument dagegen an, dass „(...) egal, wie es aussieht, allein durch die Bewegung auf solchen Plattformen zeige ich, dass ich mich mit dem Thema beschäftige, dann gibt es eine Wirkkette dahinter und das ist eine gute Voraussetzung“ (TN1, t1). So konnten sich die Teilnehmenden letztlich nicht darauf einigen, ob der Vergleich mit anderen Fahrenden oder ein individueller Vergleich wünschenswert bzw. wirksam wäre. Die beiden Positionen bildeten zum Ende des Austauschs einen Konsens über einen Dissens. Die Argumente der gegensätzlichen Positionen wurden zwar von der jeweiligen anderen Gruppe als legitim erachtet, allerdings nicht geteilt bzw. führten zu keiner Urteilsänderung.

Der auf *Frage 2.3* folgende dritte Abschnitt des Fragebogens thematisierte *Finanzielle Anreize zur Förderung eines spritsparenden Fahrstils* beinhaltete eine Frage dazu, wie PKW-Fahrende über *finanzielle Anreize motiviert werden können, ein Eco-Fahrassistenzsystem zu nutzen und spritsparend zu fahren?* Diese *Frage 3* führte vier Maßnahmen auf, die einen solchen Anreiz bilden könnten und die Teilnehmenden wurden gebeten, in ihren Kleingruppen (t1) diejenige Maßnahme auszuwählen, die sie als die *wichtigste Maßnahme* beurteilten. Alle vier Gruppen wählten in der ersten Kleingruppenphase das *Item 3 D* aus, bei dem als Maßnahme eine *Erhöhung der Ökosteuern und damit auch des Benzinpreises* vorgeschlagen wurde. Obwohl bei der Auswahl des Items Konsens bestand, meldeten einzelne Teilnehmende im ersten Plenum an, dass sie diese Auswahl gerne ergänzend kommentieren würden. Solche Kommentierungen fanden sich auch bereits in den beiden jeweiligen Fragebögen der Gruppen 2 und 3, die das mit *Sonstige* betitelte offene Feld im Fragebogen nicht wie angedacht nutzten, um mögliche eigene Maßnahmen zu formulieren, sondern um ihre Auswahl inhaltlich bereits im Fragebogen zu ergänzen. So fügte Gruppe 2 als schriftlichen Kommentar an: *Der Markt reguliert sich selbst. CO2 ist Verbrauch (Mineralölsteuer reicht)*. Gruppe 3 fügte ebenfalls einen knappen Kommentar an: *Heutiges System*.¹³⁴ An diese schriftlichen Kommentare knüpften, die im folgenden Plenum geäußerten Argumente der Teilnehmenden an. So äußerte eine Person (TN 12, t1) zu Beginn des Austauschs zu dieser Frage die Präferenz ihrer Gruppe 4, dass auch generell die Mineralölsteuer als finanzieller Anreiz ausreichen würde. Diese Person führte mit Verweis zu den weiteren Maßnahmen, die die KFZ-Steuer thematisierten, unter *Frage 3* aus, dass eine Trennung von KFZ- und Mineralölsteuer auf politische Gründe zurückzuführen sei, „(...) aber ob die Trennung sinnvoll ist, steht in Frage“ (TN12, t1). Daran anschließend erwiderte TN7, dass *Umwelt* ja auch Aspekte wie Lärmbelästigung und Fläche beinhalte, was über die Mineralölsteuer nicht abgedeckt würde, dies allerdings durch mögliche Maßnahmen abgedeckt werden sollte. TN12 erwiderte darauf, dass bestimmte Teilaspekte so oder so ausgeblendet würden. Eine weitere Person führte als Argument für das bestehende System an: *Ich wehre mich dagegen, der Politik neue Geldquellen zu eröffnen*“ (TN2, t1). Diese Person führte überdies als Argument an, dass ein solches System ambivalent sei, da der Staat weniger Einnahmen hätte, wenn der Verbrauch zurückginge. Dies führte TN10 argumentativ weiter auf

¹³⁴ Alle im Gruppendelphi zu Eco-Fahrassistenten schriftlichen Kommentare im Fragebogen (t1) finden sich auch unter den statistischen Ergebnissen im Anhang 11.13.

einen nötigen *kommunikativen* Aspekt: „Man sollte das Verständnis vom Zusammenhang *Mobilität und Umwelt bereits in den Schulen wecken*“ (TN10, t1). Die Teilnehmenden konnten sich aufbauend auf den im Plenum geäußerten Argumenten auf einen Konsens einigen, der darauf fußte, dass das heutige System aus Sicht der Teilnehmenden ausreiche und keine Erhöhung der Mineralölsteuer notwendig sei. Auch bei dieser *Frage 3* wurden damit Werturteile oder normative Argumente als legitim erachtet, die letztlich auch den Konsens über die Beibehaltung des Status Quo beförderten.

Der Wechsel der Abfrage von einer Ratingskala (t1) zu einer Rankingskala (t2) bei *Frage 5.1* wurde bereits in Abschnitt 7.4.3.1 erläutert. Da bei dieser Frage allerdings im zweiten Plenum beim Item eines „Grünen“ *Image* (5.1 G) von den Teilnehmenden Argumente geäußert wurden, die die Dimension der normativen Angemessenheit berühren, soll dieses Item nochmals in diesem Kapitel gesondert betrachtet werden. Die dem Item übergeordnete Frage fand sich im fünften Abschnitt des Fragebogens der *kommunikativen Maßnahmen zur Förderung eines spritsparenden Fahrstils / Nutzung von Eco-Fahrassistenten* zum Inhalt hatte. *Frage 5.1* beinhaltete dabei potentielle Argumente für die Nutzung eines Eco-Fahrassistenten, die zur Beurteilung durch die Expert*innen standen: *Motivierte und überzeugte Autofahrer werden den spritsparenden Fahrstil gewissenhafter praktizieren. Welche der nachfolgend genannten Argumente sind besonders geeignet, zum Praktizieren eines spritsparenden Fahrstils und zum Kauf und zur tatsächlichen Nutzung von Eco-Fahrassistenten zu motivieren?* Das mittels *Item 5.1 G* abgefragte mögliche Argument wurde in der ersten Runde von den Kleingruppen zwar konsensual positiv, aber unterschiedlich bewertet, indem die Gruppen 1, 2 und 3 dieses Item als wichtiges Thema beurteilten (10/8/8), während Gruppe 4 dieses Argument lediglich eher wichtiges Thema ansah (6).¹³⁵ Nach dem Wechsel zur Abfrage mittels Rankingskala, in der die drei wichtigsten Argumente in hierarchischer Reihenfolge durch die jeweiligen Gruppen ausgewählt werden sollten, bestand ein Dissens zwischen den Urteilen der Kleingruppen. So vergaben die Gruppen 1 und 3 für dieses Item den zweiten Platz, während Gruppe 4 dieses Item lediglich als eines von drei Items auswählte.¹³⁶ Gruppe 2 wählte dieses *Item 5.1* nicht als wichtiges Thema aus. Aufgrund des dadurch entstandenen Dissenses wurde das Item auch im zweiten Plenum diskutiert. Dabei führte TN7 bereits zu Beginn das individuelle Urteil an, dass ein, wie im Item abgefragtes, „Grünes“ *Image* eher für Firmen interessant sei, als für die Bürger*innen beim Kauf eines Fahrzeuges. Darauf bezog sich direkt eine Person mit einem Werturteil, das letztlich mit einem moralischen Urteil schloss: „Für Individuen kann das Grüne *Image* auch eine Kaufentscheidung sein, im Sinne einer Gewissensberuhigung“ (TN3, t2). Die anschließenden Argumente stellten folgend eher in Frage, wie es überhaupt möglich sei ein grünes *Image* von Eco-Fahrassistenzsystemen zu transportieren. Dabei schloss eine Person mit der individuellen Einschätzung: „Ich glaube nicht, dass das hier so relevant ist“ (TN7, t2). Mit der folgenden Argumentation wurde dieser Einschätzung widersprochen und Bezug zu anderen von Herstellern auf Fahrzeugen angebrachten Schriftzügen wie beispielsweise ABS als *Image* für Sicherheit genommen. Dem stimmte auch eine Person zu, die dieses *Image* überdies in Bezug zu einem moralischen Urteil als Vorbildfunktion stellte: „Wenn man bereit ist, da mehr zu zahlen, dann „tue Gutes und rede darüber“, so dass die Autohersteller das

¹³⁵ Item 5.1 G: v=0,2 (t1) und s=1,4 (t1).

¹³⁶ Wie bereits in Kapitel 7.4.3.1 dargestellt hat diese Gruppe 4 in der zweiten Runde kein Ranking erstellt, sondern lediglich Items ausgewählt.

auch vermarkten könnten, aggressiver vermarkten, solche Techniken, dann auch am Fahrzeug signalisieren ob das in Richtung Status geht, das scheint interessanter als Vorsorge für Nachfolgende Generationen oder Vorbildfunktion“ (TN7, t2). Letztlich blieben die individuellen Bewertungen der jeweiligen Personen aber im Gegensatz zu einander stehen. Wenn auch beide Positionen von der jeweilig anderen Position als legitim anerkannt wurden, so blieb das gegensätzliche Urteil über das in *Item 5.1 B* abgefragte „Grüne“ Image als Thema zur Förderung eines spritsparenden Fahrstils und zur Nutzung von Eco-Fahrassistenten bestehen.

Dass individuelle Urteile und persönliche Präferenzen allerdings Einfluss auf das Urteil anderer Teilnehmenden haben können, zeigt sich besonders anhand zweier Redebeiträge zu *Item 4.4 A*, in dem die Maßnahme zur Bewertung stand, ob *Nachrüstsysteme von den Herstellern angeboten werden sollten*. Dieses Item wurde im vierten Abschnitt zur Förderung der Verbreitung von Eco-Fahrassistenten durch regulative Maßnahmen unter *Frage 4.4* abgefragt, die lautete: *Angebot von Nachrüstsystemen bei Altfahrzeugen: Um einen spritsparenden Fahrstil breitenwirksam zu fördern, sollten nicht nur Neuwagen sondern auch Altfahrzeuge mit Eco-Fahrassistenten ausgerüstet werden können. Dafür sind Nachrüstsysteme zu entwickeln*. In der ersten Kleingruppenphase bestand bei diesem *Item 4.4 A* zwischen den Urteilen der Gruppen ein Dissens. Während Gruppe 1 der abgefragten Aussage voll und ganz zustimmte (10), stimmte Gruppe 3 dieser nur eher zu (6), Gruppe 2 lehnte diese eher ab (4) und Gruppe 4 lehnte sie ganz ab (2).¹³⁷ Die Ablehnung dieses Items wurde auch zu Beginn des ersten Plenums mit einer Präferenz einer Person begründet: *„Man sollte sich lieber auf neue Technologien konzentrieren als die alten zu optimieren“* (TN2, t1). Daran anschließend führte TN14 das Argument an, dass solche Systeme bereits auf dem Markt seien, aber nicht nachgefragt würden. Die nächste Person, die sich äußerte, schloss nicht an dieses Argument an, sondern bezog sich explizit auf die von TN2 geäußerte Präferenz: *„Unsere Gruppe hatte es als gut bewertet, ich sehe es jetzt aber genauso (...).“* (TN1, t1) wie TN2. Letztlich setzte sich diese Bewertung der *Ausstattung von Altfahrzeugen mit Nachrüstsystemen* als konsensuales Ergebnis durch, so dass sich die Teilnehmenden darauf einigten, diese Maßnahme abzulehnen und nicht in der zweiten Runde des Gruppendelphis nochmals zu bewerten. Diese Ablehnung wurde ebenso im Ergebnisprotokoll des Gruppendelphis festgehalten, dem nicht widersprochen wurde.

Bei diesem Gruppendelphi finden sich damit eine Vielzahl an Argumenten, die sich entweder auf den Geltungsanspruch des Gemeinwohls oder aber auf den Geltungsanspruch an normative Richtigkeit beziehen, wenn auch im Fall des Delphis zu Eco-Fahrassistenten häufig als individuelle Bewertungen und Präferenzen. Trotz der oft persönlichen Äußerungen mit Bezug zur normativen Angemessenheit wurden diese Argumente im Diskurs während beider Plenarphasen von allen Teilnehmenden als legitim erachtet und nicht als unzulässige Argumente thematisiert. So wurde keiner Person von einer anderen Person eine wertgeleitete Argumentation unterstellt, die als unzulässig wahrgenommen wurde. Auch findet sich häufig ein direkter Bezug auf die geäußerten Argumente in nachfolgenden Redebeiträgen.

¹³⁷ Item 4.4 A: v=0,5 (t1) und s=3,0 (t1).

7.4.3.4 Dimension: grundlegende Funktionslogik

Letztlich muss sich auch das Gruppendelphi zu Eco-Fahrassistenten daran bemessen lassen, ob für die zentralen Fragestellungen des Teilprojektes durch Einsatz der Methode fruchtbare Ergebnisse erzielt werden konnten. Dieses Gruppendelphi sollte besonders Fragen zu Gestaltungsmerkmalen von Eco-Fahrassistentensystemen klären, die zu einem spritsparenden Fahrstil motivieren und diesen befördern können. Auch sollten Maßnahmen zu Verbreitung von Eco-Fahrassistenten identifiziert werden und Rahmenbedingungen und begleitende Maßnahmen bestimmt werden, die überdies die Akzeptanz und aktive Nutzung von Eco-Fahrassistentensystemen zur Kraftstoffeinsparung erhöhen.

In der ersten Frage des Gruppendelphis wurde das *Spritsparpotential von Eco-Fahrassistentensystemen* mittels prozentueller Abfrage abgefragt. Die Itembatterie dieser Frage 1 umfasste dabei zwei Items:

- **Item 1 A:** *Wie viel Prozent Einsparung halten Sie bei der Nutzung von Eco-Fahrassistentensystemen für maximal möglich?*
- **Item 1 B:** *Wie viel Prozent Einsparung ist bei einem Durchschnittsautofahrer zu erwarten?*¹³⁸

Bei *Item 1 A* beurteilten die Gruppen 1 und 3 das maximale Einsparpotential in der ersten Kleingruppenphase mit 30%, während die Gruppen 2 und 4 das maximale Einsparpotential zu dieser Runde mit 20% angaben. Bei *Item 1 B* lagen die Prozentwerte der Gruppen noch näher beieinander. So schätzten die Gruppen 1,3 und 4 das Einsparpotential von Durchschnittsfahrenden auf 10% ein, während Gruppe zwei das Einsparpotential bei 13% schätzte. Dadurch, dass die Einschätzungen der einzelnen Kleingruppen nicht weit auseinanderlagen, bestand im ersten Plenum wenig Diskussionsbedarf hinsichtlich dieser Frage. So konnten sich die Teilnehmenden zügig auf das konsensuale Ergebnis einigen, dass sie für das *Item 1 A* ein Einsparpotential im Bereich zwischen den beiden Werten 20% und 30% sahen und beim *Item 1 B* ein Einsparpotential zwischen 10% und 13%. So konnte hier bereits in der ersten Runde rasch ein konsensuales Ergebnis erzielt werden.

Neben der Wirksamkeit von Eco-Fahrassistenten wurde überdies am Ende des Fragebogens mit *Item 6.1 A* auch die Bedeutsamkeit von Eco-Fahrassistentensystemen im sechsten Abschnitt zur *Einschätzung der Bedeutung von Eco-Fahrassistentensystemen und Handlungsempfehlungen für die Bundesregierung* abgefragt.

Tabelle 14: In der ersten Runde des Gruppendelphis zu Eco-Fahrassistenten abgefragte Bedeutsamkeit

Item #	Item	n	\bar{x}	s	v
6.1 A	Eco-Fahrassistentensysteme sind im Hinblick auf eine Reduktion des Kraftstoffverbrauchs und damit auch einer CO ₂ -Reduktion als bedeutsam zu erachten.	4	9,5	0,5	0,1

Die mit diesem Item zur Bewertung durch die Teilnehmenden gestellte Aussage lautete: *Eco-Fahrassistentensysteme sind im Hinblick auf eine Reduktion des Kraftstoffverbrauchs und damit*

¹³⁸ Hervorhebungen im Original.

auch einer CO₂-Reduktion als bedeutsam zu erachten. Wie in Tabelle 14 anhand der statistischen Kennzahlen erkennbar, waren sich die Teilnehmenden auch hinsichtlich der Bedeutsamkeit von Eco-Fahrassistenten einig und bewerteten in der Aussage abgefragte Bedeutsamkeit mit einem arithmetischen Mittel von $\bar{x}=9,5$ (t1) als sehr zutreffend.

So wurde die Sinnhaftigkeit des Systems sowohl von den *Items 1 A* und *1 B* zu Beginn des Fragebogens im Hinblick auf die Wirksamkeit, als auch durch *Item 6.1 A* zum Ende des Fragebogens im Hinblick auf die Bedeutsamkeit von den Teilnehmenden konsensual bestätigt und bildet damit eine Rahmung für die in den Abschnitten (2-4) abgefragten Maßnahmen und technischen Ausgestaltungen.

So konnte im Hinblick auf eine der zentralen Fragestellungen bereits in der ersten Kleingruppenphase des Gruppendelphis eine Reihe von förderlichen Gestaltungsmerkmalen sowie enthaltenen Informationssystemen (dritter Abschnitt des Fragebogens) identifiziert werden, die von den Teilnehmenden konsensual als sinnvoll beurteilt wurden und in Tabelle 15 zu finden sind.

Tabelle 15: In der ersten Runde des Gruppendelphis zu Eco-Fahrassistenten identifizierte Gestaltungsmerkmale

Item #	Item	n	\bar{x}	s	v
2.1 A	Informationen zu effizienter / ineffizienter Beschleunigung	4	8,3	1,9	0,2
2.1 D	Information zum rechtzeitig Ausrollen lassen anstatt stark zu bremsen	4	8,5	1,1	0,1
2.3 B	Wählbarkeit verschiedener Informationen (während und nach der Fahrt)	4	8,0	0,7	0,1
2.3 D	Feedback / Information ist situationsadäquat: (1) Autobahn, Landstraße, Stadtverkehr werden unterschieden; (2) durch GPS zugängliche Informationen werden berücksichtigt (z.B. Geschwindigkeitsbegrenzungen werden rechtzeitig mitgeteilt (ausrollen lassen); (3) Steigungen und Gefälle werden berücksichtigt.	4	9,0	1,2	0,1
2.3 I	Automatisierte Lösungen werden angeboten: (Eco-Knopf, der Gesamtsystem auf einen spritsparenden Fahrstil umstellt)	4	9,5	0,5	0,1
2.3 J	Ecoscore (in %) als Gesamtbewertung des Fahrstils und als Bewertung spezifischer spritsparender Fahrstilaspekte	4	8,5	1,1	0,1
2.3 O	Intuitive Bedienung, selbsterklärend	4	9,8	0,4	0,0
2.3 P	Konkrete, leicht verständliche Handlungsempfehlungen/ -hinweisen	4	9,8	0,4	0,0

Darunter finden sich ebenso technisch weniger komplexe Gestaltungsmerkmale, wie beispielsweise *Informationen zu effizienter und ineffizienter Beschleunigung (Item 2.1 A)* oder auch *Information zum rechtzeitig Ausrollen lassen anstatt stark zu bremsen (Item 2.1 D)*. Es finden sich allerdings auch komplexere technische Gestaltungsmerkmale in dieser Reihe von konsensual befürworteten Items der ersten Kleingruppenphase, wie beispielsweise das *Item 2.3 B: Feedback / Information ist situationsadäquat: (1) Autobahn, Landstraße, Stadtverkehr werden unterschieden; (2) durch GPS zugängliche Informationen werden berücksichtigt (z.B. Geschwindigkeitsbegrenzungen werden rechtzeitig mitgeteilt (ausrollen lassen); (3) Steigungen und Gefälle werden berücksichtigt.*

Neben diesen Items wurde von den Teilnehmenden im ersten Plenum ein weiteres Item vorgeschlagen, dass im Fragebogen der zweiten Kleingruppenphase abgefragt wurde: *Gestal-*

tungsmerkmale eines Eco-Fahrassistenten: Informationen als Basis für den Ecoscore, die zugeschaltet werden können (Item 2.3 T). Die Teilnehmenden formulierten dieses Item vor dem Hintergrund eine mögliche Lösung des Dilemmas zwischen Akzeptanz gegenüber dem technischen System und der Wirksamkeit durch Informationen zu einem spritsparenden Fahrstil zu gelangen. Dabei wurde der in *Item 2.3 J* abgefragte, mittels prozentueller Angabe für die Fahrzeugnutzer*innen angezeigte, *Ecoscore* durch weitere für die Nutzer*innen zuschaltbare Informationen ergänzt, um für interessierte Fahrende ein Mehr an Informationen bieten zu können. Zudem könnten so interessierte Nutzer*innen Informationen dazu erhalten, auf welcher Grundlage der *Ecoscore* errechnet wird. Dies führe zu mehr Transparenz und könne zusätzliches Einsparpotential für Nutzer*innen eröffnen. Das Item wurde in der zweiten Runde mit einem arithmetischen Mittel von $\bar{x}=8,8$, einer Standardabweichung von $s=0,4$ und einem Variationskoeffizienten von $v=0$ von den Teilnehmenden konsensual als sinnvolles Merkmal beurteilt und kann so neben den in Tabelle 15 aufgeführten Merkmalen ebenfalls als wertvolles Ergebnis zur Beantwortung einer der zentralen Fragestellungen eingestuft werden.

Während sich die Teilnehmenden bei *Frage 3*, wie im vorhergehenden Kapitel 7.4.3.3 dargestellt, nach einer Diskussion im ersten Plenum konsensual auf eine Beibehaltung des Status Quo einigten, finden sich auch im vierten Abschnitt des Fragebogens zur Förderung *der Verbreitung von Eco-Fahrassistenten durch regulative Maßnahmen* zwei Items, die bereits in der ersten Kleingruppenphase von allen vier Kleingruppen konsensual als geeignete Maßnahmen gekennzeichnet wurden. So konnten bereits zur ersten Runde mit den in Tabelle 16 aufgeführten *Items 4.1 A* und *4.3 A* ebenso Maßnahmen zu einer der zentralen Fragestellung des Gruppendelphis hinsichtlich der Fragestellung von Rahmenbedingungen identifiziert werden, die eine aktive Nutzung von Eco-Fahrassistenzsystemen zur Kraftstoffeinsparung befördern können.

Tabelle 16: In der ersten Runde des Gruppendelphis zu Eco-Fahrassistenten identifizierte Fördermaßnahmen

Item #	Item	n	\bar{x}	s	v
4.1 A	Wenn Eco-Fahrassistenzsysteme als Öko-Innovation angerechnet werden können, wird dadurch die Verbreitung dieser Systeme gefördert.	4	9,3	0,8	0,1
4.3 A	Eco-Fahrassistenzsysteme sollten verpflichtend in Neufahrzeuge eingebaut werden, um deren Verbreitung optimal zu gewährleisten	4	9,8	0,3	0,0

Neben Items, die von den Teilnehmenden konsensual befürwortet bzw. als sinnvoll beurteilt wurden, konnten in der ersten Runde des Gruppendelphis überdies zwei Items identifiziert werden, die von den Teilnehmenden konsensual abgelehnt wurden. Das erste dieser beiden Items war Teil der Itematterie unter *Frage 2.1*, die potentielle Informationen zur Bewertung stellte, die den Fahrzeugnutzer*innen während der Fahrt zur Verfügung gestellt werden sollten, um einen effizienten Fahrstil optimal zu unterstützen. Der in *Item 2.1 B* abgefragte Inhalt war dabei die Information zur erlaubten *Höchstgeschwindigkeit in der Stadt und auf der Landstraße*. Das zweite Item wurde dem Urteil der Teilnehmenden in *Frage 2.2* unterzogen, in welchem zusätzliche Informationen abgefragt wurden, die den Fahrerenden vermittelt werden sollten, um *den Spritverbrauch zu senken und nachhaltige Mobilität zu fördern*. In *Item 2.2 D* wurden *Informationen zu Mitfahrzentralen und Carsharing* zur Beurteilung gestellt.

Tabelle 17: In der ersten Runde des Gruppendelphis zu Eco-Fahrassistenten identifizierte Gestaltungsmerkmale

Item #	Item	n	\bar{x}	s	v
2.1 B	Erlaubte Höchstgeschwindigkeit in der Stadt und auf der Landstraße	4	2,0	0,7	0,4
2.2 D	Informationen zu Mitfahrzentrale und Carsharing	4	3,3	1,1	0,3

In Tabelle 17 finden sich die statistischen Werte beider Items von denen das arithmetische Mittel anzeigt, das sowohl das *Item 2.1 B* mit einem Wert von $\bar{x}=2,0$ (t_1) und das *Item 2.2 D* mit $\bar{x}=3,3$ (t_1) im Mittel als nicht wichtige Informationen eingestuft wurden. Unter Hinzuziehung der Standardabweichung ergibt sich überdies ein konsensuales Bild dieser Bewertungen, so finden sich sowohl beim *Item 2.1 B* mit $s=0,7$ als auch beim *Item 2.2 D* mit $s = 1,1$ alle Urteile der Kleingruppen auf der ablehnenden Seite der 10er Skala zur ersten Runde wieder. Die beiden relativ hohen Werte des Variationskoeffizienten ergeben sich auch bei diesen Items durch die, aufgrund der Ablehnung erfolgenden, niedrigen Werte des Divisors, bestehend aus dem jeweiligen niedrigen Wert des arithmetischen Mittels.

Unter Beachtung der Analyse unter den drei vorhergehenden Dimensionen dieses Fallbeispiels, führte der Einsatz der Methode dazu, dass zu allen zentralen Fragestellungen letztlich fruchtbare Ergebnisse erreicht werden konnten, die im Sinne von Handlungsempfehlungen an das beauftragende *BMU* weitergegeben werden konnten. So wurden konsensuale getragene Gestaltungsmerkmalen von Eco-Fahrassistenzsystemen identifiziert sowie auch Rahmenbedingungen und begleitende Maßnahmen bestimmt, die eine Akzeptanz und aktive Nutzung von Eco-Fahrassistenzsystemen zur Kraftstoffeinsparung erhöhen können. Schließen soll dieses Kapitel aber mit einer kritischen Betrachtung eines weiteren Fragetypus.

So wurden, ähnlich wie im Falle des Gruppendelphis zu Smart Metering (vgl. Kapitel 7.3.3), auch im Fragebogen dieses Fallbeispiels mit fünf offenen Abfragen eine große Anzahl dieses Fragetypus eingesetzt. Diese Fragen fanden sich sowohl im vierten, fünften und sechsten Abschnitt des Fragebogens:

- **Frage 4.2** (Durchführungsverordnung): *Was sollte eine Durchführungsverordnung berücksichtigen, um eine Anrechnung von Eco-Assistenten als Öko-Innovation zu ermöglichen?* Diese Frage wurde mit dem zusätzlichen Hinweis gekennzeichnet: *Bitte nur Stichpunkte.*
- **Frage 5.2** (Kommunikative Maßnahmen): Diese Frage teilte sich in drei Unterfragen, so dass die gesamte Anzahl der offenen Fragen letztlich auf eine Zahl von fünf ansteigt. Die übergeordnete Frage lautete: *Wie sollte die Einführung von Eco-Fahrassistenten kommunikativ begleitet werden, um ein Bewusstsein für die Bedeutung eines spritsparenden Fahrstils zu schaffen und deren Nutzung zu motivieren?* Auch bei dieser Frage fand sich ein Hinweis: *Notieren Sie bitte Ihre Ideen und Vorschläge im folgenden Feld!*
 - **Frage 5.2 – Teilfrage 1:** *Welche (zentralen) Akteure können hier gemeinsam / vernetzt aktiv werden?*
 - **Frage 5.2 – Teilfrage 2:** *Was könnte Ihre Institution dazu beitragen?*
 - **Frage 5.2 – Teilfrage 3:** *Welche kommunikativen Maßnahmen könnten sie jeweils durchführen?*

- **Frage 6.2:** *Welchen Beitrag kann die Klimaschutzinitiative / das BMU zur Akzeptanz eines spritsparenden Fahrstils und Nutzung von Eco-Fahrassistenten leisten? Auch hier wurde ein zusätzlicher Hinweis gegeben: Stichworte!*¹³⁹

Bereits die Auflistung, der in diesem Fallbeispiel verwendeten offenen Fragen zeigt den großen Umfang und die hohe Komplexität auf, die besonders im Hinblick auf die Agenda des Gruppendelphis zu Eco-Fahrassistenten die Herausforderung für die Kleingruppen verdeutlichen.¹⁴⁰ So hatten die Teilnehmenden des Gruppendelphis ein Zeitfenster von 90 Minuten für die erste Kleingruppenphase und in der zweiten Runde noch 75 Minuten für das gemeinsame Diskutieren und Ausfüllen des Fragebogens. Bereits während der ersten Kleingruppenphase zeichnete sich ab, dass die Zeit nicht für alle Kleingruppen ausreichen würde, alle Fragen des Fragebogens zu beantworten. Deshalb gab das wissenschaftliche Team nach 45 Minuten den Hinweis an die Gruppen, sie sollten die offenen Fragen erst zum Schluss bearbeiten und die Ranking- und Ratingfragen vorziehen. Daran hielten sich alle Kleingruppen, dies hatte aber auch zur Folge, dass bereits bei der offenen *Frage 4.2* für die Gruppe 3 in der ersten Runde die Zeit nicht ausreichte, diese zu beantworten. Bei *Frage 5.1* konnten neben Gruppe 3 überdies auch die Gruppen 1 und 2 nicht alle Teilfragen beantworten. Bei *Frage 6.2* fehlten Antworten der Gruppen 1 und 3. Da nicht von allen Gruppen Antworten vorlagen und auch die Zeit im ersten Plenum nicht ausreichte diese tiefergehend zu besprechen, wurde ein Austausch über die Antworten der einzelnen Kleingruppen auf die zweite Runde des Gruppendelphis verlegt und die Teilnehmenden gebeten, die offenen Fragen in der zweiten Kleingruppenphase nochmals zu bearbeiten.

In der zweiten Kleingruppenphase des Delphis konnten alle vier Kleingruppen die offenen Fragen beantworten. Hinsichtlich *Frage 4.2* und *Frage 6.2* wurden die Antworten der einzelnen Kleingruppen in der Nachbereitung des Gruppendelphis in das Protokoll übertragen und im Ergebnisprotokoll zusammengefasst. Bei *Frage 5.2* bekam jede Gruppe im zweiten Plenum die Möglichkeit ihre Antworten der einzelnen Teilfragen zu erläutern und ihre Vorstellung einer *kommunikativen Begleitung von Eco-Fahrassistenten* zu erläutern, *um ein Bewusstsein für die Bedeutung eines spritsparenden Fahrstils zu schaffen und deren Nutzung zu motivieren*. Dabei wurde nicht versucht ein Konsens zwischen den Teilnehmenden zu erzielen. Nicht nur aus diesem Grund kann der Einsatz von diesen offenen Fragen als sinnvolles Element für die Fragebogenkonstruktion dieses Fallbeispiels kritisch gesehen werden, weshalb das Element im Fall des Gruppendelphis zu Eco-Fahrassistenten zu Beginn des Zwischenfazits nochmals beleuchtet werden soll.

7.4.4 Zwischenfazit: Gruppendelphi zu Eco-Fahrerassistenten

Der Einsatz der offenen Fragen stellt sich im Hinblick auf die Eignung dieses Fragetypus für diesen methodischen Fall als problematisch dar. So war die Anzahl der offenen Fragen zu groß und die Zeit sowohl für die Beantwortung als auch für den Austausch im Plenum zu knapp bemessen, als dass im Sinne eines Gruppendelphis ein Austausch hin zu einem konsensualen Urteil hätte stattfinden können. Dies gilt für beide Runden des Gruppendelphis. So stellten die Teilnehmenden ihre Ergebnisse zu *Frage 5.2* lediglich im zweiten Plenum vor, während die

¹³⁹ Die Fragen finden sich so im Fragebogen im Anhang 11.12.

¹⁴⁰ Siehe Kapitel 7.4.2.

Fragen 4.2 und 6.2 überdies erst nach der Durchführung des Workshops in die Ergebnisaufbereitung einfließen. Hier stellt sich die Frage, ob eine offene schriftliche Abfrage in diesem Fall mit der Grundausrichtung der Methode vereinbar ist, wenn im Rahmen des Gruppendelphis keine Möglichkeit gegeben ist, die jeweiligen Argumente für die einzelnen Urteile auszutauschen. Letztlich bleibt die Einschätzung, dass der Einsatz dieses Fragetypus in diesem Fall nicht effizient war, da die schriftliche Abfrage viel Zeit in Anspruch genommen hat und die Teilnehmenden ihre Antworten im zweiten Plenum nochmals erläutert haben ohne in Austausch zu treten. So hätte diese Abfrage in schriftlicher Form auch nach dem Delphi oder als offene Abfrage am Ende des Plenums (t2) erfolgen können. Im Sinne der methodischen Ausrichtung eines Gruppendelphis hätte eine Integration der offenen Fragen über eine Vorbefragung erfolgen können, in der die Antworten für den eigentlichen Workshop in eine Rating- oder Rankingskala hätten überführt werden können.

Die Analyse der einzelnen Dimensionen hat auch bei einzelnen Fragen und Items im Fall des Gruppendelphis zu Eco-Fahrassistenten einzelne Hürden und Probleme identifiziert, die an dieser Stelle nochmals zusammenfassend dargestellt werden sollen. So wie im Falle von *Frage 3* des Fragebogens, bei der zwar ein Konsens zwischen den Urteilen der Teilnehmenden in der ersten Kleingruppenphase erzielt wurde, die Teilnehmenden sich allerdings nach dem Austausch im Plenum auf einen Konsens einigten, der darauf fußte, dass das heutige System aus Sicht der Teilnehmenden ausreiche und keine Erhöhung der Mineralölsteuer notwendig sei. So wurde zwar die Maßnahme *Erhöhung der Ökosteuer und damit auch des Benzinpreises (Item 3 D)* von den Teilnehmenden als potentielle Maßnahme identifiziert, allerdings nicht empfohlen. Im Sinne der Ausrichtung der Fragestellung durch konsensual getragene Maßnahmen finanzielle Anreize zu identifizieren, die Fahrzeugnutzer*innen zu einer Nutzung von Eco-Fahrassistenten motivieren, konnte nicht erreicht werden.

Bei *Frage 4* bestand fehlende Trennschärfe zwischen der Frageformulierung und den einzelnen Inhalten der Items. So kam sowohl, der in *Item 4.3 A* abgefragte *verpflichtende Einbau von Eco-Fahrassistenzsystemen* in Neufahrzeugen ebenso in der Überschrift sowie auch im Fragetext dieser Frage als Inhalt vor, wie auch die in *Item 4.3 D* zur Bewertung gestellte Entwicklung eines *Standards*. So wurden zwei in der Frage verknüpfte Inhalte in der Itematterie nochmals einzeln zur Bewertung gestellt, was ein Grund für den Dissens in der ersten Kleingruppenphase war. Eine solche fehlende Trennschärfe und unklare Trennung zwischen zwei inhaltlich separierten Items durch den einleitenden Fragetext sollte bei der Frageformulierung vermieden werden.

Die Probleme, die durch den Wechsel von einer Ratingskala hin zu einer Rankingskala zwischen zwei Runden entstehen können, konnten bei diesem Fallbeispiel anhand der *Frage 5.1* illustriert werden. So fand dieser Wechsel zwischen den Skalentypen deshalb statt, da einzelne Gruppen in der ersten Kleingruppenphase die 10er-skalierten Items in der Ratingskala in Relation zueinander gesetzt und bewertet hatten. Aufbauend darauf wurden die Items in der zweiten Runde mittels einer Rankingfrage nochmals dem Urteil der Teilnehmenden gestellt. Ein Indiz, dass ein *Umschalten* zwischen verschiedenen Bewertungshorizonten für die Teilnehmenden nicht unproblematisch ist, zeigt die Fehlinterpretation der Rankingskala durch bloße Auswahl der Items anstatt von Platzierungen von Gruppe 2. Auch für die Vergleichbarkeit der Ergebnisse beider Runden stellt dieses Vorgehen ein Problem dar. Um diese Problematik im Fall des Gruppendelphis zu Eco-Fahrassistenten zu vermeiden, hätte alternativ im

ersten Plenum ein Hinweis durch die Moderation an die Teilnehmenden erfolgen können, die Ratingskala auch als solche in der zweiten Runde mit dem identischen Bewertungshorizont nochmals zu bewerten. So hätte eine Vergleichbarkeit zwischen beide Runden sichergestellt werden können.

Letztlich konnten aber auch in diesem Fall, wie bereits im Kapitel 7.4.3.4 dargestellt, nutzbare und fruchtbare Ergebnisse ausgearbeitet werden, auch wenn die Effizienz der eingesetzten Fragetypen im Fragebogen dieses Gruppendelphis angezweifelt werden kann und zum Teil Trennschärfe vermissen ließ. Das folgende Fallbeispiel aus dem Projekt PROSUITE eignet sich folglich gut für eine Fokussierung der Betrachtung darauf, ob durch den (fast) exklusiven Einsatz von 10er-skalierten Items von einer höheren Effizienz ausgegangen werden kann.

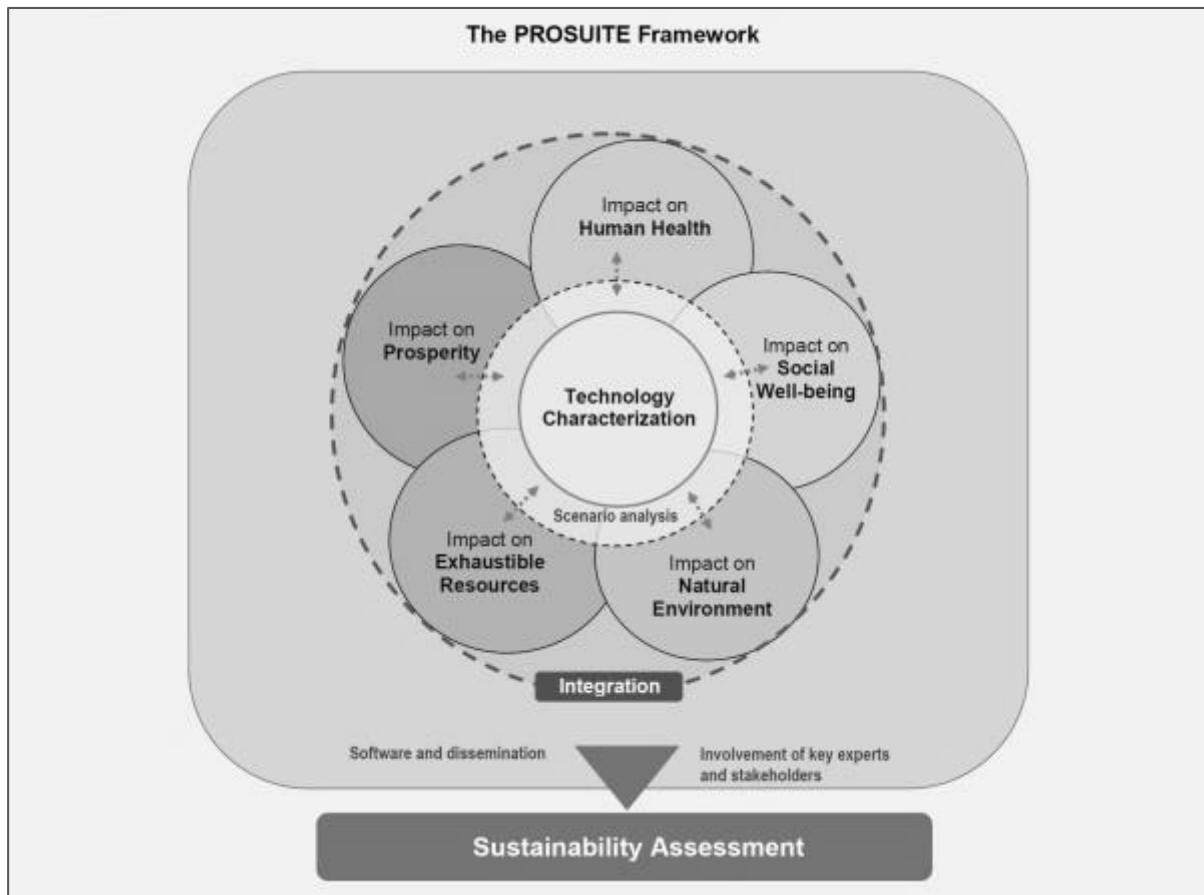
7.5 PROSUITE

Projekt: *PROSUITE – Development and application of a standardized methodology for the prospective sustainability assessment of technologies*

Institut: *DIALOGIK gemeinnützige Gesellschaft für Kommunikations- und Kooperationsforschung mbH*

Das Forschungsprojekt PROSUITE wurde im 7. EU-Forschungsrahmenprogramm gefördert und vom *Copernicus Institute, Group Energy and Resources* der *Universität Utrecht* unter Leitung von Prof. Dr. Kornelis Blok koordiniert. Ziel des aus 24 Partnerinstitutionen gebildeten Konsortiums war es, eine kohärente, wissenschaftlich fundierte Methodik zur Nachhaltigkeitsbewertung von aktuellen und zukünftigen Technologien zu schaffen, da die Einführung neuer Technologien tiefgreifende Auswirkungen sowohl auf die Gesellschaft als auch die Natur haben kann und bestehende Ansätze wie beispielsweise *Life Cycle Assessments* nicht alle Dimensionen von Nachhaltigkeit abdecken. Der Ansatz in PROSUITE hatte den Anspruch potentielle Entscheidungen über neue und existierende Produkte, Dienstleistungen und Technologien mittels retrospektiver sowie vorrausschauender Bewertungen zu unterstützen. Die Zielgruppe war dabei ebenfalls sehr umfassend, denn diese schloss sowohl Personen ein, die Entscheidungen für Unternehmen, Regierungen sowie die Gesellschaft zu treffen haben. Basierend auf einer umfassenden Literaturstudie, der Analyse von Lebenszyklen einzelner Produkte sowie verschiedener Formate der Einbindung von interdisziplinärer Expertise, wurde ein Vorschlag zu einer wissenschaftlich fundierten Methodik zur Bewertung der Nachhaltigkeit von aktuellen und zukünftigen Technologien ausgearbeitet, die sich in fünf Dimensionen aufgliedert (vgl. Blok et al. 2013: 4ff.) Das in Abbildung 47 dargestellte Messkonzept von Nachhaltigkeit wurde über die Projektlaufzeit von vier Jahren anhand von vier Beispieltechnologien ausgearbeitet. Dafür wurden *Nanotechnology*, *Multifunctional mobile devices*, *Biorefinery* sowie *Carbon (Dioxide) Capture and Storage* als Fallbeispiele gewählt.

Abbildung 47: Das Messkonzept von PROSUITE



Quelle: Ramírez et al. 2013: 2

Für diese Beispieltechnologien wurden jeweils aktuelle Technologien mit den jeweilig vermuteten Zukunftstechnologien verglichen. So wurde die konventionelle Textilproduktion mit einem erwarteten Einsatz von Nanopartikeln in der Textilproduktion im Bereich der Nanotechnologie verglichen. Bei der Mobilfunktechnologie wurden konventionelle Mobilfunkgeräte mit Smartphones kontrastiert. Kohlekraftwerke galten sowohl für Bioraffinerien als auch für Carbon Dioxid Speicherung und Einlagerung als Ausgangstechnologie, in der zum einen die anaerobe Gärung und zum anderen Kohlekraftwerke mit einer gekoppelten Speicherung und Einlagerung von Carbon Dioxid gespiegelt wurden (vgl. Gasbeek/Meijer 2013: 11).

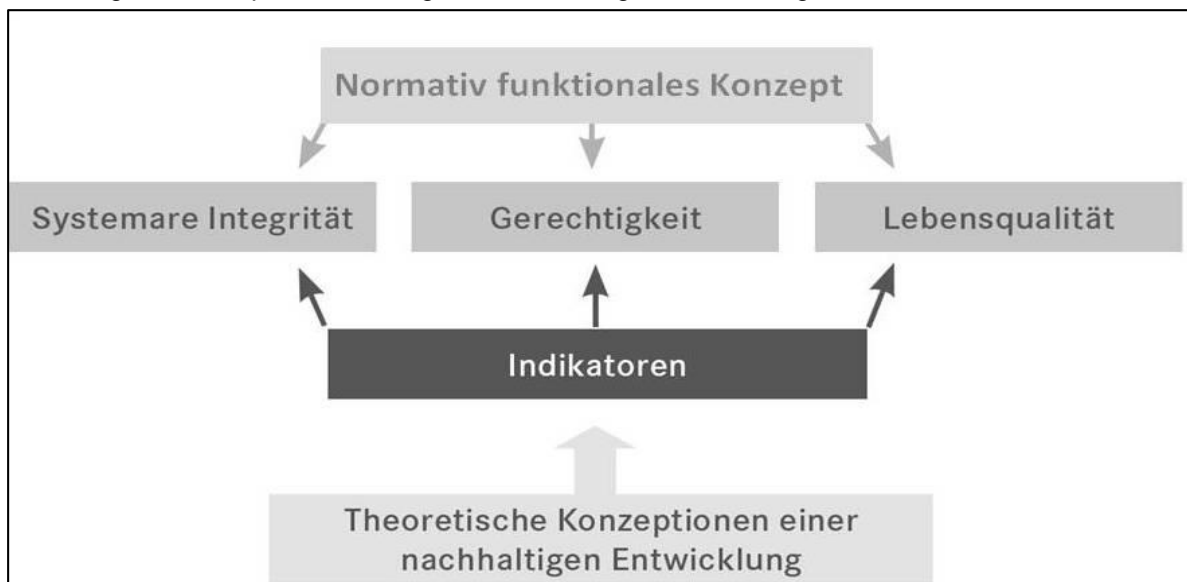
7.5.1 Fragestellung des Gruppendelphis

Für ein Gruppendelphi in PROSUITE wurde *Carbon Dioxide Capture and Storage-Technologien* (kurz CCS) als Beispieltechnologie ausgewählt. Mittels Literaturrecherche wurden in diesem Projekt aus wissenschaftlichen Artikeln, Berichten und verschiedenen externen Wissenschaftsinitiativen mehr als 600 Indikatoren zur Messung der sozialen Dimension des Messkonzeptes von Nachhaltigkeit identifiziert. Diese Anzahl der Indikatoren konnte in einem zweiten Analyseschritt auf eine Liste von 80 Indikatoren verdichtet werden, indem neben der Streichung von redundanten Indikatoren zwischen den jeweiligen Sets beispielsweise Prüfkriterien wie Messbarkeit oder Datenverfügbarkeit angewandt wurden. Diese Indikatoren sollten im

nächsten Schritt mithilfe der Einbindung von Expert*innen auf eine finale Auswahl von Indikatoren reduziert werden, die eine Messung der sozialen Dimension von Nachhaltigkeit für die in PROSUITE betrachteten Fälle ermöglichte (vgl. Renn et al. 2012c: 5). Diesen methodischen Schritt leistete das hier betrachtete Gruppendelphi.

Die Indikatoren wurden für diesen Zweck nach dem normativ-funktionalen Konzept von Nachhaltigkeit und dessen drei Dimensionen strukturiert. Die drei Dimensionen können dabei auf der funktionalen Ebene als voneinander unabhängig betrachtet werden, sie können aber überdies auch als Leitkriterien verstanden werden und beinhalten folglich ebenso eine normative Komponente. Durch eine Verknüpfung des funktionalen mit einem normativen Bemessungsraum soll verhindert werden, dass das Modell weder eine utopische Funktion noch eine rein technokratische Lösung abbildet. Da eine Vielzahl der Konzepte zur Messung und Bestimmung einer nachhaltigen Entwicklung auf einem *Drei-Säulen-Konzept* aufbauen, muss eine Verknüpfung zum normativ-funktionalen Konzept geschaffen werden. Diese erfolgt über die Zuschreibung der bestehenden Indikatoren zu den jeweiligen Dimensionen des normativ-funktionalen Konzeptes. Die erste Dimension der *Systemaren Integrität* soll dabei die Kontinuität und den Fortbestand ökologischer und sozialer Systeme abbilden, während die zweite Dimension Fragen nach der Gerechtigkeit im Hinblick auf Chancen-, Leistungs-, Bedarfs-, und Generationengerechtigkeit abbildet. Die dritte Dimension der Lebensqualität sammelt dabei nicht nur objektive Indikatoren, sondern wird um Indikatoren ergänzt, die eine Wahrnehmung und Bewertung des eigenen Lebens miteinschließen. Die Auswahl der Indikatoren sollte dabei, wie in Abbildung 48 dargestellt, auf theoretischen Konzeptionen beruhen (vgl. Jäger et al. 2007: 39-67; Kuhn et al. 2013: 42f.).

Abbildung 48: Konzept der Messung einer nachhaltigen Entwicklung



Quelle: Kuhn et al. 2013: 43.

Um soziale Indikatoren zur Messung von Nachhaltigkeit für ein umfassendes Bewertungskonzept verschiedener Technologien zu identifizieren und zu evaluieren, sollte das Gruppendelphi einen methodischen Rahmen bieten, um Bewertungen von Expert*innen zu der Auswahl der Indikatoren zu erheben und einzelne Indikatoren auf ihre Sinnhaftigkeit und Verwendbarkeit zu prüfen.

7.5.2 Merkmale des Gruppendelphis

Das Gruppendelphi in PROSUITE fand am 12. Dezember 2011 im Hotel Le Méridien in Stuttgart statt. Da vor dem CCS-Delphi eine umfassende Literaturrecherche stattgefunden hatte, wurden vom Forscher*innenteam aus der vorhandenen Liste der 80 Indikatoren für das Gruppendelphi diejenigen Indikatoren ausgewählt, die für den Fall CCS anwendbar und für ein solches Messkonzept zutreffend waren. Diese Indikatoren wurden vom Englischen ins Deutsche übersetzt und in ihrer direkten Formulierung den Expert*innen zur Bewertung mittels Fragebogen zugänglich gemacht. Eine Vorbefragung fand im Falle des Gruppendelphis in PROSUITE nicht statt. Der beim Workshop eingesetzte Fragebogen unterteilte sich in insgesamt vier Abschnitte. Der erste Abschnitt des Fragebogens führte *generelle Indikatoren* auf, die *generell zur Bemessung sozialer Nachhaltigkeit konstruiert wurden*. Im zweiten Abschnitt wurden *Organisations-spezifische Indikatoren* abgefragt, die *organisatorische Aspekte der Nachhaltigkeit erfassen sollen*. Der dritte Abschnitt umfasste *Produkt-spezifische Indikatoren*, die *Aspekte der Technik auf soziale Nachhaltigkeit erfassen sollen*. In diesen drei Abschnitten konnten die Expert*innen, die insgesamt 34 vorselektierten Indikatoren mittels 10er-skalierten Ratingfragen bewerten. In einem vierten Abschnitt wurde den Teilnehmenden mittels offener Abfrage die Möglichkeit gegeben, weitere Indikatoren für die folgende Delphirunde anzugeben. Diese Abfrage war thematisch an die vorhergehenden Abschnitte anknüpfend, in drei Unterfragen aufgeteilt, in denen die Teilnehmenden generelle Indikatoren, organisations-spezifische Indikatoren oder produkt-spezifische Indikatoren vorschlagen konnten.¹⁴¹

Die neun rekrutierten Personen nahmen an dem Gruppendelphi vollzählig teil. Diese wurden zufällig in drei Kleingruppen zu jeweils drei Personen aufgeteilt. Zur zweiten Runde des Gruppendelphis folgte eine Permutation der Teilnehmenden der einzelnen Kleingruppen. Das Gruppendelphi war als eintägiger Workshop konzipiert und gliederte sich in folgende Programmpunkte auf:

09:30 Uhr	Begrüßung und Vorstellung
09:45 Uhr	Einführung in das Gruppendelphi
10:00 Uhr	Bearbeitung des ersten Fragebogens
11:30 Uhr	Mittagspause
12:30 Uhr	Plenarsitzung: Diskussion der Ergebnisse
13:45 Uhr	Bearbeitung des zweiten Fragebogens
14:30 Uhr	Kaffeepause
15:00 Uhr	Plenarsitzung: Diskussion der Ergebnisse
16:00 Uhr	Fazit und Verabschiedung
16:15 Uhr	Ende des Gruppendelphis

Die beiden Delphirunden konnten an diesem Tag wie geplant stattfinden. Beide Plenarsitzungen wurden protokolliert und mit den statistischen Ergebnissen des Gruppendelphis zu einem Ergebnisdokument zusammengeführt. Bei diesem Gruppendelphi fand keine Feedbackschleife (z.B. mittels Versenden des Ergebnisprotokolls an die Teilnehmenden) statt.

¹⁴¹ Der im Gruppendelphi eingesetzte Fragebogen befindet sich im Anhang 11.15.

7.5.3 Analyse des Gruppendelphis in PROSUITE

Dem Ablaufschema dieser Arbeit folgend, werden nach der Beschreibung der Merkmale des Gruppendelphis in PROSUITE nun ebenso die Fragen nach einer konsensualen Annäherung sowie einer inhaltlichen Verbesserung der Urteile und letztlich einer Akzeptanz der Integration neuer Informationen in den Urteilsprozess untersucht. Startend mit der Dimension zur Evidenz.

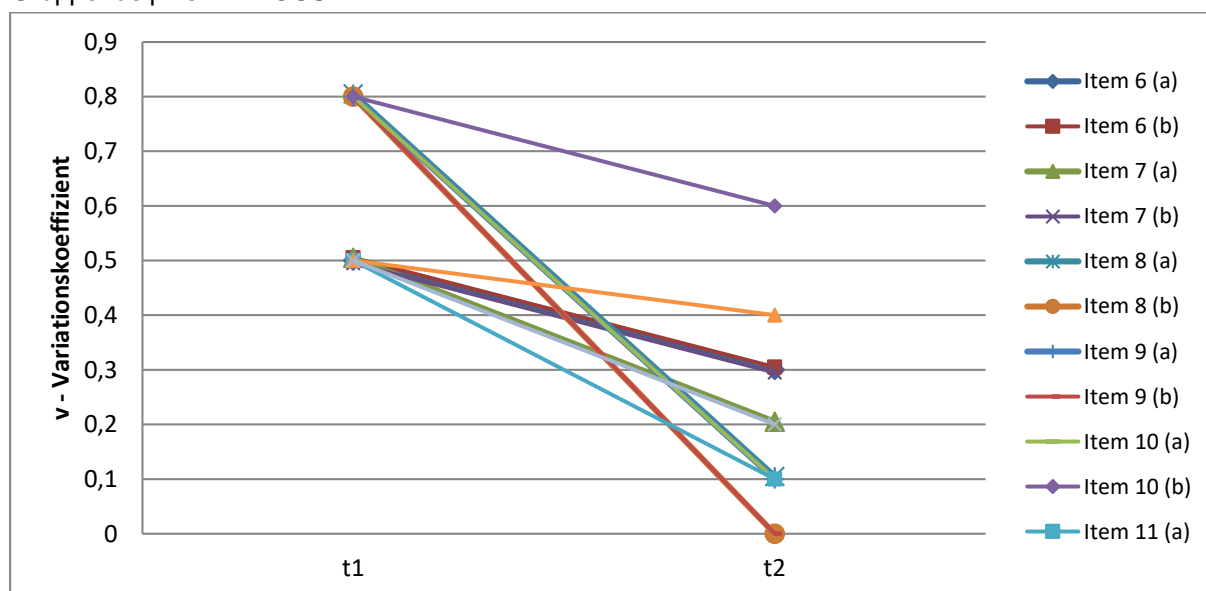
7.5.3.1 Dimension: Evidenz

Die Dimension der Evidenz bildet ebenso im Fall des Gruppendelphis in PROSUITE die zentrale Dimension im Testdesign. Die im Gruppendelphi erzielten Ergebnisse sollen hier dem Geltungsanspruch der Gültigkeit genügen. Sollten also die Qualitätsmerkmale eines handlungsleitenden Diskurses eingelöst werden, so sollten sich die in die Analyse einbezogenen Werte im Ideal in Richtung eines Konsenses bewegen oder einen eindeutigen Konsens über einen Dissens abbilden.

Im Gruppendelphi in PROSUITE wurden alle Indikatoren mittels 10er-skalierten Ratingskalen abgefragt. Im Fragebogen der ersten Kleingruppenphase wurde den Teilnehmenden zudem die Möglichkeit gegeben, mittels offener Abfrage eigene Vorschläge für zusätzliche Indikatoren einzubringen, die dann nach Konzeption des Delphis ebenso als 10er-skalierte Items in den Fragebogen der zweiten Runde integriert wurden. Folglich kann die Analyse hinsichtlich der Evidenz im Fall dieses Gruppendelphis exklusiv mit 10er-skalierten Items erfolgen. Da insgesamt 16 Items sowohl in der ersten als auch in der zweiten Runde des Gruppendelphis abgefragt wurden und für diesen Fall damit die statistischen Werte dieser 16 Items für einen Vergleich zur Verfügung stehen, findet im Falle von PROSUITE eine Betrachtung der Entwicklung der statistischen Werte strukturiert anhand der Abschnitte des Fragebogens statt. Dieses Gruppendelphi bietet überdies eine Besonderheit hinsichtlich der Abfrage der Items zur zweiten Runde des Gruppendelphis. So wurden einzelne Items der ersten Runde im Fragebogen der zweiten Runde geteilt, um sie mit jeweils anderem Bewertungshorizont zu beurteilen. Um die nachfolgende Analyse nachvollziehen zu können, ist hier ein Vorgriff auf das erste Plenum des Gruppendelphis notwendig. Bei diesem merkten die Teilnehmenden bereits beim ersten diskutierten *Item 6* an, dass sie während der ersten Kleingruppenphase ein grundsätzliches Problem bei der Bewertung einzelner Items hatten. So waren sich die Teilnehmenden in der ersten Kleingruppenphase nicht dahingehend einig, ob der per Item abgefragte Indikator grundsätzlich für die Bewertung der Nachhaltigkeit einer Technologie als wichtig zu bewerten sei, auch wenn dieser im Hinblick auf die Messung der Nachhaltigkeit des konkreten Falles CCS nicht als relevant eingeschätzt werden könne. Aus diesem Grund wurde sich im ersten Plenum auf den Vorschlag einer teilnehmenden Person geeinigt, diese Unterscheidung für die Items in den Fragebogen der zweiten Runde durchzuführen, bei denen der abgefragte Indikator jeweils anhand beider Bewertungshorizonte bearbeitet werden sollte. So kam es zur zweiten Runde zur Unterscheidung der generellen Wichtigkeit, was im Item durch den Zusatz *Wichtigkeit* gekennzeichnet und zusätzlich durch den Zusatz *a* an der Itemnummerierung vermerkt wurde. Bei Indikatoren, die auf die spezifische Bemessung der Nachhaltigkeit von CCS bewertet werden sollten, wurde der Zusatz *Impact* in die Itemformulierung aufgenommen und die Nummer des Items um den Zusatz *b* ergänzt. Nach diesem Vorgriff auf das erste Plenum kann nun die Betrachtung der Entwicklung der statistischen Werte erfolgen.

Der einleitende Fragetext der ersten Itembatterie lautete: *Im ersten Abschnitt finden sich Indikatoren die generell zur Bemessung sozialer Nachhaltigkeit konstruiert wurden. Bitte bewerten sie ob der Indikator sinnvoll zur Bewertung von CCS-Technologien ist, da diese Technologien einen direkten oder indirekten Einfluss auf den Indikator ausüben könnten.* Die Teilnehmenden wurden darum gebeten, einen Wert von 1 (-) anzugeben, wenn sie der Ansicht waren, dass dieser Indikator zur Bewertung von CCS-Technologien überhaupt nicht relevant ist und den Wert von 10 (+) anzugeben, wenn sie der Ansicht waren, dass dieser Indikator sehr relevant ist. Die Unterteilung der Items in *Wichtigkeit* und *Impact* (t2) hat bereits im ersten Abschnitt des Fragebogens Auswirkungen auf die Betrachtung der Entwicklung der statistischen Werte der einzelnen Items. So wurden in diesem Abschnitt zu den generellen Indikatoren bereits sechs der sieben Items, die über beide Runden des Gruppendelphis (t1/t2) abgefragt wurden, hinsichtlich der *Wichtigkeit* und des *Impacts* im zweiten Fragebogen gesondert abgefragt.

Abbildung 49: Werte des Variationskoeffizienten im ersten Abschnitt über die Runden (t1; t2) des Gruppendelphis in PROSUITE

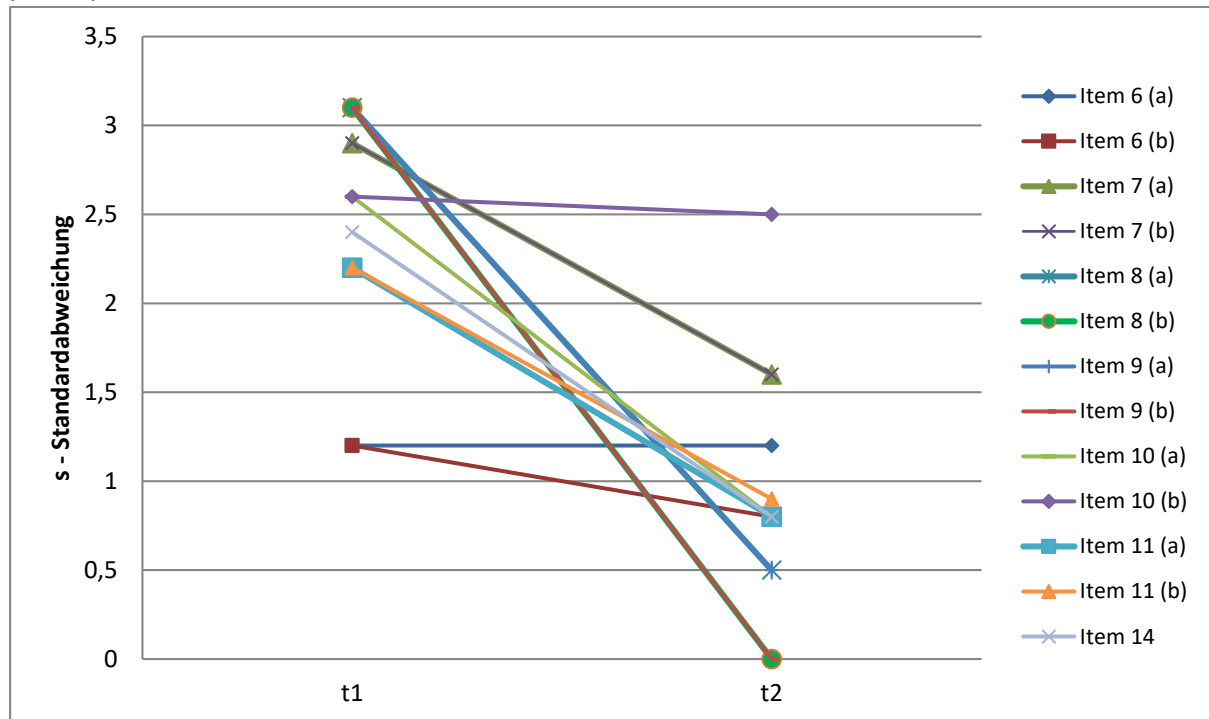


Anmerkungen: Item 7: $tn_1=4$, $tn_2=3$; Item 10b: $tn_1=3$, $tn_2=4$; alle übrigen Items: $tn_1=3$; $tn_2=3$.
Quelle: Eigene Darstellung.

Unter Betrachtung der Werte des Variationskoeffizienten in Abbildung 49 fällt bereits auf, dass der Wert von Runde 1 zu Runde 2 bei allen Items sinkt und sich dieser Maßzahl folgend hin zu einem konsensualeren Ergebnis entwickelt. Auch beim *Item 14*, in dem der Einfluss der *Technik auf die subjektiv empfundene Lebenszufriedenheit* abgefragt wurde, sinkt der Wert des Variationskoeffizienten von $v=0,5$ (t1) auf $v=0,2$ (t2). Dieses Item wurde im Gegensatz zu den anderen in dieser Abbildung dargestellten Items in der zweiten Runde nicht geteilt, sondern weiterhin exklusiv auf die Technik CCS bewertet. Ein weiteres Item auf das gesondert eingegangen werden soll, ist das *Item 10(b)*, das den Indikator *Prozentsatz der Personen deren Haushaltseinkommen unter 60 Prozent des nationalen Durchschnittseinkommens liegt* zur Bewertung stellte. In der zweiten Runde des Gruppendelphis wurde dieses Item auf seinen spezifischen *Impact* im Hinblick auf die Technik CCS als *Item 10b* abermals bewertet. Der Wert des Variationskoeffizienten sinkt zwar von $v=0,8$ (t1) auf $v=0,6$ (t2), liegt aber damit weiterhin über dem Richtwert von $v=0,5$, so dass hier anhand dieses Wertes auch in der zweiten Runde nicht von einem konsensualen Ergebnis ausgegangen werden kann. Im Kontrast dazu

hat sich das *Item 10a* (t2), in dem die generelle *Wichtigkeit* zur Bemessung der Nachhaltigkeit einer Technologie von einem Wert des *Items 10* von $v=0,8$ (t1) mit einem Wert des *Items 10a* von $v=0,1$ (t2) hin zu einer konsensualen Bewertung entwickelt.

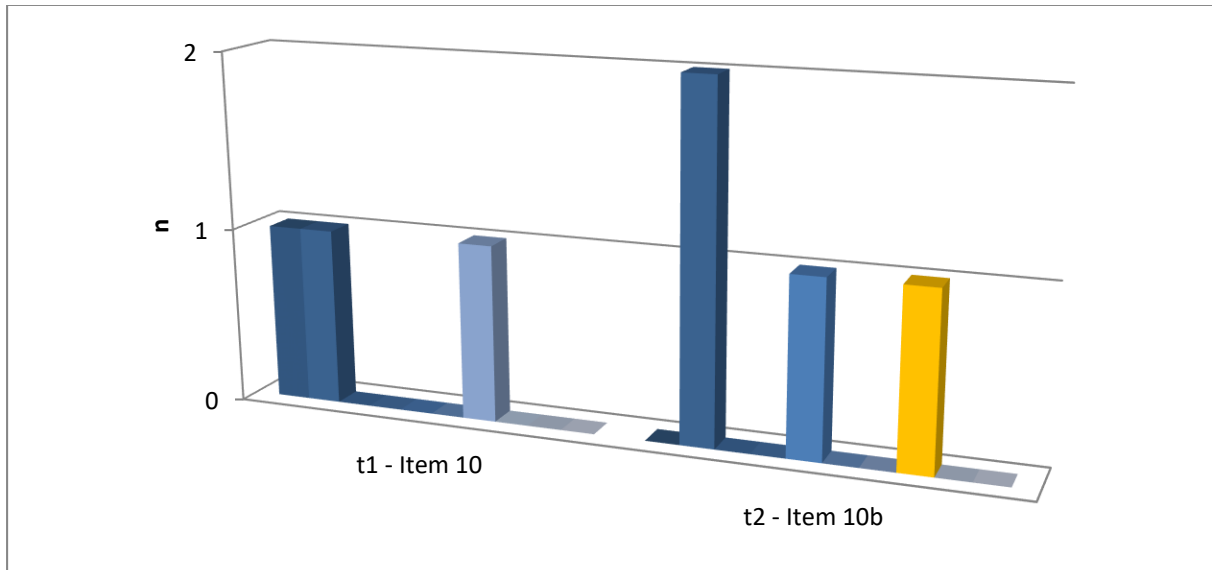
Abbildung 50: Werte der Standardabweichung im ersten Abschnitt über die Runden (t1; t2) des Gruppendelphis in PROSUITE



Anmerkungen: Item 7: $tn_1=4$, $tn_2=3$; Item 10b: $tn_1=3$, $tn_2=4$; alle übrigen Items: $tn_1=3$; $tn_2=3$. Sowohl (a) als auch (b) geben die zusätzliche Bezeichnung des Items zur zweiten Runde (t2) wieder. Quelle: Eigene Darstellung.

Diese Entwicklung spiegelt sich auch unter Beachtung der Werte der Standardabweichung, die in Abbildung 50 dargestellt sind, wider. Während der Wert der Standardabweichung beim *Item 10a* von $s=2,6$ (t1) auf $s=0,8$ (t2) sinkt, bleibt der Wert bei *Item 10b* mit $s=2,5$ (t2) auf einem hohen Niveau. Da beim *Item 10b* in der zweiten Runde zusätzlich ein Minderheitsvotum in der zweiten Kleingruppe abgegeben wurde, lohnt sich hier ein zusätzlicher Blick auf die Antwortverteilung, um prüfen zu können, ob es sich bei den Bewertungen zur zweiten Runde um einen Konsens über einen Dissens handelt. Dies wird unter Beachtung der Verteilung der einzelnen Gruppenantworten in Abbildung 51 allerdings nicht bestätigt. Während sich die Gruppen 1 und 2 während der ersten Runde des Gruppendelphis darin einig waren, dass das *Item 10* nicht wichtig sei (1/2), bewertete Gruppe 3 dieses Item als eher wichtig (7). In dieser Runde (t1) war damit eindeutiger ein Konsens über einen Dissens erkennbar, als es die Verteilung zur zweiten Runde vermuten lässt. Hier bewerteten die neu zusammengesetzten Gruppen 1 und 3 zwar das *Item 10b* als unwichtig (2/2), Gruppe 2 bewertete diesen Indikator allerdings im Mittelfeld der Skala (5), wenn auch eher als unwichtig. Zudem gab es in Gruppe 2 ein Minderheitsvotum, dass dieses *Item 10b* als wichtig beurteilte (8). So kann bei diesem Item zur zweiten Runde nicht von einem eindeutigen Ergebnis gesprochen werden, sondern die Bewertungen der einzelnen Gruppen (und Person durch das Minderheitsvotum) gehen zwischen der ersten und der zweiten Runde des Gruppendelphis auseinander.

Abbildung 51: Antwortverteilung bei Item 10 (b) - Haushaltseinkommen (Impact)



Anmerkungen: $tn_1=3$; $tn_2=4$ (Das Minderheitsvotum (t2) ist orange markiert). Quelle: Eigene Darstellung

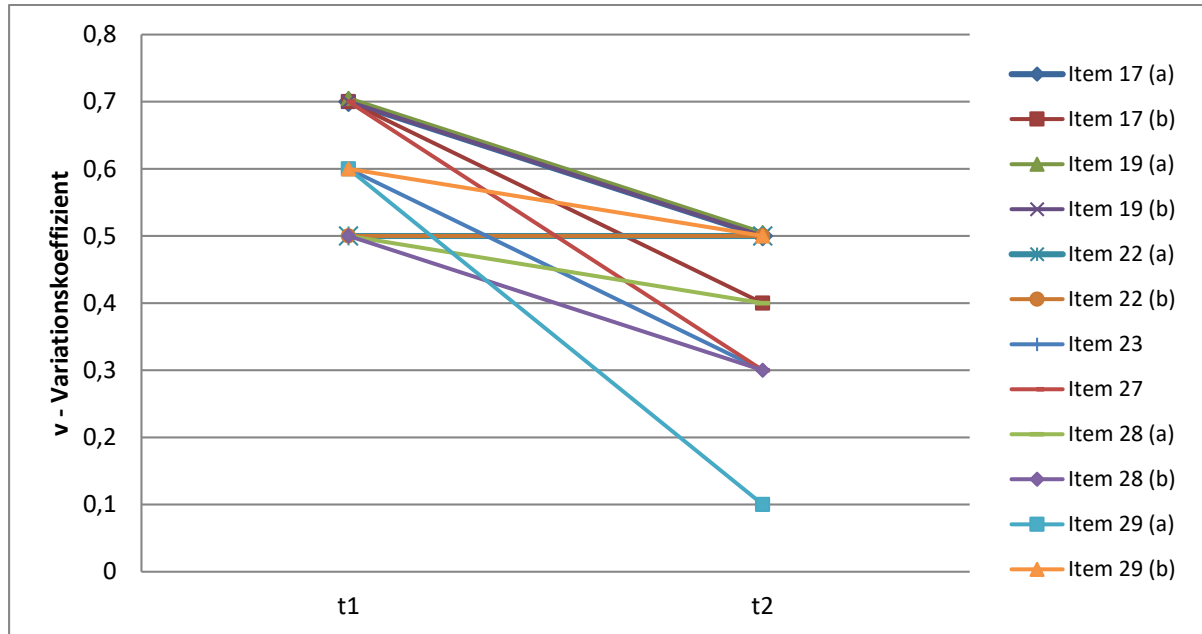
Beachtet man die in Abbildung 50 dargestellte Entwicklung der Werte der Standardabweichung, so fällt ein weiteres Item auf, bei dem sich eine genauere Betrachtung lohnt. So sank beim *Item 6(a)* der Wert des Variationskoeffizienten zwar von $v=0,5$ (t1) auf $v=0,3$ (t2), der Wert der Standardabweichung bleibt mit einem Wert von 1,2 (t1 und t2) allerdings über die Runden hinweg stabil. Dies lässt vermuten, dass die Teilnehmenden das *Item 6a* in der zweiten Runde hinsichtlich der *Wichtigkeit* des Indikators zur *Anzahl der Beschäftigten in wissensintensiven Bereichen* insgesamt als wichtiger beurteilt haben. Diese Vermutung wird unter Bezugnahme des arithmetischen Mittels bestätigt, das von einem Wert von $\bar{x}=2,7$ (t1) auf $\bar{x}=4,3$ (t2) ansteigt. Insgesamt bilden sowohl die Werte der Standardabweichung als auch die des Variationskoeffizienten für diesen ersten Abschnitt die Entwicklung hin zu einem konsensualleren Ergebnis über die Runden des Gruppendelphis ab.

Im zweiten Abschnitt des Fragebogens bildeten Items die Itembatterie, die *organisatorische Aspekte der Nachhaltigkeit* bemessen sollten. Die Teilnehmenden wurden gebeten zu bewerten, *ob die Indikatoren eine mögliche Auswirkung von CCS-Technologien auf die Organisationen messen und ob diese geeignet sind*. Dabei sollten sie einen Wert von 1 (-) angeben, wenn sie der Ansicht waren, *dass dieser Indikator zur Bewertung von CCS-Technologien überhaupt nicht relevant ist*, und eine 10 (+) anzugeben, wenn sie der Ansicht waren, *dass dieser Indikator sehr relevant ist*. Von den insgesamt 13 abgefragten Items in diesem zweiten Abschnitt, wurden die in Abbildung 52 dargestellten sieben Items auch zur zweiten Runde des Gruppendelphis abgefragt. Wie bereits weiter oben zum ersten Abschnitt des Fragebogens geschildert, wurden auch im zweiten Abschnitt Items zur zweiten Runde in ihre generelle *Wichtigkeit* für die Bemessung von Nachhaltigkeit (a) und den spezifischen *Impact* für die Bemessung der Technologie CCS (b) aufgeteilt.¹⁴² Zieht man die in Abbildung 52 dargestellten Werte des Variationskoeffizienten in die Betrachtung der Ergebnisse mit ein, so fällt auf, dass anhand der Werte bis auf *Item 22(b)* eine konsensuale Entwicklung angenommen werden kann, während dieses *Item 22(b)* auf einem hohen Wert von $v=0,5$ (t1 und t2) beständig bleibt. Nimmt man

¹⁴² Dies gilt für die Items 17, 19, 22, 28 und 29.

den erfahrungsbasierten Grenzwert von $v=0,5$ zur Bemessungsgrundlage, so liegen ebenso die *Items 19(a/b)* und *29(b)* trotz eines Absinkens des jeweiligen Wertes ebenso zur zweiten Runde bei einem Wert von $v=0,5$.

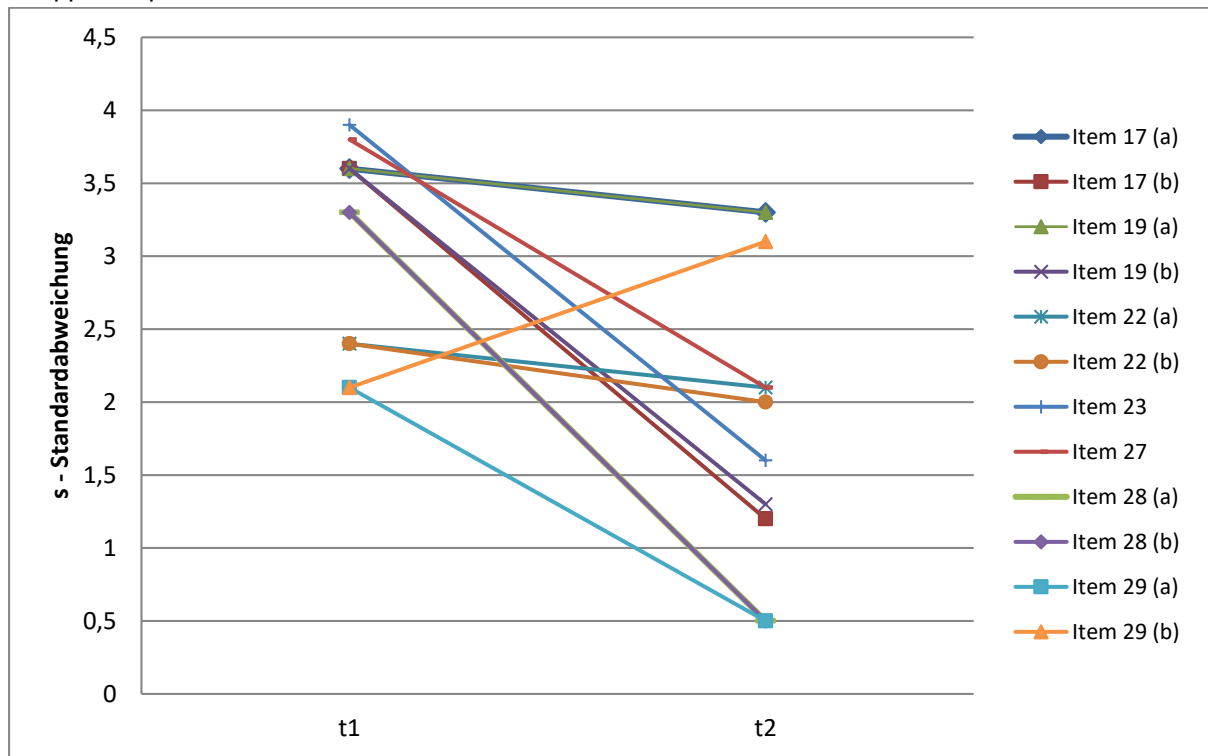
Abbildung 52: Werte des Variationskoeffizienten im zweiten Abschnitt über die Runden (t1; t2) des Gruppendelphis in PROSUITE



Anmerkungen: Item 27: $tn_1=4$; $tn_2=3$, alle anderen Items: $tn_1=3$; $tn_2=3$. Quelle: Eigene Darstellung

Insbesondere bei diesen vier Items lohnt der Einbezug der Werte der Standardabweichung, wie sie in Abbildung 53 dargestellt sind. Als erstes Item soll *Item 19* näher betrachtet werden, in dem das *Verhältnis des Grundgehalts von Männern und Frauen nach Beschäftigungsverhältnis* als Indikator zur Bewertung stand. Bei *Item 19* fällt dabei auf, dass die Entwicklung vom ursprünglich hohen Wert der Standardabweichung von $s=3,6$ (t1), die in der zweiten Runde geteilten *Items 19a* und *19b* eine unterschiedliche Entwicklung nehmen. Während der Wert von *Item 19(a)* von $s=3,6$ (t1) auf $s=3,3$ (t2) nicht erheblich sinkt, findet bei *Item 19(b)* eine deutliche Verminderung des Wertes von $s=3,6$ (t1) auf $s=1,3$ (t2) statt. Unter Einbezug des arithmetischen Mittels wird schnell deutlich worin diese unterschiedliche Entwicklung des Wertes der Standardabweichung bei identischem Wert des Variationskoeffizienten begründet liegt. So zeigt das *Item 19(a)* zur zweiten Runde im Mittel mit einem Wert von $\bar{x}=7,3$ eine Einschätzung des Indikators als relevant an. Während sowohl Gruppe 2 als auch Gruppe 1 (7) mit einem Minderheitsvotum (10) diesen Indikator als (sehr) relevant einschätzen, lehnt Gruppe 3 diesen Indikator als irrelevant ab (2). Somit findet sich hier zur zweiten Runde keine konsensuale Bewertung der generellen Wichtigkeit des *Items 19a* zwischen den Teilnehmenden. Anders beim *Item 19b*, das zur zweiten Runde einen Wert von $\bar{x}=2,8$ aufweist und bei dem alle drei Gruppen das Item als nicht relevant bewertet haben (2) sowie ein Minderheitsvotum in Gruppe 1 diesen Indikator zwar höher, aber auch als eher nicht relevant eingestuft hat (5). Somit wurde dieses *Item 19(b)* zur zweiten Runde konsensual bewertet, was ebenso durch das Absinken des Wertes der Standardabweichung bestätigt wird und unter Einbezug des niedrigen arithmetischen Mittels von $\bar{x}=2,8$ (t2) zu dem Wert des Quotienten von $v=0,5$ (t2) führt.

Abbildung 53: Werte der Standardabweichung im zweiten Abschnitt über die Runden (t1; t2) des Gruppendelphis in PROSUITE

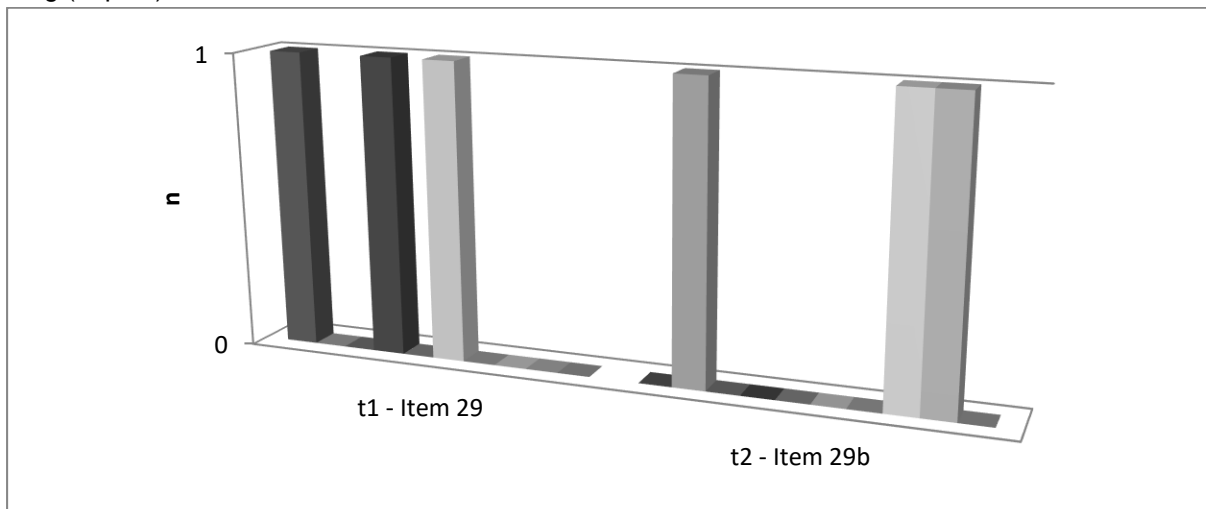


Anmerkungen: Item 27: $tn_1=4$; $tn_2=3$, alle anderen Items: $tn_1=3$; $tn_2=3$; Quelle: Eigene Darstellung

Beim Indikator zur *Anzahl von Verbraucherbeschwerden (Impact)* in *Item 22(b)* ist, bei gleichbleibendem Wert des Variationskoeffizienten von $v=0,5$ (t_1 und t_2), der Wert der Standardabweichung von $s=2,4$ (t_1) auf $s=2,0$ (t_2) nicht erheblich gesunken, was darauf hindeutet, dass auch zur zweiten Runde des Gruppendelphis kein konsensuales Ergebnis erzielt werden konnte. Das arithmetische Mittel zeigt mit einem Wert von $\bar{x}=3,8$ (t_2) eine durchschnittliche Bewertung des Indikators über die Gruppen hinweg als eher nicht relevant an. Dies bestätigt sich auch mit Blick auf die einzelnen Bewertungen von Gruppe 1 als eher nicht relevant (4) und Gruppe 2 und 3 als nicht relevant (2). Die hohen Werte der Standardabweichung und des Variationskoeffizienten, ergeben sich aber durch ein Minderheitsvotum in Gruppe 1, in der eine Person das *Item 22b* als relevant eingestuft hat (7). Aufgrund dieses Votums kann damit für dieses Item nicht von einem Konsens ausgegangen werden.

Als letztes Item wird das *Item 29(b)* unter diesem zweiten Abschnitt detaillierter betrachtet, der den Indikator *Anteil der Lieferanten die vom Unternehmen auf ihre soziale Verantwortung in den letzten Jahren überprüft wurden* enthielt. Die Entwicklung der statistischen Werte dieses Items ist besonders interessant, da hier zwar der Wert des Variationskoeffizienten von $v=0,6$ (t_1) auf $v=0,5$ (t_2) sinkt, die Standardabweichung hingegen von $s=2,1$ (t_1) auf $s=3,1$ (t_2) ansteigt. Aufgrund dieser Entwicklung lässt sich vermuten, dass dieses Item zur zweiten Runde im Mittel als relevanter eingestuft wurde, als es in der ersten Runde des Gruppendelphis der Fall war. Zieht man die Werte des arithmetischen Mittels hinzu, wird diese Vermutung durch $\bar{x}=3,7$ (t_1) zu $\bar{x}=6,3$ (t_2) bestätigt. Interessant ist bei diesem Item ebenso die Verteilung der einzelnen Urteile der Kleingruppen über die beiden Runden des Gruppendelphis (t_1 und t_2).

Abbildung 54: Antwortverteilung bei Item 29 (b) - Überprüfung der Lieferanten auf soziale Verantwortung (Impact)



Anmerkungen: $tn_1=3$; $tn_2=3$

Quelle: Eigene Darstellung

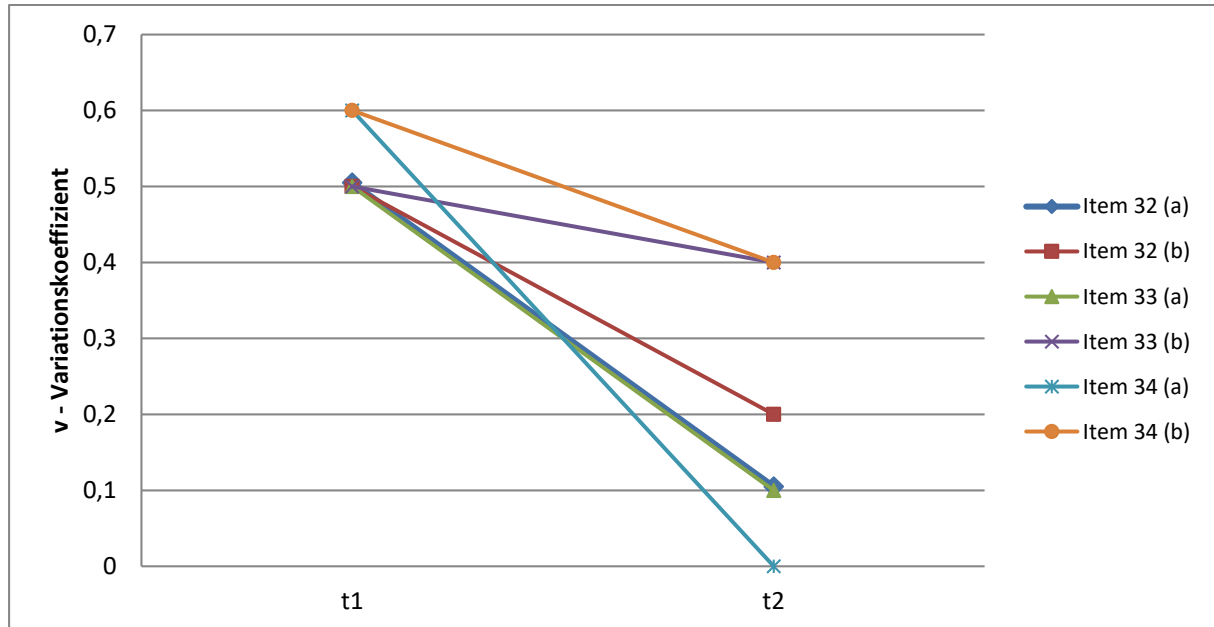
Wie in Abbildung 54 ersichtlich wird, beurteilten die Kleingruppen das *Item 29* in der ersten Runde noch überwiegend negativ, denn Gruppe 2 bewertete den abgefragten Indikator als nicht relevant (1), während Gruppe 3 diesen als eher nicht relevant (4) bewertete. Einzig Gruppe 1 bewerte diesen Indikator in der ersten Runde als eher relevant (6). In der zweiten Runde wurde der Indikator im Hinblick auf seinen *Impact (Item 29b)* sowohl von der neu zusammengesetzten Gruppe 1 als auch der Gruppe 2 als relevant eingestuft (8/9), während Gruppe 3 diesen als nicht relevant einstuft (2). Unter Einbezug von Abbildung 54 lässt sich somit anhand der Verteilungen der Antworten eine Herausbildung eines Konsenses über einen Dissens zur zweiten Runde des Gruppendelphis konstatieren.

Die letzte Itembatterie des Fragebogens der ersten Runde des Gruppendelphis in PROSUITE fand sich im dritten Abschnitt, in dem Indikatoren zur Bewertung durch die Teilnehmenden standen, die *produktspezifische Aspekte der Technik auf soziale Nachhaltigkeit erfassen sollen*. Auch in diesem Abschnitt wurden die Teilnehmenden darum gebeten, einen Wert von 1 (-) anzugeben, wenn sie der Ansicht waren, dass *dieser Indikator zur Bewertung von CCS-Technologien überhaupt nicht relevant ist* und den Wert von 10 (+) anzugeben, wenn sie der Ansicht waren, dass *dieser Indikator sehr relevant ist*. Von den insgesamt fünf abgefragten Items (t1) in diesem Abschnitt wurden drei auch in der zweiten Runde des Gruppendelphis abgefragt und stehen für eine Betrachtung der Entwicklung der statistischen Kenngrößen in diesem Kapitel zur Verfügung. Alle drei Items wurden zur zweiten Runde des Gruppendelphis, wie in den vorhergehenden Abschnitten bereits mit einer Mehrzahl der Items geschehen, in *Wichtigkeit (a)* und *Impact (b)* aufgeteilt.

Wie in Abbildung 55 bereits ersichtlich wird, zeigen alle über beide Runden abgefragten Items, die im Ideal angenommene Entwicklung hinsichtlich des Wertes des Variationskoeffizienten. Während in der ersten Runde alle Items einen Wert von $v \geq 0,5$ hatten, liegen alle Items zur zweiten Runde des Gruppendelphis mit $v \leq 0,4$ unter dem Richtwert von $v=0,5$. Besonders von *Item 34(a)* wird die idealiter angenommene Entwicklung hin zu einem konsensualen Ergebnis durch ein Absinken des Wertes von $v=0,6$ (t1) auf $v=0$ (t2) anschaulich bestätigt. Beachtet man die zwar unter dem Richtwert von $v=0,5$ liegenden, aber mit $v=0,4$ (t2) relativ hoch liegenden

Items 33b und 34b, so fällt bereits unter Einbezug des arithmetischen Mittelwertes auf, dass auch diese konsensual und von den Teilnehmenden als nicht relevant eingestuft wurden: Bei Item 33b liegt das arithmetische Mittel bei $\bar{x}=3,3$ (t2) wie auch bei Item 34b mit $\bar{x}=3,3$ (t2).

Abbildung 55: Werte des Variationskoeffizienten im dritten Abschnitt über die Runden (t1; t2) des Gruppendelphis in PROSUITE

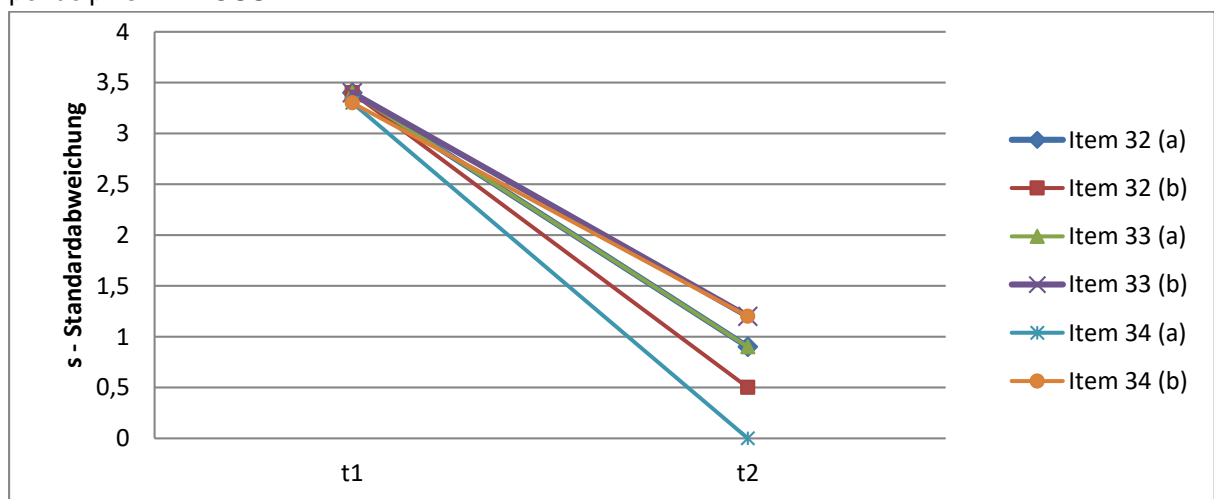


Anmerkungen: tn1=3; tn2=3

Quelle: Eigene Darstellung

Die für diesen dritten Abschnitt des Gruppendelphis in PROSUITE nahezu idealtypische Entwicklung wird überdies durch die in Abbildung 56 dargestellte Entwicklung der Werte der Standardabweichung veranschaulicht. So sinkt die Standardabweichung unter Berücksichtigung aller dargestellten Items von $s \geq 3,3$ (t1) auf $s \leq 1,2$.

Abbildung 56: Werte der Standardabweichung im dritten Abschnitt über die Runden (t1; t2) des Gruppendelphis in PROSUITE



Anmerkungen: tn1=3; tn2=3;

Quelle: Eigene Darstellung

Abschließend kann für die Dimension der Evidenz der Schluss gezogen werden, dass eine Vielzahl der hier untersuchten Items mittels der jeweiligen Entwicklung der statistischen Werte

die idealtypische Annahme bestätigen, dass auch dieses Fallbeispiel konsensuale Bewertungen befördert oder zumindest einen eindeutigen Konsens über einen Dissens ausbildet, wie es beispielsweise bei *Item 29(b)* der Fall war. Durch die Aufteilung von 14 der insgesamt 17 Items in generelle *Wichtigkeit (a)* des Items und den *Impact (b)* in Bezug auf die Technologie CCS sollte allerdings der Vergleich der Werte zwischen beiden Runden unter Berücksichtigung der jeweiligen Bewertungshorizonte vertiefend betrachtet werden, wie in der folgenden Dimension bezüglich der Verständlichkeit vorgesehen. Sollten die Inhalte der einzelnen Items durch die Spezifikation bzw. Aufteilung der Items vor einem gänzlich anderen inhaltlichen Kontext erfolgt sein, so muss dies im Rückschluss auch auf die Ergebnisse dieser Dimension kritisch betrachtet werden.

7.5.3.2 Dimension: Verständlichkeit

Im vorhergehenden Abschnitt ist bereits erläutert worden, dass einzelne Teilnehmende in der ersten Kleingruppenphase bei einigen Items zwei Bedeutungshorizonte für die jeweilige Bewertung der Indikatoren als notwendig ansahen. Auch Gruppe 2 äußerte, dass die Teilnehmenden innerhalb der Kleingruppe einzelne Indikatoren als relevant ansahen, wenn auch nicht im konkreten Fall auf CCS bezogen. Da diese Diskussion in der dritten Kleingruppen nicht stattfand und diese die Indikatoren spezifisch auf die Technik CCS bewertete, ist die Leitfrage, ob die Teilnehmenden über den Prozess des Gruppdelphis mit dem Geäußerten letztlich einen gleichen Bedeutungsinhalt verbinden, besonders für diese Unterscheidung spannend. Konnte die Methode des Gruppdelphis dem Geltungsanspruch der Zweckmäßigkeit für den Diskurs in diesem Fall entsprechen?

Im Fragebogen des Gruppdelphis in PROSUITE wurden mittels eines Einleitungstextes zu Beginn die wissenschaftlichen Vorarbeiten, die für die Auswahl der abgefragten Indikatoren durchgeführt wurden, wie auch die Ziele des Gruppdelphis erläutert: *Ziel ist es, sie in Bezug auf eine gesellschaftliche Etablierung von CCS-Technologien einzuschätzen und eine damit verbundene Analyse sozialer Nachhaltigkeit zu ermöglichen.*¹⁴³ Ebenso wurde im ersten wie auch im zweiten Abschnitt darauf hingewiesen, dass die Bewertung der Teilnehmenden danach erfolgen soll, *ob der Indikator sinnvoll zur Bewertung von CCS-Technologien ist, da diese Technologien einen direkten oder indirekten Einfluss auf den Indikator ausüben könnten.* Letztlich wurde auf den direkten Bezug der Bewertung zur Technologie ebenfalls in der Erläuterung der Skalenwerte hergestellt: *Geben Sie bitte 1 (-) an, wenn Sie der Ansicht sind, dass dieser Indikator zur Bewertung von CCS-Technologien überhaupt nicht relevant ist, geben Sie bitte 10 (+) an, wenn Sie der Ansicht sind, dass dieser Indikator sehr relevant ist.* So ist auszuschießen, dass die Trennung beider Bewertungshorizonte aufgrund eines Missverständnisses des Fragetextes oder durch bloßes *Überlesen* des Bezuges zur *CCS-Technologie* in der ersten Kleingruppenphase entstanden ist. Die Teilnehmenden, die eine Trennung der Indikatoren zwischen einer generellen *Wichtigkeit* zur allgemeinen Bemessung von Nachhaltigkeit und dem spezifischen *Impact* des Indikators zur Bemessung der CCS-Technologie befürworteten, hatten hierin ein inhaltliches Interesse, dass sie im ersten Plenum äußerten. Dieser Vorschlag wurde von den Teilnehmenden konsensual für einzelne Indikatoren als sinnvoll erachtet, da so in der zweiten Runde des Gruppdelphis auch Indikatoren als relevant bewertet werden

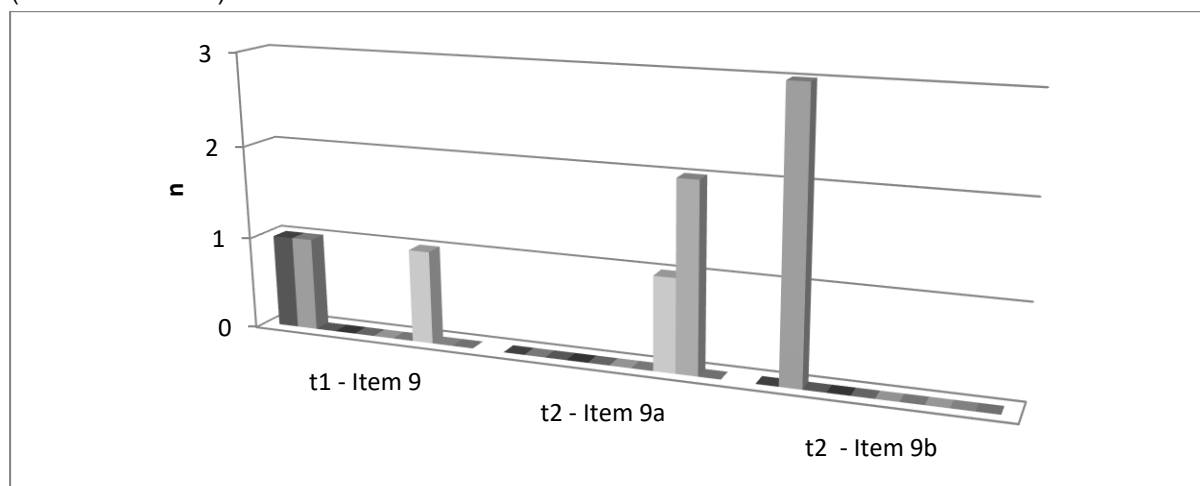
¹⁴³ Der gesamte Einleitungstext findet sich im Fragebogen des Gruppdelphis im Anhang 11.15.

konnten, die für die Bemessung der Nachhaltigkeit *wichtig*, allerdings im Falle der konkreten Beispieltechnologie nicht relevant seien. Des Weiteren wurde argumentiert, dass es durch die Trennung überdies möglich sei, Indikatoren zu beurteilen, die generell weniger relevant sein könnten, aber besonders im Falle von CCS einen großen *Impact* für die Bemessung der Nachhaltigkeit dieser Technologie hätten. Da dieser Konsens bereits sehr früh während des ersten Plenums erzielt werden konnte, bestand für die Teilnehmenden die Möglichkeit bei der Diskussion jedes Items den Vorschlag anzubringen, diesen Indikator in *Wichtigkeit (a)* und *Impact (b)* zu unterteilen und im Fragebogen der zweiten Kleingruppenphase auf diese Weise zu beurteilen. Von den 17 Items, die sowohl in der ersten wie auch in der zweiten Runde zum Urteil standen wurden 14 Items aufgeteilt, so dass diese in doppelter Summe im Fragebogen der zweiten Runde bewertet werden konnten.

Wie bereits im vorhergehenden Kapitel dargestellt, wurde die Mehrzahl der Indikatoren zur zweiten Runde des Gruppendelphis einstimmiger beurteilt, als es in der ersten Runde vor der Aufteilung der Fall war. Da die Entwicklung der einzelnen Werte von den jeweiligen Indikatoren bereits vorgestellt wurde, kann hier zusammenfassend der arithmetische Mittelwert der Standardabweichung in diese Betrachtung mit einbezogen werden. So ergibt sich für die 14 abgefragten Items in der ersten Runde ein Mittel von $\bar{x}(s) = 2,88$ (t1).¹⁴⁴ Dieser Wert sinkt sowohl bei den in *Wichtigkeit (a)* aufgeteilten Items auf $\bar{x}(s) = 1,2$ (t2)¹⁴⁵ wie auch beim *Impact (b)* auf $\bar{x}(s) = 1,2$ (t2)¹⁴⁶. Hinsichtlich Fragen der Konsensbildung besteht somit im Durchschnitt kein Unterschied zwischen den beiden Itemtypen *a* und *b* zur zweiten Runde, außer dass beide Werte hinsichtlich einer Konsensbildung im Mittel positiv verlaufen.

Interessant ist im Hinblick auf die Unterteilung nach *Wichtigkeit (a)* und *Impact (b)* allerdings auch der Einbezug der Urteile, die die Gruppen in der zweiten Runde den jeweiligen Items gegeben haben, an dieser Stelle exemplarisch das *Item 9*, das den Indikator der *Lebenserwartung zum Zeitpunkt der Geburt (nach Geschlecht)* enthielt.

Abbildung 57: Antwortverteilung bei Item 9 (a/b) - Die Lebenserwartung zum Zeitpunkt der Geburt (nach Geschlecht)



Anmerkungen: tn1=3; tn2=3

Quelle: Eigene Darstellung

¹⁴⁴ Items: 6,7,8,9,10,11,17,19,22,28,29,32,33,34.

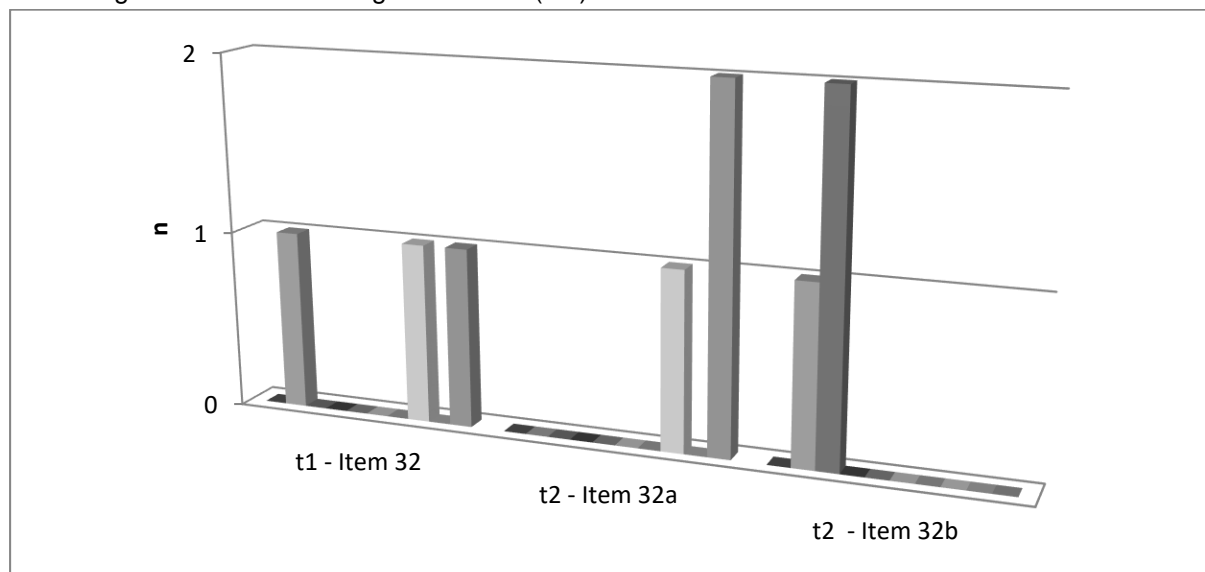
¹⁴⁵ Items: 6a,7a,8a,9a,10a,11a,17a,19a,22a,28a,29a,32a,33a,34a.

¹⁴⁶ Items: 6b,7b,8b,9b,10b,11b,17b,19b,22b,28b,29b,32b,33b,34b.

Während in der ersten Runde eine der Kleingruppen (Gruppe 1) den Indikator als relevant einstufte, beurteilen in der zweiten Runde den nach Wichtigkeit unterteilten Indikator in *Item 9a* geschlossen als relevant, während sie diesen Indikator im Hinblick auf den *Impact* zur Bemessung der Nachhaltigkeit von CCS im *Item 9b* geschlossen als nicht relevant einschätzen.

Ein ebenso gutes Beispiel für die Entwicklung, die durch die Aufteilung zwischen beiden Runden des Gruppendelphis entstanden ist, bietet das *Item 32* in dem der Indikator zur *Anzahl tödlicher Arbeitsunfälle* auf seine Relevanz zur Bemessung von Nachhaltigkeit abgefragt wurde. Dieses *Item 32* wurde in der ersten Runde mit einem arithmetischen Mittel von $\bar{x}=6,7$, mit einer Standardabweichung von $s=3,4$ und einem Variationskoeffizienten von $v=0,5$ bewertet. In der zweiten Runde beträgt das arithmetische Mittel für die generelle *Wichtigkeit* des Indikators, abgefragt mittels *Item 32a*, einen Wert von $\bar{x}=9,3$, mit einer Standardabweichung von $s=0,9$ und einem Variationskoeffizienten von $v=0,1$. Für den *Impact* des Indikators (*Item 32b*) hat das arithmetische Mittel einen Wert von $\bar{x}=2,7$, mit einer Standardabweichung von $s=0,5$ und einem Variationskoeffizienten von $v=0,2$. So beurteilen die Teilnehmenden in der zweiten Runde den nach *Wichtigkeit* unterteilten Indikator in *Item 32a* geschlossen als relevant. Im Gegensatz dazu schätzen sie diesen Indikator in *Item 32b* im Hinblick auf den *Impact* zur Bemessung der Nachhaltigkeit von CCS geschlossen als nicht relevant ein. So ergibt sich für die zweite Runde des Gruppendelphis in Abbildung 58 ein ähnliches Muster wie bereits in Abbildung 57.

Abbildung 58: Antwortverteilung bei Item 32 (a/b) - Anzahl tödlicher Arbeitsunfälle



Anmerkungen: $n_1=3$; $n_2=3$

Quelle: Eigene Darstellung

Diese beiden ausgewählten Beispiele bilden dabei keine Ausnahme. So wurde bei den unterteilten 14 Items in 13 Fällen die Wichtigkeit des Indikators zur Bemessung von Nachhaltigkeit von den Teilnehmenden in der zweiten Runde relevanter beurteilt, als der jeweilig entsprechende Impact zur Bemessung der Nachhaltigkeit von CCS. Dies zum Teil ähnlich signifikant wie im Falle des *Items 9* oder des *Items 32*. So wurde im Hinblick auf den Indikator betreffend der *Kindersterblichkeitsrate* das *Item 8a (Wichtigkeit)* mit einem arithmetischen Mittel von $\bar{x}=8,7$ (t_2) bei $s=0,5$ (t_2) als relevant beurteilt, während das *Item 8b (Impact)* mit einem Wert von $\bar{x}=2,0$ (t_2) bei $s=0$ (t_2) geschlossen als nicht relevant angesehen wurde. Auch die *Wichtigkeit* des *Gini-Index* in *Item 11a* wurde im Hinblick auf die allgemeine *Wichtigkeit* mit einem

arithmetischen Mittel von $\bar{x}=8,0$ (t2) bei $s=0,8$ (t2) als relevant angesehen, während auch hier der *Impact* bemessen auf CCS in *Item 11b* mit $\bar{x}=2,7$ (t2) bei $s=0,9$ (t2) von den Teilnehmenden als nicht relevant beurteilt wurde. So auch die *Anzahl von nicht-tödlicher Arbeitsunfälle* in *Item 33a* mit $\bar{x}=9,3$ (t2) bei $s=0,9$ (t2) als relevant und in *Item 33b* mit $\bar{x}=3,3$ (t2) bei $s=1,2$ (t2) als nicht relevant beurteilt wurde, wie ferner ebenso der Indikator zur *Anzahl der Berufskrankheiten* bei *Item 34a* mit $\bar{x}=8,0$ (t2) bei $s=0$ (t2) als relevant und in *Item 34b* mit $\bar{x}=3,3$ (t2) bei $s=1,2$ (t2) als nicht relevant eingeschätzt wurde.

Die Auflistung aller Werte des arithmetischen Mittels zu beiden Runden des Gruppendelphis sowie in der Unterteilung der Items in *Wichtigkeit* und *Impact* in Tabelle 18 geben diese Tendenz in den besagten 13 Fällen wieder.

Tabelle 18: Arithmetisches Mittel der Items unterteilt in Wichtigkeit und Impact im Gruppendelphi in PROSUITE

Item #	Frageinhalt	\bar{x} (t1)	\bar{x} (t2)
6 (a)	Beschäftigten in wissensintensiven Bereichen (Wichtigkeit)	2,7	4,3
6 (b)	Beschäftigten in wissensintensiven Bereichen (Impact)		3,0
7 (a)	Ausreichend informiert fühlen (Wichtigkeit)	5,8	8,0
7 (b)	Ausreichend informiert fühlen (Impact)		6,0
8 (a)	Kindersterblichkeitsrate (Wichtigkeit)	3,7	8,7
8 (b)	Kindersterblichkeitsrate (Impact)		2,0
9 (a)	Lebenserwartung (Wichtigkeit)	3,7	8,7
9 (b)	Lebenserwartung (Impact)		2,0
10 (a)	Haushaltseinkommen (Wichtigkeit)	3,3	8,0
10 (b)	Haushaltseinkommen (Impact)		4,3
11 (a)	Gini-Index (Wichtigkeit)	4,0	8,0
11 (b)	Gini-Index (Impact)		2,7
17 (a)	Richtlinien zur Chancengleichheit (Wichtigkeit)	5,0	7,3
17 (b)	Richtlinien zur Chancengleichheit (Impact)		3,0
19 (a)	Verhältnis des Grundgehalts (nach Geschlecht) (Wichtigkeit)	5,0	7,3
19 (b)	Verhältnis des Grundgehalts (nach Geschlecht) (Impact)		2,8
22 (a)	Anzahl von Verbraucherbeschwerden (Wichtigkeit)	5,0	4,3
22 (b)	Anzahl von Verbraucherbeschwerden (Impact)		3,8
28 (a)	Verhaltensregeln zum Schutz der Menschenrechte (Wichtigkeit)	6,0	1,3
28 (b)	Verhaltensregeln zum Schutz der Menschenrechte (Impact)		1,7
29 (a)	Überprüfung der Lieferanten (Wichtigkeit)	3,7	8,3
29 (b)	Überprüfung der Lieferanten (Impact)		6,3
32 (a)	Anzahl tödlicher Arbeitsunfälle (Wichtigkeit)	6,7	9,3
32 (b)	Anzahl tödlicher Arbeitsunfälle (Impact)		2,7
33 (a)	Anzahl nicht-tödlicher Arbeitsunfälle (Wichtigkeit)	6,7	9,3
33 (b)	Anzahl nicht-tödlicher Arbeitsunfälle (Impact)		3,3
34 (a)	Anzahl der Berufskrankheiten (Wichtigkeit)	5,7	8,0
34 (b)	Anzahl der Berufskrankheiten (Impact)		3,3

Die einzige Ausnahme bildet das *Item 28a* mit einem Wert von $\bar{x}=1,3$ (t2) bei $s=0,5$ (t2), ebenso wie das entsprechende *Item 28b* mit $\bar{x}=1,7$ (t2) bei $s=0,5$ (t2), die als nicht relevant eingeschätzt wurden. So wurde bei diesem Indikator zu *Verhaltensregeln zum Schutz der Menschenrechte* weder die *Wichtigkeit* für die generelle Bemessung von Nachhaltigkeit noch der

Impact dieses Indikators auf die Technologie CCS als relevant eingeschätzt. Dieser Indikator wurde im zweiten Plenum des Gruppendelphis nicht diskutiert, so dass keine inhaltliche Begründung der Teilnehmenden für die Ablehnung beider Indikatoren in die Analyse mit einbezogen werden kann. Allerdings ist die Ablehnung beider Items konsensual, was nach der ersten Kleingruppenphase bei *Item 28* nicht der Fall war: Das Item wurde zu diesem Zeitpunkt zwar im Mittel mit $\bar{x}=6,0$ (t1) als eher relevant eingestuft, allerdings lag mit einer Standardabweichung von $s=3,3$ (t1) und einem Variationskoeffizienten von $v=0,5$ (t1) hier kein konsensuales Ergebnis vor.

Die Aufteilung der Items zur zweiten Runde des Gruppendelphis hat ein gemeinsames Verständnis geschaffen, in welchem Bewertungshorizont die Indikatoren zu beurteilen sind, der in der ersten Kleingruppenphase noch nicht gänzlich gegeben war. Aufgrund der Trennung von *Wichtigkeit* und *Impact* konnte diese Diskrepanz aufgelöst werden.

Neben der Teilung der 14 Items kam es über dessen in diesem Fallbeispiel auch zu Umformulierungen von einzelnen Items. Zwei dieser Items sind bereits in Tabelle 18 enthalten, da sie sowohl umformuliert als auch in *Wichtigkeit* und *Impact* aufgeteilt wurden:

- **Item 7:** *Anteil der Personen die bereits Informationsmaterial über die Risiken der Technologie erhalten haben.* Dieses Item wies mit einer Standardabweichung von $s=2,9$ (t1) und einem Variationskoeffizienten von $v=0,5$ (t1)¹⁴⁷ einen Dissens zwischen den jeweiligen Urteilen der Kleingruppen auf und wurde im ersten Plenum diskutiert. Während dieser Plenarsitzung argumentierte die Person, die ein Minderheitsvotum in Gruppe 1 abgegeben hatte und den Indikator als nicht relevant bewertete (2), dass der Zugang zu Informationsmaterial in Zeiten des Internets bereits sichergestellt sei. Die zweite Kleingruppe, die den Indikator als sehr relevant bewertet hatte (10) entgegnete, dass Information über Technik immer eine hohe Relevanz hätte, da sie die Voraussetzung für eine Akzeptanz dieser Technik bilde und ebenso eine Basis für mögliche Beteiligungsprozesse schaffen könne. Dem widersprach die Person, die das Minderheitsvotum abgegeben hatte nicht, da sie auch eine Information als wichtig einschätze, nur die bloße Bereitstellung von Informationsmaterial eine Informiertheit der Betroffenen nicht garantiere sowie im Internet bereits zur Verfügung stände. Die Teilnehmenden konnten sich darauf einigen sich bei dem Indikator auf die Informiertheit der Betroffenen zu fokussieren. Da eine solche Informiertheit mit einem Indikator nicht objektiv erfasst werden könne, wurde das *Item 7 (a/b)* zur zweiten Runde wie folgt umformuliert: *Anteil der Personen, der sich subjektiv ausreichend informiert fühlt.* Das in die *Wichtigkeit* aufgeteilte neu formulierte *Item 7a* wurde zur zweiten Runde mit $s=1,6$ und $v=0,2$ mit einem arithmetischen Mittel von $\bar{x}=8,0$ konsensual als relevant beurteilt, wenn auch von Gruppe 3 als eher relevant (6) im Gegensatz zu Gruppe 1, die diesen Indikator als sehr relevant beurteilte (10). Anders bei dem neuen *Item 7b* zum *Impact*, das zwar nicht signifikant andere statistische Werte mit $s=1,6$ (t2) und $v=0,3$ (t2) hinsichtlich der Einigkeit aufwies, allerdings lagen die einzelnen Bewertungen mit einem arithmetischen Mittel von $\bar{x}=6,0$ (t2) sowohl auf der Seite der 10er-Skala, die den Indikator als relevant einstuft, wie auch auf der Seite der Skala, die dies nicht tat. So stuft Gruppe 1 dieses *Item 7b* als nicht relevant ein (4), während Gruppe 2 und Gruppe 3 es als relevant beurteilten (8/6). Aufgrund dieses Dissenses fand zu diesem *Item 7b* im zweiten Plenum nochmals ein Austausch statt.

¹⁴⁷ Da es in Gruppe 1 ein Minderheitsvotum gab, ist bei diesem Item zum Zeitpunkt t1 $n=4$.

Gruppe 1 begründete die Ablehnung damit, dass aufgrund der hohen technischen Komplexität von CCS Informationen nicht leicht vermittelbar seien und somit auch die subjektive Informiertheit nicht ein gänzlich geeigneter Indikator sei. Da Informationen in der heutigen Zeit als Bestätigung der bereits bestehenden Meinung genutzt würden, wäre der Indikator bei einer komplexen Technik wie CCS eher nicht relevant, um eine Bemessung der Nachhaltigkeit dieser Technologie durchzuführen. Die, den Indikator als relevant einstufoende Gruppe 2 führte das Argument an, dass gerade bei komplexen Technologien eine ausreichende Informiertheit der Betroffenen notwendig sei und dieser Indikator damit auch für CCS relevant sei. Diese unterschiedlichen Urteile konnten im zweiten Plenum nicht aufgelöst werden, so dass man sich auf einen Konsens über einen Dissens einigte.

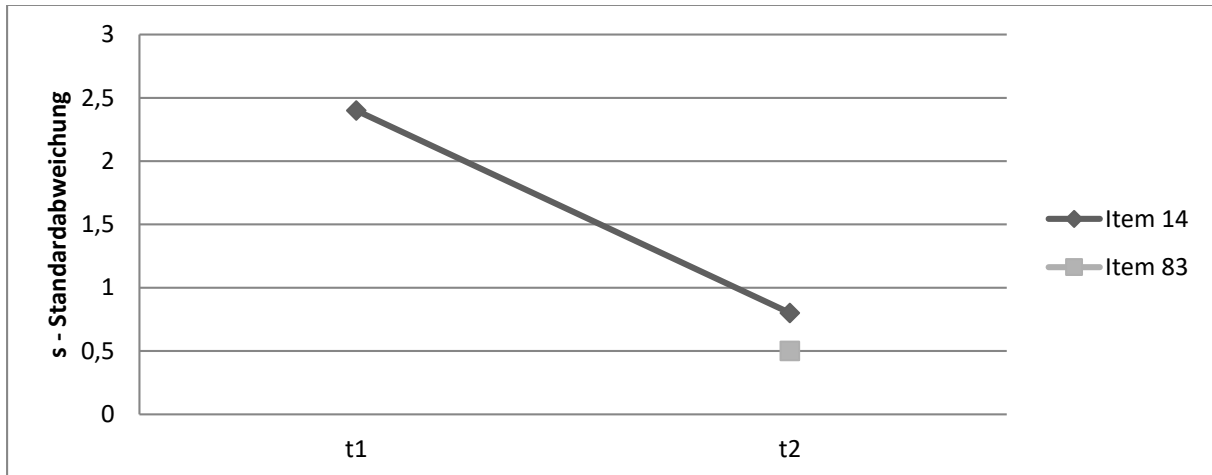
- **Item 22:** Als zweites Item wurde das *Item 22*, das den Indikator *Anzahl von Verbraucherbeschwerden* enthielt, sowohl umformuliert wie auch in *Wichtigkeit* und *Impact* aufgeteilt. In der ersten Runde des Gruppendelphis wurde das Item mit $s=2,4$ und einem Variationskoeffizienten von $v=0,5$ bei einem arithmetischen Mittel von $\bar{x}=5$ von den Kleingruppen als unterschiedlich relevant bewertet. Während Gruppe 1 den abgefragten Indikator als relevant einstufoete (8), beurteilte Gruppe 3 diesen als nicht relevant (2) und Gruppe 2 als eher nicht relevant (5). Aufgrund des bestehenden Dissenses wurde auch dieses Item im ersten Plenum diskutiert und die ablehnende dritte Gruppe begründete ihr Urteil mit dem Argument, dass der Begriff der *Kunden* und damit auch der Indikator zu unspezifisch sei, da besonders Stromkund*innen die relevante Gruppe an Personen für diesen Indikator seien. Diesem Einwand konnten auch die übrigen Teilnehmenden folgen, so dass der Indikator zur zweiten Runde umformuliert wurde und als *Item 22a* und *Item 22b* wie folgt abgefragt wurde: *Anzahl von Verbraucherbeschwerden der Stromkunden*. In der zweiten Runde des Gruppendelphis blieb der Dissens auch zwischen den neu zusammengesetzten Kleingruppen bestehen. So wies das *Item 22a* in dieser zweiten Runde einen Wert der Standardabweichung von $s=2,1$ mit $v=0,5$ bei einem arithmetischen Mittel von $\bar{x}=4,3$ auf. Ebenso das *Item 22b*, das mit $s=2,0$ mit einem Variationskoeffizienten von $v=0,5$ und einem arithmetischen Mittel von $\bar{x}=3,8$ ebenfalls nicht eindeutig und zudem im Mittel als eher nicht relevant eingestuft wurde.¹⁴⁸ Zudem gab es bei diesem Item ein Minderheitsvotum in Gruppe 1, das diesen Indikator allerdings als relevant einstufoete (7), während die Gruppe dies nicht tat (4). Aufgrund der Uneinigkeit zwischen den Bewertungen bei den beiden *Items 22a* und *22b* wurden diese im zweiten Plenum nochmals diskutiert. Die Person, die das Minderheitsvotum abgegeben hatte, führte hierzu das Argument an, dass die Einschätzungen der Kund*innen hinsichtlich einer neuen Technik wichtig seien, da diese durch mögliche Preiserhöhungen die Konsequenzen neuer Technologien spüren könnten. Als Gegenargument wurde von einer Person der dritten Gruppe, die beide Items als nicht relevant bewertet hatte, angeführt, dass die Kund*innen nur sehr selten schriftliche Beschwerden formulieren, sondern lediglich einen Wechsel des Anbieters vollziehen würden. Dieser würde von dem Indikator allerdings nicht erfasst. Wichtig seien damit die Reaktionen der Stromkund*innen. Das Plenum konnte sich konsensual darauf einigen, dass der Indikator *Reaktion der Stromkunden* lauten müsse. Damit konnte im zweiten Plenum eine konsensual getragene Formulierung des Indikators der *Items 22a* und *22b* gefunden werden. Die zwischen der ersten und zweiten Runde vorgenommene Umformulierung von *Kunden* zu *Stromkunden* zeigte

¹⁴⁸ Da es in Gruppe 1 ein Minderheitsvotum gab, ist bei diesem Item zum Zeitpunkt t2 $n=4$.

allerdings im Vorfeld des zweiten Plenums keinen konsensbildenden Effekt der Verständigung auf eine spezifische Zielgruppe.

Neben diesen geteilten Items fand beim Gruppendelphi in PROSUITE bei einem weiteren Item eine Umformulierung im Fragebogen zwischen der ersten und der zweiten Runde statt. Dieses *Item 14* beinhaltete einen Indikator zur subjektiv empfundenen Lebenszufriedenheit und wurde in der ersten Runde wie folgt abgefragt: *Die Technik beeinflusst die subjektiv empfundene Lebenszufriedenheit der Konsumenten*. Die Einschätzungen der Relevanz dieses Indikators lagen in der ersten Runde zwischen den Kleingruppen weit auseinander, so dass die Standardabweichung bei diesem Item einen Wert von $s=2,4$ bei einem Variationskoeffizienten von $v=0,5$ hatte. Der Mittelwert lag dabei mit einem Wert von $\bar{x}=5$ (t1) auf der Seite der Skala, die den Indikator als nicht relevant einstufte. Allerdings beurteilte Gruppe 2 diesen Indikator als relevant (8), während Gruppe 1 und 3 ihn als nicht relevant beurteilten (5/2). Das *Item 14* wurde während der ersten Plenarphase diskutiert. Eine Person aus der ablehnende Gruppe 3 führte als erstes für ihre Einschätzung das Argument an, dass der Indikator für CCS nicht als relevant einzustufen sei. Ergänzend dazu äußerte eine weitere Person aus dieser Gruppe, dass die Zielgruppe des Indikators zu unspezifisch für die Technik CCS sei und eine Umformulierung in *Stromkunden* notwendig sei. Ergänzend dazu wurde aus Gruppe 1 das Argument vorgetragen, dass sich *grüne* Stromkund*innen bei der Verwendung der Technik gut fühlen könnten und dies letztlich einen Einfluss auf die Lebenszufriedenheit haben könne. Alle Teilnehmenden konnten sich auch bei diesem *Item 14* darauf einigen, den Indikator hinsichtlich der Zielgruppe zu spezifizieren, so dass die neue Formulierung im Fragebogen der zweiten Runde wie folgt vorgenommen wurde: *Die Technik beeinflusst die subjektiv empfundene Lebenszufriedenheit der (Strom-)Konsumenten*. Neben dieser Umformulierung wurde im ersten Plenum hinsichtlich der Zielgruppe des Indikators ein weiteres Argument vorgetragen. So sei eine weitere wichtige Personengruppe im Falle von CCS ebenso die Anwohner*innen, die in der Nähe eines solchen Speichers leben würden. Dem konnten ebenso alle Teilnehmenden zustimmen, so dass zusätzlich zum umformulierten *Item 14* ein neues *Item 83* in den Fragebogen aufgenommen: *Die Technik beeinflusst die subjektiv empfundene Lebenszufriedenheit der Anwohner*. Im Hinblick auf *Item 14* zeigt sich durch die Spezifikation der Zielgruppe in der zweiten Runde ein konsensbildender Effekt. So wird dieser Indikator mit $s=0,8$ (t2) und einem Variationskoeffizienten von $v=0,2$ (t2) nun wesentlich einheitlicher beurteilt. Allerdings liegt der Wert des arithmetischen Mittels mit $\bar{x}=5$ (t2) immer noch knapp im als nicht relevant einzustufenden Abschnitt der Skala, so dass die neu zusammengesetzten Gruppen 1 und 3 den Indikator als eher nicht relevant einstufen (5/4) und Gruppe 2, diesen als eher relevant einstuft (6). Durch Einbezug des neuen *Items 83* löst sich dieser Konflikt allerdings auf, denn hier sind sich alle Gruppen einig, dass dieser Indikator mit $s=0,5$ (t2) und $v=0,1$ (t2) bei einem arithmetischen Mittel von $\bar{x}=9,3$ (t2) der relevante Indikator mit der notwendigen Zielgruppe ist, um mittels der *subjektiv empfundene Lebenszufriedenheit* eine Bemessung der Nachhaltigkeit von CCS vornehmen zu können. Abbildung 59 veranschaulicht nochmals die Entwicklung des Wertes der Standardabweichung bei *Item 14* sowie die Verortung des Wertes von *Item 83* zur zweiten Runde des Gruppendelphis.

Abbildung 59: Werte der Standardabweichung des umformulierten Items 14 und des neuen Items 83



Anmerkungen: Item 27: $tn1=3$; $tn2=3$

Quelle: Eigene Darstellung

Nachdem die Dimension der Verständlichkeit besonders im Falle der geteilten Items aufgezeigt hat, dass im Gruppendelphi in PROSUITE eine Abstimmung der Teilnehmenden über einen gemeinsamen Bewertungshorizont besonders notwendig war, um in der zweiten Runde des Gruppendelphis ein Urteil vor einem gemeinsamen Verständnis der Indikatoren vornehmen zu können, soll nun auch bei diesem Fallbeispiel die Dimension der normativen Angemessenheit näher betrachtet werden.

7.5.3.3 Dimension: normative Angemessenheit

In dieser Dimension richtet sich der Fokus der Analyse auf die dokumentierten Argumentationen, die die Teilnehmenden während der Plenarsitzungen geäußert haben. Ziel dieser Fokussierung ist es zu untersuchen, ob es den Teilnehmenden auch im Falle des Gruppendelphis in PROSUITE möglich war, neben den die Evidenz adressierenden Argumenten ebenso normative Einstellungen und Wertvorstellungen als Argumente anzubringen. Ferner auch, ob solche Argumente als legitime Aussagen von den anderen Teilnehmenden im Diskurs angenommen wurden und anknüpfende Argumentationen folgten. Im Falle des Gruppendelphis in PROSUITE steht für eine solche Betrachtung die Dokumentation der beiden Plenarsitzungen in Form eines tabellarischen Protokolls als Datengrundlage zur Verfügung, da kein Ergebnisprotokoll an die Teilnehmenden versandt wurde, kann ein solches nicht ergänzend in die Analyse einbezogen werden.

In der tabellarischen Dokumentation der ersten Plenarsitzung finden sich keine Argumente der Teilnehmenden, die Werturteile, Präferenzen, individuelle Bewertungen von Optionen oder moralische Urteile enthalten. Ebenso stellt sich die Situation im zweiten Plenum dar, in dem der moderierte Austausch ebenso anhand der abgefragten Items strukturiert wurde. In dieser Runde wurde allerdings anschließend eine thematische Öffnung vorgenommen, bei der die Teilnehmenden ihr persönliches Fazit des Gruppendelphis äußern konnten. In diesen Äußerungen von fünf Teilnehmenden finden sich drei Aussagen, die die Dimension der normativen

Angemessenheit adressieren.¹⁴⁹ Die erste dieser Aussagen richtete sich auf die im Gruppendelphi abgefragten Indikatoren, deren Ausrichtung zu *technokratisch* gewesen sei. Diese Aussage wurde mit einem Werturteil begründet, dass weder die *demokratische Lernfähigkeit* als auch *soziale Folgen* in den abgefragten Indikatoren enthalten seien. So würden die Indikatoren letztlich lediglich die *Folgen*, die durch den Einsatz der Technologie CCS entstehen, bewerten. Die zweite Äußerung einer weiteren Person des Teilnehmerfeldes, knüpfte inhaltlich nicht an die erste Äußerung an, sondern formulierte die Präferenz, dass ein Indikator zur *politischen Mobilisierung* oder allgemein ein *Indikator zur Konflikthaftigkeit* fehle. Dieser Präferenz fügte diese Person als Argument das Werturteil an, dass es zum Zeitpunkt des Gruppendelphis bereits *große Diskussionen* über die CCS-Technik gäbe und diese letztlich zu einem *Konflikt* in der Gesellschaft führen können. Auch die dritte Person, die eine die normative Angemessenheit adressierende Aussage tätigte, bezog sich nicht direkt auf die vorhergehenden Argumente der anderen Teilnehmenden, sondern äußerte die Präferenz eines *Indikators zur Langfristigkeit*. Diese Präferenz wurde mit dem moralischen Urteil als Argument hinterlegt, dass *folgende Generationen eine Technologie beherrschen müssten*. Dieser langfristige Bezug sei zwar im neuen Indikator zur *generationssensiblen Problemverschiebung*¹⁵⁰ enthalten, so fehlten aber noch Bemessungskriterien wie *Reversibilität* oder *Fehlerhaftigkeit*.

Letztlich bezogen sich auch die beiden von zwei weiteren Personen getätigten Aussagen nicht auf die, von den anderen Teilnehmenden vorgetragenen Argumenten, sondern adressierten zum einen die *Vulnerabilität der Technologie* und zum anderen die Dauer der *Lagernutzbarkeit* von CCS, so dass hier keine Bezugnahme auf jeweils andere Argumente in dieser Fazitrunde aus den im Protokoll dokumentierten Äußerungen veranschlagt werden kann. Anders als bei einem anhand von Items geführten Austausch im Plenum ist dies allerdings auch nicht Ziel einer Äußerung eines persönlichen Fazits der einzelnen Teilnehmenden. So wurde hier auch nicht ein gegenseitiger Austausch durch die Moderation befördert. Letztlich wäre ein Bezug der Argumente aufeinander lediglich bei entschiedenem Widerspruch einer Person zu einem Argument einer anderen Person erfolgt. Dies war in der Abschlussrunde dieses Gruppendelphis in PROSUITE allerdings nicht der Fall. Letztlich wurde die Dimension der normativen Angemessenheit damit in diesem Fallbeispiel nur im Abschluss adressiert und im Protokoll finden sich keine solcher kategorisierbaren Argumente im Austausch in den beiden Plenarsitzungen, die einen direkten Bezug zu den einzelnen Items zulassen.

Allerdings hatten die Teilnehmenden die Möglichkeit eigene Indikatoren im Fragebogen der ersten Runde des Gruppendelphis schriftlich in den Prozess einzubringen. Diese Indikatoren werden im folgenden Abschnitt zur grundlegenden Funktionslogik eines Gruppendelphis näher erläutert.

¹⁴⁹ Die kategorisierten Aussagen mit Bezug zur normativen Angemessenheit finden sich in der Zusammenstellung in Anhang 11.17.

¹⁵⁰ Dieser Indikator wurde von einer Kleingruppe unter der offenen Frage im vierten Abschnitt des Fragebogens schriftlich vorgeschlagen und zur zweiten Runde des Gruppendelphis als Item 93 abgefragt. Die Ergebnisse dieses und der übrigen von den Teilnehmenden vorgeschlagenen Indikatoren, werden im folgenden Abschnitt 7.5.3.4 näher beleuchtet.

7.5.3.4 Dimension: grundlegende Funktionslogik

In der Dimension zur grundlegenden Funktionslogik werden bei diesem Fallbeispiel zu Beginn die neuen Items betrachtet, die in Folge der ersten Runde des Gruppendelphis im zweiten Fragebogen erstmals abgefragt wurden. Im Gruppendelphi in PROSUITE gab es zwei verschiedene Typen von solchen *neuen* Items zur zweiten Runde. Zum einen hatten die teilnehmenden Expert*innen bereits in der ersten Kleingruppenphase eigene Indikatoren als Items vorgeschlagen. Diese Möglichkeit war im vierten Abschnitt des Fragebogens gegeben, in der der Fragetext zu dieser offenen Abfrage lautete: *Im vierten Abschnitt können sie Indikatoren ergänzen, die ihrer Meinung nach zur Bemessung der sozialen Nachhaltigkeit von CCS-Technologien herangezogen werden sollten, sich aber nicht in der vorhergehenden Auflistung finden.* Die Teilnehmenden hatten dabei die Möglichkeit, diese Indikatoren bereits den jeweiligen Kategorien zuzuordnen, die sich anhand der Struktur der ersten drei Abschnitte des Fragebogens aufgliederten. Somit konnten sowohl *generelle Indikatoren*, wie auch *organisations-spezifische Indikatoren* und *produkt-spezifische Indikatoren* vorgeschlagen werden.

Bereits im Kapitel zum Fallbeispiel des Gruppendelphis in SEAWIND¹⁵¹ wurden diese Art von Items beschrieben, die nicht dem typischen Ablauf eines Gruppendelphis folgen. In diesem Fall zeichneten sich die Items durch drei Aspekte aus, die eine Betrachtung hinsichtlich der generellen Funktionslogik interessant machen. So charakterisieren sich diese Items im Gegensatz zu den Items, die dem idealen Ablauf eines Gruppendelphis entsprechen, durch drei Unterschiede:

1. Die Formulierung der Items wurde von den Teilnehmenden im Fragebogen der ersten Runde (t1) durchgeführt.
2. Die Items wurden erst im Plenum (t1) diskutiert, bevor sie in der Kleingruppenphase (t2) bewertet wurden.
3. Es fand keine erneute Abfrage (auch bei Dissens) statt.

Im Gegensatz zum Gruppendelphi in SEAWIND wurden die Items in diesem Fallbeispiel allerdings nicht im ersten Plenum (t1) diskutiert, sondern direkt in den Fragebogen der zweiten Runde übernommen und somit fand ein Diskurs zu diesen Items erstmalig in den Kleingruppen statt (t2). Folglich treffen die Punkte 1 und 3 der Besonderheiten des SEAWIND-Gruppendelphis ebenso auf die in diesem Gruppendelphi zu Indikatoren der Bemessung von Nachhaltigkeit im Falle von CCS zu.

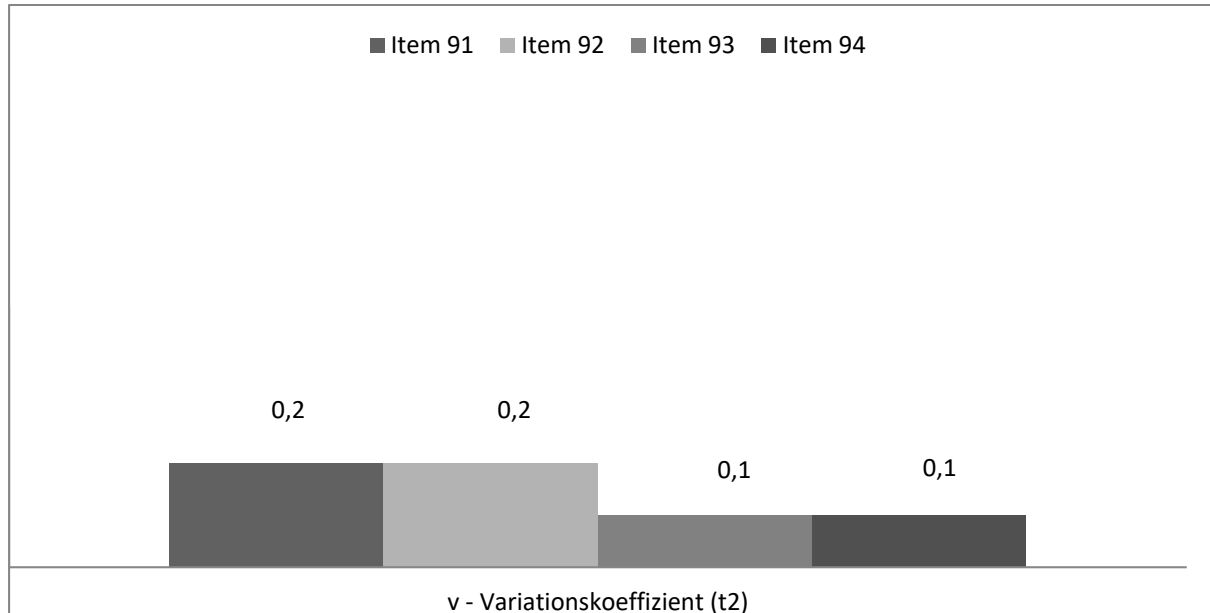
Im Gruppendelphi zu PROSUITE wurden insgesamt vier Items als *generelle Indikatoren* in der offenen Abfrage der ersten Runde vorgeschlagen, die in den Fragebogen der zweiten Runde übernommen wurden:

- **Item 91:** *Macht die Technologie Änderungen von gesetzlichen Regelungen notwendig (z.B. Planungsrecht, Raumordnungsrecht, Haushaltsrecht, etc)*
- **Item 92:** *Vertrauen in langfristige Regelungsfunktionen*
- **Item 93:** *Generationssensible Problemverschiebung (Folgekosten für zukünftige Generationen)*
- **Item 94:** *Qualität und Intensität der Bürgerbeteiligungsverfahren*

¹⁵¹ siehe Kapitel 7.5.3.4.

Diese vier Items finden sich auch in Abbildung 60, in der sich auch die jeweiligen Werte des Variationskoeffizienten jedes Items finden. Wie sich anhand dieser Werte bereits vermuten lässt, wurden alle dieser vier Items von den Kleingruppen in der zweiten Phase des Gruppendelphis konsensual beurteilt. Zudem wurden diese Items von den Kleingruppen alle geschlossen im positiven Bereich der Skala verortet, die das Item als relevant kennzeichnet.

Abbildung 60: Variationskoeffizient (t2) bei Items aus offener Abfrage (t1) im PROSUITE-Gruppendelphi



Anmerkungen: n=3.

Quelle: Eigene Darstellung

So wurde das *Item 91* mit einem arithmetischen Mittel von $\bar{x}=8,7$ (t2) allerdings mit relativ hoher Standardabweichung von $s=1,9$ (t2) trotz dessen geschlossen positiv bewertet. Dies gilt ebenso für *Item 92*, dass mit $\bar{x}=8,0$ (t2) und einer Standardabweichung von $s=1,6$ (t2) hoch bewertet wurde, allerdings hinsichtlich der *Stärke* der Relevanz auch hier unterschiedliche Urteile der verschiedenen Kleingruppen vorlagen. Anders bei den folgenden beiden Items, die im Falle von *Item 93* mit $\bar{x}=9,3$ (t2) bei $s=0,9$ (t2) und im Falle von *Item 94* mit $\bar{x}=8,7$ (t2) bei $s=0,5$ (t2) geschlossen als sehr relevant eingestuft wurden. An dieser Stelle offenbart sich allerdings ein Problem in der Konzeption dieses Gruppendelphis, denn wenn diese Items auch von den Teilnehmenden als relevant eingestuft werden, so ergibt sich für die weitere Auswertung und Verwendung dieser abgefragten Indikatoren eine inhaltliche Lücke. Da die Formulierung der Indikatoren in den Kleingruppen der ersten Runde des Gruppendelphis stattfand und der Austausch zu diesen dann in der zweiten Kleingruppenphase geschah, fand aufgrund der vorherrschenden Einigkeit hinsichtlich der einzelnen Indikatoren zur zweiten Runde letztlich ebenso kein Austausch im Plenum statt. So haben die Teilnehmenden einen inhaltlichen Austausch im geschlossenen Raum geführt, der dem Forscherteam nicht bekannt ist. Ebenso gilt dies für die Formulierung der Indikatoren in der ersten Kleingruppenphase. Was bei einer dem idealen Ablauf eines Gruppendelphis folgenden Konzeption kein Problem darstellt, da die abgefragten Items von den Durchführenden des Gruppendelphis formuliert werden und damit auch die rahmenden Inhalte bekannt sind, führt im Falle des Gruppendelphis in PROSUITE zu einer inhaltlichen Lücke für die Ergebnisauswertung. So bleiben letztlich nur die Formulierung-

gen der Indikatoren ohne deren Hintergründe und konkreten Inhalte näher beleuchten zu können. Dies hätte nur verhindert werden können, indem die Items auch im ersten Plenum (oder zumindest im zweiten Plenum) diskutiert worden wären, bevor sie in der Kleingruppenphase abgefragt wurden, wie es im Falle von SEAWIND der Fall war.

Wie bereits zu Beginn dieses Abschnitts angemerkt wurde, wurden in diesem Fallbeispiel weitere Items als *neue* Items im Fragebogen der zweiten Kleingruppenphase abgefragt. Folgt man wieder den drei Besonderheiten solcher Items, die für den Fall SEAWIND beschrieben wurden, so decken auch diese Items in PROSUITE zwei der drei Besonderheiten ab:

- Die Items wurden erst im Plenum (t1) diskutiert bevor sie in der Kleingruppenphase (t2) abgefragt wurden.
- Es fand keine erneute Abfrage (auch bei Dissens) statt.

Im Falle dieser Items wurde die Formulierung damit nicht im Fragebogen der ersten Kleingruppenphase durchgeführt, sondern die Items wurden in der ersten Plenumsdiskussion als Produkt in den Fragebogen der zweiten Runde übernommen. Auf diesem Wege wurden zur zweiten Runde folgende fünf Items abgefragt:

- **Item 81:** *Bildung*
- **Item 82:** *Regionale Dispersität*
- **Item 83:** *Subjektiv empfundene Lebenszufriedenheit (Anwohner)*
- **Item 84:** *Zertifizierung der Technologie*
- **Item 85:** *Verdrängung anderer energierelevanter Technologien*

Neben den gänzlich an neue Inhalte anknüpfenden *Item 81* hinsichtlich der *Bildung*, *Item 82* zur *regionalen Dispersität* und *Item 85* zur *Verdrängung anderer energierelevanter Technologien* wurden die beiden *Items 83* und *85* als inhaltliche Differenzierung zu bereits bestehenden Items formuliert. So wurde hinsichtlich des Indikators der *subjektiv empfundenen Lebenszufriedenheit* neben der Ergänzung in *Item 14* in dem die Stromkonsument*innen als relevante Zielgruppe, ebenfalls ein neuer Indikator durch *Item 83* im ersten Plenum vorgeschlagen und letztlich konsensual in den Fragebogen der zweiten Runde aufgenommen, der die *subjektiv empfundene Lebenszufriedenheit* der *Anwohner* als Bemessungskriterium der Nachhaltigkeit für die konkrete Technologie CCS darstellt. Ebenso im Falle des *Items 84*, das mit dem Indikator der *Zertifizierung einer Technologie* als ergänzender Indikator zu dem in *Item 23* abgefragten *Zertifizierung eines Produkts* in den Fragebogen der zweiten Runde integriert wurde. Unter Beachtung, der in Abbildung 61 veranschaulichten Werte des Variationskoeffizienten fällt das *Item 84* in der Reihe der dargestellten Items mit $v=0,5$ (t2) als nicht einheitlich bewertetes Item auf. Zudem weist dieses Item hinsichtlich der Fallzahl ein Minderheitsvotum auf, was bei den übrigen Items nicht der Fall war. So wurden in der zweiten Runde sowohl das *Item 23* mit $\bar{x}=5,3$ bei $s=1,6$ und $v=0,3$ als auch das *Item 84* mit $\bar{x}=5,3$ bei $s=2,8$ und besagtem $v=0,5$ mit jeweils einem Minderheitsvotum in Gruppe 3 nicht einheitlich von den Expert*innen im Hinblick auf seine Relevanz bewertet. Aufgrund der inhaltlichen Verknüpfung beider Items wurde der Austausch zu diesen im zweiten Plenum gemeinsam geführt. Während eine Person hier argumentierte, dass ein *Produkt* für die Bemessung der Nachhaltigkeit immer relevanter sei als die eigentliche *Technologie*, da es nicht üblich sei ganze Technologien zu *zertifizieren*,

argumentierte eine andere Person, dass dies für Entscheidungsträger*innen allerdings relevanter sei, da sie nicht zwischen einzelnen Produkten, sondern zwischen Technologien abwägen müssten.

Abbildung 61: Variationskoeffizient bei neuen Items im PROSUITE-Gruppendelphi (t2)



Anmerkungen: Item 84 n=4, Alle übrigen Items n=3.

Quelle: Eigene Darstellung

Über den Einbezug der jeweiligen Zielgruppen des jeweiligen Indikators konnte so auch ein Konsens im Plenum erreicht werden. Wenn die Zielgruppe aus Konsument*innen bestände, sei die in *Item 23* abgefragte *Zertifizierung eines Produktes* ein relevanter Indikator und wenn die Zielgruppe aus Entscheidungsträger*innen bestände, so sei die in *Item 84* abgefragte *Zertifizierung einer Technologie* ein relevanter Indikator. Die drei übrigen in Abbildung 61 dargestellten Items stellen auch unter Beachtung der übrigen statistischen Werte ein konsensuales Urteil der jeweiligen Indikatoren als relevant für die Bemessung der Nachhaltigkeit der Technologie CCS dar: So wurden *Item 82* zur *regionalen Dispersität* mit $\bar{x}=8,7$ (t2) bei $s=0,5$ (t2), *Item 83* zur *subjektiv empfundenen Lebezufriedenheit der Anwohner*innen* mit $\bar{x}=9,3$ (t2) bei $s=0,5$ (t2), als auch *Item 85* zur *Verdrängung anderer energierelevanter Technologien* mit $\bar{x}=8,7$ (t2) bei $s=1,2$ (t2) geschlossen zur zweiten Runde als relevant beurteilt. Überdies bleibt von den im Fragebogen der zweiten Runde neu aufgenommenen Items noch das *Item 81* zu diskutieren. Hier ergibt sich ein Problem, dass sich auch in den anderen Fallbeispielen bereits gezeigt hat, in diesem letzten Fall nun nochmals vertieft in der Dimension zur grundlegenden Funktionslogik exemplarisch dargestellt werden soll. So deuten die statistischen Werte auch beim *Item 81* zur zweiten Runde auf einen Konsens hin. Der Indikator *Bildung* weist hierbei mit einem arithmetischen Mittelwert von $\bar{x}=9,0$ bei einer Standardabweichung von $s=1$ und einem Variationskoeffizienten von $v=0,1$ auf eine konsensuale Einschätzung der Kleingruppen dieses Indikators als relevant hin. Allerdings haben nur zwei der drei Kleingruppen in der zweiten Runde des Gruppendelphis eine Bewertung im Fragebogen abgegeben. Da Gruppe 2 kein Votum bezüglich dieses *Items 81* vermerkt hat, beträgt die Fallzahl damit $n=2$ bei $N=3$. Durch diese Enthaltung bleibt die Zuschreibung einer konsensualen Einschätzung der Relevanz dieses Items eine bloße Vermutung. Da das Item im zweiten Plenum nicht besprochen wurde, bleibt letztlich unklar, worauf die Enthaltung der zweiten Kleingruppe basiert. Hier versteckt sich eine Gefahr eines Fehlschlusses bei Enthaltungen einen Konsens anzunehmen ohne die Gründe der Enthaltung abzufragen.

Letztlich soll dieses Gruppendelphi allerdings auch daran bemessen werden, ob es im Hinblick auf seine Fragestellung und den damit verbundenen Zielen des Einsatzes der Methode wertvolle und nutzbare Ergebnisse befördern konnte. Ziel des Gruppendelphis in PROSUITE war es aus insgesamt 34 vorselektierten Indikatoren ein Set von Indikatoren zu identifizieren, die für den Fall von *Carbon Dioxide Capture and Storage-Technologien* eine Messung der sozialen Dimension von Nachhaltigkeit dieser Technologie ermöglichen. Dabei war das Ziel nicht, diese Liste der Indikatoren so weit wie möglich zu reduzieren, wie auch im Umkehrschluss möglichst weit auszubauen, wie es der Einsatz der offenen Frage im vierten Abschnitt des Fragebogens vermuten ließe. Die Zielstellung war letztlich, dass die teilnehmenden Expert*innen sich konsensual auf ein anwendbares und sinnhaftes Set von Indikatoren einigen konnten.

So wurden für den ersten Abschnitt der *generellen Indikatoren* insgesamt sechs Indikatoren identifiziert, die für die soziale Dimension der Bemessung der Nachhaltigkeit von CCS aus Sicht der Expert*innen geeignet sind und von diesen somit konsensual als relevant bewertet wurden. Von diesen sechs Items wurden bereits drei der Items in der ersten Kleingruppenphase konsensual als relevant bewertet (*Item 1, 2 und 16*), während drei der Items zur zweiten Runde ebenso einhellig als relevant eingestuft wurden (*Items 13, 82 und 83*). Überdies wurden ebenso die vier Indikatoren (*Item 91, 92, 93 und 94*) konsensual zur zweiten Runde als relevant bewertet, auch wenn diese aufgrund fehlender zusätzlicher Erläuterung zu den Inhalten – wie oben beschrieben – nur sehr bedingt als fruchtbare Ergebnisse gewertet werden können. Im zweiten Abschnitt zu den *organisations-spezifischen Indikatoren* finden sich ebenso sechs Indikatoren, die von den Teilnehmenden geschlossen als relevant beurteilt wurden. Davon bereits fünf Indikatoren in der ersten Kleingruppenphase (*Item 20, 21, 24, 25 und 26*) und ein Indikator in der zweiten Runde des Gruppendelphis (*Item 85*). Für den dritten Abschnitt, der *produkt-spezifische Indikatoren* als Itemset enthielt, konnten letztlich drei Indikatoren konsensual als relevant ausgearbeitet werden, wovon ein Indikator bereits zur ersten Runde identifiziert wurde (*Item 24*). Zusätzlich konnte durch den Austausch im zweiten Plenum der Dissens hinsichtlich der Relevanz der Indikatoren zur *Zertifizierung des Produktes* (*Item 23*) bzw. der *Technologie* (*Item 84*) aufgelöst werden, so dass auch diese in die Liste der konsensual als relevant eingestuften Indikatoren aufgenommen werden konnten. Das Ziel aus der Liste von 34 möglichen Indikatoren diejenigen zu identifizieren, die aus Sicht der Teilnehmenden des Gruppendelphis in PROSUITE als relevant für die Bemessung der Nachhaltigkeit von CCS beurteilt wurden, konnte erreicht werden. Das Gruppendelphi hat diese Auswahl auf 14 relevante Indikatoren verdichtet, so dass nutzbare Ergebnisse für den weiteren Prozess im Rahmen des Projektes PROSUITE erzielt werden konnten.

Durch die Trennung einzelner Indikatoren in *Wichtigkeit* und *Impact* zur zweiten Runde des Gruppendelphis wurden neun zusätzliche Indikatoren von den Teilnehmenden konsensual als relevant bewertet, die losgelöst von der Technologie CCS für die generelle Bemessung von Nachhaltigkeit als relevant zu betrachten sind. Diese anhand der *Wichtigkeit* gekennzeichneten Indikatoren sind allerdings nicht direkt für die Fragestellung des Gruppendelphis relevant, da sie bereits durch die separate Abfrage keinen Bezug zu *Carbon Dioxide Capture and Storage-Technologien* haben, im Prozess des Gruppendelphis von den Teilnehmenden aber explizit vorgeschlagen wurden. Ob sich daraus Schlüsse auf die Qualität des Gesamtergebnisses des Gruppendelphis ziehen lassen, soll im nächsten Abschnitt näher beleuchtet werden.

7.5.4 Zwischenfazit: Gruppendelphi in PROSUITE

Der Fokus des Gruppendelphis lag auf einer Bemessung der Nachhaltigkeit von *Carbon Dioxide Capture and Storage-Technologien* mit dem Ziel Indikatoren zu finden, die für eine soziale Dimension dieser Bemessung relevant sein können. Durch, die im ersten Plenum vorgenommene Trennung zwischen der generellen *Wichtigkeit* und dem *Impact* auf die konkrete Technologie CCS wurde der Fokus der Fragestellung teilweise verwaschen. So wurden in der zweiten Runde des Gruppendelphis insgesamt 14 Items abgefragt, die der Fragestellung des Gruppendelphis nicht dienlich waren. Dabei spielt es weniger eine Rolle, dass die Ergebnisse nicht auch für das Gesamtprojekt genutzt werden konnten oder der Fragestellung zuwiderlaufen, sondern inwieweit diese Items die Urteile der am Impact ausgerichteten Items beeinträchtigt haben. Wie in Abschnitt 7.5.3.2 dargelegt, wurden insbesondere die Items zur generellen *Wichtigkeit* der Indikatoren zur zweiten Runde des Gruppendelphis konsensual als relevant eingestuft. Direkte Auswirkungen daraus auf die für die Fragestellungen relevanten Items zum *Impact* auf CCS zu schließen wäre ein Fehlschluss, denn es kann nicht überprüft werden, ob diese ohne die im ersten Plenum vorgenommene Trennung als relevanter eingeschätzt worden wären oder nicht. Drei problematische Felder dieser Trennung können allerdings anhand der vorliegenden Dokumentation des Gruppendelphis in PROSUITE abgesteckt werden:

- 1. Vollständigkeit:** Im ersten Plenum des Gruppendelphis wurden die Items diskutiert, bei denen nach der ersten Kleingruppenphase kein konsensuales Urteil zwischen den Kleingruppen vorlag. Von diesen 17 Items wurden insgesamt 14 Items in *Wichtigkeit (a)* und *Impact (b)* aufgeteilt. Es kann nicht ausgeschlossen werden, dass eine solche Aufteilung auch bei den verbliebenen 17 Items vorgenommen worden wäre, die in der ersten Runde konsensual als relevant eingestuft wurden. Diese wurden allerdings im ersten Plenum nicht diskutiert, so dass hier überdies ein Scheinkonsens nicht gänzlich ausgeschlossen werden kann, da einzelne Kleingruppen bereits in der ersten Kleingruppenphase einzelne Items nach der *Wichtigkeit* statt des *Impacts* beurteilt haben könnten.
- 2. Inhaltlicher Fokus:** In der ersten Plenarsitzung wurde sich bereits im Diskurs zum ersten besprochenen *Item 6* darauf verständigt, dass eine Aufteilung in *Wichtigkeit* und *Impact* konsensual von den Teilnehmenden präferiert wird. Unter Bezugnahme des tabellarischen Protokolls fällt auf, dass der Fokus der Diskussion häufig auf eine mögliche Aufteilung der Items ausgerichtet war. So findet sich im Protokoll bei sieben der Items lediglich die Anmerkung, dass dieser Indikator in *Wichtigkeit* und *Impact* aufgeteilt wurde. Für die Fragestellung des Gruppendelphis wäre ein Austausch hinsichtlich der Indikatoren mit dem Fokus auf die Bemessung der Nachhaltigkeit von CCS allerdings relevanter gewesen.
- 3. Zeit:** Der Faktor Zeit ist auch in diesem Fall relevant. So war auch diese im Falle des Gruppendelphis in PROSUITE knapp bemessen, so dass der Austausch hinsichtlich der Aufteilung der Items stattdessen für einen Diskurs über die Relevanz der jeweiligen Indikatoren zur Bemessung der Nachhaltigkeit von CCS hätte genutzt werden können oder gar sollen. Dies gilt besonders für die *Items 91, 92, 93 und 94*, die von den Teilnehmenden vorgeschlagen (t1) und konsensual als relevant bewertet wurden (t2), zu denen allerdings keine weiteren inhaltlichen Informationen vorliegen, außer deren schriftlicher Formulierung.

So müssen die Ergebnisse dieses Gruppendelphis hinsichtlich der zentralen Fragestellung unter diesen Unsicherheiten mit gewissen Einschränkungen betrachtet werden, so dass die Bewertung der Indikatoren nicht als ideales Ergebnis angesehen werden kann. Wie bereits im Falle des in Kapitel 7.2 dargestellten SEAWIND-Gruppendelphis fällt beim Gruppendelphi in PROSUITE umso mehr auf, dass sich nur wenige Beispiele finden, in denen die Teilnehmenden mittels Argumentation die Dimension der normativen Angemessenheit adressiert haben.

In diesem Fallbeispiel finden sich in beiden Plenarsitzungen keine Aussagen die als solche kategorisiert werden konnten. Lediglich in der abschließenden Fazitrunde sind entsprechende Äußerungen zu finden. Beiden Fällen ist gemein, dass die Items nicht als Aussagen formuliert waren und überdies keine Soll-Komponenten enthielten, sondern eine direkte Abfrage von Bemessungsindikatoren stattfand. Auch im Falle dieses Gruppendelphis stellt sich damit die Frage, ob eine direkte Abfrage den Raum für einen Diskurs bereits soweit begrenzt, dass der Austausch im Plenum durch zu enge Möglichkeitsräume für Argumente und Begründungen eingeschränkt ist. Letztlich ist eine Aussage darüber im Falle von PROSUITE nicht ohne Vorbehalt möglich, da das tabellarische Protokoll nicht umfassend ist, so dass nicht alle Aussagen der Teilnehmenden erfasst wurden. Da dieses Protokoll die einzige Datenquelle für eine Betrachtung des Diskurses dieses Gruppendelphis, besonders unter den Dimensionen der Verständlichkeit und der normativen Angemessenheit ist, stellt dieses Fallbeispiel keine ideale Datenquelle für eine abschließende Bewertung dar.

Zwei wichtige Erkenntnisse lassen sich allerdings aus dem letzten empirischen Fall eines Gruppendelphis ziehen. Der geschützte Raum des Diskurses der Kleingruppen birgt die Gefahr, dass bestimmte Inhalte, Hintergründe und Argumente dem durchführenden Forschungsteam verborgen bleiben, wie bei den *Items 91, 92, 93 und 94* in diesem Falle geschehen. Hier muss darauf geachtet werden, dass ein Austausch in einer der Plenarsitzungen stattfindet, auch wenn ein Konsens zwischen den Urteilen der Kleingruppen besteht. Die Fragestellung und die damit verbundenen, abzufragenden Sachverhalte sollten während der Durchführung eines Gruppendelphis immer oberste Priorität haben. So wurde zwar bereits im ersten Fragebogen zu jeder Itematterie darauf hingewiesen, dass die Indikatoren zur Messung der Nachhaltigkeit im Falle von CCS bewertet werden sollen, allerdings haben einzelne Teilnehmende diesen Bewertungshorizont eigenständig erweitert. Diesem Sachverhalt wurde mit der Aufteilung der Items zur zweiten Runde begegnet, allerdings wäre rückblickend eine Bearbeitung der ursprünglichen Fragestellung unter dem angedachten Bewertungshorizont ratsam gewesen. Hier hätte sich das durchführende Forschungsteam gegen den Vorschlag aus dem Teilnehmer*innenfeld mit Verweis auf die Beispieltechnologie stellen können und im Sinne der Qualität der Ergebnisse auch stellen müssen.

Letztlich sind auch in diesem Fallbeispiel Urteilsänderungen hin zu konsensualen Einschätzungen zu finden, die die angenommene Funktionsweise eines Gruppendelphis aufzeigen. Welche Schlüsse und Erkenntnisse sich aus den fünf Fallbeispielen ziehen lassen, soll im nächsten Kapitel näher ausgearbeitet werden.

8 Diskussion der Ergebnisse

In diesem Kapitel werden die analysierten empirischen Fälle im Hinblick auf die vorhergehenden theoretischen Annahmen sowie auf die damit verbundenen Fragestellungen kontrastiert, um ein Gesamtbild und Urteil über die Funktionsweise im Sinne eines konsensbildenden Erhebungsverfahrens abbilden zu können.

Dabei sollte zuerst auf die generelle Kritik eingegangen werden, der sich sowohl das klassische Delphiverfahren als auch das Gruppendelphi stellen müssen. Diese Kritik äußert sich darin, dass besonders dogmatische Personen, die während des Verlaufes eines Delphis stur auf ihrem Urteil beruhen, sich letztlich gegenüber weniger dogmatischen Personen durchsetzen. Die *Holdouts* würden damit über die Tendenz der *Swingers* dominieren und letztlich würde sich nicht ein rationales, sondern ein dogmatisches Ergebnis durchsetzen (vgl. Vorgrimmer/Wübben 2003: 769). Dieser Kritik aus Sicht des klassischen Delphiverfahrens von Häder mit kognitionspsychologischen Annahmen begegnet. Dabei wird dem jeweiligen Urteil zugrundeliegend, entweder zwischen chronisch verfügbaren oder temporär verfügbaren Informationen unterschieden und so als Folgen entweder Stabilität des Urteils oder aber Variabilität im Urteil postuliert. Aufgrund dieser Unterscheidung soll während des Verlaufes über die Runden eines Delphis ein Kontexteffekt hervorgerufen werden, der zu einer Verbesserung der jeweiligen Urteile der Teilnehmenden führt und letztlich eine Annäherung an das fruchtbarste Urteil bewirkt (vgl. Häder 2009: 50). Dies ist allerdings eine rein deskriptive Beschreibung eines möglichen Vorganges. Letztlich bemessen sich solche *fruchtbaren* Ergebnisse eines Delphis besonders daran, dass aus globalen Wissensbeständen unter Berücksichtigung der bestehenden Unsicherheiten rationale Urteile überführt werden können (vgl. Hammond 1996: 7). Damit zeichnet sich ein Urteil dadurch aus, dass das Ergebnis eines Delphis einem sachgerechten und wertadäquaten Urteil entspricht, also durch eine auf Hinweisreizen basierte rationale Urteilsfindung zwischen Analyse und Intention darstellt. Diese Annahmen stellen auch eine Basis für die theoretische Fundierung des Gruppendelphis dar, bieten im Falle der modifizierten Methode aufgrund der diskursiven Elemente allerdings noch weitere notwendige theoretische Ansatzpunkte, die sich auf die diskursiven Effekte eines argumentativen Austausches stützen (vgl. Kuhn et al. 2014: 67). Während bei einem klassischen Delphi im ersten Schritt die Frage des *lasse ich mich auf die Majorität ein und überdenke mein Urteil* im Raum steht, ist es im Gruppendelphi vielmehr die Frage des *überzeugt mich das Argument meines Gegenübers*, die jeweils eine Urteilsänderung der Teilnehmenden möglich machen.

Die Dimensionen des Testdesigns dieser Arbeit hatten mittels der Analyse der empirischen Fallbeispiele das Ziel, sowohl der oben geschilderten Kritik an Delphiverfahren zu begegnen, wie auch den konsensbildenden Effekt eines argumentativen Austauschs zu belegen. Die Verwendung des Begriffs eines Diskurses bedingt höhere Ansprüche an das Gruppendelphi im Vergleich zum klassischen Delphiverfahren. Durch die methodisch diskursive Erweiterung sind die Begründungen und Argumente für die jeweiligen Urteile der Teilnehmenden zentral für die Funktionsweise, da diese auf verschiedenen Ebenen rational und damit legitim sein müssen. Aufbauend auf diesem Verständnis von Rationalität wurden die Dimensionen des Testdesigns konzipiert. Dabei beinhaltet die erste zentrale Dimension die Analyse hinsichtlich des Geltungsanspruchs der Gültigkeit, da es sich auch bei einem Gruppendelphi in den diskursiven Phasen vorwiegend um den Typ eines Wissensdiskurses handelt. Letztlich sollen damit die

erzielten Ergebnisse als *rationales Urteil* eingestuft werden und ein wahrheitsanaloges Geltungskonzept anstreben, dass sowohl raumzeitliche wie auch soziale Beschränkungen beachtet (vgl. Habermas 1981: 56; Renn/Webler 1998: 51; Häder 2009: 40; Niederberger 2015a: 41).

Die Ergebnisse der einzelnen dargestellten Gruppendelphis müssen sich demnach daran bemessen lassen, ob sie den Qualitätsmerkmalen eines handlungsleitenden Diskurses entsprechen. Wenn die Ergebnisse dem ersten Geltungsanspruch nach **Gültigkeit** genügen, so müssen sich die statistischen Werte in Richtung eines Konsenses bewegen oder einen Konsens über einen Dissens markieren. Überdies durften bei dieser Betrachtung, die in den Plenarsitzungen durch die Teilnehmenden vorgebrachten Argumente für die jeweiligen Urteile nicht aus der Analyse ausgeklammert werden, besonders, wenn atypische Entwicklungen oder Ergebnisse in den empirischen Fällen eine vertiefende Betrachtung notwendig gemacht haben.

Im Fall des Gruppendelphis in **SAUBER+** standen insgesamt 16 Items für eine Analyse in der Dimension der Evidenz zur Verfügung, die über den Verlauf in jeder der Runden abgefragt wurden. Neben den beiden am Workshoptag stattfindenden Befragungen in Kleingruppen gab es in diesem Fall ebenso eine Vorbefragung. Die Analyse der statistischen Werte zeigte dabei die anhand der Vorannahmen veranschlagte Entwicklung hin zu einem konsensualen Ergebnis oder zu einem Konsens über einen Dissens auf. Insbesondere im Falle eines Konsenses über einen Dissens traten in der Analyse ebenso Muster zu Tage, die in den Plenardiskussionen auf wissensbasierten Argumentationen beruhten und Auswirkungen auf die getroffenen Ergebnisse hatten, die letztlich auch nach dem Versandt eines Protokolls mit den zentralen Ergebnissen von allen Teilnehmenden getragen wurden. Hierbei funktionierten neben den 10er-skalierten Ratingskalen auch die eingesetzten Rankingskalen, wie beispielsweise die Budgetfragen.

Auch der Fall des Gruppendelphis in **SEAWIND** machte einen Vergleich über drei Runden möglich, wobei auch hier eine Vorbefragung stattfand und auch hier unterstützen die Entwicklungen der statistischen Werte - insbesondere während des Workshoptages - die angenommene Identifikation von konsensual getragenen Urteilen der Teilnehmenden. Im Falle von SEAWIND trat allerdings erstmals die Problematik hinsichtlich der von Kleingruppen teilweise eingenommenen Enthaltungen auf. So könnten solche Enthaltungen bei exklusiver Betrachtung der statistischen Werte ohne Einbezug der Fallzahl als ein vermeintlicher Konsens interpretiert werden. Da die Enthaltung allerdings nicht als (blindes) Einverständnis zu den Urteilen der anderen Kleingruppen interpretiert werden kann, müssen auch diese Items im Plenum besprochen und mit Argumenten unterstützt werden, um einen möglichen Konsens zu solchen Items im Diskurs abzustimmen, bzw. in der folgenden Runde des Gruppendelphis nochmals abzufragen. Eine zweite Besonderheit bietet dieser Fall hinsichtlich der Formulierung der Items, da hier keine Aussagen formuliert wurden, sondern Maßnahmen direkt mittels 10er-skalierten Items abgefragt wurden, was besonders Auswirkungen auf die weiteren Dimensionen des Testdesigns hatte. Letztlich spiegelt sich aber auch im Falle der evidenzbasierten Betrachtung von SEAWIND, sowohl bei den statistischen Werten der 10er-skalierten Items sowie beispielsweise beim Fragetyp der Auswahlfragen, die im Ideal angenommene Entwicklung hin zu einem konsensualen Ergebnis wider. Dass es sich dabei um keinen Scheinkonsens während des Workshoptages gehandelt hat, zeigt auch hier der nicht stattgefundene Widerspruch hinsichtlich des Ergebnisprotokolls des Gruppendelphis. Im Falle von SEAWIND muss

allerdings trotz der Entwicklung hin zu konsensualen Urteilen eine Einschränkung hinsichtlich der Ergebnisse gemacht werden. Da nicht alle Urteile zur zweiten Plenarphase diskutiert werden konnten, leidet die Qualität der Ergebnisse unter der hohen Zahl an Enthaltungen, so dass nicht für alle abgefragten Abschnitte (und damit Forschungsinteressen) geeignete Maßnahmen identifiziert werden konnten.

Im Fallbeispiel des Gruppendelphis zu **Smart Metering** gibt es bei den 10er-skalierten Items ebenfalls Enthaltungen, die eine Überprüfung in der Dimension der Evidenz letztlich bei zwei Items erlaubt haben, da von den ursprünglich 38 abgefragten Items im ersten Fragebogen des Workshoptages lediglich vier Items auch im zweiten Fragebogen abgefragt wurden. Die darin enthaltenen zwei Items hatten allerdings jeweils Enthaltungen in der ersten Runde aufzuweisen, so dass der Wert der Standardabweichung trotz höherer Fallzahl zur zweiten Runde sank und letztlich ebenso einen Konsens markiert. Die beiden übrigen Items, die für einen Vergleich zur Verfügung stehen, bildeten überdies eine Entwicklung zu einem konsensualen Ergebnis ab, dass in Verbindung mit dem argumentativen Austausch während der Plenarphasen sowie der Integration eines neuen ergänzenden Items zu einem konsensualen Ergebnis geführt hat. Neben den 10er-skalierten Items wurden auch zwei prozentuale Ratingfragen eingesetzt, die bereits einen Konsens in der ersten Runde markierten und durch weitere Argumentation in einen Möglichkeitsraum für eine potentielle Einsparung eingegrenzt werden konnten. Im Falle des Gruppendelphis zu Smart Metering fand keine Vorbefragung statt, da das Gruppendelphi den methodischen Teil eines Testdesigns bildete, der auf eine umfassende Literaturrecherche sowie Fokusgruppen mit potentiellen Nutzer*innen folgte. Trotz dieser Vorkenntnisse wählten die Teilnehmenden letztlich bei der Auswahlfrage keine der im ersten Fragebogen vorgeschlagenen Maßnahmen aus, sondern nutzten die offene Kategorie *Sonstige* um eigene Maßnahmen in den jeweiligen Kleingruppen zu formulieren und diese in der zweiten Runde konsensual auszuwählen. Hier hätte eine Vorbefragung bereits potentielle Maßnahmen identifizieren können, die von den Expert*innen priorisiert oder zumindest vorgeschlagen werden und so letztlich auch Zeit am eigentlichen Workshoptag eingespart werden können.

Das Gruppendelphi zu **Eco-Fahrassistenten** bildete ein Teilprojekt desselben Gesamtprojektes, indem auch das Gruppendelphi zu Smart Metering angesiedelt war. So war auch dieses Gruppendelphi Teil eines Testdesigns, dem eine Literaturrecherche und Fokusgruppen vorausgingen, so dass auch hier auf eine Vorbefragung der Teilnehmenden verzichtet wurde. In diesem Fallbeispiel bildet nicht die ungewöhnlich hohe Zahl an den von den Teilnehmenden vorgeschlagenen Maßnahmen eine Besonderheit, sondern die vom Forschungsteam ungewöhnlich hohe eingesetzte Anzahl von fünf offenen Fragen, die sich nach Analyse dieses Falles aufgrund ihrer Konstruktion und schwer vergleichbarer Ergebnisse als nicht für eine konsensuale Verständigung während eines Gruppendelphis geeignet darstellen. Überdies hat das Zeitfenster der Kleingruppenphasen nicht ausgereicht, dass die Gruppen diese hätten zufriedenstellend beantworten und ausfüllen können. Problematisch für einen auf Evidenz basierenden Diskurs ist es, wenn die Teilnehmenden ihrem Urteil verschiedene Bewertungshorizonte zugrunde legen. Dies ist durch den Wechsel von einer Ratingskala in der ersten Runde hin zu einer Rankingskala zur zweiten Runde in diesem Fallbeispiel geschehen. Hier verknüpfen sich die Dimensionen des gemeinsamen Verständnisses mit der Evidenz und mussten über einen langwierigen Austausch im zweiten Plenum aufgelöst werden, da eine Kleingruppe die Ran-

kingskala als Ratingskala behandelt hatte. Letztlich konnte hier nur ein argumentativer Austausch einen gemeinsamen Bewertungshorizont schaffen, auf dem letztlich mittels weiterer evidenzbasierter Argumente ein konsensuales Ergebnis erreicht werden konnte. Im Gruppendelphi zu Eco-Fahrassistenzsystemen standen für eine Überprüfung in der Dimension der Evidenz anhand der statistischen Werte der 10er-skalierten Items letztlich drei, der in der ersten Runde abgefragten 48 Items zur Verfügung. Von diesen drei Items bestätigten zwei der Items die im Ideal angenommene Entwicklung hin zu einem konsensualen Ergebnis. Das dritte Item zeigt nach zusätzlicher Betrachtung der einzelnen Antwortangaben eine Entwicklung zu einem Konsens über einen Dissens auf, bei dem aber auch dem Geltungsanspruch der Gültigkeit folgend immer die Argumente der jeweiligen Positionen integriert werden sollten, um von einem Einverständnis über die geteilten Einschätzungen auszugehen. Ansonsten würde ein Dissens als ein Konsens über diesen verkauft, der keine gemeinsam geteilte Grundlage hat.

Im Kontrast zu den beiden vorhergehenden Fallbeispielen bot letztlich das Gruppendelphi in **PROSUITE** mit 17 Items, die sowohl in der ersten, wie auch in der zweiten Runde abgefragt wurden, wieder einen umfassenderen Vergleich der statistischen Werte der 10er-skalierten Items an. Dabei wurde dieser Skalentyp im Falle von PROSUITE nahezu exklusiv abgefragt, da ansonsten nur eine offene Abfrage durch potentielle Nennung von zusätzlichen Items durch die Teilnehmenden Teil des Fragebogens der ersten Runde war. Diese zusätzlichen Items wurden im Fragebogen der zweiten Delphirunde in 10er-skalierte Items überführt. Die gesamt abgefragten Items bestätigen nach analytischer Betrachtung die angenommene Entwicklung hin zu einem konsensualen Ergebnis oder wenn dies nicht erreicht werden konnte, hin zu einem vermuteten Konsens über einen Dissens, der nach Betrachtung der Häufigkeitsverteilung und weiterer argumentativer Fundierung so auch bestätigt werden konnte, wenn ein entsprechender Austausch zum zweiten Plenum zu diesen stattgefunden hatte. Das Gruppendelphi in PROSUITE bietet allerdings eine Besonderheit, denn 14 dieser 17 Items wurden zur zweiten Runde des Gruppendelphis aufgeteilt, wobei die generelle Wichtigkeit des Items und ihr konkreter Impact auf die Technologie abgefragt wurden. Da nur die zweite Abfrage nach dem Impact für die Beantwortung der Fragestellung des Gruppendelphis relevant war, wurden die Ressourcen (besonders natürlich die Zeit) hinsichtlich der Zielsetzung des Gruppendelphis nicht effizient verwendet, da zwar brauchbare Ergebnisse erzielt wurden, aber nicht gänzlich mit Argumenten gestützt werden konnten, wohingegen weitere Erkenntnisse hinsichtlich der generellen Wichtigkeit erzielt wurden, die allerdings nicht der Fragestellung des Gruppendelphis dienlich waren.

Letztlich bietet die Analyse der Fallbeispiele in der Dimension der Evidenz allerdings anhand der empirischen Daten zwei Erkenntnisse hinsichtlich der Fragestellung dieser Arbeit. So ist durchaus eine **Bewegung zwischen den Urteilen der teilnehmenden Expert*innen** hin zu einem Konsens oder einem Konsens über einen Dissens zu beobachten, der anhand der statistischen Werte bereits nachvollzogen werden kann und darauf hindeutet, dass die Teilnehmenden von ihrem Gegenüber im Prozess auch überzeugen lassen und sich letztlich nicht bloß die stoischste, auf ihrem Urteil beruhende Persönlichkeit durchsetzt. Dies wird durch den Einbezug durch die Betrachtung des **argumentativen Austauschs** während der Plenarsitzungen zusätzlich gestützt und zeigt, dass auch die **Gültigkeit einzelner Argumente** während eines Gruppendelphis in Frage gestellt und mittels eines **diskursiven Austauschs zwischen den teilnehmenden Expert*innen** überprüft und letztlich zu einem **Konsens** geführt werden

kann. Zusätzlich kann auf diesem Wege ein bestehender **Konsens über einen Dissens** bestätigt werden. Argumente sind letztlich also „(...) *diejenigen Mittel, mit deren Hilfe die intersubjektive Anerkennung für den zunächst hypothetisch erhobenen Geltungsanspruch eines Proponenten herbeigeführt und damit Meinung in Wissen transformiert werden kann*“ (Habermas 1981: 48).

Wie bereits der Aufbau des Testdesigns vermittelt, geht es aber im Prozess eines Gruppendelphis nicht exklusiv darum, dem Geltungsanspruch der Gültigkeit zu entsprechen. Dies ist im Hinblick auf die jeweiligen Fragestellungen, die mit dem Einsatz der Methode verbunden werden, das prioritäre Ziel. Aufgrund der Diskursivität einzelner Elemente bietet ein Gruppendelphi aber auch die Chance und Notwendigkeit für die Teilnehmenden in einem argumentativen Austausch ein gemeinsames Verständnis hinsichtlich der verwendeten Begriffe sowie des Bewertungshorizontes einzelner Sachverhalte zu bilden. Denn um kommunikative Aussagen gemeinsam bewerten zu können, müssen diese Aussagen daraufhin überprüft werden, ob die Teilnehmenden mit dem abgefragten Aussagen auch den gleichen Bedeutungsinhalt verbinden. Die abgefragten Sachverhalte müssen damit sowohl verständlich, wie auch eindeutig sein. Ein kommunikativer Aussagetyt ist somit, dann für das Geltungskriterium der Verständlichkeit zulässig, wenn er dem Geltungsanspruch der Zweckmäßigkeit für den Diskurs genügen kann (vgl. Renn/Webler 1998: 50; Renn 1999: 72f.). Ob die empirischen Fälle dem Geltungskriterium der **Verständlichkeit** entsprechen und somit dem Geltungsanspruch der Zweckmäßigkeit für den Diskurs genügen konnten, wurde mithilfe eines Vergleichs der Fragebögen in Ergänzung mit den Entwicklungen der statistischen Kenngrößen überprüft. Auch hier wurden die zusätzlichen im Plenum getätigten Argumente der Teilnehmenden in die Auswertung mit einbezogen.

Das erste Fallbeispiel des Gruppendelphis in **SAUBER+** zeigte bereits, dass ein solcher Vergleich möglich ist, wenn eine Umformulierung zwischen dem Fragebogen der ersten Runde und des Bogens der zweiten Runde vorgenommen wurde. So finden sich bei diesem Fall vier Items, bei denen im ersten Plenum eine Umformulierung vorgenommen wurde. Diese Items waren als Aussagen formuliert und die Teilnehmenden äußerten im ersten Plenum, dass sie der Aussage unter der gegebenen Formulierung nicht zustimmen konnten. Hier zeigte sich, dass eine semantische Klärung hinsichtlich einzelner Begriffe oder hinsichtlich der Präzision der formulierten Aussagen aus Sicht der Teilnehmenden notwendig war, um diesen zustimmen zu können. Solche Änderungen wurden im Plenum konsensual entschieden und zeigten ihre Wirkung dadurch, dass den Aussagen, wie die Annahme ideal vermuten lässt, vor einem gemeinsam getragenen Bewertungshorizont geschlossen zugestimmt werden konnte. Das Gruppendelphi in SAUBER+ zeigt allerdings auch, dass die Teilnehmenden trotz konkret formulierter Aussagen unterschiedliche Schwerpunkte in der Interpretation der Aussagen legen oder gar einzelne Begriffe bzw. Formulierungen in ihrem Urteil nicht berücksichtigen. So können in den Aussagen einzelne Spezifizierungen enthalten sein, die von den Teilnehmenden in ihrem Urteil nicht einbezogen werden. Im Falle von SAUBER+ finden sich drei solcher Items, bei denen es in der ersten Runde zu unterschiedlichen Bewertungen der einzelnen Gruppen kam, weil einzelne Kleingruppen einzelne Aspekte der Aussage außeracht gelassen hatten. Durch den Austausch im Plenum konnte ein Verweis auf diese Aspekte erfolgen und nachdem sie zusätzlich optisch im Fragebogen der zweiten Runde hervorgehoben wurden, führte auch diese Änderung zu einer gemeinsamen Bewertung des identischen Sachverhaltes. Ebenso

findet sich in SAUBER+ ein Beispiel für eine notwendige Präzisierung des einleitenden Fragetextes, da hier in der ersten Runde ein unterschiedliches Verständnis der Kleingruppen zum geographischen Fokus sowie zur nicht-exklusiven technischen Ausrichtung der Bewertung einzelner Items bei dieser Frage vorlag. Dies konnte durch einen argumentativen Austausch im Plenum und einer darauf basierenden Ergänzung von Hinweisen im Fragetext zur zweiten Runde aufgelöst werden. So kann auch die Ergänzung eines einzelnen Begriffes einen Dissens hinsichtlich einzelner Maßnahmen auflösen sowie deren Reichweite näher bestimmen und ein zur zweiten Runde in die Formulierung eingefügtes *bedarfsorientiert* letztlich eine konsensuale Bewertung befördern.

Das Gruppendelphi in **SEAWIND** bot im Gegensatz zur idealtypischen Konzeption eine Besonderheit hinsichtlich der Itemformulierung. So wurden diese hier nicht als formulierte Aussagen, sondern als direkte Nennung der Sachverhalte abgefragt. In diesem Fallbeispiel fiel auf, dass die Teilnehmenden sowohl Verständigungsbedarf hinsichtlich der Formulierung einzelner Items hatten, wie auch besonders an den einleitenden Fragetexten zu den Itembatterien. So wurde bereits zu Beginn des ersten Plenums eine Umformulierung des Fragetextes der ersten Frage beschlossen, da ein Teil der Teilnehmenden die abgefragten Akteursgruppen danach beurteilte, welche geeignete Kommunikationsakteure und andere Teilnehmende, welche die am besten geeigneten Akteure für eine solche Kommunikation seien. Die gemeinsam beschlossene Umformulierung zur zweiten Runde des Gruppendelphis, die am besten geeigneten Akteure zu bewerten, konnte letztlich einen Konsens über einen Dissens markieren. Neben dieser Umformulierung fanden bei vier weiteren Fragen aufgrund des Austauschs im ersten Plenum Umformulierungen der einleitenden Fragetexte statt. Ebenso wurden zusätzlich weitere Items diskutiert, die nach Äußerungen der Teilnehmenden im ersten Plenum im Fragebogen der zweiten Runde in ihrer Formulierung geändert wurden. Die Teilnehmenden begründeten die Umformulierungen im SEAWIND-Gruppendelphi allerdings weniger mit semantischen Missverständnissen, sondern führten das Argument an, dass die Fragen nicht klar formuliert gewesen wären. Da die Teilnehmenden einzelne Fragepassagen allerdings in der ersten Kleingruppenphase bewusst nach ihrer individuellen Prioritätensetzung bereits hinsichtlich der jeweiligen Bewertung umgedeutet hatten, ist hier auch ein taktisches Handeln einzelner Teilnehmender nicht auszuschließen. Überdies lassen sich in diesem Fall versteckte Argumentationsmuster einzelner Teilnehmender ebenso nicht ausschließen, die im Falle der häufigen Umformulierungen thematisch umgelenkt bzw. vorgeschoben wurden. Da die Fragen sowie Items in der Vorbefragung in identischer Formulierung abgefragt wurden, wie sie auch im ersten Workshopfragebogen formuliert waren, drängt sich dieser Verdacht auf, da in dieser Vorbefragung keine Änderungswünsche und Anmerkungen durch die Befragungsteilnehmenden geäußert wurden.

Im Fallbeispiel des Gruppendelphis zu **Smart Metering** kam es im Gegensatz zu den vorhergehenden Fällen zu keiner Umformulierung, weder hinsichtlich eines Items, noch hinsichtlich einer Frageformulierung. Wenn zusätzliche Spezifikationen und andere Bewertungshorizonte in der ersten Runde als notwendig erachtet wurden, so kam es zu einer Integration einer zusätzlichen Frage, da die Teilnehmenden die Notwendigkeit sahen, die Akteursgruppen weiter zu spezifizieren. Hier wurde eine offene Frage zur Nennung der spezifischen Akteure integriert. Neben dieser neuen Frage wurde auch die Rahmung des Gruppendelphis diskutiert, da die Teilnehmenden bereits im ersten Plenum der Ansicht waren, dass Smart Metering in einem

breiten Kontext aus Energieeinsparung und Energiewende betrachtet werden müsse. So wurde dem Fragebogen der zweiten Runde eine offene Frage vorausgestellt, in der die Teilnehmenden eine solche Rahmung, der im Gruppendelphi abgefragten Urteile und Bewertungen, als Präambel formulieren konnten. Da bereits im ersten Plenum ein Austausch der Argumente, die die Teilnehmenden konsensual teilten, stattfand, unterschieden sich die Präambeln zwar in der konkreten Formulierung aber nicht in den von den Expert*innen geteilten Inhalten, so dass die von den vier Kleingruppen formulierten Präambeln jeweils von den anderen Kleingruppen als geeignete Rahmung des Gruppendelphis eingeschätzt wurden und auf diesem Wege dem Gruppendelphi ex post ein gemeinsamer Bewertungshorizont gegeben wurde.

Das Gruppendelphi zu **Eco-Fahrassistenten** ist in der Häufigkeit der getätigten Umformulierungen zwischen dem Fragebogen der ersten und dem der zweiten Runde vergleichbar mit dem Gruppendelphi in SAUBER+. So kam es zu einer Umformulierung einer einleitenden Frage sowie einer Umformulierung von zwei weiteren Items. Auch hier wurde mit einer optischen Hervorhebung von geänderten und relevanten Inhalten innerhalb der Fragedarstellung gearbeitet. Die Umformulierung eines Items bietet allerdings weitere Erkenntnisse, die in den vorhergehenden Fallbeispielen nicht zu beobachten waren. So wurde ein Item zu *Basisinformationen zu spritsparenden Fahrstilaspekten* durch die Umformulierung von *ständig* auf *dezen*te Rückmeldung umformuliert und damit auch die thematische Ausrichtung des Items geändert, indem eine Abschwächung der Aussage vorgenommen wurde, die auch den Bewertungshorizont der Teilnehmenden veränderte. Dieser veränderte Bewertungshorizont wurde von diesen im zweiten Plenum argumentativ thematisiert, da durch die Verschiebung eine klare Ablehnung einer Kleingruppe die Folge war. Letztlich führte diese Ablehnung zu einem Konsens über einen Dissens, da folgende Argumente zwei gegenläufige Positionen identifizieren konnten: Akzeptanz versus ständige Information.

Im Gruppendelphi in **PROSUITE** wurden drei Items zwischen der ersten und der zweiten Runde umformuliert. Von diesen drei Items konnte ein Item mittels eines argumentativen Austausch im zweiten Plenum nach der Umformulierung in einem Konsens über einen Dissens aufgelöst werden und ein weiteres der umformulierten Items zeigte bereits anhand der statistischen Kenngrößen zur zweiten Runde einen Konsens auf. Bei einem dieser umformulierten Items verteilten sich die Antworten der einzelnen Kleingruppen allerdings so, dass auch nach der Umformulierung kein eindeutiges Urteil als Ergebnis ableitbar war. Neben diesen unterschiedlichen Entwicklungen der Items nach der Umformulierung im Fragebogen der zweiten Delphirunde, bietet das Fallbeispiel des PROSUITE-Gruppendelphis zudem eine Besonderheit. So wurden auch bei diesem Gruppendelphi keine formulierten Aussagen als Items abgefragt, sondern die aus der vorhergehenden Recherche ermittelten und übersetzten Indikatoren direkt verwendet. Es fand damit eine direkte Abfrage statt, wie sie auch im Fall des SEAWIND-Gruppendelphis stattgefunden hatte. Die Besonderheit des Gruppendelphis in PROSUITE bildet allerdings die Aufteilung von insgesamt 14, der in der zweiten Runde abgefragten 17 Items in die generelle Wichtigkeit und den jeweiligen Impact auf die Technologie CCS. Dies ist im Hinblick auf den Geltungsanspruch der Verständlichkeit insofern besonders relevant, da in der ersten Abfrage zur ersten Runde im Fragebogen eindeutig und wiederholend der Bewertungshorizont des Impacts auf CCS betont wurde. Auch wenn hier zur zweiten Runde Umformulierungen durch die Teilung der Items stattfanden, so hat dies nicht auf einem fehlenden Ver-

ständnis gegenüber der einzelnen Formulierungen basiert, sondern auf einer Prioritätensetzung der Teilnehmenden hinsichtlich des Bewertungshorizontes einzelner Items. Dies führt zu einem Effekt, der über die fünf Fälle des Gruppendelphis zu beobachten ist.

So zeigen die empirischen Fälle der Gruppendelphis spezifische Besonderheiten, in den im Plenum geäußerten Argumenten sowie den getätigten Umformulierungen, die besonders auf das geteilte Verständnis eines **gemeinsamen Bewertungshorizontes** hinauslaufen. Damit findet eine Verknüpfung der **Dimension der Verständlichkeit** mit der **Dimension der Evidenz** statt, da ein geteilter Hintergrund die Grundlage für die Gültigkeit von Ergebnissen bildet, die sich auf den Geltungsanspruch der Evidenz berufen und letztlich zudem individuelle Präferenzen der Teilnehmenden einschließen. Überdies bietet das Gruppendelphi, wie in den Vorüberlegungen bereits angenommen, einen Raum dafür, einen gemeinsamen Bedeutungsinhalt für die strittigen **Begriffsdefinitionen, Klassifikationsschemata** und **Formulierungen von Aussagen** zu erzielen. So können interdisziplinär unterschiedlich konzeptuierte Fachtermini, semantische Missverständnisse oder durch das Forscherteam nicht klar formulierte Aussagen im Fragebogen zu unterschiedlichen Interpretationen durch die Teilnehmenden führen, die letztlich mittels einer konsensual getroffenen Umformulierung im folgenden Fragebogen aufgelöst werden können. Aber gerade die in allen empirischen Fällen zu beobachtende Suche nach einem gemeinsamen Bewertungshorizont bildet letztlich auch die Suche nach kommunikativer Rationalität ab, indem die Teilnehmenden „(...) *sich dank der Gemeinsamkeit vernünftig motivierter Überzeugungen gleichzeitig der Einheit der objektiven Welt und der Intersubjektivität ihres Lebenszusammenhangs vergewissern*“ (Habermas 1981: 28).

Um eine umfassende Bewertung von Handlungsoptionen vornehmen zu können, müssen auch Werte und normative Einstellungen in die angenommenen Konsequenzen und wahrscheinlichen Implikationen einzelner Maßnahmen mit einfließen können (vgl. Failing et al. 2004: 16). Das Gruppendelphi muss damit in den diskursiven Phasen einen Rahmen bereitstellen, in dem die Teilnehmenden die Möglichkeit haben auch ihre Wertvorstellungen und normativen Einstellungen mittels Argumentation einzubringen. Um aber dem Geltungskriterium der **normativen Angemessenheit** zu entsprechen, müssen die evaluativen Aussagen den Kriterien der Kohärenz und der Verallgemeinerungsfähigkeit sowie die normativen Aussagen den Kriterien des Gebots der Universalisierung und der Kompatibilität mit Gesetzen entsprechen (vgl. Renn 1999: 72). So kann auch während eines Gruppendelphis eine Argumentation erfolgen, die den Geltungsansprüchen des Gemeinwohls und der normativen Richtigkeit folgt. Für die Betrachtung der empirischen Fälle ergaben sich damit mehrere Fokusse, die logisch aufeinander aufbauten. So wurde in den dokumentierten diskursiven Phasen eine Kategorisierung der Aussagen vorgenommen, die sich in ihrer Äußerung direkt auf den Geltungsanspruch nach Gemeinwohl oder normativer Richtigkeit bezogen. Ob eine solche Argumentation von den Teilnehmenden als legitimes Argument angenommen wurde, wurde dahingehend geprüft, ob einzelnen Personen von anderen Teilnehmenden eine wertgeleitete Argumentation als nicht legitimes Argument unterstellt wurde oder ob sich anknüpfende Erwiderungen auf den geäußerten Geltungsanspruch bezogen bzw. sogar auf diesem aufbauten. Letztlich kann so herausgestellt werden, dass wenn diese Argumente im Diskurs als gleichwertige Argumentationen anerkannt werden, diese letztlich auch für den Austausch der Expert*innen während des konkreten Falles eines Gruppendelphis als legitim erachtet wurden.

Im Gruppendelphi zu **SAUBER+** standen für die Inhaltsanalyse in der Dimension der normativen Angemessenheit die Protokolle der beiden durchgeführten Plenarsitzungen und ein Ergebnisprotokoll des Gruppendelphis als inhaltliche Datenquellen zur Verfügung. Das Fallbeispiel bietet dabei von den Teilnehmenden geäußerte und in der Analyse kategorisierte Aussagen, die sowohl Äußerungen von individuellen Präferenzen enthalten sowie auch Aussagen, die den Geltungsanspruch des Gemeinwohls oder den Geltungsanspruch an normative Richtigkeit in Anspruch nehmen. Bei den Äußerungen der Präferenz einer Bewertungsposition konnte diese in ein konsensuales Bewertungsmuster für die zweite Runde des Gruppendelphis überführt werden, so dass die Teilnehmenden in der zweiten Runde, dem von einer Person geäußerten Leitsatz folgten, *nicht kleckern, sondern klotzen*. Ebenso führten geäußerte normative Vorstellungen oder Wertvorstellungen dazu, dass diese im Diskurs als legitimes Argument angenommen wurden und nach Erwiderung bzw. Anschluss zu einem Konsens geführt bzw. einen eindeutigen Konsens über einen Dissens markiert haben, der im Fall dieses Gruppendelphis besonders auf zukünftigen Forschungsbedarf begründet werden konnte. Ob es an der Besonderheit lag, dass die Teilnehmenden in dem Gruppendelphi zu SAUBER+ ihre Rolle selbst im Verständnis von Stakeholdern in einer Reihe von Stakeholder-Workshops sahen und damit eine normative bzw. wertgeleitete Argumentation nicht als unzulässig angesehen wurde, kann im Kontrast zu den folgenden Fällen ausgearbeitet werden, da in diesen Fallbeispielen die Teilnehmenden im Sinne der klassischen Teilnehmendenbezeichnung vorwiegend als Expert*innen angesprochen wurden.

Für diese Analyse des Gruppendelphis in **SEAWIND** standen das Ergebnisprotokoll des Gruppendelphis sowie die Veröffentlichung der Ergebnisse des Gruppendelphis in Hilpert et al. (vgl. ebd. 2013: 11-32) zur Verfügung. Dabei zeigte sich in diesem Fall, dass bei diesem Gruppendelphi kaum Argumente geäußert wurden, die den Geltungsanspruch nach normativer Angemessenheit und Gemeinwohl adressierten. Wenn an solche Äußerungen direkte Anknüpfungen erfolgen, so wurden diese mit Bezug zur Evidenz gekoppelt, indem ein direkter Bezug zu dem ursprünglichen Argument widersprechenden wissenschaftlichen Daten und Studien herausgestellt wurde. Nach Kategorisierung der Aussagen fiel zudem auf, dass die einzelnen Items weniger Spielraum für einen Diskurs boten, als es beispielsweise im Gruppendelphi in SAUBER+ der Fall war, da sie nicht als Aussagen formuliert waren, sondern eine direkte Abfrage bildeten. Auch zeigt sich in Betrachtung des Diskurses eine Dominanz einzelner Teilnehmenden, die sich nur teilweise auf den Diskurs einließen und ihrer eigenen Einschätzung widersprechende Argumente letztlich ohne weiteren Bezug als *falsch* bezeichneten. Dies ist natürlich höchst problematisch bei einem argumentativen Austausch und zeigt sich letztlich auch in einem illustrativen Beispiel dieses Falles. So wurden zwei Positionen in SEAWIND identifiziert, die als normative Leitsätze für die Kommunikation von Strahlungsrisiken fungieren können: Erstens eine reaktive Kommunikation der Einschätzung der IARC (*right to know*) und zweitens eine unbedingt aktive Kommunikation der Einschätzung der IARC (*need to know*). Die Teilnehmenden des Gruppendelphis erkannten zwar die jeweils andere Position als legitimen Geltungsanspruch an, konnten sich aber hier letztlich nicht auf einen Konsens über Dissens einigen, da eine Person dafür plädierte eine Abstimmung durchzuführen, bei der letztlich eine Majorität und eine Minorität abgebildet wurden. Die Positionen standen damit in einem hierarchischen Verhältnis zueinander, wie es in einem Gruppendelphi als Ergebnis nicht angelegt und im Sinne eines konsensbildenden Austauschs kontraproduktiv ist. Somit bildet der Fall SEAWIND einen Kontrast zum Fall SAUBER+. Ob dies ebenso auf die Frage hinsichtlich

des Rollenverständnisses der Teilnehmenden als Expert*innen statt als Stakeholder zurückzuführen ist, kann mit dem nächsten Fall weiterverfolgt werden.

Für die Inhaltsanalyse des Gruppendelphis zu **Smart Metering** standen sowohl ein Protokoll beider Plenarsitzungen sowie das Ergebnisprotokoll als Datenquellen zur Verfügung. Überdies konnten auch die von den Teilnehmenden in der zweiten Kleingruppenphase im Fragebogen formulierten Präambeln in die Analyse integriert werden. Anderes als es in SEAWIND der Fall war, finden sich unter den teilnehmenden Expert*innen im Gruppendelphi zu Smart Metern in beiden Plenarsitzungen Äußerungen, die in ihrem Argumentationsmuster den Geltungsanspruch nach Gemeinwohl und normativer Richtigkeit direkt adressiert haben. Ebenfalls wurden solche Argumente gegenüber Argumenten, die den Geltungsanspruch der Gültigkeit adressierten, von den Teilnehmenden als gleichwertig angenommen und nicht wie im Falle von SEAWIND als falsch bezeichnet. So konnte in diesem empirischen Fall überdies ein Konsens darüber erzielt werden, bei einzelnen Fragen sowohl die Perspektive des wahrscheinlichen *Ist*-Zustandes sowie eine Verteilung nach einem wünschenswerten *Soll*-Zustandes einzunehmen. So einigten sich die Teilnehmenden im argumentativen Austausch während des Plenums letztlich darauf, dass das Budget zum einen danach verteilt werden sollte, wie es verteilt sein *sollte* und zum zweiten, dass Budget so zu verteilen, wie es unter den gegebenen Rahmenbedingungen verteilt werden *wird*. So konnte ein bestehender Dissens zwischen den Einschätzungen der Expert*innen letztlich durch gegenseitige Einnahme beider Perspektiven in einen Konsens aufgelöst werden. Die Integration der Präambel stellt auch im Fokus der Dimension der normativen Angemessenheit eine Besonderheit dar. So einigten sich die Teilnehmenden im ersten Plenum darauf dem Gruppendelphi und den mit der Methode erzielten Ergebnissen eine thematisch wie normative Rahmung zu geben. Im zweiten Plenum fand eine konsensuale Verständigung hinsichtlich dieser Rahmung statt, so dass den Ergebnissen des Gruppendelphis damit ex post ein normativer Rahmen gegeben wurde, in dem die Ergebnisse zu verstehen seien.

Der Fall des Gruppendelphi zu **Eco-Fahrassistenten** bot für die Inhaltsanalyse der Dimension der normativen Angemessenheit die umfassendste Dokumentation, der in dieser Arbeit betrachteten Fälle. So wurden in beiden Plenardiskussionen die einzelnen Redebeiträge mit Unterstützung von Audioaufnahmen als Transkript im Protokoll in anonymisierter Form aufgenommen sowie ein umfassendes Ergebnisprotokoll erstellt. So konnten in diesem Fall nicht nur die einzelnen aufeinander folgenden Argumente in die Analyse integriert, sondern die Aussagen konnten auch den jeweiligen sie äßernden Personen zugeteilt werden. Im Gruppendelphi zu Eco-Fahrassistenten finden sich viele argumentative Äußerungen, die entweder den Geltungsanspruch nach normativer Angemessenheit oder dem nach Gemeinwohl adressieren, allerdings auch häufig als subjektive Bewertungen von Optionen oder als persönliche Präferenzen geäußert wurden. Trotz dieses individuellen Bezugs wurden auch solche Äußerungen während des diskursiven Austauschs in den Plenarsitzungen mittels anknüpfender Äußerungen erwidert oder bekräftigt, so dass diese Kategorien von Argumenten in diesem Fallbeispiel letztlich auch als legitime Argumente im Diskurs anerkannt wurden. Keiner Person wurde von den anderen Teilnehmenden eine wertgeleitete Argumentation unterstellt, die als unzulässig wahrgenommen wurde. Die hohe Zahl an kategorisierten individuellen Bewertungen und

Präferenzen ergibt sich allerdings auch aus der umfassenden Protokollierung dieses empirischen Falles, da solche Äußerungen in einem Ergebnisprotokoll, das die Argumentationsstränge hin zu einem konsensualen Ergebnis abbildet, häufig nicht gänzlich enthalten sind.

Das Fallbeispiel des Gruppendelphis in **PROSUITE** bietet als Datengrundlage ein knappes, tabellarisches Protokoll. Weil nach diesem Gruppendelphi kein Ergebnisprotokoll an die Teilnehmenden versandt wurde, war dies für diesen Fall die exklusive Datengrundlage für eine inhaltsanalytische Betrachtung. Da im ersten Plenum bereits zu Beginn der Vorschlag geäußert wurde, die Items bei Notwendigkeit in Impact und Wichtigkeit aufzuteilen, dominierte diese konsensual beschlossene optionale Teilung den Austausch im ersten Plenum. Auch in der zweiten Plenarphase knüpften die Aussagen häufig an diese Teilung der Items an, so dass auch hier keine Argumente dokumentiert sind, die Werturteile, Präferenzen, individuelle Bewertungen von Optionen oder moralische Urteile enthalten. Äußerungen, die den Geltungsanspruch nach normativer Angemessenheit oder dem nach Gemeinwohl adressieren, finden sich im Falle von PROSUITE lediglich in den Äußerungen des persönlichen Fazits der Teilnehmenden am Ende des Gruppendelphis. Da hier allerdings kein konsensuales Ergebnis forciert wurde, fand kein Bezug der jeweilig geäußerten Redebeiträge zueinander statt, so dass für das Gruppendelphi in PROSUITE keine tiefergehende Analyse der Dimension der normativen Angemessenheit erfolgen konnte.

So zeigt sich nach Betrachtung der einzelnen Fälle ein Muster, dass nicht auf der Bezeichnung der Teilnehmenden entweder als Expert*innen oder als Stakeholder aufbaut, sondern an der Form der Itemabfrage auszumachen ist. So bilden die beiden Fälle SEAWIND und PROSUITE Beispiele, bei denen Argumente, die den Geltungsanspruch nach **normativer Angemessenheit** kaum bzw. keine **legitime Rolle im diskursiven Austausch** der Teilnehmenden gespielt haben. Beiden Gruppendelphis ist gemein, dass hier keine Aussagen formuliert wurden, die eine *Sollens*-Komponente in ihrer Formulierung beinhalteten, sondern die Items sowohl in SEAWIND als auch in PROSUITE per direkter Abfrage des zu bewertenden Sachverhaltes erfolgten. Im Gegensatz dazu enthielten die Fragebögen der drei Gruppendelphis in SAUBER+, zu Smart Metering und zu Eco-Fahrassistenten größtenteils Items, die als **Aussage formuliert** waren und einen Möglichkeitsraum des *wie etwas sein sollte*, also eine **Sollens-Komponente** enthielten. Letztlich boten solche Aussagen nicht nur einen **Möglichkeitsraum für einen diskursiven Austausch** im Hinblick auf einen auf Entwicklung von Wissen und Verständnis ausgelegten Diskurs, sondern beförderten auch einen weitergehenden Austausch zu den Werteeinstellungen und normativen Vorstellungen der Teilnehmenden, wenn diese eine Grundlage für das jeweilige subjektive Urteil bildeten. Dies ist insofern für einen diskursiven Austausch innerhalb des Gruppendelphis entscheidend, da auch solche Argumente offengelegt werden dürfen und müssen, wenn letztlich ein **konsensual getragenes Ergebnis** erzielt werden soll. So muss ein diskursiver Rahmen existieren, in dem die Teilnehmenden ihre normativen Einstellungen und Wertvorstellungen offenlegen können, um auch nach rationalen Gesichtspunkten qualitativ hochwertige Ergebnisse erzielen zu können. *„Wer sich in seinen Einstellungen und Bewertungen so privatistisch verhält, daß sie durch Appelle an Wertstandards nicht erklärt und plausibel gemacht werden können, der verhält sich nicht rational“* (Habermas 1981: 37).

Unter der **grundlegenden Funktionslogik** können im Vergleich der einzelnen Fallbeispiele einige Bereiche identifiziert werden, die für den Prozess der Methode des Gruppendelphis hin

zu konsensual getragenen Ergebnissen als Stolpersteine oder wichtige zu erfüllende Bedingungen herausgestellt werden können.

Als erstes soll an dieser Stelle der **Variationskoeffizient** als Bemessungsgröße eines konsensualen Ergebnisses diskutiert werden. So wird bei einem Gruppendelphi davon ausgegangen, dass Items, die mit einem Wert von unter $v=0,5$ bewertet wurden, als konsensuales Ergebnis eingestuft werden können. Dieser empirische Richtwert - verschärft ausgedrückt auch als Grenzwert eines Konsenses - basiert auf den Erfahrungen von vergangenen Anwendungen der Methode des Gruppendelphis (vgl. Kuhn et al. 2014: 72f.). Da durch die 10er-skalierten Items eine quasi-metrische Skalierung erreicht werden kann, lässt sich der Variationskoeffizient neben dem arithmetischen Mittel und der Standardabweichung in die Interpretation der Ergebnisse mit einbeziehen. Dabei bilden die Standardabweichung und das arithmetische Mittel ein übliches Streuungsmaß ab, während der Variationskoeffizient als relatives Streuungsmaß interpretiert werden kann, der als Quotient aus Standardabweichung und arithmetischem Mittel gebildet wird. Der Variationskoeffizient bildet damit ein Verhältnis des arithmetischen Mittels zur Häufigkeitsverteilung ab und bietet folgend den Vorteil, dass dieser Wert auch für sich genommen interpretiert werden kann, womit der Variationskoeffizient als ein von \bar{x} bereinigtes Streuungsmaß verstanden werden kann (vgl. Hartung 1989: 47; Benninghaus 2002: 62; Diekmann 2009: 678ff.). Nun birgt eine solche exklusive Betrachtung des Wertes des Variationskoeffizienten innerhalb der Methode des Gruppendelphis allerdings einige Besonderheiten.

So ist bei einem Konsens über einen Dissens von hohen Werten der Standardabweichung auszugehen. Da der Wert der Standardabweichung den Dividenden des Variationskoeffizienten bildet, ergibt sich folglich auch ein höherer Wert des Variationskoeffizienten. Da gerade ein Konsens über einen Dissens allerdings auch im diskursiven Austausch mittels Argumentation gestützt werden sollte, ergibt sich diesbezüglich kein Problem hinsichtlich des Richtwertes von $v=0,5$. Bei der Interpretation der Ergebnisse sollte für die Zuschreibung eines Konsenses über einen Dissens allerdings zusätzlich die Häufigkeitsverteilung der einzelnen Antworten berücksichtigt werden, um einen nicht konsensualen Dissens auszuschließen. Bei einem niedrigen Wert des arithmetischen Mittels ergibt sich ein weiteres Problem bei exklusiver Betrachtung des Variationskoeffizienten. So kann es vorkommen, dass der Wert des Variationskoeffizienten über dem Richtwert von $v=0,5$ liegt, obwohl die Kleingruppen sich in ihrem Urteil einig sind, allerdings verschiedene niedrige Werte auf der 10er-Skala gewählt haben. So würde bei exklusiver Betrachtung des Variationskoeffizienten von einem Dissens ausgegangen, der allerdings nicht vorläge. Da diese Fehlinterpretation während eines Austauschs im Plenum leicht aufgedeckt werden kann, wird diese Abhängigkeit von arithmetischem Mittel und Wert des Variationskoeffizienten im Umkehrschluss zu einem Problem, wenn der Wert des arithmetischen Mittels groß ist. So kann auch bei hohem Wert der Standardabweichung im ungünstigen Fall bei zeitgleich hohem Wert des arithmetischen Mittels ein Wert des Variationskoeffizienten unter $v=0,5$ bestehen, der bei exklusiver Auslegung als konsensuales Ergebnis interpretiert würde, dies aber anhand der Verteilung der einzelnen Antworten der jeweiligen Kleingruppen nicht widerspiegelt. So besteht bei der Interpretation des Variationskoeffizienten die Tendenz, dass hohe Werte auf der 10-er Skala eher als konsensuales Ergebnis abgebildet werden. Des-

halb sollte bei der Auswertung (insbesondere im engen Zeitrahmen während eines Gruppendelphis) immer auch ein Einbezug der Werte der Standardabweichung erfolgen und im Zweifel die einzelnen Antworten der Kleingruppen direkt verglichen werden.

Letztlich besteht noch eine weitere Gefahr bei exklusiver Auslegung des Wertes des Variationskoeffizienten. Wenn die Fallzahl der Antworten (n) kleiner als die Menge der gebildeten Kleingruppen (N) ist, dann kann der Wert des Variationskoeffizienten ein konsensuales Ergebnis markieren, dass aber nicht auf den Antworten aller beteiligten Kleingruppen basiert. Eine Enthaltung kann allerdings nicht als Zustimmung zum Ergebnis der anderen Kleingruppen gezählt werden, bevor dieser Sachverhalt im Plenum besprochen wurde. Liegt der Wert des Variationskoeffizienten aber unter $v=0,5$ und die Fallzahl wird nicht beachtet, geschieht in der Konzeption eines Gruppendelphis genau das. Also muss ebenso immer die Fallzahl bei der Betrachtung des Variationskoeffizienten mit einbezogen werden. Bei einem Minderheitsvotum besteht diese Gefahr nicht, so dass der Wert des Variationskoeffizienten nur dann als Richtwert interpretiert werden kann, wenn gilt: $n \geq N$.

Enthaltungen stellen allerdings nicht nur im Hinblick auf die Auslegung des Wertes des Variationskoeffizienten ein Problem dar. Das Ziel des Gruppendelphis ist eine konsensual getragene Bewertung von Maßnahmen und Sachverhalten durch alle Teilnehmenden zu erhalten. Enthält sich eine oder gar mehrere Kleingruppen einem Votum hinsichtlich einzelner abgefragter Items, so können hier keine zufriedenstellenden Ergebnisse erzielt werden. Die zwei Ursachen solcher Enthaltungen konnten in den Fallbeispielen dieser Arbeit identifiziert werden. Zum einen zeigte das Beispiel des Gruppendelphis in SEAWIND, dass die Expert*innen bei Uneinigkeit in ihrer Kleingruppe den Entschluss fassten, das Item nicht zu bewerten, anstatt Minderheitsvoten abzugeben. Das Beispiel in SEAWIND ist dabei das einzige Gruppendelphi bei dem zu Beginn des Workshoptages nicht darauf hingewiesen wurde, dass die Kleingruppen bei Uneinigkeit die Möglichkeit haben, ein Minderheitsvotum im Fragebogen zu vermerken. So konnte in den übrigen Fällen verhindert werden, dass Minderheitsvoten aufgrund von Uneinigkeit auftreten. Hier blieben die Bewertungen nur aus, wenn die Teilnehmenden das Item nicht bewerten konnten, weil ihnen beispielsweise das Verständnis hinsichtlich der Aussage fehlte oder aber die Zeit zu knapp war. Der Faktor Zeit spielt hinsichtlich der Enthaltungen eine weitere wichtige Rolle. So sind gerade in den Fällen der Gruppendelphis zu Smart Metering und Eco-Fahrassistenten, aber auch in SAUBER+ einige Items nicht bewertet worden, weil am Ende der Kleingruppenphase die Zeit nicht mehr ausgereicht hat alle Items in der Kleingruppe zu besprechen und sich auf ein Urteil zu verständigen. Hier handelt es sich bei strenger Auslegung des Begriffs nicht um Enthaltungen, allerdings haben die **fehlenden Antworten** hier den gleichen negativen Effekt auf die Interpretation der Ergebnisse, wie es auch die Enthaltungen haben.

Dem Faktor **Zeit** kann dabei nur mit einem angemessenen Fragebogen begegnet werden. Ist der Fragebogen zu umfassend, können auch Verweise an die Kleingruppen auf die verbleibende Zeit nur bedingt fehlenden Beantwortungen verhindern. Alle fünf empirischen Fallbeispiele zeigen auf, das Gruppendelphis häufig zeitlich knapp bemessen sind, so dass als Abbruchkriterium weniger eine Stabilität der Urteile als eher die Zeit greift. Besonders eintägige Gruppendelphis leiden unter dieser zeitlichen Beschränkung, allerdings ist dies meist die einzige Möglichkeit eine genügend große Anzahl an Expert*innen zu rekrutieren, da eine Rekrutierung für eine zweitägige Veranstaltung ungleich schwerer ist. Ebenso kann ein solcher

Workshop nur bedingt flexibel sein, da eine Agenda für die Teilnehmenden schon bezüglich der persönlichen Zeit- und Reiseplanung unerlässlich ist. Gerade bei national oder international anreisenden Expert*innen kann nicht ein flexibles Ende bis hin zum konsensualen Ergebnis angesetzt werden. Das im Ideal angenommene natürliche Ende eines Diskurses ist damit nicht nur im Hinblick auf die Entscheidung, wann alle Argumente eingebracht worden sind, schwer zu erreichen, sondern bereits alleine aufgrund der Zeit beschränkt, so dass das Ideal nicht gänzlich erreicht werden kann. *„Faktisch beenden wir, unter günstigen Bedingungen, eine Argumentation erst dann, wenn sich die Gründe im Horizont bisher unproblematisch gebliebener Hintergrundannahmen zu einem kohärenten Ganzen soweit verdichten, daß über die Akzeptabilität des strittigen Geltungsanspruchs ein zwangloses Einverständnis zusammen kommt“* (Habermas 1992: 278).

Die empirischen Beispiele zeigen allerdings, dass die fehlende Zeit zwar ein Problem darstellen kann, aber bei einer angemessenen Flexibilität im Rahmen der Agenda des Tages sowie durch Einsatz eines angemessenen Fragebogens fruchtbare Ergebnisse erzielt werden können. So wurde, bis auf den Fall des Gruppendelphis in PROSUITE, in den übrigen Fällen nach dem Workshop jeweils ein **Ergebnisprotokoll** an die Teilnehmenden versandt, dem in keinem der Fälle widersprochen wurde. Es kann also davon ausgegangen werden, dass die Ergebnisse auch nach der Durchführung eines Gruppendelphis stabil bleiben und negative Effekte, wie beispielsweise ein Scheinkonsens aufgrund fehlender Motivation zur Argumentation oder fehlender Zeit ausgeschlossen werden kann. Wichtig sind in diesem Hinblick die Plenarphasen, die ausreichend Raum für einen umfassenden argumentativen Austausch hinsichtlich der jeweiligen Bewertungen ermöglichen. So muss letztlich eine sinnvolle Ausgewogenheit zwischen den statistischen Analysen und dem Austausch über die jeweiligen Positionen ermöglicht werden. *„Auch wenn der standardisierte Fragebogen statistische Analysen zulässt, die u. a. als Diskussionsinput für den Workshop dienen, geht es nicht um die Frage wie viele Experten die eine oder andere Meinung vertreten. Wichtig ist allein, dass die gesamte Bandbreite an Urteilen vertreten ist und die Argumente der verschiedenen Positionen transparent werden“* (Niederberger 2015b: 127).

Dadurch, dass die **Kleingruppen** während eines Gruppendelphis als ein geschützter Raum konzipiert sind, in dem die Teilnehmenden sich ohne den Druck einer Dokumentation offen austauschen können, kommt dem **Plenum** eine wichtige Funktion zu. Hier müssen die Argumente, die den jeweiligen Gruppenurteilen zu Grunde liegen im gesamten Teilnehmendenkreis transparent gemacht werden. Dabei geht es nicht darum alle genannten Argumente und Begründungen für die jeweiligen individuellen Positionen in ihrem Verlauf in den Kleingruppen auszutauschen, sondern letztlich die entscheidenden Argumente zu identifizieren, die für das Urteil der Kleingruppe ausschlaggebend waren. So ist bereits das Plenum ergebnisorientiert, was die Behandlung der jeweiligen Abstimmungsprozesse in den Kleingruppen betrifft und knüpft an diese Ergebnisorientierung an, indem in einem zweiten auf Konsensbildung ausgelegten Prozess die Argumente der jeweiligen Gruppen zu einem gemeinsam getragenen Urteil überführt werden sollen, welches in der folgenden Runde einer weiteren Überprüfung unterzogen wird. So spielt die **Dokumentation** der im Plenum geäußerten Argumente für die jeweiligen Einschätzungen eine entscheidende Rolle. Gerade bei einem Konsens über einen Dissens ist es für die Ergebnisqualität eines Gruppendelphis entscheidend beide Positionen mit den entsprechenden Argumenten zu dokumentieren, um den zukünftigen Forschungsbedarf

inhaltlich einzugrenzen und die unterschiedlichen Positionen zu begründen. „Es geht ja in diesem Schritt darum, Konsens und Dissens so klar und aussagekräftig wie irgend möglich zu dokumentieren“ (Renn/Webler 1998: 80). Nur so kann ein Dissens auch als Konsens über einen Dissens festgehalten werden, indem nicht nur die Uneinigkeit dokumentiert wird, sondern auch die Argumente für die jeweiligen Positionen der Expert*innen. Nur durch diese inhaltliche Ergänzung und Eingrenzung kann die Ergebnissicherung eines Gruppendelphis einen Dissens in einen *vernünftigen Dissens* (vgl. Gottschalk 1999: 32) bzw. *rationalen Dissens* (vgl. Miller 1992: 39) überführen, der durch eine Kalibrierung der Expert*innenurteile im Prozess erreicht werden konnte. Bei der Dokumentation der Plenarsitzung muss also besonders darauf geachtet werden, dass alle mit dem Ergebnis verknüpften Argumente, die letztlich den Konsens über einen Dissens (sowie ebenso den Konsens) begründen, protokollarisch erfasst wurden. Hierbei kann eine Audioaufnahme des Austauschs im Plenum eine gute Datengrundlage bilden, um auch nachträglich Argumente identifizieren zu können. Im Falle des Gruppendelphis zu Eco-Fahrassistenten wurde eine solche Aufnahme angefertigt, transkribiert und in ein anonymisiertes Protokoll überführt. So standen für die Auswertung dieser Arbeit wesentlich mehr Aussagen für die Inhaltsanalyse zur Verfügung als beispielsweise im Fallbeispiel in PROSUITE, bei dem ein knappes tabellarisches Protokoll angefertigt wurde. Auch wenn nicht jedes Gruppendelphi eine Inhaltsanalyse des Diskurses innerhalb des Plenums, wie in dieser Arbeit geschehen, notwendig macht, so war die Datenlage im Falle von PROSUITE zu gering, um jeden anhand der statistischen Werte erzielten Konsens und auch Konsens über einen Dissens mit inhaltlichen Begründungen und Argumenten stützen zu können. Dies ist gerade beim Konsens über einen Dissens aus zweierlei Hinsicht problematisch: Zum einen lässt sich nicht eindeutig konstatieren, ob es sich um einen *tatsächlichen* Konsens über einen Dissens handelt und zum zweiten fehlen die inhaltlichen Begründungen der jeweiligen Positionen um zukünftigen Forschungsbedarf klar zu identifizieren. Die Fälle der Gruppendelphis zu SAUBER+, SEAWIND und Smart-Metering zeigen, dass ein umfassendes schriftliches **Protokoll** die Aufzeichnung der notwendigen Inhalte sicherstellen kann, eine Audioaufnahme bietet aber überdies eine zusätzliche Absicherung. Zumal der Fall des Gruppendelphis zu Eco-Fahrassistenten trotz der Aufnahme keinen Hinweis auf einen Scheinkonsens liefert, der eventuelle Hemmnisse durch die Aufnahme des Plenums angedeutet hätte. Letztlich ist neuere technische Ausstattung für die Diskussion des Plenums nicht mehr so aufdringlich, als dass dieses kleine Gerät während des Austauschs - auch von Expert*innen - nicht schnell in Vergessenheit gerät und sich der Diskurs frei entfalten kann. Die Zusage zur Anonymisierung und Löschung der Aufzeichnung nach der Auswertung können zusätzlich einen freien Austausch befördern. Für die **Auswertung** dieser Arbeit wäre eine umfassendere Dokumentation der weiteren Fälle hilfreich gewesen, da so eine vergleichbarere Datenlage für die **Inhaltsanalyse** zur Verfügung gestanden hätte. Gerade im Falle der Dimension der normativen Angemessenheit kann nicht ausgeschlossen werden, dass neben den analysierten Protokollen nicht mehr Argumente an den Geltungsanspruch nach Gemeinwohl oder normativer Angemessenheit angeknüpft haben, da diese Protokolle ergebnisorientiert erstellt wurden.

Überdies bieten die empirischen Fälle hinsichtlich der **Fragebogenkonstruktion** weitere praxisrelevante Hinweise. Genuin ist die nicht überraschende und wiederkehrende Erkenntnis, dass der Fragebogen aufgrund der zeitlichen Beschränkungen, denen der Einsatz der Methode unterliegt, nicht zu ambitioniert und zu umfassend sein sollte. Überdies eignen sich die

verschiedenen **Fragetypen** unterschiedlich gut für den Einsatz während eines Gruppendelphis. Aufgrund der 10er-skalierten Abfrage bieten Ratingfragen, den für ein Gruppendelphi am geeignetsten Abfragetyp, da auf diesem Wege Aussagen abgefragt werden können, deren Bewertungen sich im Rahmen eines Gruppendelphis anhand ihrer statistischen Kenngrößen ideal für eine strukturierte Diskussion im Plenum aufbereiten lassen. Müssen einzelne Sachverhalte im Vergleich zu anderen direkt ausgewählt werden, so bietet sich eine Auswahlfrage an und sollen diese in eine hierarchische Beziehung gebracht werden, können Rankingfragen mittels einfachem Ranking oder über ein fiktives Budget eingesetzt werden. Hier sollte der Einsatz aber begrenzt erfolgen, da diese Fragetypen mehr Zeit in Anspruch nehmen, als es die 10er-skalierten Ratingfragen tun. Nach Analyse der empirischen Fälle kann auf drei Fehler hingewiesen werden, die bei der Konstruktion und Anpassung eines Fragebogens vermieden werden sollten.

- *Verknüpfung von mehreren Fragetypen und Inhalten:* Im Falle des Gruppendelphis zu Smart Metering wurden innerhalb einer Frage komplexe Szenarien inklusive der jeweiligen Kommunikationsakteure mittels 10er-skalierten Items abgefragt und simultan sollte eine Zuschreibung per Nennung der entsprechenden Zielgruppen erfolgen. Eine solche Verknüpfung von Inhalten und verschiedenen Fragetypen ist auch für die teilnehmenden Expert*innen eines Gruppendelphis zu komplex - besonders da während der Beantwortung ein konsensbildender Austausch unter zeitlicher Beschränkung erfolgen soll.
- *Wechsel von Skalentypen:* Im Falle des Gruppendelphis zu Eco-Fahrassistenten wurde von der in der ersten Runde eingesetzten Ratingskalierung zur zweiten Runde auf eine Rankingskala gewechselt. Dies führte nicht nur zu Problemen der Vergleichbarkeit der Ergebnisse beider Runden, sondern auch dazu, dass eine Kleingruppe die Rankingskala nicht als solche beantwortete, sondern eine *bloße* Auswahl von Items vornahm.
- *Offene Fragen:* Die besonders in den Fällen zu Smart-Metering und Eco-Fahrassistenten vermehrt eingesetzten offenen Fragen, bieten keinen idealen Fragetypus für ein Gruppendelphi. Zum einen sind auch diese Fragen in der Beantwortung sehr zeitintensiv, da sich die Teilnehmenden nicht nur auf ein gemeinsames Urteil verständigen, sondern überdies dieses auch gemeinsam formulieren müssen und zum zweiten, weil die Auswertung und die gemeinsame Verständigung zwischen den Gruppen im Plenum im Rahmen eines Gruppendelphis kaum umsetzbar sind. Letztlich bleibt die Frage, ob eine offene Abfrage überhaupt im Sinne einer konsensualen Ergebnisfindung dienlich ist, da individuelle Formulierungen miteinander in Einklang gebracht werden müssen, bevor überhaupt ein Diskurs über die gemeinsamen Inhalte erfolgen kann.

Auch wenn es sich bei der **Kategorie Sonstige(s)** ebenfalls im methodischen Sinne um eine offene Abfrage handelt, so bieten die empirischen Fälle hier durchaus Beispiele für ein Funktionieren des Einsatzes einer solchen. Der Unterschied zu einer Abfrage in einer offenen Frage ist, dass bei der Kategorie bereits der Fragekontext sowie der Rahmen in dem sich die neue Nennung bewegen kann, vorgegeben sind. So konnten, die von den Teilnehmenden unter dieser Kategorie genannten Items in einer Vielzahl der empirischen Fälle in ein konsensuales Ergebnis überführt werden. Damit die Abfrage solcher Vorschläge zur zweiten Runde des Gruppendelphis erfolgen kann, ist es allerdings wichtig einen Austausch über diese **neuen Items** während der ersten Plenarsitzung - oder zumindest zum zweiten Plenum - zu führen. Ansonsten hat man als Ergebnis zwar konsensual bewertete Items vorliegen, allerdings keine

weiteren Informationen zu den Inhalten oder den Argumenten, die diese Auswahl begründen, wie es im Falle von PROSUITE geschehen ist. Die Kategorie *Sonstiges* sollte bei der Fragebogenkonstruktion generell kritisch betrachtet werden, da diese auch Nennungen durch die Teilnehmenden provozieren kann und so eine Beschäftigung mit den auf vorhergehenden Analysen basierenden Aussagen und Items aus dem Fokus gerät. Auch wenn diese Kategorie funktional ist, sollte sie nur eingesetzt werden, wenn neue Items auch im Rahmen der Fragestellung des Gruppendelphis gewünscht sind.

Inwieweit die **Vorbefragung** einen konsensfördernden Effekt für den eigentlichen Workshoptag hat, lässt sich aus den Ergebnissen der betrachteten Fälle nicht ableiten. So zeigt sich sowohl im Fallbeispiel zu SAUBER+ als auch in SEAWIND, in denen eine Vorbefragung durchgeführt wurde, keine eindeutige Entwicklung hin zu einem konsensualen Ergebnis zwischen der Vorbefragung und der ersten Runde des Gruppendelphis. Dadurch, dass der Fragebogen der Vorbefragung individuell von den Teilnehmenden beantwortet wird, liegt die Fallzahl wesentlich höher als am Workshoptag, in dem die Kleingruppen jeweils einen gemeinsamen Fragebogen beantworten, so dass die Vergleichbarkeit der statistischen Ergebnisse nicht gänzlich gegeben ist. Die Vermutung, dass alleine aufgrund der geringeren Fallzahl auch konsensuälere Ergebnisse in der ersten Runde des Gruppendelphis zu beobachten seien, bestätigt sich allerdings nicht. Überdies wird eine weitere Annahme anhand der beiden empirischen Fälle nicht bestätigt. So wurden bei beiden Vorbefragungen strittige bzw. unklare Formulierungen der Items nicht bereits durch Kommentierung in den Fragebögen der Vorbefragung identifiziert. In beiden Fällen fand eine Diskussion hinsichtlich einzelner Formulierungen im ersten Plenum statt. Diese Vorschläge für Umformulierungen basierten meist auf einem Austausch in den Kleingruppen, aus denen der Wunsch nach Umformulierung im Plenum vorgetragen wurde. Ebenso führte der Austausch in der Kleingruppe oder im Plenum dazu, dass die Teilnehmenden eine fehlende Verständlichkeit einzelner Items (oder deren Teilen) identifizierten und daraufhin eine Umformulierung vorgeschlagen wurde. So scheint erst der diskursive Austausch eine Präzisierung zu befördern.

Letztlich zeigt ein Beispiel aus SEAWIND, dass es trotz der strukturellen Bedingungen, die ein Gruppendelphi methodisch bietet, auch eine **Kultur des Diskurses** bzw. ein Einlassen auf die Logik eines konsensbildenden Prozesses braucht. So wurde dies in diesem Beispiel negativ belegt, dass bei einem bestehenden Dissens von einzelnen Teilnehmenden eine Abstimmung präferiert wurde. Dies widerspricht dem Ziel eines Gruppendelphis und verweist die unterschiedlichen Positionen in Majorität und Minorität aus, statt die Argumente für die jeweilige Position auszutauschen und im Falle eines Konsenses über einen Dissens auch gegenseitig anzuerkennen. Ob eine solche Kultur des Diskurses für den Workshop bereits durch die Formulierung der Items befördert werden kann, soll zu Beginn des folgenden Fazits diskutiert werden.

9 Fazit

Im Vergleich unterscheiden sich die empirischen Fallbeispiele der Gruppendelphis in SEAWIND und PROSUITE von den übrigen drei Fällen dadurch, dass die Items nicht als formulierte Aussagen abgefragt wurden, die überwiegend eine *Sollensaussage* enthielten, sondern per direkter Abfrage des interessierenden Sachverhaltes erfolgten. Durch diese direkte Abfrage wirkt der Möglichkeitsraum für das Vorbringen der Argumente jeweiliger Positionen in beiden Fällen eingeschränkter, als es in den übrigen Fallbeispielen der Fall ist, insbesondere was die normative Angemessenheit adressierende Äußerungen betrifft. Dies führt nicht nur zu Rückschlüssen hinsichtlich der funktionellen Formulierung von Items für ein Gruppendelphi, sondern bietet überdies die Möglichkeit eine weitere Spezifizierung der Anwendungsfelder des methodischen Ansatzes zu unternehmen.

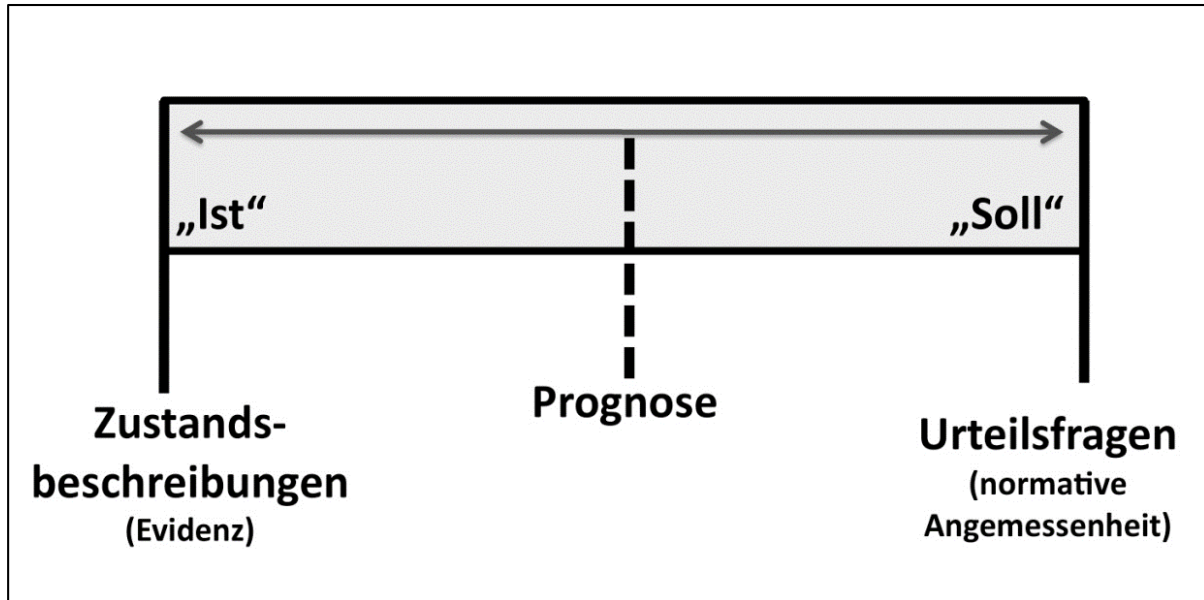
So zeigte die Dimension der normativen Angemessenheit bereits, inwieweit die Bandbreite des Einflusses der persönlichen Einstellungen der Expert*innen in die Urteilsbildung einfließen können bzw. dürfen. Zwar spricht sich Niederberger dafür aus, dass persönliche Einstellungen und Werte der Teilnehmenden nicht in die Diskussion während eines Gruppendelphis mit einfließen sollten. Dies betrifft besonders Grundsatzfragen bei denen ein grundlegender Dissens erkennbar ist (vgl. ebd. 2015b: 132). Auch Renn und Webler beschreiben eine Neigung von Expert*innen auf Basis ihres genuinen Wissens bereits die politischen Schlussfolgerungen mitzuliefern, was nicht ihrer Aufgabe entspricht und aufgrund fehlender politischer Legitimation problematisch sein dürfte (vgl. ebd. 1998: 80). Damit sind Aussagen über den Soll-Zustand im Verlauf eines Gruppendelphis aufgrund der beschriebenen Grenzen komplizierter einzustufen. Die Fälle zeigen aber auf, dass ein solcher Bezug nicht gänzlich auszuschließen ist und von Teilnehmenden durchaus als legitim erachtet wird. Damit sollten die Ansichten der Teilnehmenden nicht grundsätzlich für ein Delphi als Verfahren insgesamt ausgeschlossen werden. Auch in Häders Typenbildung¹⁵² ist die Anwendung des Ansatzes zur *Ermittlung von Expertenmeinungen* ein legitimer Anwendungstypus. Die Qualität lässt sich dabei daran bemessen, „(...) inwieweit die Ansichten aller Teilnehmer mithilfe des Instrumentes methodisch einwandfrei abgebildet wurden und ob erwartet werden kann, dass es im Verlauf der Delphi-Befragung tatsächlich zu einer Verbesserung der Urteile gekommen ist“ (Häder 2009: 33). Damit die Teilnehmenden in diesem Prozess aber nicht die Rolle der Entscheidungsträger*innen einnehmen, sollten die im Delphi abgefragten Aussagen so formuliert sein, dass sie sich auf entscheidungsrelevante Wissensbestände der Expert*innen beschränken. So können auch normative Betrachtungen unter dem Gesichtspunkt der umfassenden Bewertung in die angenommenen Konsequenzen und wahrscheinlichen Implikationen einzelner Maßnahmen mit einfließen (vgl. Renn/Webler 1998: 80; Failing et al. 2004: 16).

Abgefragte Aussagen können folglich in Aussagen über den *Ist-Zustand* und Aussagen über den *Soll-Zustand* unterschieden werden. Aussagen, die den Ist-Zustand betreffen, beziehen sich im Rahmen der Bewertung vorwiegend auf das Geltungskriterium der Evidenz, wobei Aussagen über den Soll-Zustand als Bezugsgröße vorwiegend das Geltungskriterium der normativen Angemessenheit heranziehen. Während es sich somit bei Fragen zum Ist-Zustand

¹⁵² vgl. Kapitel 3.2

um Zustandsbeschreibungen handelt, die an den derzeitigen wissenschaftlichen Erkenntnissen der Teilnehmenden bemessen werden können, handelt es sich bei den Soll-Aussagen um Urteilsfragen, die ebenso die normativen und wertgeleiteten Vorstellungen der Teilnehmenden miteinschließen können. Beide Aussagetypen bilden die Ränder eines Kontinuums, wie in Abbildung 62 dargestellt.

Abbildung 62: Mögliche im Gruppendelphi abgefragte Aussagen und deren Bewertungskontinuum



Quelle: Eigene Darstellung

Aufgrund der prognostischen Ausrichtung der im Gruppendelphi eingesetzten Items, können die Prognosen sich entweder eher auf Zustandsbeschreibungen und deren evidenzbasierter Ableitung oder aber eher auf Urteilsfragen nach normativer Angemessenheit berufen. So bildet sich die Zielgröße der Fragen bei einem Gruppendelphi letztlich auf einem Kontinuum ab, dass ein auf dem Ist-Zustand basierendes Urteil entlang der prognostischen Ausrichtung erwartet, die wie in der Abbildung dargestellt ein gewisses Maß an Urteilsfragen beinhaltet, die den Geltungskriterien der Richtigkeit unterliegen. Was dabei als richtig zu beurteilen ist, kann und soll durch einen diskursiven Austausch der Teilnehmenden erreicht werden, der soweit möglich einen auf Evidenz basierten Wissensdiskurs abbildet. Allerdings beinhaltet das Kontinuum auch die Komponente der Unwissenheit, die den Einsatz der Methode erst initiiert und ebenso eine Abwägung zwischen Zielgrößen (Reflexionsdiskurs) und eine Auswahl der gewollten Richtung (Gestaltungsdiskurs) integriert. Letztlich können sich die bei einem Gruppendelphi getroffenen Prognosen zur Sinnhaftigkeit einzelner Aussagen und Maßnahmen immer nur auf einem Kontinuum zwischen Zustandsbeschreibung und Urteilsfragen bewegen und sollten aufgrund der Ausrichtung im Ideal soweit an einen der Pole reichen, wie es aus Sicht der Teilnehmenden notwendig ist. Dies kann nur mittels eines argumentativen Austauschs transparent geschehen, wie es die diskursiven Elemente eines Gruppendelphis ermöglichen.

Aufgrund der Ausrichtung und der Herausforderungen eines inhaltlichen Diskurses auf diesem Kontinuum, bildet die Theorie des kommunikativen Handelns für ein Gruppendelphi eine unverzichtbare theoretische Grundlage und schließt damit den argumentativen Kreis zu den wissenschaftstheoretischen Grundlagen. Denn so kann eine Verknüpfung von einem auf Wissen ausgehenden über den Austausch der potentiellen Konsequenzen basierenden Konsens mit

den Werten verknüpft werden, die in eben diesem Diskurs gelöst werden sollen (vgl. Renn 2000a: 20). Das Gruppendelphi bietet aufgrund seiner Integration der diskursiven Elemente dem mit der Modifikation des klassischen Delphiverfahrens angestrebten Anspruch Rechnung, indem die Methode in Verknüpfung mit den strukturierenden Elementen der Befragung und den damit einhergehenden statistischen Kenngrößen einen intensiven Austausch von Begründungen und Argumenten befördert, wie in den empirischen Fällen aufgezeigt werden konnte: **So können die Qualitätsmerkmale handlungsleitender Diskurse von der Methode des Gruppendelphis eingelöst werden.** Die Ergebnisse eines Gruppendelphis entsprechen sowohl dem engen, als auch dem weiten Verständnis von Rationalität, indem sie vernünftige Gründe für die getroffenen Einschätzungen vorlegen.

Ein Gruppendelphi bietet den Rahmen für einen kommunikativen Austausch, bei dem besonders bei vorherrschender Kritik an den jeweiligen Positionen Argumente für die jeweiligen Einschätzungen vorgebracht werden. Dafür ist es unabdingbar einem Diskurs einen methodisch sinnvollen und funktionalen Rahmen zu geben, denn damit „(...) Diskurse über die ihnen aus der Lebenswelt gelieferten synthetisierenden Leistungen verfügen können, müssen sie deshalb daran gehindert werden, die ihnen eigentümliche expansive Tendenz voll zu entfalten. Diskurse können nicht ihrem Ideal entsprechen, sie müssen gleichsam gegen sich selber geschützt werden“ (Giegel 1992b: 70). Um einen solchen Rahmen zu definieren, dienen in einem Gruppendelphi nicht nur die standardisierten Elemente, wie der Fragebogen oder die quantitativen Analysen, sondern bilden auch die anhand dieser Diskussionsinputs strukturierten Runden des Austauschs in den Kleingruppen sowie im Plenum wichtige Elemente. Dabei muss aufgrund der limitierten möglichen Teilnehmendenzahl eines Gruppendelphis darauf geachtet werden, dass „(...) die gesamte Bandbreite an Urteilen vertreten ist und die Argumente der verschiedenen Positionen transparent werden“ (Niederberger 2015b: 127). So werden auch bei einem Gruppendelphi die konstitutiven Sprechhandlungen zu einer kommunikativen Praxis, „(...) die vor dem Hintergrund einer Lebenswelt auf die Erzielung, Erhaltung, und Erneuerung von Konsens angelegt ist, und zwar eines Konsenses, der auf der intersubjektiven Anerkennung kritisierbarer Geltungsansprüche beruht“ (Habermas 1981: 37).

Nun folgen Modifikationen von etablierten Verfahren und Methoden häufig mit dem hehren Ziel, die Schwächen der ursprünglichen Anwendung auszugleichen und letztlich eine qualitativ bessere und/oder auch effizientere Lösung zu entwickeln. Der Vorwurf, der dem klassischen Delphiverfahren häufig zur Kritik gereicht wird, liegt in der quantitativen Anwendung des Verfahrens begründet, dass von einer Tendenz hin zur Mehrheit von Runde zu Runde beeinflusst werden soll und letztlich damit nur die *schwachen Swingers* gegenüber den *beständigen Holdouts* nachgeben: „*However, recent reviews and research on the Delphi process suggest that to achieve improvements in accuracy from round to round, experts must be provided with rationales to accompany the feedback they receive about the responses from other group members, and that in the absence of these rationales, their responses will tend to converge only towards a majority position*“ (McBride et al. 2012: 907). Dieser Vorwurf trifft sicherlich auf einzelne Anwendungen des klassischen Delphiverfahrens zu, kann aber nicht als generelle Kritik ausgelegt werden. So zeigt beispielsweise Häder auf, dass auch beim klassischen Delphiverfahren eine Urteilsänderung nicht zwangsläufig auf einer Tendenz zur Majorität, sondern durchaus auf kognitionspsychologische Prozesse begründet werden kann (vgl. ebd. 2009: S. 43-53). Auch die Ergebnisse sowie die zugrundeliegenden Prozesse der Urteilsänderung sind

evaluierbar (vgl. Häder 2009: S. 191-209), wenn Häder auch für das klassische Verfahren weiteren Forschungsbedarf und weitere Verifikation der Funktionsweise des Verfahrens identifiziert (vgl. Häder 2009: S. 211-213). Um diese theoretische Grundlage, die vorwiegend beschreibende Elemente für die Funktionsweise des Delphiverfahrens bereitstellt weiter zu festigen, wurden in dieser Arbeit zusätzlich theoretische Annahmen aus der Social Judgment Theory mit ihrem Linsenmodell integriert (vgl. Hammond 1996: 86f.; Goldstein 2004: 41f.). Diese wissenschaftstheoretischen Grundlagen wurden aufgrund der Diskursivität des Untersuchungsgegenstandes dieser Arbeit für eben dieses Gruppendelphi mit theoretischen Annahmen aus der Theorie des kommunikativen Handelns ergänzt (vgl. Habermas 1981), um diese in Dimensionen eines Testdesigns zu überführen. Die wissenschaftstheoretischen Grundlagen wurden folglich für die konkrete Methode des Gruppendelphis geschaffen.

Der bestehende Forschungsbedarf für das Delphiverfahren bleibt damit aufgrund der vielfältigen Einsatzfeldern allgemein bestehen, da das Verfahren von einer Ideenaggregation, der Bestimmung eines Sachverhaltes, der Ermittlung von Expert*innenmeinungen bis hin zur Konsensbildung jeweilig spezifische Ziele verfolgt (vgl. Häder 2009: 36). Die nach diesen Zielen aufgestellte Typologie von Anwendungsfeldern des Delphiverfahrens ist überdies im Hinblick auf das Gruppendelphi nicht trennscharf anwendbar, da das Gruppendelphi das Ziel verfolgt einen unbekanntem Sachverhalt, aufgrund der Ermittlung von Expert*innenurteilen konsensual zu ermitteln, womit bereits drei der Typen abgedeckt wären. Alleine dieser Umstand verhindert einen direkten Vergleich zwischen einem Gruppendelphi und dem klassischen Delphiverfahren, da jeder Einsatz eines Verfahrens bzw. einer Methode mit spezifischen Fragestellungen und den dazugehörigen Einsatzfeldern verknüpft ist. Wenn es aber darum geht auch die Argumente und Begründungen hinter konsensualen Bewertungen (besonders bei Konsens über Dissens) zu erfassen, so zeigt sich das Gruppendelphi als besonders geeignet. Überdies bestehen die bereits vor der Analyse der empirischen Fallbeispiele konstatierten Vorteile¹⁵³ auch nach der Betrachtung der empirischen Daten, besonders was die in Forschungsvorhaben nicht trivialen Aspekten wie Einsparung von Zeit und Geld betrifft.

Gerade die meist knappe Zeitspanne, die für ein Gruppendelphi zur Verfügung steht, knüpft an die Frage eines sinnvollen Abbruchkriteriums an, das im Ideal der Diskurstheorie erst damit erreicht wäre, wenn die Akzeptanz über einen zuvor strittigen Geltungsanspruch durch ein zwangloses Einverständnis erreicht würde (vgl. Habermas 1992: 278). Dabei wird bei praktischen Delphianwendungen häufig eine Definition von Konsens als Abbruchkriterium festgelegt, die meist auf prozentualen oder auf proportionalen Zustimmungswerten der Teilnehmenden in einem spezifischen Range basiert. Damit wird ein spezifisch erreichter Konsens letztlich auch als Abbruchkriterium festgelegt (vgl. der Gracht 2012: 1529; Diamond et al. 2014: 405). Dies ist insofern problematisch, da das vorrangige Ziel des Endes eines Delphis nicht die reine Orientierung am Konsens sein sollte. „*In fact numerous researchers have illustrated why stability should be tested first, before terminating Delphi rounds based on consensus*“ (von der Gracht 2012: 1533). Dies stellt auch das Gruppendelphi vor eine Herausforderung, da die Situation eines Workshops meist nicht den zeitlichen Rahmen bietet eine zusätzliche Runde durchzuführen, um neben dem erreichten Konsens auch die Stabilität eben dieses Ergebnisses zu prüfen. Aus diesem Grund kommt dem Ergebnisprotokoll hier nicht nur die Aufgabe zu,

¹⁵³ siehe Kapitel 4.3

beispielsweise mögliche Scheinkonsense aufzudecken, sondern auch die Stabilität der Ergebnisse zu verifizieren, wie auch in den dargestellten Fällen zu beobachten war.

Eine exklusive Fokussierung auf bloße Konsensbildung schließt überdies ein zweites legitimes Ergebnis eines Delphis aus. Eine Anwendung innerhalb des klassischen Delphiverfahrens zielt dabei besonders auf die Identifikation von Dissens ab und wird als *dissensus* oder *Policy Delphi* bezeichnet (vgl. Diamond et al. 2014: 401). Bei einer solchen Anwendung ist das Ziel besonders die strittigen Standpunkte und verschiedene Meinungen herauszustellen. „*Future research could concentrate more on the characteristics in dissent-orientated Delphi design in order to best serve the purposes of maximizing variance and different argumentations*“ (von der Gracht 2012: 1534). Dies schließt an die Argumentation an, dass das Ergebnis eines Diskurses nicht unbedingt Einigkeit bei allen Sachfragen bedeuten sollte, allerdings mehr Klarheit erreicht werden muss (vgl. Renn 2007: 180). Gerade, wenn sich der Fokus eines Delphiverfahrens auch auf die Identifikation strittiger Sachverhalte verschiebt, wird es nicht nur aus diskurstheoretischer Sicht unabdingbar, auch die Argumente für die strittigen Positionen zu identifizieren und herauszustellen. Hier bietet gerade das Gruppendelphi aufgrund der diskursiven Elemente einen Vorteil gegenüber dem klassischen Verfahren. An dieser Stelle findet sich die Urteilsbildung zwischen einer analytisch-sachgerechten Abbildung der Realität und der normen- und wertgeleiteten Abwägung wieder. Auf dieser Grundlage können die Teilnehmenden eines Gruppendelphis ein rationales Urteil aus globalen Wissensbeständen ableiten, das Lösungswege für die Zukunft bereitstellt. Solche Urteile müssen nicht zwangsläufig konsensual sein, wenn sich die Urteile unterscheiden, so zeigen sie zwei mögliche Lösungsansätze hinsichtlich der gegebenen Fragestellung, die entweder auf sachgerechter Ebene weiteren Forschungsbedarf identifizieren oder sich in ihrer Wünschbarkeit unterscheiden.

Aus dieser Perspektive ist es allerdings wichtig, eine spezifische Eingrenzung des Begriffes eines Dissenses vorzunehmen. Wie bereits in der Begriffsbestimmung zu einem Konsens geschehen,¹⁵⁴ ist nicht jeder Dissens auch aus diskurstheoretischer Perspektive eine Zielgröße, die akzeptabel ist. Eine Methode muss die Rahmenbedingungen bereitstellen nicht bloß einen Dissens, sondern einen *rationalen Dissens* (vgl. Miller 1992: 39), einen *vernünftigen Dissens* (vgl. Gottschalk 1999: 36) oder wie in dieser Arbeit als Begriff verwendeten *Konsens über einen Dissens* zu erzielen, der für die einzelnen Positionen immer auch jeweilige Gründe und Argumente beinhaltet, die die Basis dieser unterschiedlichen Einschätzungen bilden. So kann anhand unterschiedlicher Positionen konkreter Forschungsbedarf identifiziert und eingegrenzt werden und dem Ideal nach Habermas folgend, der *Kette substanzieller Gründe* auch nach der eingesetzten Methode im weiteren Feld eine Möglichkeit eröffnet werden (vgl. Habermas 1992: 278). So bemisst sich auch ein Gruppendelphi nicht bloß daran bei wie vielen Sachverhalten ein konsensuales Ergebnis erreicht werden konnte, sondern auch an der Identifizierung von Konsens über Dissens hinsichtlich weiterer Sachverhalte. Die Ergebnisse eines Gruppendelphis können damit als *Kalibrierung* von Expert*innenurteilen verstanden werden (vgl. Niederberger/Renn 2018: 1), womit das Gruppendelphi eher als *Erkenntnisinstrument* verstanden werden kann, denn ein *Erhebungsinstrument* darstellt (vgl. Niederberger/Renn 2018: 182).

¹⁵⁴ siehe Kapitel 5.2.1

Dies führt zurück zur Frage, inwieweit das Gruppendelphi über eine rein technische Bewertung hinausgehen kann? So sollte eine Bewertung von Maßnahmen und Sachverhalten, die gesellschaftlich wichtige Themen und zukunftsrelevante Entscheidungen betreffen, nicht auf einer rein wissenschaftlich-technischen Bewertung erfolgen: *„Probleme der gesellschaftlichen Akzeptanz von durch wissenschaftlich-technischen Entwicklungen ausgehenden Folgen auf Leben und Zusammenleben der Gesellschaftsmitglieder können unserer Auffassung nach nur bedingt auf rein wissenschaftlich technischer Ebene gelöst werden; mit teilweise vielschichtigen psychischen, sozialen und kulturellen, mit politisch-rechtlichen, moralischen und religiösen Aspekten handelt es sich im Kern um Wert- und Normenprobleme, die allein mit wissenschaftlichen Methoden nicht angegangen werden können. Somit bleiben z.B. auch innertechnische Orientierungswerte wie Funktionstüchtigkeit, Effizienz oder Perfektion, außertechnischen Wertorientierungen untergeordnet“* (Schenk/Döbler 1996: 80). Wie die empirischen Fälle aufzeigen, kann das Gruppendelphi diese außertechnischen Wertorientierungen aufnehmen und zum argumentativen Gegenstand der Bewertungen der Teilnehmenden machen. Allerdings bleibt dies auf die Teilnehmenden des Prozesses - seien es je nach Begriffsbestimmung Expert*innen oder Stakeholder - begrenzt. Die Inklusion anderer Standpunkte und Interessen kann über das Verfahren nicht gänzlich erfolgen. Dies führt allerdings nur dann zwangsläufig zu einem Legitimationsproblem, wenn der Einsatz des Gruppendelphis exklusiv bleibt. So empfiehlt sich bei komplexen Sachverhalten der Einsatz des Gruppendelphis in einem *between-method-Design*, wie es in den empirischen Fallbeispielen in unterschiedlicher Form angewandt wurde. So können unterschiedliche Perspektiven auf den interessierenden Sachverhalt eingenommen werden und durch die Kombination von aufeinander aufbauenden Methoden ein umfassenderes Bild des Untersuchungsgegenstandes erlangt werden (vgl. Kuhn 2012: 139). Legitimation sollte der Einsatz des Gruppendelphis in einem solchen Methodenaufbau dabei zum einen dadurch erhalten, dass er entsprechend der Fragestellung sinnvoll ist und zum zweiten die richtige Zielgruppe involviert. So eignet sich ein Gruppendelphi besonders dann, wenn die Urteile von Expert*innen oder Stakeholdern zu ambiguen und komplexen Sachverhalten benötigt werden (vgl. Renn 2008: 280) und sowohl funktionalistische Ziele der Wissensintegration wie auch deliberative Ziele des Arguing verfolgt werden (vgl. Alcántara et al. 2016: 68f.). Wenn es allerdings darum geht politische Entscheidungen vorzubereiten und gesellschaftlich relevante Fragen zu behandeln, so kann das Gruppendelphi nur ein Teil dieses Prozesses sein und besonders im Bereich der Problemdefinition einen Beitrag zu den Phasen eines komplexen Politikzyklus und Entscheidungsbildungsprozesses leisten (vgl. Alcántara et al. 2016: 143).

Bei der Entscheidung eines Einsatzes der Methode des Gruppendelphis sollte letztlich dem Grundsatz gefolgt werden, der bei einer Entscheidung zum Einsatz einer Methode immer im Vordergrund stehen sollte: *„(...) the most important advice to be given to the potential user is to suit the method to the problem, not the problem to the method“* (Lindstone 1985: 627)¹⁵⁵.

¹⁵⁵ Das Originalzitat von Lindstone beginnt mit einem „perhaps“ (Lindstone 1985: 627). Diese sprachliche Abmilderung wurde für dieses Fazit bewusst ausgelassen.

10 Literatur

- Abromeit, Heidrun (2002): Wozu braucht man Demokratie? Die postnationale Herausforderung der Demokratietheorie. Opladen: Leske und Budrich.
- Adelman, Leonard; Steward, Thomas R.; Hammond, Kenneth R. (1975): A Case History of the Application of Social Judgment Theory to Polica Formulation. In: Policy Sciences 6, S. 137-159.
- Alcántara, Sophia; Bach, Nicolas; Kuhn, Rainer; Ullrich, Peter (2016): Demokratietheorie und Partizipationspraxis. Analyse und Anwendungspotentiale deliberativer Verfahren. In: Klein, A.; Kleinfeld, R.; Krimmer, H.; Rehder, B.; Teune, S.; Walk, H.; Zimmer, A.: Schriftenreihe Bürgergesellschaft und Demokratie. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ammon, Ursula (2009): Delphi-Befragung. In: Kuhl, S; Strodtolz, P.; Taffertsdorf, A. (Hrsg.): Handbuch Methoden der Sozialforschung. Quantitative und Qualitative Methoden. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 458-476.
- Baron, Jonathan (2004): Normative Models of Judgment and Decision Making. In: Koehler, D.J.; Harvey, N. (Hrsg.): Blackwell Handbook of Judgment & Decision Making. Malden, Oxford, Carlton: Blackwell Publishing, S. 19-36.
- Benninghaus, Hans (2002): Deskriptive Statistik: Eine Einführung für Sozialwissenschaftler. Wiesbaden: Westdeutscher Verlag.
- Blok, Kornelis; Huijbregts, Mark; Roes, Lex; van Haaster, Berthe; Patel, Martin; Hertwich, Edgar; Wood, Richard; Hauschild, Michal Z.; Sellke, Piet; Antunes, Paula; Hellweg, Stefanie; Ciroth, Andreas, Harmelink, Mirjam (2013): A Novel Methology for the Sustainable Impact Assessment of New Technologies. A Report prepared within the EC 7th framework project. Online verfügbar: <http://46.105.145.85/web/guest/the-prosuite-framework>.
- Bowles, Nick (1999): The Delphi technique. Nursing Standard 13, 45, S.32-36.
- Brosi, Walter; Krekel, Elisabeth M.; Ulrich, Joachim Gerd (2003): Delphi Erhebung zur Identifikation von Forschungs- und Entwicklungsaufgaben in der beruflichen Aus- und Weiterbildung. *Stichworte: Zukunftsstudien, Prognose, Mythos des Orakels, Definition, Delphis in Deutschland, Integration von Workshops.*
- Cuhls, Kerstin (2000): Wie kann ein Foresight-Prozess in Deutschland organisiert werden? Gutachten, Domitra, M. (Hrsg.). Düsseldorf, Friedrich-Ebert-Stiftung.
- Cuhls, Kerstin; Blind, Knut; Grupp, Hariolf (1998): Delphi '98 Studie zur globalen Entwicklung von Wissenschaft und Technik. Fraunhofer Institut für Systemtechnik und Innovationsforschung.
- Cuhls, Kerstin; Kimpeler, Simone (2008): Delphi-Report: Zukünftige Informations- und Kommunikationstechniken. In: Fazit-Schriftenreihe, Forschungsbericht Band 10, S. 1-160.
- Cuhls, Kerstin; Möhrle, Martin G. (2005): Unternehmensstrategische Auswertung der Delphi-Berichte. In: Möhrle, Martin G.; Isenmann, Ralf: Technologie-Roadmapping. Zukunftsstrategien für Technologieunternehmen. Berlin u.a.: Springer, S.103-130.

- Dalkey, Norman C. (1969): The Delphi Method: An experimental Study of Group Opinion. Report prepared for United States Air Force Project RAND. RAND Corporation, Santa Monica.
- Dalkey, Norman C.; Helmer, Olaf (1962): An experimental application of the Delphi method to the use of Experts. Report prepared for United States Air Force Project RAND. RAND Corporation, Santa Monica.
- Deuschle, Jürgen; Sonnberger, Marco (2009): Das Gruppendelphi im Adipositas-Projekt. In: Schulz, M.; Renn, O.: Das Gruppendelphi: Konzept und Fragebogenkonstruktion. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 65-77.
- Diamond, Ivan R.; Grant, Robert C.; Feldman, Brian M.; Pencharz, Paul B.; Ling, Simon C.; Moore, Aileen M.; Wales, Paul W. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology* 67, S. 401–409.
- Diaz-Bone, Rainer (2002): Kulturwelt, Diskurs und Lebensstil. Eine diskurstheoretische Erweiterung der Bourdieuschen Distinktionstheorie. Reihe "Forschung Soziologie", Bd. 164. Opladen: Leske + Budrich.
- Diekmann, Andreas (2009): Empirische Sozialforschung. In: König, B. (Hrsg.): Rowohlt's Enzyklopädie. Reinbek: Rowohlt Taschenbuch Verlag.
- Dreyer, M.; Kuhn, R. (2014): Arzneimittelrückstände im Gewässerkreislauf: Partizipative Untersuchungsmethoden in einem jungen Risikofeld. In: DIALOGIK (Hrsg.): Innovativ und partizipativ: Einblicke in die Arbeit von DIALOGIK: ein Beitrag zum 10-jährigen Jubiläum von DIALOGIK. Stuttgarter Beiträge zur Risiko- und Nachhaltigkeitsforschung, Nr. 30, S. 48-59.
- Fach, Wolfgang (1974): Diskurs und Herrschaft - Überlegungen zu Habermas' Legitimationslogik. In: *Zeitschrift für Soziologie*, Jg. 3, Heft 3, S. 221-228.
- Failing, Lee; Horn, Graham; Higgins, Paul (2004): Using Expert Judgement and Stakeholder Values to Evaluate Adaptive Management Options. In: *Ecology and Society*, Vol. 9 (1). Online verfügbar: <http://www.ecologyandsociety.org/vol9/iss1/art13>
- Gallego Carrera, Diana (2009a): Der Einsatz des Delphiverfahrens zur Identifikation von Indikatoren zur Bewertung der Nachhaltigkeit von Stromerzeugungstechnologien in der Schweiz. In: Schulz, M.; Renn, O.: Das Gruppendelphi: Konzept und Fragebogenkonstruktion. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 95-109.
- Gallego Carrera, Diana (2009b): Der Einsatz des Delphiverfahrens zur Identifikation von Indikatoren zur Messung sozialer Effekte von Stromerzeugungstechnologien in Europa. In: Schulz, M.; Renn, O.: Das Gruppendelphi: Konzept und Fragebogenkonstruktion. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 79-93.
- Gasbeek, Anne; Meijer, Ellen (2013): Handbook on a novel methodology for the sustainability impact assessment of new technologies. Report prepared within the EC 7th Framework. Online verfügbar: <http://46.105.145.85/web/guest/the-prosuite-framework>.
- Giebel, Marion (2001): Das Orakel von Delphi – Geschichte und Texte. Reclam, Stuttgart.
- Giegel, Hans-Joachim (1992a): Einleitung. Kommunikation und Konsens in modernen Gesellschaften. In: Giegel, H.-J. (Hrsg.): Kommunikation und Konsens in modernen Gesellschaften. Frankfurt am Main: Suhrkamp Verlag, S. 7-17.

- Giegel, Hans-Joachim (1992b): Diskursive Verständigung und systemische Selbststeuerung. In: Giegel, H.-J. (Hrsg.): Kommunikation und Konsens in modernen Gesellschaften. Frankfurt am Main: Suhrkamp Verlag, S. 59-112.
- Gigerenzer, Gerd; Hoffrage, Ulrich; Kleinbolting, Hans (1991): Probabilistic Mental Models: A Brunswikian Theory of Confidence. In: Psychological Review, Vol. 98 (4), S. 506 – 528.
- Goldstein, William M. (2004): Social Judgment Theory: Applying and Extending Brunswik's Probabilistic Functionalism. In: Koehler, D.J.; Harvey, N. (Hrsg.): Blackwell Handbook of Judgment & Decision Making. Malden, Oxford, Carlton: Blackwell Publishing, S. 37-61.
- Goodman, Claire M. (1987): The Delphi technique: a critique. In: Journal of Advanced Nursing, Nr.12, S.729-734.
- Gottschalk, Niels (1999): Diskursethik. Begründungs- und Anwendungsfragen. Wissenschaftliche Arbeit zur Erlangung des Dr. phil. Universität Stuttgart.
- Gregory, Robin S. (2016): The Troubling Logic of Inclusivity in Environmental Consultations. In: Science, Technology, & Human Values. Vorabveröffentlichung online verfügbar: <http://sth.sagepub.com/content/early/2016/08/23/0162243916664016.full.pdf+html>
- Grunwald, Armin (1998): Das prädiskursive Einverständnis. Wissenschaftlicher Wahrheitsbegriff und prozedurale Rechtfertigung. In: Journal for General Philosophy of Science, Nr. 29, S. 205-223.
- Habermas, Jürgen (1973): Wahrheitstheorien. In: Fahrenbach, H. (Hrsg.): Festschrift für Walter Schulz, Wirklichkeit und Reflexion. Pfullingen, S. 211 - 265.
- Habermas, Jürgen (1981): Theorie des kommunikativen Handelns (Band 1). Handlungsrationalität und gesellschaftliche Rationalisierung. 9. Auflage. Frankfurt am Main: Suhrkamp Verlag.
- Habermas, Jürgen (1984): Theory of Communicative Action. Vol. 1: Reason and the Rationalization of Society. Boston: Beacon Press.
- Habermas, Jürgen (1992): Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats. Frankfurt am Main: Suhrkamp Verlag.
- Häder, Michael (2000): Die Expertenauswahl bei Delphibefragungen. In: ZUMA How-to-Reihe, Nr. 5, S. 1-15.
- Häder, Michael (2009): Delphibefragungen. Ein Arbeitsbuch. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Häder, Michael; Häder, Sabine (1995): Delphi und Kognitionspsychologie: Ein Zugang zur theoretischen Fundierung der Delphi-Methode. In: ZUMA-Nachrichten Nr. 37, S. 8-34.
- Häder, Michael; Häder, Sabine (2000): Die Delphimethode als Gegenstand methodischer Forschungen. In: Häder, M.; Häder, S. (Hrsg.): Die Delphitechnik in den Sozialwissenschaften: Methodische Forschungen und innovative Anwendungen. Opladen: Westdeutscher Verlag, S. 11-31.
- Hammond, Kenneth R. (1996): Human Judgment and Social Policy. Irreducible Uncertainty, Inevitable Error, Unavoidable Injustice. New York: Oxford University Press.
- Hartung, Joachim (1989): Statistik. München: R. Oldenbourg Verlag.

- Helmer, Olaf (1963): The systematic Use of Expert Judgment in Operation Research. The RAND Corporation, Santa Monica.
- Hilpert, Jörg; Kuhn, Rainer; Schetula, Viola; Renn, Ortwin (2013): Kommunikation von Risiken und Unsicherheiten der Strahlung kabelloser Netzwerke. Ergebnisse eines Gruppendelphis zur Risikokommunikation im Rahmen des EU Projektes „Sound exposure and risk assessment of wireless network devices“ (SEAWIND). In: Stuttgarter Beiträge zur Risiko- und Nachhaltigkeitsforschung, Nr. 28.
- Hitzler, Ronald (1994): Wissen und Wesen des Experten: Ein Annäherungsversuch – zur Einleitung. In: Hitzler, R.; Honer, A.; Maeder, Ch. (Hrsg.): Expertenwissen – Die institutionalisierte Kompetenz zur Konstruktion von Wirklichkeit. Westdeutscher Verlag, Opladen.
- Hsu, Chia-Chien; Sandford, Brian A. (2007): The Delphi Technique: Making Sense of Consensus. In: Practical Assessment, Research & Evaluation - A peer-reviewed electronic journal. (12/10), S. 1 – 8.
- Jäger, Alexander; Deuschle, Jürgen; Renn, Ortwin (2007): Ein normativ-funktionales Konzept für Nachhaltige Entwicklung. In: Renn, O.; Deuschle, J.; Jäger, A.; Weimer-Jehle, W.: Leitbild Nachhaltigkeit. Eine normativ-funktionale Konzeption und ihre Umsetzung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 39-71.
- Jäger, Magarete (2008): Diskursanalyse: Ein Verfahren zur kritischen Rekonstruktion von Machtbeziehungen. In: Becker, R.; Kortendiek, B. (Hrsg.): Handbuch Frauen- und Geschlechterforschung. Theorie, Methoden, Empirie. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 378-383.
- Jungermann, Helmut; Pfister, Hans-Rüdiger; Fischer, Katrin (2005): Die Psychologie der Entscheidung. Eine Einführung. München: Spektrum Akademischer Verlag.
- Juwana, Iwan; Perera, B.J. Chris; Muttill, Nitin (2010): A Water Sustainability Index for West Java – Part 2: Refining the Conceptual Framework Using Delphi Technique. In: Water Science and Technology, Nr. 62 (7), S. 1641-1652.
- Keller, Reiner (2011): Diskursforschung. Eine Einführung für SozialwissenschaftlerInnen. 4. Auflage. In: Bohnsack, R.; Flick, U.; Lüders, C.; Reichertz, J.: Qualitative Sozialforschung Band 14. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Keller, Reiner; Hirseland, Andreas; Schneider, Werner; Viehöver, Willy (2011): Zur Aktualität sozialwissenschaftlicher Diskursanalyse – Eine Einführung. In: Keller, R.; Hirseland, A.; Schneider, W.; Viehöver, W. (Hrsg.): Handbuch sozialwissenschaftliche Diskursanalyse. Band 1: Theorien und Methoden. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 7-34.
- Kerber, Heide; Schmitt, Bea; Schramm, Engelbert (2015a): Arzneimittel im Abwasser als „geteiltes Risiko“: Beobachtungen aus dem Stakeholder-Dialog des Projekts SAUBER+. In: Pinnekamp, J.; Palmowski, L.; Kümmerer, K.; Schramm, E. (Hrsg.): Abwasser aus Einrichtungen des Gesundheitswesens – Charakterisierung, Technologien, Kommunikation und Konzepte. Aachen: Gewässerschutz, Wasser, Abwasser (238), S. 101 – 115.
- Kerber, Heide; Schramm, Engelbert; Winker, Martina (2015b): Zwei Zukunftsszenarien zur Verringerung der Gewässerbelastung mit Humanarzneimittelrückständen. In: Pinnekamp, J.; Palmowski, L.; Kümmerer, K.; Schramm, E. (Hrsg.): Abwasser aus Einrichtungen des Gesundheitswesens – Charakterisierung, Technologien, Kommunikation und Konzepte. Aachen: Gewässerschutz, Wasser, Abwasser (238), S. 305 – 315.

- Kleimann, Bernd (1996): Das Dilemma mit den Experten – Ein Expertendilemma? Literaturbericht. In: Nennen, H-U.; Garbe, D.: Das Expertendilemma – Zur Rolle wissenschaftlicher Gutachter in der öffentlichen Meinungsbildung. Berlin, Heidelberg: Springer-Verlag, S.183-214.
- Knorr-Cetina, Karin; Amann, Klaus (1992): Konsensprozesse in der Wissenschaft. In: Giegel, H.-J. (Hrsg.): Kommunikation und Konsens in modernen Gesellschaften. Frankfurt am Main: Suhrkamp Verlag, S. 212-235.
- Kuckartz, Udo; Dresing, Thorsten; Rädiker, Stefan; Stefer, Claus (2008): Qualitative Evaluation: Der Einstieg in die Praxis. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kuhn, Rainer (2012): Fokusgruppen innerhalb eines between-method-Designs zur Bestimmung der Risikowahrnehmung von potenziellen Terroranschlägen. In: Schulz, M.; Mack, B.; Renn, O.: Fokusgruppen in der empirischen Sozialwissenschaft. Von der Konzeption bis zur Auswertung. Wiesbaden: VS Verlag für Sozialwissenschaften, S.134-147.
- Kuhn, Rainer; Weimer-Jehle, Wolfgang; Ruddat, Michael (2013): Soziale Aspekte nachhaltiger Entwicklung am Beispiel der Energiewende. In: Uhl, E.; Görtz, H.-D.; Brümmer, F. (Hrsg.): Nachhaltigkeit. Problemanalysen, Lösungsansätze, Perspektiven, Materialien (13). IZKT, Universität Stuttgart.
- Kuhn, Rainer; Tampe-Mai, Karolin; Mack, Birgit (2014): Das Gruppendelphi - eine diskursive Methode zur Erhebung von Expertenurteilen. Veranschaulicht am Beispiel eines Projekts zu Smart Metering. In: Zeitschrift für Umweltpsychologie, Nr. 34, 1/2014, S. 61-83.
- Kuhn, Rainer; Dreyer, Marion; Winker, Martina (2015): Handlungsbedarf und Maßnahmenbewertung zur Minderung des Arzneimittel eintrags in die Gewässer aus Stakeholderperspektive: Ergebnisse eines Gruppendelphis. In: Pinnekamp, J.; Palmowski, L.; Kümmerer, K.; Schramm, E. (Hrsg.): Abwasser aus Einrichtungen des Gesundheitswesens – Charakterisierung, Technologien, Kommunikation und Konzepte. Aachen: Gewässerschutz, Wasser, Abwasser (238), S. 317 – 323.
- Kusche, Isabell (2008): Politikberatung und die Herstellung von Entscheidungssicherheit im politischen System. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lagnado, David A.; Sloman, Steven A. (2004): Inside and Outside Probability Judgment. In: Koehler, D.J.; Harvey, N. (Hrsg.): Blackwell Handbook of Judgment & Decision Making. Malden, Oxford, Carlton: Blackwell Publishing, S. 157-176.
- Linstone, Harold A.; Turoff, Murray (1975): The Delphi Method: Techniques and Applications. London: Addison-Wesley.
- Linstone, Harold A. (1985): The Delphi Technique. In: Fowles, J. (Hrsg.): Handbook for Futures Research. Greenwood: Westport.
- Mack, Birgit; Tampe-Mai Karolin (2012): Konzeption, Diskussionsleitfaden und Stimuli einer Fokusgruppe am Beispiel eines BMU-Projekts zur Entwicklung von Smart Meter Interfaces und begleitenden einführenden Maßnahmen zur optimalen Förderung des Stromsparens im Haushalt. In: Schulz, M.; Mack, B.; Renn, O. (Hrsg.): Fokusgruppen in der empirischen Sozialwissenschaft. Von der Konzeption bis zur Auswertung. Wiesbaden, Springer VS, S. 66-87.
- Martin, Tara G.; Burgman, Mark A.; Fidler, Fiona; Kuhnert, Petra M.; Low-Choy, Samantha; McBride, Marissa; Mengersen, Kerrie (2011): Eliciting Expert Knowledge in Conservation Science. In: Conservation Biology, Vol. 26, Nr. 1, S. 29-38.

- Martschinke, Sabine (2001): Aufbau mentaler Modelle durch bildliche Darstellungen. Eine experimentelle Studie über die Bedeutung der Merkmalsdimensionen Elaboriertheit und Strukturiertheit im Sachunterricht der Grundschule. Münster, New York, München, Berlin: Waxmann Verlag.
- Mayer, Horst Otto (2006): Interview und schriftliche Befragung. München. Oldenbourg Verlag.
- Mayring, Philipp (2002): Einführung in die qualitative Sozialforschung: Eine Anleitung zu qualitativem Denken. Weinheim und Basel: Beltz Verlag.
- McBride, Marissa F.; Garnett, Stephen T; Szabo, Judit K.; Burbidge, Allan H.; Butchart, Stuart H.M.; Christidis, Les; Dutson, Guy; Ford, Hugh A.; Loyn, Richard H.; Watson, David M.; Burgman, Mark A. (2012): Structured elicitation of expert judgments for threatened species assessment: a case study on a continental scale using email. In: *Methods in Ecology and Evolution* (3/5), S.906-920.
- Miller, Max (1992): Rationaler Dissens. Zur gesellschaftlichen Funktion sozialer Konflikte. In: Giegel, H.-J. (Hrsg.): *Kommunikation und Konsens in modernen Gesellschaften*. Frankfurt am Main: Suhrkamp Verlag, S. 31-58.
- Niederberger, Marlen; Kuhn, Rainer (2013): Das Gruppendelphi als Evaluationsinstrument. In: *Zeitschrift für Evaluation* (1/13), S. 53-77.
- Niederberger, Marlen (2015a): Methoden der Experteneinbindung. In: Niederberger, M.; Wassermann, S. (Hrsg.): *Methoden der Experten- und Stakeholdereinbindung in der sozialwissenschaftlichen Forschung*. Springer VS, Wiesbaden, S.33-50.
- Niederberger, Marlen (2015b): Das Gruppendelphi. In: Niederberger, M.; Wassermann, S. (Hrsg.): *Methoden der Experten- und Stakeholdereinbindung in der sozialwissenschaftlichen Forschung*. Springer VS, Wiesbaden, S.117-140.
- Niederberger, Marlen; Wassermann, Sandra (2015): Einleitung. In: Niederberger, M.; Wassermann, S. (Hrsg.): *Methoden der Experten- und Stakeholdereinbindung in der sozialwissenschaftlichen Forschung*. Springer VS, Wiesbaden, S.11-14.
- Niederberger, Marlen; Renn, Ortwin (2018): *Das Gruppendelphi-Verfahren. Vom Konzept bis zur Anwendung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- NISTEP (2010): *The 9th Science and Technology Foresight – Contribution of Science and Technology to Future Society – The 9th Delphi Survey (Summary)*. A report on study supported by Special Coordination Funds for Promoting Science and Technology, FY2009. NISTEP report No.140.
- Okoli, Chitu; Pawlowski, Suzanne D. (2004): The Delphi Method as a Research Tool: An Example, Design Considerations and Applications. In: *Information and Management* (42/1): S. 15 – 29.
- Pinnekamp, Johannes; Adomßent, Maik; Bastian, Daniel; Bäumlisberger, Mathias; Dreyer, Marion; Esser-Schmittmann, Wolfgang; Fiehn, Rainer; Firk, Julia; Gebhardt, Wilhelm; Herrel, Markus; Herrmann, Manuel; John, Steffen; Jonas, Daniel; Kerber, Heide; Kuhn, Rainer; Kümmerer, Klaus; Meyer, Andreas; Lyko, Sven; Montag, David; Mousel, Danièle; Nafo, Issa; Olsson, Oliver; Palmowski, Laurence; Scheele, Sandra; Schmidt, Bea; Schramm, Engelbert; Steube, Tobias; Winker, Martina (2015): Bewertung und Minderung der Emissionen aus Einrichtungen des Gesundheitswesens in die aquatische Umwelt. In: Pinnekamp, J.; Palmowski, L.; Kümmerer, K.; Schramm, E. (Hrsg.): *Abwasser aus Einrichtungen des Gesundheitswesens – Charakterisierung, Technologien, Kommunikation und Konzepte*. Aachen: Gewässerschutz, Wasser, Abwasser (238), S. 357-366.

- Pitz, Gordon F.; Sachs, Natalie J. (1984): Judgment and Decision: Theory and Application. In: Annual Review of Psychology (35): S. 139-163.
- Powell, Catherine (2003): The Delphi technique: myths and realities. In: Journal of Advanced Nursing (41/4), S. 376-382.
- Przyborski, Aglaja; Wohlrab-Sahr, Monika (2008): Qualitative Sozialforschung: Ein Arbeitsbuch. München: Oldenbourg Verlag.
- Quade, Ed S. (1970): Cost-Effectiveness: Some Trends in Analysis. A Report prepared for United States Air Force Project RAND. RAND Corporation, Santa Monica.
- Ramírez, Andrea; Blok, Kornelius; Hauschild, Michael; Wood, Richard; Hellweg, Stefanie; Antunes, Paula; Ciroth, Andreas (2013): Looking back at four years of research in PROSUITE. Lessons learned and challenges ahead. Online verfügbar: <http://46.105.145.85/web/guest/the-prosuite-framework>.
- Renn, Ortwin; Webler, Thomas (1994): Konfliktbewältigung durch Kooperation in der Umweltpolitik - theoretische Grundlagen und Handlungsvorschläge. In: oikos (Hrsg.): Kooperationen für die Umwelt - im Dialog zum Handeln. Chur: Rüegger, S. 11-52.
- Renn, Ortwin; Webler, Thomas (1998): Der kooperative Diskurs – Theoretische Grundlagen, Anforderungen, Möglichkeiten. In: Renn, O.; Kastenholz, H.; Schild, P.; Wilhelm, U. (Hrsg.): Abfallpolitik im kooperativen Diskurs. Bürgerbeteiligung bei der Standortsuche für eine Deponie im Kanton Aargau. ETH Zürich: Hochschulverlag AG, Zürich, S. 3-103.
- Renn, Ortwin (1999): Ethische Anforderungen an den Diskurs. In: Grunwald, A.; Saupe, S. (Hrsg.): Ethik in der Technikgestaltung. Praktische Relevanz und Legitimation. Springer, Berlin, Heidelberg, New York: Springer-Verlag, S. 63-94.
- Renn, Ortwin; Schrimpf, Monika; Büttner, Thomas; Carius, Rainer; Köberle, Sabine; Oppermann, Bettina; Schneider, Elke; Zöller, Katharina (1999): Abfallwirtschaft 2005. Bürger planen ein regionales Abfallkonzept. In: Renn, O.; Kastenholz, H.; Wachlin, K.D. (Hrsg.): Forum Kooperative Politik, Band 2. Baden-Baden: Nomos Verlag.
- Renn, Ortwin (2000a): The Challenge of Integrating Deliberation and Expertise: Participation and Discourse in Risk Management. Accepted Contribution to the 2000 Risk Symposium Volume: Risk and Governance. Edited by T.L. MaDaniels and M. Small. S. 1-83.
- Renn, Ortwin (2000b): Diskursive Verfahren der Technikfolgenabschätzung. In: Petermann, T.; Coenen, R. (Hrsg.): Technikfolgen-Abschätzung in Deutschland. Bilanz und Perspektiven. Campus Verlag, Frankfurt, New York, S. 115-130.
- Renn, Ortwin (2007): Diskursive Verfahren zur Lösung von Ziel- und Transformationskonflikten. In: Renn, O.; Deuschle, J.; Jäger, A.; Weimer-Jehle, W.: Leitbild Nachhaltigkeit. Eine normativ-funktionale Konzeption und ihre Umsetzung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 169-188.
- Renn, Ortwin (2008): Risk Governance. Coping with uncertainty in a complex world. Earthscan: London.
- Renn, Ortwin; Mack, Birgit; Tampe-Mai, Karolin; Heins, Bernd; Klöpffer, Carsten; Stelzer, Bertine (2012a): Soziale, öffentliche, ökonomische und politische Akzeptanz von Fördermaßnahmen zur Verbesserung der Effizienz im Bereich Elektrizitätsnutzung und von nachhaltiger Mobilität bei privaten Haushalten. Schlussbericht Teil 1 zum Teilprojekt Elektrizitätsnutzung: Smart Metering. Stuttgart, Oldenburg.

- Renn, Ortwin; Mack, Birgit; Tampe-Mai, Karolin; Heins, Bernd; Klöpfer, Carsten; Stelzer, Bertine (2012b): Soziale, öffentliche, ökonomische und politische Akzeptanz von Fördermaßnahmen zur Verbesserung der Effizienz im Bereich Elektrizitätsnutzung und von nachhaltiger Mobilität bei privaten Haushalten. Schlussbericht Teil 2 zum Teilprojekt Elektrizitätsnutzung: Eco-Fahrassistenten. Stuttgart, Oldenburg.
- Renn, Ortwin; Kuhn, Rainer; Sellke, Piet (2012c): Final Selection of Social Indicators. Report prepared within the EC 7th framework project PROSUITE (D 4.3), Online verfügbar: <http://46.105.145.85/web/guest/public-deliverables>
- Rescher, Nicholas (1969): Delphi and Values. A Report prepared for United States Air Force Project RAND. RAND Corporation, Santa Monica.
- Rowe, Gene; Wright, George; Bolger, Fergus (1991): Delphi, A reevaluation of research and theory. *Technological Forecasting and Social Change* (39/3), S.235-251.
- Ruddat, Michael; Beninghaus, Christina (2009): OSIRIS Projekt. In: Schulz, M.; Renn, O.: *Das Gruppendelphi: Konzept und Fragebogenkonstruktion*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 57-77.
- Sackman, Harold (1974): Delphi Assessment: Expert Opinion, Forecasting, and Group Response. A Report prepared for United States Air Force Project RAND. RAND Corporation, Santa Monica.
- Scheibe, M.¹⁵⁶; Skutch, Margaret; Schofer, Joseph L. (1975): Experiments in Delphi Methodology. In: Linstone, H.A.; Turoff, M. (Hrsg.): *The Delphi Method: Techniques and Applications*. London: Addison and Wesley, S.257-281.
- Schenk, Michael; Döbler, Thomas (1996): Das Expertendilemma – Ein Kommentar. In: Nennen, H-U.; Garbe, D.: *Das Expertendilemma – Zur Rolle wissenschaftlicher Gutachter in der öffentlichen Meinungsbildung*. Berlin, Heidelberg: Springer-Verlag, S.75-82.
- Schulz, Marlen; Renn, Ortwin (2009). *Das Gruppendelphi: Konzept und Fragebogenkonstruktion*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schulz, Marlen; Renn, Ortwin; Daschkeit, Achim (2009): WASKlim Projekt. In: Schulz, M.; Renn, O.: *Das Gruppendelphi: Konzept und Fragebogenkonstruktion*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 47-55.
- Schulz, Marlen; Wassermann, Sandra (2010): Expertenbewertung einer dauerhaften Speicherung von Kohlendioxid. Ergebnisse des Delphis über den Stand des Wissens zur Speicherung von Kohlendioxid. Kurzbericht mit den zentralen Ergebnissen eines vom BMBF geförderten Projektes, Universität Stuttgart. Online verfügbar: http://www.zirn-info.de/pdf/zirn_kurzbericht_delphi_ccs.pdf (Letzter Zugriff: 11.12.14)
- Schulz, Marlen (2012): Quick and easy?! Fokusgruppen in der angewandten Sozialwissenschaft. In: Schulz, M.; Mack, B.; Renn, O.: *Fokusgruppen in der empirischen Sozialwissenschaft. Von der Konzeption bis zur Auswertung*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 9-22.
- Schlemm, Annette: Mentale Modelle – Ein zweiter Beitrag. In: http://www.opentheory.org/kf_mentales_modell/text.phtml, zugegriffen am 02.04.2015.

¹⁵⁶ Die Vornamen werden in dem originalen Beitrag in *The Delphi Method: Techniques and Applications* nicht angegeben. Während die Vornamen von Skutch, Margaret und Schofer, Joseph L. dank Bibliografien recherchiert werden konnten, blieb der Vorname von Scheibe, M. für mich trotz intensiver Recherche unbekannt.

- Schweitzer, Pia-Johanna E. (2007): Diskursive Risikoregulierung. Diskurstheorie im Vergleich. Dissertationsschrift zur Erlangung der Würde eines Doktors der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.), Universität Stuttgart.
- Seeger, Thomas (1979): Die Delphimethode: Expertenbefragung zwischen Prognose und Gruppenmeinungsbildungsprozessen: Überprüft am Beispiel von Delphi-Befragungen im Gegenstandsbereich Informationen und Dokumentation. Freiburg: Hochschul Verlag, Dissertationsschrift.
- Sorkin, Robert D.; Luan, Shenghua; Itzkowitz, Jesse (2004): Group Decision and Deliberation: A Distributed Detection Process. In: Koehler, D.J.; Harvey, N. (Hrsg.): Blackwell Handbook of Judgment & Decision Making. Malden, Oxford, Carlton: Blackwell Publishing, S. 464-484.
- Urban, Dieter; Mayerl, Jochen (2008): Regressionsanalyse: Theorie, Technik und Anwendung. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Von der Gracht, Heiko A. (2012): Consensus measurement in Delphi studies. Review and implications for future quality assurance. In: Technological Forecasting and Social Change, 79, S. 1525-1536.
- Von Oertzen, Jürgen; Cuhls, Kerstin; Kimpeler, Simone (2006): Wie nutzen wir Informations- und Kommunikationstechnologien im Jahr 2020? Ergebnisse einer Delphibefragung. In: Fazit-Schriftenreihe, Forschungsbericht, Band 3, S. 7-117.
- Vorgrimler, Daniel; Wübben, Dirk (2003): Die Delphi-Methode und ihre Eignung als Prognoseinstrument. In: Wirtschaft und Statistik Nr.8. Statistisches Bundesamt, S. 763-774.
- Wassermann, Sandra (2015): Expertendilemma. In: Niederberger, M.; Wassermann, S. (Hrsg.): Methoden der Experten- und Stakeholdereinbindung in der sozialwissenschaftlichen Forschung. Wiesbaden: Springer VS, S.15-32.
- Webler, Thomas; Levine, Debra; Rakel, Horst; Renn, Ortwin (1991): The Group Delphi: A Novel Attempt at Reducing Uncertainty. Technological Forecasting and Social Change, 39(3), S. 253-263.
- Weinrich, Harald (1973): System, Diskurs, Didaktik und die Diktatur des Sitzfleisches. In: Maciejewski, F. (Hrsg.): Theorie der Gesellschaft und Sozialtechnologie, Theorie-Diskussion, Supplement 1. Frankfurt am Main: Suhrkamp, S. 145 - 161.
- Wittmann, Werner W. (2002): Brunswik-Symmetrie: Ein Schlüsselkonzept für erfolgreiche psychologische Forschung. In: Myrtek, M. (Hrsg.): Die Person im biologischen und sozialen Kontext. Göttingen: Hogrefe, S.163-186.
- Woudenberg, Fred (1991): An Evaluation of Delphi. In: Technological Forecasting and Social Change 40, S. 131-150.

Online:

- Homepage der Dialogik gemeinnützigen Gesellschaft für Kommunikations- und Kooperationsforschung mbH: <http://www.dialogik-expert.de>, zugegriffen am 24.11.2016.
- Homepage des ZIRIUS - Zentrum für Interdisziplinäre Risiko- und Innovationsforschung der Universität Stuttgart: <http://www.zirius.eu>, zugegriffen am 05.11.2018.

11 Anhang

11.1 Fragebogen der Vorbefragung des Gruppendelphis in SAUBER+



GRUPPENDELPHI


Vorbefragung zum 4. Stakeholder-Workshop

Abschnitt I: Handlungsbedarf bei verschiedenen Kontaminanten

Frage 1

Es gibt eine Vielzahl von Kontaminanten, die zu einer Belastung von Oberflächengewässern und von Grundwasser führen können. Ebenso bestehen für viele dieser Kontaminanten bereits Regulierungen und Maßnahmen. Sehen Sie zusätzlichen Handlungsbedarf?

Stellen Sie sich vor, die Bundesregierung stellt Ihnen ein Gesamtbudget von 10 Millionen Euro (fiktives Budget) zur Verfügung. Mit diesem Geld sollen neue Maßnahmen finanziert werden, um Gewässerbelastungen zu verringern. Ihre Aufgabe ist es, dieses Geld auf Maßnahmen zur Verringerung der Gewässerbelastung durch die Kontaminanten zu verteilen. Bitte tragen Sie jeweils die Summe ein, die Sie aus dem Budget von 10 Millionen Euro den unten aufgeführten Kontaminanten für neue Maßnahmen zuweisen möchten. Es zählt dabei nur Ihre Meinung unabhängig davon, ob Ihre Vorschläge politisch durchsetzbar wären oder nicht.

	Kontaminanten	Summe		Kontaminanten	Summe
01	Krankheitserreger	_____ €	06	Humanarzneimittel	_____ €
02	Multiresistente Keime	_____ €	07	Tierarzneimittel	_____ €
03	Nährstoffe	_____ €	08	Diagnostika	_____ €
04	Pestizide	_____ €	09	Organische Lösungsmittel	_____ €
05	Schwermetalle	_____ €	10	Streusalz	_____ €
			11	Sonstige:	 _____

Abschnitt II: Maßnahmen im Handlungsfeld „Arzneimittelentwicklung“

Frage 2

Im Weiteren geht es um **Humanarzneimittelwirkstoffe und Diagnostika** im Zusammenhang mit Gewässerbelastung. Wenn folgend (der Kürze halber) nur von „Arzneimitteln“ die Rede ist, sind immer Humanarzneimittelstoffe und Diagnostika gemeint.

Ein Handlungsfeld für Maßnahmen in Bezug auf Arzneimittelrückstände im Gewässerkreislauf ist die Arzneimittelentwicklung.

Ihre Aufgabe ist es nun anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesem Handlungsfeld zustimmen. Bitte tragen Sie jeweils einen Wert ein. Geben Sie 1 an, wenn Sie die Aussage ganz und gar ablehnen und 10, wenn Sie der Aussage voll und ganz zustimmen.

	Aussage	1	2	3	4	5	6	7	8	9	10
12	Bei der Zulassung von Humanarzneimitteln sollte das Umweltrisiko in der Nutzen-Risiko-Analyse des Arzneimittels berücksichtigt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	Forschungseinrichtungen und Arzneimittelhersteller sollten durch staatliche Förderprogramme bei der Entwicklung „Grüner Arzneimittelwirkstoffe“ gezielt unterstützt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	Eine Verlängerung von Patentlaufzeiten für „grüne Arzneimittelwirkstoffe“ sollte als Entwicklungsanreiz für eine „grüne Produktpolitik“ der Arzneimittelhersteller eingeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	Es sollte eine auf Umweltverträglichkeit bezogene Produktkennzeichnung von Medikamenten als Entwicklungsanreiz für eine „grüne Produktpolitik“ der Arzneimittelhersteller eingeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	Es sollten staatliche Programme aufgelegt werden, die Innovationsallianzen zwischen Pharma- und Biotechnologieunternehmen fördern.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abschnitt III: Maßnahmen im Handlungsfeld „Umgang mit Arzneimitteln“

Frage 3.1

Ein weiteres Handlungsfeld für Maßnahmen in Bezug auf Arzneimittelrückstände im Gewässerkreislauf ist der Umgang mit Arzneimitteln.

Ihre Aufgabe ist es nun anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesem Handlungsfeld zustimmen. Bitte tragen Sie jeweils einen Wert ein. Geben Sie wieder 1 an, wenn Sie die Aussage ganz und gar ablehnen und 10, wenn Sie der Aussage voll und ganz zustimmen.

	Aussage	1	2	3	4	5	6	7	8	9	10
17	Das Angebot an Arzneimitteln in variablen Packungsgrößen sollte erweitert und die Packungsgröße an die jeweilige Therapie besser angepasst werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	Ärzte sollten bei ähnlicher Wirksamkeit eher andere Darreichungsformen ¹⁵⁷ statt Salben verabreichen oder verschreiben, um den Wirkstoffeintrag durch Textil- und Körperwäsche zu mindern.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	Die Gesetzlichen Krankenkassen sollten die Teilnahme an gesundheitsfördernden Aktionen über ein erweitertes Bonuskartensystem verstärkt unterstützen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	Ein bundesweit einheitlicher und verbindlicher Standard für die Entsorgung unverbrauchter Arzneimittel über die Apotheken sollte geschaffen werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21	Ein bundesweit einheitlicher und verbindlicher Standard für die Entsorgung unverbrauchter Arzneimittel über die Apotheken, bei dem die nicht frei verkäuflichen Medikamente einem Pfandsystem unterliegen, sollte erprobt werden (der Arzneimittelnutzer erhält bei Rückgabe des angebrochenen Medikaments oder der Verpackung einen monetären Wert zurück, den er beim Kauf des Arzneimittels mit entrichtet hat).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

¹⁵⁷ Beispiele für andere Darreichungsformen sind Tabletten, Kapseln, Tropfen, Injektionslösungen und transdermale Pflaster.

Frage 3.2

Weitere Maßnahmenbereiche im Handlungsfeld „Umgang mit Arzneimitteln“ sind Informations-, Kommunikations- und Bildungsmaßnahmen.

Ihre Aufgabe ist es nun anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesen Bereichen zustimmen. Bitte tragen Sie jeweils einen Wert ein. Geben Sie wieder 1 an, wenn Sie die Aussage ganz und gar ablehnen und 10, wenn Sie der Aussage voll und ganz zustimmen.

	Aussage	1	2	3	4	5	6	7	8	9	10
22	Es sollte eine bundesweite Aufklärungskampagne zur Thematik „Arzneimittelrückstände im Gewässerkreislauf“ durchgeführt werden, die sowohl Fragen des Umgangs mit Arzneimitteln wie auch Fragen der Entsorgung unverbrauchter Arzneimittel behandelt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23	Es sollte eine bundesweite Aufklärungskampagne zur sachgerechten Entsorgung unverbrauchter Arzneimittel durchgeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24	Es sollte eine internetgestützte, öffentlich zugängliche Umweltklassifikation für Arzneimittel etabliert werden („schwedisches Modell“).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25	Beipackzettel von Arzneimitteln sollten mit einem Hinweis auf die Umweltverträglichkeit der Wirkstoffe versehen werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26	Eine Gesundheitsberatung, die Bewegung und Ernährung als Handlungsoptionen berücksichtigt, sollte als Standardleistung der gesetzlichen Krankenkassen eingeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27	Eine Pharma-Datenbank, die alle medizinischen Studien im Zuge einer Zulassung offen legt und detailliert über patientenindividuelle Dosis-Wirkungs-Verhältnisse informiert, sollte eingerichtet werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28	Es sollten gezielte Dialogangebote für Arzneimittelnutzer zum Thema Gewässer- und Umweltrelevanz von Arzneimitteln erfolgen; mögliche Anbieter könnten Apotheken, Verbraucherzentralen und Patientenorganisationen sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29	Präventive Aspekte und Ansätze sollten in der Ausbildung von Ärzten stärkeres Gewicht erhalten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30	Das Thema Gewässerbelastung durch Arzneimittelwirkstoffe sollte fester Bestandteil der Fort- und Weiterbildung von Ärzten und Apothekern werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31	Bundesweit sollten Apotheker und Ärzte zu „Antibiotikabeauftragten“ weitergebildet werden, die im Anschluss in Krankenhäusern als Berater tätig sind (Bsp. Niedersachsen).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abschnitt IV: Maßnahmen im Handlungsfeld „Emissionsmanagement“

Frage 4

Ein weiteres Handlungsfeld für Maßnahmen, die auf Arzneimittelrückstände im Gewässerkreislauf zielen, ist das Emissionsmanagement.

Ihre Aufgabe ist es anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesem Handlungsfeld zustimmen. Bitte tragen Sie jeweils einen Wert ein. Geben Sie wieder 1 an, wenn Sie die Aussage ganz und gar ablehnen und 10, wenn Sie der Aussage voll und ganz zustimmen.

	Aussage	1	2	3	4	5	6	7	8	9	10
32	Es sollte bedarfsorientiert eine vierte Reinigungsstufe in Kläranlagen eingerichtet werden, bei der erweiterte Verfahren der Abwasserreinigung zur Elimination von Arzneimittelstoffen und multiresistenten Keimen zum Einsatz kommen (analog der Schweiz).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33	Krankenhäuser und Spezialkliniken sollten ihr Abwasser generell vorbehandeln und erst nach dieser Vorbehandlung in das kommunale Kanalnetz einleiten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34	Gesundheitseinrichtungen mit vergleichsweise hohen Frachten an problematischen Stoffen im Abwasser sollten eine getrennte Erfassung und Behandlung des Toilettenabwassers vor seiner Einleitung in das kommunale Kanalnetz vornehmen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35	Die separate Erfassung, spezielle Behandlung und Entsorgung von mit Diagnostika belasteten Abwasserströmen (wie z.B. Urin) sollte über weitere Pilotprojekte erprobt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36	Alle Wasserwerke, die ihr Trinkwasser aus Uferfiltrat, Flusswasser oder angereichertem Grundwasser gewinnen, sollten erweiterte Reinigungstechniken wie beispielsweise die Technik der Aktivkohlefiltration einsetzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
37	Die Liste prioritärer Stoffe der EU-Wasserrahmenrichtlinie sollte um multiresistente Keime erweitert werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
38	Kommunen bzw. Kanalnetzbetreiber sollten das Abwasserkanalnetz so sanieren, dass die Menge an Wasser, die durch undichte Abwasserkanäle austritt, verringert wird.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
39	Die novellierte deutsche Klärschlammverordnung sollte ein generelles Verbot des Aufbringens von Klärschlamm auf Lebensmittelböden (Anbauflächen für Nahrungs- und Futtermittel) umfassen. Für Energieböden (nachwachsende Rohstoffe) sollten andere Bestimmungen gelten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abschnitt V: Forschungsförderung in Bezug auf Arzneimittelrückstände im Gewässerkreislauf

Frage 5

Stellen Sie sich bitte vor, die Bundesregierung stellt Ihnen 100 Millionen Euro (fiktives Budget) für Forschungsförderprogramme zur Verfügung. Mit diesen Programmen soll die Forschung im Bereich „Arzneimittelrückstände im Gewässerkreislauf“ vorangetrieben werden.

Ihre Aufgabe ist es, dieses Geld auf die unten angeführten Forschungsfelder zu verteilen. Bitte tragen Sie jeweils die Summe ein, die Sie aus dem Budget von 100 Millionen Euro den unten genannten Forschungsfeldern zuteilen möchten. Es zählt dabei nur Ihre Meinung, unabhängig davon, ob Ihre Vorschläge politisch durchgesetzt werden könnten oder nicht.

	Forschungsfeld	Summe
40	Forschung und Entwicklung im Bereich umweltverträgliche Arzneimittelwirkstoffe	_____ €
41	Forschung- und Entwicklung im Bereich Verlustminderung pharmakologisch aktiver Substanzen, z.B. durch Optimierung der Formulierung ¹⁵⁸ und damit teilweise verbundene Änderung der Darreichungsform	_____ €
42	Forschung zum konsequenten Befolgen ärztlicher Therapien durch die Patienten (Gute Compliance)	_____ €
43	Forschung und Entwicklung im Bereich nicht-medikamentöser Behandlungsformen	_____ €
44	Forschung zur Entstehung, Verbreitung und Weitergabe von Antibiotika-Resistenzen, multiresistenten Keimen und Antibiotikaresistenzeigenschaften	_____ €
45	Forschung zu Abbau- und Transformationsprodukten von Arzneimittelwirkstoffen und ihrer Metaboliten	_____ €


¹⁵⁸ Der Begriff Formulierung bezeichnet die Zusammensetzung einer Darreichungsform (beispielsweise kann eine Tablette als Hilfsmittel Füllmittel wie Lactose oder Cellulose enthalten).

Abschnitt VI: Dringlichkeit des Handlungsbedarfs nach Eintragsquellen

Frage 6

Wie beurteilen Sie die Dringlichkeit des Handlungsbedarfs bei verschiedenen Eintragsquellen? „Handlungsbedarf“ kann sich auf alle der oben angeführten Maßnahmen einschließlich Forschungsförderung und weitere mögliche Maßnahmen beziehen.

Geben Sie 1 an, wenn Sie der Ansicht sind, dass bei der jeweiligen Eintragsquelle kein Handlungsbedarf besteht, geben Sie 10 an, wenn Sie der Ansicht sind, dass bei der jeweiligen Eintragsquelle hoher Handlungsbedarf besteht.

Eintragsquellen		1	2	3	4	5	6	7	8	9	10
46	Arzneimittelproduktionsstätten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
47	Allgemeine Krankenhäuser	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
48	Spezielle Einrichtungen des Gesundheitswesens z.B. Ärztehäuser, psychiatrische Kliniken, Sportkliniken etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
49	Pflegeeinrichtungen und Betreuungsheime, Seniorenresidenzen, Hospize etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
50	Privathaushalte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
51	Weitere Quellen (z.B. undichte Kanäle, Klärschlamm, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
52	Sonstiges: 										

Abschnitt VII: Anmerkungen

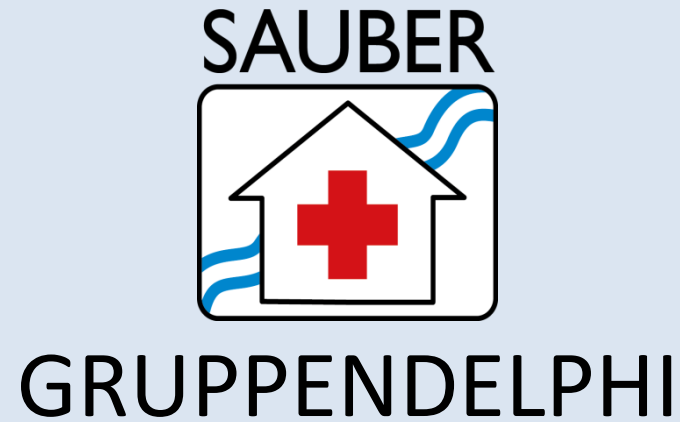
Frage 7

Falls Sie Ideen und Anmerkungen zum Fragebogen haben, würden wir uns freuen, wenn Sie uns diese mitteilen. Damit helfen sie uns bei der Weiterentwicklung des Fragebogens. Der so angepasste Fragebogen kommt dann im Rahmen des 4. Stakeholder-Workshops zur Anwendung.



Vielen Dank für die Beantwortung des Fragebogens

11.2 Fragebogen des Gruppendelphis in SAUBER+



Vierter Stakeholder-Workshop Erste Kleingruppenphase




Abschnitt I: Handlungsbedarf bei verschiedenen Kontaminanten

Frage 1

Es gibt eine Vielzahl von Kontaminanten, die zu einer Belastung von Oberflächengewässern und von Grundwasser führen können. Ebenso bestehen für viele dieser Kontaminanten bereits Regulierungen und Maßnahmen. Sehen Sie zusätzlichen Handlungsbedarf?

Stellen Sie sich vor, die Bundesregierung stellt Ihnen ein Gesamtbudget von 10 Millionen Euro (fiktives Budget) zur Verfügung. Mit diesem Geld sollen neue Maßnahmen finanziert werden, um Gewässerbelastungen zu verringern. Ihre Aufgabe ist es, dieses Geld auf Maßnahmen zur Verringerung der Gewässerbelastung durch die Kontaminanten zu verteilen. Bitte tragen Sie jeweils die Summe ein, die Sie aus dem Budget von 10 Millionen Euro den unten aufgeführten Kontaminanten für neue Maßnahmen zuweisen möchten. Es zählt dabei nur Ihre Meinung unabhängig davon, ob Ihre Vorschläge politisch durchsetzbar wären oder nicht.

	Kontaminanten	Summe		Kontaminanten	Summe
01	Krankheitserreger	_____ €	06	Humanarzneimittel	_____ €
02	Multiresistente Keime	_____ €	07	Tierarzneimittel	_____ €
03	Nährstoffe	_____ €	08	Diagnostika	_____ €
04	Pestizide	_____ €	09	Organische Lösungsmittel	_____ €
05	Schwermetalle	_____ €	10	Streusalz	_____ €
			11	Sonstige:	 _____

Abschnitt II: Maßnahmen im Handlungsfeld „Arzneimittelentwicklung“

Frage 2

Im Weiteren geht es um **Humanarzneimittelwirkstoffe und Diagnostika** im Zusammenhang mit Gewässerbelastung. Wenn folgend (der Kürze halber) nur von „Arzneimitteln“ die Rede ist, sind immer Humanarzneimittelstoffe und Diagnostika gemeint.

Ein Handlungsfeld für Maßnahmen in Bezug auf Arzneimittelrückstände im Gewässerkreislauf ist die Arzneimittelentwicklung.

Ihre Aufgabe ist es nun anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesem Handlungsfeld zustimmen. Bitte tragen Sie jeweils einen Wert ein. Geben Sie 1 an, wenn Sie die Aussage ganz und gar ablehnen und 10, wenn Sie der Aussage voll und ganz zustimmen.

	Aussage	1	2	3	4	5	6	7	8	9	10
12	Bei der Zulassung von Humanarzneimitteln sollte das Umweltrisiko in der Nutzen-Risiko-Analyse des Arzneimittels berücksichtigt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	Forschungseinrichtungen und Arzneimittelhersteller sollten durch staatliche Förderprogramme bei der Entwicklung „Grüner Arzneimittelwirkstoffe“ gezielt unterstützt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	Eine Verlängerung von Patentlaufzeiten für „grüne Arzneimittelwirkstoffe“ sollte als Entwicklungsanreiz für eine „grüne Produktpolitik“ der Arzneimittelhersteller eingeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	Es sollte eine auf Umweltverträglichkeit bezogene Produktkennzeichnung von Medikamenten als Entwicklungsanreiz für eine „grüne Produktpolitik“ der Arzneimittelhersteller eingeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16	Es sollten staatliche Programme aufgelegt werden, die Innovationsallianzen zwischen Pharma- und Biotechnologieunternehmen fördern.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abschnitt III: Maßnahmen im Handlungsfeld „Umgang mit Arzneimitteln“

Frage 3.1

Ein weiteres Handlungsfeld für Maßnahmen in Bezug auf Arzneimittelrückstände im Gewässerkreislauf ist der Umgang mit Arzneimitteln.

Ihre Aufgabe ist es nun anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesem Handlungsfeld zustimmen. Bitte tragen Sie jeweils einen Wert ein. Geben Sie wieder 1 an, wenn Sie die Aussage ganz und gar ablehnen und 10, wenn Sie der Aussage voll und ganz zustimmen.

	Aussage	1	2	3	4	5	6	7	8	9	10
17	Das Angebot an Arzneimitteln in variablen Packungsgrößen sollte erweitert und die Packungsgröße an die jeweilige Therapie besser angepasst werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	Ärzte sollten bei ähnlicher Wirksamkeit eher andere Darreichungsformen ¹⁵⁹ statt Salben verabreichen oder verschreiben, um den Wirkstoffeintrag durch Textil- und Körperwäsche zu mindern.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	Die Gesetzlichen Krankenkassen sollten die Teilnahme an gesundheitsfördernden Aktionen über ein erweitertes Bonuskartensystem verstärkt unterstützen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	Ein bundesweit einheitlicher und verbindlicher Standard für die Entsorgung unverbrauchter Arzneimittel über die Apotheken sollte geschaffen werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21	Ein bundesweit einheitlicher und verbindlicher Standard für die Entsorgung unverbrauchter Arzneimittel über die Apotheken, bei dem die nicht frei verkäuflichen Medikamente einem Pfandsystem unterliegen, sollte erprobt werden (der Arzneimittelnutzer erhält bei Rückgabe des angebrochenen Medikaments oder der Verpackung einen monetären Wert zurück, den er beim Kauf des Arzneimittels mit entrichtet hat).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

¹⁵⁹ Beispiele für andere Darreichungsformen sind Tabletten, Kapseln, Tropfen, Injektionslösungen und transdermale Pflaster.

Frage 3.2

Weitere Maßnahmenbereiche im Handlungsfeld „Umgang mit Arzneimitteln“ sind Informations-, Kommunikations- und Bildungsmaßnahmen.

Ihre Aufgabe ist es nun anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesen Bereichen zustimmen. Bitte tragen Sie jeweils einen Wert ein. Geben Sie wieder 1 an, wenn Sie die Aussage ganz und gar ablehnen und 10, wenn Sie der Aussage voll und ganz zustimmen.

	Aussage	1	2	3	4	5	6	7	8	9	10
22	Es sollte eine bundesweite Aufklärungskampagne zur Thematik „Arzneimittelrückstände im Gewässerkreislauf“ durchgeführt werden, die sowohl Fragen des Umgangs mit Arzneimitteln wie auch Fragen der Entsorgung unverbrauchter Arzneimittel behandelt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23	Es sollte eine bundesweite Aufklärungskampagne zur sachgerechten Entsorgung unverbrauchter Arzneimittel durchgeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24	Es sollte eine internetgestützte, öffentlich zugängliche Umweltklassifikation für Arzneimittel etabliert werden („schwedisches Modell“).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25	Beipackzettel von Arzneimitteln sollten mit einem Hinweis auf die Umweltverträglichkeit der Wirkstoffe versehen werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26	Eine Gesundheitsberatung, die Bewegung und Ernährung als Handlungsoptionen berücksichtigt, sollte als Standardleistung der gesetzlichen Krankenkassen eingeführt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27	Eine Pharma-Datenbank, die alle medizinischen Studien im Zuge einer Zulassung offen legt und detailliert über patientenindividuelle Dosis-Wirkungs-Verhältnisse informiert, sollte eingerichtet werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28	Es sollten gezielte Dialogangebote für Arzneimittelnutzer zum Thema Gewässer- und Umweltrelevanz von Arzneimitteln erfolgen; mögliche Anbieter könnten Apotheken, Verbraucherzentralen und Patientenorganisationen sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29	Präventive Aspekte und Ansätze sollten in der Ausbildung von Ärzten stärkeres Gewicht erhalten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30	Das Thema Gewässerbelastung durch Arzneimittelwirkstoffe sollte fester Bestandteil der Fort- und Weiterbildung von Ärzten und Apothekern werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31	Bundesweit sollten Apotheker und Ärzte zu „Antibiotikabeauftragten“ weitergebildet werden, die im Anschluss in Krankenhäusern als Berater tätig sind (Bsp. Niedersachsen).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abschnitt IV: Maßnahmen im Handlungsfeld „Emissionsmanagement“

Frage 4

Ein weiteres Handlungsfeld für Maßnahmen, die auf Arzneimittelrückstände im Gewässerkreislauf zielen, ist das Emissionsmanagement.

Ihre Aufgabe ist es anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesem Handlungsfeld zustimmen. Bitte tragen Sie jeweils einen Wert ein. Geben Sie wieder 1 an, wenn Sie die Aussage ganz und gar ablehnen und 10, wenn Sie der Aussage voll und ganz zustimmen.

	Aussage	1	2	3	4	5	6	7	8	9	10
32	Es sollte bedarfsorientiert eine vierte Reinigungsstufe in Kläranlagen eingerichtet werden, bei der erweiterte Verfahren der Abwasserreinigung zur Elimination von Arzneimittelstoffen und multiresistenten Keimen zum Einsatz kommen (analog der Schweiz).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33	Krankenhäuser und Spezialkliniken sollten ihr Abwasser generell vorbehandeln und erst nach dieser Vorbehandlung in das kommunale Kanalnetz einleiten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34	Gesundheitseinrichtungen mit vergleichsweise hohen Frachten an problematischen Stoffen im Abwasser sollten eine getrennte Erfassung und Behandlung des Toilettenabwassers vor seiner Einleitung in das kommunale Kanalnetz vornehmen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35	Die separate Erfassung, spezielle Behandlung und Entsorgung von mit Diagnostika belasteten Abwasserströmen (wie z.B. Urin) sollte über weitere Pilotprojekte erprobt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36	Alle Wasserwerke, die ihr Trinkwasser aus Uferfiltrat, Flusswasser oder angereichertem Grundwasser gewinnen, sollten erweiterte Reinigungstechniken wie beispielsweise die Technik der Aktivkohlefiltration einsetzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
37	Die Liste prioritärer Stoffe der EU-Wasserrahmenrichtlinie sollte um multiresistente Keime erweitert werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
38	Kommunen bzw. Kanalnetzbetreiber sollten das Abwasserkanalnetz so sanieren, dass die Menge an Wasser, die durch undichte Abwasserkanäle austritt, verringert wird.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
39	Die novellierte deutsche Klärschlammverordnung sollte ein generelles Verbot des Aufbringens von Klärschlamm auf Lebensmittelböden (Anbauflächen für Nahrungs- und Futtermittel) umfassen. Für Energieböden (nachwachsende Rohstoffe) sollten andere Bestimmungen gelten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abschnitt V: Forschungsförderung in Bezug auf Arzneimittelrückstände im Gewässerkreislauf

Frage 5

Stellen Sie sich bitte vor, die Bundesregierung stellt Ihnen 100 Millionen Euro (fiktives Budget) für Forschungsförderprogramme zur Verfügung. Mit diesen Programmen soll die Forschung im Bereich „Arzneimittelrückstände im Gewässerkreislauf“ vorangetrieben werden.

Ihre Aufgabe ist es, dieses Geld auf die unten angeführten Forschungsfelder zu verteilen. Bitte tragen Sie jeweils die Summe ein, die Sie aus dem Budget von 100 Millionen Euro den unten genannten Forschungsfeldern zuteilen möchten. Es zählt dabei nur Ihre Meinung, unabhängig davon, ob Ihre Vorschläge politisch durchgesetzt werden könnten oder nicht.

	Forschungsfeld	Summe
40	Forschung und Entwicklung im Bereich umweltverträgliche Arzneimittelwirkstoffe	_____ €
41	Forschung- und Entwicklung im Bereich Verlustminderung pharmakologisch aktiver Substanzen, z.B. durch Optimierung der Formulierung ¹⁶⁰ und damit teilweise verbundene Änderung der Darreichungsform	_____ €
42	Forschung zum konsequenten Befolgen ärztlicher Therapien durch die Patienten (Gute Compliance)	_____ €
43	Forschung und Entwicklung im Bereich nicht-medikamentöser Behandlungsformen	_____ €
44	Forschung zur Entstehung, Verbreitung und Weitergabe von Antibiotika-Resistenzen, multiresistenten Keimen und Antibiotikaresistenzeigenschaften	_____ €
45	Forschung zu Abbau- und Transformationsprodukten von Arzneimittelwirkstoffen und ihrer Metaboliten	_____ €


¹⁶⁰ Der Begriff Formulierung bezeichnet die Zusammensetzung einer Darreichungsform (beispielsweise kann eine Tablette als Hilfsmittel Füllmittel wie Lactose oder Cellulose enthalten).

Abschnitt VI: Dringlichkeit des Handlungsbedarfs nach Eintragsquellen

Frage 6

Wie beurteilen Sie die Dringlichkeit des Handlungsbedarfs bei verschiedenen Eintragsquellen? „Handlungsbedarf“ kann sich auf alle der oben angeführten Maßnahmen einschließlich Forschungsförderung und weitere mögliche Maßnahmen beziehen.

Geben Sie 1 an, wenn Sie der Ansicht sind, dass bei der jeweiligen Eintragsquelle kein Handlungsbedarf besteht, geben Sie 10 an, wenn Sie der Ansicht sind, dass bei der jeweiligen Eintragsquelle hoher Handlungsbedarf besteht.

Eintragsquellen		1	2	3	4	5	6	7	8	9	10
46	Arzneimittelproduktionsstätten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
47	Allgemeine Krankenhäuser	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
48	Spezielle Einrichtungen des Gesundheitswesens z.B. Ärztehäuser, psychiatrische Kliniken, Sportkliniken etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
49	Pflegeeinrichtungen und Betreuungsheime, Seniorenresidenzen, Hospize etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
50	Privathaushalte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
51	Weitere Quellen (z.B. undichte Kanäle, Klärschlamm, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
52	Sonstiges: 										

Abschnitt VII: Wirksamkeit und Umsetzbarkeit ausgewählter Maßnahmen

Im Folgenden finden Sie jeweils sechs Maßnahmen aus den in den Abschnitten II-IV abgefragten Handlungsfeldern. Aus jedem Handlungsfeld („Arzneimittelentwicklung“, „Umgang mit Arzneimitteln“ und „Emissionsmanagement“) wurden die in der Vorbefragung jeweils zwei am besten bewerteten Maßnahmen ausgewählt.

Ihre Aufgabe ist es nun anzugeben, inwieweit Sie den folgenden Aussagen über Maßnahmen aus diesem Handlungsfeld zustimmen. Hierbei unterscheiden Sie bitte zwischen der Wirksamkeit und der Umsetzbarkeit der Maßnahme. Geben Sie 1 an, wenn Sie die Wirksamkeit bzw. Umsetzbarkeit der Maßnahme als sehr gering einschätzen und 10, wenn Sie die Wirksamkeit bzw. Umsetzbarkeit der Maßnahme als sehr hoch einschätzen.

Aussage	#	Wirksamkeit der Maßnahme										#	Umsetzbarkeit der Maßnahme									
		1	2	3	4	5	6	7	8	9	10		1	2	3	4	5	6	7	8	9	10
Bei der Zulassung von Humanarzneimitteln sollte das Umweltrisiko in der Nutzen-Risiko-Analyse des Arzneimittels berücksichtigt werden.	53	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	54	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Forschungseinrichtungen und Arzneimittelhersteller sollten durch staatliche Förderprogramme bei der Entwicklung „Grüner Arzneimittelwirkstoffe“ gezielt unterstützt werden.	55	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	56	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ein bundesweit einheitlicher und verbindlicher Standard für die Entsorgung unverbrauchter Arzneimittel über die Apotheken sollte geschaffen werden.	57	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	58	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Es sollte eine bundesweite Aufklärungskampagne zur sachgerechten Entsorgung unverbrauchter Arzneimittel durchgeführt werden.	59	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	60	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gesundheitseinrichtungen mit vergleichsweise hohen Frachten an problematischen Stoffen im Abwasser sollten eine getrennte Erfassung und Behandlung des Toilettenabwassers vor seiner Einleitung in das kommunale Kanalnetz vornehmen.	61	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	62	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die separate Erfassung, spezielle Behandlung und Entsorgung von mit Diagnostika belasteten Abwasserströmen (wie z.B. Urin) sollte über weitere Pilotprojekte erprobt werden.	63	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	64	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vielen Dank für die Beantwortung des Fragebogens

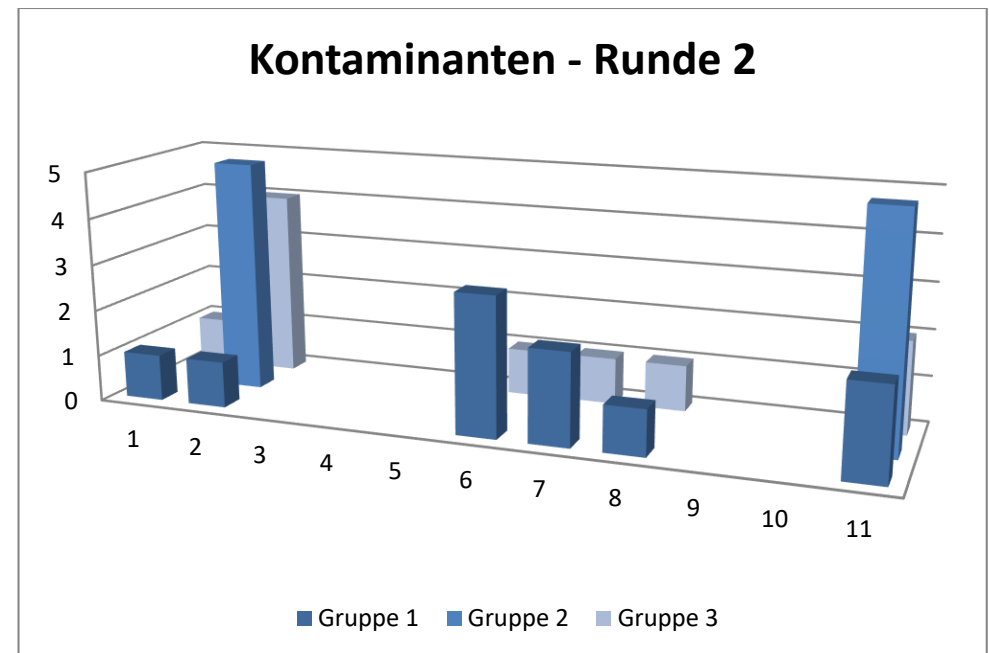
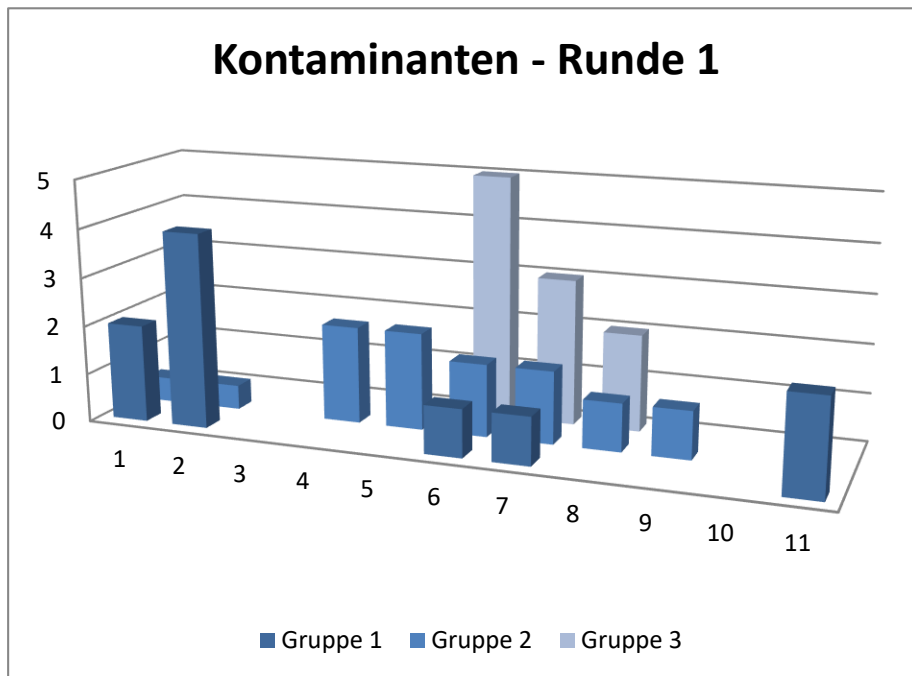
11.3 Statistische Ergebnisse des Gruppendelphis in SAUBER+

Ergebnisse: Erste Kleingruppenphase											
#	Frageinhalt	Anzahl der Antworten	Gruppe 1	MV 1	Gruppe 2	MV 2	Gruppe 3	MV 3	x	s	VKoeff.
12	Zulassung von Humanarzneimitteln	4	9	1	1	-	8	-	4,8	3,8	0,8
13	Forschungseinrichtungen und Arzneimittelhersteller	3	6	-	6	-	9	-	7,0	1,4	0,2
14	Eine Verlängerung von Patentlaufzeiten	3	6	-	6	-	10	-	7,3	1,9	0,3
15	Umweltverträglichkeit bezogene Produktkennzeichnung	4	9	1	1	-	10	-	5,3	4,3	0,8
16	Staatliche Programme	3	4	-	8	-	10	-	7,3	2,5	0,3
17	Angebot an Arzneimitteln in variablen Packungsgrößen	3	10	-	3	-	3	-	5,3	3,3	0,6
18	Ähnliche Wirksamkeit eher andere Darreichungsformen	4	7	1	4	-	2	-	3,5	2,3	0,7
19	Gesetzlichen Krankenkassen	3	9	-	10	-	10	-	9,7	0,5	0,0
20	Bundesweit einheitlicher und verbindlicher Standard: Entsorgung	3	5	-	10	-	10	-	8,3	2,4	0,3
21	Bundesweit einheitlicher und verbindlicher Standard: Pfand	3	1	-	1	-	10	-	4,0	4,2	1,1
22	Bundesweite Aufklärungskampagne zur Thematik „Arzneimittelrückstände im Gewässerkreislauf“	3	8	-	10	-	10	-	9,3	0,9	0,1
23	Bundesweite Aufklärungskampagne zur sachgerechten Entsorgung unverbrauchter Arzneimittel	3	9	-	10	-	10	-	9,7	0,5	0,0
24	Umweltklassifikation für Arzneimittel	4	10	5	7	-	10	-	8,0	2,1	0,3
25	Beipackzettel von Arzneimitteln	4	10	5	7	-	6	-	7,0	1,9	0,3
26	Gesundheitsberatung	3	10	-	10	-	10	-	10,0	0,0	0,0
27	Pharma-Datenbank	2	7	-	-	-	10	-	8,5	1,5	0,2
28	Dialogangebote für Arzneimittelnutzer	3	8	-	8	-	4	-	6,7	1,9	0,3
29	Präventive Aspekte und Ansätze	3	8	-	10	-	10	-	9,3	0,9	0,1
30	Gewässerbelastung durch Arzneimittelwirkstoffe sollte fester Bestandteil der Fort- und Weiterbildung	3	8	-	10	-	10	-	9,3	0,9	0,1
31	Antibiotikabeauftragte	3	10	-	8	-	1	-	6,3	3,9	0,6
32	Bedarfsorientiert eine vierte Reinigungsstufe	4	8	2	10	-	10	-	7,5	3,3	0,4

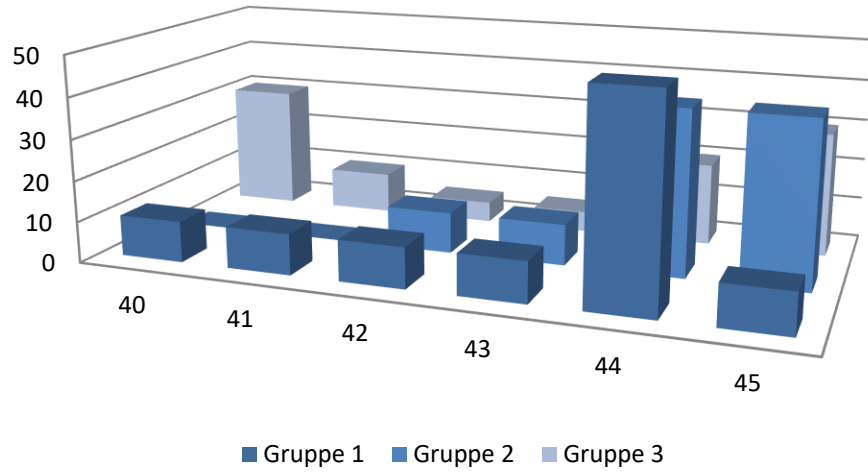
33	Krankenhäuser und Spezialkliniken sollten ihr Abwasser generell vorbehandeln	4	3	10	1	-	9	-	5,8	3,8	0,7
34	Gesundheitseinrichtungen mit vergleichsweise hohen Frachten	3	8	-	1	-	2	-	3,7	3,1	0,8
35	Separate Erfassung, spezielle Behandlung und Entsorgung	3	9	-	10	-	10	-	9,7	0,5	0,0
36	Wasserwerke, die ihr Trinkwasser aus Uferfiltrat, Flusswasser oder angereichertem Grundwasser gewinnen	3	3	-	3	-	10	-	5,3	3,3	0,6
37	Liste prioritärer Stoffe der EU-Wasserrahmenrichtlinie	3	2	-	1	-	9	-	4,0	3,6	0,9
38	Kommunen bzw. Kanalnetzbetreiber sollten das Abwasserkanalnetz	3	8	-	10	-	10	-	9,3	0,9	0,1
39	Novellierte deutsche Klärschlammverordnung	3	7	-	5	-	3	-	5,0	1,6	0,3
46	Arzneimittelproduktionsstätten	3	2	-	8	-	3	-	4,3	2,6	0,6
47	Allgemeine Krankenhäuser	3	5	-	3	-	7	-	5,0	1,6	0,3
48	Spezielle Einrichtungen des Gesundheitswesens z.B. Ärzthäuser, psychiatrische Kliniken, Sportkliniken etc.	3	7	-	2	-	7	-	5,3	2,4	0,4
49	Pflegeeinrichtungen und Betreuungsheime, Seniorenresidenzen, Hospize etc.	3	7	-	1	-	7	-	5,0	2,8	0,6
50	Privathaushalte	3	7	-	1	-	10	-	6,0	3,7	0,6
51	Weitere Quellen (z.B. undichte Kanäle, Klärschlamm, etc.)	3	5	-	10	-	10	-	8,3	2,4	0,3
52	Sonstiges	Nennung: Forschungseinrichtungen									
53	Wirksamkeit: Nutzen-Risiko-Analyse des Arzneimittels	4	6	1	1	-	9	-	4,3	3,4	0,8
54	Umsetzbarkeit: Nutzen-Risiko-Analyse des Arzneimittels	4	9	3	1	-	7	-	5,0	3,2	0,6
55	Wirksamkeit: Grüne Arzneimittelwirkstoffe	3	5	-	1	-	6	-	4,0	2,2	0,5
56	Umsetzbarkeit: Grüne Arzneimittelwirkstoffe	3	5	-	10	-	9	-	8,0	2,2	0,3
57	Wirksamkeit: Entsorgung unverbrauchter Arzneimittel	3	5	-	10	-	10	-	8,3	2,4	0,3
58	Umsetzbarkeit: Entsorgung unverbrauchter Arzneimittel	3	2	-	10	-	10	-	7,3	3,8	0,5
59	Wirksamkeit: Aufklärungskampagne zur sachgerechten Entsorgung	3	7	-	7	-	6	-	6,7	0,5	0,1
60	Umsetzbarkeit: Aufklärungskampagne zur sachgerechten Entsorgung	3	10	-	10	-	10	-	10,0	0,0	0,0
61	Wirksamkeit: Getrennte Erfassung und Behandlung	3	8	-	2	-	10	-	6,7	3,4	0,5
62	Umsetzbarkeit: Getrennte Erfassung und Behandlung	4	9	4	5	-	6	-	6,0	1,9	0,3
63	Wirksamkeit: Pilotprojekte	2	8	-	-	-	10	-	9,0	1,0	0,1
64	Umsetzbarkeit: Pilotprojekte	3	8	-	10	-	10	-	9,3	0,9	0,1

Ergebnisse: Zweite Kleingruppenphase											
#	Frageinhalt	Anzahl der Antworten	Gruppe 1	MV 1	Gruppe 2	MV 2	Gruppe 3	MV 3	x	s	VKoeff.
12	Zulassung von Humanarzneimitteln	3	1	-	1	-	2	-	1,3	0,5	0,4
15	Umweltverträglichkeit bezogene Produktkennzeichnung	3	5	-	2	-	1	-	2,7	1,7	0,6
16	Staatliche Programme	3	4	-	8	-	10	-	7,3	2,5	0,3
20	Bundesweit einheitlicher und verbindlicher Standard: Entsorgung	4	10	-	1	10	10	-	7,8	3,9	0,5
21	Bundesweit einheitlicher und verbindlicher Standard: Pfand	4	1	-	1	10	1	-	3,3	3,9	1,2
31	Antibiotikabeauftragte	3	8	-	10	-	10	-	9,3	0,9	0,1
32	Bedarfsorientiert eine vierte Reinigungsstufe	3	10	-	10	-	10	-	10	0	0
33	Krankenhäuser und Spezialkliniken sollten ihr Abwasser bedarfsorientiert vorbehandeln (<i>Formulierung neu</i>)	3	10	-	8	-	5	-	7,7	2,1	0,3
37	Multiresistente Keime (<i>Formulierung neu</i>)	3	5	-	1	-	10	-	5,3	3,7	0,7
39	Novellierte deutsche Klärschlammverordnung (<i>Formulierung neu</i>)	3	10	-	5	-	7	-	7,3	2,1	0,3
46	Arzneimittelproduktionsstätten	3	3	-	1	-	8	-	4,0	2,9	0,7
47	Allgemeine Krankenhäuser	3	5	-	7	-	8	-	6,7	1,2	0,2
48	Spezielle Einrichtungen des Gesundheitswesens z.B. Ärztehäuser, psychiatrische Kliniken, Sportkliniken etc.	3	7	-	7	-	8	-	7,3	0,5	0,1
49	Pflegeeinrichtungen und Betreuungsheime, Seniorenresidenzen, Hospize etc.	3	7	-	7	-	8	-	7,3	0,5	0,1
50	Privathaushalte	3	7	-	10	-	8	-	8,3	1,2	0,1
51	Weitere Quellen (z.B. undichte Kanäle, Klärschlamm, etc.)	3	5	-	8	-	10	-	7,7	2,1	0,3
52	Forschungseinrichtungen	4	2	8	10	-	8	-	7,0	3,0	0,4
53	Wirksamkeit: Nutzen-Risiko-Analyse des Arzneimittels	3	10	-	1	-	8	-	6,3	3,9	0,6
54	Umsetzbarkeit: Nutzen-Risiko-Analyse des Arzneimittels	3	1	-	1	-	1	-	1,0	0,0	0,0
59	Wirksamkeit: Aufklärungskampagne zur sachgerechten Entsorgung	3	5	-	7	-	5	-	5,7	0,9	0,2
60	Umsetzbarkeit: Aufklärungskampagne zur sachgerechten Entsorgung	3	9	-	10	-	10	-	9,7	0,2	0,0
63	Wirksamkeit: Pilotprojekte	2	9	-	-	-	5	-	7,0	2,0	0,3
64	Umsetzbarkeit: Pilotprojekte	3	9	-	6	-	10	-	8,3	1,7	0,2

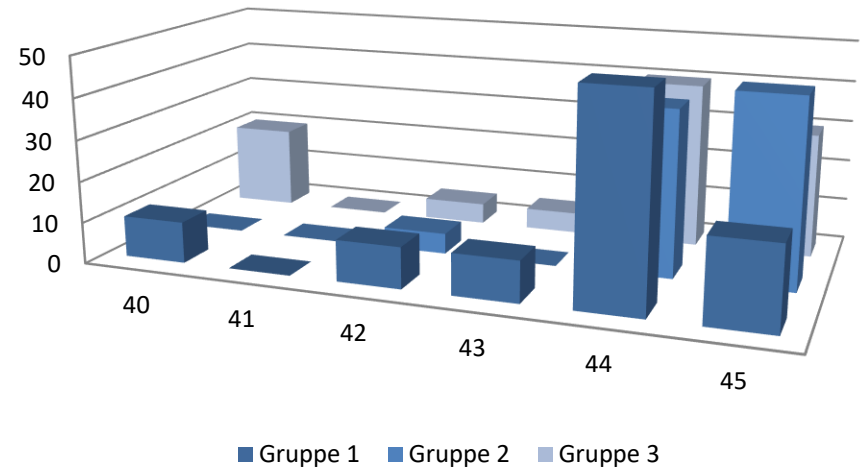
65	Wirksamkeit: Patentlaufzeit	3	5	-	1	-	7	-	4,3	2,5	0,6
66	Umsetzbarkeit: Patentlaufzeit	3	3	-	6	-	4	-	4,3	1,2	0,3
67	Wirksamkeit: Gesundheitsberatung	3	7	-	4	-	5	-	5,3	1,2	0,2
68	Umsetzbarkeit: Gesundheitsberatung	3	10	-	10	-	10	-	10,0	0,0	0,0
69	Wirksamkeit: Abwasserkanalnetz	3	5	-	10	-	10	-	8,3	2,4	0,3
70	Umsetzbarkeit: Abwasserkanalnetz	3	6	-	3	-	8	-	5,7	2,1	0,4



Forschungsförderung - Runde 1



Forschungsförderung - Runde 2



11.4 Ausgewertete Aussagen: Normative Dimension in SAUBER+

	Erste Plenarsitzung	Zweite Plenarsitzung	Ergebnisprotokoll
Handlungsbedarf bei verschiedenen Kontaminanten (Abschnitt I).	Man sollte sich auf drei wesentliche Bereiche beschränken und keine ‚Gießkannen- Taktik‘ verfolgen. ‚Nicht kleckern sondern klotzen‘		Während die drei Kleingruppen in der ersten Runde des Delphis unterschiedliche Schwerpunkte bei der Verteilung des Budgets setzten, einigte man sich für die zweite Runde bei der Verteilung des Budgets auf den Leitsatz: „Nicht kleckern sondern klotzen“
Bei der Zulassung von Humanarzneimitteln sollte das Umweltrisiko in der Nutzen-Risiko-Analyse des Arzneimittels berücksichtigt werden. (Item 12 (53/54))	Die eigentliche Fragestellung ist jedoch, dass sich etwas ändern soll. Zulassung auch unter Ausnahmen und/ oder Auflagen wird bereits berücksichtigt.	Es ist wirksam, jedoch nicht umsetzbar. Gründe für die Nicht-Umsetzbarkeit sind ethische Gründe. Der Patient- Nutzen wird immer der Umwelt vorgelagert sein. Vorgeschlagene Lösung: Vergleiche mit Arzneimitteln, die die gleiche Wirksamkeit haben, aber umweltverträglicher sind. Gegenargument: Ethische Sache: Beispielsweise ein Epileptiker, dessen Einstellung auf ein Medikament einen langen Prozess darstellt.	Als Begründung für die Nicht-Umsetzbarkeit wurde angeführt, dass bei der Arzneimittelzulassung Gesundheitsaspekte Priorität vor Umweltaspekten haben müssten. Der Patient/innen-Nutzen und die Patient/innen-Gesundheit müssten immer Vorrang vor dem Umweltschutz haben. Ein weiterer Vorschlag in diesem Zusammenhang war, neue umweltschädliche Arzneimittel zwar zunächst zuzulassen sie aber dann wieder vom Markt zu nehmen, wenn umweltverträglichere Alternativen zur Verfügung stünden. Dagegen stand das Argument, dass eine Umstellung auf ein alternatives Medikament nicht immer problemlos erfolgen könne. Als Beispiel wurden Epileptiker genannt, deren Einstellung auf ein Medikament einen langen Prozess darstelle. Es gab eine allgemeine Zustimmung zu dem Grundsatz: Der Versagungsgrund sollte nicht die Umwelt sein, wenn das Medikament sinnvoll für die Patient/innen ist.
Es sollte eine auf Umweltverträglichkeit bezogene Produktkennzeichnung von Medikamenten als Entwicklungsanreiz für eine ‚grüne‘		Das Finden neuer Wirkstoffen wäre gut, die Produktion sollte gefördert werden.	Für kleinere Unternehmen und Start-up-Unternehmen könnte eine solche Produktkennzeichnung aber sehr wohl ein gutes Marketing-Instrument sein und sei deshalb nicht generell abzulehnen.

<p>Produktpolitik' der Arzneimittelhersteller eingeführt werden. (Item 15)</p>			
<p>Ein bundesweit einheitlicher und verbindlicher Standard für die Entsorgung unverbrauchter Arzneimittel über die Apotheken, bei dem die nicht frei verkäuflichen Medikamente einem Pfandsystem unterliegen, sollte erprobt werden (der Arzneimittelnutzer erhält bei Rückgabe des angebrochenen Medikaments <u>oder der Verpackung</u> einen monetären Wert zurück, den er beim Kauf des Arzneimittels mit entrichtet hat). (Item 21)</p>	<p>Ablehnung: Ein Totschlagargument: ist das nicht eine Verführung für den Patienten, dass dieser die Medikamenten dann nicht zu sich nimmt.</p> <p>[Verständnisproblem: Nur die leere Verpackung soll zurückgegeben werden]</p>		<p>Die andere Gruppe begründete ihre Ablehnung mit der Befürchtung, dass dies für die Patient/innen einen Anreiz schaffen könnte, die Medikamenteneinnahme vorzeitig abzubrechen. Hier lag allerdings ein Missverständnis vor, denn die Maßnahme sieht vor, dass auch die leeren Verpackungen zur Einlösung des Pfands zurückgegeben werden können.</p> <p>In der Diskussion, die sich an diese Stellungnahme angeschlossen, wurde von anderen Teilnehmern die Ablehnung eines pfandgestützten Systems erneut betont. Ein solches System sei nicht verbraucherfreundlich, und die Verbraucher würden sich nach einer Kosten/Aufwand-Nutzen-Kalkulation dagegen entscheiden.</p> <p>Andere Stakeholder betonten, dass dem Patienten bewusst gemacht werden müsse, dass Medikamente ein besonderes Produkt seien. Dafür wäre die Rückgabe über die Apotheken ein geeignetes didaktisches Mittel: „Kommt aus der Apotheke raus, muss in die Apotheke rein“.</p>
<p>Die novellierte deutsche Klärschlammverordnung sollte ein generelles Verbot des Aufbringens von Klärschlamm auf Lebensmittelböden (Anbauflächen für Nahrungs- und Futtermittel) umfassen. Für Energieböden (nachwachsende Rohstoffe) sollten andere Bestimmungen gelten. (Item 39)</p>		<p>Nutzung des Potentials zur Düngung</p>	<p>So war ein Standpunkt, dass eine Unterscheidung in diesem Kontext deshalb keinen Sinn mache, da die Aufbringung von Klärschlamm immer eine Grundwassergefährdung beinhalten würde. In der Diskussion kam die Frage auf, inwiefern solche Alternativen bestünden bzw. entwickelt werden sollten, bei denen die im Klärschlamm enthaltenen Nährstoffe gewonnen und als Dünger genutzt werden könnten. Die Gruppe einigte sich auf folgende Umformulierung der Maßnahme: Die novellierte deutsche Klärschlammverordnung sollte ein generelles Verbot des Aufbringens von Klärschlamm auf landwirtschaftlich genutzten Böden umfassen.</p>

<p>Es sollte eine bundesweite Aufklärungskampagne zur sachgerechten Entsorgung unverbrauchter Arzneimittel durchgeführt werden. (Item 59)</p>		<p>Wenn jedoch kein Standard realisiert werden konnte, dann muss klar kommuniziert werden, dass es sich um Umweltschutzmaßnahmen und nicht um Maßnahmen zum Schutz des Menschen geht.</p>	<p>Generell müsse klar kommuniziert werden, dass es bei der Frage der Entsorgung vorrangig um Umweltschutz und nicht um Gesundheitsschutz gehe.</p>
<p>Forschungsförderung in Bezug auf Arzneimittelrückstände im Gewässerkreislauf (Abschnitt V)</p>			<p>Einzelne Beiträge der Stakeholder mündeten vielmehr rasch in dem Konsens, dass auch hier der Leitsatz „Nicht kleckern sondern klotzen“ gelten sollte, in dem Sinne, dass das Geld vor allem dort hinfließen sollte, wo es notwendig ist, Forschung aufzubauen.</p> <p>Die Bedeutung des Feldes wurde dabei unterschiedlich eingeschätzt. Einige Stakeholder befanden die Forschung zu „Guter Compliance“ wichtig, erachteten aber den finanziellen Aufwand als eher gering. Andere Stakeholder bezeichneten diese Forschung als „nice to have“, schätzten also die Bedeutung niedriger ein.</p>

Legende: (1) Werturteile; (2) Präferenzen; (3) (individuelle) Bewertungen von Optionen ; (4) moralische Urteile; Aussagekontext

11.5 Fragebogen der Vorbefragung des Gruppendelphis zu SEAWIND



The
SEAWIND Project



Experten-Delphi

Sound Exposure & Risk Assessment of Wireless Network Devices

Sehr geehrte Dame, sehr geehrter Herr,

die gemeinnützige DIALOGIK GmbH bearbeitet gegenwärtig unter Leitung von Prof. Dr. Dr. h.c. Renn ein von der Europäischen Kommission gefördertes Projekt namens „*Sound exposure and risk assessment of wireless network devices*“ (SEAWIND; <http://seawind-fp7.eu/>). In SEAWIND haben wir es uns zum Ziel gesetzt, den Einfluss kabelloser Netzwerke wie z.B. WLAN auf mögliche Auswirkungen in humanen Zellen interdisziplinär zu erforschen, wobei die Ergebnisse dieser Forschungen von Studien zur Risikowahrnehmung flankiert werden. Aufgabe von DIALOGIK¹⁶¹ ist es hierbei, einen Maßnahmenkatalog für politische Entscheidungsträger mit Empfehlungen zur Risk-Governance im Umgang mit den Einflüssen kabelloser Netzwerke auf Zellstrukturen zu entwickeln.

Auch wenn Sie leider am Gruppen-Delphi leider nicht persönlich teilnehmen können, möchten wir Sie bitten, den beigefügten Fragebogen auszufüllen und uns bis zum 31.08.2012 zukommen zu lassen. Ziel ist eine umfassende Integration der Kompetenzen aller von uns angefragten Experten. Diesen Fragebogen erhalten Teilnehmer und Nichtteilnehmer im Vorfeld des Gruppen-Delphis. Die Beantwortung der Fragen gibt uns wichtige Hinweise auf die Identifikation strittiger Fragen, die dann vertiefend von den Teilnehmern des Gruppen-Delphis diskutiert werden.

Wir würden uns sehr freuen, wenn Sie mit Ihrem Expertenwissen zum Gelingen des Delphis schon im Vorfeld beitragen!

Für Rückfragen stehen wir gerne zur Verfügung!

Kontakt:

Jörg Hilpert

E-Mail: Hilpert@dialogik-expert.de

Tel.: 0711 685 84539

¹⁶¹In Kooperation mit: „Foundation for Research on Information Technologies in Society“; Interdisciplinary Institute for BroadBand Technology, Ghent-Ledeberg; University of Aalborg, Department of Electronic Systems; Aristotle University of Thessaloniki, Department of Physics; Fraunhofer-Institute of Toxicology and Experimental Medicine, Division Toxicology and Environmental Hygiene, Hannover; University of Basel, Department of Biomedicine; Schmid & Partner Engineering AG, Zürich

Abschnitt I: Kommunikationsakteure

Frage 1.1: Information der Öffentlichkeit

Im Jahr 2011 hat die International Agency for Research on Cancer (IARC) ihre Einschätzung des Krebsrisikos durch elektromagnetische Felder (EMF) bekanntgegeben. Die Einschätzung lautet: „möglicherweise krebserregend“ (Risikogruppe 2b)¹⁶². Welche der nachfolgend aufgelisteten Akteursgruppen sollten Ihrer Meinung nach die Öffentlichkeit über diese Einstufung der Risiken unter die Kategorie 2b informieren? Bitte bewerten Sie diesbezüglich nachstehend die Akteursgruppen. Die Ziffer 1 (-) bedeutet hierbei, dass diese Akteursgruppe die Öffentlichkeit auf keinen Fall über diese Einstufung und ihre Implikationen informieren sollte und die Zahl 10 (+) bedeutet, dass diese Gruppe auf jeden Fall die Öffentlichkeit über diese Einstufung und ihre Implikationen informieren sollte.

		Akteursgruppe	1	2	3	4	5	6	7	8	9	10
01	A	Internationale Organisationen, wie z.B. die Europäische Kommission										
02	B	Nationale Organisationen, wie z.B. die Bundeszentrale für gesundheitliche Aufklärung										
03	C	Nationale Behörden, wie z.B. das Bundesministerium für Gesundheit, das Bundesamt für Strahlenschutz oder das Bundesamt für Risikobewertung										
04	D	Journalisten										
05	E	Wissenschaftler										
06	F	Organisierte Öffentlichkeit, wie z.B. Umweltschutzgruppen										
07	G	Hardwarehersteller, z.B. von Routern oder Multi-Media										
08	H	Content Provider										
09	I	Politiker auf nationaler Ebene										
10	J	Stiftungen										
11	K	Sonstige Akteursgruppe (bitte angeben):	<hr/>									

¹⁶² IARC Risikogruppen: (1) krebserregend, (2a) wahrscheinlich krebserregend, (2b) möglicherweise krebserregend, (3) nicht klassifizierbar für Menschen, (4) wahrscheinlich nicht krebserregend.

Frage 1.2: Zusammenarbeit von Akteuren

Bei der Kommunikation von Risiken mit hohen Unsicherheiten kann eine Zusammenarbeit verschiedener Akteure sinnvoll sein.


Bitte kreuzen Sie nachstehend die drei wichtigsten Akteure an, die Ihrer Meinung nach zusammenarbeiten sollten, um die Risiken elektromagnetischer Felder an die Öffentlichkeit zu kommunizieren.

Akteursgruppe	Auswahl
Internationale Organisationen wie z.B. die Europäische Kommission	
Nationale Organisationen, wie z.B. die Bundeszentrale für gesundheitliche Aufklärung	
Nationale Behörden, wie z.B. das Bundesministerium für Gesundheit, das Bundesamt für Strahlenschutz oder das Bundesamt für Risikobewertung	
Journalisten	
Wissenschaftler	
Organisierte Öffentlichkeit, wie z.B. Umweltschutzgruppen	
Hardwarehersteller, z.B. von Routern oder Multi-Media	
Provider	
Sonstige Akteursgruppe: _____	

Abschnitt II: Kommunikation

Frage 2: Geeignete Kommunikationsmedien/ -formen

Bei der Kommunikation von Risiken mit hohen Unsicherheiten durch elektromagnetische Felder können unterschiedliche Kommunikationsmedien/ -formen, wie z.B. Tageszeitung oder das Internet eingesetzt werden. Bitte bewerten Sie die nachstehend aufgeführten Kommunikationsmedien/ -formen im Hinblick auf ihre Eignung zur Kommunikation von Risiken mit hoher Unsicherheit an die allgemeine Öffentlichkeit (keine Spezialisten in dem Fach).
Bitte tragen Sie hierfür einen der nachfolgenden Werte ein: Geben Sie den Wert 1 (-) an, wenn Sie der Ansicht sind, dass sich ein Medium / eine Form ganz und gar nicht für die Kommunikation von Risiken mit hoher Unsicherheit eignet und geben Sie die Ziffer 10 (+) an, wenn Sie ein Medium / eine Form als ganz und gar geeignet beurteilen.

		Kommunikationsmedium/ -form	1	2	3	4	5	6	7	8	9	10
12	A	Strahlenrisiko-Label (Ampel-System) auf emitierende Geräte wie Toaster oder Handies										
13	B	Informationsbroschüre(n)										
14	C	Artikel in Fachzeitschrift(en)										
15	D	Artikel in Tageszeitung(en)										
16	E	Fernsehreportage(n)										
17	F	Radiosendung(en)										
18	G	Informationsveranstaltung(en)										
19	H	Bürgerdiskussionsrunde(n)										
20	I	Internet: Web-Präsenzen										
21	J	Internet: Social Communities (Twitter, Facebook, etc.), Blogs										
22	K	Computerspiele/-tools										
23	L	Sonstiges Kommunikationsmedium sollte genutzt werden (bitte angeben)	 _____									


Abschnitt III: Kommunikationsinhalte

Frage 3.1: Inhalte der Kommunikation

Die Einschätzung der Agency for Research on Cancer (IARC), dass elektromagnetische Felder „möglicherweise krebserregend“ (Risikogruppe 2b) sind, kann mittels unterschiedlicher Inhalte kommuniziert werden. Was denken Sie: Welche Inhalte sollten bei solch einem Risiko mit hoher Unsicherheit an die allgemeine Bevölkerung kommuniziert werden?

Bitte geben Sie für die nachstehenden Aussagen jeweils einen Wert von 1 (-) bis 10 (+) an. Die Ziffer 1 (-) bedeutet hierbei, dass Sie der Ansicht sind, dass dieser Kommunikationsinhalt ganz und gar nicht kommuniziert werden sollte und die Zahl 10 (+) dass dieser Inhalt unbedingt kommuniziert werden sollte.

		Kommunikationsinhalte	1	2	3	4	5	6	7	8	9	10
24	A	Die <u>Datengrundlage</u> zur Festlegung der Risikogruppe.										
25	B	Die Methode zur <u>Datenbewertung</u> .										
26	C	<u>Handlungsempfehlungen</u> für die Vermeidung von Risiken.										
27	D	<u>Unsicherheiten</u> in der Risikobewertung.										
28	E	<u>Nutzungssituationen, die mit Risiken verbunden sind.</u>										
29	F	<u>Fachliche Expertise / Ausrichtung</u> , die der Risikobewertung zugrunde liegt.										
30	G	Die <u>Unabhängigkeit</u> der fachlich bewertenden Institution.										
31	H	Das angestrebte <u>Risikomanagement</u> .										
32	I	Zur besseren Vergleichbarkeit sollten <u>Forschungsergebnisse mit vergleichbaren Risiken</u> dargestellt werden (etwa andere Risiken in der Kategorie 2b).										
33	J	Das <u>Schema der Risikoklassifikation</u> (IARC Risikogruppen: (1) krebserregend, (2a) wahrscheinlich krebserregend, (2b) möglicherweise krebserregend, (3) nicht klassifizierbar für Menschen, (4) wahrscheinlich nicht krebserregend) sowie die Kriterien für die Zuordnung.										

34	K	Es sollte Folgendes kommuniziert werden (bitte angeben):	 _____
----	---	--	---

Frage 3.2: Kommunikation von Untersuchungsergebnissen

Wie sollen Ergebnisse von Strahlungsuntersuchungen kommuniziert werden, die mit hoher Unsicherheit behaftet sind und bei welchen bisher keine verlässlichen Aussagen über gesundheitliche Konsequenzen möglich sind?

Frage 3.3: Ziele der Kommunikation

Kommunikation zu Risiken mit hoher Unsicherheit kann zum Erreichen von unterschiedlichen Zielen eingesetzt werden. Welche Ziele sollten Ihrer Meinung nach durch eine Kommunikation zu Risiken mit hoher Unsicherheit erreicht werden?

Nachstehend finden Sie einige Aussagen dazu, was Kommunikation leisten kann. Bitte kreuzen Sie nachstehend die drei Ziele an, die Ihrer Meinung nach durch eine entsprechende Risikokommunikation auf jeden Fall erreicht werden sollten. Bei den übrigen Zielen geben Sie bitte an, ob sie angestrebt werden sollen (+), ob Sie diese als indifferent ansehen (0) oder ob Sie eher der Meinung sind, dass dieses Ziel nicht angestrebt werden soll (-).

	Kommunikationsziele	Auswahl	Rating
A	Das <u>Vertrauen in Risikobewertungsverfahren</u> steigern.		
B	Das <u>Vertrauen in die kommunizierenden Institutionen</u> steigern.		
C	<u>Einseitige Risiko-Interpretationen vermeiden</u> .		
D	<u>Unsicherheitsquellen in der Risikobewertung</u> identifizieren.		
E	Den Grad der <u>Evidenz der Risikoanalysen</u> charakterisieren und bewerten.		
F	Das <u>Management durch zuständige Behörden</u> im Umgang mit dem Risiko darlegen.		
G	Schaffung von <u>Transparenz</u> in der Risikobewertung und im Risikomanagement.		
H	Hervorhebung der <u>Kompetenzen</u> der risikobewertenden <u>Institutionen</u> .		
I	Nachvollziehbarkeit der <u>Datengrundlage</u> zur Risikobewertung für jedermann.		
J	Durch eine gezielte Kommunikation sollte: _____		

Frage 3.4: Vermittlung von Schutzkonzepten

Auf welchen Wegen können Ihrer Meinung nach Schutzkonzepte (wie z.B. Vorsorge) vermittelt werden, ohne die Risikowahrnehmung durch die Öffentlichkeit negativ zu verstärken?

Frage 3.5: Informationsbedürfnis der Bevölkerung

Welche Bedürfnisse der Bevölkerung sollten Ihrer Ansicht nach bei der Kommunikation von Risiken bzgl. elektromagnetischer Strahlung vorrangig berücksichtigt werden?

Abschnitt IV: Kommunikation von Messdaten

Frage 4.1: Wege der Übermittlung

Grundlage für die Beurteilung eines Risikos sind Messdaten. Die Kommunikation von Messdaten und deren Implikationen kann sich auf verschiedene Art und Weise an unterschiedliche Akteursgruppen richten. Damit eine geeignete Kommunikationsweise ermittelt werden kann, ist es wichtig vorab zu definieren, an welche Zielgruppe sich die Kommunikation richten soll. Nachfolgend finden Sie verschiedene Zielgruppen, für die Messdaten relevant sein könnten:

(1) Wirtschaftliche Stakeholder (Provider, Hersteller von Hardware, etc.); **(2)** Politische Stakeholder (NGOs, Strahlensensible, etc.); **(3)** Wissenschaftliche Akteure; **(4)** Politik (national/international); **(5)** Medien; **(6)** Öffentlichkeit junge Leute, d.h. Personen bis ca. 40 Jahre; **(7)** Öffentlichkeit ältere Leute, d.h. Personen ab ca. 40 Jahren.

Bitte geben Sie per Nennung der oben aufgeführten Zahl an, welche Zielgruppen Sie für die jeweils dargestellte Kommunikationsweise passend finden (Mehrfachnennungen sind möglich). Bitte bewerten Sie auch die generelle Effektivität der dargestellten Kommunikationsweise, wobei Effektivität hier ein Beurteilungskriterium darstellen soll, mit dem sich beschreiben lässt, in welchem Ausmaß eine Kommunikationsweise dafür geeignet ist, die entsprechende Akteursgruppen anzusprechen. 1 (-) bedeutet, dass Sie diesen Weg als ineffektiv bewerten, 10 (+) bedeutet, dass Sie diesen Weg als sehr effektiv bewerten.

		Kommunikationsweise	Relevante Zielgruppe(n) (Nennung per Zahl)	Generelle Effektivität											
				1	2	3	4	5	6	7	8	9	10		
35	A	Kommunikation der Messdaten auf Pressekonferenzen													
36	B	Kommunikation der Messdaten in informellen Gesprächsgruppen													
37	C	Virtuelle Treffen oder Online-Konferenzen im Internet													
38	D	Fachliche Beiträge auf Konferenzen und Workshops													
39	E	Diskussionen in Blogs und Foren													
40	F	Videos auf Youtube und ähnlichen Portalen													
41	G	Berichte in TV und Radio													
42	H	Folgendes Kommunikationsmedium halte ich für geeignet (bitte angeben):	<input type="text"/>												

Frage 4.2: Glaubwürdigkeit der Übermittlungswege

Neben der Effektivität können sich die verschiedenen Wege der Übermittlung von Messdaten auch in ihrer Glaubwürdigkeit unterscheiden. Bitte geben Sie für die nachstehenden Aussagen jeweils einen Wert von 1 bis 10 an. 1 (-) bedeutet, dass Sie diesen Kommunikationsweg als überhaupt nicht glaubwürdig bewerten, 10 (+) bedeutet, dass Sie diesen Kommunikationsweg als sehr glaubwürdig bewerten.

		Wege der Übermittlung	Glaubwürdigkeit											
			1	2	3	4	5	6	7	8	9	10		
43	A	Kommunikation der Messdaten auf Pressekonferenzen												
44	B	Kommunikation der Messdaten in Bürger-Gesprächsgruppen												
45	C	Virtuelle Treffen oder Online-Konferenzen im Internet												
46	D	Fachliche Beiträge auf Konferenzen und Workshops												
47	E	Diskussionen in Blogs und Foren												
48	F	Videos auf Youtube und ähnlichen Portalen												
49	G	Berichte in TV und Radio												
50	H	Folgendes Kommunikationsmedium halte ich für glaubwürdig (bitte angeben):												

Abschnitt V: Konkrete Maßnahmen

Frage 5: Maßnahmen zur Steigerung der Aufmerksamkeit

Im Folgenden finden Sie eine Reihe von Maßnahmen, die auf das Problem der potentiellen Strahlung von elektromagnetischen Feldern aufmerksam machen sollen.

Bitte tragen Sie hierfür einen Wert ein. Geben Sie den Wert 1 (-) an, wenn Sie der Ansicht sind, dass sich diese Maßnahme ganz und gar nicht eignet und geben Sie die Ziffer 10 (+) an, wenn sich die Maßnahme voll und ganz eignet.

		Maßnahmen	1	2	3	4	5	6	7	8	9	10
5 1	A	Apps, die die individuelle Strahlenbelastung eines Individuums messen bzw. simulieren										
5 2	B	Labels auf allen Geräten, die eine Bandbreite von möglichen Risiken aufzeigen										
5 3	C	Mobile Dienstleister, die Messungen vor Ort zur Überprüfung der Strahlenintensität durchführen										
5 4	D	Sonstiges und zwar: _____										

Abschnitt VI: Weitere Aspekte

Frage 6.1: Widersprüchliche Ergebnisse

Wie sollte Ihrer Meinung nach die Politik mit widersprüchlichen wissenschaftlichen Ergebnissen umgehen? (Strahlung gefährlich vs. nicht gefährlich)

Frage 6.2: Faktoren der Wahrnehmung von Risiko

Welche Faktoren beeinflussen Ihrer Meinung nach die Risikowahrnehmung in der Öffentlichkeit?

Bitte geben Sie für die nachfolgenden Faktoren jeweils einen Wert von 1 (-) bis 10 (+) an. Die Ziffer 1 (-) bedeutet hierbei, dass Sie der Ansicht sind, dass dieser Faktor ganz und gar nicht relevant ist und die Zahl 10 (+), dass dieser Faktor sehr relevant ist.

		Faktoren der Risikowahrnehmung	1	2	3	4	5	6	7	8	9	10
55	A	Alter										
56	B	Geschlecht										
57	C	Bildung										
58	D	Vertrauen in die Kontrollinstanzen										
59	E	Grundeinstellung zu Technik und Wissenschaft										
60	F	Ausstattung mit Technik im eigenen Haushalt										
61	G	Sonstiges und zwar: _____										

Frage 6.3: Vermittlung widersprüchlicher Ergebnisse

Stellen Sie sich vor, Sie sind in einer Podiumsdiskussion und ein Experte betont die Sicherheit von IT-Geräten, die elektromagnetische Strahlung aussenden, während der andere Experte die möglichen Risiken und vor allem die Unsicherheiten betont. Dann steht ein Publikumsgast auf und fragt Sie: „Was stimmt denn nun: Sind die Geräte, die ich benutze, gefährlich oder nicht? Darauf will ich eine klare Antwort“. Was würde Sie ihm antworten?

Vielen Dank für die Beantwortung des Fragebogens!

11.6 Fragebogen des Gruppendelphis zu SEAWIND



Experten-Delphi

Sound Exposure & Risk Assessment of Wireless Network Devices

Erste Runde



gemeinnützige
Gesellschaft für
Kommunikations- und
Kooperationsforschung
mbH

Abschnitt I: Kommunikationsakteure

Frage 1.1: Information der Öffentlichkeit

Im Jahr 2011 hat die International Agency for Research on Cancer (IARC) ihre Einschätzung des Krebsrisikos durch elektromagnetische Felder (EMF) bekanntgegeben. Die Einschätzung lautet: „möglicherweise krebserregend“ (Risikogruppe 2b)¹⁶³. Welche der nachfolgend aufgelisteten Akteursgruppen sollten Ihrer Meinung nach die Öffentlichkeit über diese Einstufung der Risiken unter die Kategorie 2b informieren? Bitte bewerten Sie diesbezüglich nachstehend die Akteursgruppen. Die Ziffer 1 (-) bedeutet hierbei, dass diese Akteursgruppe die Öffentlichkeit auf keinen Fall über diese Einstufung und ihre Implikationen informieren sollte und die Zahl 10 (+) bedeutet, dass diese Gruppe auf jeden Fall die Öffentlichkeit über diese Einstufung und ihre Implikationen informieren sollte.

		Akteursgruppe	1	2	3	4	5	6	7	8	9	10
01	A	Internationale Organisationen, wie z.B. die Europäische Kommission										
02	B	Nationale Organisationen, wie z.B. die Bundeszentrale für gesundheitliche Aufklärung										
03	C	Nationale Behörden, wie z.B. das Bundesministerium für Gesundheit, das Bundesamt für Strahlenschutz oder das Bundesamt für Risikobewertung										
04	D	Journalisten										
05	E	Wissenschaftler										
06	F	Organisierte Öffentlichkeit, wie z.B. Umweltschutzgruppen										
07	G	Hardwarehersteller, z.B. von Routern oder Multi-Media										
08	H	Content Provider										
09	I	Politiker auf nationaler Ebene										
10	J	Stiftungen										
11	K	Verbraucherschutzorganisationen										
12	L	Sonstige Akteursgruppe (bitte angeben):	<input type="text"/>									

¹⁶³ IARC Risikogruppen: (1) krebserregend, (2a) wahrscheinlich krebserregend, (2b) möglicherweise krebserregend, (3) nicht klassifizierbar für Menschen, (4) wahrscheinlich nicht krebserregend.

Abschnitt II: Kommunikationsinhalte

Frage 2.1: Inhalte der Kommunikation

Die Einschätzung der Agency for Research on Cancer (IARC), dass elektromagnetische Felder „möglicherweise krebserregend“ (Risikogruppe 2b) sind, kann mittels unterschiedlicher Inhalte kommuniziert werden. Was denken Sie: Welche Inhalte sollten bei solch einem Risiko mit hoher Unsicherheit an die allgemeine Bevölkerung kommuniziert werden?

Bitte geben Sie für die nachstehenden Aussagen jeweils einen Wert von 1 (-) bis 10 (+) an. Die Ziffer 1 (-) bedeutet hierbei, dass Sie der Ansicht sind, dass dieser Kommunikationsinhalt ganz und gar nicht kommuniziert werden sollte und die Zahl 10 (+) bedeutet, dass dieser Inhalt unbedingt kommuniziert werden sollte.

		Kommunikationsinhalte	1	2	3	4	5	6	7	8	9	10
13	A	Die <u>Datengrundlage</u> zur Festlegung der Risikogruppe.										
14	B	Die Methode zur <u>Datenbewertung</u> .										
15	C	<u>Handlungsempfehlungen</u> für die Vermeidung von Risiken.										
16	D	<u>Unsicherheiten</u> in der Risikobewertung.										
17	E	<u>Nutzungssituationen</u> , die mit Risiken verbunden sind.										
18	F	<u>Fachliche Expertise / Ausrichtung</u> , die der Risikobewertung zugrunde liegt.										
19	G	Die <u>Unabhängigkeit</u> der fachlich bewertenden Institution.										
20	H	Das angestrebte <u>Risikomanagement</u> .										
21	I	Zur besseren Vergleichbarkeit sollten <u>Forschungsergebnisse mit vergleichbaren Risiken</u> dargestellt werden (etwa andere Risiken in der Kategorie 2b).										
22	J	Das <u>Schema der Risikoklassifikation</u> (IARC Risikogruppen: (1) krebserregend, (2a) wahrscheinlich krebserregend, (2b) möglicherweise krebserregend, (3) nicht klassifizierbar für Menschen, (4) wahrscheinlich nicht krebserregend) sowie die Kriterien für die Zuordnung.										

Frage 2.2: Ziele der Kommunikation

Kommunikation zu Risiken mit hoher Unsicherheit kann zum Erreichen von unterschiedlichen Zielen eingesetzt werden. Welche Ziele sollten Ihrer Meinung nach durch eine Kommunikation zu Risiken mit hoher Unsicherheit erreicht werden?


Nachstehend finden Sie einige Aussagen dazu, was Kommunikation leisten kann. Bitte kreuzen Sie nachstehend die drei Ziele an, die Ihrer Meinung nach durch eine entsprechende Risikokommunikation auf jeden Fall erreicht werden sollten.

	Kommunikationsziele	Auswahl
A	Das <u>Vertrauen in Risikobewertungsverfahren</u> steigern.	
B	Das <u>Vertrauen in die kommunizierenden Institutionen</u> steigern.	
C	<u>Einseitige Risiko-Interpretationen vermeiden</u> .	
D	<u>Unsicherheitsquellen in der Risikobewertung identifizieren</u> .	
E	Den Grad der <u>Evidenz der Risikoanalysen</u> charakterisieren und bewerten.	
F	Das <u>Management durch zuständige Behörden</u> im Umgang mit dem Risiko darlegen.	
G	Schaffung von <u>Transparenz</u> in der Risikobewertung und im Risikomanagement.	
H	Hervorhebung der <u>Kompetenzen</u> der risikobewertenden <u>Institutionen</u> .	
I	Nachvollziehbarkeit der <u>Datengrundlage</u> zur Risikobewertung für jedermann.	
J	Den <u>Umgang mit (unvermeidbaren) Risiken</u> verdeutlichen.	
K	Durch eine gezielte Kommunikation sollte: z	

Abschnitt III: Kommunikation von Messdaten

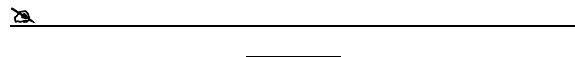
Frage 3.1: Wege der Übermittlung

Grundlage für die Beurteilung eines Risikos sind Messdaten. Die Kommunikation von Messdaten und deren Implikationen kann auf verschiedene Art und Weise erfolgen. Bitte bewerten Sie die generelle Effektivität der dargestellten Kommunikationsweise, wobei Effektivität hier ein Beurteilungskriterium darstellen soll, mit dem sich beschreiben lässt, in welchem Ausmaß eine Kommunikationsweise dafür geeignet ist, bestimmte Messdaten zu kommunizieren. 1 (-) bedeutet, dass Sie diesen Weg als ineffektiv bewerten, 10 (+) bedeutet, dass Sie diesen Weg als sehr effektiv bewerten.

		Kommunikationsinhalte	Effektivität										
			1	2	3	4	5	6	7	8	9	10	
23	A	Kommunikation der Messdaten auf Pressekonferenzen											
24	B	Kommunikation der Messdaten in Bürger-Gesprächsgruppen											
25	C	Virtuelle Treffen oder Online-Konferenzen im Internet											
26	D	Fachliche Beiträge auf Konferenzen und Workshops											
27	E	Diskussionen in Blogs und Foren											
28	F	Videos und Podcasts auf Youtube und ähnlichen Portalen											
29	G	Berichte in TV und Radio											
30	H	Wissenschaftliche Publikationen											
31	I	Bürgerforen und Bürgerparlamente											
32	J	Informationsbroschüren											
33	K	Anerkannte Zeitungen und Zeitschriften											
34	L	Folgendes Kommunikationsmedium halte ich für geeignet (bitte angeben):	 _____										

Frage 3.2: Glaubwürdigkeit der Übermittlungswege

Neben der Effektivität können sich die verschiedenen Wege der Übermittlung von Messdaten auch in ihrer Glaubwürdigkeit unterscheiden. Bitte geben Sie für die nachstehenden Aussagen jeweils einen Wert von 1 bis 10 an. 1 (-) bedeutet, dass Sie diesen Kommunikationsweg als überhaupt nicht glaubwürdig bewerten, 10 (+) bedeutet, dass Sie diesen Kommunikationsweg als sehr glaubwürdig bewerten.

		Wege der Übermittlung	Glaubwürdigkeit											
			1	2	3	4	5	6	7	8	9	10		
35	A	Kommunikation der Messdaten auf Pressekonferenzen												
36	B	Kommunikation der Messdaten in Bürger-Gesprächsgruppen												
37	C	Virtuelle Treffen oder Online-Konferenzen im Internet												
38	D	Fachliche Beiträge auf Konferenzen und Workshops												
39	E	Diskussionen in Blogs und Foren												
40	F	Videos auf Youtube und ähnlichen Portalen												
41	G	Berichte in TV und Radio												
42	H	Wissenschaftliche Publikationen												
43	I	Bürgerforen und Bürgerparlamente												
44	J	Informationsbroschüren												
45	K	Anerkannte Zeitungen und Zeitschriften												
46	L	Folgendes Kommunikationsmedium halte ich für glaubwürdig (bitte angeben):												

Abschnitt V: Kommunikation von Untersuchungsergebnissen

Frage 5: Kommunikation bei hoher Unsicherheit

Wie sollen Ergebnisse von Strahlungsuntersuchungen kommuniziert werden, die mit hoher Unsicherheit behaftet sind und bei welchen bisher keine verlässlichen Aussagen über gesundheitliche Konsequenzen möglich sind?

Vielen Dank für die Beantwortung des Fragebogens!

Expertendelphi

Sound Exposure & Risk Assessment of Wireless Network Devices

Statistische Ergebnisse

Runde 1: 10er Skalen

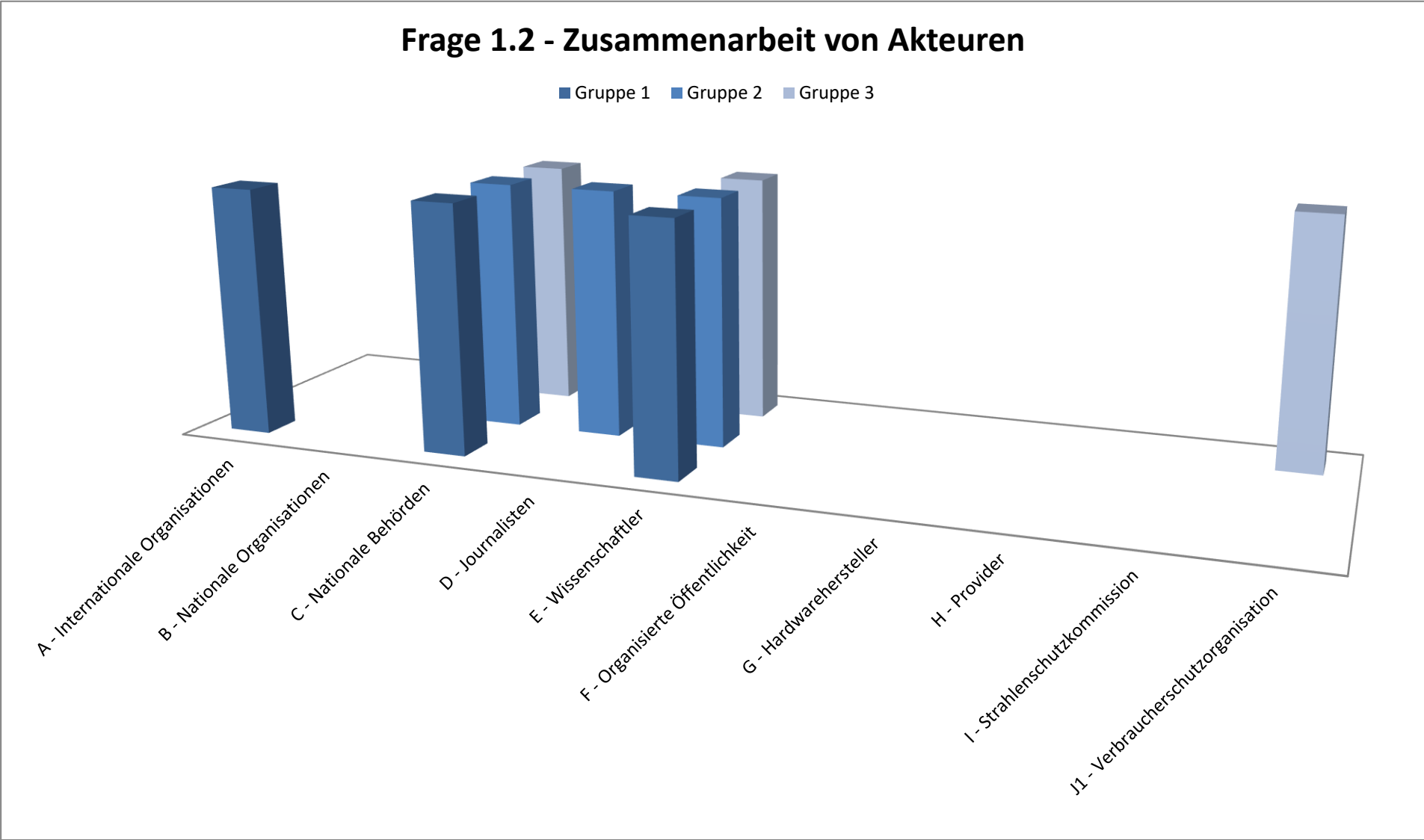
Frage-Nr.	Thematischer Bereich der Frage	Frageinhalt	Skala	Antwort Gruppe 1	Antwort Gruppe 2	Antwort Gruppe 3	Anzahl Antworten	Arithmetisches Mittel	Standardabweichung	Varianz	Variationskoeffizient
1	I: Kommunikationsakteure	Internationale Organisationen	10	10	-	5	2	7,5	2,5	6,3	0,3
2	I: Kommunikationsakteure	Nationale Organisationen	10	10	-	3	2	6,5	3,5	12,3	<u>0,5</u>
3	I: Kommunikationsakteure	Nationale Behörden	10	10	-	10	2	10,0	0,0	0,0	0,0
4	I: Kommunikationsakteure	Journalisten	10	10	-	5	2	7,5	2,5	6,3	0,3
5	I: Kommunikationsakteure	Wissenschaftler	10	10	-	8	2	9,0	1,0	1,0	0,1
6	I: Kommunikationsakteure	Organisierte Öffentlichkeit	10	10	-	5	2	7,5	2,5	6,3	0,3
7	I: Kommunikationsakteure	Hardwarehersteller	10	8	2	1	3	3,7	3,1	9,6	<u>0,8</u>
8	I: Kommunikationsakteure	Content Provider	10	3	-	1	2	2,0	1,0	1,0	<u>0,5</u>
9	I: Kommunikationsakteure	Politiker auf nationaler Ebene	10	10	1	1	3	4,0	4,2	18,0	<u>1,1</u>
10	I: Kommunikationsakteure	Stiftungen	10	10	-	3	2	6,5	3,5	12,3	<u>0,5</u>
11	I: Kommunikationsakteure	Verbraucherschutzorganisationen	10	10	-	8	2	9,0	1,0	1,0	0,1
12	I: Kommunikationsakteure	Sonstige Akteursgruppen	Netzbetreiber, fachbezogene Vereine (Organisationen), berufsständische Organisationen im Gesundheitsbereich, Fachgesellschaften								

13	II: Kommunikationsinhalte	Datengrundlage	10	10	1	10	3	7,0	4,2	18,0	<u>0,6</u>
14	II: Kommunikationsinhalte	Datenbewertung	10	10	1	10	3	7,0	4,2	18,0	<u>0,6</u>
15	II: Kommunikationsinhalte	Handlungsempfehlungen	10	10	10	10	3	10,0	0,0	0,0	0,0
16	II: Kommunikationsinhalte	Unsicherheiten	10	10	3	10	3	7,7	3,3	10,9	0,4
17	II: Kommunikationsinhalte	Nutzungssituation	10	10	10	10	3	10,0	0,0	0,0	0,0
18	II: Kommunikationsinhalte	Fachlicher Expertise / Ausrichtung	10	7	5	10	3	7,3	2,1	4,2	0,3
19	II: Kommunikationsinhalte	Unabhängigkeit	10	10	10	10	3	10,0	0,0	0,0	0,0
20	II: Kommunikationsinhalte	Risikomanagement	10	10	7	10	3	9,0	1,4	2,0	0,2
21	II: Kommunikationsinhalte	Forschungsergebnisse mit vgl. Risiken	10	8	5	10	3	7,7	2,1	4,2	0,3
22	II: Kommunikationsinhalte	Schema der Risikoklassifikation	10	10	3	10	3	7,7	3,3	10,9	0,4
23	III: Kommunikation von Messdaten	Pressekonferenzen (Effektivität)	10	5	-	5	2	5,0	0,0	0,0	0,0
24	III: Kommunikation von Messdaten	Bürger-Gesprächsgruppen (Effektivität)	10	5	2	1	3	2,7	1,7	2,9	<u>0,6</u>
25	III: Kommunikation von Messdaten	Virtuelle Treffen und Online Konferenzen (Effektivität)	10	4	-	5	2	4,5	0,5	0,3	0,1
26	III: Kommunikation von Messdaten	Konferenzen und Workshops (Effektivität)	10	10	-	8	2	9,0	1,0	1,0	0,1
27	III: Kommunikation von Messdaten	Blogs und Foren (Effektivität)	10	4	-	1	2	2,5	1,5	2,3	<u>0,6</u>
28	III: Kommunikation von Messdaten	Youtube und andere Portale (Effektivität)	10	7	-	1	2	4,0	3,0	9,0	<u>0,8</u>
29	III: Kommunikation von Messdaten	TV und Radio (Effektivität)	10	9	-	1	2	5,0	4,0	16,0	<u>0,8</u>
30	III: Kommunikation von Messdaten	Wissenschaftl. Publikationen (Effektivität)	10	10	10	10	3	10,0	0,0	0,0	0,0

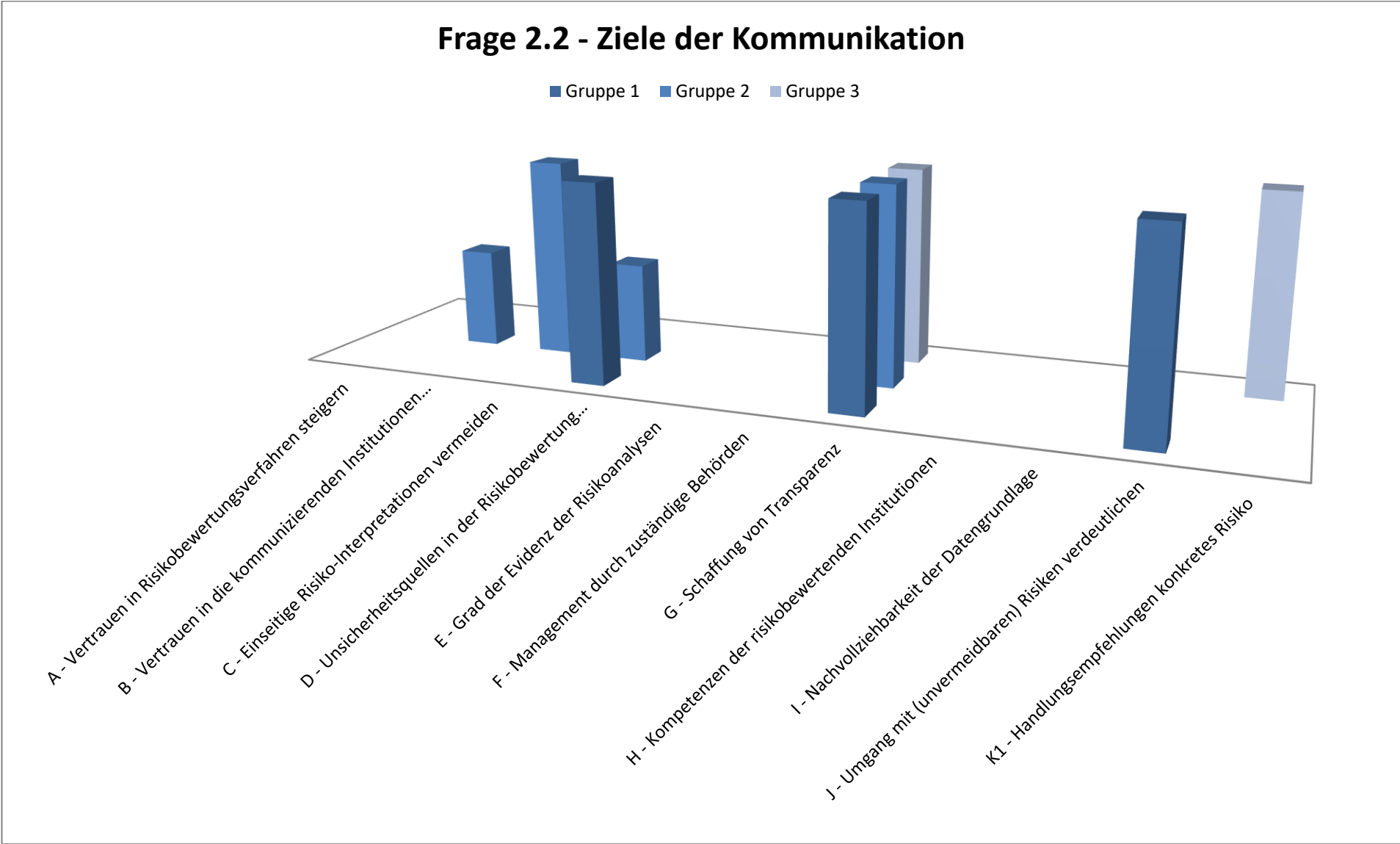
31	III: Kommunikation von Messdaten	Bürgerforen und Bürgerparlamente (Effektivität)	10	5	2	1	3	2,7	1,7	2,9	<u>0,6</u>
32	III: Kommunikation von Messdaten	Informationsbroschüren (Effektivität)	10	5	-	8	2	6,5	1,5	2,3	0,2
33	III: Kommunikation von Messdaten	Zeitungen und Zeitschriften (Effektivität)	10	9	-	8	2	8,5	0,5	0,3	0,1
34	III: Kommunikation von Messdaten	Sonstiges Kommunikationsmedium (Effektivität)	Websites (eigene)								
35	III: Kommunikation von Messdaten	Pressekonferenzen (Glaubwürdigkeit)	10	7	-	5	2	6,0	1,0	1,0	0,2
36	III: Kommunikation von Messdaten	Bürger-Gesprächsgruppen (Glaubwürdigkeit)	10	8	-	5	2	6,5	1,5	2,3	0,2
37	III: Kommunikation von Messdaten	Virtuelle Treffen und Online Konferenzen (Glaubwürdigkeit)	10	5	-	5	2	5,0	0,0	0,0	0,0
38	III: Kommunikation von Messdaten	Konferenzen und Workshops (Glaubwürdigkeit)	10	10	-	8	2	9,0	1,0	1,0	0,1
39	III: Kommunikation von Messdaten	Blogs und Foren (Glaubwürdigkeit)	10	5	-	5	2	5,0	0,0	0,0	0,0
40	III: Kommunikation von Messdaten	Youtube und andere Portale (Glaubwürdigkeit)	10	5	-	5	2	5,0	0,0	0,0	0,0
41	III: Kommunikation von Messdaten	TV und Radio (Glaubwürdigkeit)	10	8	-	5	2	6,5	1,5	2,3	0,2
42	III: Kommunikation von Messdaten	Wissenschaftl. Publikationen (Glaubwürdigkeit)	10	10	-	10	2	10,0	0,0	0,0	0,0
43	III: Kommunikation von Messdaten	Bürgerforen und Bürgerparlamente (Glaubwürdigkeit)	10	5	-	5	2	5,0	0,0	0,0	0,0
44	III: Kommunikation von Messdaten	Informationsbroschüren (Glaubwürdigkeit)	10	7	-	5	2	6,0	1,0	1,0	0,2
45	III: Kommunikation von Messdaten	Zeitungen und Zeitschriften (Glaubwürdigkeit)	10	9	-	8	2	8,5	0,5	0,3	0,1

46	III: Kommunikation von Messdaten	Sonstiges Kommunikationsmedium (Glaubwürdigkeit)	Eigene Websites, gestufte Informationen								
47	IV: Konkrete Maßnahmen	Apps	10	8	5	1	3	4,7	2,9	8,2	<u>0,6</u>
48	IV: Konkrete Maßnahmen	Labels	10	7	8	1	3	5,3	3,1	9,6	<u>0,6</u>
49	IV: Konkrete Maßnahmen	Messungen	10	3	7	1	3	3,7	2,5	6,2	<u>0,7</u>
50	IV: Konkrete Maßnahmen	Messgeräte	10	3	5	1	3	3,0	1,6	2,7	<u>0,5</u>
51	IV: Konkrete Maßnahmen	Sonstige	-								

Runde 1: Frage 1.2 - Zusammenarbeit von Akteuren



Runde 1: Frage 2.2 - Ziele der Kommunikation



Runde 2: 10er Skalen

Frage-Nr.	Thematischer Bereich der Frage	Frageinhalt	Skala	Antwort Gruppe 1	Antwort Gruppe 2	Antwort Gruppe 3	Anzahl Antworten	Arithmetisches Mittel	Standardabweichung	Varianz	Variationskoeffizient
80	I: Kommunikationsakteure	Einschätzung der IARC kommunizieren	10	-	2	10	2	6,0	4,0	16,0	<u>0,7</u>
1	I: Kommunikationsakteure	Internationale Organisationen	10	10	6	8	2	9,0	1,0	1,0	0,1
2	I: Kommunikationsakteure	Nationale Organisationen	10	1	7	1	3	3,0	2,8	8,0	<u>0,9</u>
3	I: Kommunikationsakteure	Nationale Behörden	10	10	5	10	3	8,3	2,4	5,6	0,3
4	I: Kommunikationsakteure	Journalisten	10	10	10	5	3	8,3	2,4	5,6	0,3
5	I: Kommunikationsakteure	Wissenschaftler	10	7	10	5	3	7,3	2,1	4,2	0,3
6	I: Kommunikationsakteure	Organisierte Öffentlichkeit	10	1	4	5	3	3,3	1,7	2,9	<u>0,5</u>
7	I: Kommunikationsakteure	Hardwarehersteller	10	1	1	5	3	2,3	1,9	3,6	<u>0,8</u>
8	I: Kommunikationsakteure	Content Provider	10	1	-	-	1	1,0	0,0	0,0	0,0
9	I: Kommunikationsakteure	Politiker auf nationaler Ebene	10	1	2	5	3	2,7	1,7	2,9	<u>0,6</u>
10	I: Kommunikationsakteure	Stiftungen	10	1	1	5	3	2,3	1,9	3,6	<u>0,8</u>
11	I: Kommunikationsakteure	Verbraucherschutzorganisationen	10	1	10	10	3	7,0	4,2	18,0	<u>0,6</u>
81	I: Kommunikationsakteure	Berufsständische Organisationen im Gesundheitsbereich	10	10	-	10	2	10,0	0,0	0,0	0,0
82	I: Kommunikationsakteure	Fachgesellschaften	10	1	2	10	3	4,3	4,0	16,2	<u>0,9</u>

84	I: Kommunikationsakteure	Netzbetreiber	10	1	2	2	3	1,7	0,5	0,2	0,3
85	I: Kommunikationsakteure	Fachbezogene Vereine / Organisationen	10	10	2	5	3	5,7	3,3	10,9	<u>0,6</u>
13	II: Kommunikationsinhalte	Datengrundlage	10	5	6	10	3	7,0	2,2	4,7	0,3
14	II: Kommunikationsinhalte	Datenbewertung	10	5	6	10	3	7,0	2,2	4,7	0,3
15	II: Kommunikationsinhalte	Handlungsempfehlungen	10	10	10	10	3	10,0	0,0	0,0	0,0
16	II: Kommunikationsinhalte	Unsicherheiten	10	7	6	10	3	7,7	1,7	2,9	0,2
17	II: Kommunikationsinhalte	Nutzungssituation	10	10	10	10	3	10,0	0,0	0,0	0,0
18	II: Kommunikationsinhalte	Fachlicher Expertise / Ausrichtung	10	5	6	10	3	7,0	2,2	4,7	0,3
19	II: Kommunikationsinhalte	Unabhängigkeit	10	10	10	10	3	10,0	0,0	0,0	0,0
20	II: Kommunikationsinhalte	Risikomanagement	10	5	9	10	3	8,0	2,2	4,7	0,3
21	II: Kommunikationsinhalte	Forschungsergebnisse mit vgl. Risiken	10	5	5	10	3	6,7	2,4	5,6	0,4
22	II: Kommunikationsinhalte	Schema der Risikoklassifikation	10	7	3	10	3	6,7	2,9	8,2	0,4
23	III: Kommunikation von Messdaten	Pressekonferenzen (Effektivität)	10	10	10	10	3	10,0	0,0	0,0	0,0
24	III: Kommunikation von Messdaten	Bürger-Gesprächsgruppen (Effektivität)	10	3	3	10	3	3,3	0,5	0,2	0,1
25	III: Kommunikation von Messdaten	Virtuelle Treffen und Online Konferenzen (Effektivität)	10	1	3	10	3	1,7	0,9	0,9	<u>0,6</u>
26	III: Kommunikation von Messdaten	Konferenzen und Workshops (Effektivität)	10	1	10	4	3	6,3	3,9	14,9	<u>0,6</u>
27	III: Kommunikation von Messdaten	Blogs und Foren (Effektivität)	10	3	3	1	3	2,3	0,9	0,9	0,4
28	III: Kommunikation von Messdaten	Youtube und andere Portale (Effektivität)	10	2	2	2	2	2,0	0,0	0,0	0,0

29	III: Kommunikation von Messdaten	TV und Radio (Effektivität)	10	10	7	10	3	9,0	1,4	2,0	0,2
30	III: Kommunikation von Messdaten	Wissenschaftl. Publikationen (Effektivität)	10	1	10	2	3	7,0	4,2	18,0	<u>0,6</u>
31	III: Kommunikation von Messdaten	Bürgerforen und Bürgerparlamente (Effektivität)	10	1	4	10	3	2,0	1,4	2,0	<u>0,7</u>
32	III: Kommunikation von Messdaten	Informationsbroschüren (Effektivität)	10	5	8	10	3	6,0	1,4	2,0	0,2
33	III: Kommunikation von Messdaten	Zeitungen und Zeitschriften (Effektivität)	10	10	10	1	3	10,0	0,0	0,0	0,0
86	III: Kommunikation von Messdaten	Eigene Websites (Effektivität)	10	10	-	10	2	10,0	0,0	0,0	0,0
35	III: Kommunikation von Messdaten	Pressekonferenzen (Glaubwürdigkeit)	10	8	10	8	3	8,7	0,9	0,9	0,1
36	III: Kommunikation von Messdaten	Bürger-Gesprächsgruppen (Glaubwürdigkeit)	10	-	3	5	2	4,0	1,0	1,0	0,3
37	III: Kommunikation von Messdaten	Virtuelle Treffen und Online Konferenzen (Glaubwürdigkeit)	10	-	2	1	2	1,5	0,5	0,3	0,3
38	III: Kommunikation von Messdaten	Konferenzen und Workshops (Glaubwürdigkeit)	10	9	10	10	3	9,7	0,5	0,2	0,0
39	III: Kommunikation von Messdaten	Blogs und Foren (Glaubwürdigkeit)	10	3	5	1	3	3,0	1,6	2,7	<u>0,5</u>
40	III: Kommunikation von Messdaten	Youtube und andere Portale (Glaubwürdigkeit)	10	3	2	3	3	2,7	0,5	0,2	0,2
41	III: Kommunikation von Messdaten	TV und Radio (Glaubwürdigkeit)	10	8	5	6	3	6,3	1,2	1,6	0,2
42	III: Kommunikation von Messdaten	Wissenschaftl. Publikationen (Glaubwürdigkeit)	10	10	10	10	3	10,0	0,0	0,0	0,0
43	III: Kommunikation von Messdaten	Bürgerforen und Bürgerparlamente (Glaubwürdigkeit)	10	-	2	5	2	3,5	1,5	2,3	0,4
44	III: Kommunikation von Messdaten	Informationsbroschüren (Glaubwürdigkeit)	10	-	6	5	2	5,5	0,5	0,3	0,1

45	III: Kommunikation von Messdaten	Zeitungen und Zeitschriften (Glaubwürdigkeit)	10	10	10	8	3	9,3	0,9	0,9	0,1
83	III: Kommunikation von Messdaten	„gestufte“ Informationen der zuständigen Behörden in Print und Internet (Glaubwürdigkeit)	10	-	9	9	2	9,0	0,0	0,0	0,0
87	III: Kommunikation von Messdaten	Eigene Websites (Glaubwürdigkeit)	10	-	-	8	1	8,0	0,0	0,0	0,0
47	IV: Konkrete Maßnahmen	Apps	10	5	2	1	3	2,7	2,9	2,9	<u>0,6</u>
48	IV: Konkrete Maßnahmen	Labels	10	-	5	1	2	3,0	4,0	4,0	<u>0,7</u>
49	IV: Konkrete Maßnahmen	Messungen	10	1	3	1	3	1,7	0,9	0,9	<u>0,6</u>
50	IV: Konkrete Maßnahmen	Messgeräte	10	1	1	1	3	1,0	0,0	0,0	0,0

11.8 Ausgewertete Aussagen: Normative Dimension in SEAWIND

	Argument
<p>Item 13: Die Datengrundlage zur Festlegung der Risikogruppe.</p>	<p>Es wurde empfohlen, nicht die Datengrundlage zu veröffentlichen, sondern zu erklären, wie die Bewertungen zustande gekommen wären, um den Bewertungsprozess transparent zu gestalten. (Quelle: Hilpert et al. 2013: 17)</p>
<p>Item 80: Ist es Ihrer Meinung nach notwendig, die Einschätzung der IARC in die breite Öffentlichkeit zu kommunizieren?</p>	<p>Eine Gruppe war der Meinung, dass die Einschätzungen des IARC in die Öffentlichkeit kommuniziert werden müssten, eine andere Gruppe lehnte dies ab und eine weitere enthielt sich der Bewertung (...). Die Gruppe die sich für eine aktive Kommunikation aussprach, begründete dies mit der Relevanz der IARC als Autorität in diesem Gebiet und der großen Relevanz der Nachricht. Die Inhalte müssten jeweils auf die Gruppe abgestimmt werden. Die Gruppe die sich gegen eine aktive Kommunikation aussprach, begründete dies damit, dass eine aktive Kommunikation nicht notwendig sei, weil man damit die potentiellen Ansprechpartner nur verunsichern würde. Den meisten Menschen könne man die Relevanz der Einstufung nicht plausibel vermitteln. Nur auf Nachfrage solle entsprechend reagiert werden (Right to know, but not need to know). (Quelle: Hilpert et al. 2013: 14.f.)</p>
<p>Frage 5: Kommunikation bei hoher Unsicherheit</p>	<p>(...) Risiken sollten vielmehr darauf abgeklöpft werden, ob sie evtl. kontraproduktiv wirken. Was könnte möglicherweise der Effekt auf die Gesamtpopulation sein? Verzicht auf (WLAN) Strahlung hat wahrscheinlich größere Auswirkungen als der Aufruf auf den Verzicht. Keine Bewusste Kommunikation um etwas „zurückzunehmen“ sondern um einen öffentlichen Diskurs zu führen. Man sollte eher auf die Diskussion reagieren. So ist im Bereich EMF der Fall, dass der Anfangsverdacht seit 10 – 15 Jahren in der Bevölkerung bekannt ist, aber die Einstufung der IARC kein Mehrwert darstellt. Es sollte die Frage nach dem Mehrwert hinter einer Meldung gestellt werden. (Quelle: Ergebnisprotokoll)</p> <p>Als Argument gegen die Kommunikation wurde von einem teilnehmenden Experten eine Studie zum Thema „risk avoidance“ angeführt. Die Ergebnisse würden dagegen sprechen, in einem so frühen Stadium Risiken zu kommunizieren. Die Kommunikation von Risiken dürfe nicht kontraproduktiv sein und müsse immer auf ihren Mehrwert überprüft werden. Die Experten sprachen sich jedoch auch nicht einhellig gegen eine aktive Kommunikation unter Unsicherheit aus. Es sei entscheidend Vertrauen und Transparenz zu schaffen und einen offenen Diskurs über die Chancen und Risiken zu führen. (Quelle: Hilpert et al. 2013: 32)</p>

Legende: (1) Werturteile; (2) Präferenzen; (3) (individuelle) Bewertungen von Optionen ; (4) moralische Urteile; Aussagekontext

Expertendelphi

Stromsparen im Haushalt fördern durch Smart Meter

„Soziale, öffentliche, ökonomische und politische Akzeptanz von Fördermaßnahmen zur Verbesserung der Effizienz im Bereich Elektrizitätsnutzung und von nachhaltiger Mobilität bei privaten Haushalten“

gefördert durch das Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (BMU) im Rahmen der nationalen Klimaschutzinitiative

Abschnitt I: Derzeitige Gesetzeslage

Frage 1.1:

Mit der Richtlinie über Endenergieeffizienz und Energiedienstleistungen (EDL-Richtlinie) hat die Europäische Union im Jahr 2006 die Grundlage für die Einführung von intelligenten Zählern (Smart Meter – kurz: SM) geschaffen. Die Richtlinie sieht vor, dass individuelle Zähler, die den tatsächlichen Energieverbrauch des Endkunden und die tatsächliche Nutzungszeit anzeigen, zu wettbewerbsorientierten Preisen eingebaut werden sollen. Abrechnungen für die Verbraucher sollen auf Grundlage des tatsächlichen Verbrauchs so häufig erstellt werden, dass der Kunde die Möglichkeit hat, seinen Energieverbrauch bewusst zu steuern. Ziel dieser Vorgabe ist es, die Grundlage für ein den Energieverbrauch senkendes Verhalten zu schaffen.

Die EDL-Richtlinie wurde in Deutschland im Jahr 2008 durch das Gesetz zur Öffnung des Messwesens umgesetzt, mit dem das Energiewirtschaftsgesetz (EnWG) novelliert wurde. §21b des EnWG sieht vor, dass Messstellenbetreiber ab dem 01.01.2010 bei Gebäuden, die neu an das Energieversorgungsnetz angeschlossen werden oder einer grundlegenden Renovierung unterzogen werden, Messeinrichtungen einbauen, die den tatsächlichen Energieverbrauch und die tatsächliche Nutzungszeit widerspiegeln. Ferner sollen die Messstellenbetreiber allen übrigen Kunden ebenfalls solche Messeinrichtungen anbieten. Die Bestandskunden können dieses Angebot ablehnen. §40 des EnWG sieht vor, dass die Energieversorgungsunternehmen (EVUs) für alle Verbraucher (unabhängig davon, ob bei ihnen bereits ein intelligenter Stromzähler verbaut wurde) bis zum 30.12.2010 einen Tarif anbieten, der einen Anreiz zur Energieeinsparung oder zur Steuerung des Energieverbrauchs setzt – insbesondere lastvariable oder tageszeitabhängige Tarife.

Das Gesetz schreibt eine Funktion zur Datenübertragung nicht vor. Monatliche Abrechnungen des tatsächlichen Verbrauchs sind für den Anbieter verpflichtend, sofern der Kunde dies wünscht. Persönliche Feedbackinstrumente, die eine Auswertung und graphische Aufbereitung des Verbrauchs und ggf. der Kosten für verschiedene Zeiträume ermöglichen, sind nicht vorgeschrieben. Persönliche Feedbackinstrumente sind etwa Auswertungsprogramme für den PC, Internetportale, in denen der Kunde seinen Verbrauch analysieren kann, mobile Home-Displays oder Smartphone-Applikationen. Entsprechende Lösungen werden bereits von vielen Anbietern als Zusatzangebot angeboten und sind z.T. mit zusätzlichen Kosten für den Kunden verbunden. Voraussetzung für monatliche Abrechnungen und persönliche Feedbackinstrumente ist eine Datenübertragung, die auf verschiedenen Wegen erfolgen kann, etwa über eine DSL-Leitung oder GPRS (Mobilfunk).

Im Folgenden finden Sie allgemeine Bewertungen der Gesetzeslage.

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Durch diese gesetzliche Regelung wird die <u>Durchschlagskraft</u> im Hinblick auf die Verbreitung von Smart Metern erhöht.										
B	Diese Maßnahme ist im Hinblick auf das Stromsparen im Haushalt als <u>effektiv</u> zu bewerten.										

Abschnitt II: Stromsparerpotential von Smart Metern; Produktvarianten

Frage 2.1:

In dieser Frage geht es darum, das Stromsparerpotential von Smart-Metern zu bewerten.
Bitte tragen Sie einen Wert in Prozent ein.

		Einsparerpotential in %
A	Wie viel Prozent Einsparung halten Sie bei der Nutzung von Smart Metern für <u>maximal</u> möglich?	_____ %
B	Wie viel Prozent Einsparung ist in einem <u>Durchschnittshaushalt</u> zu erwarten?	_____ %

Frage 2.2:

Die Umsetzung von Smart Metering kann in verschiedenen Varianten erfolgen. Für wie erfolgversprechend bewerten Sie die einzelnen Varianten bezogen auf die Effektivität im Hinblick auf das Stromsparen?

Bitte tragen Sie einen Wert ein. Geben Sie 1 (-) an, wenn Sie die Effektivität der Variante im Hinblick auf das Stromsparen niedrig bewerten, geben sie 10 (+) an, wenn Sie diese hoch bewerten.

	Varianten	-	2	3	4	5	6	7	8	9	+
A	Einfache Smart Meter <u>ohne Datenübertragungsfunktion</u>										
B	Smart Meter <u>mit Datenübertragungsfunktion</u> und <u>monatlicher informativer Abrechnung</u>										
C	<p>Smart Meter <u>mit Datenübertragungsfunktion</u> und <u>persönlichen Feedbackinstrumenten</u> über Internetportal, PC-Software, Home-Display oder Smartphone-App.</p> <p>Die Feedbackinstrumente beinhalten folgende Informationen:</p> <ul style="list-style-type: none"> - Lastgänge zum Gesamtstromverbrauch - Gerätespezifisches Verbrauchsfeedback - Bewertung des Verbrauchs durch Vergleich des Stromverbrauchs mit strukturähnlichen Haushalten bzw. den energieeffizientesten Geräten am Markt - Kombination von Verbrauchsfeedback mit bereichsspezifischen Handlungsempfehlungen zu energieeffizienten Geräten und Nutzungsweisen - Strategien zum Setzen von eigenen Stromsparzielen und zur Erfolgskontrolle durch den Smart Meter - Erinnerungshilfen 										

Abschnitt III: Verpflichtende Einführung von Smart Metern für alle Haushalte

Frage 3.1:

Smart Meter haben über die gesetzlich geforderten Basislösungen in Neubauten und bei grundlegenden Sanierungen hinaus bislang lediglich eine sehr eingeschränkte Verbreitung gefunden. Wie sich in den Fokusgruppen gezeigt hat und weitere Studien belegen, sind ein wesentlicher Grund dafür die Kostenstrukturen (Einmalkosten für Geräte und Einbau, im Vergleich zu den Standardtarifen höhere monatliche Grundpreise, höhere kWh-Preise außerhalb der Schwachlastzeiten). Eingesparte Kosten durch Verbrauchsreduktionen werden durch diese Mehrkosten oftmals aufgezehrt.

Im Folgenden finden Sie einen weitergehenden Ausbau von Smart Metern als bisher vorgesehen:

Smart Meter mit Datenübertragungsfunktion müssen verpflichtend von jedem Energieversorger bzw. Netzbetreiber bei allen Kunden zu dem Zeitpunkt eingebaut werden, wenn der alte Zähler geeicht werden müsste. Mechanische Zähler müssen je nach Typ alle 12 oder 16 Jahre geeicht werden.

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Diese Maßnahme ist <u>politisch durchsetzbar</u> .										

Wie würden Sie die zusätzlich entstehenden Kosten verteilen. Bitte verteilen Sie 100% auf die drei möglichen Akteure

	Staat	Energieversorgungsunternehmen <i>Es wird gesetzlich festgelegt, dass dieser Anteil durch die EVUs getragen wird</i>	Privathaushalte
B	_____ %	_____ %	_____ %

Frage 3.2:

Im Folgenden finden Sie einen weitergehenden Ausbau von Smart Metern als bisher vorgesehen:
Smart Meter mit Datenübertragungsfunktion werden verpflichtend eingebaut.

Wie beurteilen Sie eine Kombination mit persönlichen Feedbackinstrumenten über Internetportal, PC-Software, Home-Display oder Smartphone-App als zusätzliche Maßnahme?

Die Feedbackinstrumente beinhalten folgende Informationen (vgl. Einführungsvortrag):

- Lastgänge zum Gesamtstromverbrauch
- Gerätespezifisches Verbrauchsfeedback
- Bewertung des Verbrauchs durch Vergleich des Stromverbrauchs mit strukturähnlichen Haushalten bzw. den energieeffizientesten Geräten am Markt
- Kombination von Verbrauchsfeedback mit bereichsspezifischen Handlungsempfehlungen zu energieeffizienten Geräten und Nutzungsweisen
- Strategien zum Setzen von eigenen Stromsparzielen und zur Erfolgskontrolle durch den Smart Meter
- Erinnerungshilfen

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Die Ergänzung des Smart Meters mit Datenübertragungsfunktion <u>durch persönliche Feedbacklösungen</u> erachte ich für wichtig.										
B	Die <u>Effektivität</u> im Hinblick auf das Stromsparen wird durch diese Kombination gesteigert.										
C	Diese Kombination ist <u>politisch durchsetzbar</u> .										

Wie würden Sie die zusätzlich entstehenden Kosten verteilen? Bitte verteilen Sie 100% auf die drei möglichen Akteure.

	Staat	Energieversorgungsunternehmen <i>Es wird gesetzlich festgelegt, dass dieser Anteil durch die EVUs getragen wird</i>	Privathaushalte
D	_____ %	_____ %	_____ %

Abschnitt IV: Förderung der Verbreitung von Smart Metern durch ein Markteinführungsprogramm und empfohlene Tarifgestaltungen (keine verpflichtende Einführung für alle Haushalte)

Frage 4.1:

Smart Meter haben über die gesetzlich geforderten Basislösungen in Neubauten und bei grundlegenden Sanierungen hinaus bislang lediglich eine sehr eingeschränkte Verbreitung gefunden. Wie sich in den Fokusgruppen gezeigt hat und weitere Studien belegen, sind ein wesentlicher Grund dafür die Kostenstrukturen (Einmalkosten für Geräte und Einbau, im Vergleich zu den Standardtarifen höhere monatliche Grundpreise, höhere kWh-Preise außerhalb der Schwachlastzeiten). Eingesparte Kosten durch Verbrauchsreduktionen werden durch diese Mehrkosten oftmals aufgezehrt.

Welche Maßnahmen halten Sie für am besten geeignet, um die Verbreitung von Smart Metern (kombiniert mit monatlicher informativer Abrechnung und Feedbackinstrumenten) im Rahmen eines Markteinführungsprogramms zu fördern?

Begleitende kommunikative Maßnahmen werden hier vorausgesetzt, die detaillierte Diskussion darüber erfolgt in der nächsten Frage.

Die **drei** wichtigsten, erfolgversprechendsten Maßnahmen im Hinblick auf Verbreitung bitte ankreuzen.

Maßnahmen		Kreuzen Sie die drei wichtigsten Vorschläge an!
A	Subvention der einmaligen Einbaukosten für EVUs unter der Bedingung, dass eine vorgegebene Mindestanzahl SM pro Jahr eingebaut wird	
B	Subvention der einmaligen Einbaukosten für Verbraucher (analog zur „Abwrackprämie“ im Kfz-Bereich)	
C	Subvention des monatlichen Grundpreises für Verbraucher	
D	Steuerliche Begünstigungen des Stromverbrauchs für die SM nutzenden Endverbraucher, d.h. der Strompreis sinkt für diese Verbraucher	
E	Steuerliche Begünstigungen für die SM anbietenden Unternehmen unter der Bedingung, dass eine vorgegebene Mindestanzahl SM pro Jahr eingebaut wird	
F	Ergänzendes Bonus- oder Prämienangebot, z.B. kostenlose Nutzung eines Elektromobils für einen Tag	
G	Leasingmodelle für die Anschaffung der Geräte; Gesamtkosten über die Laufzeit dürfen nicht höher sein als die Kosten bei Einmalzahlung	
H	An SM gekoppelte Tarife mit einem Kontingent an Inklusiv-KWh, wobei sich die Inklusivstrommenge dabei an einem sehr stromsparenden Durchschnittsverbraucher orientiert	
I	Progressive Tarife gekoppelt an den SM: deutlicher Preisanstieg bei überdurchschnittlichem Verbrauch, dafür aber geringere Preise im Bereich unterhalb des Durchschnittsverbrauchs	
J	Sonstiges: _____	

Abschnitt V: Tarifstrukturen gebunden an Smart Meter, die das Stromsparen fördern können

Frage 5.1:

Smart Meter schaffen Transparenz bzgl. des Stromverbrauchs. Ergänzende geeignete Tarifstrukturen können dazu beitragen, einen Anreiz zu schaffen, den Stromverbrauch effektiv zu reduzieren.

Welche an Smart Meter gekoppelte Tarifstrukturen können am besten dazu beitragen, das Stromsparen mit Hilfe von Smart Metern zu fördern?

Die **zwei** wichtigsten, erfolgversprechendsten Tarifstrukturen im Hinblick auf erfolgreiches Energiesparen bitte ankreuzen.

Maßnahmen		Kreuzen Sie die zwei wichtigsten Vorschläge an!
A	Progressive Tarife gekoppelt an den SM: deutlicher Preisanstieg bei überdurchschnittlichem Verbrauch, dafür aber geringere Preise im Bereich unterhalb des Durchschnittsverbrauchs	
B	Anreiz zum Stromsparen durch last- oder zeitvariable Stromtarife mit deutlicher Spreizung (Teurer Spitzenlasttarif fördert Sparen, wenn man nicht auf die Schwachlastzeiten ausweichen möchte/kann)	
C	Anreizschaffung durch Prepaid-Funktionalität (Bewusstsein schaffen für Stromverbrauch, der Strom kommt nicht mehr einfach aus der Steckdose, sondern muss aktiv eingekauft werden)	
D	Gestaffeltes Bonussystem oder Prämienzahlung bei deutlicher Verbrauchsreduktion als Tarifergänzung	
E	Sonstiges: _____	

Abschnitt VI: Wie kann man EVUs motivieren, Anreize zum Stromsparen zu schaffen?

EVUs sind geeignete Partner, wenn es darum geht, das Stromsparen in Haushalten zu fördern, denn sie haben direkten Kontakt zu ihren Kunden. Allerdings spricht es gegen ihr derzeitiges Geschäftsmodell, Stromeinsparungen bei Kunden zu fördern, denn damit sind Umsatzverluste für sie verbunden. In den Fokusgruppen hat sich zudem gezeigt, dass auch die Glaubwürdigkeit der EVUs, sich ernsthaft für das Stromsparen zu engagieren, von vielen Teilnehmern angezweifelt wird. Dass EVUs ihre Kunden zum Stromsparen motivieren, obgleich ihr Firmenziel der Verkauf von Strom ist, erscheint den meisten nicht plausibel. Beispiele aus anderen Ländern zeigen, dass durch geeignete staatliche Maßnahmen aber sowohl die Motivation der EVUs, sich beim Stromsparen der Haushalte zu engagieren, als auch die Glaubwürdigkeit von EVUs gegenüber der Bevölkerung, wenn sie Stromsparmaßnahmen fördern, erhöht werden kann.

Im Folgenden finden Sie zwei Beispiele für solche Maßnahmen.

Frage 6.1:

Gesetzlich festgelegte Verpflichtung der EVUs zur Einsparung vorgegebener Energiemengen in einem festgelegten Zeitraum (Einsparquote). Kosten der EVUs für Einsparmaßnahmen bei den Verbrauchern werden in die Strompreise einkalkuliert.

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Die <u>Durchsetzbarkeit</u> dieser Verpflichtung gegenüber den EVUs ist als hoch zu bewerten.										
B	Diese Verpflichtung ist im Hinblick auf das Stromsparen als <u>effektiv</u> zu bewerten.										
C	Diese Verpflichtung wird von den <u>Kunden akzeptiert</u> werden.										

Frage 6.2:

Gesetzlich festgelegte Verpflichtung der EVUs zur Einsparung vorgegebener Energiemengen in einem festgelegten Zeitraum (Einsparquote). Ergänzend dazu wird ein Handelssystem für Einsparzertifikate (sog. weiße Zertifikate) eingerichtet. Zertifikate werden in diesem System über Energieeffizienzmaßnahmen generiert, die EVUS bei ihren Kunden durchführen. Flexibilität für die EVUs wird darüber erreicht, dass sie entscheiden können, ob sie zur Erfüllung der Einsparquote ganz oder teilweise selber Einsparmaßnahmen ergreifen und / oder Einsparzertifikate von anderen EVUs kaufen, die ihre Einsparquote übererfüllt haben. Einsparungen finden in diesem System an der Stelle statt, wo sie am kosteneffizientesten durchzuführen sind.

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Ein Handelssystem mit weißen Zertifikaten ist als <u>durchsetzbar</u> gegenüber den EVUs zu bewerten.										
B	Ein Handelssystem mit weißen Zertifikaten ist im Hinblick auf das Stromsparen <u>effektiv</u> .										
C	Ein Handelssystem mit weißen Zertifikaten wird von den Kunden <u>akzeptiert</u> werden.										

Abschnitt VII: Kommunikative Maßnahmen

Frage 7.1:

Neben einem Markteinführungsprogramm zur besseren Verbreitung der Smart Meter und Maßnahmen, die Anreize zum Stromsparen über an SM gekoppelte Tarife setzen, sind begleitende kommunikative Maßnahmen von großer Bedeutung, um das Stromsparen mittels Smart Metern im Haushalt zu fördern. Insbesondere solche kommunikativen Maßnahmen können vom BMU im Rahmen der Nationalen Klimaschutzinitiative gefördert werden.

Im Folgenden finden Sie einige Szenarien, die bezüglich der kommunikativen Vermittlung von Smart Metern denkbar wären.

Vorausgesetzt wird hier, dass veränderte Rahmenbedingungen herrschen, in denen die Verbraucher über die Nutzung der Smart Meter beim Stromsparen auch Kosten sparen.

Bitte bewerten Sie diese Szenarien bezüglich ihres Erfolgs hinsichtlich der Verbreitung von Smart Metern. Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass dieses Szenario überhaupt nicht erfolversprechend ist, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass dieses Szenario sehr erfolversprechend ist.

Für welche Zielgruppen ist das jeweilige Szenario besonders gut geeignet? Bitte tragen sie in die Spalte der Zielgruppen die Zahl der Zielgruppe(n) ein, für die Sie das Szenario als erfolversprechend bewerten.

Zielgruppen:

1. 18-30 jährige Paare / Singles ohne Kinder im Haushalt
2. 31-50 jährige Paare / Alleinerziehende mit Kindern im Haushalt
3. 51-65 jährige Paare / Singles ohne Kinder im Haushalt
4. Rentner
5. Hartz IV-Empfänger

Bitte beachten Sie auch die Szenarien auf der folgenden Seite.

Szenario	Szenarienbeschreibung	-	2	3	4	5	6	7	8	9	+	Für welche Zielgruppen (1-5) ist diese Maßnahme besonders gut geeignet?
7A	<p>Klassische Medien: Fernsehen, Rundfunk, Zeitung, Flyer Problem- und handlungsorientierte Information sowie Vorbilder nutzen bei Kommunikationsangeboten:</p> <ul style="list-style-type: none"> • Bekannte Persönlichkeiten aus Politik und Gesellschaft werben für SM; • Videoclips mit Modellfamilien, WGs usw., überzeugte Nutzer kommen zu Wort <p>Akteure: BMU, DENA, Energieagenturen, EVUs</p>											

		-	2	3	4	5	6	7	8	9	+	
7B	<p>Soziale Netzwerke im Internet Einrichtung einer „Fan-Plattform“(Facebook o.a.) / eines Blogs, Bereitstellung von Informationen, Kommunikation zwischen den „Fans“</p> <p>Akteure: BMU, Kommunikationsagenturen, Multiplikatoren, engagierte Privatpersonen</p>											
7C	<p>Persönliche Beratung / Einführung durch Servicepersonal beim Einbau und im Vorfeld des Kaufs</p> <p>Akteure: Kundenservice von EVUs, Stadtwerke, Netz- und Messstellenbetreiber, Verbraucherzentralen</p>											
7D	<p>Soziale Settings für Kommunikationsangebote nutzen: Schulen, Kindergärten, Universitäten Infoveranstaltungen, Vorführung, praktische Nutzung, Beratungsangebote, schriftliche Materialien</p> <p>Akteure: BMU, Schulen, Kindergärten, Kommunen, Universitäten</p>											
7E	<p>Soziale Settings für Kommunikationsangebote nutzen: Wohnungsbaugesellschaften und Wohnheime Begleitende Informationskampagne zur Einführung: Demonstration, Beratung, schriftliche Materialien, Wettbewerbe</p> <p>Akteure: BMU, DENA, Wohnungswirtschaft</p>											
7F	<p>Kommunikationsangebote an Bürger durch Kommunen</p> <ul style="list-style-type: none"> • Kommunale Kampagnen mit Plakaten, Flyern u.a. • Meldestellen: Kommunale Beratung zu SM beim Wohnungswechsel: schriftlich und persönlich • Informationsbroschüren für Eigentümer beim Wohnungskauf und für Mieter <p>Akteure: BMU, Kommunen, Stadtwerke</p>											
7G	<p>Halten Sie andere Szenarien für wichtig? Wenn ja, welche ?</p>											

Frage 7.2:

Welche Inhalte sollten bei kommunikativen Maßnahmen zu Smart Metern thematisiert werden, um zum Kauf und zur Nutzung zu motivieren und das Strom sparen im Haushalt optimal zu unterstützen? Für wie wichtig erachten Sie folgende Themen:
 Geben Sie 1 (-) an, wenn Sie die Themen für überhaupt nicht wichtig erachten und 10 (+), wenn Sie die Themen für sehr wichtig erachten.

Problemorientierte Informationen		-	2	3	4	5	6	7	8	9	+
A	Klimarelevanz, CO2-Einsparung, Ressourcenschonung durch Energieeinsparung										
B	Vorsorge für nachfolgende Generationen, globale Gerechtigkeit										
C	Energieversorgung der Zukunft – Smart Meter in den Kontext setzen von : Erhöhung des Anteils erneuerbarer Energien im Energiemix, Einführung von Smart Grids, Lastmanagement										
D	Datenschutz gewährleisten										
E	Selbstwirksamkeit / Kontrolle des eigenen Verbrauchs / der Kosten										
F	Good- / Best-Practice-Beispiele aus anderen Ländern										
G	Transparenz schaffen hinsichtlich des Umfangs der tatsächlichen Nutzung von Smart Metern und der erzielten Verbrauchseinsparungen in der Bundesrepublik (Verbraucher erkennt sich als Teil einer gemeinsamen Anstrengung)										
H	Sonstiges: <hr/>										

Frage 7.2 (Fortsetzung):

Welche Inhalte sollten bei kommunikativen Maßnahmen zu Smart Metern thematisiert werden, um zum Kauf und zur Nutzung zu motivieren und das Strom sparen im Haushalt optimal zu unterstützen? Für wie wichtig erachten Sie folgende Themen:
 Geben Sie 1 (-) an wenn Sie die Themen für überhaupt nicht wichtig erachten und 10 (+) wenn Sie die Themen für sehr wichtig erachten.

	Handlungsorientierte Informationen	-	2	3	4	5	6	7	8	9	+
I	Die größten Stromsparpotentiale im Haushalt – Vorschläge für wirksame Maßnahmen										
J	Unterstützung von Kaufentscheidungen durch Checks für vorhandene elektrische Haushaltsgeräte und Ausweisung der stromeffizientesten Geräte am Markt										
K	Unterstützung des stromsparenden Nutzungsverhaltens durch Stromspartipps und aufzeigen von Strategien, wie man Hemmnisse leicht überwinden kann										
L	Stromsparpotential aufzeigen durch Vergleich mit strukturell ähnlichen Haushalten										
M	Vermittlung von Handlungsstrategien, wie Smart Meter zur Unterstützung des Stromsparens einfach genutzt werden können										
N	Vermittlung von Handlungsstrategien, wie andere Haushaltsmitglieder (z.B. Kinder, Jugendliche, Wohngemeinschaftsmitbewohner) zum Stromsparen motiviert und in Stromsparbemühungen eingebunden werden können.										
O	Sonstiges: <hr/>										

Abschnitt VIII: Einschätzung der Bedeutung von Smart Metering, Handlungsempfehlungen für die Bundesregierung

Frage 8.1:

Aller Einsparbemühungen und Effizienzsteigerungen zum Trotz: der Stromverbrauch in privaten Haushalt steigt weiter. Obwohl die Elektrogeräte in den Haushalten immer energieeffizienter werden, verhindern eine steigende Anzahl an Geräten und veränderte Nutzungsgewohnheiten einen Rückgang des Stromverbrauchs.

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Eine zügige Einführung von SM kombiniert mit monatlicher informativer Abrechnung und persönlichen Feedbackinstrumenten (vgl. Einführungsvortrag) ist im Hinblick auf das Stromsparen als bedeutsam zu erachten.										

Frage 8.2:

Welche Handlungsempfehlungen würden Sie der Bundesregierung geben, um die **Verbreitung** von Smart Metern zu fördern?

Frage 8.3:

Welche Handlungsempfehlungen würden Sie der Bundesregierung geben, um Smart Meter im Hinblick auf das Stromsparen **effizienter** zu machen?

Vielen Dank für die Beantwortung des Fragebogens !

Expertendelphi

Stromsparen im Haushalt fördern durch Smart Meter

Statistische Ergebnisse

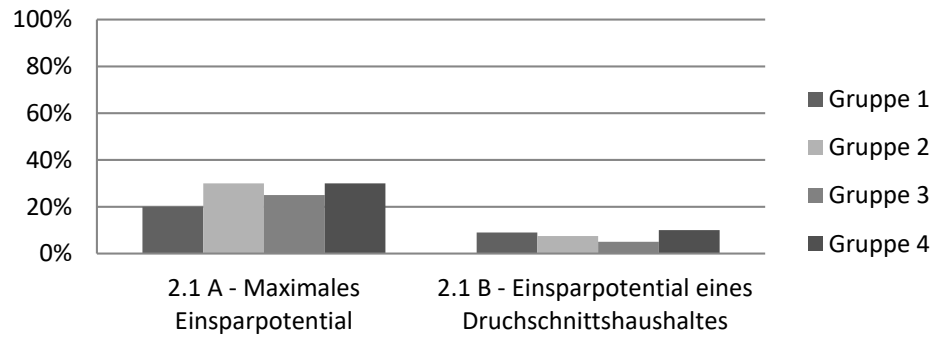
Runde 1: 10er Skalen

Frage-Nr.	Frageinhalt	Skala	Antwort Gruppe 1	Antwort Gruppe 2	Antwort Gruppe 3	Minderheit Gruppe 3	Antwort Gruppe 4	Minderheit Gruppe 4	N	Arithmetisches Mittel	s	s ²	Variationskoeffizient
1.1 A	Durchschlagskraft im Hinblick auf die Verbreitung von Smart Metern	10	2	1	2		2		4	1,8	0,4	0,2	0,2
1.1 B	Effektivität bzgl. Stromsparen im Haushalt	10	2	4	1		2		4	2,3	1,1	1,2	0,5
2.2 A	Erfolg der Varianten: Einfache Smart Meter	10	2	4	1	2	2		5	2,2	1,0	1,0	0,4
2.2 B	Erfolg der Varianten: Smart Meter mit Datenübertragungsfunktion	10	4	7	8	10	5		5	6,8	2,1	4,6	0,3
2.2 C	Erfolg der Varianten: Smart Meter mit Datenübertragungsfunktion und persönlichen Feedbacklösungen	10	7	10	8	10	7		5	8,4	1,4	1,8	0,2
3.1 A	Verpflichtender Einbau von Smart Meter bei Zählereichung: Politische Durchsetzbarkeit	10	10	10	8	10	5		5	8,6	2,0	3,8	0,2
3.2 A	Smart Meter mit Datenübertragungsfunktion werden verpflichtend eingebaut: Persönliche Feedbacklösungen werden als wichtig erachtet	10	10	10	10		7		4	9,3	1,3	1,7	0,1
3.2 B	Smart Meter mit Datenübertragungsfunktion werden verpflichtend eingebaut: Effektivität wird durch diese Kombination gesteigert	10	7	10	10		5		4	8,0	2,1	4,5	0,3
3.2 C	Smart Meter mit Datenübertragungsfunktion werden verpflichtend eingebaut: Diese Kombination ist politisch durchsetzbar	10	10	10	8		4		4	8,0	2,4	6,0	0,3
6.1 A	Gesetzliche Verpflichtung der E-VUs: Durchsetzbarkeit der Verpflichtung	10	1	10	6		3		4	5,0	3,4	11,5	0,7
6.1 B	Gesetzliche Verpflichtung der E-VUs: Effektivität der Verpflichtung	10	1	10	9		6		4	6,5	3,5	12,3	0,5

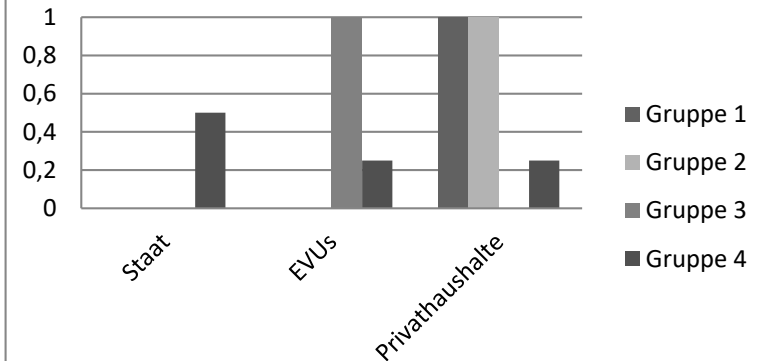
6.1 C	Gesetzliche Verpflichtung der EVUs: Kundenakzeptanz der Verpflichtung	10	1	10	8		2		4	5,3	3,8	14,7	0,7
6.2 A	Einsparquote (Weiße Zertifikate): Durchsetzbarkeit gegenüber den EVUs	10	10		5		2		3	5,7	3,3	10,9	0,6
6.2 B	Einsparquote (Weiße Zertifikate): Effektivität hinsichtlich des Stromsparens	10	1		8		3		3	4,0	2,9	8,7	0,7
6.2 C	Einsparquote (Weiße Zertifikate): Kundenakzeptanz eines Handlungssystems mit weißen Zertifikaten	10	1		7		5		3	4,3	2,5	6,2	0,6
7.1 A	Szenario: Klassische Medien	10	3	2	6		9		4	5,0	2,7	7,5	0,5
7.1 B	Szenario: Soziale Netzwerke im Internet	10	3	3	2		7		4	3,8	1,9	3,7	0,5
7.1 C	Szenario: Persönliche Beratung / Einführung	10	7	10	9		6		4	8,0	1,6	2,5	0,2
7.1 D	Szenario: Soziale Settings für Kommunikationsangebote nutzen (Schulen Kindergärten, Universitäten)	10	5	8	4		8		4	6,3	1,8	3,2	0,3
7.1 E	Szenario: Soziale Settings für Kommunikationsangebote nutzen (Wohnungsbaugesellschaften und Wohnheime)	10	5	6	8		8		4	6,8	1,3	1,7	0,2
7.1 F	Szenario: Kommunikationsangebote an Bürger durch Kommunen	10	4	3	4	6	6		5	4,6	1,2	1,4	0,3
7.1 G	Szenario: Sonstiges	10							0				
7.2 A	Problemorientierte Informationen: Klimarelevanz, CO2-Einsparung, Ressourcenschonung durch Energieeinsparung	10	9	4			8		3	7,0	2,2	4,7	0,3
7.2 B	Problemorientierte Informationen: Vorsorge für nachfolgende Generationen, globale Gerechtigkeit	10	9	6			7		3	7,3	1,2	1,6	0,2
7.2 C	Problemorientierte Informationen: Energieversorgung der Zukunft	10	10	6			9		3	8,3	1,7	2,9	0,2
7.2 D	Problemorientierte Informationen: Datenschutz gewährleisten	10	8	5			8		3	7,0	1,4	2,0	0,2

7.2 E	Problemorientierte Informationen: Selbstwirksamkeit / Kontrolle des eigenen Verbrauchs / der Kosten	10	6	9			9		3	8,0	1,4	2,0	0,2
7.2 F	Problemorientierte Informationen: Good- / Best-Practice-Beispiele aus anderen Ländern	10	5	3			4		3	4,0	0,8	0,7	0,2
7.2 G	Problemorientierte Informationen: Transparenz	10	3	6			6		3	5,0	1,4	2,0	0,3
7.2 H	Problemorientierte Informationen: Sonstiges	10							0				
7.2 I	Handlungsorientierte Informationen: Die größten Stromfresser im Haushalt – Vorschläge für wirksame Maßnahmen	10	9	10			9		3	9,3	0,5	0,2	0,1
7.2 J	Handlungsorientierte Informationen: Unterstützung von Kaufentscheidungen durch Checks	10	7	10			8		3	8,3	1,2	1,6	0,1
7.2 K	Handlungsorientierte Informationen: Unterstützung des stromsparenden Nutzungsverhaltens durch Stromspartipps	10	4	10			8		3	7,3	2,5	6,2	0,3
7.2 L	Handlungsorientierte Informationen: Stromsarpotential aufzeigen durch Vergleich mit strukturell ähnlichen Haushalten	10	5	10			9		3	8,0	2,2	4,7	0,3
7.2 M	Handlungsorientierte Informationen: Vermittlung von Handlungsstrategien, wie Smart Meter genutzt werden können	10	6	10			8		3	8,0	1,6	2,7	0,2
7.2 N	Handlungsorientierte Informationen: Vermittlung von Handlungsstrategien, wie andere Haushaltsmitglieder zum Stromsparen motiviert werden können	10	6	10			8		3	8,0	1,6	2,7	0,2
7.2 O	Handlungsorientierte Informationen: Sonstiges	10							0				
8.1 A	Bedeutsamkeit einer zügigen Einführung von Smart Metern	10	7	7			3	10	4	6,8	2,5	6,2	0,4

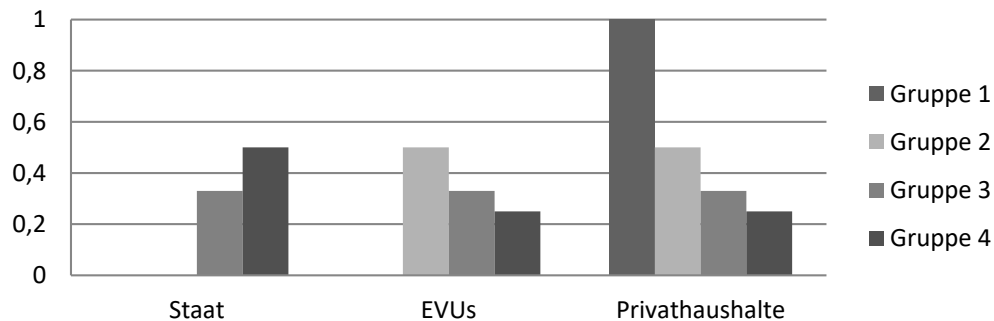
Frage 2.1 - Einsparpotential von Smart Metern



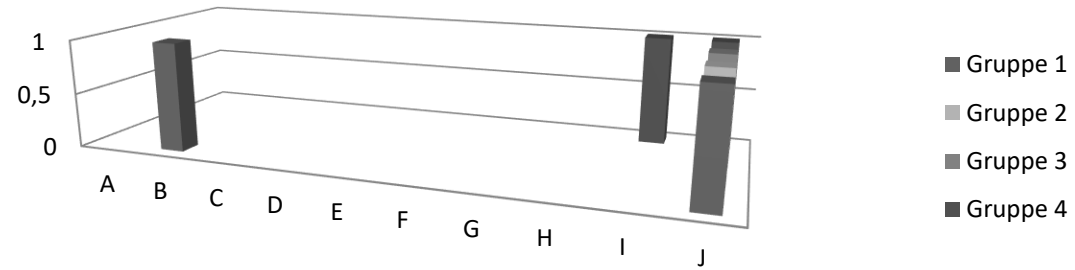
Frage 3.1 B - Kostenverteilung (SM mit Datenübertragungsfunktion)



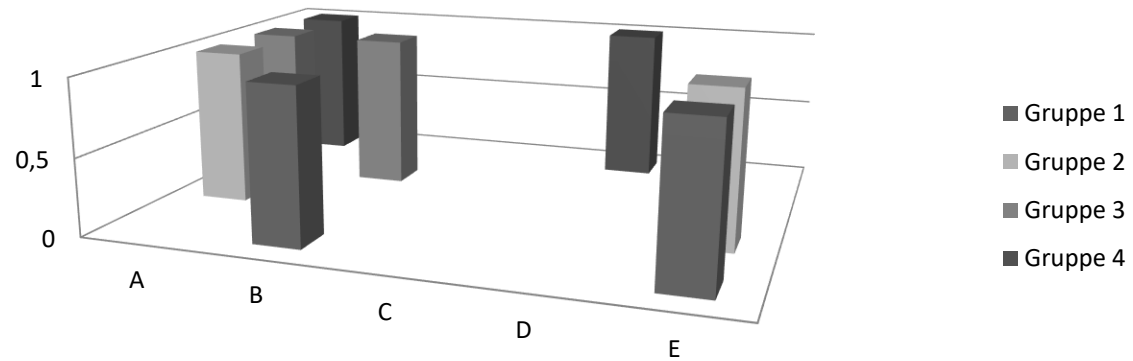
Frage 3.2 D - Kostenverteilung (Kombination mit Feedbacklösungen)



Frage 4.1 - Maßnahmen Förderung der Verbreitung von Smart Metern



Frage 5.1 Maßnahmen bzgl. Tarifstrukturen



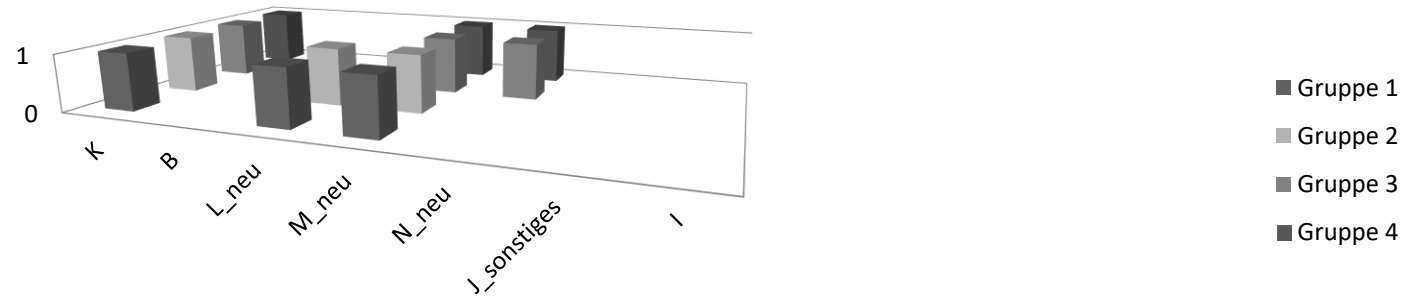
Runde 1: Szenarien und Zielgruppenverteilung

Item Nr.	Szenario	Arithmetisches Mittel	s	Zielgruppen (Gruppe1)	Zielgruppen (Gruppe2&3)		Zielgruppen (Gruppe 4)
7.1 A	Szenario: Klassische Medien	5,0	2,7	1,2,3,4,5	-		1,2,3,4,5
7.1 B	Szenario: Soziale Netzwerke im Internet	3,8	1,9	1	-		1,2
7.1 C	Szenario: Persönliche Beratung / Einführung	8,0	1,6	2,3	-		2,3,4,5
7.1 D	Szenario: Soziale Settings für Kommunikationsangebote nutzen (Schulen Kindergärten, Universitäten)	6,3	1,8	1,2,3	-		1,2,4
7.1 E	Szenario: Soziale Settings für Kommunikationsangebote nutzen (Wohnungsbaugesellschaften und Wohnheime)	6,8	1,3	1,2,3,4	-		1,2,3,4,5
7.1 F	Szenario: Kommunikationsangebote an Bürger durch Kommunen	4,6	1,2	1,2,3,4	-		1,2,3,4,5
7.1 G	Szenario: Sonstiges			Leuchtturm öffentliche Einrichtungen	(Gruppe 2) -	(Gruppe 3) Stromsparchecks für alle	Aufklärung zu Smart Grid im Kontext zu Smart Metering

Runde 2: 10er Skalen

Frage-Nr.	Frageinhalt	Neues Item	Skala	Antwort Gruppe 1	Minderheit Gruppe 1	Antwort Gruppe 2	Antwort Gruppe 3	Antwort Gruppe 4	Minderheit Gruppe 4	N	Arithmetisches Mittel	s	s ²	Variationskoeffizient
2.2 B	Erfolg der Varianten : Smart Meter mit Datenübertragungsfunktion und monatlicher informativer Abrechnung	nein	10	7		6	7	6		4	6,5	0,5	0,3	0,1
2.2 C	Erfolg der Varianten: Smart Meter mit Datenübertragungsfunktion und persönlichen Feedbacklösungen	nein	10	7		7	6	8	4	5	6,4	1,4	1,8	0,2
2.2 D_neu	Erfolg der Varianten: Kopplung von monatlicher Abrechnung und Feedback	ja	10	9	8	8	9	9	8	6	8,5	0,5	0,3	0,1
7.2 K	Handlungsorientierte Informationen: Unterstützung des stromsparenden Nutzungsverhaltens durch Stromspartipps	nein	10	10		7		6		3	7,7	1,7	2,9	0,2
8.1 A	Verpflichtender Einbau von Smart Meter bei Zählereichung: Politische Durchsetzbarkeit	nein	10	9	8	8	9	9	8	6	8,5	0,5	0,3	0,1
8.1 A_neu	Smart Meter mit Datenübertragungsfunktion werden verpflichtend eingebaut: Persönliche Feedbacklösungen werden als wichtig erachtet	ja	10	7	8	6	7			4	7,0	0,7	0,5	0,1

Frage 4.1 - Maßnahmen Förderung der Verbreitung von Smart Metern



3.3_neu - Relevante Akteure:

- Endverbraucher
- Messdienstleister
- Messstellenbetreiber
- Netzbetreiber (ÜNB/VNB)
- Vertrieb
- Erzeuger (regenerativ/konventionell)
- Bund
- Land
- Kommune
- EU
- Energiehandel

11.11

Ausgewertete Aussagen: Normative Dimension (Smart Meter)

	Erste Plenarsitzung	Zweite Plenarsitzung	Ergebnisprotokoll
Frage 4.1 (Förderung der Verbreitung von Smart Metern durch ein Markteinführungsprogramm und empfohlene Tarifgestaltungen)	<p>Stromeinsparung sei hingegen nur „nice to have“ und rangiere allenfalls an dritter Stelle hinter Lastverlagerung und Kostensparpotential. Dieser Meinung schlossen sich in der Diskussion mehrere Teilnehmer an.</p> <p>Gruppe 3 stellt die Kostenseite in den Mittelpunkt. Derzeit werden die Kosten von den Kunden getragen, SM-Lösungen sollten für die Kunden aber kostenneutral sein. Hierzu kam aus dem Plenum der Einwand, dass der Kunde letztendlich immer zahle, entweder individuell oder sozialisiert.</p>		
Frage 3.1 B und 3.2 D (Kostenverteilung auf Staat, EVUs und Privathaushalte) & 3.3 Relevante Akteure (t2)	<p>Wichtig ist nach Meinung des Plenums die zeitliche Dimension: in der Einführungsphase sei ein zeitlich begrenzter Anschub durch den Staat durchaus sinnvoll.</p>	<p>Die Kostenverteilung sollte sich danach richten, wer profitiert und dann eher entsprechende Kosten (mit-)tragen kann.</p>	
Frage 5.1 Tarifstrukturen, die das Stromsparen fördern können - Maßnahme A (Progressive Tarife)	<p>Zu Maßnahme A (Progressive Tarife) wurde diskutiert, ob eine Unterscheidung zwischen „gutem Strom“ (geringer Verbrauch) und „bösem Strom“ (darüber hinausgehender Verbrauch) sinnvoll sei. So könne es durchaus Sondersituationen geben, in denen ein erhöhter Stromverbrauch auftritt, der sich aber nicht durch Energiesparen lösen lässt, wie etwa Nachwuchs, Besuch oder die Hausinfrastruktur. Progressive Tarife können zu Ungerechtigkeiten führen. In der Formulierung der Maßnahme A ist von „überdurchschnittlichem“ Verbrauch die Rede. Hierzu wurde gefragt, was ein durchschnittlicher Verbrauch ist und wie dieser ermittelt werde. Unter Berücksichtigung der sozialen Dimension seien Konstrukte wie Durchschnittsverbräuche fragwürdig.</p>		<p>Die Experten wiesen an dieser Stelle darauf hin, dass das Stromsparen nicht im Vordergrund steht, sondern an erster Stelle die Integration der erneuerbaren Energien.</p>
Frage 6.1 (Einsparquote)	<p>Als Argument für die Ablehnung wurde genannt, dass die Versorger nicht für die Einsparungen bei</p>		

	<p>den Kunden verantwortlich gemacht werden sollten, der Staat solle hier nicht „über Bande spielen“. Vielmehr müsse der Staat andere Instrumente entwickeln, um direkt bei den Kunden zu wirken. Zielsetzung der Anbieter sei es, Strom zu verkaufen, Stromsparen liege nicht in ihrem Interesse. Für die Anbieter ergebe sich ein Glaubwürdigkeitsproblem, wenn sie plötzlich zum Stromsparen animieren sollten. Die Netzbetreiber sollten hier über das Regulierungskonto Ausgleichsmöglichkeiten für Mindereinnahmen erhalten.</p> <p>Vorgeschlagen wurden für den Akteur Netzbetreiber auch pauschale Tarife (Flatrates) für die Dienstleistung des Transports mit einem festen Preis für die Durchleitung unabhängig von der Menge.</p>		
<p>Frage 7.2 (Inhalte von kommunikativen Maßnahmen)</p>	<p>Die im Fragebogen vorgesehene Differenzierung müsse aufrecht erhalten bleiben. Im Mittelpunkt sollte dabei die Vermittlung einer Vision stehen: die Energiewende, für die Smart Meter zwingend notwendig sind.</p>		
<p>Frage 7.1 (Kommunen)</p>		<p>Kommunen sind zwar nah an den Bürgern, es stellt sich aber die Frage, ob Kommunen Mittel zur Flyererstellung erhalten sollten oder ob dafür nicht andere Akteure (Profiteure von SM) herangezogen werden sollten.</p> <p>Kommunen sind andererseits aber für neutrale Informationen wichtig.</p> <p>Kommunen gehen selbst häufig schlampig mit Energie um. Sie sollten selber mit der Installation von SM vorangehen (Vorbildwirkung).</p> <p>Kommunen sind zur Neutralität verpflichtet (keine Werbung für Anbieter).</p>	

Präambeln	<p>Smart Metering ist in einem Gesamtkontext aus Smart Grid und Smart Home zu sehen. Insbesondere sind dabei die Integration und Ausbau von erneuerbaren Energien, Energieeffizienzerhöhung und Energieeinsparung in dem Gesamtkontext zu berücksichtigen.</p> <p>Der Staat hat Interesse an CO2-Ansparung, der Kunde hat Interesse, Kosten zu sparen und die Energieversorger ein Interesse optimal Erlöse zu gewinnen. Nur wenn sich alle diese Interessen gemeinsam realisieren, ist der Einsatz von Smart Metering flächendeckend erfolgreich.</p>
	<p>Politischer Nutzen: Unterstützung der Umwelt und Klimaschutzziele durch die langfristige Integration von erneuerbare Energien mit ihrer volatilen Verfügbarkeit in die Netzinfrastruktur und Lastflüsse. → Zielerreichung durch zeitliche Lastverteilung</p> <p>Volkswirtschaftlicher Nutzen und Nutzen beim regulierten natürlichen Monopol (Netzbetreiber / "Messtellenbetreiber"): Reduzierung von Netzinfrastrukturinvestitionen durch verbesserte Auslastung der heutigen Netze und Reduzierung der anstehenden Neuinvestitionen. → Zielerreichung durch zeitliche und räumliche Lastverteilung (dezentrale Einspeisung) („smart Grid“)</p> <p>Nutzen für Lieferanten und Kunden: Stromkosteneinsparung durch Berücksichtigung des tatsächlichen Lastverlauf beim Einkauf und setzen von Preissignalen und Preisanreize für den Kunden. → Zielerreichung durch Bilanzierbarkeit des tatsächlich gemessenen Lastprofils Anstelle von Standardlastprofile</p> <p>Kundennutzen und politischer Nutzen: Energieeinsparung durch transparente Darstellung des Energieverbrauches und damit verbunden Verhaltensänderung beim Kunden. → Zielerreichung durch transparente Darstellung des Energieverbrauchs</p>
	<p>Es ist ein durchgängiges Konzept von der politischen Zielstellung bis zum Nutzen für den Endkunden durch aktive Einbindung aller Akteure zu schaffen. Es müssen Anreize durchgängig geschaffen werden für Endkunden, Lieferant/Vertrieb, MSB/MDL, Netzbetreiber, Erzeuger / dezentraler Erzeuger (Anlagenbetreiber).</p> <p>Es müssen dynamische Transfers durchgängig abbildbar sein:</p> <ul style="list-style-type: none"> • Smart Grid i.w.S. (e-Mobility, Smart Home) – Fähigkeit <p>Es müssen weitere Sparten (Wärme, Wasser etc.) berücksichtigt werden</p>
	<p>Kommunikation zu SM immer im Zusammenhang mit SG und Integration EE. Mehr Transparenz, mehr Ehrlichkeit, was kann ein SM wirklich für wen tun. Keine reine Kosteneinsparkommunikation. SM in Zusammenhang mit einer Gesamtstrategie stellen: Konsumer → Prosumer</p>

Legende: (1) Werturteile; (2) Präferenzen; (3) (individuelle) Bewertungen von Optionen ; (4) moralische Urteile; Aussagekontext

Expertendelphi

Kraftstoffeinsparung durch Eco-Fahrassistentensysteme fördern

„Soziale, öffentliche, ökonomische und politische Akzeptanz von Fördermaßnahmen zur Verbesserung der Effizienz im Bereich Elektrizitätsnutzung und von nachhaltiger Mobilität bei privaten Haushalten“

Abschnitt I: Spritsparpotential von Eco-Fahrassistenten

Frage 1:

Bewerten Sie das Spritsparpotential von Eco-Fahrassistentensystemen.
Bitte tragen Sie einen Wert in Prozent ein.

		Einsparpotential in %
A	Wie viel Prozent Einsparung halten Sie bei der Nutzung von Eco-Fahrassistentensystemen für <u>maximal</u> möglich?	_____ %
B	Wie viel Prozent Einsparung ist bei einem <u>Durchschnittsautofahrer</u> zu erwarten?	_____ %

Abschnitt II: Gestaltung von Eco-Fahrassistenten

Im Folgenden werden verschiedene Fragen thematisiert:

Welche Informationen sollten durch einen Eco-Fahrassistenten vermittelt werden, welche Qualität sollten diese Informationen haben und welche Gestaltungsmerkmale sollten ihn auszeichnen, um einen effizienten Fahrstil optimal zu motivieren und zu unterstützen?

Welche weiteren Informationen können außerdem dazu beitragen den Spritverbrauch zu senken und nachhaltige Mobilität zu fördern?

Frage 2.1:

Welche Informationen sollten dem Fahrer durch einen Eco-Fahrassistenten vermittelt werden, um einen effizienten Fahrstil optimal zu unterstützen?

Geben Sie bitte 1 (-) an, wenn Sie der Ansicht sind, dass diese Information überhaupt nicht wichtig ist, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass diese Information sehr wichtig ist.

Informationen während der Fahrt		-	2	3	4	5	6	7	8	9	+
A	Informationen zu effizienter / ineffizienter Beschleunigung										
B	Erlaubte Höchstgeschwindigkeit in der Stadt und auf der Landstraße										
C	Effiziente / ineffiziente Geschwindigkeit auf der Autobahn: Bei höheren Geschwindigkeiten als 130 km/h werden beide Geschwindigkeiten hinsichtlich Zeit und Verbrauch pro Strecke verglichen										
D	Information zum rechtzeitig Ausrollen lassen anstatt stark zu bremsen										
E	Informationen zur Gleichmäßigkeit des Tempos beim Fahren										
F	Sonstige:										

Frage: 2.2

Welche zusätzlichen Informationen sollten an den Fahrer vermittelt werden, um den Spritverbrauch zu senken und nachhaltige Mobilität zu fördern? Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass diese Information überhaupt nicht wichtig ist, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass diese Information sehr wichtig ist.

Informationen zur Auswahl der Route und des Verkehrsmittels (vor und während der Fahrt)		-	2	3	4	5	6	7	8	9	+
A	Informationen zur spritsparendsten Route zum Zielort: Die spritsparendste Route wird mit alternativen Routen hinsichtlich Zeit und Verbrauch verglichen										
B	Informationen zu multimodaler Verkehrsmittelnutzung (P & R, ...)										
C	Informationen über die Verfügbarkeit von Parkplätzen, um unnötige Weg bei der Suche zu sparen.										
D	Informationen zu Mitfahrzentrale und Carsharing										
E	Sonstige:										
Informationen zu effizienten Verhaltensweisen, die den Spritverbrauch senken		-	2	3	4	5	6	7	8	9	+
F	Klimaanlage ausschalten, Spriteinsparung veranschaulichen										
G	Zusammenhang von Spritverbrauch und unnötigem Ballast und Aufbauten wie Dachgepäckträger herstellen										
H	Zusammenhang von Spritverbrauch und regelmäßiger Wartung herstellen										
I	Sonstige:										

Frage 2.3:

Gestaltungsmerkmale eines Eco-Fahrassistenten, die einen spritsparenden Fahrstil unterstützen und / oder die Motivation, spritsparend zu fahren, positiv beeinflussen.

Geben Sie bitte 1 (-) an, wenn Sie der Ansicht sind, dass diese Gestaltungsmerkmale überhaupt nicht wichtig sind, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass diese Gestaltungsmerkmale sehr wichtig sind.

		-	2	3	4	5	6	7	8	9	+
A	Abschaltbarkeit aller Eco-Fahrassistentenfunktionen										
B	Wählbarkeit verschiedener Informationen (während und nach der Fahrt)										
C	Basisinformationen zu spritsparenden Fahrstilaspekten (Beschleunigen, Ausrollen, gleichmäßig Fahren, effiziente Geschwindigkeit) werden während der Fahrt permanent rückgemeldet und sind nicht abschaltbar.										
D	Feedback / Information ist <u>situationsadäquat</u> : (1) Autobahn, Landstraße, Stadtverkehr werden unterschieden; (2) durch GPS zugängliche Informationen werden berücksichtigt (z.B. Geschwindigkeitsbegrenzungen werden rechtzeitig mitgeteilt (ausrollen lassen)); (3) Steigungen und Gefälle werden berücksichtigt.										
E	Car-to-car und car-to-infrastructure Informationen wie Ampelphasen werden berücksichtigt										
F	Informationskanal (visuell, akustisch) ist wählbar										
G	Haptische Information (übers Gaspedal oder Lenkrad) wird zur Feinjustierung von Beschleunigen, Ausrollen und gleichmäßig Fahren eingesetzt										
H	Integration des Eco-Fahrassistenten in ein Multifunktionsgerät, z.B. Navi, das hohe Akzeptanz erfährt										
I	Automatisierte Lösungen werden angeboten: (Eco-Knopf, der Gesamtsystem auf einen spritsparenden Fahrstil umstellt)										
J	Ecoscore (in %) als Gesamtbewertung des Fahrstils und als Bewertung spezifischer spritsparender Fahrstilaspekte										
K	Vergleich mit vorhergehenden Fahrten (Spriteinsparung, Kosteneinsparung, ggf. routenbezogen)										

		-	2	3	4	5	6	7	8	9	+
L	Information zur CO ₂ Einsparung pro Fahrt										
M	Vergleich mit anderen Fahrern (desselben Fahrzeugs oder Fahrzeugtyps)										
N	Symbolische Belohnung (Smileys, blühende Blumen, Lob durch Assistent...) als Ergänzung eines Ecoscores										
O	Intuitive Bedienung, selbsterklärend										
P	Konkrete, leicht verständliche Handlungsempfehlungen / -hinweise										
Q	Ein Eco-Fahrassistent sollte für Fahrer, die den spritsparenden Fahrstil lernen wollen, ein Training anbieten, das verschiedene Themen eines spritsparenden Fahrstils gesondert behandelt (z.B. Thema vorausschauend fahren, ausrollen lassen).										
R	Für Fortgeschrittene sollte der Eco-Fahrassistent nur noch über „Fahrfehler“ informieren.										
S	Sonstige:										

Abschnitt III: Finanzielle Anreize zur Förderung eines spritsparenden Fahrstils

Frage 3:

Wie können PKW-Fahrer über finanzielle Anreize motiviert werden, ein Eco-Fahrassistenzsystem zu nutzen und spritsparend zu fahren?

Bitte kreuzen Sie die wichtigste Maßnahme an.

Maßnahmen		Kreuzen Sie die wichtigste Maßnahme an!
A	Eine reine CO ₂ -bezogene KFZ-Steuer, die das jetzige System ablöst. Der CO ₂ -Ausstoß könnte z.B. über ein Messsystem direkt im Auto ermittelt werden.	
B	Eine reine CO ₂ -bezogene KFZ-Steuer, bei der eine progressive Steuerzahlung auferlegt wird, z.B. Steuerbefreiung bis x-Gramm CO ₂ -Ausstoß, danach ansteigend.	
C	Sammeln von Bonuspunkten, anhand eines Ecoscores, der das Fahrverhalten bewertet. Diese Bonuspunkte können dann bei Tankstellen, Werkstätten oder Versicherungen geltend gemacht werden.	
D	Erhöhung der Ökosteuer und damit auch des Benzinpreises	
E	Sonstiges:	

Abschnitt IV: Fördern der Verbreitung von Eco-Fahrassistenten durch regulative Maßnahmen

Frage 4.1: Verordnung des EU-Parlamentes

Mit der Verordnung (EU) 443/2009 hat EU-Parlament und Rat eine Vorgabe zur Einhaltung der durchschnittlichen CO₂-Emissionen der Neuwagenflotte eines Herstellers von 120g/km gegeben.
 130 g CO₂/km sollen durch Verbesserungen bei der Motorentechnik sowie innovative Technologien erreicht werden.
 Im Rahmen des Gesamtkonzepts der Gemeinschaft werden durch weitere Maßnahmen pauschal 10g/km angerechnet (Artikel 1).
 Hersteller, deren flottendurchschnittlicher CO₂-Ausstoß über diesem Grenzwert liegt, müssen bis zu 95€ je g CO₂/km und PKW der Flotte als Abgabe wegen Emissionsüberschreitung bezahlen.

Auf Antrag können zusätzlich bis zu 7g CO₂/km durch Öko-Innovationen zur Reduktion der Zielvorgabe für die durchschnittlichen spezifischen Emissionen jedes Herstellers angerechnet werden (Artikel 12).

Im Folgenden finden Sie allgemeine Bewertungen der Gesetzeslage.

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Wenn Eco-Fahrassistenzsysteme als Öko-Innovation angerechnet werden können, wird dadurch die Verbreitung dieser Systeme gefördert.										

Frage 4.2: Durchführungsverordnung

Was sollte eine Durchführungsverordnung berücksichtigen, um eine Anrechnung von Eco-Assistenten als Öko-Innovation zu ermöglichen?
 (Bitte nur Stichpunkte)

Frage 4.3: Verpflichtender Einbau von Eco-Fahrassistenzsystemen in Neufahrzeuge

Bei der Verbreitung von Eco-Fahrassistenzsystemen stellt sich die Frage, ob es einen verpflichtenden Einbau in Neuwagen seitens der Hersteller geben soll. Dadurch soll die Verbreitung von Eco-Fahrassistenzsystemen sichergestellt werden. Außerdem könnte damit auch eine Standardisierung von Eco-Fahrassistenzsystemen erreicht werden.
Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Eco-Fahrassistenzsysteme sollten verpflichtend in Neufahrzeuge eingebaut werden, um deren Verbreitung optimal zu gewährleisten										
B	Diese Maßnahme ist politisch durchsetzbar.										
C	Diese Maßnahme ist erfolgversprechender als die Förderung der Verbreitung durch die Anrechnung der Eco-Fahrassistenten als Öko-Innovation im Rahmen der oben genannten EU Verordnung 443/2009										
D	Für Eco-Fahrassistenzsysteme sollte ein Standard entwickelt werden.										

Frage 4.4: Angebot von Nachrüstsystemen bei Altfahrzeugen

Um einen spritsparenden Fahrstil breitenwirksam zu fördern, sollten nicht nur Neuwagen sondern auch Altfahrzeuge mit Eco-Fahrassistenten ausgerüstet werden können. Dafür sind Nachrüstsysteme zu entwickeln.
Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Nachrüstsysteme sollten von den Herstellern angeboten werden.										

Abschnitt V: Kommunikative Maßnahmen zur Förderung eines spritsparenden Fahrstils / Nutzung von Eco-Fahrassistenten

Frage 5.1: Argumente

Motivierte und überzeugte Autofahrer werden den spritsparenden Fahrstil gewissenhafter praktizieren. Welche der nachfolgend genannten Argumente sind besonders geeignet, zum Praktizieren eines spritsparenden Fahrstils und zum Kauf und zur tatsächlichen Nutzung von Eco-Fahrassistenten zu motivieren?

Geben Sie 1 (-) an wenn Sie die Themen für überhaupt nicht wichtig erachten und 10 (+) wenn Sie die Themen für sehr wichtig erachten.

	Mögliche Argumente	-	2	3	4	5	6	7	8	9	+
A	Klimarelevanz und CO ₂ -Einsparung										
B	Resourcenschonung durch Energieeinsparung										
C	Vorsorge für nachfolgende Generationen, globale Gerechtigkeit										
D	Persönliche (finanzielle) Vorteile durch Sprit- und Kosteneinsparung										
E	Zusammenhang zwischen Verkehrssicherheit und spritsparendem Fahrstil										
F	Vorbildfunktion für Kinder/Andere										
G	„Grünes“ Image für Firmen und öffentliche Einrichtungen bei Ausrüstung der Dienstwagenflotte										

Mögliche Argumente		-	2	3	4	5	6	7	8	9	+
		H	Neuberechnung der KFZ-Steuer anhand des CO ₂ -Ausstoßes. Dabei sollen die Anteile von Fahrzeuggewicht, Hubraum und Fahrstil am CO ₂ -Ausstoß transparent gemacht werden.								
I	Wirksamkeit von Eco-Fahrassistenten und eines spritsparenden Fahrstils zu anderen Maßnahmen nachhaltiger Mobilität ins Verhältnis setzen										
J	Sonstiges: <hr/>										

Frage 5.2: Kommunikative Maßnahmen

Wie sollte die Einführung von Eco-Fahrassistenten kommunikativ begleitet werden, um ein Bewusstsein für die Bedeutung eines spritsparenden Fahrstils zu schaffen und deren Nutzung zu motivieren?

Welche (zentralen) Akteure können hier gemeinsam / vernetzt aktiv werden?

Was könnte Ihre Institution dazu beitragen?

Welche kommunikativen Maßnahmen könnten sie jeweils durchführen?

Notieren Sie bitte Ihre Ideen und Vorschläge im folgenden Feld!

Welche (zentralen) Akteure können hier gemeinsam / vernetzt aktiv werden?

Was könnte Ihre Institution dazu beitragen?

Welche kommunikativen Maßnahmen könnten sie jeweils durchführen?

Frage 5.3: Kommunikative Begleitmaßnahmen

Im Folgenden geht es um die Bedeutung kommunikativer Maßnahmen

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Kommunikative Maßnahmen, die von verschiedenen Akteuren abgestimmt die Einführung begleiten, sind als bedeutsam zu erachten.										
B	Eco-Fahrassistenten sollten nicht nur im Fahrzeug eingebaut sein und in der Betriebsanleitung beschrieben werden. Es sollten ansprechende, motivierende Informationen zum Assistenten und Fahrstil im Fahrzeug bereitgestellt werden.										

Abschnitt VI: Einschätzung der Bedeutung von Eco-Fahrassistenzsystemen, Handlungsempfehlungen für die Bundesregierung

Frage 6.1:

Geben Sie 1 (-) an, wenn Sie der Ansicht sind, dass die Aussage überhaupt nicht zutrifft, geben Sie 10 (+) an, wenn Sie der Ansicht sind, dass sie sehr zutrifft.

		-	2	3	4	5	6	7	8	9	+
A	Eco-Fahrassistenzsysteme sind im Hinblick auf eine Reduktion des Kraftstoffverbrauchs und damit auch einer CO ₂ -Reduktion als bedeutsam zu erachten.										

Frage 6.2:

Welchen Beitrag kann die Klimaschutzinitiative / das BMU zur Akzeptanz eines spritsparenden Fahrstils und Nutzung von Eco-Fahrassistenten leisten?
Stichworte!

Vielen Dank für die Beantwortung des Fragebogens !

Expertendelphi

Kraftstoffeinsparung durch Eco-Fahrassistentensysteme fördern

Statistische Ergebnisse

Runde 1 – 10er Skalen

Frage-Nr.	Frageinhalt	Skala	Antwort Gruppe 1	Minderheit Gruppe 1	Antwort Gruppe 2	Minderheit Gruppe 2	Antwort Gruppe 3	Minderheit Gruppe 3	Antwort Gruppe 4	Minderheit Gruppe 4	Anzahl Antworten	Arithmetisches Mittel	Standardabweichung	Varianz	Variationskoeffizient
2.1 A	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Informationen zu effizienter / ineffizienter Beschleunigung	10	10	-	5	-	9	-	9	-	4	8,3	1,9	3,7	0,2
2.1 B	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Erlaubte Höchstgeschwindigkeit in der Stadt und auf der Landstraße	10	2	-	2	-	3	-	1	-	4	2,0	0,7	0,5	0,4
2.1 C	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Effiziente / ineffiziente Geschwindigkeit auf der Autobahn	10	8	-	9	-	6	-	6	-	4	7,3	1,3	1,7	0,2
2.1 D	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Information zum rechtzeitig Ausrollen lassen anstatt stark zu bremsen	10	8	-	10	-	9	-	7	-	4	8,5	1,1	1,3	0,1
2.1 E	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Informationen zur Gleichmäßigkeit des Tempos beim Fahren	10	5	-	10	-	7	-	6	-	4	7,0	1,9	3,5	0,3
2.2 A	Gestaltung von Eco-Fahrassistenten - Informationen zur Auswahl der Route und des Verkehrsmittels: Informationen zur spritsparendsten Route zum Zielort	10	9	-	8	-	3	-	2	-	4	5,5	3,0	9,3	0,6

2.2 B	Gestaltung von Eco-Fahrassistenten - Informationen zur Auswahl der Route und des Verkehrsmittels: Informationen zu multimodaler Verkehrsmittelnutzung	10	9	-	7	-	5	-	5	-	4	6,5	1,7	2,8	0,3
2.2 C	Gestaltung von Eco-Fahrassistenten - Informationen zur Auswahl der Route und des Verkehrsmittels: Informationen über die Verfügbarkeit von Parkplätzen	10	4	-	6	-	7	-	6	-	4	5,8	1,1	1,2	0,2
2.2 D	Gestaltung von Eco-Fahrassistenten - Informationen zur Auswahl der Route und des Verkehrsmittels: Informationen zu Mitfahrzentrale und Carsharing	10	3	-	3	-	5	-	2	-	4	3,3	1,1	1,2	0,3
2.2 F	Gestaltung von Eco-Fahrassistenten - Informationen zu effizienten Verhaltensweisen: Klimaanlage ausschalten, Spriteinsparung veranschaulichen	10	6	-	5	-	8	-	8	-	4	6,8	1,3	1,7	0,2
2.2 G	Gestaltung von Eco-Fahrassistenten - Informationen zu effizienten Verhaltensweisen: Ballast und Aufbauten	10	8	-	10	-	8	-	2	-	4	7,0	3,0	9,0	0,4
2.2 H	Gestaltung von Eco-Fahrassistenten - Informationen zu effizienten Verhaltensweisen: Regelmäßige Wartung	10	3	-	7	-	6	-	1	-	4	4,3	2,4	5,7	0,6
2.3 A	Gestaltungsmerkmale eines Eco-Fahrassistenten: Abschaltbarkeit aller Eco-Fahrassistentenzfunktionen	10	8	2	6	-	3	-	6	-	5	5,0	2,2	4,8	0,4
2.3 B	Gestaltungsmerkmale eines Eco-Fahrassistenten: Wählbarkeit verschiedener Infos	10	8	-	8	-	7	-	9	-	4	8,0	0,7	0,5	0,1
2.3 C	Gestaltungsmerkmale eines Eco-Fahrassistenten: Basisinformationen zu spritsparenden Fahrstilaspekten	10	1	-	2	-	4	-	8	-	4	3,8	2,7	7,2	0,7

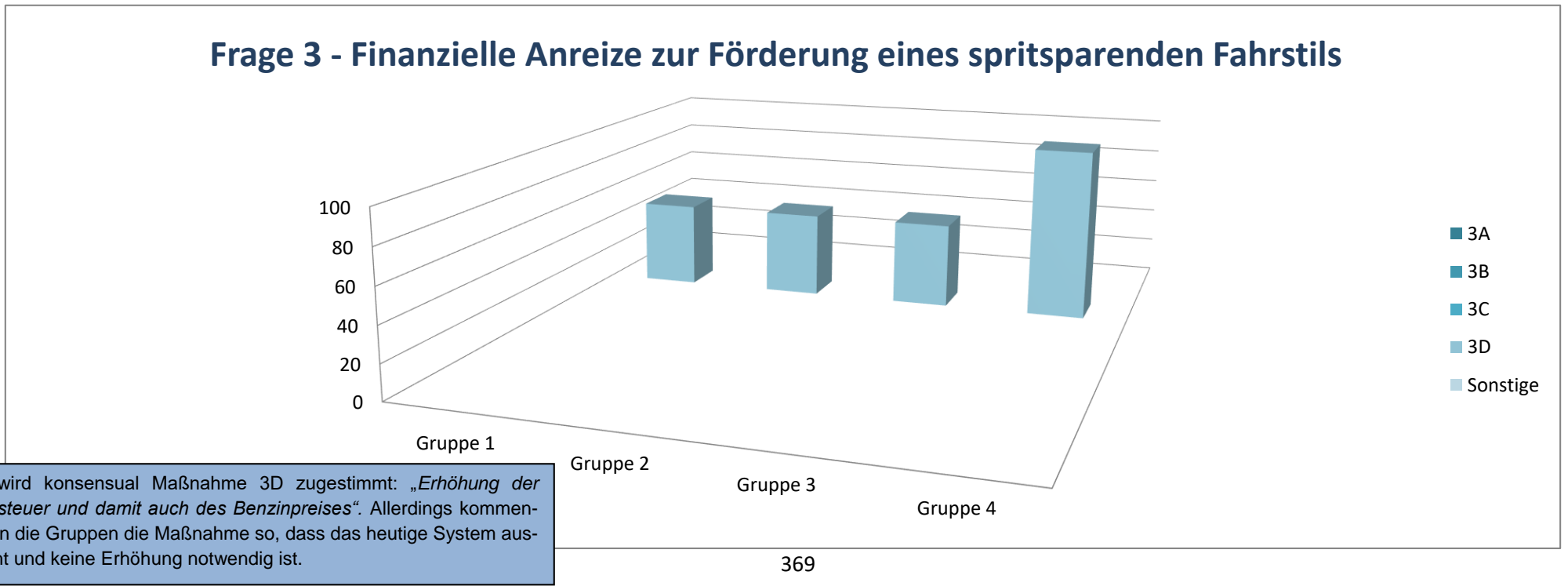
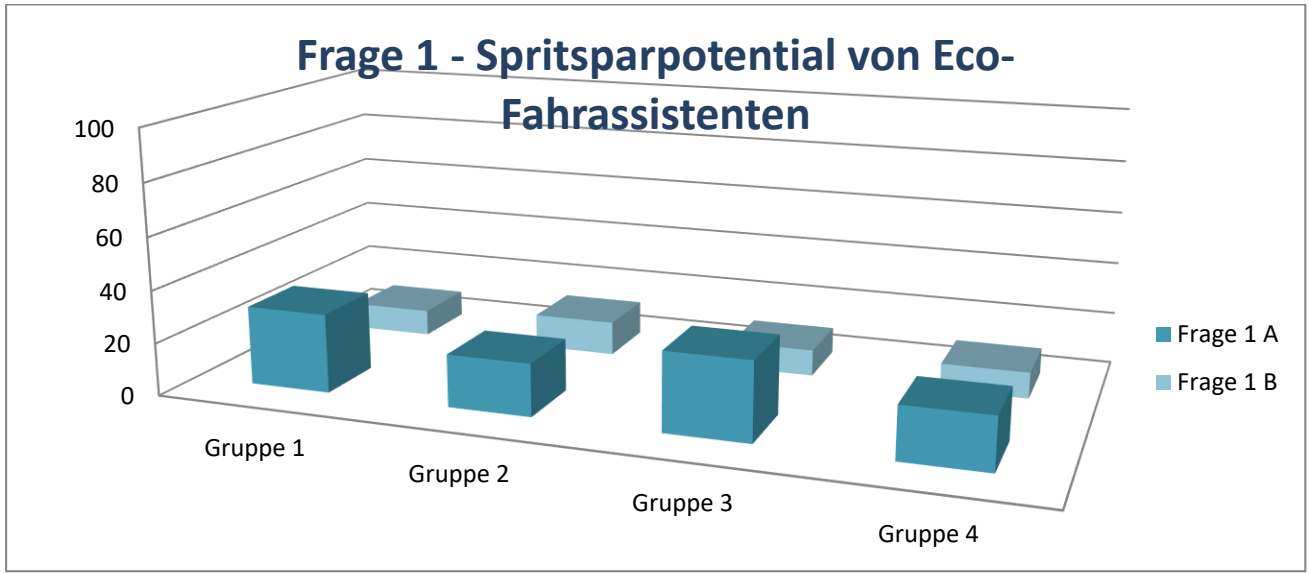
2.3 D	Gestaltungsmerkmale eines Eco-Fahrassistenten: Feedback / Information ist situationsadäquat	10	10	-	10	-	9	-	7	-	4	9,0	1,2	1,5	0,1
2.3 E	Gestaltungsmerkmale eines Eco-Fahrassistenten: Car-to-car und car-to-infrastructure Informationen	10	5	-	10	-	6	-	7	-	4	7,0	1,9	3,5	0,3
2.3 F	Gestaltungsmerkmale eines Eco-Fahrassistenten: Informationskanal	10	8	-	8	-	3	-	8	-	4	6,8	2,2	4,7	0,3
2.3 G	Gestaltungsmerkmale eines Eco-Fahrassistenten: Haptische Info	10	8	-	8	-	7	-	2	6	5	6,2	2,2	5,0	0,4
2.3 H	Gestaltungsmerkmale eines Eco-Fahrassistenten: Integration des Eco-Fahrassistenten in Multifunktionsgerät	10	10	-	9	-	8	-	2	-	4	7,3	3,1	9,7	0,4
2.3 I	Gestaltungsmerkmale eines Eco-Fahrassistenten: Automatisierte Lösungen werden angeboten	10	10	-	10	-	9	-	9	-	4	9,5	0,5	0,3	0,1
2.3 J	Gestaltungsmerkmale eines Eco-Fahrassistenten: Ecoscore	10	9	-	10	-	7	-	8	-	4	8,5	1,1	1,3	0,1
2.3 K	Gestaltungsmerkmale eines Eco-Fahrassistenten: Vergleich mit vorhergehenden Fahrten	10	8	-	8	-	5	-	6	-	4	6,8	1,3	1,7	0,2
2.3 L	Gestaltungsmerkmale eines Eco-Fahrassistenten: Information zur CO2 Einsparung pro Fahrt	10	7	-	5	-	-	-	5	-	3	5,7	0,9	0,9	0,2
2.3 M	Gestaltungsmerkmale eines Eco-Fahrassistenten: Vergleich mit anderen Fahrern	10	8	-	5	8	8	-	1	-	5	6,0	2,8	7,6	0,5
2.3 N	Gestaltungsmerkmale eines Eco-Fahrassistenten: Symbolische Belohnung	10	5	-	9	-	8	-	3	-	4	6,3	2,4	5,7	0,4

2.3 O	Gestaltungsmerkmale eines Eco-Fahrassistenten: Intuitive Bedienung	10	10	-	10	-	9	-	10	-	4	9,8	0,4	0,2	0,0
2.3 P	Gestaltungsmerkmale eines Eco-Fahrassistenten: Konkrete, leicht verständliche Handlungsempfehlungen	10	10	-	10	-	9	-	10	-	4	9,8	0,4	0,2	0,0
2.3 Q	Gestaltungsmerkmale eines Eco-Fahrassistenten: Training	10	10	-	-	-	7	-	5	-	3	7,3	2,1	4,2	0,3
2.3 R	Gestaltungsmerkmale eines Eco-Fahrassistenten: Über Fahrfehler informieren	10	4	-	2	-	5	-	1	-	4	3,0	1,6	2,5	0,5
4.1 A	Regulative Maßnahmen: Eco-Fahrassistenten als Ökoinnovationen	10	10	-	10	-	8	-	9	-	4	9,3	0,8	0,7	0,1
4.3 A	Verpflichtender Einbau: Eco-Fahrassistentensysteme verpflichtend in Neufahrzeuge	10	10	-	10	-	9	-	10	-	4	9,8	0,4	0,2	0,0
4.3 B	Verpflichtender Einbau: Maßnahme ist politisch durchsetzbar	10	3	-	10	-	9	-	7	-	4	7,3	2,7	7,2	0,4
4.3 C	Verpflichtender Einbau: Maßnahme ist erfolgversprechender	10	3	-	10	-	9	-	10	-	4	8,0	2,9	8,5	0,4
4.3 D	Verpflichtender Einbau: Standard	10	10	-	1	-	7	-	8	-	4	6,5	3,4	11,3	0,5
4.4 A	Angebot von Nachrüstsystemen bei Altfahrzeugen : Hersteller	10	10	-	4	-	6	-	2	-	4	5,5	3,0	8,8	0,5
5.1 A	Kommunikative Maßnahmen - Argumente: Klimarelevanz und CO2-Einsparung	10	7	-	5	-	4	-	3	-	4	4,8	1,5	2,2	0,3
5.1 B	Kommunikative Maßnahmen - Argumente: Ressourcenschonung durch Energieeinsparung	10	8	-	6	-	7	-	2	-	4	5,8	2,3	5,2	0,4
5.1 C	Kommunikative Maßnahmen - Argumente: Vorsorge für nachfolgende Generationen, globale Gerechtigkeit	10	7	-	4	-	2	-	1	-	4	3,5	2,3	5,3	0,7

5.1 D	Kommunikative Maßnahmen - Argumente: Persönliche (finanzielle) Vorteile durch Sprit- und Kosteneinsparung	10	9	-	9	-	8	-	10	-	4	9,0	0,7	0,5	0,1
5.1 E	Kommunikative Maßnahmen - Argumente: Zusammenhang zwischen Verkehrssicherheit und spritsparendem Fahrstil	10	5	-	8	-	2	8	1	-	5	4,8	2,9	8,6	0,6
5.1 F	Kommunikative Maßnahmen - Argumente: Vorbildfunktion für Kinder/Andere	10	-	-	5	-	6	-	2	-	3	4,3	1,7	2,9	0,4
5.1 G	Kommunikative Maßnahmen - Argumente: „Grünes“ Image	10	10	-	8	-	8	-	6	-	4	8,0	1,4	2,0	0,2
5.1 H	Kommunikative Maßnahmen - Argumente: Neuberechnung der KFZ-Steuer	10	5	10	7	-	3	-	1	-	5	5,2	3,1	9,8	0,6
5.1 I	Kommunikative Maßnahmen - Argumente: Wirksamkeit von Eco-Fahrassistenten	10	3	-	-	-	-	-	1	-	2	2,0	1,0	1,0	0,5
5.3 A	Kommunikative Begleitmaßnahmen: Die von verschiedenen Akteuren die Einführung abgestimmt begleiten, sind als bedeutsam zu erachten.	10	9	-	9	-	8	-	10	-	4	9,0	0,7	0,5	0,1
5.3 B	Kommunikative Begleitmaßnahmen: Ansprechende, motivierende Informationen zum Assistenten	10	9	-	9	-	7	-	7	-	4	8,0	1,0	1,0	0,1
6.1 A	Handlungsempfehlungen: Eco-Fahrassistentensysteme sind im Hinblick auf eine Reduktion des Kraftstoffverbrauchs und damit auch einer CO ₂ -Reduktion als bedeutsam zu erachten.	10	9	-	10	-	9	-	10	-	4	9,5	0,5	0,3	0,1

Anmerkungen durch die Experten (im Fragebogen)					
Frage Nr.	Frageinhalt	Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4
2.1 C	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Effiziente / ineffiziente Geschwindigkeit auf der Autobahn			Zeit + Verbrauch ist gleichzeitig nicht möglich	
2.1 D	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Information zum rechtzeitig Ausrollen lassen anstatt stark zu bremsen			„segeln“	
2.1 E	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Informationen zur Gleichmäßigkeit des Tempos beim Fahren		Kein Bremsweg?	Mehr betroffen von Simulation	
2.1 F	Gestaltung von Eco-Fahrassistenten - Informationen während der Fahrt: Sonstige				Überschaubare Zusammenfassung auf einen Blick, nicht zu viele Details während der Fahrt. → Schaltempfehlung
2.2 E	Gestaltung von Eco-Fahrassistenten - Informationen zur Auswahl der Route und des Verkehrsmittels: Sonstiges		Kostenschätzung mit Nebenkosten (Parken)		Verlässliche Stauinformationen
2.2 F	Gestaltung von Eco-Fahrassistenten - Informationen zu effizienten Verhaltensweisen: Klimaanlage ausschalten, Spriteinsparung veranschaulichen		(Komfort, Sicherheit)		Nicht unmittelbar während des Fahrens
2.2 G	Gestaltung von Eco-Fahrassistenten - Informationen zu effizienten Verhaltensweisen: Ballast und Aufbauten				Während der Fahrt schwierig, generell den Zusammenhang an den Fahrer vermitteln

2.2 I	Gestaltung von Eco-Fahrassistenten - Informationen zu effizienten Verhaltensweisen: Sonstiges		Nebenverbraucher (Sitzheizung, Scheibenheizung)		Reifendruck
2.3 A	Gestaltungsmerkmale eines Eco-Fahrassistenten: Abschaltbarkeit aller Eco-Fahrassistentenfunktionen		Funktionsabhängig		Bei dezentem Design
2.3 E	Gestaltungsmerkmale eines Eco-Fahrassistenten: Car-to-car und car-to-infrastructure Informationen			Resultat	
2.3 F	Gestaltungsmerkmale eines Eco-Fahrassistenten: Informationskanal				Visuell bevorzugt
2.3 J	Gestaltungsmerkmale eines Eco-Fahrassistenten: Ecoscore	Wenn intelligent			
2.3 L	Gestaltungsmerkmale eines Eco-Fahrassistenten: Information zur CO2 Einsparung pro Fahrt			kontrovers	
2.3 M	Gestaltungsmerkmale eines Eco-Fahrassistenten: Vergleich mit anderen Fahrern			Nur freiwillig	Eher im Internet
2.3 Q	Gestaltungsmerkmale eines Eco-Fahrassistenten: Training		Ist abhängig von der Umsetzung		„Tip of the Day“ (Audi): kann nicht so viel bieten wie Fahrerlehrer
3 E	Finanzielle Anreize: Sonstiges		Der Markt reguliert sich selbst. CO2 ist Verbrauch (Mineralölsteuer reicht)	Heutiges System	
4.4 A	Angebot von Nachrüstsystemen bei Altfahrzeugen: Hersteller				Technisch möglich?

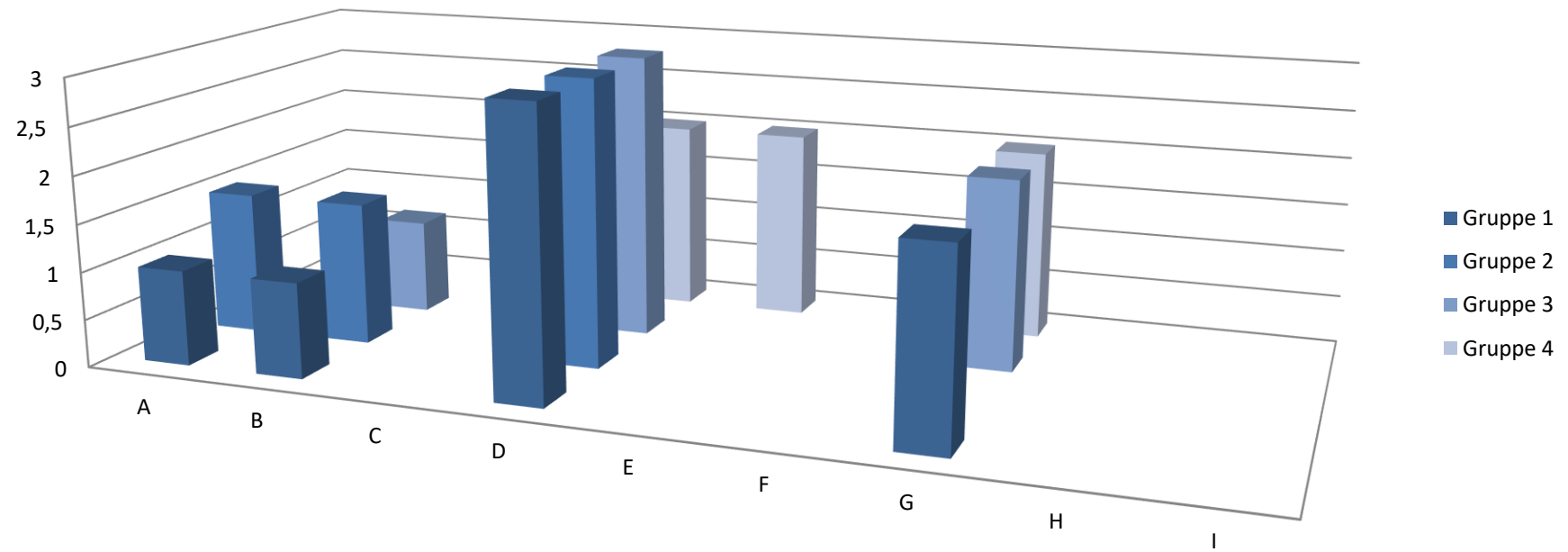


Es wird konsensual Maßnahme 3D zugestimmt: „Erhöhung der Ökosteuer und damit auch des Benzinpreises“. Allerdings kommentieren die Gruppen die Maßnahme so, dass das heutige System ausreicht und keine Erhöhung notwendig ist.

Runde 2 – 10er Skalen

Frage-Nr.	Frageinhalt	Neues	Umformuliertes	Skala	Antwort Gruppe 1	Minderheit Gruppe 1	Antwort Gruppe 2	Minderheit Gruppe 2	Antwort Gruppe 3	Minderheit Gruppe 3	Antwort Gruppe 4	Minderheit Gruppe 4	Anzahl Antworten	Arithmetisches Mittel	Standardabweichung	Varianz	Variationskoeffizient
2.3 C	Gestaltungsmerkmale eines Eco-Fahrassistenten: Grundinformationen zu spritsparenden Fahrstilaspekten permanent aber dezent	nein	ja	10	3	-	9	-	9	-	10	-	4	7,8	2,8	7,7	0,4
2.3 M	Gestaltungsmerkmale eines Eco-Fahrassistenten: Vergleich mit anderen Fahrern	nein	nein	10	6	-	-	-	5	-	9	-	3	6,7	1,7	2,9	0,3
2.3 R	Gestaltungsmerkmale eines Eco-Fahrassistenten: Nur noch Optimierungsmöglichkeiten für fortgeschrittene Fahrer	nein	ja	10	2	6	-	-	1	-	2	-	4	2,8	1,9	3,7	0,7
2.3 T	Gestaltungsmerkmale eines Eco-Fahrassistenten: Informationen als Basis für den Ecoscore, die zugeschaltet werden können	ja	nein	10	9	-	9	-	8	-	9	-	4	8,8	0,4	0,2	0,0
4.3 B	Verpflichtender Einbau: Basissystems ist politisch durchsetzbar	nein	ja	10	7	-	10	-	9	-	10	-	4	9,0	1,2	1,5	0,1
4.3 C	Verpflichtender Einbau: Maßnahme ist erfolgversprechender	nein	ja	10	-	-	10	-	9	-	8	3	4	7,5	2,7	7,3	0,4
4.3 D	Verpflichtender Einbau: Minimalstandard	nein	ja	10	7	-	9	-	9	-	7	-	4	8,0	1,0	1,0	0,1

Frage 5.1 - Kommunikative Maßnahmen: Mögliche Argumente



11.14 Ausgewertete Aussagen: Normative Dimension (Eco-Fahrassistenten)

	Erste Plenarsitzung	Zweite Plenarsitzung	Ergebnisprotokoll
Item 2.3.C	<p>TN10: Basisdaten stehen im Sinne von rohen Daten zur Verfügung und sind zuschaltbar nach der Fahrt, Sichtbar ist nur was Dezentos. Das wären die Fragen: „wollen wir alles → ja/ nein“ Wollen wir was kontinuierlich angezeigtes Dezentos haben? – dann wären wir konsensfähig TN12: Wenn es so weit abzuschalten ist, dass man es nicht mehr wahrnimmt, dann sage ich, nein, das braucht man dann auch nicht, das kann man vergessen. Die Drehzahlanzeige ist auch immer da. Es muss im Blickpunkt des Fahrers sein.</p>	<p>TN3: Akzeptanz käme nur, wenn es so dezent ist, dass man es übersehen kann, dann bringt es aber nichts, man kann es ignorieren, dann bringt es aber nichts mehr. TN12: Wenn ich ganz abschalten kann, brauche ich das System nicht.</p>	<p>Hier stellt sich als nicht auflösbares Dilemma dar, dass Akzeptanz gegenüber dem Eco-Fahrassistenten nur erreicht werden kann, wenn dieser sehr dezent ist. Die Funktion der Förderung einer spritsparenden Fahrweise ist dann jedoch stark eingeschränkt. Als Folge bleibt bei der Entscheidung über die Gestaltung des Eco-Fahrassistenten die Wahl zwischen hoher Akzeptanz bei der Einführung durch dezentes Design oder aber hohe Funktionalität durch weniger dezentes Design.</p>
Item 2.3 H	<p>TN10: Nachrüstnavis nicht vom Hersteller sondern Nachrüstgerät von Drittherstellern, da ist es, auch wenn es nicht hundertprozentig ist, besser, als es nicht zu haben TN2: Daten aus dem Navi sind extrem wichtig, z.B. Umgebungsdaten insbesondere Steigung, potentielle Bremsstellen, Kurven sind extrem wichtig, Navi als Datenquelle zu nutzen wäre extrem wichtig, ...</p>		

<p>Item 2.3 M</p>	<p>TN8: Im pädagogischen Sinne muss man fragen, welche Norm angesprochen wird. Hier ist es eine soziale Norm, vielleicht wäre eine individuelle Norm besser in dem Sinne wie war ich vor einem Monat, das erleichtert vielleicht eher ein Lernen.</p> <p>TN10: Der soziale Vergleich hat durchaus auch in anderen Ländern einen hohen Wert „der Beste zu sein“</p> <p>TN7: Es gibt auch Untersuchungen, dass die die besser als die Durchschnittswerte gemeldet bekommen, sich dann der Norm angleichen, dann also weniger effizient werden. Nur die die drunter liegen haben eine Motivation</p> <p>TN1: Grundsätzlich, egal, wie es aussieht, allein durch die Bewegung auf solchen Plattformen zeige ich, dass ich mich mit dem Thema beschäftige, dann gibt es eine Wirkkette dahinter und das ist eine gute Voraussetzung.</p>		<p>Die Frage des Vergleiches mit anderen Fahrern wurde kontrovers diskutiert. Der selbstbestimmte Austausch der Fahrer in sozialen Netzwerken wird eher als kommunikative Maßnahme, die Bereitstellung der Information im Fahrzeug als datenschutzrechtlich bedenklich angesehen. Die Sinnhaftigkeit des Vergleichs mit einer sozialen Norm wird angezweifelt, es könnte mehr Sinn machen, mit einer individuellen Norm, also vorausgegangenen eigenen Fahrten zu vergleichen. Der Vergleich mit anderen kann darüber hinaus auch negative Auswirkungen in dem Sinne haben, dass nicht die sparsame Fahrweise sondern bspw. Geschwindigkeiten verglichen werden. Zu den Prioritäten der Nutzer sehen die Experten noch weiteren Forschungsbedarf.</p>
<p>Item 3 D</p>	<p>TN12: (Gr.4) für uns reicht generell die Mineralölsteuer als finanzieller Anreiz aus. Es gibt politische Gründe, warum es KFZ- und Mineralölsteuer gibt, aber ob die Trennung sinnvoll ist, steht in Frage</p> <p>TN7: Umwelt ist Lärm, ist Fläche... das wird nicht über die Mineralölsteuer reingenommen</p> <p>TN12: Bestimmte Teilaspekte blenden wir sowieso aus</p> <p>TN2: Ich wehre mich dagegen, der Politik neue Geldquellen zu eröffnen</p>		

	<p>TN2: Das System ist ambivalent. Der Staat verliert Einnahmen, wenn der Verbrauch zurückgeht.</p> <p>TN10: Der kommunikative Aspekt: man sollte das Verständnis vom Zusammenhang Mobilität Umwelt bereits in den Schulen wecken.</p>		
Item 4.4 A	<p>TN2: man sollte sich lieber auf neue Technologien konzentrieren als die alten zu optimieren</p> <p>TN14: so etwas ist bereits auf dem Markt, wird aber nicht nachgefragt</p> <p>TN1: unsere Gruppe hatte es als gut bewertet, ich sehe es jetzt aber genauso wie Herr TN2</p>		
Item 5.1 G	<p>TN7: Der Punkt Image ist eher für Firmen interessant für die Bürger beim Kauf weniger.</p> <p>TN3: Für Individuen kann das Grüne Image auch eine Kaufentscheidung sein, im Sinne einer Gewissensberuhigung.</p> <p>TN7: Wie transportiere ich ein Image von EFA? Hat der dann hinten ein EA drauf? Ich glaube nicht, dass das hier so relevant ist.</p> <p>TN3: In den 80ern hatte Mercedes ABS als Zusatzinformation auf dem Schriftzug hatten hinten drauf, Volvo hatte Lambda-Sonde auf dem Kühlergrill stehen, für eine gewisse Zeit hat das Prestige gebracht. Es wird ja immerhin im Auto dargestellt</p> <p>TN12: Wenn man bereit ist, da mehr zu zahlen, dann „tue Gutes und rede darüber“, so dass die Autohersteller das auch vermarkten könnten, aggressiver vermarkten, solche Techniken, dann auch am Fahrzeug signalisieren ob das in Richtung Status geht, das scheint interessanter als Vorsorge für nachfolgende Generationen oder Vorbildfunktion</p>		
Frage 5.2	<p>TN1: Das deckt sich mit unserer Untersuchung mit ökonomischer Fahrweise hatten wir im Schnitt ca. 1km/h höhere Geschwindigkeit</p>		

		<p>TN7: viele Fahrer denken das aber anders, die geben vor der Ampel nochmal Gas und effiziente, die tats. In der Realität vielleicht nicht schneller aber gleich sind, trotzdem gibt es sehr viele Autofahrer, die meinen, ich bin schneller, wenn ich viel beschleunige</p> <p>TN13: Das Bild das die Fahrer haben ist falsch, das kann man kommunizieren, man muss das Bild, dass die Fahrer haben, wegnimmt!</p>	
Offene Fragen (5.2) - ÖPNV		<p>TN2: Busse haben Vorzugsschaltung, priv. PKWs kriegen rot geschaltet. Was ist Ziel der Politik, Nahverkehr fördern oder Emissionen reduzieren. Normaler Verkehr wird ausgebremst, kostet CO₂ aber man möchte die Leute in die öffentlichen Verkehrsmittel zwingen. Hier sind unterschiedliche Zielsetzungen, weiß nicht wer das regeln kann, diese Widersprüche aufzulösen.</p> <p>TN10: Wenn man viele in einen Bus setzt, dann kann man viele Autos ersetzen, die Infragestellung der Bevorzugung von ÖPNV würde ich nicht gegeben sehen</p> <p>TN1: Zum Verständnis: ist dieses Konzept dazu da die Autofahrer zu vergraulen, damit er sich in den Bus setzt? Aus Emissionsgesichtspunkten gibt es einen Bus der stehen bleibt vs. 30 Fahrzeuge</p> <p>TN13: Das ist ein wenig push and pull, hier sind wir auf einer anderen Ebene da geht es um Verkehrskonzepte insgesamt, wenn man die Leute vom privaten Verkehr in den ÖPNV bringen will muss man diese Priorisierung machen, da geht es um langfristige Systemänderung.</p> <p>TN3: Vor jedem dieser Punkte steht „prüfen“, das heißt das soll geprüft werden, nicht als Empfehlung „Priorisierung des ÖPNV aufheben“ man muss sehen, ob Optimierungspotential bei den Prozessen zwischen Bund Land und Gemeinden zu identifizieren ist.</p>	

Offene Frage (6.2 – NKI)			<p>Durchsetzung der Verwertung von Straßendaten mittels automatischer Übermittlung der Daten (über Topografie und Straßenzustände, wie Kurvenradien, Kreisverkehre, Steigungen, Höheninformationen) an die Kartenlieferanten, dies ist derzeit Länder- und Kreissache und sollte gesetzlich geregelt werden.</p> <p>Grundlegende politische Positionierung (ähnliche Positionierung wie bei Energiespeichertechnologien)</p>
-----------------------------	--	--	--

Legende: (1) Werturteile; (2) Präferenzen; (3) (individuelle) Bewertungen von Optionen ; (4) moralische Urteile; Aussagekontext



-Expertendelphi-

Soziale Indikatoren der Nachhaltigkeit am Fallbeispiel von Carbon Dioxide Capture and Storage-Technologien

Dieses Gruppendelphi findet im Rahmen des EU-Projekts *PROSUITE* statt.

Ziel des Projektes ist es, neue Technologien im Hinblick auf ihre Nachhaltigkeit einzuschätzen. In unserem Teilprojekt wird dabei ein Modell sozialer Nachhaltigkeit angestrebt, das *Indikatoren der gesellschaftlichen Ebene von Nachhaltigkeit* mit *Indikatoren aus dem Bereich der produktbezogenen Lebenszyklusanalyse* verbindet. Dazu wurden im Vorfeld mehr als 80 internationale und nationale Indikatorensets betrachtet. Weiterhin wurde eine Vorauswahl der Indikatoren entlang verschiedener Kriterien (etwa ihrer empirischen Anwendbarkeit, ihrer Eindeutigkeit oder ihrer inhaltlichen Relevanz) getroffen. Die verbleibenden Indikatoren sind der Gegenstand des heutigen Delphis. Ziel ist es, sie in Bezug auf eine gesellschaftliche Etablierung von *CCS-Technologien* einzuschätzen und eine damit verbundene Analyse sozialer Nachhaltigkeit zu ermöglichen.

Abschnitt I: Generelle Indikatoren

Im ersten Abschnitt finden sich Indikatoren die generell zur Bemessung sozialer Nachhaltigkeit konstruiert wurden. Bitte bewerten sie ob der Indikator sinnvoll zur Bewertung von CCS-Technologien ist, da diese Technologien einen direkten oder indirekten Einfluss auf den Indikator ausüben könnten.

Geben Sie bitte 1 (-) an, wenn Sie der Ansicht sind, dass dieser Indikator zur Bewertung von CCS-Technologien überhaupt nicht relevant ist, geben Sie bitte 10 (+) an, wenn Sie der Ansicht sind, dass dieser Indikator sehr relevant ist.

	Indikator	-	2	3	4	5	6	7	8	9	+
1	Anzahl zivilgesellschaftlicher Organisationen pro tausend Einwohner										
2	Anteil der Personen, die in der Nähe von Produktionsstätten wohnen und aufgrund dieser ihren Wohnort verlassen wollen										
3	Anteil der Personen die zufrieden oder sehr zufrieden mit ihrer Arbeitsstelle sind										
4	Generelle Lebenszufriedenheit der Personen (Likert Skala)										
5	Anzahl innovativer Unternehmen										

	Indikator	-	2	3	4	5	6	7	8	9	+
6	Anzahl der <i>Beschäftigten in wissensintensiven Bereichen</i>										
7	Anteil der Personen die bereits <i>Informationsmaterial über die Risiken</i> der Technologie erhalten haben										
8	<i>Kindersterblichkeitsrate</i> pro 1000 Geburten										
9	Die <i>Lebenserwartung</i> zum Zeitpunkt der Geburt (nach Geschlecht)										
10	Prozentsatz der Personen deren <i>Haushaltseinkommen</i> unter 60 Prozent des nationalen Durchschnittseinkommens liegt										
11	<i>Gini-Index</i> als Indikator der Einkommensungleichverteilung										
12	<i>Arbeitslosenquote</i> (nach Altersgruppe und Geschlecht)										
13	Anteil der Personen die bereits an wichtigen <i>Planungsentscheidungen</i> beteiligt waren										
14	Die Technik beeinflusst die <i>subjektiv empfundene Lebenszufriedenheit</i> der Konsumenten										
15	Durch den Einsatz der Technologie werden <i>Grün- und Brachflächen</i> verringert										
16	Die Bevölkerung nimmt den <i>Einsatz der Technik als Risiko</i> wahr										

Abschnitt II: Organisations-spezifische Indikatoren

Im zweiten Abschnitt finden sich Indikatoren die die organisatorischen Aspekte der Nachhaltigkeit erfassen sollen. Zu bewerten ist, ob die Indikatoren eine mögliche Auswirkung von CCS-Technologien auf die Organisationen messen und ob diese geeignet sind.

Geben Sie bitte 1 (-) an, wenn Sie der Ansicht sind, dass dieser Indikator zur Bewertung von CCS-Technologien überhaupt nicht relevant ist, geben Sie bitte 10 (+) an, wenn Sie der Ansicht sind, dass dieser Indikator sehr relevant ist.

	Indikator	-	2	3	4	5	6	7	8	9	+
17	Vorhandensein von formalen <i>Richtlinien zur Chancengleichheit</i>										
18	Gesamtzahl an <i>Vorfällen von Diskriminierung</i> und dagegen ergriffene <i>Maßnahmen</i>										
19	<i>Verhältnis des Grundgehalts</i> von Männern und Frauen nach Beschäftigungsverhältnis										
20	Vorhandensein einer <i>formalen Politik in Bezug auf Gesundheit und Sicherheit</i>										
21	Anzahl der (schweren/nicht-schwerwiegende) Occupational Safety and Health Administration <i>OSHA-Verstöße</i> (Occupational Safety and Health Administration) innerhalb der letzten 3 Jahre										
22	<i>Anzahl von Verbraucherbeschwerden</i>										
23	<i>Zertifizierung des Produktes</i> durch die Organisation										
24	Existieren klare <i>Managementsysteme</i> um sicherzustellen, dass eine klare <i>Nutzerinformation</i> geleistet wird										
25	Verfügt die Organisation über ein <i>zertifiziertes Umweltmanagement-System</i>										
26	Die Organisation hat sich verpflichtet, den <i>Prinzipien des Global Compact</i> zu entsprechen und kommuniziert jährlich ihren Fortschritt										
27	<i>Investitionen in Technologieentwicklung und Technologietransfer</i>										

	Indikator	-	2	3	4	5	6	7	8	9	+
28	Präsenz von expliziten <i>Verhaltensregeln</i> , die die <i>Menschenrechte</i> von Arbeitnehmern schützen										
29	Anteil der <i>Lieferanten</i> die vom Unternehmen auf ihre <i>soziale Verantwortung</i> in den letzten Jahren überprüft wurden										

Abschnitt III: Produkt-spezifische Indikatoren

In diesem dritten Abschnitt finden sich Indikatoren die die produktspezifischen Aspekte der Technik auf soziale Nachhaltigkeit erfassen sollen.

Geben Sie bitte 1 (-) an, wenn Sie der Ansicht sind, dass dieser Indikator zur Bewertung von CCS-Technologien überhaupt nicht relevant ist, geben Sie bitte 10 (+) an, wenn Sie der Ansicht sind, dass dieser Indikator sehr relevant ist.

	Indikator	-	2	3	4	5	6	7	8	9	+
30	<i>Anzahl dauerhafter Arbeitsplätze</i> die durch die Technologie geschaffen werden										
31	<i>Anzahl befristeter Arbeitsplätze</i> die durch die Technologie geschaffen werden										
32	<i>Anzahl tödlicher Arbeitsunfälle</i>										
33	<i>Anzahl nicht-tödlicher Arbeitsunfälle</i>										
34	<i>Anzahl der Berufskrankheiten</i>										

Abschnitt IV: Indikatoren

Im vierten Abschnitt können sie Indikatoren ergänzen, die ihrer Meinung nach zur Bemessung der sozialen Nachhaltigkeit von CCS-Technologien herangezogen werden sollten, sich aber nicht in der vorhergehenden Auflistung finden. Bitte notieren sie diese in den dafür vorgesehenen Feldern

		Indikator
36	Genereller Indikator	<hr/> <hr/> <hr/> <hr/>
37	Organisations-spezifischer Indikator	<hr/> <hr/> <hr/> <hr/>
38	Produkt-spezifischer Indikator	<hr/> <hr/> <hr/> <hr/>

Vielen Dank für die Beantwortung des Fragebogens!



-Expertendelphi-

Soziale Indikatoren der Nachhaltigkeit am Fallbeispiel Nano-Technologien

Statistische Ergebnisse

Runde 1

Frage-Nr.	Frageinhalt	Thematischer Bereich	Skala	Antwort Gruppe 1	Minderheit Gruppe 1	Antwort Gruppe 2	Minderheit Gruppe 2	Antwort Gruppe 3	Minderheit Gruppe 3	Anzahl Antworten	Arithmetisches Mittel	Standard-	Varianz	Variations-Koeffizient
1	Anzahl zivilgesellschaftlicher Organisationen	I: Generelle Indikatoren		7	-	8	-	8	-	3	7,7	0,5	0,2	0,1
2	Wohnort verlassen	I: Generelle Indikatoren		10	-	10	-	9	-	3	9,7	0,5	0,2	0,0
3	Zufrieden mit Arbeitsstelle	I: Generelle Indikatoren		1	-	2	-	2	-	3	1,7	0,5	0,2	0,3
4	Generelle Lebenszufriedenheit	I: Generelle Indikatoren	10	5	-	3	-	2	-	3	3,3	1,2	1,6	0,4
5	Anzahl innovativer Unternehmen	I: Generelle Indikatoren	10	4	-	6	-	4	-	3	4,7	0,9	0,9	0,2
6	Beschäftigten in wissensintensiven Bereichen	I: Generelle Indikatoren	10	4	-	3	-	1	-	3	2,7	1,2	1,6	0,5
7	Informationsmaterial über die Risiken	I: Generelle Indikatoren	10	6	2	10	-	5	-	4	5,8	2,9	8,2	0,5
8	Kindersterblichkeitsrate	I: Generelle Indikatoren	10	8	-	1	-	2	-	3	3,7	3,1	9,6	0,8
9	Lebenserwartung	I: Generelle Indikatoren	10	8	-	1	-	2	-	3	3,7	3,1	9,6	0,8
10	Haushaltseinkommen	I: Generelle Indikatoren	10	1	-	2	-	7	-	3	3,3	2,6	6,9	0,8
11	Gini-Index	I: Generelle Indikatoren	10	2	-	3	-	7	-	3	4,0	2,2	4,7	0,5
12	Arbeitslosenquote	I: Generelle Indikatoren	10	8	-	3	-	7	-	3	6,0	2,2	4,7	0,4
13	Planungsentscheidungen	I: Generelle Indikatoren	10	6	-	10	-	8	-	3	8,0	1,6	2,7	0,2
14	Subjektiv empfundene Lebenszufriedenheit	I: Generelle Indikatoren	10	5	-	8	-	2	-	3	5,0	2,4	6,0	0,5
15	Verringerung von Grün- und Brachflächen	I: Generelle Indikatoren	10	8	-	6	-	5	-	3	6,3	1,2	1,6	0,2

16	Wahrnehmung des Einsatzes der Technik als Risiko	I: Generelle Indikatoren	10	10	-	10	-	9	-	3	9,7	0,5	0,2	0,0
17	Richtlinien zur Chancengleichheit	II: Organisations-spezifische Indikatoren	10	3	-	10	-	2	-	3	5,0	3,6	12,7	0,7
18	Vorfällen und Maßnahmen bei Diskriminierung	II: Organisations-spezifische Indikatoren	10	3	-	1	-	2	-	3	2,0	0,8	0,7	0,4
19	Verhältnis des Grundgehalts (nach Geschlecht)	II: Organisations-spezifische Indikatoren	10	3	-	10	-	2	-	3	5,0	3,6	12,7	0,7
20	Formale Politik in Bezug auf Gesundheit und Sicherheit	II: Organisations-spezifische Indikatoren	10	7	-	10	-	9	-	3	8,7	1,2	1,6	0,1
21	OSHA-Verstöße	II: Organisations-spezifische Indikatoren	10	8	-	10	-	8	-	3	8,7	0,9	0,9	0,1
22	Anzahl von Verbraucherbeschwerden	II: Organisations-spezifische Indikatoren	10	8	-	5	-	2	-	3	5,0	2,4	6,0	0,5
23	Zertifizierung des Produktes	II: Organisations-spezifische Indikatoren	10	8	-	10	-	1	-	3	6,3	3,9	14,9	0,6
24	Managementsysteme zur Nutzerinformation	II: Organisations-spezifische Indikatoren	10	8	-	10	-	6	-	3	8,0	1,6	2,7	0,2
25	Zertifiziertes Umweltmanagement-System	II: Organisations-spezifische Indikatoren	10	8	-	10	-	9	-	3	9,0	0,8	0,7	0,1
26	Prinzipien des Global Compact	II: Organisations-spezifische Indikatoren	10	8	-	8	-	9	-	3	8,3	0,5	0,2	0,1
27	Investitionen in Technologieentwicklung und Technologietransfer	II: Organisations-spezifische Indikatoren	10	1	-	10	-	2	8	4	5,3	3,8	14,7	0,7
28	Verhaltensregeln zum Schutz der Menschenrechte	II: Organisations-spezifische Indikatoren	10	6	-	10	-	2	-	3	6,0	3,3	10,7	0,5
29	Überprüfung der Lieferanten auf soziale Verantwortung	II: Organisations-spezifische Indikatoren	10	6	-	1	-	4	-	3	3,7	2,1	4,2	0,6
30	Anzahl dauerhafter Arbeitsplätze	III: Produkt-spezifische Indikatoren	10	8	-	7	-	7	-	3	7,3	0,5	0,2	0,1
31	Anzahl befristeter Arbeitsplätze	III: Produkt-spezifische Indikatoren	10	8	-	3	-	8	-	3	6,3	2,4	5,6	0,4
32	Anzahl tödlicher Arbeitsunfälle	III: Produkt-spezifische Indikatoren	10	10	-	2	-	8	-	3	6,7	3,4	11,6	0,5
33	Anzahl nicht-tödlicher Arbeitsunfälle	III: Produkt-spezifische Indikatoren	10	10	-	2	-	8	-	3	6,7	3,4	11,6	0,5
34	Anzahl der Berufskrankheiten	III: Produkt-spezifische Indikatoren	10	8	-	1	-	8	-	3	5,7	3,3	10,9	0,6

Runde 2

Nr	Frage-Nr.	Frageinhalt	Änderung ¹⁶⁴	Thematischer Bereich	Skala	Antwort Gruppe 1	Minderheit Gruppe 1	Antwort Gruppe 2	Minderheit Gruppe 2	Antwort Gruppe 3	Minderheit Gruppe 3	Anzahl Antworten	Arithmetisches Mittel	Standard-	Varianz	Variations-koeffizient
1	6a	Beschäftigten in wissensintensiven Bereichen (Wichtigkeit)	A	I: Generelle Indikatoren	10	6		4		3		3	4,3	1,2	1,6	0,3
2	6b	Beschäftigten in wissensintensiven Bereichen (Impact)	A	I: Generelle Indikatoren	10	4		2		3		3	3,0	0,8	0,7	0,3
3	81	Bildung	N	I: Generelle Indikatoren	10	8		-		10		2	9,0	1,0	1,0	0,1
4	7a	Anteil der Personen die sich subjektiv ausreichend informiert fühlen (Wichtigkeit)	U, A	I: Generelle Indikatoren	10	10		8		6		3	8,0	1,6	2,7	0,2
5	7b	Anteil der Personen die sich subjektiv ausreichend informiert fühlen (Impact)	U, A	I: Generelle Indikatoren	10	4		8		6		3	6,0	1,6	2,7	0,3
6	8a	Kindersterblichkeitsrate (Wichtigkeit)	A	I: Generelle Indikatoren	10	8		9		9		3	8,7	0,5	0,2	0,1
7	8b	Kindersterblichkeitsrate (Impact)	A	I: Generelle Indikatoren	10	2		2		2		4	2,0	0,0	0,0	0,0
8	9a	Lebenserwartung (Wichtigkeit)	A	I: Generelle Indikatoren	10	8		9		9		3	8,7	0,5	0,2	0,1
9	9b	Lebenserwartung (Impact)	A	I: Generelle Indikatoren	10	2		2		2		3	2,0	0,0	0,0	0,0
10	10a	Haushaltseinkommen (Wichtigkeit)	A	I: Generelle Indikatoren	10	7		8		9		3	8,0	0,8	0,7	0,1
11	10b	Haushaltseinkommen (Impact)	A	I: Generelle Indikatoren	10	2		5	8	2		4	4,3	2,5	6,2	0,6
12	11a	Gini-Index (Wichtigkeit)	A	I: Generelle Indikatoren	10	7		8		9		3	8,0	0,8	0,7	0,1
13	11b	Gini-Index (Impact)	A	I: Generelle Indikatoren	10	2		4		2		3	2,7	0,9	0,9	0,4

¹⁶⁴ A= Der Indikator wurde in seine generelle Wichtigkeit und seinen Impact auf CCS aufgeteilt; U= Der Indikator wurde umformuliert, N = Der Indikator ist ein von den Experten vorgeschlagener Indikator

14	14	Subjektiv empfundene Lebenszufriedenheit (Konsumenten: Strom)	U	I: Generelle Indikatoren	10	5		6		4		3	5,0	0,8	0,7	0,2
15	83	Subjektiv empfundene Lebenszufriedenheit (Anwohner)	N	I: Generelle Indikatoren	10	9		9		10		3	9,3	0,5	0,2	0,1
16	82	Regionale Disparität	N	I: Generelle Indikatoren	10	9		9		8		3	8,7	0,5	0,2	0,1
17	17a	Richtlinien zur Chancengleichheit (Wichtigkeit)	A	II: Organisations-spezifische Indikatoren	10	7	10	10		2		4	7,3	3,3	10,7	0,5
18	17b	Richtlinien zur Chancengleichheit (Impact)	A	II: Organisations-spezifische Indikatoren	10	2	5	3		2		4	3,0	1,2	1,5	0,4
19	19a	Verhältnis des Grundgehalts (nach Geschlecht) (Wichtigkeit)	A	II: Organisations-spezifische Indikatoren	10	7	10	10		2		4	7,3	3,3	10,7	0,5
20	19b	Verhältnis des Grundgehalts (nach Geschlecht) (Impact)	A	II: Organisations-spezifische Indikatoren	10	2	5	2		2		4	2,8	1,3	1,7	0,5
21	22a	Anzahl von Verbraucherbeschwerden durch Stromkunden (Wichtigkeit)	U, A	II: Organisations-spezifische Indikatoren	10	7		4		2		3	4,3	2,1	4,2	0,5
22	22b	Anzahl von Verbraucherbeschwerden durch Stromkunden (Impact)	U, A	II: Organisations-spezifische Indikatoren	10	4	7	2		2		4	3,8	2,0	4,2	0,5
23	23	Zertifizierung des Produktes		II: Organisations-spezifische Indikatoren	10	4		5		4	8	4	5,3	1,6	2,7	0,3
24	84	Zertifizierung der Technologie	N	II: Organisations-spezifische Indikatoren	10	8		2		3	8	4	5,3	2,8	7,7	0,5
25	27	Förderung regionaler Wertschätzung	U	II: Organisations-spezifische Indikatoren	10	8		10		5		3	7,7	2,1	4,2	0,3
26	85	Verdrängung anderer energierelevanter Technologien	N	II: Organisations-spezifische Indikatoren	10	9		10		7		3	8,7	1,2	1,6	0,1

27	28a	Verhaltensregeln zum Schutz der Menschenrechte (Wichtigkeit)	A	II: Organisations-spezifische Indikatoren	10	1		1		2		3	1,3	0,5	0,2	0,4
28	28b	Verhaltensregeln zum Schutz der Menschenrechte (Impact)	A	II: Organisations-spezifische Indikatoren	10	1		2		2		3	1,7	0,5	0,2	0,3
29	29a	Überprüfung der Lieferanten auf soziale Verantwortung (Wichtigkeit)	A	II: Organisations-spezifische Indikatoren	10	8		9		8		3	8,3	0,5	0,2	0,1
30	29b	Überprüfung der Lieferanten auf soziale Verantwortung (Impact)	A	II: Organisations-spezifische Indikatoren	10	8		9		2		3	6,3	3,1	9,6	0,5
31	32a	Anzahl tödlicher Arbeitsunfälle (Wichtigkeit)	A	III: Produkt-spezifische Indikatoren	10	10		10		8		3	9,3	0,9	0,9	0,1
32	32b	Anzahl tödlicher Arbeitsunfälle (Impact)	A	III: Produkt-spezifische Indikatoren	10	3		3		2		3	2,7	0,5	0,2	0,2
33	33a	Anzahl nicht-tödlicher Arbeitsunfälle (Wichtigkeit)	A	III: Produkt-spezifische Indikatoren	10	10		10		8		3	9,3	0,9	0,9	0,1
34	33b	Anzahl nicht-tödlicher Arbeitsunfälle (Impact)	A	III: Produkt-spezifische Indikatoren	10	3		5		2		3	3,3	1,2	1,6	0,4
35	34a	Anzahl der Berufskrankheiten (Wichtigkeit)	A	III: Produkt-spezifische Indikatoren	10	8		8		8		3	8,0	0,0	0,0	0,0
36	34b	Anzahl der Berufskrankheiten (Impact)	A	III: Produkt-spezifische Ind.	10	3		5		2		3	3,3	1,2	1,6	0,4
37	91	Macht die Technologie Änderungen von gesetzlichen Regelungen notwendig (z.B. Planungsrecht, Raumordnungsrecht, Haushaltsrecht, etc)	N	IV: Vorgeschlagene Indikatoren	10	10		10		6		3	8,7	1,9	3,6	0,2
38	92	Vertrauen in langfristige Regelungsfunktionen	N	IV: Vorgeschlagene Indikatoren	10	10		6		8		3	8,0	1,6	2,7	0,2
39	93	Generationensensible Problemverschiebung (Folgekosten für zukünftige Generationen)	N	IV: Vorgeschlagene Indikatoren	10	10		10		8		3	9,3	0,9	0,9	0,1
40	94	Qualität und Intensität der Bürgerbeteiligungsverfahren	N	IV: Vorgeschlagene Indikatoren	10	9		9		8		3	8,7	0,5	0,2	0,1

11.17 Ausgewertete Aussagen: Normative Dimension (PROSUITE)

Zweite Plenarsitzung	
Abschlussdiskussion im zweiten Plenum	Kritisiert wird die zu technokratische Abfrage. Die demokratische Lernfähigkeit sei nicht vertreten. Zudem wird nicht nach den sozialen Voraussetzungen gefragt, sondern nur die Folgen bewertet.
	Der Punkt „politische Mobilisierung“, bzw. ein Indikator zur Konflikthaftigkeit fehle. Schon jetzt gebe es große Diskussion über die Technik, was zu einem Konflikt führen könne.
	Es fehlt ein Indikator zur Langfristigkeit. Also wie viele Generationen eine Technologie beherrschen müssen. Dies ist zwar in dem neuen Indikator „Generationssensible Problemverschiebung“ enthalten. Es fehlt aber noch etwas wie Reversibilität oder Fehlerhaftigkeit.

Legende: (1) Werturteile; (2) Präferenzen; (3) (individuelle) Bewertungen von Optionen ; (4) moralische Urteile; Aussagekontext

11.18 Erklärung zur Eigenständigkeit dieser Arbeit

Hiermit erkläre ich, dass diese Arbeit - abgesehen von der Beratung durch meinen Doktorvater Prof. Dr. Dr. h. c. Ortwin Renn - nach Inhalt und Form mein eigenes Werk darstellt. Bei der Anfertigung habe ich keine anderen als die angegebenen Quellen und Hilfsmittel verwendet. Die Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis entstanden. Ich habe die Arbeit weder ganz noch zu Teilen bereits an einer anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegt, veröffentlicht oder zu Veröffentlichung eingereicht. Ich habe zuvor noch keinen Promotionsversuch unternommen.

Stuttgart, den 06.02.2020

Rainer Kuhn