

.....

**Inhaltsbezogene Kompetenzen von Schülerinnen und  
Schülern im Fach Naturwissenschaft und Technik in der  
gymnasialen Oberstufe**

.....

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

Vorgelegt von

**Marcus Brändle**

aus Saulgau, jetzt Bad Saulgau

Hauptberichter: Prof. Dr. Bernd Zinn (Universität Stuttgart)

Mitberichter: Prof. Dr. Ronny Nawrodt (Universität Stuttgart)

Tag der mündlichen Prüfung: 20. Juli 2023

Institut für Erziehungswissenschaft

Abteilung Berufspädagogik mit Schwerpunkt Technikdidaktik (BPT)  
der Universität Stuttgart

2023



## Danksagung

Ich möchte mich recht herzlich bei all jenen bedanken, die in den Jahren 2018 bis 2023 zum Gelingen dieser Dissertationsschrift beitrugen!

An erster Stelle ist hier mein Doktorvater, Herr Prof. Dr. Bernd Zinn, Leiter der Abteilung Berufspädagogik mit Schwerpunkt Technikdidaktik (BPT) des Instituts für Erziehungswissenschaft der Universität Stuttgart, zu nennen, der diese Arbeit ermöglichte und mir mit konstruktivem, fachlichem und kritischem Rat stets vertrauensvoll zur Seite stand. Mein Dank gilt auch Herrn Prof. Dr. Ronny Nawrodt für die Zweitbegutachtung der Arbeit, der Vector Stiftung für die Förderung des Forschungsprojekts *NwT-Kursstufe*, allen Schüler:innen für die Teilnahme an den Befragungen sowie den Lehrpersonen der beteiligten Schulen, die mir Unterrichtszeit für die Erhebungen zur Verfügung stellten und sämtliche weiteren Anfragen zu Interviews oder zu unterrichtspraktischen Aspekten zeitnah und kompetent beantworteten.

Bedanken möchte ich mich bei allen Kolleg:innen für ihre tatkräftige Unterstützung. Mit Miras Hilfe konnte ich erste Erfahrungen im Bereich der Bildungsforschung sammeln. Tobias war nicht nur ein Sparringpartner im fachlichen Diskurs, sondern auch mein Freund und Wegbegleiter. Mit ihrer Expertise und im Rahmen zahlreicher Fachgespräche standen mir Carolin, Christina, Katharina, Matthias W. und Matthias H. im Team zur Seite. Meine wissenschaftlichen Hilfskräfte Laura und Lena haben mich tatkräftig bei der Datensammlung und -auswertung unterstützt.

Besonders bedanken möchte ich mich bei meiner Familie und meinen Eltern dafür, dass sie hinter meinen Entscheidungen standen und mir so den beruflichen Weg ebneten. Meiner Ehefrau möchte ich an dieser Stelle sagen: Barbara, du hattest immer ein offenes Ohr und warst deshalb die wichtigste Unterstützerin bei diesem für mich so bedeutenden Projekt.



## Kurzfassung

Die technische Allgemeinbildung gewinnt angesichts der sich beschleunigenden Entwicklung technologischer Trends immer mehr an Bedeutung. Das Schulfach Naturwissenschaft und Technik (NwT) wurde bereits im Schuljahr 2007/08 in Baden-Württemberg eingeführt, um einen Zugang zur technischen Bildung am allgemeinbildenden Gymnasium zu schaffen und Schüler:innen zur technischen Mündigkeit zu befähigen. Die Bildungsplanrevision im Jahr 2016 legte dazu einen stärkeren Fokus auf inhalts- und prozessbezogene Kompetenzen im Kontext der Technikwissenschaften in der Sekundarstufe I. Mit den Schulversuchen des Basisfachs und vor allem des Leistungsfachs in der Kursstufe eröffnen sich Möglichkeiten für eine Vertiefung wissenschaftspropädeutischer technischer Kompetenzen von Lernenden. Gleichzeitig stellt sich im Bezugsfeld eines Bildungsmonitorings die Frage nach der Wirkung des interdisziplinären naturwissenschaftlich-technischen Fachs auf die Entwicklung der inhaltsbezogenen Kompetenzen von Schüler:innen. Der Fokus dieser Arbeit liegt daher auf der Generierung eines umfänglichen Beschreibungswissens. Dazu werden zunächst Testinstrumente für eine valide und reliable Messung inhaltsbezogener Kompetenzen des Leistungsfachs entwickelt. Sie ermöglichen die Erfassung der Voraussetzungen der Lernenden in Form des Vorwissens aus der Sekundarstufe I und des erreichten Fachwissens im Leistungsfach. In Bezug auf die Ausprägung der inhaltsbezogenen Kompetenzen der Schüler:innen wird darüber hinaus die Frage nach unterschiedlichen Kompetenzniveaus geklärt und abschließend überprüft, welche kognitiven Merkmale die Fachwissensentwicklung unterstützen und inwiefern sie die Kompetenzniveaus charakterisieren.

Die Ergebnisse zeigen, dass sich über den Verlauf der Sekundarstufe II bei den Lernenden schulübergreifend eine Entwicklung der inhaltsbezogenen Kompetenzen des Bildungsplans vollzogen hat. Die Kompetenzniveaumodellierung ermöglicht die Kontrastierung der Schüler:innen anhand ihrer unterschiedlichen Kompetenzausprägung in drei Niveaustufen. Die Analyse der Zusammenhänge der kognitiven Merkmale *Vorwissen* und *fluide Intelligenz* belegen, dass ein anknüpfungsfähiges Vorwissen aus der Mittelstufe die Kompetenzentwicklung im Leistungsfach begünstigt und signifikante Unterschiede zwischen dem niedrigsten und dem höchsten Niveau bestehen.



## Abstract

General technical education is becoming increasingly important considering the accelerating development of technological trends. The school subject Naturwissenschaft und Technik (NwT; *Science and Technology*) was introduced in Baden-Württemberg in the school year 2007/08 to create access to technical education and to foster pupils' technology and engineering literacy. The revision of the education plan in 2016 focusses on content-related and process-related competencies in the context of technology and engineering sciences at the lower secondary level. The implementation of the basic subject and above all the advanced subject in grades 11 and 12, unlocked further possibilities for the development of scientific-propaedeutic technical competencies of K-12 learners. At the same time the question arises as to what the effects are of the interdisciplinary science and technology subject on the development of content-related competencies of students. The thesis' focus is therefore on the generation of a comprehensive descriptive knowledge. To this end, test instruments are first developed for a valid and reliable measurement of content-related competencies in the performance subject. They can provide information about the learners' preconditions in form of prior knowledge from the lower secondary school and about subject knowledge achieved in grades 11 and 12. Furthermore, the question of different competency levels is clarified regarding the development of the students' content-related competencies. Conclusively, it is examined which cognitive characteristics support the development of subject knowledge and to what extent they characterize the competency levels.

Results show a development of content-related competencies over the course of the upper secondary level among the students across schools. Modelling of competency levels enables the division into three contrasting groups of students according to their different competencies. The correlation and mediation analysis of cognitive characteristics *prior knowledge* and *fluid intelligence* show that a connectable prior knowledge from the middle school favors the development of competencies in the performance subject and that there are significant differences between lowest and highest level.





## Inhaltsverzeichnis

<b>Danksagung</b> .....	<b>I</b>
<b>Kurzfassung</b> .....	<b>III</b>
<b>Abstract</b> .....	<b>V</b>
<b>Inhaltsverzeichnis</b> .....	<b>VII</b>
<b>Abbildungsverzeichnis</b> .....	<b>XIII</b>
<b>Tabellenverzeichnis</b> .....	<b>XVII</b>
<b>Abkürzungsverzeichnis</b> .....	<b>XIX</b>
<b>1 Einleitung</b> .....	<b>1</b>
1.1 Genese und Bildungswert des Schulfachs NwT .....	2
1.2 Relevanz und Ziele der Untersuchung .....	4
1.3 Aufbau der Arbeit .....	4
<b>2 Theoretischer Rahmen</b> .....	<b>9</b>
2.1 Schulisches Lernen.....	9
2.1.1 Unterrichtliche Rahmenbedingungen – die Angebotsebene.....	10
2.1.2 Affektive Lernendenmerkmale – die Nutzungsebene.....	11
2.1.3 Lernergebnisse.....	11
2.2 Einflussfaktoren der Angebots- und Nutzungsebene auf die Kompetenzentwicklung .....	11
2.2.1 Einfluss unterrichtlicher Rahmenbedingungen .....	11
2.2.2 Einfluss individueller Lernendenmerkmale.....	13
2.3 Kompetenzbegriff und Bildungsstandards zur Beschreibung von Lernergebnissen .....	15
2.4 Beschreibung der Bildungsstandards im Fach NwT .....	15
2.4.1 Bildungsstandards für die Mittelstufe.....	16
2.4.2 Bildungsstandards für die Kursstufe (Basis- und Leistungsfach) .....	17
2.5 Kompetenzdiagnostik.....	21
2.6 Kompetenzmodelle in naturwissenschaftlichen Fächern .....	25
2.7 Erfassung naturwissenschaftlich-technischer Kompetenz .....	26
2.8 Zwischenfazit zum theoretischen Hintergrund .....	29
<b>3 Forschungsstand</b> .....	<b>33</b>
3.1 Kompetenzen in naturwissenschaftlichen Fächern .....	33
3.2 Kompetenzen im Kontext der technischen Allgemeinbildung .....	35
3.3 Stand der Forschung zum Fach NwT in Mittel- und Kursstufe .....	38
3.4 Zwischenfazit zum Forschungsstand .....	40

---

<b>4</b>	<b>Fragestellung und Anlage der Untersuchung .....</b>	<b>43</b>
4.1	Anlage der Untersuchung .....	43
4.1.1	Rahmenbedingungen des Leistungsfachs .....	45
4.1.2	Entwicklung und Pilotierung der Testinstrumente .....	45
4.1.3	Deskriptive Untersuchung des Fachwissens.....	46
4.1.4	Kompetenzniveaumodellierung & Zusammenhangsanalyse kognitiver Merkmale.....	46
<b>5</b>	<b>Methodische Grundlagen .....</b>	<b>49</b>
5.1	Forschungsansatz im Projekt <i>NwT-Kurstufe</i> – Design-Based Research .....	49
5.2	Exploration des Forschungsfelds .....	50
5.2.1	Expert:inneninterviews .....	50
5.2.2	Qualitative Inhaltsanalyse .....	51
5.2.3	Gütekriterien qualitativer Inhaltsanalyse.....	53
5.3	Aspekte der Testgestaltung zur Erfassung von Fachwissen .....	54
5.3.1	Testplanung.....	54
5.3.2	Itemkonstruktion .....	56
5.4	Gütekriterien quantitativer Testinstrumente .....	57
5.4.1	Objektivität, Reliabilität und Validität.....	57
5.4.2	Weitere Gütekriterien für die Testentwicklung .....	59
5.4.3	Statistische Auswertungskriterien für quantitative Testinstrumente .....	60
5.5	Klassische Testtheorie.....	61
5.5.1	Grundannahmen der Klassischen Testtheorie.....	61
5.5.2	Möglichkeiten der Reliabilitätsüberprüfung in der Klassischen Testtheorie .....	62
5.6	Probabilistische Testtheorie .....	65
5.6.1	Item Response Theory (IRT) .....	65
5.6.2	Das Rasch-Modell .....	65
5.6.3	Annahmen und Bedingungen für die Modellgültigkeit.....	66
5.6.4	Schätzung der Itemschwierigkeit und Personenfähigkeit.....	67
5.6.5	Prüfung der Modellgültigkeit.....	69
5.6.6	Übersicht der Kennwerte des Rasch-Modells .....	71
5.7	Explorative Datengruppierung .....	72
5.7.1	Clusteranalyse.....	72
5.8	Untersuchung von Zusammenhängen .....	73
5.9	Zwischenfazit zu den methodischen und statistischen Grundlagen.....	76
<b>6</b>	<b>Inhaltliche und kontextuelle Rahmenbedingungen des Leistungsfachs NwT .....</b>	<b>81</b>

6.1	Untersuchungsdesign und Stichprobe .....	81
6.2	Untersuchungsergebnisse .....	84
6.2.1	Demographische Rahmenbedingungen und individuelle Unterrichtserfahrung .....	84
6.2.2	Motivation und fachliche Voraussetzungen .....	84
6.2.3	Unterrichtsorganisation .....	85
6.2.4	Unterrichtsvorbereitung .....	85
6.2.5	Spiralcurriculare Kompetenzteilbereiche .....	86
6.2.6	Unterrichtsinhalte in der Kursstufe .....	86
6.2.7	Unterrichtsmethodik .....	87
6.2.8	Ausstattung der Schule und verfügbare Ressourcen .....	88
6.2.9	Kompetenzerwerb der Schüler:innen .....	89
6.2.10	Prüfungsformate .....	90
6.2.11	Qualifikation der Lehrpersonen .....	91
6.2.12	Wahrgenommene Unterstützung, zeitliche Faktoren und Deputat .....	92
6.3	Zwischenfazit zur Interviewstudie .....	92
6.4	Implikation der Lehrpersonenbefragung für die Testentwicklung .....	94
<b>7</b>	<b>Testentwicklung zur Erfassung inhaltsbezogener Kompetenzen im Leistungsfach NwT .....</b>	<b>95</b>
7.1	Testinstrument zur Erfassung des Vorwissens aus der Mittelstufe .....	96
7.1.1	Bezug zu den Kompetenzteilbereichen des Profilfachs der Mittelstufe .....	96
7.1.2	Inhaltliche Validierung der Testaufgaben zum Vorwissen .....	96
7.1.3	Testgüte des Vorwissenstests .....	99
7.2	Testinstrument zur Erfassung des Fachwissens in der Kursstufe .....	102
7.2.1	Bezug zu den Kompetenzteilbereichen des Leistungsfachs .....	102
7.2.2	Inhaltliche Validierung der Testaufgaben zum Fachwissen im Leistungsfach ....	102
7.2.3	Testgüte des Fachwissenstests .....	102
7.3	Zwischenfazit zur Testentwicklung und -pilotierung .....	105
<b>8</b>	<b>Fachwissen von Schüler:innen im Jahrgang 2019 – 2021 (L1) .....</b>	<b>107</b>
8.1	Studiendesign .....	107
8.2	Datenauswertung .....	108
8.3	Beschreibung der Stichprobe .....	108
8.4	Deskriptive Auswertung des Vorwissens .....	108
8.5	Fachwissen im ersten Leistungsfachdurchgang (L1) .....	109
8.5.1	Deskriptive Auswertung des Fachwissens – erster Testzeitpunkt (L1) .....	109
8.5.2	Deskriptive Auswertung des Fachwissens – zweiter Testzeitpunkt (L1) .....	110

8.5.3	Deskriptive Auswertung der Kompetenzteilbereiche <i>Grundlagen der Elektronik und Datenkommunikation</i> (L1) .....	111
8.6	Zwischenfazit zu den Ergebnissen aus Studie L1 .....	112
<b>9</b>	<b>Fachwissen von Schüler:innen im Jahrgang 2020 – 2022 (L2) .....</b>	<b>115</b>
9.1	Studiendesign .....	115
9.2	Datenauswertung.....	115
9.3	Beschreibung der Stichprobe .....	116
9.4	Deskriptive Auswertung des Vorwissens .....	116
9.5	Fachwissen im zweiten Leistungsfachdurchgang (L2).....	118
9.5.1	Deskriptive Auswertung des Fachwissens – erster Testzeitpunkt (L2) .....	118
9.5.2	Deskriptive Auswertung des Fachwissens – zweiter Testzeitpunkt (L2).....	118
9.5.3	Deskriptive Auswertung der Kompetenzteilbereiche <i>Grundlagen der Elektronik und Datenkommunikation</i> (L2) .....	120
9.6	Zwischenfazit zu den Ergebnissen aus Studie L2 .....	120
<b>10</b>	<b>Betrachtung der kombinierten Stichprobe L1 &amp; L2.....</b>	<b>123</b>
10.1	Zwischenfazit zu den deskriptiven Befunden .....	126
<b>11</b>	<b>Schätzen der Personenfähigkeiten.....</b>	<b>129</b>
11.1	Skalierung des Vorwissenstests.....	130
11.2	Skalierung des Fachwissenstests.....	132
11.3	Zwischenfazit der IRT-Skalierung.....	137
<b>12</b>	<b>Kompetenzniveaumodellierung .....</b>	<b>139</b>
12.1	Standard-Setting mittels Clusteranalyse.....	139
12.2	Kompetenzniveaustufen .....	140
12.3	Zwischenfazit zur Kompetenzniveaumodellierung.....	141
<b>13</b>	<b>Zusammenhangsanalyse kognitiver Voraussetzungen mit dem Fachwissen im Leistungsfach.....</b>	<b>143</b>
13.1	Korrelative Zusammenhänge der Kompetenzteilbereiche .....	145
13.2	Beschreibung der Kompetenzniveaus unter Einbezug kognitiver Merkmale .....	147
13.3	Zwischenfazit zu den Zusammenhangsanalysen kognitiver Voraussetzungen mit dem Fachwissen im Leistungsfach .....	149
<b>14</b>	<b>Resümee.....</b>	<b>151</b>
14.1	Zusammenfassung .....	151
14.2	Diskussion und Implikationen .....	155
14.3	Limitationen .....	157
14.4	Ausblick .....	158
	<b>Literaturverzeichnis.....</b>	<b>161</b>

<b>Referierte Publikationen .....</b>	<b>178</b>
<b>Anhänge .....</b>	<b>179</b>
Anhang 1: Kategoriensystem der Interviewstudie zu inhaltlichen und kontextuellen Rahmenbedingungen.....	179
Anhang 2: Codierungsrichtlinie der qualitativen Interviewauswertung.....	201
Anhang 3: Kommentiertes Aufgabenverzeichnis .....	203
Anhang 4: Ergebnisse der Expert:innenbefragung (Leistungsfach).....	217
Anhang 5: Pilotierungsergebnisse der Fachwissenstests (Leistungsfach) .....	226
Anhang 6: Ergebnisse der IRT-Skalierung.....	235
Rasch-Modell <i>Vorwissen Energie, Mobilität, Stoffe &amp; Produkte</i> .....	235
Rasch-Modell <i>Vorwissen Informationsaufnahme und -verarbeitung</i> .....	238
Rasch-Modell <i>Energieversorgung und Elektrische Antriebstechnik</i> .....	241
Rasch-Modell <i>Technische Mechanik, Produktentwicklung und Technikfolgenab-     schätzung</i> .....	244
Rasch-Modell <i>Messtechnik und Regelungstechnik</i> .....	248
Rasch-Modell <i>Grundlagen der Elektronik</i> .....	251
Anhang 7: Weitere Materialien.....	255



## Abbildungsverzeichnis

Abbildung 1: Angebots-Nutzungs-Modell unter Einbezug von Struktur- und Prozessparadigma (Seidel, 2014; unverändert übernommen).....	10
Abbildung 2: Prozess- und inhaltsbezogene Kompetenzen im Bildungsplan des Profulfachs NwT in der Mittelstufe (entnommen aus Brändle, 2023).....	16
Abbildung 3: Prozess- inhaltsbezogene Kompetenzen im Bildungsplan des Basis- und Leistungsfachs NwT in der Kursstufe (entnommen aus Brändle, 2023). ....	17
Abbildung 4: Modell naturwissenschaftlicher Kompetenz nach Kauertz et al. (Unverändert entnommen aus Kauertz et al., 2010, S. 145).....	26
Abbildung 5: Übersicht über den geplanten Ablauf des Schulversuchs <i>NwT-Kursstufe</i> . ....	43
Abbildung 6: Abfolge der Arbeitsschritte, Studien und Analysen zur Beantwortung der Forschungsfragen.....	44
Abbildung 7: Schematische Darstellung des DBR-Ansatzes im Kontext eines Bildungsmonitorings (Eigene Abbildung; Zitat: Anderson & Shattuck, 2012). ....	49
Abbildung 8: Exemplarische Darstellung einer Itemcharakteristikkurve (eigene Abbildung). ....	66
Abbildung 9: Mediatormodell mit drei Variablen (angelehnt an Hayes, 2022, S. 85). ....	76
Abbildung 10: Verortung der Studie zu den Rahmenbedingungen im Erhebungsablaufplan. ....	81
Abbildung 11: Verortung der Testentwicklung und Pilotierungsstudie im Erhebungsablaufplan. ....	95
Abbildung 12: Vorgehensweise bei der Entwicklung von Testaufgaben.....	95
Abbildung 13: Beispielaufgabe bestehend aus zwei Items zur statischen Beanspruchung von Brückenbauteilen aus dem Bereich <i>Eigenschaften von Stoffen</i> .....	101
Abbildung 14: Beispielaufgabe aus dem Bereich <i>Elektrische Antriebstechnik</i> des Testhefts zur Erfassung des Fachwissens aus der Kursstufe.....	105
Abbildung 15: Verortung der Studie zum Fachwissen im Leistungsfachdurchgang L1 im Erhebungsablaufplan. ....	107
Abbildung 16: Testergebnis des Vorwissenstests an den Pilotschulen im Leistungsfachdurchgang L1. Der rote Punkt in diesem und allen weiteren Boxplots entspricht dem Mittelwert. ....	109
Abbildung 17: Gesamtlösungsquote des Fachwissenstests an den Pilotschulen zum ersten Testzeitpunkt im Leistungsfachdurchgang L1. ....	110
Abbildung 18: Gesamtlösungsquote des Fachwissenstests an den Pilotschulen des Leistungsfachdurchgangs L1 am zweiten Testzeitpunkt vor dem Abitur. ....	111
Abbildung 19: Gesamtlösungsquote des Testteils <i>Grundlagen der Elektronik und Datenkommunikation</i> im Leistungsfachdurchgang L1. ....	111
Abbildung 20: Entwicklung der Gesamtlösungsquote über den Verlauf zweier Testzeitpunkte im Leistungsfachdurchgang L1: a) nach einem dreiviertel Schuljahr und b) vor dem Abitur. ....	113
Abbildung 21: Verortung der Studie zum Fachwissen im Leistungsfachdurchgang L2 im Erhebungsablaufplan. ....	115

---

Abbildung 22: Testergebnis des Vorwissenstests an den Pilot- und Modellschulen im Leistungsfachdurchgang L2. ....	116
Abbildung 23: Übersicht über die Lösungsquoten der Pilot- und Modellschulen im Vorwissenstest. Die Schüler:innen verteilen sich zu $n_{Modell} = 65$ und $n_{Pilot} = 22$ .....	117
Abbildung 24: Gesamtlösungsquote des Fachwissenstests an den Pilot- und Modellschulen zum ersten Testzeitpunkt im Leistungsfachdurchgang L2. ....	118
Abbildung 25: Gesamtlösungsquote des Fachwissenstests an den Pilot- und Modellschulen des Leistungsfachdurchgangs L2 am zweiten Testzeitpunkt vor dem Abitur. ....	119
Abbildung 26: Übersicht über die Lösungsquoten der Pilot- und Modellschulen im Fachwissenstest des Leistungsfachdurchgangs L2 vor der Abiturprüfung. Die erfassten Ergebnisse verteilen sich zu $n_{Modell} = 56$ auf die Modellschulen und $n_{Pilot} = 18$ auf die Pilotschulen.....	119
Abbildung 27: Gesamtlösungsquote des Testteils <i>Grundlagen der Elektronik</i> und <i>Datenkommunikation</i> im Leistungsfachdurchgang L2. ....	120
Abbildung 28: Vergleich der Gesamtlösungsquoten der beiden Testzeitpunkte nach einem dreiviertel Schuljahr (a) und vor dem Abitur (b) im Leistungsfachdurchgang L2.....	121
Abbildung 29: Verortung der kombinierten Betrachtung des Fachwissens der Leistungsfachdurchgänge L1 und L2 im Erhebungsablaufplan.....	123
Abbildung 30: Gesamtlösungsquote des Vorwissenstests der kombinierten Stichprobe aus L1 und L2. ....	123
Abbildung 31: Vergleich der Lösungsquoten von Pilot- und Modellschulen im Vorwissenstest. Der Zahlenwert vor den Boxplots entspricht der jeweiligen Anzahl der Schüler:innen in den Gruppen. ....	124
Abbildung 32: Vergleich der Lösungsquoten von Pilot- und Modellschulen im Fachwissen vor der Abiturprüfung. Der Zahlenwert vor den Boxplots entspricht der jeweiligen Anzahl der Schüler:innen in den Gruppen.....	125
Abbildung 33: Übersicht des Fachwissensfortschritts der Schüler:innen in den einzelnen Kompetenzteilbereichen des Bildungsplans zwischen den beiden Messzeitpunkten nach einem dreiviertel Schuljahr und vor der Abiturprüfung.....	125
Abbildung 34: Person-Item Maps der Dimensionen (a) <i>Energie, Mobilität, Stoffe &amp; Produkte</i> und (b) <i>Informationsaufnahme und -verarbeitung</i> . ....	131
Abbildung 35: Finale Skalen der Dimensionen <i>Energieversorgung und Elektrische Antriebstechnik; Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung</i> sowie <i>Messtechnik und Regelungstechnik</i> .....	135
Abbildung 36: Person-Item Map des Kompetenzteilbereichs <i>Grundlagen der Elektronik</i> .....	136
Abbildung 37: Vorwissen als zweidimensionales Konstrukt aus den Kompetenzbereichen <i>Energie, Mobilität, Stoffe und Produkte</i> sowie <i>Informationsaufnahme und -verarbeitung</i> .....	137
Abbildung 38: Fachwissen als vierdimensionales Konstrukt aus den Kompetenzteilbereichen <i>Energieversorgung und Elektrische Antriebstechnik; Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung; Messtechnik und Regelungstechnik</i> sowie <i>Grundlagen der Elektronik</i> . ....	137
Abbildung 39: Kompetenzniveaumodell der geschätzten Personenfähigkeiten der Lernenden im Leistungsfach NwT. Für jedes Niveau sind die jeweiligen Medianwerte	



$\theta_M$ dargestellt. Die Bildung der Niveaus erfolgte mit den hellblau umrandeten Fachwissensdimensionen.....	140
Abbildung 40: Korrelative Zusammenhänge zwischen fluider Intelligenz (IQ) sowie der mittleren Personenfähigkeiten des Vorwissens (Vw.) und Fachwissens (Fw.).....	143
Abbildung 41: Mediationsanalyse des Einflusses der fluiden Intelligenz (IQ) auf Vorwissen (Vw.) und Fachwissen (Fw.). Alle Pfadkoeffizienten sind standardisiert. Das Modell klärt $R^2 = 21\%$ der Varianz des Fachwissens auf. Die Interpretation der Effektstärken orientiert sich an Cohen (1988). .....	144
Abbildung 42: Darstellung der erwarteten und tatsächlich signifikanten Korrelationen zwischen den Vor- und Fachwissensdimensionen auf inhaltlicher Basis und der kognitiven Grundvoraussetzung. Die Korrelationswerte werden aus Übersichtsgründen nicht in die Abbildung integriert und sind in Tabelle 26 dargestellt. ....	146
Abbildung 43: Beschreibung der Kompetenzniveaus durch kognitive Auszeichnungsmerkmale <i>Personenfähigkeit Vorwissen</i> und <i>fluide Intelligenz (IQ)</i> . Abgetragen sind die jeweiligen Medianwerte. Die Bildung der Niveaus erfolgte mit den hellblau umrandeten Fachwissensdimensionen. ....	147



## Tabellenverzeichnis

Tabelle 1: Übersicht über die vorliegende Arbeit.....	6
Tabelle 2: Möglichkeiten zur Analyse des codierten Materials im Rahmen der inhaltlich strukturierenden qualitativen Inhaltsanalyse (Auswahl nach Kuckartz & Rädiker, 2022, S. 147-154).....	53
Tabelle 3: Übersicht über die relevanten Kennwerte zur Rasch-Modellierung. ....	71
Tabelle 4: Kurzdefinition der deduktiven Kategorien des Interviewleitfadens. ....	81
Tabelle 5: Verteilung der Codes auf das Kategoriensystem mit zugehöriger Interraterreliabilität.....	83
Tabelle 6: Inhaltliche Validierung der Aufgaben durch die Lehrpersonen. Die angegebenen Werte repräsentieren den Median der Stichprobe. Ausgegraute Aufgaben wurden nach der Expertenbefragung verworfen. ....	97
Tabelle 7: Statistische Kennwerte der Testanalyse nach der Klassischen Testtheorie. Die finale Trennschärfe bezieht sich auf die Skala mit dem höchsten Cronbachs Alpha Wert. ....	99
Tabelle 8: Übersicht der wichtigsten Testkennwerte der einzelnen Skalen zu den Kompetenzteilbereichen der Kursstufe. ....	103
Tabelle 9: Übersicht über die Normalverteilung der einzelnen Skalen zu den Kompetenzteilbereichen der Kursstufe. ....	104
Tabelle 10: Detaillierte Betrachtung der Lösungsquoten des Vorwissenstests der Studie 1 (Gesamtstichprobe).....	109
Tabelle 11: Status der Kompetenzteilbereiche beim ersten Testzeitpunkt in L1. ....	110
Tabelle 12: Ausführliche Betrachtung der Testergebnisse des Fachwissenstests der abiturrelevanten Kompetenzteilbereiche am Zeitpunkt der ersten Testung in L1. ....	110
Tabelle 13: Übersicht der Testergebnisse des Fachwissenstests der zweiten Testung der abiturrelevanten Kompetenzteilbereiche in L1. ....	110
Tabelle 14: Übersicht der Testergebnisse der Kompetenzteilbereiche <i>Grundlagen der Elektronik</i> und <i>Datenkommunikation</i> in L1.....	112
Tabelle 15: Ausführliche Betrachtung der Lösungsquoten des Vorwissenstests der Studie 2 (Gesamtstichprobe).....	117
Tabelle 16: Aufstellung der Testergebnisse des Fachwissenstests der abiturrelevanten Kompetenzteilbereiche am Zeitpunkt der ersten Testung im Leistungsfachdurchgang L2.....	118
Tabelle 17: Aufstellung der Ergebnisse der zweiten Testung des Fachwissens der Schüler:innen zu den abiturrelevanten Kompetenzteilbereichen im Leistungsfachdurchgang L2. ....	119
Tabelle 18: Übersicht der Ergebnisverteilung der Kompetenzteilbereiche <i>Grundlagen der Elektronik</i> und <i>Datenkommunikation</i> in L2. ....	120
Tabelle 19: Detaillierte Übersicht des Fachwissens der kombinierten Stichprobe der Pilot- und Modellschulen aus L1 und L2 vor der Abiturprüfung. ....	125
Tabelle 20: IRT-Skalierung des Vorwissens anhand der Dimensionen <i>Energie, Mobilität, Stoffe &amp; Produkte</i> und <i>Informationsaufnahme und -verarbeitung</i> .....	130

Tabelle 21: Korrelative Zusammenhänge der prozentualen Lösungsquoten in den zu modellierenden Kompetenzteilbereichen. ....	133
Tabelle 22: Modellparameter der Fachwissensdimensionen <i>Energieversorgung und Elektrische Antriebstechnik; Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung</i> sowie <i>Messtechnik und Regelungstechnik</i> .....	134
Tabelle 23: Kennwerte der IRT-Skalierung der Fachwissensdimension <i>Grundlagen der Elektronik</i> .....	136
Tabelle 24: Verteilung der mittleren Personenfähigkeit innerhalb der Cluster.....	139
Tabelle 25: ANOVA der geschätzten Personenfähigkeit in der Fachwissensdimension <i>Grundlagen der Elektronik</i> .....	140
Tabelle 26: Korrelative Zusammenhänge der kognitiven Auszeichnungsmerkmale (Personenfähigkeiten Vor- und Fachwissen sowie fluide Intelligenz). ....	145
Tabelle 27: ANOVAs der kognitiven Auszeichnungsmerkmale.....	148
Tabelle 28: Kategoriensystem der Interviewstudie. ....	179
Tabelle 29: Verzeichnis der Aufgaben des Vorwissenstests mit Herkunftsnachweis. ....	203
Tabelle 30: Verzeichnis der Aufgaben des Fachwissenstests mit Herkunftsnachweis. ....	206
Tabelle 31: Medianwerte der Expert:innenbefragung zu <i>Technikfolgenabschätzung</i> ( $n = 5$ ). ...	217
Tabelle 32: Medianwerte der Expert:innenbefragung zu <i>Energieversorgung</i> ( $n = 5$ ). ....	218
Tabelle 33: Medianwerte der Expert:innenbefragung zu <i>Elektrische Antriebstechnik</i> ( $n = 5$ ). .....	219
Tabelle 34: Medianwerte der Expert:innenbefragung zu <i>Technische Mechanik</i> ( $n = 6$ ). ....	220
Tabelle 35: Medianwerte der Expert:innenbefragung zu <i>Produktentwicklung</i> ( $n = 5$ ). ....	221
Tabelle 36: Medianwerte der Expert:innenbefragung zu <i>Grundlagen der Elektronik</i> ( $n = 5$ ). ....	222
Tabelle 37: Medianwerte der Expert:innenbefragung zu <i>Aufnahme und Verarbeitung von Signalen in der Messtechnik</i> ( $n = 5$ ).....	223
Tabelle 38: Medianwerte der Expert:innenbefragung zu <i>Regelungstechnik</i> ( $n = 5$ ).....	224
Tabelle 39: Medianwerte der Expert:innenbefragung zu <i>Datenkommunikation</i> ( $n = 6$ ). ....	225
Tabelle 40: Pilotierungsergebnisse der Skala <i>Technikfolgenabschätzung</i> . ....	226
Tabelle 41: Pilotierungsergebnisse der Skala <i>Energieversorgung</i> . ....	227
Tabelle 42: Pilotierungsergebnisse der Skala <i>Elektrische Antriebstechnik</i> .....	228
Tabelle 43: Pilotierungsergebnisse der Skala <i>Technische Mechanik</i> . ....	229
Tabelle 44: Pilotierungsergebnisse der Skala <i>Produktentwicklung</i> . ....	230
Tabelle 45: Pilotierungsergebnisse der Skala <i>Grundlagen der Elektronik</i> . ....	231
Tabelle 46: Pilotierungsergebnisse der Skala <i>Aufnahme und Verarbeitung von Signalen in der Messtechnik</i> .....	232
Tabelle 47: Pilotierungsergebnisse der Skala <i>Regelungstechnik</i> .....	233
Tabelle 48: Pilotierungsergebnisse der Skala <i>Datenkommunikation</i> . ....	234

---

## Abkürzungsverzeichnis

---

<b>ANOVA</b>	Einfaktorielle Varianzanalyse ( <i>engl.</i> Analysis of Variance)
<b>BF</b>	Bewegung und Fortbewegung
<b>BP</b>	Bildungsplan
<b>bspw.</b>	beispielsweise
<b>bzw.</b>	beziehungsweise
<b>C<sub>α</sub></b>	Cronbachs Alpha
<b>CAD</b>	Rechnergestütztes Konstruieren ( <i>engl.</i> computer-aided design)
<b>CML</b>	Conditional Maximum Likelihood
<b>cov</b>	Kovarianz ( <i>engl.</i> covariance)
<b>curr.</b>	curricular
<b><i>d</i></b>	Effektstärke
<b>DBR</b>	Design-Based Research
<b>DIF</b>	Differential Item Functioning
<b>E</b>	Energie in Natur und Technik
<b>EA</b>	Elektrische Antriebstechnik
<b>ebd.</b>	ebenda
<b>EAP</b>	Expected A Posteriori
<b>EigS</b>	Eigenschaften von Stoffen
<b>En</b>	Energie
<b>engl.</b>	englisch
<b>EPA</b>	Einheitliche Prüfungsanforderungen in der Abiturprüfung
<b>ES</b>	Elektronische Schaltungen
<b>EVA</b>	Eingabe-Verarbeitung-Ausgabe
<b>Ev</b>	Energieversorgung
<b>Evs</b>	Energieversorgungssysteme
<b>Fw</b>	Fachwissen
<b>GAD</b>	Gewinnung und Auswertung von Daten
<b>GE</b>	Grundlagen der Elektronik
<b>ggf.</b>	gegebenenfalls
<b>la</b>	Informationsaufnahme durch Sinne und Sensoren

---

<b>ICC</b>	Itemcharakteristikkurve ( <i>engl.</i> Item Characteristic Curve)
<b>IGLU</b>	Internationale Grundschul-Lese-Untersuchung
<b>Infa</b>	Informationsaufnahme und -verarbeitung
<b>inkl.</b>	inklusive
<b>IPN</b>	Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik
<b>IRT</b>	Item Response Theory
<b>IQ</b>	Intelligenzquotient
<b>ITEEA</b>	International Technology and Engineering Educators Association
<b>IQB</b>	Institut zur Qualitätsentwicklung im Bildungswesen
<b>lv</b>	Informationsverarbeitung
<b>JML</b>	Joint Maximum Likelihood
<b>k. A./kA</b>	keine Angabe
<b>KI</b>	Konfidenzintervall
<b>KIT</b>	Karlsruher Institut für Technologie
<b>KMK</b>	Kultusministerkonferenz
<b>KTT</b>	Klassische Testtheorie
<b>L</b>	Leistungsfachdurchgang
<b>LCA</b>	Latent Class Analysis
<b>LRT</b>	Likelihood Ratio Test
<b>Lp</b>	Lehrperson
<b>LQ</b>	Lösungsquote
<b>m</b>	männlich
<b>Max./max.</b>	Maximum/maximal
<b>Med.</b>	Median
<b>Min./min.</b>	Minimum/minimal
<b>MKJS</b>	Ministerium für Kultus, Jugend und Sport
<b>MNSQ</b>	Mean square
<b>Mo</b>	Mobilität
<b>modif.</b>	modifiziert
<b>Mt</b>	Messtechnik

<b>MW</b>	Mittelwert
<b>n</b>	Stichprobengröße
<b>NEPS</b>	National Educational Panel Study
<b>NIMBY</b>	Not in my backyard
<b>NwT</b>	Naturwissenschaft und Technik (Schulfach in Baden-Württemberg)
<b>o. Ä.</b>	oder Ähnliches
<b>o. a.</b>	oben angeführt
<b>OECD</b>	Organisation für wirtschaftliche Zusammenarbeit und Entwicklung ( <i>engl.</i> Organization for Economic Co-operation and Development)
<b>p</b>	Signifikanz
<b>Pe</b>	Produktentwicklung
<b>PISA</b>	Programme for International Student Assessment
<b>Pr</b>	Produkt(e)
<b>Q</b>	Quartil
<b>R<sup>2</sup></b>	Determinationskoeffizient (Varianzaufklärung)
<b>r</b>	Produkt-Moment-Korrelationskoeffizient
<b>Relev.</b>	Relevanz
<b>Rel</b>	Reliabilität
<b>Rt</b>	Regelungstechnik
<b>S.</b>	Seite
<b>s.</b>	siehe
<b>SD/sd</b>	Standardabweichung ( <i>engl.</i> Standard deviation)
<b>Sek.</b>	Sekundarstufe
<b>sic</b>	Fehlerhinweis in Zitaten ( <i>lateinisch</i> sic erat scriptum)
<b>sog.</b>	sogenannte
<b>s. o.</b>	siehe oben
<b>SP</b>	Systeme und Prozesse
<b>SsV</b>	Stoffströme und Verfahren
<b>St</b>	Stoffe
<b>StP</b>	Statische Prinzipien in Natur und Technik

---

<b>SuS</b>	Schülerinnen und Schüler
<b>Ta</b>	Technikfolgenabschätzung
<b>TIMSS</b>	Trends in International Mathematics and Science Study
<b>TM</b>	Technische Mechanik
<b>u. a.</b>	unter anderem
<b>v. a.</b>	vor allem
<b>vgl.</b>	vergleiche
<b>VERA</b>	Vergleichsarbeiten
<b>VIF</b>	Varianz-Inflations-Faktor ( <i>engl.</i> Variance Inflation Factor)
<b>Vw</b>	Vorwissen
<b>W</b>	Shapiro-Wilk-Test
<b>w</b>	weiblich
<b>WLE</b>	Weighted Likelihood Estimators
<b>z. B.</b>	zum Beispiel
<b>Zsha</b>	Zusammenhangsanalyse
<b>zw.</b>	zwischen

---



## 1 Einleitung

Nationale und internationale Institutionen (z. B. European Commission. Directorate General for Education, Youth, Sport and Culture & ECORYS., 2020; International Technology and Engineering Educators Association [ITEEA], 2020; Stifterverband für die Deutsche Wissenschaft e. V., 2021) fordern einschlägige Kompetenzen, um Bürger:innen eine zukunftsfähige gesellschaftliche, wissenschaftliche und wirtschaftliche Teilhabe zu ermöglichen. Die übergeordneten Leitziele sind in Rahmenwerken wie bspw. den *21st Century Skills* (z. B. Kennedy & Sundberg, 2020; van Laar, van Deursen, van Dijk & Haan, 2017) oder den *Future Skills 2021* (Stifterverband für die Deutsche Wissenschaft e. V., 2021) dargestellt. Die technische Literalität soll aus diesen Rahmenwerken in Form grundlegender technischer Kompetenzen hervorgehen und gewinnt durch den sich beschleunigenden digitalen und technologischen Transformationsprozess immer mehr an Bedeutung. Technologische Grundkompetenz steht im digitalen Zeitalter längst in Zusammenhang mit jedem Lebensbereich. Die ITEEA (2020, S. 9) differenziert die technische Dimension weiter aus und fordert eine aus acht Bereichen bestehende (s. Kapitel 2.7) obligatorische Grundbildung. Die Forderung nach elaborierten technischen Grundkompetenzen beeinflusst seit einigen Jahren auch die Diskussion um eine zukunftsfähige Schulbildung (s. z. B. OECD, 2019, 2020, 2022) und betrifft damit auch den allgemeinbildenden Bereich. Etablierte naturwissenschaftliche Fächer fördern technische Kompetenzen nur wenig (z. B. Theuerkauf, 2009) und es existieren diesbezüglich kaum spezifische Fächerangebote an allgemeinbildenden Schulen (Zinn, 2018c, S. 232). In Deutschland wirft die Diskussion zu PISA 2015 (Reiss, Sälzer, Schiepe-Tiska, Klieme & Köller, 2016) zudem auf, dass naturwissenschaftlicher Unterricht wenig „Hands-on“-Aktivierung bietet, die aber u. a. zur Förderung der technischen Grundbildung in der unterrichtlichen Umsetzung (z. B. *Making and Doing*) vonnöten ist (vgl. ITEEA, 2020, S. 11). Es mangelt an Fächern, die Technik in Form eines ganzheitlichen Ansatzes betrachten (vgl. VDMA, 2019, S. 27). Meist sind solche Fächer in den Schulformen Haupt- oder Realschule und seltener am Gymnasium verortet; wenn sie für den allgemeinbildenden Bereich in den Bundesländern überhaupt existieren (vgl. ebd., S. 27). Damit besitzt eine Betrachtung interdisziplinärer und handlungsorientierter Fächeransätze im Kontext naturwissenschaftlich-technischer Allgemeinbildung am Gymnasium und daraus ableitbare Implikationen eine gewisse Relevanz für das nationale Bildungssystem. Ein solches Fach ist *Naturwissenschaft und Technik (NwT)* am allgemeinbildenden Gymnasium in Baden-Württemberg, dessen

Genese und Bildungswert im Kontext technischer Allgemeinbildung im Folgenden betrachtet wird.

### **1.1 Genese und Bildungswert des Schulfachs NwT<sup>1</sup>**

Die Schaffung des Fachs NwT soll den Schüler:innen am allgemeinbildenden Gymnasium neben einer naturwissenschaftlichen Grundbildung eine interdisziplinär ausgerichtete technische Literalität bis zum Abschluss des Abiturs zugänglich machen und die damit verbundene Interessen- und Kompetenzentwicklung ermöglichen. NwT wurde zu diesem Zweck im Schuljahr 2007/08 in Baden-Württemberg innerhalb des MINT-Fächerclusters als interdisziplinäres Fach für die 8. bis 10. Jahrgangsstufe an allgemeinbildenden Gymnasien flächendeckend eingeführt<sup>2</sup>. Eine Besonderheit des Fachs ist seine formale Stellung als Ersatz für eine dritte Fremdsprache. Das bedeutet, dass neben dem Fach NwT alle naturwissenschaftlichen Fächer ohne Einschränkungen im Sinne einer MINT-Profilierung gewählt werden können (vgl. Offermann & Schäfer, 2012, S. 191). Von verschiedenen Interessenvertreterinnen und -vertretern einer allgemein technischen Bildung (z. B. Offermann & Schäfer, 2012; Pittschellis, 2012) wurden bei der Einführung des neuen Schulfachs NwT fünf Argumente vertreten: „(1.) den nicht hinreichend befriedigten Bedarf an Ingenieuren, (2.) im Einklang mit technikphilosophischen und bildungstheoretischen Argumenten [...] die Durchdringung gesellschaftlicher Praxis durch Technologien und die Notwendigkeit, diese zu reflektieren und zu gestalten, (3.) positive Effekte eines in allen Schularten implementierten Technikunterrichts auf die Durchlässigkeit des Schulsystems, (4.) positive Rückkopplungen auf die Motivationsentwicklung in Mathematik und den Naturwissenschaften und schließlich (5.) die erweiterte Option, den eigenen Bildungsweg spezifisch zu akzentuieren“ (Mokhonko, Stefanica & Nickolaus, 2014, S. 104).

Bereits im Schuljahr 2008/09 erfolgte die Einführung als zweistündiges Fach in der Kursstufe in Form eines Schulversuchs an ausgewählten Modellversuchsschulen (Zinn, 2018c, S. 233). Seit dem Wintersemester 2010/11 ist es möglich, NwT als Lehramtsfach an den Universitäten *Karlsruher Institut für Technologie*, *Universität Stuttgart*, *Eberhard-Karls-Universität Tübingen* und *Universität Ulm* zu studieren. Mit der Revision des Bildungsplans erfolgte ab dem Schuljahr 2016/17 die wissenschaftliche Begleitung des zweistündigen Oberstufenfachs. Im

---

<sup>1</sup> Teile dieses Kapitels entstammen in identischer oder leicht abgewandelter Form einer bereits erschienenen Publikation: Brändle (2023)

<sup>2</sup> Aufgrund der immer drängenderen Problematik des Fachkräftemangels im MINT-Bereich wurde bereits 1995 von Offermann und Schäfer (2012, S. 183) ein Antrag zur Einrichtung des Fachs NwT an das Ministerium für Kultus, Jugend und Sport (MKJS) Baden-Württemberg gestellt.

Schuljahr 2018/19 startete der Schulversuch des zunächst vierstündigen Kursstufenfachs (ohne Abitur) an sechs Pilotschulen. Ein Schuljahr darauf erfolgte die Reform der Kursstufe, wodurch NwT als Leistungsfach an fünf Pilotschulen fünfständig mit Abitur angeboten werden konnte. Seit dem Schuljahr 2020/21 sind sieben weitere Modellschulen in den Schulversuch eingemündet, die das fünfständige Leistungsfach inkl. Abitur anbieten. Im Juli 2022 konnte bereits der zweite Jahrgang und insgesamt mehr als 100 Schüler:innen das Leistungsfach NwT mit Abitur abschließen. Das Fach zählt formal nur inhaltlich zum MINT-Bereich (Ministerium für Kultus, Jugend und Sport Baden-Württemberg [MKJS], 2022, S. 6), wird aber in der Kurswahl nicht als naturwissenschaftliches bzw. technisches Fach berücksichtigt (ebd., S. 7). Dadurch ergeben sich Einschränkungen, die z. B. eine Kombination mit Wirtschaft oder Informatik (das sich ebenfalls im Schulversuch befindet) als Leistungsfach aktuell nicht zulassen. Van Laar et al. (2017; 2020) stellen z. B. fest, dass grundlegende technische und informations- bzw. kommunikationstechnische Kompetenzen im Kontext der *21st century skills* vernetzt sind. Insgesamt betrachtet hat sich die technische Bildung trotz bestehender Herausforderungen an allgemeinbildenden Gymnasien in Baden-Württemberg in Mittel- und Kursstufe etabliert.

Das Bedeutungsspektrum der technischen Allgemeinbildung ist breit, wobei die Begründung aus einer bildungstheoretischen, soziologischen, wissenschaftstheoretischen und berufspraktischen Perspektive erfolgen kann (vgl. Zinn, 2014, S. 27). NwT adressiert als Schulfach am allgemeinbildenden Gymnasium diese Bereiche, indem es (1.) ein lebensweltnahes Verständnis von Technik kompetenzorientiert fördert, welches Schüler:innen (2.) die Teilhabe in einer durch Technologie beeinflussten und sich dynamisch wandelnden Gesellschaft u. a. im Sinne der *21st Century Skills* ermöglicht. Das Fach vertritt (3.) die Interessen einer allgemeinen Technikwissenschaft im Sinne einer interdisziplinären Korrespondenz natur- und ingenieurwissenschaftlicher Domänen (vgl. Zinn, 2014, S. 28) und leistet (4.) seinen Beitrag zur Vorbereitung auf Beruf und Studium innerhalb natur- und technikwissenschaftlicher Domänen, indem es Denk- und Arbeitsweisen fördert, die den Kompetenzerwerb im beruflichen und hochschulischen Ausbildungsbereich unterstützen (s. z. B. Mokhonko et al., 2014, S. 104; Zinn, 2014, S. 29). Darüber hinaus soll die interdisziplinäre Themenvielfalt des Fachs NwT eine Entwicklung von Interesse, Motivation und beruflicher Orientierung im Kontext naturwissenschaftlich-technischer Domänen bei den Lernenden in Mittel- und Oberstufe ermöglichen. Dass NwT zur technischen Grundbildung beiträgt, ist unbestritten (z. B. Zinn & Latzel, 2017; Zinn, Latzel & Ariali, 2017).

## 1.2 Relevanz und Ziele der Untersuchung

Nach dem kurzen Überblick über die Wirkungsziele des Fachs NwT in Bezug zur technischen Allgemeinbildung stellt sich die Frage nach einer aktuellen Bestandsaufnahme der tatsächlichen Kompetenzentwicklung der Lernenden. Fleischer, Koeppen, Kenk, Klieme und Leutner (2013, S. 6) vertreten die Meinung, dass Kompetenzmessung und damit evidenzorientiert angestoßene Optimierungsprozesse zur Qualitätssicherung und -entwicklung im Bildungswesen beitragen. Sie postulieren:

„Prosperität, soziale Kohäsion und Entwicklungschancen einer Gesellschaft hängen in großem Maße von den erworbenen Kompetenzen ihrer Mitglieder ab.“ (Fleischer et al., 2013, S. 6)

Insbesondere zum Bildungoutput des im Schulversuch befindlichen Leistungsfachs liegen aber bislang keine Erkenntnisse vor. Die wissenschaftliche Begleitung des Leistungsfachs im Kontext des Forschungsprojekts *NwT-Kurstufe* und die damit verbundene Generierung empirischer Erkenntnisse zum Bildungoutput adressiert daher nicht nur eine Forschungslücke, sondern ermöglicht bereits während der Einführungsphase des Schulfachs eine formative Prozessoptimierung. Die vorliegende Arbeit verfolgt als Hauptziel die Generierung eines Beschreibungswissens zu den inhaltsbezogenen Kompetenzen bei Lernenden im Leistungsfach NwT. Die Evaluationsergebnisse können mitunter bedeutsam für die schulische Praxis sein, weil sie Aussagen über den Wissensstand der Schüler:innen im Kontext vergleichbarer Voraussetzungen der Lernenden für die schriftliche Abiturprüfung ermöglichen. Eine summative Untersuchung hinsichtlich der inhaltsbezogenen Kompetenzniveaus und die Beschreibung der niveaustufenspezifischen Auszeichnungsmerkmale der Lernenden kann Ansatzpunkte für zukünftige Optimierungen im Leistungsfach eröffnen.

## 1.3 Aufbau der Arbeit

Die Arbeit besteht aus insgesamt 14 Kapiteln und fokussiert dabei als Hauptziel die Generierung eines Beschreibungswissens zum Leistungsfach NwT im Rahmen des Schulversuchs *NwT-Kurstufe*. Die Einleitung umreißt dazu die Notwendigkeit technischer Allgemeinbildung am Gymnasium, gibt einen Einblick in die Genese und den Bildungswert des Schulfachs NwT und stellt die Relevanz und übergeordneten Ziele der Untersuchung kurz dar. Die darauffolgenden Kapitel setzen sich mit dem theoretischen Rahmen und dem Forschungsstand auseinander, um Forschungsdesiderate zu identifizieren. Aus den Forschungsdesideraten ergibt sich die konkrete Fragestellung und Anlage der Untersuchung. In den methodischen und statistischen Grundlagen wird dargestellt, wie die Konstrukte im

Rahmen des Erkenntnisinteresses operationalisiert und statistisch überprüft werden können. Es werden inhaltliche und kontextuelle Rahmenbedingungen untersucht, valide und reliable Testinstrumente zur Operationalisierung inhaltsbezogener Kompetenzen entwickelt sowie zu deren Erfassung in Form von Vor- und Fachwissen der Schüler:innen in zwei Leistungsfachdurchgängen eingesetzt. Eine Kompetenzniveaumodellierung und die Untersuchung von Einflüssen kognitiver Merkmale auf das Fachwissen schließen den Forschungsteil der Arbeit ab. In einem Resümee werden die zentralen Befunde zusammengefasst, diskutiert und Implikationen für die schulische Praxis sowie zukünftige Forschungsvorhaben abgeleitet. Tabelle 1 stellt die einzelnen Kapitel übersichtlich dar.

**Tabelle 1:** Übersicht über die vorliegende Arbeit.

<b>Kapitel 1: Einleitung</b>		
1.1 Genese und Bildungswert des Schulfachs NwT	1.2 Relevanz und Ziele der Untersuchung	1.3 Aufbau der Arbeit
<b>Kapitel 2 &amp; 3: Theoretischer Rahmen und Forschungsstand</b>		
<b>2 Theoretischer Rahmen</b> 2.1 Schulisches Lernen 2.2 Einflussfaktoren der Angebots- und Nutzungsebene auf die Kompetenzentwicklung 2.3 Kompetenzbegriff und Bildungsstandards zur Beschreibung von Lernergebnissen 2.4 Beschreibung der Bildungsstandards im Fach NwT 2.5 Kompetenzdiagnostik 2.6 Kompetenzmodelle in naturwissenschaftlichen Fächern 2.7 Erfassung naturwissenschaftlich-technischer Kompetenz 2.8 Zwischenfazit zum theoretischen Hintergrund	<b>3 Forschungsstand</b> 3.1 Kompetenzen in naturwissenschaftlichen Fächern 3.2 Kompetenzen im Kontext der technischen Allgemeinbildung 3.3 Stand der Forschung zum Fach NwT in Mittel- und Kursstufe 3.4 Zwischenfazit zum Forschungsstand	
<b>Kapitel 4 &amp; 5: Fragestellung, Anlage der Untersuchung und Methodik</b>		
<b>4 Fragestellung und Anlage der Untersuchung</b> <ul style="list-style-type: none"> <li>• Darstellung des Forschungsanliegens der Arbeit.</li> <li>• Beschreibung der Vorgehensweise bei der Anlage der Untersuchung unter Einbezug der Fragestellung sowie der methodischen und statistischen Grundlagen.</li> </ul>	<b>5 Methodische Grundlagen</b> 5.1 Forschungsansatz im Projekt NwT-Kursstufe 5.2 Exploration des Forschungsfelds 5.3 Aspekte der Testgestaltung zur Erfassung von Fachwissen 5.4 Gütekriterien quantitativer Testinstrumente 5.5 Klassische Testtheorie 5.6 Probabilistische Testtheorie 5.7 Explorative Datengruppierung 5.8 Untersuchung von Zusammenhängen 5.9 Zwischenfazit zu den methodischen und statistischen Grundlagen	

<b>Kapitel 6 – 10: Rahmenbedingungen und deskriptive Analysen</b>			
<b>6 Inhaltliche und kontextuelle Rahmenbedingungen des Leistungsfachs NwT</b> 6.1 Untersuchungsdesign und Stichprobe 6.2 Untersuchungsergebnisse 6.3 Zwischenfazit zur Interviewstudie 6.4 Implikation der Lehrpersonenbefragung für die Testentwicklung	<b>7 Testentwicklung zur Erfassung inhaltsbezogener Kompetenzen im Leistungsfach NwT</b> 7.1 Testinstrument zur Erfassung des Vorwissens aus der Mittelstufe 7.2 Testinstrument zur Erfassung des Fachwissens in der Kursstufe 7.3 Zwischenfazit zur Testentwicklung	<b>8 Fachwissen von Schüler:innen im Jahrgang 2019 – 2021 (L1)</b>  8.1 Studiendesign 8.2 Datenauswertung 8.3 Beschreibung der Stichprobe 8.4 Deskriptive Auswertung des Vorwissens 8.5 Fachwissen im ersten Leistungsfachdurchgang (L1) 8.6 Zwischenfazit zu den Ergebnissen aus Studie L1	<b>9 Fachwissen von Schüler:innen im Jahrgang 2020 – 2022 (L2)</b>  9.1 Studiendesign 9.2 Datenauswertung 9.3 Beschreibung der Stichprobe 9.4 Deskriptive Auswertung des Vorwissens 9.5 Fachwissen im zweiten Leistungsfachdurchgang (L2) 9.6 Zwischenfazit zu den Ergebnissen aus Studie L2
		<b>10 Betrachtung der kombinierten Stichprobe L1 &amp; L2</b>	

<b>Kapitel 11 – 13: Kompetenzniveaumodellierung</b>		
<b>11 Schätzen der Personenfähigkeiten</b>  11.1 Skalierung des Vorwissentests 11.2 Skalierung des Fachwissentests 11.3 Zwischenfazit der IRT-Skalierung	<b>12 Kompetenzniveaumodellierung</b>  12.1 Standard-Setting mittels Clusteranalyse 12.2 Kompetenzniveaustufen 12.3 Zwischenfazit zur Kompetenzniveaumodellierung	<b>13 Zusammenhangsanalyse kognitiver Voraussetzungen mit dem Fachwissen im Leistungsfach</b>  13.1 Korrelative Zusammenhänge der Kompetenzteilbereiche 13.2 Beschreibung der Kompetenzniveaus 13.3 Zwischenfazit zu den Zusammenhangsanalysen kognitiver Voraussetzungen mit dem Fachwissen im Leistungsfach

<b>Kapitel 14: Resümee</b>			
<b>14.1 Zusammenfassung</b>	<b>14.2 Diskussion und Implikationen</b>	<b>14.3 Limitationen</b>	<b>14.4 Ausblick</b>





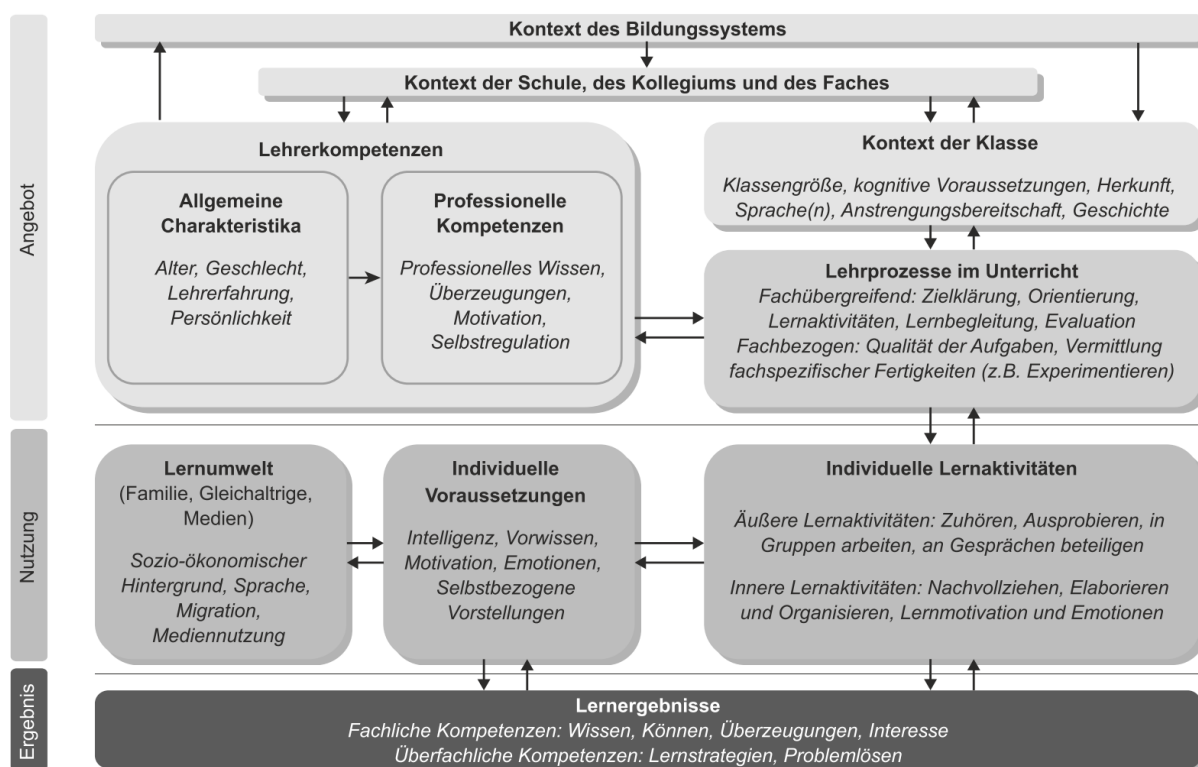
## **2 Theoretischer Rahmen**

In Anbetracht der Zielsetzung in der Arbeit erscheint es im theoretischen Rahmen zunächst wichtig, das schulische Lernen (Abschnitt 2.1) und die damit verbundenen Einflussfaktoren auf die Lernergebnisse (Abschnitt 2.2) zu betrachten. Abschnitt 2.3 stellt dazu den Kompetenzbegriff und seine Verortung in den Bildungsstandards in Deutschland dar. Es schließt sich eine Beschreibung der Bildungsstandards im Fach Naturwissenschaft und Technik (Abschnitt 2.4) an. Ausgehend davon wird in Abschnitt 2.5 Kompetenzdiagnostik als ein Element im Bereich der Evaluation von Bildungsprozessen (z. B. Fleischer et al., 2013) erörtert und durch den Einbezug naturwissenschaftlicher Kompetenzmodelle eingeordnet (Abschnitt 2.6). Abschließend wird eine Möglichkeit zur Operationalisierung der durch das Schulfach NwT geförderten inhaltsbezogenen Kompetenzen vorgenommen (Abschnitt 2.7). Der letzte Abschnitt 2.8 diskutiert den dargestellten theoretischen Rahmen und stellt einen Bezug zum Forschungsvorhaben der Arbeit her.

### **2.1 Schulisches Lernen**

In der Forschungsliteratur wird zur Darstellung des schulischen Lernens unter Berücksichtigung unterrichtlicher und individueller Rahmenbedingungen zwischen Struktur- und Prozessparadigma unterschieden (vgl. z. B. Seidel, 2014). Dem Strukturparadigma liegen die Annahmen zu Grunde, dass sich einzelne Elemente der Angebotsebene (z. B. schulische Rahmenbedingungen, Lehrpersonenkompetenz usw.) auf die Lernergebnisse der Schüler:innen auswirken. Der forschungsmethodische Feldzugang erfolgt über eine Akquise repräsentativer Stichproben, in denen eine Einschätzung des Unterrichts durch die Lernenden selbst oder externe Beobachter in Relation zu den erreichten Lernergebnissen gesetzt wird (ebd.). Das Prozessparadigma fokussiert die internen Lernprozesse auf der Nutzungsebene und wie deren Aktivierung in Zusammenhang mit den unterrichtlichen Rahmenbedingungen stehen. Aus der Modellierung von Lernprozessen anhand psychologischer Theorien kann z. B. mittels standardisierten Testinstrumenten überprüft werden, inwiefern die in der Theorie als relevant betrachteten Prozesse in Relation zu Unterrichtsmerkmalen zu erfolgreichem Lernen führen (vgl. Seidel, 2014, S. 854 f). Einflüsse affektiver Lernendenmerkmale auf den Lernerfolg können zum Prozessparadigma zugeordnet werden (ebd., S. 856). Bekannte Ansätze in diesem Kontext sind u. a. Theorien zur motivationalen Orientierung von Deci und Ryan (1993) und zum Einfluss des Interesses von Krapp und Prenzel (1992).

Seidel (2014) geht jedoch davon aus, dass „[...] sich die komplexen Phänomene des Lehrens und Lernens im Unterricht nicht durch eine einzige psychologische Theorie erklären [lassen]“ (Seidel, 2014, S. 857) und schlägt vor, beide Paradigmen in einem Angebots-Nutzungs-Modell (Abbildung 1) zu vereinen. Dieses Modell ermöglicht gleichzeitig einen Überblick über die komplexe Vernetzung von einzelnen Unterrichtselementen und deren Nutzung durch Lernende (vgl. auch Seidel & Reiss, 2014) sowie eine Einordnung von Forschungsvorhaben, die sich aus Gründen der Forschungsökonomie meist nur auf einzelne Facetten fokussieren können (Seidel, 2014). Im Folgenden werden die Ebenen des integrierten Angebots-Nutzungs-Modells ausführlicher beschrieben und deren Zusammenhänge erläutert.



**Abbildung 1:** Angebots-Nutzungs-Modell unter Einbezug von Struktur- und Prozessparadigma (Seidel, 2014; unverändert übernommen).

### 2.1.1 Unterrichtliche Rahmenbedingungen – die Angebotsebene

Die Angebotsebene zeigt nach Seidel (2014) eine Vernetzung der Lehrpersonenkompetenz, bestehend aus zwei Subgruppen, den demographischen Voraussetzungen wie z. B. Alter oder Persönlichkeit und Professionalisierung, mit den Lehrprozessen im Unterricht (ebd.). Die Umsetzbarkeit der Lehrprozesse hängt wiederum mit dem Kontext der Klasse zusammen (ebd.). Dem Ganzen übergeordnet

steht der schulische Kontext, der Kontext des Kollegiums sowie des unterrichteten Unterrichtsfachs (ebd.). Als oberste Instanz der Angebotsebene stehen die kontextuellen Vorgaben des Bildungssystems (ebd.).

### **2.1.2 Affektive Lernendenmerkmale – die Nutzungsebene**

Die Angebotsebene ist über die Lernaktivitäten der Schüler:innen mit der Nutzungsebene verknüpft. Die Lernaktivitäten werden von individuellen Lernendenmerkmalen beeinflusst, die im Zusammenhang mit der Lebenswelt der Lernenden steht (Seidel, 2014). Abschließend sieht Seidel (2014) einen Zusammenhang dieser individuellen Voraussetzungen und Lernaktivitäten mit den Lernergebnissen, die sich als fachliche Kompetenzen der Schüler:innen darstellen.

### **2.1.3 Lernergebnisse**

Die Lernergebnisse der Schüler:innen stellen sich bivalent dar. Die Fachkompetenz in Form von Wissen und Können (vgl. Kapitel 2.3) repräsentiert die kognitiven Facetten (ebd.). Demgegenüber stehen nicht-kognitive Facetten, wie die Überzeugungen und das Interesse als fachliche sowie Lernstrategien und das Problemlösen als überfachliche Kompetenzen (Seidel, 2014).

## **2.2 Einflussfaktoren der Angebots- und Nutzungsebene auf die Kompetenzentwicklung**

Im vorherigen Abschnitt wurde das Angebots-Nutzungs-Modell nach Seidel (2014) vorgestellt. Im Folgenden soll eine Betrachtung bereits empirisch belegter Zusammenhänge der einzelnen Aspekte der Angebots- und Nutzungsebene mit den Lernergebnissen vorgenommen werden. Da der Hauptfokus der vorliegenden Arbeit auf den inhaltsbezogenen Kompetenzen von Schüler:innen (Ergebnisebene) liegt, werden Aspekte der Angebotsebene nur oberflächlich angeschnitten. Die Aspekte der Nutzungsebene werden durch ihre unmittelbare Auswirkung auf die Lernergebnisse als bedeutsamer erachtet und deshalb ausführlicher beschrieben.

### **2.2.1 Einfluss unterrichtlicher Rahmenbedingungen**

Als erster Aspekt der Angebotsebene wird die Kompetenz der Lehrperson betrachtet. Darunter sind als allgemeine Charakteristika nach Seidel (2014) Alter, Geschlecht, Lehrerfahrung und Lehrpersönlichkeit gelistet. In der Literatur finden sich zu den Charakteristika der Lehrperson Erkenntnisse, die belegen, dass Lehrpersonenqualität sowie die Persönlichkeit in Bezug auf Lehrpersonen-Lernenden-Beziehung einen Effekt auf die Leistung der Schüler:innen ausübt (Hattie, 2020, S. 151). Ein weiterer Fokus liegt auf dem Einfluss der professionellen Kompetenzen der Lehrpersonen. Seidel (2014, S. 857 ff) listet unter diesem Punkt die

Aspekte der professionellen Handlungskompetenz von Lehrpersonen nach Baumert und Kunter (2006) auf, für die sie einen Zusammenhang mit Lehrprozessen im Unterricht vermutet. Das Modell von Baumert und Kunter (2006) beschreibt die professionelle Kompetenz durch die Aspekte *Motivationale Orientierungen, selbstregulative Fähigkeiten, Überzeugungen und Werthaltungen* sowie *Professionswissen* (vgl. ebd., S. 482). Das Professionswissen orientiert sich dabei an den von Shulman (1986) definierten Wissensbereichen *allgemeines pädagogisches Wissen, Fachwissen, fachdidaktisches Wissen, Organisationswissen* und *Beratungswissen* (vgl. Baumert & Kunter, 2006). Die wichtigsten Facetten des fachübergreifenden *allgemeinen pädagogischen Wissens* stellen „(1.) konzeptuelles bildungswissenschaftliches Grundwissen, (2.) allgemeindidaktisches Konzeptions- und Planungswissen, (3.) Unterrichtsführung und Orchestrierung von Lerngelegenheiten sowie (4.) fachübergreifende Prinzipien des Diagnostizierens, Prüfens und Bewertens“ (Baumert & Kunter, 2006, S. 485) dar. Für den fachbezogenen Kontext ist das *Fachwissen* und das *fachdidaktische Wissen* relevant. Baumert und Kunter (2006) konstatieren unter Einbezug von empirischen Erkenntnissen, dass fachdidaktisches Handeln durch das Fachwissen der Lehrpersonen beeinflusst wird. Neben dem *professionellen Wissen* stehen die *Überzeugungen* als Komplex, der sich aus „(1.) Wertebindungen, (2.) epistemologischen Überzeugungen, (3.) subjektiven Theorien über Lehren und Lernen und (4.) Zielsystemen für Curriculum und Unterricht“ (Baumert & Kunter, 2006, S. 497) zusammensetzt. Als funktionelle Kompetenzen in Bezug auf berufliches Handeln beinhalten *Motivation* und *Selbstregulation* „(1.) Kontrollüberzeugungen und Selbstwirksamkeitserwartungen, (2.) intrinsische motivationale Orientierung: [Lehrpersonen]enthusiasmus und (3.) Selbstregulation: Engagement und Distanzierungsfähigkeit“ (Baumert & Kunter, 2006, S. 502).

Seidel (2014) nennt in ihrem Modell als zusätzliche Einflussfaktoren Kontexte des *Bildungssystems*, der *Schule* und der *Klasse*. Schrader und Helmke (2008, S. 287) ordnen diesen Faktoren einen indirekten Einfluss zu. Diese Aspekte wurden auch von Hattie (2009) diskutiert. Es kristallisiert sich heraus, dass die Klassengröße neben der Lehrpersonenkompetenz einen Einfluss auf die Lernleistung ausübt (Brühwiler & Blatchford, 2011). Darüber hinaus sind es das Klassenklima und die Beeinflussung durch die Peer-Groups der Lernenden sowie die Häufigkeit des Auftretens von Unterrichtsstörungen, die im Kontext der Klasse den größten Einfluss ausüben (Hattie, 2020, S. 128). In Bezug auf Kontexte der Schule und des Bildungssystems sind die Erstellung angemessener Curricula und die Strategien der Lehrpersonen zur Umsetzung als wichtige Einflussfaktoren zu erachten (Hattie, 2020, S. 190).

## 2.2.2 Einfluss individueller Lernendenmerkmale

Den Einfluss individueller Lernendenmerkmale gliedert Seidel (2014) in die drei Bereiche *Lernumwelt*, *individuelle Voraussetzungen* und *individuelle Lernaktivitäten* (Abbildung 1).

In Bezug auf die Lernumwelt zählt Seidel (2014) den *sozio-ökonomischen Hintergrund*, *Sprache*, *Migration* und *Mediennutzung* auf. Gemäß Ditton und Maaz (2015, S. 240 f) übt die soziale Herkunft in Deutschland im Vergleich zu internationalen Bildungssystemen einen überdurchschnittlichen Einfluss aus. Neben der sozialen Herkunft determiniert Migration und Sprachgebrauch innerhalb der Familie zusätzlich den Lernerfolg (Walter, 2009). Gleichzeitig stellt er fest, dass die Kompetenzniveaus von Lernenden mit zwei im Ausland geborenen Elternteilen deutlich geringer ausgeprägt sind als bei Schüler:innen mit einem migrierten Elternteil oder Eltern ohne Migrationshintergrund (Walter, 2009, S: 165). In Bezug auf die Mediennutzung spezifiziert Seidel (2014) nicht näher, welche Technologien unter dem Begriff subsummiert sind. Hattie (2020, S. 259-273) expliziert in seiner Vergleichsstudie Effekte für eine Reihe von mediengestützten Lernformaten. Die Effektstärken rangieren zwischen  $d = 0.18$  für webbasiertes Lernen über  $d = 0.37$  bei computergestütztem Lernen bis hin zu  $d = 0.52$  für Lernvideos (Hattie, 2020, S. 259 ff).

Unter den individuellen Voraussetzungen der Lernenden subsummiert Seidel (2014) *Intelligenz*, *Vorwissen*, *Motivation*, *Emotionen* sowie *selbstbezogene Vorstellungen*. Hattie (2009, S. 41) stellt in seiner Metastudie fest, dass sowohl Intelligenz als auch Vorwissen einen großen Effekt auf die Lernergebnisse ausüben. Für den Zusammenhang von Intelligenz und Lernergebnissen gibt er eine Effektstärke von  $d = 1.19$  an (Hattie, 2009, S. 41). Für den Einfluss von Vorwissen auf die Lernergebnisse ergibt sich eine Effektstärke von  $d = 0.67$  (ebd., S. 41 f). Im Bezugsfeld der beruflichen Bildung wurde der Einfluss fachspezifischen Vorwissens und der kognitiven Grundfähigkeit in Form der Betrachtung als fluide und kristalline Intelligenz (nach Cattell, 1963) auf Fachwissen bereits in unterschiedlichen technischen Kontexten untersucht und nachgewiesen (z. B. Hedrich, 2021; Nickolaus, Abele, Gschwendtner, Nitzschke & Greiff, 2012; Nickolaus, Abele & Albus, 2015; Wyrwal, 2020). Gleichzeitig weisen neuere Studien darauf hin, dass nach wie vor Forschungsdesiderate bezüglich des Effekts von Vorwissen auf die Lernergebnisse bestehen (z. B. Simonsmeier, Flaig, Deiglmayr, Schalk & Schneider, 2022).

Die von Seidel (2014) aufgelisteten selbstbezogenen Vorstellungen sind als vielschichtiges Konstrukt zu begreifen. Als selbstbezogene Vorstellungen „wird [...]

die Gesamtheit selbstbezogener Kognitionen bzw. kognitiver Repräsentationen eigener Fähigkeiten verstanden“ (Förster, 2016, S. 23). Diesbezüglich stellt Hattie (2020, S. 52 f) mit  $d = 1.44$  den stärksten Effekt seiner Metastudie für die „Selbsteinschätzung des eigenen Leistungsniveaus“ (ebd., S. 52) der Lernenden fest. Einen mittleren Effekt ( $d = 0.43$ ) findet er für den Einfluss des Selbstkonzepts (ebd., S. 55 ff). Dessen Ausprägungen fasst er als „Primärmotive [auf], die dann zur Aktivierung verschiedener situationsspezifischer Orientierungen des Selbst, wie z. B. Selbstwirksamkeitsüberzeugung, Ängste, Leistung oder Lernorientierung führen“ (Hattie, 2020, S. 55).

In Bezug auf die Motivation belegt die Vergleichsstudie von Hattie (2020, S. 57) einen mittelgroßen Effekt von  $d = 0.46$ . Er führt eine Arbeit von Dörnyei (2001) an, die aufzeigt, dass Lernende besonders motiviert sind, wenn sie „kompetent sind, über ausreichend Autonomie verfügen, sich lohnende Ziele setzen, Feedback bekommen und Bestätigung von anderen erhalten“ (Hattie, 2020, S. 58). Seidel (2014) ergänzt diese Auflistung um die soziale Einbindung in die Lerngemeinschaft als weiteres Grundbedürfnis. Ein höheres Maß an intrinsischer Motivation resultiert aus der Überzeugung, das eigene Lernen beeinflussen zu können und führt dazu, dass intrinsisch motivierte Personen bessere akademische Leistungen erzielen (vgl. ebd., S. 58). Dabei scheint es in Bezug auf den Lernerfolg außerdem wichtig zu sein, dass Lernende die Steigerung ihrer Lernleistung als positives Resultat aus einer Kombination ihrer Anstrengung und des persönlichen Interesses wahrnehmen (vgl. Hattie, 2020, S. 58; Krapp & Prenzel, 1992, 2011). Gleichzeitig sollte das Lernen, angelehnt an Deci und Ryan (2002), „mit positiven Emotionen und einem vertieften Durchdringen von Lerninhalten verbunden“ (Seidel, 2014, S. 856) sein.

Die individuellen Lernaktivitäten gliedert Seidel (2014) in *äußere* und *innere Lernaktivitäten*. In Anlehnung an Mandl und Friedrich (1992) fasst Seidel (2014, S. 859) zusammen, dass die inneren Lernaktivitäten gegenüber den äußeren einen entscheidenderen Einfluss ausüben. Unter inneren Lernaktivitäten erläutert sie, dass die Lernenden u. a. fähig sein müssen, behandelte Inhalte nachzuvollziehen und mit bestehendem Vorwissen zu verknüpfen sowie das bestehende Vorwissen strukturell durch das hinzukommende Wissen zu erweitern und neu zu organisieren (vgl. Seidel, 2014, S. 859).

In der vorliegenden Arbeit liegt der inhaltliche Schwerpunkt auf der Ergebnisebene des Angebots-Nutzungs-Modells nach Seidel (2014). In den folgenden Abschnitten erfolgt daher eine eingehendere Auseinandersetzung mit dem Kompetenzbegriff, den Bildungsstandards und der Kompetenzdiagnostik.

### **2.3 Kompetenzbegriff und Bildungsstandards zur Beschreibung von Lernergebnissen**

In der Arbeit wird von der Kompetenzdefinition nach Weinert (2001) ausgegangen. Er definiert Kompetenzen als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, 2001, S. 27). Seine Definition impliziert, dass Kompetenzen - in Abgrenzung zur Intelligenz als relativ stabiles Konstrukt - erworben werden können (vgl. Hartig & Klieme, 2006, S. 130). Es ist von einer Kontextabhängigkeit auszugehen, sodass einzelne Individuen zunächst nicht zwangsläufig über Kompetenzen in einem Bereich verfügen, jedoch in der Lage sind, die notwendigen Fähigkeiten und Fertigkeiten zu erlernen (Hartig & Klieme, 2006, S. 129 ff). Auf den Kompetenzerwerb wirken z. B. im Rahmen schulischen Lernens die Einflussfaktoren der unterschiedlichen Ebenen des Angebots-Nutzungs-Modells, wodurch zu erwarten ist, dass sich messbare Unterschiede bei den durch die Schüler:innen erreichten Kompetenzausprägungen zeigen. Anhand von Aufgabenstellungen mit unterschiedlichen Anforderungen lässt sich demnach z. B. eine Einordnung in Niveaustufen vornehmen, die die unterschiedliche kognitive Leistung der Lernenden repräsentieren (Köller, 2011) und über das Lernergebnis indirekt die individuellen Merkmale der Lernenden widerspiegeln (vgl. Definition nach Weinert, 2001). Dieses Kompetenzverständnis griff ebenfalls die KMK in der 2003 bzw. 2004 beschlossenen Einführung der Bildungsstandards für das deutsche Bildungssystem auf. Die zu erreichenden Bildungsziele sind „als Can-do-Statements (Kann-Beschreibungen) formuliert“ (Köller, 2011, S. 181) und beziehen sich auf die zuvor dargestellte Kompetenzdefinition nach Weinert (2001) unter der Annahme einer Aneignung von Fähigkeiten und Fertigkeiten in einem domänenspezifischen Kontext (Fleischer et al., 2013).

### **2.4 Beschreibung der Bildungsstandards im Fach NwT<sup>3</sup>**

Eine Gemeinsamkeit der Unterrichtsfächer im natur- und technikwissenschaftlichen Bildungsbereich ist die Differenzierung der Bildungsstandards in prozess- und inhaltsbezogene Kompetenzen. Im Fach NwT liegen Bildungsstandards für

---

<sup>3</sup> Teile dieses Kapitels und seiner Unterkapitel 2.4.1 und 2.4.2 entstammen in identischer oder leicht abgewandelter Form einer bereits erschienenen Publikation: Brändle (2023).

das Profilfach in der Mittelstufe (MKJS, 2016) sowie für das Basis- und Leistungsfach in einer Schulversuchsfassung (Autorengruppe Bildungsplan NwT, 2019a, 2019b) vor und werden im Folgenden näher vorgestellt.

### 2.4.1 Bildungsstandards für die Mittelstufe

Die Kompetenzbereiche des Mittelstufencurriculums unterscheiden sich in inhalts- und prozessbezogenen Kompetenzen (Abbildung 2).



Eigene Abbildung, basierend auf MKJS (2016) und Zinn(2018c, S. 233 f).

**Abbildung 2:** Prozess- und inhaltsbezogene Kompetenzen im Bildungsplan des Profilfachs NwT in der Mittelstufe (entnommen aus Brändle, 2023).



Es wurde bereits in anderen Veröffentlichungen eine Betrachtung der Kompetenzbereiche des Mittelstufencurriculums vorgenommen, weshalb hier nicht näher auf einzelne Kompetenzfacetten eingegangen wird. Abbildung 2 orientiert sich an der Zusammenfassung von Zinn (2018c, S. 233 f). Eine detaillierte Ausführung der Kompetenzbereiche ist im Bildungsplan 2016 (MKJS, 2016) einsehbar.

### 2.4.2 Bildungsstandards für die Kursstufe (Basis- und Leistungsfach)

Der Bildungsplan der Kursstufe (Basis- und Leistungsfach) wird an dieser Stelle ausführlicher betrachtet, da hierzu eine Neuentwicklung von Testinstrumenten erfolgt. Einen Überblick über die einzelnen Kompetenzbereiche liefert Abbildung 3. Die prozessbezogenen Kompetenzen sind identisch mit dem Curriculum der Mittelstufe und im Kontext dieser Arbeit zweitrangig. Der Fokus liegt hauptsächlich auf den inhaltsbezogenen Kompetenzen im Leistungsfach.



Eigene Abbildung, basierend auf Autorengruppe Bildungsplan NwT (2019a; 2019b). \* Nicht im Bildungsplan des Basisfachs. \*\* Im Bildungsplan des Basisfachs subsummiert unter „Messen, Steuern, Regeln“.

**Abbildung 3:** Prozess- inhaltsbezogene Kompetenzen im Bildungsplan des Basis- und Leistungsfachs NwT in der Kursstufe (entnommen aus Brändle, 2023).

Analog zur Mittelstufe gliedern sich die inhaltsbezogenen Kompetenzen in vier Kompetenzbereiche: *Denk- und Arbeitsweisen in Naturwissenschaft und Technik; Energie und Antrieb; Technische Mechanik und Produktentwicklung* sowie *Elektro- und Informationstechnik* (Autorengruppe Bildungsplan NwT, 2019a, 2019b). Die aufeinander abgestimmten prozess- und inhaltsbezogenen Kompetenzen in Mittel- und Kursstufe weisen auf einen spiralcurricularen Charakter des Fachs hin, der in den *Leitgedanken zum Kompetenzerwerb* erläutert wird (Autorengruppe Bildungsplan NwT, 2019b, S. 7f). Nachfolgend werden die inhaltsbezogenen Kompetenzteilbereiche des Leistungsfachs (angelehnt an Autorengruppe Bildungsplan NwT, 2019b) kurz umrissen:

- **Systeme und Prozesse**  
Dieser Kompetenzteilbereich wird vernetzend „an geeigneten Stellen des Unterrichts in Verbindung mit den inhaltsbezogenen Kompetenzen der anderen Kompetenzbereiche vertieft“ (ebd., S. 14). Zentraler Aspekt ist der Systemgedanke, in dessen Kontext die Schüler:innen in die Lage versetzt werden, Gesamtsysteme zu analysieren, in Teilsysteme zu untergliedern und deren Wechselwirkungen und die dabei ablaufenden Prozesse zu erkennen und zu erläutern sowie dieses Wissen im Rahmen von Entwicklungs- und Optimierungsvorhaben einzusetzen (ebd., S. 14).
- **Technische und wissenschaftliche Handlungskompetenzen**  
Zur Umsetzung der Forschungs- und Produktentwicklungsaufgaben im Rahmen der Kompetenzteilbereiche des Leistungsfachs benötigen die Lernenden „[e]in eigenständiges wissenschaftspropädeutisch orientiertes technisches Handeln“ (ebd., S. 15). Darunter werden grundlegende Kompetenzen zur Untersuchung von wissenschaftlichen Fragestellungen und damit einhergehend Gütekriterien wissenschaftlicher Studien, grundlegendes Projektmanagement sowie technische Kommunikationsmethoden subsummiert (ebd., S. 15).
- **Technikfolgenabschätzung**  
Technologie beeinflusst auf unterschiedliche Art und Weise Gesellschaft, Wirtschaft und Umwelt. Es ist daher wichtig, dass Lernende „[d]ie Folgen technischer Innovationen [...] nach vielfältigen Kriterien [...] und im Kontext eines ethischen Handelns begründet gegeneinander abwägen [können], um sich selbst auch als technikemanzipierte und technikmündige Bürgerinnen und Bürger in entsprechende Meinungs- und Entscheidungsprozesse einbringen zu können“ (ebd., S. 16). Dabei ist es entscheidend, dass die Schüler:innen als Methode die ethische Fallanalyse kennenlernen

und diese im technologischen Bezugsfeld zur Konfliktanalyse und Entscheidungsfindung anwenden (ebd., S. 16).

- **Energieversorgung**

Im Kontext der „[...] Notwendigkeit einer nachhaltigen Energieversorgung erkennen die Schülerinnen und Schüler die Herausforderungen, die sich beim Zusammenwirken von Bereitstellung, Speicherung und Transport elektrischer Energie in Versorgungsnetzen ergeben“ (ebd., S. 17). Die Lernenden erlangen dazu in diesem Kompetenzteilbereich grundlegendes Wissen über technische Systeme zur Energiewandlung und -speicherung sowie die dazu benötigten Komponenten mit dem Fokus auf erneuerbare Energien und die Energiewende (ebd., S. 17).

- **Elektrische Antriebstechnik**

Als „zentrale[...] Baustein[e] der Elektromobilität in einem nachhaltigen und klimaschonenden Verkehrssystem“ (ebd., S. 18) lernen die Schülerinnen und Schüler, welche Komponenten sie für den Aufbau eines elektrischen Antriebssystems benötigen, wie diese zu charakterisieren sind und wie sie aufeinander abgestimmt werden müssen (ebd., S. 18).

- **Technische Mechanik**

Für die „[...] Lösung statischer Problemstellungen sowie zur Dimensionierung technischer Konstruktionen“ (ebd., S. 19) erlangen die Schüler:innen grundlegende Kompetenzen im Bereich der technischen Mechanik sowie der Festigkeitslehre und sind in der Lage, dieses Wissen im Konstruktionsprozess zur Entwicklung und Optimierung von Artefakten anzuwenden (ebd., S. 19).

- **Produktentwicklung**

Bei Projektumsetzungen im Leistungsfach „bearbeiten die Schülerinnen und Schüler zunehmend komplexe Entwicklungsaufgaben, in welche sie die Kenntnisse aus verschiedenen Naturwissenschafts- und Technikdisziplinen einbringen“ (ebd., S. 20). Dazu erlangen sie fortgeschrittene Kompetenzen beim technischen Zeichnen als Kommunikationsmittel und müssen in der Lage sein, computergestützte Konstruktions- und Fertigungsmethoden (CAD-CAM-Kopplung) anzuwenden (ebd., S. 20).

- **Grundlagen der Elektronik**

Anknüpfend an die Mittelstufe werden die Grundlagen der Elektronik unterrichtet, da „[d]ie Energietechnik und die Messtechnik in Klassenstufe 11/12 eine Erweiterung und Vertiefung der Grundlagen [erfordern]“ (ebd., S. 21). Dabei lernen die Schüler:innen grundlegende Techniken für Platinenlayout und -fertigung kennen sowie die Untersuchung und den Einsatz dafür notwendiger elektronischer Komponenten (ebd., S. 21).

- **Aufnahme und Verarbeitung von Signalen in der Messtechnik**  
Im naturwissenschaftlich-technischen Unterricht hat die Erfassung und Auswertung von Messwerten einen zentralen Stellenwert. „Auf dem Weg von der Erfassung von Messgrößen über die Verarbeitung und Darstellung der Signale erwerben die Schülerinnen und Schüler ein vertieftes Verständnis für die Funktionsweise von Messgeräten“ (ebd., S. 22). Dazu müssen sie in der Lage sein, vollwertige Messvorrichtungen zu entwickeln und dabei v. a. elektrische Signalaufnahme und -verstärkung umzusetzen (ebd., S. 22).
- **Regelungstechnik**  
Die Schüler:innen „[...] erkennen die Bedeutung geregelter Vorgänge für natürliche und technische Systeme“ (ebd., S. 23). Dazu erlernen sie die Funktionsweise unterschiedlicher Regelstrecken und die Charakteristik ihrer jeweiligen Kenngrößen. Darüber hinaus sollen die Lernenden eigene Regelstrecken unter Einbezug des Mikrocontrollers entwickeln und aufbauen (ebd., S. 23).
- **Datenkommunikation**  
Im Rahmen der Datenkommunikation werden Datenübertragungsprinzipien und die Funktionsweise elektronischer Komponenten in Kommunikationssystemen vermittelt. „Die Schülerinnen und Schüler lernen [dabei] allgemeine Strukturen der Kommunikation kennen“ (ebd., S. 24). In einer lebensweltlichen Kontextualisierung werden Grundprinzipien der Programmierung und „[...] Datensicherheit, zum Beispiel im Kontext von *Internet of Things*“ (ebd., S. 24) vermittelt.

Laut den Bildungsplänen der Kursstufe unterscheiden sich Basis- und Leistungsfach inhaltlich hinsichtlich der Kompetenzteilbereiche *Messen, Steuern, Regeln* sowie *Grundlagen der Elektronik* (Autorengruppe Bildungsplan NwT, 2019a, 2019b). Dem Leistungsfach wird ein höherer Mathematisierungsgrad und eine tiefere, inhaltlich breitere Ausgestaltung der Kompetenzteilbereiche zugesprochen (vgl. Autorengruppe Bildungsplan NwT, 2019a, S. 7).

Zusammengefasst ermöglichen die Kompetenzbereiche des Bildungsplans eine Orientierung an technikwissenschaftlichen Disziplinen und Fragestellungen; auch mit dem Ziel, Schüler:innen eine wissenschaftspropädeutische technische Allgemeinbildung<sup>4</sup> zu vermitteln. Für die fachdidaktische Forschung und die

---

<sup>4</sup> Hahn (2013, S. 162) definiert: „Ein wissenschaftspropädeutischer Unterricht zielt also *mindestens* auf eine exemplarische Einführung in den wissenschaftlichen Sprachgebrauch, wissenschaftlicher Denk- und Arbeitsweisen sowie auf eine wissenschaftstheoretische Einordnung grundlegender Erkenntnismethoden.“ (Hervorhebungen im Original)

Lehrpersonenbildung ist es u. a. von Interesse, mit welchen Kenntnissen und Fähigkeiten die SuS in die Kursstufe einmünden und welche Kompetenzen die Schüler:innen im Zuge der zweijährigen Kursstufe erwerben.

## 2.5 Kompetenzdiagnostik

Nachdem die Messung von Kompetenzen in Form von Bildungsmonitorings mit der Einführung der Bildungsstandards in den Fokus der Bildungsforschung rückte, entwickelten sich zwei zentrale Modellansätze zur Beschreibung von Kompetenzen: Kompetenzstrukturmodelle und Kompetenzniveaumodelle. Kompetenzstrukturmodelle befassen sich mit der Ausdifferenzierung von Kompetenzdimensionen (Fleischer et al., 2013, S. 8). Demgegenüber fokussieren Kompetenzniveaumodelle Unterschiede in der Stärke der Ausprägung einer Kompetenz (ebd., S. 8). Insgesamt setzt ein Bildungsmonitoring zunächst voraus, „dass ein Kompetenzbegriff definiert und operationalisiert wird, der anhand von umfassend entwickelten Testaufgaben als zuverlässige[r] Leistungs[indikator] gemessen werden kann“ (Sälzer, 2016, S. 33). Eine solche Operationalisierung wird u. a. in nationalen und internationalen Schulleistungsstudien umgesetzt.

Zur Generierung eines Überblicks werden im Nachfolgenden die relevantesten Schulleistungsstudien, an denen auch Lernende in Deutschland teilnehmen, kurz beschrieben und deren Definition des Kompetenzbegriffs vorgestellt (vgl. Sälzer, 2016). Sälzer (2016) führt insgesamt sechs Studien auf, von denen sich drei im nationalen (IQB-Bildungstrend, VERA, NEPS) und drei im internationalen (PISA, TIMSS, IGLU) Raum verorten lassen. VERA wird in diesem Kontext ausgeklammert und nicht näher erläutert, da es sich durch den Rückmeldecharakter zum Lernstand der Schüler:innen an die Lehrpersonen um eine Lernstandserhebung und aufgrund einer fehlenden Definition des Kompetenzbegriffs nicht um ein Bildungsmonitoring handelt (Sälzer, 2016). Die Evaluation des NEPS orientiert sich an einer längsschnittlichen Erfassung der Kompetenzentwicklung über den Lebensverlauf und ist in diesem Kontext ebenfalls weniger relevant (ebd., S. 37 ff). IGLU (Internationale Grundschul-Lese-Untersuchung) bezieht sich auf Lesekompetenz in der Primarstufe und liefert daher im Rahmen naturwissenschaftlicher Kompetenzdiagnostik in der Sekundarstufe II kaum Anknüpfungspunkte (ebd., S. 20 ff).

Während sich die internationalen Studien PISA, TIMSS und IGLU dem Kompetenzbegriff in Form einer Literalität (*engl.* Literacy) (OECD, 2013; Sälzer & Prenzel, 2013) „als Konzept einer funktionalen Grundbildung im Sinne einer Allgemeinbildung“ (Sälzer, 2016, S. 10) mit einem Fokus auf problembasierter Anwendung

nähern, orientiert sich das national aufgelegte Programm des IQB am Kompetenzbegriff der Bildungsstandards (vgl. Kapitel 2.3). Die Kompetenzauffassung der Studien stimmt auf einer übergeordneten Ebene mit der Definition von Weinert (2001) (s. Kapitel 2.3) überein. Welche Bereiche speziell die Kompetenz in den naturwissenschaftlich-technischen Fächern auszeichnen, lohnt einer eingehenderen Betrachtung und wird im nachfolgenden Kapitel 2.6 dargestellt.

Alle vier Studien (PISA, TIMSS, IGLU und IQB) vereint die Operationalisierung von Kompetenz als mehrdimensionales Konstrukt und der Einsatz von standardisierten Testinstrumenten zu dessen Erfassung (z. B. Sälzer, 2016). Als national angelegtes Bildungsmonitoring mit Fokus auf den allgemeinbildenden Bereich und konkretem Bezug zum deutschen Bildungssystem liefert die Methodik des IQB-Bildungstrends für die Kompetenzmessung bei Schüler:innen im Leistungsfach NwT Anhaltspunkte zur Umsetzung der Evaluationsmaßnahmen.

Die Kompetenzdiagnostik des IQB-Bildungstrends erfolgt mit einer Mischung aus offenen, halboffenen und geschlossenen Antwortformaten. Die Aufgaben wurden u. a. in Zusammenarbeit mit Lehrpersonen als Bildungsexpert:innen in mehreren Iterationszyklen entwickelt und erprobt (Mahler, Schipolowski & Weirich, 2019, S. 99). Eine Niveaudifferenzierung wird durch Aufgaben unterschiedlicher Schwierigkeit erreicht (ebd., S. 101). In den Naturwissenschaften erfassen die Testinstrumente die von Kauertz, Fischer, Mayer, Sumfleth und Walpuski (2010) als empirisch unterscheidbar dargestellten Kompetenzbereiche (ausführlicher in Kapitel 2.7) *Fachwissen* und *Erkenntnisgewinnung*. Die Erfassung des Fachwissens orientiert sich dabei an den inhaltsbezogenen Kompetenzteilbereichen der Bildungspläne für die Naturwissenschaften (Mahler et al., 2019, S. 100). Die Erfassung des prozessbezogenen Kompetenzbereichs *Erkenntnisgewinnung* ermöglichen die Aufgaben der Testinstrumente über die Abbildung der Kompetenzteilbereiche zur naturwissenschaftlichen Modellbildung, zu naturwissenschaftlichen Untersuchungen und zu deren wissenschaftstheoretischen Reflexion (vgl. ebd., S. 100). Die Testhefte setzen sich in ihrer finalen Version aus sechs Aufgabenblöcken zu jeweils 20 Minuten Bearbeitungszeit zusammen (ebd., S. 106 f). Ein einzelner Aufgabenblock besteht dabei aus Items, die einen Kompetenzteilbereich des Bildungsplans für ein bestimmtes Fach abbilden (vgl. ebd., S. 106).

Insgesamt fokussieren alle hier dargestellten Schulleistungsstudien eine Beschreibung von Kompetenzniveaus bei Lernenden. Zur Generierung von Aussagen über die Ausprägung von Kompetenzen streben Kompetenzniveaumodelle

die Segmentierung des kontinuierlichen Gesamtkonstrukts in eine diskrete Einteilung an und ermöglichen dadurch eine Beschreibung der individuellen Kompetenzausprägung von Lernenden (Hartig & Klieme, 2006; Wilson, 2004). Dazu werden sog. *Standard-Setting-Verfahren* eingesetzt, die im Weiteren ausführlicher beschrieben sind.

Die für eine Ergebnisinterpretation erforderliche Einteilung der Niveaustufen kann anhand unterschiedlicher test- oder personenzentrierter *Standard-Setting-Verfahren* umgesetzt werden (Pant, Tiffin-Richards & Köller, 2010). Eine Unterscheidung nach qualitativen und quantitativen Verfahren sowie der beiden möglichen Einteilungsvarianten nach Personenfähigkeit (personenbezogen) oder Aufgabenschwierigkeit (testbezogen) erscheint notwendig, um eine begründete Auswahl eines geeigneten Verfahrens treffen zu können. Zwei testzentrierte Ansätze finden häufigere Anwendung (vgl. Pant et al., 2010, S. 176). Es handelt sich dabei um die *Bookmark-Methode* und das *Angoff-Verfahren* (ebd., S. 176). Die Bookmark-Methode nutzt die berechneten empirischen Itemschwierigkeiten (zugänglich z. B. durch IRT-Skalierung) und sortiert die Aufgaben nach aufsteigendem Schwierigkeitsgrad in einer Mappe (Zieky, 2012, S. 24). Im weiteren Prozess erfolgt die Einteilung der Niveaustufen anhand eines Expert:innengremiums (ebd., S. 24). Die Expert:innen setzen an zwei Punkten Lesezeichen in der Mappe und legen damit fest, wo die Grenzen zwischen mittleren und einfachen Aufgaben (Lösungswahrscheinlichkeit  $> 0.67$ ) sowie mittleren und schweren Aufgaben (Lösungswahrscheinlichkeit  $< 0.67$ ) liegen (ebd., S. 24). Beim Angoff-Verfahren sollen Expert:innen einschätzen, mit welcher Wahrscheinlichkeit eine imaginäre Person aus ihrer Sicht das vorgelegte Item noch lösen kann (ebd., S. 21). Die aggregierten Bewertungen der Expert:innen ergeben die Niveauschnitte (Cut-Scores) (Pant et al., 2010).

Den beiden testzentrierten Ansätzen steht die personenzentrierte Kontrastgruppenmethode (*Contrasting-groups-Methode*) gegenüber (Pant et al., 2010, S. 176). Sie orientiert sich an den finalen Testergebnissen der Schüler:innen. Die bewertenden Bildungsexpert:innen müssen dahingehend ausgewählt werden, dass sie die Leistungen der Testteilnehmer:innen gut einschätzen können (ebd., S. 176). Sie unterteilen anhand repräsentativer Sub-Stichproben verschiedener Ergebnisbereiche (es wird nicht die gesamte Stichprobe betrachtet) die Testteilnehmer:innen in Gruppen von Personen mit unterschiedlichem Fähigkeitsniveau (Zieky, 2012, S. 22).

Da es sich bei den bisher vorgestellten Methoden um rein qualitative Verfahren handelt, stellt demgemäß „das Setzen von Cut-Scores durch Expert[:innen]urteile [...] *per se* einen bewertenden Vorgang dar“ (Pant et al., 2010, S. 177, Hervorhebungen im Original). Pant et al. (2010) erläutern weiter:

„Die Festlegung der Cut-Scores stellt [...] ein besonders kritisches Verbindungsglied zwischen den evidenzbezogenen, empirisch gut untersuchbaren Aspekten des Gesamtsystems und den konsequenzbezogenen, eher normativen und praxisrelevanten Aspekten dar“ (Pant et al., 2010, S. 178).

Pant und Kolleg:innen erkennen im Standard-Setting ein kritisches Moment in Bezug auf die abzuleitenden Implikationen für die Praxis. Eine datenbezogene quantitative Setzung bietet dahingehend Vorteile, indem sie (1.) die Gesamtheit der Daten einer Stichprobe betrachten kann (nicht nur eine begrenzte Auswahl aufgrund forschungsökonomischer Gesichtspunkte) und (2.) unabhängig von subjektiven Einschätzungen agiert. Zur objektiveren Festlegung von Niveaustufen eignen sich daher u. a. explorative hierarchisierende oder partitionierende Verfahren wie z. B. *Cluster-* oder *Latent-Class-Analysen (LCA)* (z. B. Sireci, Robin & Patelis, 1999; Templin & Jiao, 2012), die in älteren und aktuelleren Studien eingesetzt werden (z. B. Hess, Subhiyah & Giordano, 2007; Schlüter, 2015; Thomas, Cassady & Finch, 2018; Tseng, Chiou & Sung, 2017; Violato, Marini & Lee, 2003). Hess et al. (2007) und Tseng et al. (2017) erzielen in ihrer Gegenüberstellung der Clusteranalyse und der Angoff-Methode vergleichbare Ergebnisse bei der Segmentierung der Stichprobe. Schlüter (2015) nutzt die Clusteranalyse erfolgreich zur Identifizierung von Anforderungsdimensionen der von ihr entwickelten Aufgaben. Thomas et al. (2018, S. 503) erhalten mit den beiden Verfahren Clusteranalyse und LCA ebenfalls einheitliche Ergebnisse, die ihre Stichprobe in drei Profile einteilen. Violato et al. (2003) bestätigen, dass die Clusteranalyse vergleichbare Ergebnisse bei der Einstufung von Testteilnehmern hervorbringt. Sie untersuchen dazu die Konvergenz der Verfahren von Nedelsky und Ebel als qualitative Ansätze mit dem quantitativen Ansatz der Clusteranalyse (Violato et al., 2003, S. 60).

Zusammenfassend lässt sich konstatieren, dass die Clusteranalyse als quantitatives und rein objektives Verfahren im Rahmen der Kompetenzdiagnostik eine Möglichkeit bietet, Testteilnehmer:innen in unterschiedliche gegeneinander abgrenzbare, aber in sich homogene Niveaugruppen zu differenzieren. Sie kann an den vorliegenden Daten variabel eingesetzt werden und ermöglicht dadurch die explorative Gegenüberstellung verschiedener Gruppierungsansätze. Gleichzeitig ist festzustellen, dass die Kompetenzdiagnostik und die Setzung von Schwellen-

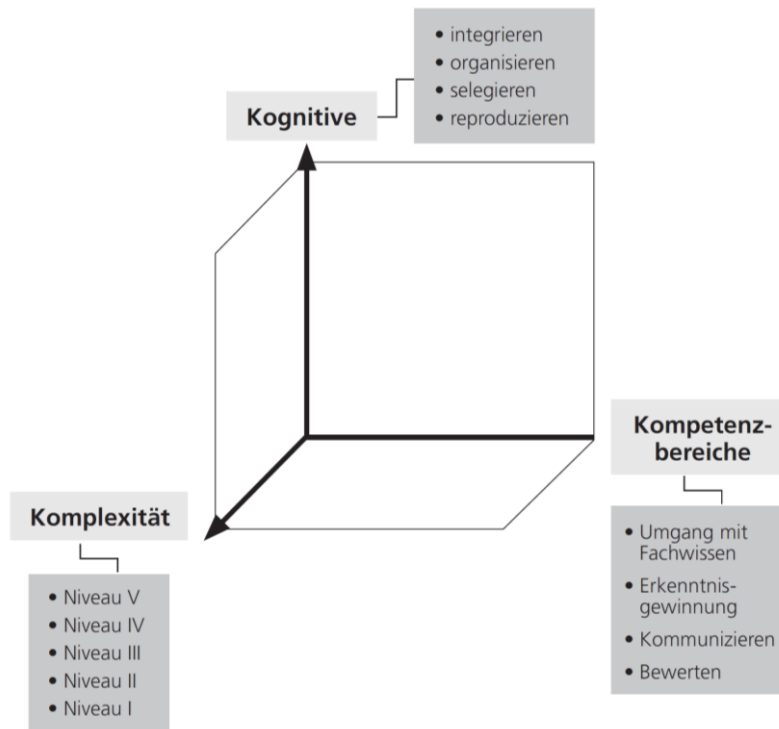


werten (Cut-Scores) für die Festlegung von Kompetenzniveaus neben psychometrischen Gütekriterien auch unter Berücksichtigung fachdidaktischer Aspekte erfolgen muss, um der Domänenspezifität gerecht zu werden. Das nachfolgende Kapitel befasst sich aus diesem Grund mit Kompetenzmodellen in naturwissenschaftlichen Fächern.

## 2.6 Kompetenzmodelle in naturwissenschaftlichen Fächern

In den Bildungsstandards der naturwissenschaftlichen Fächer findet sich eine normative Unterscheidung in die vier Kompetenzbereiche *Fachwissen*, *Erkenntnisgewinnung*, *Kommunikation* und *Bewerten* (vgl. Parchmann, 2010, S. 138). Ziel der fächerübergreifend geltenden Standards in den naturwissenschaftlichen Fächern ist es „vielfältige Anknüpfungspunkte, die sowohl zur vertikalen Vernetzung von Inhalten innerhalb eines naturwissenschaftlichen Fachs und damit zum Aufbau einer kohärenten Wissensstruktur beitragen als auch zur horizontalen Vernetzung zwischen den naturwissenschaftlichen Fächer[n] und darüber hinaus“ (Pant et al., 2013, S. 38) zu ermöglichen. Aufgrund der Domänenspezifität von Kompetenzen (Klieme et al., 2003) muss für die Förderung und Diagnostik deren kontextuelle Orientierung berücksichtigt werden.

Ein bedeutsamer Aspekt bei der Förderung und Überprüfung von Kompetenzen im schulischen Rahmen sind in allen (naturwissenschaftlichen) Fächern u. a. Lern- und Leistungsaufgaben (z. B. Kleinknecht, Bohl, Maier & Metz, 2013). Die Anforderungen für Aufgaben werden in den Bildungsstandards mittels Operatoren in drei Bereiche mit unterschiedlichem kognitiven Potenzial untergliedert (vgl. Kleinknecht et al., 2013, S. 15). Laut Neumann, Kauertz, Lau, Notarp und Fischer (2007) reicht es jedoch nicht aus, die in den Bildungsstandards festgelegten Niveaustufen als Vorgabe für die Schwierigkeit einer Aufgabe heranzuziehen. Mit dieser Problematik setzen sich Kauertz (2008), Kauertz et al. (2010) und Bernholt (2010) auseinander. Als Ausgangspunkt für die vorliegende Arbeit wird aufgrund seiner empirischen Betrachtung und Fundierung das Modell nach Kauertz et al. (2010) ausgewählt. Das Modell (Abbildung 4) unterscheidet das kognitive Aufgabenpotenzial anhand von vier Niveaus und gliedert sich nach Kauertz et al. (2010) in die Bereiche *Reproduktion*, *Selektion*, *Organisation* und *Integration* (s. Abbildung 4). Außerdem führt er fünf Komplexitätsstufen für Aufgaben an. Sie „basieren [...] auf den zwei Prinzipien *Anzahl an Elementen* und (funktionale) *Verknüpfung zwischen Elementen*“ (Kauertz et al., 2010, 142f). Diese Komplexitätsstufen beeinflussen genauso wie die kognitive Anforderung die Aufgabenschwierigkeit mit geringen Effektstärken (Kauertz et al., 2010, S. 149).



**Abbildung 4:** Modell naturwissenschaftlicher Kompetenz nach Kauertz et al. (Unverändert entnommen aus Kauertz et al., 2010, S. 145).

Da das Fach NwT über die naturwissenschaftlichen Kompetenzbereiche hinaus einen technischen Aspekt fördert, schließt sich im nachfolgenden Abschnitt eine Überlegung zur Anpassung des naturwissenschaftlichen Kompetenzmodells an.

## 2.7 Erfassung naturwissenschaftlich-technischer Kompetenz

Mangels eines auf das Fach NwT zugeschnittenen Kompetenzmodells, das die Ableitung einer spezifischen Diagnostik in Bezug auf interdisziplinäre naturwissenschaftlich-*technische* Kompetenzen ermöglicht, wurde als Grundlage das Modell von Kauertz et al. (2010) für naturwissenschaftliche Fächer (Kapitel 2.6) betrachtet. Für die Integration der technischen Literalität erfolgt zunächst eine Klärung des Begriffsverständnisses.

Die ITEEA (2020) definiert acht zentrale und übergeordnete Standards technischer Literalität (wörtlich aus ITEEA, 2020, S. 7):

1. Nature and Characteristics of Technology and Engineering (ebd., S. 7).
2. Core Concepts of Technology and Engineering (ebd., S. 7).
3. Integration of Knowledge, Technologies, and Practices (ebd., S. 7).
4. Impacts of Technology (ebd., S. 7).
5. Influence of Society on Technological Development (ebd., S. 7).
6. History of Technology (ebd., S. 7).
7. Design in Technology and Engineering Education (ebd., S. 7).
8. Applying, Maintaining, and Assessing Technological Products and Systems (ebd., S. 7).

Dieses Verständnis der technischen Literalität wird in den Kompetenzteilbereichen der Bildungspläne für Profil-, Basis- und Leistungsfach (Autorengruppe Bildungsplan NwT, 2019a, 2019b; MKJS, 2016) durch die inhalts- und prozessbezogenen Kompetenzen über verschiedene technikwissenschaftliche Kontexte adressiert. Zinn (2018b) stellt diesbezüglich für das Bezugsfeld der Allgemeinbildung fest:

„Der technische Unterricht muss eine reflexive Auseinandersetzung mit Themen der allgemeinen Technikwissenschaft sowie anwendungsorientierter Fragestellungen der Technik ermöglichen. Eine technische Allgemeinbildung muss daher in der Lage sein, zum einen ein grundlegendes Wissensfundament zur Technik zu begründen, um gesellschaftliche Entscheidungen, Entwicklungen und den Einsatz von Technik im Hinblick auf die intendierten und nicht intendierten Folgen und Unwägbarkeiten wissenschaftsbasiert zu bewerten, zum anderen muss sie es Jugendlichen ermöglichen, basale technische Kompetenzen zu erwerben, um angemessen mit technischen Artefakten in privaten, gesellschaftlichen und beruflichen Situationen sach- und fachgerecht umzugehen“ (Zinn, 2014, S. 120).

Gemäß den Bildungsstandards der MINT-Bildung in Baden-Württemberg unterscheidet das Fach NwT ebenfalls nach den in den Naturwissenschaften etablierten Kompetenzbereichen *Fachwissen* (inhaltsbezogene Kompetenzen), *Erkenntnisgewinnung*, *Bewerten*, *Kommunikation* (prozessbezogene Kompetenzen) (MKJS, 2016). Zusätzlich beinhalten die prozessbezogenen Kompetenzen einen weiteren technisch orientierten Kompetenzteilbereich *Entwicklung und Konstruktion* (MKJS, 2016). Da die Grundstrukturierung sich damit im Wesentlichen an den Naturwissenschaften orientiert, erscheint es möglich, das in Kapitel 2.6 vorgestellte Kompetenzmodell für die Erfassung naturwissenschaftlich-*technischer* Kompetenz im Leistungsfach NwT zu verwenden und um den prozessbezogenen Kompetenzbereich *Entwicklung und Konstruktion* zu erweitern.

Kauertz et al. (2010, S. 137) geben zu bedenken, dass Fachwissen auch in den anderen aufgeführten Kompetenzbereichen benötigt wird, weshalb eine empirische Abgrenzung der Kompetenzbereiche noch zu bestätigen ist. Sie deuten jedoch an, dass es in den Naturwissenschaften „[f]ür die Kompetenzbereiche *Fachwissen* und *Erkenntnisgewinnung* [...] erste Hinweise [gibt], dass eine Unterscheidung empirisch haltbar ist“ (Kauertz et al., 2010, S. 137). Diese Unterscheidung wurde u. a. im IQB-Bildungstrend adaptiert (Mahler et al., 2019, S. 100). Darüber hinaus postuliert Kauertz, dass Fachwissen für eine zielführende Anwendung prozessbezogener Kompetenzen benötigt wird (Kauertz et al., 2010). Für den empirischen Zugang zur Erfassung einer allgemein technischen Kernkompetenz im Leistungsfach NwT ergibt sich daraus die Möglichkeit einer Operationa-

lisierung der Fachkompetenz anhand des Fachwissens. Es wird dabei eine Reduktion der gesamten naturwissenschaftlich-technischen Kompetenzen auf kognitive Fähigkeiten in Bezug auf die Ausprägung der inhaltsbezogenen Kompetenzen des Bildungsplans der Kursstufe unter der Annahme vorgenommen, dass Fachwissen als Basis für die zielgerichtete Anwendung im Sinne des in Kapitel 2.3 definierten Kompetenzbegriffs dient.

Für die Erfassung technischer Kompetenz und insbesondere die inhaltsbezogene Kompetenzfacette liefert die *Einheitliche Prüfungsanforderung in der Abiturprüfung Technik (EPA-Technik)* umfassende Ansatzpunkte (Autorengruppe der Kultusministerkonferenz, 2006). Sie definiert die nachfolgenden Sachkompetenzen:

„Technische Probleme analysieren, Wirkungszusammenhänge ermitteln, Lösungen entwickeln und deren Wirksamkeit beurteilen“ (Autorengruppe der Kultusministerkonferenz, 2006, S. 6).

„Technische Aufgabenstellungen und Lösungen unter den Aspekten ihrer Zweckbestimmung, Funktionalität und Übertragbarkeit analysieren und die Folgen unter Beachtung humaner, ökonomischer und ökologischer Aspekte bewerten“ (ebd., S. 6).

„Optimieren von Lösungen für technische Aufgabenstellungen durch Strukturieren von Lösungswegen, Feststellen möglicher Lösungsvarianten, Vergleichen der Lösungsvarianten, Auswählen einer Variante und Darstellen des Kompromisscharakters der bevorzugten Lösung“ (ebd., S. 6).

„Theorien und Gesetzmäßigkeiten sowie möglicher Analyse- und Syntheseverfahren mittels Reduktion technischer Sachverhalte auf Modelldarstellungen unter Berücksichtigung ihres Geltungsbereichs bereitstellen und anwenden“ (ebd., S. 6).

Für die Berücksichtigung konkreter fachlicher Inhalte im Rahmen der Sachkompetenz erläutert die Autorengruppe der Kultusministerkonferenz (2006), dass die Sachkompetenzen anhand der curricularen Ländervorgaben nachzuweisen sind und sich für das Fach NwT demzufolge an den inhaltsbezogenen Kompetenzen des Bildungsplans orientieren. Zur Festlegung des kognitiven Aufgabenpotenzials dienen die vorgeschlagenen Anforderungsbereiche und Operatoren der EPA-Technik (Autorengruppe der Kultusministerkonferenz, 2006, S. 10 ff).

Neuere Studien legen diesbezüglich nahe, dass sich die Schwierigkeit von Aufgaben anhand der Anforderungsniveaus zutreffend abbilden lässt. Florian, Sandmann und Schmiemann (2014) zeigen in ihrer Untersuchung mit  $n = 2385$  Prüfungsteilnehmenden, dass die Anforderungsniveaus aus der EPA für das Fach Biologie die Aufgabenschwierigkeit anhand der vorgeschlagenen Operatoren

adäquat abbilden. Mithilfe der den Anforderungsbereichen zugeordneten Operatoren kann also bereits in der Aufgabenentwicklung eine erste Niveauabstufung vorgenommen werden.

Eine weitere Möglichkeit zur Einschätzung schwierigkeitsbestimmender Merkmale im Zuge der Aufgabenentwicklung ist das Heranziehen von Bildungsexpert:innen. Diese sollen auf Basis ihrer Unterrichtserfahrung u. a. den Komplexitätsgrad einzelner Aufgaben einschätzen. Unter Einbezug der zuvor dargestellten Erkenntnisse zu den Anforderungsbereichen (Florian et al., 2014) kann (in Anlehnung an das Kompetenzmodell nach Kauertz (2008)) von einem Zusammenhang zwischen Komplexitätsgrad, Anforderungsniveaus bzw. deren zugeordneten Operatoren und der tatsächlichen Aufgabenschwierigkeit ausgegangen werden (s. z. B. Florian et al., 2014; Woitkowski & Riese, 2017).

## **2.8 Zwischenfazit zum theoretischen Hintergrund**

Zur Einordnung des Forschungsvorhabens in den Kontext schulischen Lernens dient das Angebots-Nutzungs-Modell nach Seidel (2014). Gleichzeitig stellt es einen Orientierungsrahmen für die Zielsetzung der Arbeit bezüglich des Zusammenwirkens unterschiedlicher Einflussfaktoren auf Lernprozesse dar. Die Operationalisierung der inhaltsbezogenen Kompetenzen erscheint im Rahmen der Einführung eines Leistungsfachs mit Abiturprüfung bedeutsam für Forschung und schulische Praxis. Auf Basis der Definition des Kompetenzbegriffs nach Weinert (2001) und nach einer allgemeinen Betrachtung zur Kompetenzdiagnostik wird im Rahmen dieser Arbeit die Erfassung des Fachwissens fokussiert. Der Ansatz orientiert sich dabei an der Vorgehensweise zur Messung der naturwissenschaftlichen Kompetenz im Sinne des IQB-Bildungstrends (Mahler et al., 2019). Die zu erstellenden Testaufgaben decken dazu die inhaltlichen Vorgaben der Bildungsstandards für das Leistungsfach NwT ab. In diesem Kontext interessiert v. a., welche Kompetenzniveaus die Lernenden erreichen, welche strukturellen Merkmale das Fachwissen in Bezug zu den inhaltsbezogenen Kompetenzteilbereichen aufweist und welche kognitiven Einflussfaktoren eine Wissensentwicklung positiv begünstigen. Hattie (2009) betrachtet in seiner Metastudie unterschiedliche Einflüsse auf die Lernergebnisse und stellt fest, dass u. a. Vorwissen und weitere kognitive Merkmale der Lernenden eine positive Entwicklung begünstigen können. Für ein Beschreibungswissen zum Fachwissen im Kontext technischer Allgemeinbildung ist die Erfassung der kognitiven Merkmale bedeutsam, da ihr Einfluss im beruflich-technischen Aus- und Weiterbildungsbereich bereits vielfältig nachgewiesen wurde (z. B. Nickolaus et al., 2015). Das Angebots-Nutzungs-Modell zeigt aber gleichermaßen die Grenzen der eigenen Studien auf und welche

weiteren Untersuchungen notwendig sind, um das Beschreibungswissen hinsichtlich individueller Einflussfaktoren (z. B. selbstbezogene Vorstellungen oder Motivation) auf das Lernergebnis der Schüler:innen im Leistungsfach NwT auszubauen.

Kauertz et al. (2010) fordern für die Erfassung von Kompetenzen eine Aufschlüsselung des kognitiven Aufgabenpotenzials in vier Stufen. Demgegenüber stehen die in den Bildungsstandards bzw. der EPA-Technik formulierten drei Anforderungsniveaus. Ergebnisse von Florian et al. (2014) belegen diesbezüglich, dass die Anforderungsniveaus in der EPA die Aufgabenschwierigkeit für die Abiturprüfung adäquat abbilden können und sich drei Anforderungsbereiche mit statistisch belegbaren Schwierigkeitsniveaus abzeichnen. Im Rahmen dieser Arbeit erfolgt die Festlegung des kognitiven Aufgabenpotenzials bei der Aufgabenerstellung gemäß des Ansatzes der Schwierigkeitsbestimmung von Florian et al. (2014). Durch die Verwendung der Operatoren der EPA-Technik entsprechen die Anforderungsniveaus der Aufgaben den Bildungsstandards und ermöglichen damit eine bessere Vermittelbarkeit der Implikationen in die Bildungspraxis. Die Aufgaben bilden alle inhaltsbezogenen Kompetenzteilbereiche des Leistungsfachbildungsplans (Autorengruppe Bildungsplan NwT, 2019b) ab. Ein probates Vorgehen, das z. B. von Woitkowski und Riese (2017) gewählt wird, schlägt zur Einschätzung der Komplexität und inhaltlichen Passung der Aufgaben zusätzlich eine Expert:innenbefragung vor. Dieser Ansatz greift die von Hartig, Frey, Nold und Klieme (2012) und Kauertz et al. (2010) vertretene These auf, dass über rein kognitive Prozesse hinaus ebenfalls inhaltliche Aufgabenelemente die Schwierigkeit eines Items determinieren. Die Expert:innen überprüfen die Testitems auf eine möglichst fachpraxisnahe inhaltliche Validität sowie deren curriculare Relevanz und geben eine Einschätzung bezüglich des Komplexitätsgrads der Aufgabe ab. Zusammengefasst erscheint für die Operationalisierung der inhaltsbezogenen Kompetenzen eine Anwendung des Modells nach Kauertz möglich, wenn die Aufgaben unter Verwendung der Operatoren aus der EPA-Technik zur Festlegung des kognitiven Niveaus einer Expert:innenbefragung zur Feststellung des Komplexitätsgrads unterworfen und nach der finalen Überarbeitung im Anschluss an die Pilotierung nicht weiter verändert werden. Damit wird die Aufgabenkomplexität und das kognitive Aufgabenpotenzial über alle Erhebungen als konstant angenommen, sofern sich die Zielgruppe der Befragten nicht wesentlich zur Pilotierungsstichprobe unterscheidet.

An die Testentwicklung und die deskriptiven Studien zum Fachwissen lassen sich weiterführende Analysen anschließen. Zur Kompetenzniveaumodellierung wird

als Standard-Setting-Verfahren die Kontrastgruppenmethode (Ziely & Livingston, 1977) gewählt. Die Identifikation möglicher Kontrastgruppen wird durch eine Clusteranalyse (vgl. Sireci et al., 1999) in Abhängigkeit der Personenfähigkeit durchgeführt. Die erhaltenen Kompetenzniveaus können durch eine zusätzliche Betrachtung weiterer Konstrukte beschrieben werden, um bekannte Wirkzusammenhänge auf das Fachwissen zu untersuchen. Dazu werden die kognitiven Merkmale *fluide Intelligenz* und *Vorwissen* der Lernenden erfasst. Die Messung der fluiden Intelligenz erfolgt mittels des standardisierten CFT 20-R Testinstruments (Weiß, 2019). Das Testinstrument basiert auf der Intelligenztheorie nach Cattell (1963), die von einer kognitiven Grundfähigkeit in Form von kristalliner und fluiden Intelligenz ausgeht. Die Vorgehensweise bei der Testerstellung für die Erfassung des Vorwissens ist analog zur Erfassung des Fachwissens und bezieht sich auf die inhaltsbezogenen Kompetenzen des Mittelstufencurriculums.





### 3 Forschungsstand

Für die Verortung der Arbeit im Forschungsfeld liefert der nachfolgend ausgeführte Forschungsstand zunächst eine Übersicht über die Kompetenzen von Lernenden in naturwissenschaftlichen Fächern (Kapitel 3.1). Des Weiteren gibt er einen nationalen sowie internationalen Einblick zu Kompetenzen und zum Begriffsverständnis der technischen Allgemeinbildung (Kapitel 3.2) und schließt mit einer spezifischen Betrachtung des Forschungsstands zum Fach Naturwissenschaft und Technik in Baden-Württemberg (Kapitel 3.3) sowie einem Zwischenfazit (Kapitel 3.4) ab.

#### 3.1 Kompetenzen in naturwissenschaftlichen Fächern

Für die Betrachtung des Forschungsstands in Bezug auf die Einordnung der Kompetenzen von Schüler:innen in den Naturwissenschaften im deutschen Bildungssystem findet sich ein Zugang über die internationalen Large-Scale-Assessments TIMSS/III sowie PISA im naturwissenschaftlichen Kontext und größere nationale Schulleistungsstudien wie den IQB-Bildungstrend.

TIMSS/III offenbart im Vergleich zu anderen westlichen Industriestaaten defizitäre mathematische und naturwissenschaftliche Kompetenzen bei Schüler:innen an Gymnasien in Deutschland (Baumert, Bos & Lehmann, 2000, S. 52). Der Forschungsgruppe um Baumert et al. (2000) fällt in diesem Kontext auf, dass sich das Defizit über den Verlauf der Sek. I und Sek. II aufweitet und sogar die leistungsstärksten Schüler:innen im Vergleich mit der Spitze anderer europäischer Länder nicht konkurrieren können (Baumert et al., 2000, S. 52). Dieses Ergebnis führte u. a. dazu, dass in Deutschland die Überprüfung der Unterrichtsqualität an Schulen bildungspolitische Aufmerksamkeit erlangte (vgl. Spiegelhauer, 2017, S. 72).

Die darauffolgenden Studien PISA 2006 (OECD, 2007) und PISA 2015 (OECD, 2016) mit Fokus auf den Naturwissenschaften als Hauptdomäne zeigen, dass die naturwissenschaftliche Literalität von Lernenden in Deutschland im internationalen Vergleich nun auf einem höheren Niveau rangiert (Reiss et al., 2016, S. 4). Demgegenüber fällt jedoch auf, dass im Vergleich der Ergebnisse des Jahres 2015 mit der Durchführung von PISA 2006 am Gymnasium wieder ein Rückgang naturwissenschaftlicher Kompetenz zu verzeichnen ist (vgl. Reiss et al., 2016, S. 5). Der Unterricht zeichnet sich laut Reiss et al. (2016) „vor allem durch sogenannte «*Minds-on*»-Aktivitäten aus, die auf eine kognitive Anregung der Schüler:innen abzielen. «*Hands-on*»-Aktivitäten wie die Entwicklung eigener Experi-

mente oder die Durchführung strukturierter Laborexperimente kommen hingegen seltener vor“ (Reiss et al., 2016, S. 6, Hervorhebungen im Original). Dahingehend zeigt sich der naturwissenschaftliche Unterricht gegenüber der Erfassung im Jahr 2006 wenig verändert (vgl. ebd., S. 6). Reiss et al. (2016) stellen diesbezüglich klar, „dass sich ein guter Naturwissenschaftsunterricht durch eine Kombination hoher kognitiver Anregung [...], regelmäßiger «Hands-on»-Aktivitäten sowie dem Aufzeigen von Anwendungsbezügen auszeichnet“ (Reiss et al., 2016, S. 6, Hervorhebungen im Original).

Die Untersuchung des IQB-Bildungstrends 2018 widmet sich der Analyse mathematischer und naturwissenschaftlicher Kompetenzen (Stanat, Schipolowski, Mahler, Weirich & Henschel, 2019) und ist gleichzeitig die jüngste Studie in der hier dargestellten Chronologie. Die Ergebnisse deuten an, dass die Lernenden zu 45 % in Mathematik und zu je 71 % in Biologie, 56 % in Chemie und 69 % in Physik über dem KMK-Regelstandard einzuordnen sind, wenngleich eine Heterogenität der erreichten Kompetenzstandards für die unterschiedlichen Bundesländer vorliegt (vgl. Stanat et al., 2019, S. 430 ff). In der Trendanalyse zeigt sich in der deutschlandweiten Gesamtbetrachtung keine Veränderung der Ergebnisse beim Erreichen der Bildungsstandards für Mathematik und in den naturwissenschaftlichen Fächern unter Rückbezug auf die Befundlage des Jahres 2012 (ebd., S. 433). Bei einer spezifischen Betrachtung der Gymnasien zeichnen sich jedoch ungünstige Entwicklungen ab. Die Lernenden erzielen in allen Naturwissenschaften v. a. beim Fachwissen signifikant geringere Werte im Vergleich der beiden Messzeitpunkte (ebd., S. 235). In gewisser Weise ergibt dieser Sachverhalt ein gemischtes Bild, indem er eine Stabilisierung des Niveaus für Deutschland insgesamt indiziert, gleichzeitig aber ungünstige Entwicklungen für die Gymnasien belegt. Inwiefern es sich dabei nach wie vor um ein international konkurrenzfähiges Niveau handelt, wird PISA 2025 mit erneutem Fokus auf die naturwissenschaftlichen Domänen zeigen.

Die Frage nach der Wirkung von fächerübergreifendem MINT-Unterricht auf die Entwicklung naturwissenschaftlicher Kompetenzen wird von den großen Studien bislang nicht beantwortet. In den letzten Jahren haben sich jedoch einige Arbeiten mit dieser Thematik beschäftigt (z. B. Busch, 2016; Hoffmann, 2021; Labudde, 2014; Labudde & Schecker, 2021). Brändle (2023) stellt dazu (in Anlehnung an Labudde (2014)) fest:

„Interdisziplinär organisierte naturwissenschaftliche Fächer werden bezüglich ihrer Kompetenzvermittlung (sic) zudem häufiger in Zweifel gezogen, obschon die Literatur einige Argumente zugunsten fächerübergreifenden Unterrichts liefert (Labudde, 2014, S. 13).“ (Brändle, 2023, S. 3)

Eine von Busch (2016, S. 122) durchgeführte Evaluationsstudie zeigt, dass Schüler:innen im interdisziplinären Unterricht eine höhere Kompetenz aufweisen, kann diese aber nicht mit Sicherheit der Interdisziplinarität zuschreiben. Labudde und Schecker (2021) führen einige Beispiele gelungener Umsetzungen ins Feld und liefern empirische Befunde über die Wirksamkeit der jeweiligen Schulversuche. Sie berichten bspw. über einen Ansatz aus der Schweiz, in dem das Fach *MINT* in der gymnasialen Oberstufe fächerverbindend unterrichtet wird (Labudde & Schecker, 2021, S. 449 ff). Eine Evaluation über drei Erhebungszyklen belegt, dass die Schüler:innen im Fach *MINT* „ein hohes Interesse und ein überdurchschnittliches Selbstkonzept bezüglich der Naturwissenschaften und Mathematik“ (Labudde & Schecker, 2021, S. 452) besitzen. Lediglich die Unterrepräsentation technischer Inhalte wird von Schüler:innenseite bemängelt (vgl. ebd., S. 452). Labudde und Schecker führen noch weitere Beispiele an. Deren Evaluationen erscheinen aufgrund der Spezifität und des Stichprobenumfangs jedoch nicht als generalisierbar. Hoffmann (2021) kann diesbezüglich einen umfangreicheren Überblick liefern und fasst zusammen:

„[...] Lernende im fächerübergreifenden naturwissenschaftlichen Unterricht [zeichnen sich] durch eine größere Selbstständigkeit, ein umfassenderes Methodenrepertoire, ein größeres Fachinteresse sowie ein größeres Interesse an naturwissenschaftlichen Arbeitsweisen und Fragestellungen [aus].“ (Hoffmann, 2021, S. 36)

Er ergänzt außerdem, dass im bisherigen Forschungsstand zum Fachwissen „keine signifikanten Unterschiede zwischen Lernenden im disziplinorientierten und fächerübergreifenden Unterricht festzustellen“ (Hoffmann, 2021, S. 36) sind; ein positiv zu bewertender Befund, wenn die Steigerung der affektiven und motivationalen Merkmale der Lernenden bei gleichem Fachwissenserwerb gegenüber disziplinärem Unterricht erreicht wird.

Es existieren für den MINT-Bereich bereits einige Ansätze fächerübergreifenden Unterrichts in Sek. I und Sek. II. Das Fach NwT ermöglicht z. B. in Baden-Württemberg den Zugang zu einer natur- und technikwissenschaftlichen Allgemeinbildung. Zu Kompetenzen im Kontext der technischen Allgemeinbildung können die bisher diskutierten Studien keine Auskunft geben. Der nachfolgende Abschnitt betrachtet daher Kompetenzen im Kontext technischer Allgemeinbildung im nationalen und internationalen Raum.

### **3.2 Kompetenzen im Kontext der technischen Allgemeinbildung**

Durch die Allgegenwärtigkeit und Interdisziplinarität von Technik finden sich drei übergeordnete Bezugfelder technischer Bildung (vgl. Zinn, Tenberg & Pittich,

2018, Kapitel 2). Technische Bildung gliedert sich in den *beruflichen*, den *ingenieurwissenschaftlich-hochschulischen* und den *allgemeinbildenden* Kontext. Der im Folgenden dargestellte Forschungsstand fokussiert sich auf die Befundlage zur technischen Allgemeinbildung, in dem sich das Fach NwT verorten lässt. Im allgemeinbildenden Kontext wird häufig der Begriff *technische Mündigkeit* (z. B. Fletcher, 2020) oder *technische Literalität* (z. B. Zinn, 2018b) angetroffen. Fletcher (2020) setzt sich mit dem Begriffsverständnis auseinander und schlägt für die technische Mündigkeit folgende Definition vor:

„Die Fähigkeit und Bereitschaft, auf der Basis von technikorientiertem Konzept-, Alltags- und Bewertungswissen erfolgreich typische technische Handlungsformen in höchst unterschiedlichen Anwendungskontexten erfolgreich ausführen und die Folgen für sich und die Gesellschaft abschätzen zu können.“ (Fletcher, 2020, S. 139)

Im nationalen Bezugsfeld technischer Allgemeinbildung zeigt sich einerseits, dass Lernende bezüglich ihrer technischen Mündigkeit am Ende der Sekundarstufe I Defizite aufweisen (vgl. Fletcher, 2020, S. 140). Die Studie von Fletcher, Vries und Max (2018) belegt, dass Schüler:innen zwischen 13 und 16 Jahren eine optimierungsbedürftige technische Mündigkeit besitzen und sich zudem Fehlkonzepte verfestigt haben (Fletcher et al., 2018, S. 49 f). Andererseits scheinen sich diese und andere grundlegende technische Kompetenzen über den Verlauf der Sekundarstufe II nicht zu verbessern. Studienanfänger:innen weisen nur unzureichende Voraussetzungen für ein Studium im MINT-Bereich auf, was sich v. a. in hohen Wechsel- und Abbruchquoten von ca. 53 % in natur- und ingenieurwissenschaftlichen Studiengängen niederschlägt (vgl. acatech & Joachim-Herz-Stiftung, 2022, S. 24). Gleichzeitig erweist sich die Passung von Studieninhalten und Interessen der Lernenden als divergent (ebd., S. 24). Eine Förderung technischer Kompetenzen über eigenständige (interdisziplinäre) Fächer im nationalen Bereich erscheint u. a. zur Entwicklung anschlussfähiger Voraussetzungen für ein technikwissenschaftliches Studium notwendig (Zinn, 2014). Vereinzelt Studien zu nationalen Programmen, die technische Allgemeinbildung bei Lernenden adressieren, liegen im Kontext von Energiemündigkeit (Fletcher & Deutsch, 2016) vor. Weitere Abhandlungen beschäftigen sich mit Problemlösekompetenzen im technischen Kontext (Stemmann & Lang, 2014; Stemmann, 2018) sowie zum technischen Verständnis als Basiskompetenz bei Lernenden (Duismann & Meschenmoser, 2009). Zinn et al. (2017) entwickeln in ihrer Abhandlung erstmals ein Testinstrument zur spezifischen Erfassung des technischen Wissens von Schüler:innen im Fach NwT.

Fletcher und Deutsch (2016) stellen in ihrer Studie fest, dass die befragten Schüler:innen nur eingeschränkt kompetent sind, um mündig am gesellschaftlichen

Diskurs zur Energiethematik teilzunehmen (vgl. Fletcher & Deutsch, 2016, S. 106). Stemmann und Lang (2014) entwickeln ein Testinstrument zur Erfassung technischer Problemlösekompetenz. Diesbezüglich stellt Stemmann (2018) fest, dass sich das Konzept der technischen Problemlösekompetenz von der allgemeinen fächerübergreifenden Problemlösekompetenz abgrenzt und sich kontextabhängig darstellt (Stemmann, 2018, S. 57 ff). Mit der reinen Förderung allgemeiner Problemlösekompetenz lässt sich demnach nicht im gleichen Zug die technische Problemlösekompetenz steigern (ebd., S. 57 ff). In einer früheren Studie stellen Duismann und Meschenmoser (2009) bereits fest, dass die Kompetenzentwicklung bezüglich des technischen Verständnisses hauptsächlich auf außerschulische und informelle Lernkontexte zurückzuführen ist (Duisman & Meschenmoser, 2009, S. 114 f).

Im internationalen Raum entspricht das Begriffsverständnis der technischen Mündigkeit (s. Fletcher, 2020, S. 139) dem der *Technology and Engineering Literacy* (vgl. ITEEA, 2020). Hierzu finden sich einige Studien, die (ohne Anspruch auf Vollständigkeit) einen Einblick in den internationalen Forschungskonsens ermöglichen (z. B. Avsec & Jamšek, 2016, 2018; Castillo, 2010; Ebrahiminejad, 2017; Krupczak & Disney, 2013; National Academy of Engineering and National Research Council, 2006; Vries et al., 2018; Walach, 2015).

Es existieren im internationalen Raum einige Ansätze zur Operationalisierung technischer Literalität (Avsec & Jamšek, 2016; Krupczak & Disney, 2013; National Academy of Engineering and National Research Council, 2006; Walach, 2015). Walach (2015) sowie Avsec und Jamšek (2016) entwickeln bspw. Testinstrumente, die sich an den Kompetenzfacetten der ITEEA (2007) orientieren. In einer Testung mit Schüler:innen zeigen Avsec und Jamšek (2018), dass technische Wahlfächer die technische Kompetenz der Lernenden mit einer großen Effektstärke im Vergleich zu einer Kontrollgruppe verbessern (vgl. Avsec & Jamšek, 2018, S. 145). Darüber hinaus stellen sie anhand einer Pfadanalyse fest, dass die wichtigsten Einflussfaktoren auf die technische Kompetenz *kritisches Denken* und *Entscheidungsfindung* darstellen und dass die technische Literalität als ein signifikanter Prädiktor für die weitere Bildungswegentscheidung mit Orientierung an technischen Domänen wirkt (vgl. ebd., S. 145). Diese Befunde stehen im Konsens mit der Studie von Castillo (2010). Er stellt ebenfalls fest, dass ein auf Standards basierender Technikunterricht einen signifikanten Einfluss auf die technische Literalität ausübt (Castillo, 2010, Pos. 15.1196.2). Diese Wirkung ist bedeutsam, denn in einer Literaturübersicht stellt Ebrahiminejad (2017) fest, dass

Ingeneur:innen häufig die Basiskompetenzen *kritisches Denken* und *Problemlösen* weder aufweisen noch über den Studienverlauf entwickeln. Diese Befundlage spricht für die möglichst früh im Bildungsverlauf angesiedelte Entwicklung einer anknüpfungsfähigen allgemeinen technischen Kompetenz (z. B. Adenstedt, 2018, S. 26 f).

Ein interdisziplinärer Ansatz zur Förderung naturwissenschaftlicher und technischer Kompetenzen stellt das Schulfach NwT in Baden-Württemberg dar. Die Befundlage zum Fach NwT wird im Folgenden betrachtet.

### **3.3 Stand der Forschung zum Fach NwT in Mittel- und Kursstufe<sup>5</sup>**

Es bestehen keine Zweifel an der Notwendigkeit einer technischen Grundbildung, wenngleich sich der Forschungsstand in Bezug auf die Wirkung fächerübergreifenden naturwissenschaftlich-technischen Unterrichts am allgemeinbildenden Gymnasium noch ausbaufähig darstellt (vgl. Zinn, 2014, S. 37 ff, 2018c, S. 235). Es erscheint daher interessant, sich mit den Wirkeffekten des Schulfachs NwT zu beschäftigen.

Es liegen Studien u. a. zur Ausbildung der Lehrpersonen für das Fach, den förderlichen räumlichen und sächlichen Voraussetzungen des Fachunterrichts, schulübergreifenden inhaltlichen Schwerpunkten, Organisationsformen, dem Verhältnis von NwT zum traditionellen Fächergefüge und zur Unterrichtsmethodik des Fachs in der Mittelstufe sowie dem zweistündigen Kursstufenfach vor (Mokhonko et al., 2014; Zinn & Latzel, 2017). Die Studie von Mokhonko et al. (2014) in der Mittelstufe zeigt, dass Lehrpersonen im Fach (ohne NwT Lehramtsstudium) hauptsächlich über Fortbildungsveranstaltungen die notwendige Fachkompetenz für den Unterricht erwerben. Die Frage nach einer ausreichenden Fähigkeit zur Förderung inhalts- und prozessbezogener Kompetenzen durch diese Weiterbildungsmaßnahmen stellt sich insbesondere aufgrund der festgestellten Abhängigkeit von Frage- und Problemstellungen von der Qualifikation der Lehrpersonen (Mokhonko et al., 2014, S. 123 ff). In Bezug auf räumliche und sächliche Untersuchungsaspekte wird von der Forschergruppe zudem festgestellt, dass die tradierten Naturwissenschaften eine bessere räumliche Voraussetzung, aber keine Vorteile bei der Ausstattung haben. Lediglich kostenintensive technische Geräte bspw. zur computergestützten Fertigung (CAM; z. B. CNC-Fräse, 3D-Drucker) seien an den Schulen nicht in ausreichender Menge vorhanden (ebd., S.

---

<sup>5</sup> Teile dieses Kapitels entstammen in identischer oder leicht abgewandelter Form einer bereits erschienenen Publikation: Brändle (2023).

112). Mit der sächlichen Ausstattung in Zusammenhang steht die Behandlungsintensität von Kompetenzteilbereichen (ebd., S. 124). In Bezug zur Organisationsform des Fachs wird häufig ein Wechsel der Lehrperson in einem Semester- oder Trimestermodell vollzogen. In diesem Kontext wird außerdem festgestellt, dass sich Themenschwerpunkte herauskristallisieren, die sich stark an der jeweils grundständig studierten Naturwissenschaft der Lehrpersonen zu orientieren scheinen (ebd., S. 125). In Mittel- (ebd.) und Kursstufe (Zinn & Latzel, 2017, S. 3) orientiert sich der Unterricht an der Methode des problembasierten Lernens und setzt dabei auf technische Schülerinnen- und Schülerexperimente. Darüber hinaus wird Projekten als handlungsorientierte Unterrichtsmethode (vgl. Zendler, 2018, S. 22) ein hoher Stellenwert zur Kompetenzförderung eingeräumt.

Im Unterschied zur Mittelstufe zeigt sich bei den Rahmenbedingungen für das zweistündige Wahlfach, dass NwT in der Kursstufe vermehrt von nur einer Lehrperson unterrichtet wird (Zinn & Latzel, 2017, S. 2). Bei der sächlichen Ausstattung werden ebenfalls Großgeräte wie z. B. 3D-Drucker und CNC-Fräse in flächendeckender Verfügbarkeit vermisst (ebd., S. 2 f). Als qualitativ ausreichend und in hoher Zahl vorhandene Geräte und Materialien werden Mikrocontroller, elektronisches Zubehör sowie die Ausstattung für das computergestützte Konstruieren (CAD) genannt (ebd., S. 3).

Ein Forschungsdesiderat bei Mokhonko et al. (2014) ist die Kompetenzentwicklung bei Lernenden in NwT, welches später von Zinn et al. (2017) aufgegriffen wurde. Im Zuge der Studie dieser Forschungsgruppe wurden valide und reliable Instrumente zur Erfassung der Kompetenzen der Schüler:innen zu Beginn der Kursstufe entwickelt und zur Generierung eines Beschreibungswissens der erreichten Kompetenzen der Schüler:innen eingesetzt (Zinn et al., 2017). Das Testinstrument erfasst dabei das NwT-Fachwissen für die Kompetenzteilbereiche *Automatisierungstechnik, Bautechnik, Erneuerbare Energien, Robotik, Schall und Lärmtechnik* sowie *Wetter und Klima* bei n = 763 Schüler:innen (Zinn & Latzel, 2017, S. 3). Die Testung belegt Unterschiede bei Lernenden zwischen den einzelnen Schulen und innerhalb der Klassen (ebd., S. 3). Außerdem wurden affektive und motivationale Merkmale bei Schüler:innen der Jahrgangsstufe I im zweistündigen Fach sowie die Rahmenbedingungen des Schulversuchs von Zinn und Latzel (2017) untersucht. Es zeigt sich, dass die Schüler:innen NwT in der Kursstufe aufgrund „bereits bestehendem fachlichen Interesse an naturwissenschaftlich-technischen Fragestellungen“ (ebd., S. 6) wählen. Damit einhergehend konnte, v. a. bei Jungen, eine hohe identifizierte, intrinsische und interessierte Motivation bei gleichzeitig niedriger extrinsischer Motivation und Amotivation

der Lernenden festgestellt werden (ebd., S. 6). Die Forschungsgruppe um Zinn attestiert dem NwT-Unterricht in der Kursstufe außerdem vielfältige Ansatzpunkte zur Förderung der persönlichen Interessenbildung und einschlägigen beruflichen Orientierung (ebd.).

### 3.4 Zwischenfazit zum Forschungsstand<sup>6</sup>

Über Ansätze fächerübergreifenden MINT-Unterrichts in Deutschland liegt bislang kein umfängliches Beschreibungswissen zu den Wirkeffekten vor. Die bisherigen Studien deuten darauf hin, dass Schüler:innen bei mindestens gleichem Kompetenzgewinn ein stärkeres Interesse an Naturwissenschaften und eine höhere Motivation im Unterricht aufweisen. Die Bereiche *Informatik* und *Technik* sind im allgemeinbildenden MINT-Fächerkanon unterrepräsentiert. Bislang wurden diese Bereiche den Naturwissenschaften im allgemeinbildenden Bereich angehängt. In Baden-Württemberg haben sich in den letzten Jahren zwei fächerübergreifende Ansätze *Informatik, Mathematik, Physik (IMP)* (s. z. B. Bahr & Zinn, 2023) und *Naturwissenschaft und Technik (NwT)* etabliert, die den Wunsch nach einer eigenständigen und fächerübergreifenden technischen und informationstechnischen Allgemeinbildung adressieren. Auch wenn im mathematischen und naturwissenschaftlichen Bereich im Rahmen größerer Schulleistungsstudien national und international bereits vielfältig Erkenntnisse zur Kompetenzausprägung von Schüler:innen generiert wurden, gestaltet sich der Forschungsstand zur technischen Allgemeinbildung und dabei insbesondere zu Fächern am allgemeinbildenden Gymnasium als lückenhaft. Zusammengefasst signalisieren die vorliegenden nationalen Forschungsansätze einen Bedarf an schulintegrierten Programmen bzw. Schulfächern im Kontext technischer Allgemeinbildung, um zielgerichtet diese Kompetenzen zu fördern und über informelle und außerschulische Lernprozesse hinaus eine technische Allgemeinbildung institutionell zu vermitteln. Zinn (2018a) fasst zutreffend zusammen:

„Es kann davon ausgegangen werden, dass Kompetenzen [...] sich aus einem Zusammenwirken von Fachwissen, Fertigkeiten, Erfahrungen, Routinen und der Bereitschaft und Überzeugung zur kritischen Überprüfung der Handlungsweisen entwickeln.“ (Zinn, 2018a, S. 68)

Der internationale Forschungsstand gibt dahingehend Aufschluss, dass er zum einen den Beitrag technischen Unterrichts auf die Entwicklung einer technischen Mündigkeit belegen kann. Zum anderen liefert er Hinweise, dass technischer Unterricht möglichst früh im Bildungsverlauf angesiedelt werden sollte und bis zum

---

<sup>6</sup> Teile dieses Kapitels entstammen in identischer oder leicht abgewandelter Form einer bereits erschienenen Publikation: Brändle (2023).



Schulabschluss fortgeführt wird. Schulfächer wie NwT, die mit dem Ziel der Förderung einer allgemeinen technischen Kompetenz am allgemeinbildenden Gymnasium etabliert werden, können diesem Anspruch möglicherweise gerecht werden. Es gilt jedoch deren Wirkung zu untersuchen.

Diesbezüglich lässt sich feststellen, dass NwT sich nicht zuletzt mit der Einführung von Basis- und Leistungsfach in der Kursstufe weiterentwickelt hat, woraus sich die nachfolgend genannten Forschungsdesiderate ergeben. Durch das Einmünden grundständig studierter oder spezifisch weitergebildeter Lehrpersonen in den Schuldienst sowie durch die Bildungsplanrevision 2016 besteht das Desiderat nach einer aktuellen empirischen Bestandsaufnahme zu den Rahmenbedingungen und zum Bildungoutput. Vor allem im Kontext des im Schulversuch befindlichen Leistungsfachs liegen noch keine Erkenntnisse zu dessen Umsetzung und Wirkung vor. Es fehlen geeignete und auf die aktuellen Rahmenbedingungen angepasste empirische Testinstrumente zur Erfassung des Vorwissens bei Einmündung in die Kursstufe und zur Erfassung der inhaltsbezogenen Kompetenzen (Autorengruppe Bildungsplan NwT, 2019b) der Schüler:innen im Leistungsfach. Damit einhergehend sind bislang keine Aussagen zum Bildungoutput, zu den Fachwissensdimensionen und zum erreichten Kompetenzniveau der Lernenden möglich. Es ist zudem von bedeutsamen Einflüssen kognitiver Voraussetzungen (z. B. fachspezifisches Vorwissen und fluide Intelligenz) auf das Fachwissen auszugehen. Die theoretische Basis für diese Annahme liefern Befunde aus der beruflich-technischen Aus- und Weiterbildung, die das prädiktive Potenzial dieser Konstrukte belegen (z. B. Nickolaus et al., 2015). Zur Beschreibung des Bildungoutputs des Leistungsfachs NwT stellt die Untersuchung dieser Einflussgrößen gegenwärtig ein weiteres Desiderat dar und könnte den Forschungsstand hinsichtlich der Wirkungszusammenhänge für die technische Allgemeinbildung ergänzen. Gleichzeitig könnten aus der Zusammenhangsanalyse Implikationen für die schulische Praxis abgeleitet werden. Es wären z. B. Ansatzpunkte für evidenzorientierte Fördermaßnahmen oder empirische Aussagen zu Zusammenhängen der Kompetenzdimensionen des Vorwissens mit dem Fachwissen im Sinne eines Spiralcurriculums denkbar.



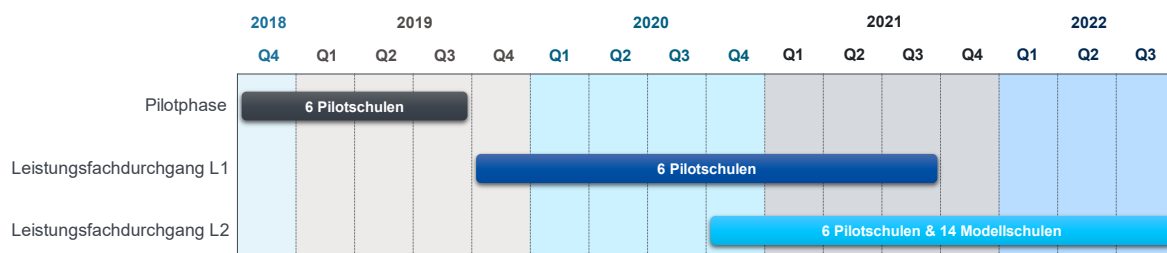
## 4 Fragestellung und Anlage der Untersuchung

Für einen Zugang zur technischen Kompetenz von Schüler:innen in der Oberstufe und insbesondere im Leistungsfach des allgemeinbildenden Gymnasiums ergeben sich ausgehend von den Forschungsdesideraten folgende Forschungsfragen:

- F1** Wie können inhaltsbezogene Kompetenzen im Leistungsfach NwT valide und reliabel gemessen werden?
- F2** Über welches Vorwissen verfügen Schüler:innen bei Einmündung in die Kursstufe?
- F3** Über welches Fachwissen verfügen Schüler:innen des Leistungsfachs NwT am Ende der gymnasialen Kursstufe?
- F4** Welche unterschiedlichen Kompetenzniveaus der Schüler:innen lassen sich feststellen?
- F5** Welche Zusammenhänge zwischen den kognitiven Merkmalen *Vorwissen* sowie *fluide Intelligenz* und dem Fachwissen der Schüler:innen sind nachweisbar?
- F6** Inwiefern eignen sich die kognitiven Merkmale *fluide Intelligenz* und *Vorwissen* zur Einteilung der Lernenden in Kompetenzniveaus anhand ihrer Eingangsvoraussetzungen?

### 4.1 Anlage der Untersuchung<sup>7</sup>

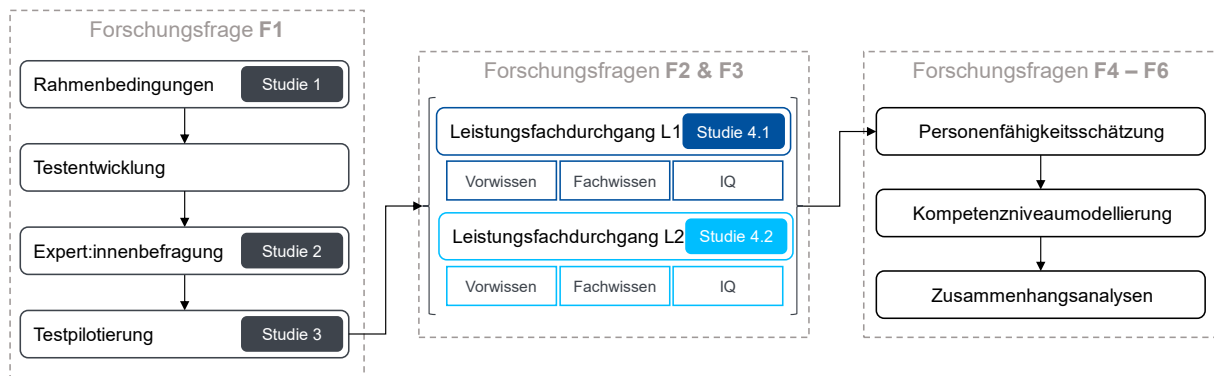
Bei der Anlage der Untersuchung ist zu berücksichtigen, dass diese in die wissenschaftliche Begleitung des Schulversuchs *NwT-Kursstufe* eingebunden ist. Hierdurch wird gleichzeitig ein Rahmen für das Vorgehen in der Arbeit vorgegeben. Die Gesamtuntersuchung orientiert sich methodisch an einem Design-Based-Research Ansatz mit regelmäßigen formativen Rückmeldezyklen an die Lehrpersonen (s. Kapitel 5.1). Die Evaluation des Schulversuchs gliedert sich dabei in drei übergeordnete Phasen (Abbildung 5).



**Abbildung 5:** Übersicht über den geplanten Ablauf des Schulversuchs *NwT-Kursstufe*.

<sup>7</sup> Teile dieses Kapitels entstammen in identischer oder leicht abgewandelter Form einer bereits erschienenen Publikation: Brändle (2023).

Im ersten Schuljahr der Einführung (Schuljahr 2018/19) entwickeln die Lehrpersonen im noch vierstündigen Kursstufenfach Unterrichtskonzepte und -materialien (Pilotphase). Daran schließen sich im Evaluationsprojekt zwei Schulversuchsdurchgänge im fünfstündigen Leistungsfach an. Im ersten Durchgang über die Schuljahre 2019/20 und 2020/21 (Leistungsfachdurchgang L1) legen die Schüler:innen der insgesamt sechs Pilotschulen das erste mögliche Abitur im Leistungsfach ab. Während die erste Durchführungswelle läuft, ist parallel dazu im Schuljahr 2020/21 die Einmündung von insgesamt 14 weiteren Modellschulen in den Schulversuch geplant. Im zweiten Durchgang absolvieren die Lernenden an den insgesamt 20 Pilot- und Modellschulen über die Schuljahre 2020/21 und 2021/22 (Leistungsfachdurchgang L2) ihr Abitur. Das Forschungsvorhaben der Arbeit orientiert sich am übergeordneten Ablauf des Schulversuchs und umfasst die in Abbildung 6 dargestellten Studien.



**Abbildung 6:** Abfolge der Arbeitsschritte, Studien und Analysen zur Beantwortung der Forschungsfragen.

Die geplanten Studien ergänzen sich bzw. bauen aufeinander auf: Es werden (1.) durch eine qualitativ-inhaltsanalytisch angelegte Studie inhaltliche und kontextuelle Aspekte des Unterrichts in der Kursstufe aus Sicht von Lehrpersonen (Angebotsebene) als Bildungsexpert:innen erfasst. Die daraus resultierenden explorativen Erkenntnisse unterstützen (2.) die Erstellung von reliablen und validen Testinstrumenten für die Untersuchung des Fachwissens (im Sinne der Repräsentation inhaltsbezogener Kompetenzen des NwT-Bildungsplans) als bedeutenden Teil der Fachkompetenz (Ergebnisebene) der Schüler:innen in der NwT Kursstufe. Diesbezüglich lassen sich außerdem standortübergreifend Vergleiche zwischen der Gruppe der Pilot- und Modellschulen vornehmen. Die gleichzeitige Rückmeldung dieser Ergebnisse an die Bildungsexpert:innen verschafft der schulischen Praxis einen Überblick über die Lernergebnisse der Schüler:innen in den einzelnen Kompetenzteilbereichen des Bildungsplans im Sinne eines Bildungs-

monitorings. Des Weiteren ergibt sich (3.) im Rahmen der Erhebung weiterer relevanter Facetten für die Charakterisierung der unterschiedlichen Kompetenzniveaus ein quantitatives Beschreibungswissen über kognitive Merkmale wie Intelligenz und Vorwissen der Schüler:innen (ausgewählte Aspekte der Nutzungsebene).

#### **4.1.1 Rahmenbedingungen des Leistungsfachs**

Die inhaltlichen und kontextuellen Rahmenbedingungen des Leistungsfachs werden über teilstandardisierte Interviews mit Lehrpersonen der Pilotierungsphase erfasst (Studie 1). Die Interviews dienen erstens zur aktuellen Bestandsaufnahme, da vermutet wird, dass sich das Fach NwT gegenüber der Befundlage im Forschungsstand (Kapitel 3.3) weiterentwickelt hat. Zweitens sollen die Interviews einen Feldzugang ermöglichen, um Implikationen für die zielgerichtete Entwicklung der benötigten Testinstrumente zu erhalten. In diesen Kontext besteht u. a. Erkenntnisinteresse daran welche Kompetenzen aus der Mittelstufe (Vorwissen) bedeutsam für das Leistungsfach sind und wie die Umsetzung der Inhalte des Leistungsfachbildungsplans aktuell und zukünftig geplant ist.

Die Erkenntnisse aus den Interviews mit einbeziehend, schließt sich die Aufgabenentwicklung zur Erstellung geeigneter Testinstrumente an. Neben den qualitativen Interviews werden die Lehrpersonen gebeten Übungs- und Prüfungsaufgaben einzureichen, um eine möglichst hohe inhaltliche, curriculare und unterrichtsbezogene Validität der Testinstrumente zu erreichen.

#### **4.1.2 Entwicklung und Pilotierung der Testinstrumente**

Zur Erfassung des Vorwissens und zu jedem der neun Kompetenzteilbereiche des Bildungsplans (Kapitel 2.4.2) wird ein Aufgabenpool erstellt, der von den bereits zuvor interviewten Bildungsexpert:innen hinsichtlich Aufgabenkomplexität sowie curricularer und unterrichtlicher Relevanz bewertet wird (Studie 2).

Die hinsichtlich der o. a. Kriterien der Expert:innenbefragung überarbeiteten Aufgaben werden zu Testskalen für die jeweiligen Kompetenzteilbereiche und zum Vorwissen zusammengestellt. Die Erprobung der Testinstrumente erfolgt mit den Schüler:innen der Pilotphase (Studie 3). Aufgrund einer zu erwartenden geringen Stichprobengröße werden Gütekriterien der klassischen Testtheorie zugrunde gelegt. Die geplante Vorgehensweise bis zu diesem Punkt adressiert die Forschungsfrage **F1**.

#### 4.1.3 Deskriptive Untersuchung des Fachwissens

Die aus der Pilotierung resultierenden Testinstrumente werden zur deskriptiven Untersuchung des Fach- und Vorwissens der Schüler:innen in den Leistungsfachdurchgängen L1 (Studie 4.1) und L2 (Studie 4.2) verwendet. Zu Beginn der Kursstufe erfolgt jeweils die Erfassung des Vorwissen der Schüler:innen aus der Mittelstufe. Damit wird das Desiderat der Forschungsfrage **F2** einer Klärung zugeführt. Die Erhebung des Fachwissens wird in sieben abiturrelevanten Kompetenzteilbereichen (*Technikfolgenabschätzung, Energieversorgung, Elektrische Antriebstechnik, Technische Mechanik, Produktentwicklung, Aufnahme und Verarbeitung von Signalen in der Messtechnik* und *Regelungstechnik*) an zwei Erhebungszeitpunkten über den Verlauf der Kursstufe vorgenommen. Die erste Erhebung findet nach einem dreiviertel Schuljahr statt. Die zweite Erhebung kurz vor dem Abitur. Die Kompetenzteilbereiche *Grundlagen der Elektronik* und *Datenkommunikation* werden auf einen Messzeitpunkt nach dem Abitur ausgelagert. Die Auslagerung von *Datenkommunikation* resultiert aus der geplanten Durchführung des Themas an allen Schulen nach dem Abitur. *Grundlagen der Elektronik* werden ausgelagert, um das erste Testinstrument nicht zu überladen. Begründet wird die Auslagerung durch den hohen Anteil an prozessbezogenen Kompetenzen im Bildungsplan (Autorengruppe Bildungsplan NwT, 2019b, S. 21), die im Abitur nicht abgeprüft werden. Eine genaue Erläuterung der Vorgehensweise erfolgt im Rahmen der einzelnen Studien in den Kapiteln 8 und 9. Ziel der Erhebungsdurchgänge L1 und L2 ist die Beantwortung der Forschungsfrage **F3**.

#### 4.1.4 Kompetenzniveaumodellierung & Zusammenhangsanalyse kognitiver Merkmale

Nach Abschluss der beiden Erhebungsdurchgänge L1 und L2 kann mit der kombinierten Stichprobe das Analysepotenzial der Skalen mittels IRT-Skalierung ausgebaut werden. Über die erhaltene Personenfähigkeit wird eine Segmentierung der Stichprobe nach den erreichten Kompetenzniveaustufen in Form von Kontrastgruppen vorgenommen und adressiert damit das Desiderat der Forschungsfrage **F4**. Der theoretische Hintergrund deutet auf einen Zusammenhang kognitiver Voraussetzungen (Vorwissen, fluide Intelligenz) von Schüler:innen mit deren Lernergebnis an (Hattie, 2009, S. 41 f). Gleichzeitig deuten z. B. Befunde von Nikolaus et al. (2015) in diversen Modellen zum technischen Fachwissen in unterschiedlichen Kontexten einen medierten Zusammenhang der fluiden Intelligenz (IQ) auf das Fachwissen über fachrelevantes Vorwissen an. Zur Beantwortung der Forschungsfrage **F5** werden daher die Zusammenhänge der Konstrukte *Fachwissen, Vorwissen* und *fluide Intelligenz (IQ)* betrachtet und die nachfolgenden Hypothesen geprüft:

**H1** Es besteht ein signifikant positiver Zusammenhang zwischen Vorwissen und Fachwissen der Schüler:innen im Leistungsfach NwT.

**H2** Es besteht ein signifikant positiver Zusammenhang zwischen Vorwissen und fluider Intelligenz der Schüler:innen im Leistungsfach NwT.

**H3** Es besteht ein signifikant positiver Zusammenhang zwischen Fachwissen und fluider Intelligenz der Schüler:innen im Leistungsfach NwT.

Auf Basis der bekannten Theorien von u. a. Nickolaus et al. (2015) wird zudem angenommen, dass das Fachwissen im Leistungsfach nicht direkt durch die fluide Intelligenz determiniert wird sondern indirekt über das Vorwissen wirkt. Es wird dahingehend folgende Hypothese geprüft:

**H4** Der Einfluss der fluiden Intelligenz auf das Fachwissen wird durch das Vorwissen vollständig mediiert.

Ein weiteres Erkenntnisinteresse besteht hinsichtlich der Eignung der fluiden Intelligenz und des Vorwissens zur Einordnung der Lernenden in Kompetenzniveaus anhand ihrer Eingangsvoraussetzungen. Forschungsfrage **F6** soll dahingehend mittels Varianzanalysen beantwortet werden.



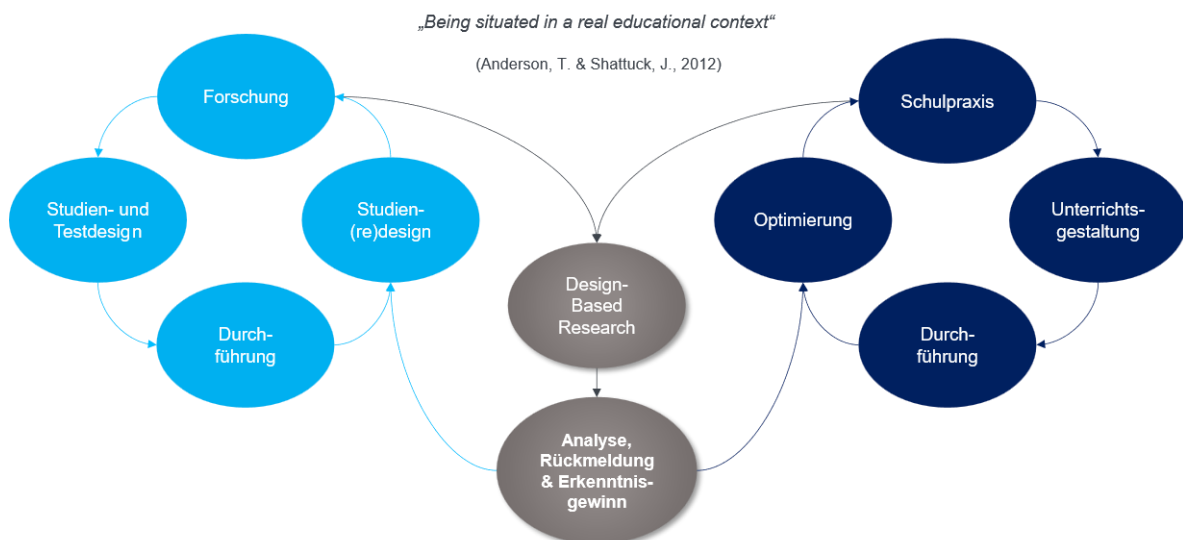


## 5 Methodische Grundlagen

Im Folgenden werden die methodischen und statistischen Grundlagen für die vorliegende Arbeit dargestellt. Zuerst wird hierzu der Forschungsansatz des Gesamtprojekts *NwT-Kursstufe* beschrieben (Kapitel 5.1). Die darauffolgenden Abschnitte thematisieren qualitative Forschungsmethoden (Kapitel 5.2) für die Exploration des Forschungsfelds sowie die Testgestaltung zur quantitativen Erfassung von Fachwissen (Kapitel 5.3). Anschließend werden die wissenschaftlichen Gütekriterien für quantitative Testinstrumente in Kapitel 5.4 dargestellt, um darauffolgend in Kapitel 5.5 und 5.6 die Skalenanalyse der Testinstrumente mittels Klassischer und Probabilistischer Testtheorie zu umreißen. Als Grundlage für die Niveaumodellierung beschäftigt sich Kapitel 5.7 mit explorativer Datengruppierung. Die Methodik zur Untersuchung von Zusammenhängen wird in Kapitel 5.8 dargestellt.

### 5.1 Forschungsansatz im Projekt *NwT-Kursstufe* – Design-Based Research

Anderson und Shattuck (2012) betonen, dass es für Forschende wichtig ist, in realen Bildungskontexten verortet zu sein, um den Nutzen für die Wissenschaft und für die Schulpraxis gleichermaßen zu maximieren und zwei getrennte Ökosysteme (s. Abbildung 7) zu verknüpfen. Dieses Vorgehen hat sich im Bezugsfeld der Bildungsforschung als *Design-Based Research* (DBR) etabliert (Anderson & Shattuck, 2012; Fischer, Waibel & Wecker, 2005; Reinmann, 2005).



**Abbildung 7:** Schematische Darstellung des DBR-Ansatzes im Kontext eines Bildungsmonitorings (Eigene Abbildung; Zitat: Anderson & Shattuck, 2012).

Der DBR-Ansatz weist Gemeinsamkeiten mit der Methodik formativer Evaluationsforschung auf, die die Interventionsoptimierung zum Ziel hat (vgl. Fischer et al., 2005, S. 435). Sie geht gleichzeitig noch einen Schritt weiter und verfolgt ein

zur formativen Optimierung „gleichrangiges Ziel, eine von Theorien geleitete Forschung zum Lernen in komplexen Systemen zu betreiben und Theorien des Lehrens und Lernens zu entwickeln“ (Fischer et al., 2005, S. 435). Der DBR-Ansatz eignet sich damit für den Einsatz im Rahmen von Bildungsmonitorings<sup>8</sup>, mit denen eine formative Rückmeldung an die kollaborierenden Bildungsinstitutionen und Stakeholder in mehreren Iterationszyklen angestrebt sowie gleichzeitig ein Beitrag zur einschlägigen Evidenz im Forschungsfeld und der darauf basierenden Anpassung des Studien- und Testdesigns geleistet wird (z. B. Merk, Poindl & Bohl, 2018).

## **5.2 Exploration des Forschungsfelds**

Bei der Exploration des Forschungsfelds im Rahmen von Pilotstudien führen aufgrund der geringen Anzahl von Personen oder zielgerichteten Auswahl der Stichprobe oftmals Auswertungen auf quantitativer Basis zu nur wenig repräsentativen Aussagen, weshalb auf qualitative Ansätze zurückgegriffen wird (z. B. Flick, 2021, S. 26 f). Im Kontext der Schulversuchsforschung erscheint es sinnvoll, sich zunächst einen Überblick über die Umsetzung des zu untersuchenden Fachs zu verschaffen und zu diesem Zweck praktizierende Lehrpersonen in ihrer Funktion als Bildungsexpert:innen zu befragen.

### **5.2.1 Expert:inneninterviews**

Durch Expert:inneninterviews<sup>9</sup> (ebd., S. 214 ff) im Kontext qualitativer Forschung ergibt sich forschungsmethodisch ein Feldzugang zu kleinen Stichproben, die hinreichende Einblicke in der explorativen Phase der Erschließung des Forschungsfeldes und im Rahmen der Instrumentenentwicklung für den weiteren Forschungsprozess geben können (Bogner & Menz, 2002, S. 37; Flick, 2021, S. 218). Nach Bogner und Menz erfüllen Expert:innen in ihrer Funktion folgende Anforderungen:

„Der Experte verfügt über technisches Prozess- und Deutungswissen, das sich auf sein spezifisches professionelles oder berufliches Handlungsfeld bezieht. Insofern besteht das Expertenwissen nicht allein aus systematisiertem, reflexiv zugänglichem Fach- oder Sonderwissen, sondern es weist zu großen Teilen den Charakter von Praxis- oder Handlungswissen auf, in das verschiedene und durchaus disparate Handlungsmaximen und individuelle Entscheidungsregeln, kollektive Orientierungen und soziale Deutungsmuster einfließen“ (Bogner & Menz, 2002, S. 46).

---

<sup>8</sup> Definition nach Rürup, Fuchs und Weishaupt (2016, S. 411): „Die Begriffe „Bildungsberichterstattung/Bildungsmonitoring“ stehen für Verfahren der systematischen Informationsgewinnung und Informationsaufbereitung über das Bildungswesen, mit denen Gestaltungsentscheidungen zielgerichtet vorgenommen werden sollen.“

<sup>9</sup> Eine Gegenüberstellung unterschiedlicher Interviewarten findet sich bei Flick (2021, S. 193-278).

Darüber hinaus betonen Bogner und Menz (2002), dass Expert:innen innerhalb ihrer Organisation in der Lage sein müssen, erstens Wissen in der Praxis umzusetzen und zweitens mit ihrer Expertise gleichzeitig einen Einfluss auf die Handlungen weiterer Akteur:innen auszuüben (ebd., S. 46). Als Expert:innen im Bezugsfeld der vorliegenden Arbeit sind somit u. a. Lehrpersonen zu verstehen, die maßgeblich an der frühen Gestaltungsphase des Schulversuchs *NwT-Kurststufe* mitwirken und damit u. a. in die Entwicklung des Bildungsplans oder von Weiterbildungskonzepten für zukünftige Akteur:innen eingebunden sind.

### 5.2.2 Qualitative Inhaltsanalyse

Zur Auswertung des Interviewmaterials eignet sich die qualitative Inhaltsanalyse (z. B. Kuckartz & Rädiker, 2022; Mayring, 2022), die darauf abzielt, den Umfang des vorliegenden Textes anhand von Kategorien zu strukturieren und auf wesentliche Punkte zu reduzieren (vgl. Flick, 2021, S. 409). Nach Mayring (2022) gibt es für die qualitative Inhaltsanalyse einige grundlegende Regeln. Die zentralen Punkte für die Auswertungen im Rahmen dieser Arbeit sind (1.) ein Vorgehen, das regelgeleitet und systematisch erfolgt, (2.) die Orientierung der Analyse an der Theorie und am Forschungsstand, (3.) die Erarbeitung eines Kategoriensystems, das das Zentrum der Strukturierung bildet, (4.) die Auswertung des strukturierten Materials unter Einbezug diverser qualitativer und quantitativer Methoden sowie (5.) die Einhaltung wissenschaftlicher Gütekriterien (vgl. Mayring, 2022, S. 49-52).

Unter (1.) ist die „Orientierung an vorab festgelegten Regeln der Textanalyse“ (Mayring, 2022, S. 49) zu verstehen. Da es sich bei der qualitativen Inhaltsanalyse nicht um ein starr definiertes Instrument handelt, kann und muss eine Anpassung an das jeweilige Material erfolgen. Kuckartz und Rädiker (2022) nennen drei gängige Ansätze der qualitativen Inhaltsanalyse: *Inhaltlich strukturierend*, *evaluativ* sowie *typenbildend* (s. Kuckartz & Rädiker, 2022, S. 129-195). Die Anpassung der Analysemethodik an das Material orientiert sich (2.) am Forschungsstand und an den jeweiligen Forschungsfragen (Kuckartz & Rädiker, 2022, S. 63; Mayring, 2022, S. 52). Weil es aufgrund der Verschiedenheit des Materials, an dem die qualitative Inhaltsanalyse angewandt werden kann, keinen einheitlichen Standard gibt, ist es wichtig, die Vorgehensweise des Forschenden nachvollziehen zu können und es auch anderen Bewertern zu ermöglichen, das Textmaterial in ähnlicher Art und Weise zu bearbeiten<sup>10</sup> (Mayring, 2022, S. 50).

---

<sup>10</sup> Die inhaltlich strukturierende qualitative Inhaltsanalyse (s. Kuckartz & Rädiker, 2022, S. 129-156) eignet sich gut für die Analyse von Material, das durch leitfadengestützte Interviews generiert wird. Die nachfolgenden Ausführungen beziehen sich daher auf dieses Verfahren.

Dazu wird (3.) ein Kategoriensystem ausgearbeitet, das das Zentrum der qualitativ inhaltsanalytischen Vorgehensweise bildet (vgl. ebd.). Die Kategorien dienen durch ihre Strukturierung zur Beantwortung der Forschungsfragen und erfassen dazu vollumfänglich die wichtigen Facetten aus dem Analysematerial (vgl. Kuckartz & Rädiker, 2022, S. 63). Gleichzeitig ist zu beachten, dass die Kategorien und Subkategorien für Außenstehende verständlich dargestellt werden, indem auf komplizierte Sprache und Fachbegriffe verzichtet wird (vgl. ebd., S. 65). Zudem ist auf deren Eindeutigkeit und Trennschärfe zu achten und darauf, dass die Subkategorien stets als Ausprägung auf ihre Hauptkategorie zurückzuführen sind (vgl. ebd., S. 64). Bei der Bildung von Subkategorien ist es außerdem wichtig, dass das Kategoriensystem nicht zu feingliedrig wird und durch seinen Umfang an Übersichtlichkeit verliert (vgl. ebd., S. 135). Zur Organisation des Kategoriensystems eignen sich hauptsächlich drei Strukturen. Es kann linear, hierarchisch oder netzwerkartig angelegt werden (vgl. Kuckartz & Rädiker, 2022, S. 61). Für die wissenschaftliche Anwendung ist die hierarchische Anordnung nach Oberkategorien und untergeordneten Subkategorien etabliert. Es gibt keine festgelegte Grenze für die Anzahl an Ebenen. In der Praxis reicht meist eine zwei- bis vierstufige Strukturierung aus, um eine entsprechende Interpretationstiefe bei gleichzeitigem Erhalt der Übersichtlichkeit zu erreichen (Rädiker & Kuckartz, 2019, S. 97). Im Erstellungsprozess des Kategoriensystems raten Kuckartz und Rädiker (2022, S. 132 ff) für die inhaltlich strukturierende qualitative Inhaltsanalyse zu einer Vorgehensweise, bei der zunächst die Hauptkategorien aus den Forschungsfragen sowie dem Interviewleitfaden gebildet werden und ein Teil des Materials mit den Hauptkategorien codiert wird, um deren Verwendbarkeit zu überprüfen. Nach dem ersten Materialdurchgang können induktiv Subkategorien als Ausprägungen der Hauptkategorien gebildet werden, um die Datenbasis in einem zweiten Bewertungsdurchgang facettenreicher zu strukturieren (Kuckartz & Rädiker, 2022, S. 133 ff). Nach Abschluss der Codierung schließt sich (4.) die Auswertung des strukturierten Materials unter Einbezug diverser qualitativer und quantitativer Methoden an. Die Zusammenstellung in Tabelle 2 liefert eine kurze Übersicht über gängige Auswertungsmethoden.

**Tabelle 2:** Möglichkeiten zur Analyse des codierten Materials im Rahmen der inhaltlich strukturierenden qualitativen Inhaltsanalyse (Auswahl nach Kuckartz & Rädiker, 2022, S. 147-154).

Analyseform	Beschreibung
Kategorienbasierte Analyse entlang der Hauptkategorien	Umfasst die Betrachtung der generellen Ergebnisse zu jeder Hauptkategorie. Für den Fall, dass sich eine Hauptkategorie in Subkategorien aufschlüsselt, können diese dargestellt und die Häufigkeit der Zuordnung berichtet werden.
Zusammenhänge zwischen den Subkategorien einer Hauptkategorie	Einerseits können Zusammenhänge von Subkategorien innerhalb einer Hauptkategorie analysiert werden und andererseits können Querverbindungen zu anderen Subkategorien erkannt werden (darunter fällt die häufige gemeinsame Nennung zweier Subkategorien im Interviewmaterial kategorieintern oder -übergreifend).
Paarweise Zusammenhänge zwischen Kategorien	Spezifischere Form der Zusammenhangsanalyse zwischen Subkategorien, die v. a. das Verhältnis des Zusammenhangs zwischen zwei Kategorien betrachtet und nicht nur dessen Auftreten.
Mehrdimensionale Konfigurationen von Kategorien	Erweiterung der Zusammenhangsanalyse über die paarweise Betrachtung hinaus erlaubt die Erfassung komplexerer Zusammenhänge (z. B. Auswirkungen des Budgets auf die Ausstattung und deren Einfluss auf die unterrichtliche Umsetzungsmöglichkeit eines Schulfachs).

Im Rahmen der qualitativen Forschung ist (5.) die Einhaltung von wissenschaftlichen Gütekriterien als genauso wichtig zu erachten wie für quantitative Methoden. Im nachfolgenden Abschnitt werden die angesprochenen Gütekriterien im Kontext der qualitativen Inhaltsanalyse umfangreicher beschrieben.

### 5.2.3 Gütekriterien qualitativer Inhaltsanalyse

In der Literatur ist die Diskussion der Übertragbarkeit der klassischen Gütekriterien in der quantitativen Forschung auf die qualitative Inhaltsanalyse nicht abgeschlossen (Flick, 2021; Kuckartz & Rädiker, 2022). Dennoch besteht die Forderung einer Qualitätssicherung des Auswertungsprozesses (Mayring, 2022). Zur Erfüllung dieser Forderung schlagen Flick, Kuckartz und Rädiker, Mayring und andere eine Spezifizierung und Anpassung der klassischen Gütekriterien der quantitativen Forschung an die qualitative Methodik vor.

Zur Überprüfung der Objektivität eignet sich die *Intercoderübereinstimmung* (vgl. Mayring, 2022, S. 119). Für deren Ermittlung wird das vorliegende Material von zwei unabhängigen Forschenden beurteilt und die Übereinstimmung (z. B.

*Cohens  $\kappa$*  berechnet (Wirtz & Caspar, 2002, S. 56 ff). Mayring (2022) schlägt zur Ermittlung der Reliabilität die *Intracoderübereinstimmung* vor, in deren Rahmen die Forschenden mit zeitlichem Abstand das eigene Material erneut codieren, um die Stabilität, Reproduzierbarkeit und Exaktheit ihrer vorhergehenden Analyse und des verwendeten Kategoriensystems zu verifizieren (Mayring, 2022, S. 119-124). Die Intracoderübereinstimmung wird in der Forschungspraxis selten eingesetzt (s. Göhner & Krell, 2020). Häufiger wird die *Intercoder-Reliabilität* (z. B. Gläser-Zikuda, 2015, S. 126 f) verwendet, unter der der Objektivitäts- und Reliabilitätsanspruch gleichermaßen subsummiert wird, indem eine unabhängige Person mit der Nutzung desselben Kategoriensystems (Reliabilität im Sinne der Reproduzierbarkeit des Ergebnisses durch eine zweite Person; Schreier, 2012, S. 34) zu einem möglichst identischen Ergebnis (Objektivität durch Intersubjektivität; s. Früh, 2004, S. 42) kommt. Validität untergliedert sich im Kontext der qualitativen Inhaltsanalyse nach Material-, Ergebnis- und Prozessorientierung (vgl. Mayring, 2022, S. 121). Die Materialorientierung umfasst die Stichprobengültigkeit, die sich an den üblichen Kriterien zur Stichprobenziehung orientiert (ebd.). Die ergebnisorientierte Validität kann im Anschluss an die Auswertung z. B. durch korrelative Zusammenhänge zum Forschungsstand und zur Theorie überprüft werden (ebd.). Die am Prozess orientierte „Konstruktvalidität lässt sich [...] durch einige Kriterien überprüfen wie

- Bisherige Erfolge mit ähnlichen Konstrukten und/oder Situationen
- Erfahrungen mit dem Kontext des vorliegenden Materials
- Etablierte Theorien und Modelle
- Repräsentative Interpretationen und Experten“ (Mayring, 2022, S. 122).

Eine weitere Möglichkeit der prozessorientierten Validierung stellt die kommunikative Validierung dar, die in einem diskursiven Verfahren zwischen Forschenden und Beforschten eine Konsensfindung über die Interpretation anstrebt (ebd.).

### **5.3 Aspekte der Testgestaltung zur Erfassung von Fachwissen**

Zur Erfassung des Fachwissens der Schüler:innen werden quantitative Testinstrumente entwickelt und eingesetzt. Die Entwicklung der Testinstrumente erfordert eine standardisierte Vorgehensweise bei der Testplanung, Aufgabenerstellung und der statistischen Überprüfung der Aufgaben- und Testgüte.

#### **5.3.1 Testplanung**

Um ein für die Zielgruppe angemessenes Testinstrument zu erstellen, müssen Vorüberlegungen zum generellen Testaufbau, zur Länge des Tests und zur administrativen Umsetzung erfolgen. Es ist zu berücksichtigen, ob das zu messende

Konstrukt qualitativ oder quantitativ erfasst werden soll. Bei qualitativen Messungen existieren diskrete Kategorien, in die sich die Testpersonen unterteilen lassen (Jonkisz, Moosbrugger & Brandt, 2012, S. 28). Die erhaltenen Daten sind nominalskaliert. Quantitative Messungen generieren ordinal- oder intervallskalierte Daten, die eine Abstufung der Merkmalsausprägung ermöglichen (vgl. ebd.). Im Rahmen der vorliegenden Arbeit werden Leistungstests entwickelt und eingesetzt. Diese lassen sich in zwei Gruppen – Geschwindigkeits- oder Niveautests – unterteilen (s. ebd., S. 30). Zur Erfassung kognitiver Merkmale (z. B. fluide Intelligenz) werden meist Geschwindigkeitstests eingesetzt, die eine Differenzierung der Stichprobe über die Anzahl richtig gelöster Aufgaben innerhalb einer begrenzten Bearbeitungszeit ermöglichen (ebd.). Die Aufgaben sind eher einfach gestellt und für alle Testpersonen lösbar (ebd.). Demgegenüber stehen Niveautests, die schwierigere Aufgaben beinhalten (ebd.). Die Leistungsdifferenzierung erfolgt damit über das Aufgabenniveau, das gewährleistet, dass einzelne Aufgaben selbst bei potenziell unbegrenzter Testzeit nicht von allen Proband:innen gelöst werden können (vgl. ebd.). An die Festlegung der Testart schließt sich die Definition des Geltungsbereichs an (ebd., S. 32). Dabei ist die inhaltliche Validität der Testaufgaben zu berücksichtigen (vgl. Kapitel 5.4 in dieser Arbeit), um zu gewährleisten, dass das zu messende Merkmal in gewünschter Weise durch die Testaufgaben abgebildet wird (ebd., S. 32 f). Die Zielgruppe des Tests ist ebenfalls entscheidend für die Merkmalsabbildung, da sie die inhaltliche Breite beeinflusst (ebd.). Je größer der zu erfassende Personenkreis wird, desto höher werden die Anforderungen in Bezug auf Schwierigkeitsbereich und inhaltlichem Wirkungsgrad der Aufgaben (ebd.).

Die Vorüberlegungen bezüglich der Testart, des Geltungsbereichs und der Zielgruppe geben zusammen den strukturellen Testaufbau vor (vgl. ebd., S. 34). Generell bestehen psychometrische Testinstrumente aus mehreren Aufgaben, die ein einzelnes Konstrukt erfassen, um die Testreliabilität (ausführlicher in Kapitel 5.5.2 in dieser Arbeit) und damit die Zuverlässigkeit der Messung statistisch ermitteln zu können (ebd.). Wenn in multidimensionalen Testinstrumenten mehrere Konstrukte erfasst werden sollen, gliedert sich der Gesamtest in mehrere unidimensional gestellte Subtests (ebd.). Abschließend wird die Festlegung der Testlänge und Testzeit betrachtet. Die Testlänge versteht sich dabei als die Anzahl der Aufgaben innerhalb eines Tests (ebd.). Die Testzeit entspricht der Gesamtzeit, die Proband:innen für die Bearbeitung der einzelnen Aufgaben benötigen (ebd.). Nach Jonkisz et al. (2012, S. 34 f) erscheint es sinnvoll, einen Kompromiss aus Testlänge und Bearbeitungszeit anzustreben, um zu gewährleisten, dass bei den Testteilnehmer:innen einerseits Konzentration und Motivation bis

zum Ende des Tests hoch bleiben und andererseits genügend Items in der Skala eine hohe Reliabilität der Messung gewährleisten.

### 5.3.2 Itemkonstruktion

Für die Itemkonstruktion wird zwischen Aufgabentypen mit freiem und gebundenem Antwortformat unterschieden (vgl. Jonkisz et al., 2012, S. 39). Während bei Aufgaben mit freiem Antwortformat keine Lösungsmöglichkeiten vorgegeben werden und die Antwort von den Testpersonen eigenständig formuliert oder ergänzt wird, können bei gebundenen Antwortformaten vorgegebene Lösungsmöglichkeiten ausgewählt werden (ebd., S. 39 ff). Die vorgegebenen Lösungen sollen durch Distraktoren mit Ähnlichkeit zur richtigen Lösung gewährleisten, dass die Ratewahrscheinlichkeit bei der Erkennung der richtigen Antwortmöglichkeit minimiert wird (ebd., S. 45). Der Einsatz von diskret abgestuften Ratingskalen wie z. B. mehrstufige Likert-Skalen (vgl. ebd., S. 51) eignen sich weniger für die Erfassung von Fachwissen und werden daher in dieser Arbeit nicht weiter thematisiert.

Darüber hinaus ist bei der Itemkonstruktion zu beachten, dass systematische Fehlerquellen, die bei der Aufgabenbearbeitung durch Proband:innen auftreten können, durch das Aufgabendesign gering gehalten werden. Dazu zählen nach Jonkisz et al. (2012, S. 57) soziale Erwünschtheit, Akquieszenz sowie kognitive Prozesse. Soziale Erwünschtheit und Akquieszenz werden im Rahmen der Fachwissenstestung als Fehlerquellen mit geringerem Einfluss erachtet. Wenngleich die intrinsische Motivation der Schüler:innen zur Testteilnahme wünschenswert ist, sind die Möglichkeiten zur Abgabe sozial erwünschter Antworten in Bezug auf das Fachwissen durch Aufgabenformat und -inhalt an sich begrenzt. Wichtiger sind kognitive Prozesse, die bei der Beantwortung der Aufgaben einen Einfluss ausüben. Darunter sind u. a. die in Kapitel 2.6 diskutierten schwierigkeitsbestimmenden Merkmale für naturwissenschaftliche Aufgabenstellungen und allgemeine kognitive Stadien bei der Aufgabenbeantwortung (vgl. ebd., S. 58) zu betrachten. Während die Schwierigkeitsanforderung der Aufgabe die Stichprobe in gewünschter Weise nach Leistungsniveau differenziert, beeinflussen allgemeine kognitive Prozesse die Aufgabenbeantwortung möglicherweise unterschiedlich. Nach Podsakoff, MacKenzie, Lee und Podsakoff (2003) ist der kognitive Antwortprozess in die fünf Stadien *Verständnis*, *Abruf*, *Urteil*, *Antwortwahl* und *Antwortabgabe* unterteilt. Die Hauptfehlerquelle beim Verständnis resultiert aus einer möglichen Itemmehrdeutigkeit (vgl. Jonkisz et al., 2012, S. 58), durch die sich Proband:innen am Kontext anderer Items orientieren oder willkürlich ohne kontextuelle Orientierung antworten (vgl. ebd.). An das Aufgabenverständnis



schließt sich der Informationsabruf an, der z. B. durch intrapersonale Merkmale (Stimmungslage, Ermüdung o. Ä.) beeinflusst werden kann (vgl. ebd.). Nach dem Abruf der Informationen erfolgt das Urteil über die Angemessenheit, Vollständigkeit und Richtigkeit durch die getestete Person (vgl. ebd.). Die Abgabe der Antwort kann an dieser Stelle des Prozesses ebenfalls durch den Kontext bereits bearbeiteter Aufgaben beeinflusst werden, wenn Proband:innen das vorliegende Item nicht als unabhängig einschätzen (vgl. ebd.).

Zusammenfassend ist für die Itemkonstruktion zu berücksichtigen, dass bei der Formulierung auf sprachliche Verständlichkeit in Form klarer Satzstrukturen geachtet wird und die Formulierung bestenfalls ohne die Verwendung von Akronymen auskommt (vgl. ebd., S. 64 ff). Zudem muss ein Item eindeutig gestellt sein, sodass kein Interpretationsspielraum in Bezug auf die erwartete Antwort besteht (ebd., S. 65). Ein Item soll sich dabei ausschließlich auf eine einzelne Facette des Konstrukts beziehen (ebd.). Eine Verknüpfung von mehreren Facetten innerhalb eines Items erlaubt bei der Auswertung keine ausreichende Differenzierung, ob sich die Antwort der Testperson auf einen oder beide Aspekte bezieht (vgl. ebd., S. 66).

#### **5.4 Gütekriterien quantitativer Testinstrumente**

Die Definition von Gütekriterien für quantitative Testinstrumente steht nicht in unmittelbarem Zusammenhang mit einer Testtheorie (vgl. Stumpf, 1996, S. 416). Dieser Abschnitt gliedert sich in eine eingehendere Betrachtung der wissenschaftlichen Gütekriterien *Objektivität*, *Reliabilität* und *Validität* sowie in einen Überblick über weitere Gütekriterien, die im Rahmen der Testentwicklung berücksichtigt werden.

##### **5.4.1 Objektivität, Reliabilität und Validität**

Objektivität als Gütekriterium gliedert sich nach Moosbrugger und Kelava (2012, S. 8) in drei Facetten, die sich am Testungs- und Auswertungsprozess orientieren. An erster Stelle steht die *Durchführungsobjektivität*, zu deren Einhaltung gewährleistet sein muss, dass das Testergebnis nicht durch Testleiter:innen beeinflusst wird (vgl. Moosbrugger & Kelava, 2012, S. 9). Eine hohe Standardisierung der Testinstrumente und genaue Anweisungen in den Testheften tragen laut Moosbrugger und Kelava (2012, S. 9) dazu bei, konstante Bedingungen für die Durchführung zu schaffen. Daran knüpft die *Auswertungsobjektivität* an, die dann erfüllt ist, wenn das Ergebnis einer Testaufgabe nicht von der bewertenden Person beeinflusst wird (vgl. ebd., S. 9f). Während die Auswertungsobjektivität

bei geschlossenen Aufgabenformaten größtenteils durch den Aufgabencharakter gewährleistet wird, müssen für offene Aufgabenformate Regeln wie Musterlösung oder Erwartungshorizont vorliegen, deren Einhaltung sich zudem empirisch überprüfen lässt (vgl. ebd., S. 10). Abschließend ist die *Interpretationsobjektivität* zu beachten, die impliziert, dass unterschiedliche Anwender:innen des Testinstruments zur gleichen wissenschaftlichen Interpretation gelangen, wenn unterschiedliche Testpersonen dieselben Testwerte erreichen (vgl. ebd., S. 10).

Die Reliabilität definiert sich allgemein als Messgenauigkeitskoeffizient und ermöglicht eine Aussage über das Zustandekommen eines Testergebnisses in Bezug auf seine Messfehlerbehaftung (vgl. Moosbrugger & Kelava, 2012, S. 11). Der Bereich, in dem der Reliabilitätswert rangiert, liegt zwischen Null und Eins (vgl. Schermelleh-Engel & Werner, 2012, S. 122). Ein Wert von Null repräsentiert dabei ein Testergebnis, das ausschließlich aus Messfehlern besteht (ebd., S. 121 f). Ein Wert von Eins entspricht einer vollständig fehlerfreien Messung, bei der die Varianz des Messfehlers Null wird (vgl. ebd., S. 121). Spezifische Anpassungen und Möglichkeiten zur Berechnung des Reliabilitätswerts werden für die Klassische (Kapitel 5.5) sowie Probabilistische Testtheorie (Kapitel 5.6) im jeweiligen Abschnitt erläutert.

Die Generalisierbarkeit von Testergebnissen hängt von der Validität des Testinstruments ab (Moosbrugger & Kelava, 2012, S. 13). Die Anforderungen an die Testgültigkeit ergeben sich aus den vier Facetten *Inhalts-, Augenschein-, Konstrukt-, und Kriteriumsvalidität* (vgl. Moosbrugger & Kelava, 2012, S. 15). Diese Unterteilung erfordert die Betrachtung eines Testinstruments aus unterschiedlichen Perspektiven. Die Inhaltsvalidität setzt sich mit der Frage nach der tatsächlichen Erfassung eines Merkmals mittels einer Testskala bzw. der einzelnen Items auf inhaltlicher Ebene auseinander (vgl. Hartig, Frey & Jude, 2012, S. 148). Bei einer Operationalisierung von Merkmalen, bspw. der Prüfung des Fachwissens als Output der inhaltsbezogenen Kompetenzfacetten eines Bildungsplans, ist ein Testinstrument inhaltlich valide, wenn sich anhand der Testergebnisse verallgemeinernde Aussagen über die Lernzielerreichung der Schüler:innen treffen lassen (ebd., S. 149). Im Rahmen von Messungen schulischer Lernziele wird der Begriff *curriculare Validität* verwendet, wenn eine Repräsentation der Kompetenzfacetten des Bildungsplans durch das Testinstrument erfolgen soll (ebd., S. 150). Die Inhaltsvalidität eines Tests bzw. einer Testaufgabe wird optimalerweise von Expert:innen auf dem jeweiligen Gebiet eingeschätzt, um eine repräsentative Erfassung des untersuchten Merkmals zu gewährleisten (Moosbrugger & Kelava,

2012, S. 15). Die Augenscheinvalidität bezieht sich auf die Akzeptanz der Testpersonen und wird laut Moosbrugger und Kelava (2012, S. 15) häufiger mit der Inhaltsvalidität verwechselt bzw. vermengt. Das Testinstrument ist erst dann augenscheinlich valide, wenn es von Proband:innen als tauglich zur Erfassung eines Merkmals eingestuft wird (vgl. ebd., S. 15). Für einen Fachwissenstest wäre dieses Kriterium bspw. erfüllt, wenn die Schüler:innen davon ausgehen, dass das Testinstrument in der Lage ist, die jeweiligen inhaltsbezogenen Kompetenzfacetten eines Kompetenzteilbereichs abzubilden. In Ergänzung zu diesen eher qualitativ ausgerichteten Facetten will die Untersuchung der Konstruktvalidität einen empirischen Zugang mittels statistischer Kennwerte schaffen (vgl. ebd., S. 16). Moosbrugger und Kelava (2012, S. 16) unterscheiden dazu struktursuchende und strukturprüfende Zugänge. Als struktursuchende Zugänge eignen sich exploratorische Faktorenanalysen (vgl. ebd., S. 17). Als strukturprüfende Zugänge sind konfirmatorische Faktorenanalysen oder IRT-Skalierungen (s. hierzu Kapitel 5.6.1) verbreitet (vgl. ebd., S. 17). Die Kriteriumsvalidität lässt sich mit der Übertragbarkeit des Testergebnisses auf eine reale Gegebenheit prüfen (vgl. ebd., S. 18). Für einen Fachwissenstest gilt die Kriteriumsvalidität als bestätigt, wenn Proband:innen, die im Test gut abschneiden, entsprechend z. B. höhere Leistungen im geprüften Fach erbringen als Schüler:innen, die schlechtere Testergebnisse erzielen oder sich die Ergebnisse durch bekannte Zusammenhänge aus der Theorie erklären lassen (ebd., S. 18).

#### **5.4.2 Weitere Gütekriterien für die Testentwicklung**

Weitere für die Entwicklung und Durchführung psychometrischer Leistungstests relevante Gütekriterien sind *Skalierung*, *Normierung*, *Testökonomie* und *Unverfälschbarkeit* (Moosbrugger & Kelava, 2012, S. 18 ff). Darüber hinaus werden von Moosbrugger und Kelava (2012, S. 18 ff) *Nützlichkeit*, *Zumutbarkeit* und *Fairness* aufgeführt<sup>11</sup>. Ein Leistungstest soll die Stichprobe anhand des Testwertes charakterisieren (ebd., S. 18). Das Gütekriterium der Skalierung gilt dann als erfüllt, wenn damit die tatsächliche Differenzierung nach der Leistungsfähigkeit der Testpersonen erreicht wird (ebd., S. 18). Die Normierung dient der Generierung von Vergleichswerten, um getestete Personen gegenüberzustellen oder zu gruppieren (ebd., S. 19). Die Testökonomie bilanziert Kosten und Nutzen eines Tests und ist im Sinne eines Gütekriteriums erfüllt, wenn der Erkenntnisgewinn über das Instrument in einem angemessenen Verhältnis der benötigten Ressourcen

---

<sup>11</sup> Die Gütekriterien *Nützlichkeit*, *Zumutbarkeit* und *Fairness* werden der Vollständigkeit halber genannt. Sie werden aber im Rahmen dieser Arbeit aufgrund ihrer allgemeinverständlichen Formulierung nicht weiter ausgeführt. Eine Definition findet sich bei Moosbrugger und Kelava (2012, S. 18 ff).

steht (vgl. ebd., S. 21). Ein Testinstrument muss zudem das Gütekriterium der Unverfälschbarkeit erfüllen, indem es gewährleistet, dass Testpersonen durch ihr Verhalten im Test das Ergebnis nicht verzerren können (z. B. durch Abschreiben) (vgl. ebd., S. 23).

#### **5.4.3 Statistische Auswertungskriterien für quantitative Testinstrumente**

Die deskriptivstatistischen Gütekriterien für wissenschaftliche Testinstrumente geben Richtwerte vor, um die objektive Messung eines Merkmals vornehmen zu können. Im Folgenden werden die zentralen Kennwerte für Leistungstests *Lösungsquote* bzw. *Schwierigkeitsindex*, *Itemvarianz* sowie *Trennschärfe* kurz umrissen und eine Abgrenzung zwischen *Niveau-* und *Geschwindigkeitstest* vorgenommen.

Die Lösungsquote bzw. der Schwierigkeitsindex ist der Kennwert für die Verteilung des Antwortverhaltens der Stichprobe in Bezug auf ein einzelnes Item (Kelava & Moosbrugger, 2012, S. 78 ff). Hierbei ist eine Unterscheidung zwischen Geschwindigkeits- und Niveautests bezüglich der Behandlung von nicht gelösten Aufgaben notwendig. Während in Niveautests nicht zwischen ausgelassenen und unbearbeiteten Aufgaben differenziert wird, ist dies für Geschwindigkeitstest eine notwendige Unterscheidung, weil Aufgaben aufgrund von Zeitmangel unbearbeitet bleiben können (ebd.). Aus Zeitgründen unbearbeitete Aufgaben fließen bei Geschwindigkeitstests nicht in den zu bestimmenden Schwierigkeitsindex ein (ebd.). Für Niveautests gilt die Annahme, dass ausreichend Zeit für die Bearbeitung aller Aufgaben zur Verfügung steht. Eine Unterscheidung zwischen Aufgaben, die aus Zeit- oder Schwierigkeitsgründen nicht bearbeitet wurden, wird somit obsolet (ebd.). Die Lösungsquote resultiert dann aus dem Quotienten der richtig gelösten Aufgaben gegenüber der maximalen Aufgabenanzahl (ebd.).

Die Differenzierungsfähigkeit eines Items wird durch die Itemvarianz abgebildet (Kelava & Moosbrugger, 2012, S. 81 ff). Insgesamt zeigt sich, dass Items mit sehr hohen bzw. sehr niedrigen Schwierigkeitsgraden eine schlechte Differenzierung der Stichprobe ermöglichen (ebd.). Für dichotome Items mit einer Lösungsquote von 50 %, also einer mittleren Itemschwierigkeit, resultiert die beste Differenzierung (ebd.). Gleichermäßen müssen, wie o. a., die Randbereiche für eher gute bzw. eher schlechte Proband:innen ausreichend abgedeckt sein (ebd.).

Die Trennschärfe wird ermittelt, um die Differenzierungsfähigkeit eines einzelnen Items der Differenzierungsfähigkeit der kombinierten Items in einer Skala gegenüberzustellen (Kelava & Moosbrugger, 2012, S. 85 ff). Es wird dazu eine Korrela-

tion zwischen der Lösungsquote des spezifischen Items und der Gesamtlösungsquote aller Testitems für eine Testperson berechnet (ebd.). Der korrelative Zusammenhang zeigt auf, ob Personen, die das Item gelöst haben, anhand ihrer Merkmalsausprägung oder durch zufällige Effekte zu dieser Lösung gekommen sind (ebd.). Der Zahlenwert der Trennschärfe lässt sich anhand des soeben beschriebenen Sachverhalts interpretieren. Trennschärfen, die sich dem Wert  $r = 1$  nähern, deuten darauf hin, dass Personen mit hoher Merkmalsausprägung zur richtigen Lösung kommen, Personen mit niedriger Merkmalsausprägung jedoch nicht (ebd.). Eine Trennschärfe zwischen .4 und .7 gilt als optimal (vgl. ebd., S. 86). Je weiter sich ein Item einer Trennschärfe von  $r = 0$  annähert, desto weniger ist es in der Lage, die verschiedenen Niveaus der Merkmalsausprägungen gegeneinander abzugrenzen (ebd.). Wird die Trennschärfe negativ, kann der Wert im Fall einer Testung affektiver Merkmale (Persönlichkeitsmerkmale, o. Ä.) auf eine Invertierung eines Items hindeuten (ebd.). Im Fall von Geschwindigkeits- oder Niveautests bedeutet die Invertierung, dass Probanden mit insgesamt guten Testwerten dieses Item falsch oder nicht beantworten, Personen mit insgesamt schlechten Testwerten das Item aber richtig lösen (ebd.). Items mit einer negativen Trennschärfe sollten in jedem Fall genauer betrachtet und ggf. überarbeitet werden (ebd.).

## 5.5 Klassische Testtheorie

Die Klassische Testtheorie (KTT) ermöglicht als etabliertes und zugleich forschungsökonomisch sparsames Verfahren die statistische Fundierung von psychometrischen Testinstrumenten (z. B. Moosbrugger, 2012; Petrillo, Cano, McLeod & Coon, 2015). Sie erfüllt laut Petrillo et al. (2015, S. 25) die Anforderungen an die Entwicklung von deskriptiven Instrumenten und wird daher im Rahmen dieser Arbeit als Grundlage für die Testkonstruktion verwendet.

### 5.5.1 Grundannahmen der Klassischen Testtheorie

Die Klassische Testtheorie stützt sich auf drei Axiome: *Existenzaxiom*, *Verknüpfungaxiom* sowie *Unabhängigkeitsaxiom* und zusätzliche Grundannahmen, darunter die Unabhängigkeit der verwendeten Items sowie die Unabhängigkeit der untersuchten Personen (vgl. DeMars, 2018, S. 49 ff; Moosbrugger, 2012, S. 104 ff). Das Existenzaxiom definiert, dass der wahre Wert für ein Item, das von einem Probanden gelöst wird, existiert (Moosbrugger, 2012, S. 104). Das Verknüpfungaxiom unterstellt, dass jede vorgenommene Messung von einem zufälligen Messfehler beeinflusst wird und sich das erhaltene Ergebnis damit aus dem wahren Wert und dem Messfehler zusammensetzt (ebd., S. 105). Aus dem Unabhän-

gigkeitsaxiom geht hervor, dass keine Korrelation zwischen Messfehler und wahrem Wert bei beliebigen Items und beliebigen Personen vorliegt (vgl. ebd., S. 105). Die Unabhängigkeitsannahmen werden für die Berechnung der Reliabilität (s. Kapitel 5.5.2) benötigt (vgl. ebd., S. 106). Im Kontext der Unabhängigkeit von Items muss die Unkorreliertheit des Messergebnisses aus beliebigen Aufgaben des Tests gewährleistet sein (vgl. ebd., S. 105). Das bedeutet, dass Aufgaben so gestellt sein müssen, dass der Erfolg ihrer Bearbeitung nicht von anderen bereits gelösten oder nicht gelösten Items abhängt (ebd., S. 105). Die Unabhängigkeit von Personen ist dann erfüllt, wenn die Fehlerwerte bei der Messung unterschiedlicher Personen mit demselben Item nicht korrelieren (ebd., S. 105). Für die Testpraxis bedeutet dies, dass zu gewährleisten ist, dass die Proband:innen bei der Testbearbeitung z. B nicht voneinander abschreiben (vgl. ebd., S. 105).

### 5.5.2 Möglichkeiten der Reliabilitätsüberprüfung in der Klassischen Testtheorie

Die Reliabilität ist in der Klassischen Testtheorie das zentrale Gütekriterium und beschreibt die Messgenauigkeit eines Testinstruments in Form des Quotienten der Varianz des wahren Messwerts zur Varianz des erhaltenen Messergebnisses (Moosbrugger, 2012, S. 109 f). Schermelleh-Engel und Werner (2012) führen vier verbreitete Methoden zur Reliabilitätsbestimmung auf: *Retest-*, *Paralleltest-* und *Splithalf-Reliabilität* sowie *Interne Konsistenz* (vgl. Schermelleh-Engel & Werner, 2012, S. 122).

Die Bestimmung der Retest-Reliabilität setzt die zweimalige Durchführung eines Tests an derselben Stichprobe voraus, um Korrelationen der Testwerte berechnen zu können (Schermelleh-Engel & Werner, 2012, S. 122). Je höher die Korrelation des Ergebnisses der beiden Testzeitpunkte ausfällt, desto höher wird die Retest-Reliabilität (vgl. ebd., S. 123). Prinzipiell eignet sich dieses Verfahren für jede Art von Test, setzt aber voraus, dass Testpersonen für zwei Testungen zur Verfügung stehen (ebd., S. 123). Zudem sind Veränderungen bei den Merkmalsausprägungen zwischen den beiden Testzeitpunkten unter bestimmten Umständen problematisch für die Bestimmung der Retest-Reliabilität (ebd., S. 123). Es muss zwischen einer systematischen und unsystematischen Veränderung der wahren Werte unterschieden werden. Während „[s]ystematische Veränderungen der wahren Werte, die bei allen getesteten Personen gleich ausfallen, [...] grundsätzlich kein Problem bei der Bestimmung der Retest-Reliabilität [sind]“ (Schermelleh-Engel & Werner, 2012, S. 123), können unsystematische Veränderungen die Korrelation der Messergebnisse aus den beiden Testzeitpunkten negativ beeinflussen (ebd., S. 123). Unsystematische Veränderungen können bspw. Übungseffekte sein, die bei verschiedenen Testpersonen unterschiedlich stark

ausgeprägt sind (ebd., S. 123). Vor allem im Rahmen von Leistungstests ist dieser Einfluss als problematisch zu erachten (ebd., S. 123). Den gegenteiligen Einfluss können Erinnerungseffekte ausüben, indem Testpersonen bei beiden Testzeitpunkten exakt dieselben Antworten geben und dadurch die Reliabilität überschätzt wird (ebd., S. 124 f).

Eine Eliminierung der Erinnerungseffekte ermöglicht das Verfahren zur Ermittlung der Paralleltest-Reliabilität (ebd., S. 127). Es setzt jedoch voraus, dass sich zwei Parallelformen eines Testinstruments erstellen lassen (ebd., S. 126). Für Leistungstests, bei denen sich die Aufgaben leicht abwandeln lassen (z. B. durch Ändern von Zahlenwerten bei Rechenaufgaben) ist die Paralleltest-Reliabilität forschungsökonomisch sinnvoll einsetzbar (ebd., S. 126). Demgegenüber geht für Tests zur Erfassung von Persönlichkeitsmerkmalen oder Leistungstests mit komplexerer Ausgestaltung der Aufgaben ein hoher Aufwand für die Entwicklung gleichwertiger Items einher (ebd., S. 126). Die Ermittlung der Paralleltest-Reliabilität erfolgt in möglichst kurzem Zeitabstand an derselben Stichprobe (ebd., S. 126 ff). Der Reliabilitätswert ergibt sich dann aus der Korrelation der Testergebnisse der parallelen Tests (vgl. ebd., S. 126). Zunächst muss jedoch über ein geeignetes Prüfverfahren verifiziert werden, dass es sich tatsächlich um parallele Skalen handelt (ebd., S. 126 f). Dazu eignen sich einerseits deskriptive Kriterien wie identische Mittelwerte und Streuungen und andererseits das Verfahren der konfirmatorischen Faktorenanalyse, mit dem gleichzeitig die Paralleltest-Reliabilität über die Einführung einer gemeinsamen latenten Variablen ermittelt werden kann (vgl. ebd., S. 127). In der Praxis erweist sich das Paralleltest-Verfahren anfällig gegenüber unterschiedlichen Formulierungen von Items (ebd., S. 127), wodurch sich die Anwendbarkeit eher auf Skalen mit einfach abzuändernden, geschlossenen Aufgabenformaten beschränkt.

Umfangreichere Tests mit ausreichender Itemanzahl ermöglichen die Reliabilitätsbestimmung mit der Splithalf-Methode (Testhalbierung), zu deren Umsetzung eine Einteilung der Items aus der Gesamtskala in zwei möglichst gleichwertige Testhälften erfolgt (ebd., S. 128). Da sich die Reliabilität mit zunehmender Testlänge unter Erfüllung des Kriteriums der Itemhomogenität erhöht, kann zur Kompensation der halbierten Skala die Spearman-Brown-Korrektur berechnet werden, um die Reliabilität für den Gesamttest zu schätzen (vgl. ebd., S. 128). Die Bildung tatsächlich paralleler Testhälften ist nicht unproblematisch, da sich mit steigender Itemanzahl die Zuordnungsmöglichkeiten zu den jeweiligen Testhälften

ten schnell erhöhen (ebd., S. 130). Weiterhin problematisch ist die mögliche Heterogenität der Items in den Testhälften, die sich ebenfalls bei der Bestimmung der internen Konsistenz auswirkt (ebd., S. 130).

Die Vorgehensweise bei der Ermittlung der internen Konsistenz erweitert die Splithalf-Methode, indem sie jedes einzelne Item als separaten Testteil zur Erfassung eines Merkmals betrachtet (vgl. ebd., S. 130). Bei einer reinen Übertragung des Parallelitätskonzepts auf die Ebene einzelner Items müsste jedoch gewährleistet sein, dass die Fehlervarianzen gleich und die Items alle identisch schwer sind (vgl. ebd., S. 131). Vor allem bei Fachwissenstests ist eine Ausgestaltung mit Items identischer Schwierigkeit nicht wünschenswert, um Boden- und Deckeneffekte zu vermeiden. Variieren Items in ihrer Schwierigkeit, ersetzt das Konzept der *essentiellen  $\tau$ -Äquivalenz* das Parallelitätskonzept (vgl. ebd., S. 131). Die essentielle  $\tau$ -Äquivalenz postuliert, dass die wahren Werte eines Items sich aus dem tatsächlichen wahren Wert (z. B. Fachwissen eines Kompetenzbereichs) und einer itemspezifischen Konstante (z. B. Schwierigkeit der Aufgabe) additiv zusammensetzen (vgl. ebd., S. 131). Unter diesen Voraussetzungen wurde von Cronbach (1951) eine Berechnungsgrundlage geschaffen und sein Koeffizient *Cronbachs Alpha* hat sich weitgehend als Standard der Konsistenzanalyse in der klassischen Testtheorie etabliert (Schermelleh-Engel & Werner, 2012, S. 130). Cronbachs Alpha drückt aus, dass „die interne Konsistenz eines Test (sic) [...] umso höher [ist], je höher die Korrelationen zwischen den Items im Durchschnitt sind“ (Schermelleh-Engel & Werner, 2012, S. 131). Die Berechnung erfolgt anhand der nach Cronbach (1951, S. 321) und Schermelleh-Engel und Werner (2012, S. 131) adaptierten Gleichung 1.

$$\alpha = \frac{m}{m-1} \left( 1 - \frac{\sum_{i=1}^m V_i}{V_{ges}} \right) \quad \mathbf{1}$$

$V_i$  entspricht der Varianz eines Items an der Stelle  $i$ ,  $V_{ges}$  bezeichnet die Varianz des Gesamttests und  $m$  entspricht der Gesamtzahl der Testitems (Cronbach, 1951, S. 321; Schermelleh-Engel & Werner, 2012, S. 131). Für die Interpretation von  $\alpha$  gilt, dass ab einem Wert von  $\alpha \geq 0.7$  die Testskala eine zufriedenstellende Reliabilität aufweist (Moosbrugger & Kelava, 2012, S. 11). Die Genauigkeit der Reliabilitätsschätzung hängt dabei von den Items ab, bei deren Entwicklung gewährleistet werden muss, dass sie tatsächlich dasselbe Merkmal erfassen (Schermelleh-Engel & Werner, 2012, S. 132). Heterogene Items führen ggf. zu einer Unterschätzung der tatsächlichen Reliabilität (vgl. ebd., S. 132). Außerdem ist zu



beachten, dass eine hohe Interkorrelation der Items und ein damit zusammenhängender guter Cronbachs Alpha Wert kein Beleg für eine Eindimensionalität der Skala darstellt (z. B. Miller, 1995; Schermelleh-Engel & Werner, 2012, S. 133). Zur Eindimensionalitätsprüfung eignen sich Verfahren wie z. B. die Rasch-Modellierung (s. Kapitel 5.6.2) im Rahmen der Probabilistischen Testtheorie.

## 5.6 Probabilistische Testtheorie

Für weiterführende Analysen ist es empfehlenswert, die nach der Klassischen Testtheorie entwickelten Testinstrumente bei ausreichender Stichprobengröße mittels probabilistischer Methoden zu überprüfen und damit deren Analysepotenzial auszubauen (vgl. Petrillo et al., 2015).

### 5.6.1 Item Response Theory (IRT)

In der Item-Response-Theory gibt es mehrere Ansätze, die sich in dichotome und polytome Modelle unterscheiden (z. B. Andrich & Marais, 2019). Zu den bekanntesten polytomen Modellen gehört das *Partial-Credit-Modell* (Andrich & Marais, 2019, S. 233 ff). Der bekannteste dichotome Ansatz, der auch in größeren Schulleistungsstudien häufiger eingesetzt wurde, ist das *Rasch-Modell* (z. B. Sälzer, 2016, S. 43). Die Modelle unterscheiden sich hinsichtlich ihrer Anforderungen an die Testlänge und Stichprobengröße (Sahin & Anil, 2017). In dieser Abhandlung wird aufgrund der zu untersuchenden Zielgruppe und deren voraussichtlichen Größe von  $n < 200$  Proband:innen ausschließlich das Rasch-Modell angewandt und daher eingehender betrachtet.

### 5.6.2 Das Rasch-Modell<sup>12</sup>

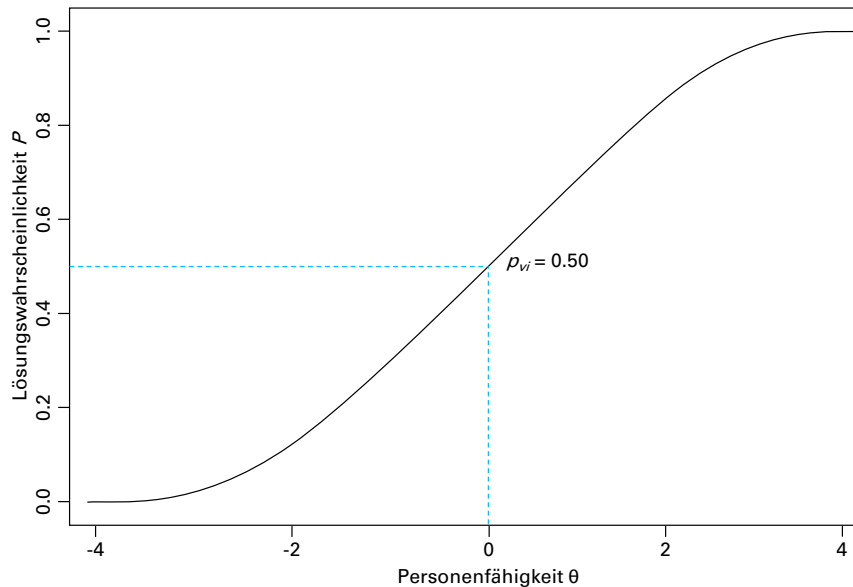
Die Rasch-Modellierung für dichotome Items erfolgt anhand der nachfolgend dargestellten Gleichung 2 (Koller, Alexandrowicz & Hatzinger, 2012, S. 9).

$$P(X_{vi} = x_{vi} | \theta_v, \varphi_i) = \frac{e^{x_{vi}(\theta_v - \varphi_i)}}{1 + e^{\theta_v - \varphi_i}} = p_{vi} \quad 2$$

Die Funktion beschreibt die Lösungswahrscheinlichkeit  $p_{vi}$  in Abhängigkeit der Personenfähigkeit  $\theta_v$ , sowie der Schwierigkeit  $\varphi_i$  des Items. Der Parameter  $x_{vi}$  entspricht hierbei dem Wert der Antwortvariablen  $X_{vi}$  und kann im dichotomen Fall  $x_{vi} = 0$  oder  $x_{vi} = 1$  betragen. Die Lösungswahrscheinlichkeit für eine Aufgabe

<sup>12</sup> Im Rahmen dieser Arbeit wird ein deskriptiver Überblick über das Rasch-Modell und dessen Ansatz zur Bestimmung und Überprüfung der Personen- und Itemparameter gegeben. Eine mathematische Auseinandersetzung mit der Modellgleichung und verschiedenen Schätzmethoden ist ausführlich in den Werken von Koller, Alexandrowicz und Hatzinger (2012), Strobl (2015) sowie Trendtel, Pham und Yanagida (2016) einsehbar.

wird durch die item characteristic curve (ICC; dt.: Itemcharakteristikkurve) abgebildet. Deren Verlauf wird durch die Exponentialfunktion bestimmt und steht somit in direktem Zusammenhang mit der Personenfähigkeit und der Itemschwierigkeit (Koller et al., 2012). Am Punkt der größten Steigung im Verlauf der ICC, an dem die Lösungswahrscheinlichkeit gleichzeitig  $p_{vi} = 0.50 = 50\%$  beträgt, wird die Itemschwierigkeit abgelesen (Abbildung 8) und das Item in der Rangfolge der Gesamtskala verortet (Andrich & Marais, 2019).



**Abbildung 8:** Exemplarische Darstellung einer Itemcharakteristikkurve (eigene Abbildung).

Es gilt demzufolge, dass Testteilnehmer:innen mit einer höheren Personenfähigkeit eine höhere Lösungswahrscheinlichkeit für das exemplarische Item in Abbildung 8 haben.

### 5.6.3 Annahmen und Bedingungen für die Modellgültigkeit

Die Gültigkeit des Rasch-Modells ist an grundlegende Annahmen und Bedingungen auf Item- und Testebene gebunden (Koller et al., 2012; Strobl, 2015). Als allgemeine Annahme gilt zunächst, dass Personen mit einer höheren Personenfähigkeit im Sinne einer suffizienten Statistik mit einer höheren Wahrscheinlichkeit mehr Items des Tests lösen können als Personen mit einer niedrigeren Personenfähigkeit (Strobl, 2015). Des Weiteren ist die lokale stochastische Unabhängigkeit der Testaufgaben zu berücksichtigen. Die lokale stochastische Unabhängigkeit unterstellt, dass „für eine Person mit einer bestimmten [Personenfähigkeit] die Wahrscheinlichkeit ein Item zu lösen nur von dem Item selbst und nicht von der Lösungswahrscheinlichkeit eines anderen Items abhängt“ (Koller et al., 2012, S. 17) und könnte in zwei möglichen Fällen verletzt werden. Zum einen, wenn Aufgabeninhalte unterschiedlicher Items aufeinander aufbauen und zum anderen,

wenn Proband:innen während der Testung voneinander abschreiben (Koller et al., 2012; Strobl, 2015). In diesem Kontext ist zusätzlich die spezifische Objektivität zu berücksichtigen, die „gewährleistet, dass Aussagen über die Fähigkeiten von zwei Personen nicht davon abhängen, anhand welcher Aufgabe sie verglichen werden“ (Strobl, 2015, S. 20). Tritt der Fall ein, dass eine Aufgabe die Personenfähigkeit zweier Gruppen mit gleichem Leistungsniveau unterschiedlich bewertet, ist *Differential Item Functioning* (DIF) zu unterstellen (vgl. Strobl, 2015). Das betroffene Item ist ungeeignet für die Verwendung im Testinstrument und muss überarbeitet oder ausgeschlossen werden (Strobl, 2015). Als weitere Bedingung ist die Eindimensionalität der Skala zu berücksichtigen. Eine Skala gilt dann als eindimensional, wenn gewährleistet wird, dass sie lediglich ein einzelnes, unabhängiges Konstrukt misst (ebd.). Es ist zu berücksichtigen, dass beispielsweise bei Aufgaben zur Überprüfung von technischer Kompetenz keine Störgrößen (u. a. sprachliche Kompetenz mit Bezug auf das Textverständnis oder eingeschränktes Farbsehen im Sinne der Barrierefreiheit) einen Einfluss auf die Lösungswahrscheinlichkeit ausüben (Koller et al., 2012; Strobl, 2015). Wenn die grundlegenden Bedingungen erfüllt sind, kann die Schätzung der Modellparameter vorgenommen werden.

#### **5.6.4 Schätzung der Itemschwierigkeit und Personenfähigkeit**

Die Schätzung der Itemschwierigkeit  $\varphi_i$  und der Personenfähigkeit  $\theta_v$  erfolgt anhand der in Kapitel 5.6.2 erläuterten Gleichung 1. Die beiden zunächst unbekannt Parameter werden auf Basis der Annahmen geschätzt, dass Personen, die viele Aufgaben lösen, eine hohe Personenfähigkeit  $\theta_v$  und Items, die oft gelöst werden, eine niedrige Itemschwierigkeit  $\varphi_i$  aufweisen. Zusätzlich spielt bei der Schätzung der Parameter die Bedingung der lokalen stochastischen Unabhängigkeit (s. Kapitel 5.6.3) eine Rolle, indem sie die Multiplikation von Wahrscheinlichkeiten grundlegend ermöglicht (Koller et al., 2012). Im übergeordneten Kontext des Rasch Modells wird von *Likelihood* gesprochen. Ziel der Rasch Modellierung ist es, die Likelihood als Kombination der Personenfähigkeit  $\theta_v$  mit der Itemschwierigkeit  $\varphi_i$  unter Multiplikation aller vorliegender Fälle im Datensatz zu maximieren. Zu diesem Zweck gibt es unterschiedliche Verfahren (s. Strobl, 2015, 26 ff), von denen in diesem Abriss die *Gemeinsame Maximum-Likelihood-Schätzung* (engl. *Joint Maximum Likelihood; JML*) als Grundlage sowie die *Bedingte Maximum-Likelihood-Schätzung* (engl. *Conditional Maximum Likelihood; CML*) aufgrund ihrer auswertungspraktischen Relevanz eingehender ausgeführt werden (Koller et al., 2012).

Die JML-Schätzung versucht anhand eines Suchalgorithmus iterativ die maximale Likelihood aus der bestmöglichen Kombination der Lösungswahrscheinlichkeit einzelner Items zu ermitteln (ebd.). Das Maximum der Likelihood-Funktion repräsentiert in diesem Kontext die bestmögliche Parameterschätzung für Personenfähigkeit und Aufgabenschwierigkeit (ebd.). Als Nachteil tritt bei dieser Schätzmethode das *Incidental Parameter Problem* (s. Koller et al., 2012, S. 36) auf, das durch die konstante Anzahl an Items aber einer möglicherweise variablen Stichprobengröße verursacht wird. Prinzipiell gilt, dass die Schätzung der Itemparameter mit größerer Stichprobe genauer wird und umgekehrt die Schätzung der Personenparameter durch die Erhöhung der Itemanzahl (ebd.). Dadurch bedingt ist die Schätzung der Itemparameter im JML-Ansatz genauer als die der Personenparameter (ebd.). Dem Incidental Parameter Problem könnte durch eine gleichzeitige Erhöhung der Itemanzahl begegnet werden, die in der Praxis forschungsökonomisch kaum umsetzbar ist (ebd.). Faktisch enthalten dadurch die Itemparameter in der JML-Methode immer einen Schätzfehler (vgl. Koller et al., 2012, S. 36).

Dem Incidental Parameter Problem begegnet die CML-Schätzung durch die Eliminierung des Einflusses der Personenfähigkeit auf die Itemschwierigkeit (ebd., S. 38). Entscheidend für die Schätzung der Itemparameter ist dann nur noch die Lösungshäufigkeit der einzelnen Aufgaben (vgl. Koller et al., 2012, S. 38-41). Diese Methode ermöglicht außerdem den Vergleich von Aufgaben unabhängig von den antwortenden Personen und erfüllt dadurch gleichzeitig die Bedingung der spezifischen Objektivität (Koller et al., 2012). Nachteilig wirken sich Items aus, die von keiner oder von allen Personen gelöst werden konnten (ebd., S. 41). Solche Items sollten in jedem Fall überarbeitet oder ausgeschlossen werden (ebd., S. 41).

Die Schätzung der Personenparameter bei der CML-Methode erfolgt nach einer Maximum-Likelihood-Methode, die der JML-Methode ähnelt (vgl. Koller et al., 2012, S. 43). Da die Itemschwierigkeit  $\varphi_i$  bei dieser Schätzmethode bereits bekannt ist und davon ausgegangen wird, dass die Anzahl der gelösten Items einer suffizienten Statistik entspricht, erhalten alle Personen, die die gleiche Anzahl an Items lösen eine identische Personenfähigkeit (Koller et al., 2012, S. 43). Dieser Ansatz verhindert dadurch die Vermehrung an Personenparametern, die sonst mit der Vergrößerung der Stichprobe einhergeht (ebd.). Insgesamt können bei dieser Schätzung nur so viele unterschiedliche Personenparameter vorkommen, wie Items in der Skala sind (plus ein weiterer), wodurch dem Incidental Parameter Problem entgegengewirkt wird (Koller et al., 2012, S. 42 ff).

Abschließend werden zwei Möglichkeiten zur Normierung der Itemparameter betrachtet. Die Wahl eines Bezugspunktes ist notwendig, da es sich bei der „Messung der Itemschwierigkeit [...] nicht um eine absolute Messung sondern nur um eine relative [handelt]“ (Koller et al., 2012, S. 42). Die erste Möglichkeit wählt willkürlich ein Item aus, dem der Nullwert zugeordnet wird und es damit ermöglicht, die Schwierigkeiten der anderen Aufgaben relativ zu diesem Bezugspunkt zu ermitteln (vgl. Koller et al., 2012, S. 42). Als zweite Option findet die *Summe-Null-Normierung* häufiger Anwendung, bei der die geschätzten Schwierigkeiten aller Items aufsummiert werden, sodass Null resultiert (ebd.). Für alle Items, deren Schwierigkeit kleiner als Null geschätzt wird, gilt die Annahme, dass sie leichter sind als die Aufgaben im Mittel (ebd.). Für eine geschätzte Schwierigkeit größer Null gilt äquivalent, dass sie schwieriger sind (ebd.). Eine Interpretation der Itemschwierigkeit ist somit ohne den Bezugspunkt eines Referenzitems möglich (ebd.).

### 5.6.5 Prüfung der Modellgültigkeit

Zur Prüfung der Modellgültigkeit unter Erfüllung der in 5.6.3 dargestellten Bedingungen eignet sich eine Reihe statistischer Methoden. Grundlegend wird für die Gültigkeit des Rasch-Modells auf Subgruppeninvarianz und Itemhomogenität untersucht (Koller et al., 2012; Strobl, 2015). Zusätzlich ist die Testreliabilität der modellierten Skala von Interesse. Im Folgenden werden gängige Testverfahren zur Überprüfung der Gültigkeit von Modellparametern überblicksweise umrissen.

Die Subgruppeninvarianz kann für das Modell global mittels *Andersen Likelihood Ratio Test (Andersen LRT)* überprüft werden (Koller et al., 2012). Der Andersen LRT vergleicht die geschätzten Modellparameter anhand zweier oder mehrerer Personengruppen, die durch ein zuvor festgelegtes Teilungskriterium aus der Gesamtstichprobe gebildet werden (ebd.). Die Aufteilung in Subgruppen überprüft die grundlegende Annahme, dass sich bei Gültigkeit des Rasch-Modells die geschätzten Itemparameter zwischen den Personen der Gruppen und in Relation zur Schätzung anhand der Gesamtgruppe nicht unterscheiden dürfen (Strobl, 2015). Die Gültigkeit des Rasch-Modells ist dann abzulehnen, wenn der Andersen LRT ein signifikantes Ergebnis liefert (Koller et al., 2012). Analog zum Verfahren der globalen Modellüberprüfung kann die Subgruppeninvarianz durch Teilung der Stichprobe auch für einzelne Aufgaben untersucht werden, um auffällige Items in Bezug auf DIF zu identifizieren (ebd.). Dazu eignet sich der *Wald-Test*, der die statistische Berechnungsgrundlage zur Untersuchung einer Divergenz

der Itemparameter zwischen den Subgruppen liefert (Strobl, 2015). Itemschwierigkeiten, die sich gemäß dem Ergebnis des Wald-Tests statistisch signifikant zwischen den Gruppen unterscheiden, können in einem iterativen Prozess aus der Skala ausgeschlossen werden (Koller et al., 2012). Koller und Kolleg:innen (2012) raten überdies dazu, in diesem Prozess nur das jeweils am stärksten von DIF betroffene Item auszuschließen und dann das Rasch-Modell neu zu berechnen, um sicherzustellen, dass andere Items nicht beeinflusst und ggf. fälschlicherweise eliminiert werden (vgl. Koller et al., 2012, S. 79).

Zur Untersuchung der Itemhomogenität kann der *Martin Löff Test* durchgeführt werden (Koller et al., 2012). Bei diesem Testverfahren werden im Gegensatz zum Andersen LRT die Items in Subgruppen eingeteilt. Die grundlegende Annahme ist, dass bei einer homogenen Skala die Schätzung der Personenparameter aus den gebildeten Subskalen annähernd identisch ausfallen sollte (ebd.). Da dieses Testverfahren laut Koller und Kolleg:innen (2012) aufgrund der Kombinationsmöglichkeiten der einzelnen Items der Subskalen für eine statistische Belastbarkeit der erforderlichen  $\chi^2$ -Verteilung relativ große Stichproben erfordert, erscheint es sinnvoll, zusätzlich zum Ergebnis des Martin Löff Tests Fit-Statistiken der einzelnen Items zu Rate zu ziehen (s. Boone, 2016). Dadurch können erstens einzelne inhomogene Items identifiziert und zweitens das Ergebnis des Martin Löff Tests verifiziert werden. In der Literatur finden sich Hinweise als Fit-Statistik den *Mean Square Infit (MNSQ Infit)* sowie den *Mean Square Outfit (MNSQ Outfit)* zu betrachten (Boone, 2016; Trendtel, Pham & Yanagida, 2016). Ames und Penfield (2015) schlagen vor, als Grenzwerte für den Itemfit minimal 0.5 und maximal 1.5 zu akzeptieren (vgl. Ames & Penfield, 2015, S. 45). Darüber hinaus legen sie nahe, dem Infit größere Bedeutung beizumessen, da dieser robuster gegen Verzerrung durch extreme Antwortmuster ist (ebd.). In der Literatur zu größeren Schulleistungsstudien finden sich konservativere Werte von minimal 0.8 und maximal 1.2 für den MNSQ Infit (Adams & Wu, 2003), die auch im Rahmen dieser Arbeit außer in ggf. inhaltlich zu begründenden Einzelfällen angelegt werden.

Sofern das Rasch-Modell nach der Überprüfung von Subgruppeninvarianz und Itemhomogenität seine Gültigkeit behält, können Reliabilitätswerte der erhaltenen Schätzwerte zur Betrachtung der Testgüte ermittelt werden (z. B. Adams, 2005; Wyrwal, 2020). Dazu werden die *Expected A Posteriori Reliabilität (EAP-Reliabilität)* (vgl. Trendtel et al., 2016, S. 203 f) und die *Weighted Likelihood Estimators Reliabilität (WLE-Reliabilität)* (vgl. Trendtel et al., 2016, S. 202 f) berechnet. Trendtel et al. (2016) erläutern:

„Für eine konkrete Person mit zugehörigem Testscore schätzt der WLE die wahrscheinlichste Fähigkeit. Der EAP hingegen sagt für eine (beliebige) Person in der Population für einen bestimmten Testscore sowie weitere Informationen [...] die wahrscheinlichste Fähigkeit einer solchen Person voraus“ (Trendtel et al., 2016, S. 205).

Eine Betrachtung der jeweiligen WLE- bzw. EAP-Reliabilität beinhaltet damit die Überprüfung der Itemparameter (basierend auf der EAP-Schätzung) sowie der Personenparameter (basierend auf der WLE-Schätzung) und kann die gewünschten Informationen über die Güte der Skala liefern (vgl. Wyrwal, 2020, S. 92).

### 5.6.6 Übersicht der Kennwerte des Rasch-Modells

**Tabelle 3:** Übersicht über die relevanten Kennwerte zur Rasch-Modellierung.

Kennwert	Zielbereich/Grenzwert	Literaturhinweis
Geschätzte Personenfähigkeit $\hat{\theta}$	Methode: Summe-Null-Normierung $\hat{\theta} > 0$ : Person ist fähiger als alle Personen im Durchschnitt. $\hat{\theta} < 0$ : Person ist unfähiger als alle Personen im Durchschnitt.	(Koller et al., 2012; Strobl, 2015)
Geschätzte Itemschwierigkeit $\hat{\phi}$	Methode: Summe-Null-Normierung $\hat{\phi} > 0$ : Item ist schwerer als Items im Durchschnitt. $\hat{\phi} < 0$ : Item ist leichter als Items im Durchschnitt.	(ebd.)
Andersen LRT	$p > .05$ : Keine Varianz zwischen den Subgruppen.	(ebd.)
Wald-Test	$p > .05$ : Kein DIF.	(ebd.)
Martin Löff Test <sup>a</sup>	$p > .05$ : Itemhomogenität bestätigt.	(ebd.)
MNSQ Infit	Maximalbereich: $0.5 < MNSQ\ Infit < 1.5$ Konservativer Bereich: $0.8 < MNSQ\ Infit < 1.2$	(Adams & Wu, 2003; Ames & Penfield, 2015)
EAP Rel <sup>b</sup>	$EAP > 0.7$	(Adams, 2005; Trendtel et al., 2016; Wyrwal, 2020)
WLE Rel <sup>b</sup>	$WLE > 0.7$	(ebd.)

<sup>a</sup> Statistische Belastbarkeit aufgrund notwendiger  $\chi^2$  Verteilung nur bei großen Stichproben gewährleistet. <sup>b</sup> Wert wird angegeben, um die Messfehlerbehaftung der Ergebnisse einer Skala einzuschätzen.

Zusammenfassend bietet Tabelle 3 einen Überblick über die wichtigen Modellkennwerte und deren Zielbereich.

## 5.7 Explorative Datengruppierung

Die Festlegung von Kompetenzniveaustufen erfolgt anhand von Standard-Setting Verfahren. Quantitative Analysen bieten gegenüber den qualitativen Ansätzen u. a. aufgrund ihrer Objektivität und aus forschungsökonomischen Gesichtspunkten einen Mehrwert (Diskussion verschiedener Ansätze in Kapitel 2.5). Sie sind zudem zeitlich flexibel im Analyseprozess einsetzbar und mehrfach durchführbar, um Ergebnisse iterativ zu generieren; ein Vorteil bei kleineren Stichproben, die über mehrere Erhebungswellen ausgebaut werden. Als quantitative Methode des Standard-Setting eignet sich daher das nachfolgend vorgestellte Verfahren der Clusteranalyse. Sie wurde bereits 1999 von Sireci et al. für Standard Settings eingesetzt, ist jedoch nach wie vor in aktuelleren Studien anzutreffen (s. Kapitel 2.5).

### 5.7.1 Clusteranalyse<sup>13</sup>

Hinter dem Sammelbegriff *Clusteranalyse* stehen verschiedene technische Ansätze (Backhaus, Erichson, Weiber & Plinke, 2015, S. 455 ff; Bortz & Schuster, 2010, S. 453 ff), die über Mustererkennung versuchen, eine Stichprobe in mehrere in sich möglichst homogene Subgruppen einzuteilen, die ihrerseits eine maximale Distanz zueinander aufweisen (vgl. Wiedenbeck & Züll, 2010, S. 525). Zur Mustererkennung in den Daten existieren diverse Algorithmen (z. B. Theodoridis & Koutroumbas, 2010), deren Anwendbarkeit jedoch von einigen zu berücksichtigenden Parametern abhängt (vgl. Wiedenbeck & Züll, 2010, S. 526). In der Forschungspraxis sind *hierarchische*<sup>14</sup> und *partitionierende* Verfahren (Backhaus et al., 2015, S. 476) etabliert. Wiedenbeck und Züll (2010, S. 526) schlagen für ausschließlich quantitative Variablen den *k-Means* Algorithmus vor, der den partitionierenden Verfahren angehört. Dessen Einsatz ist unter der Bedingung möglich, dass die Anzahl an Clustern bereits im Vorfeld eingegrenzt werden kann (ebd., S. 534). Diese Eingrenzung lässt sich entweder durch theoretische Fundierung ableiten, mittels statistischer Kennwerte abschätzen oder als Kombination der beiden Ansätze vornehmen (ebd., S. 526).

---

<sup>13</sup> Dieser Abschnitt liefert einen Überblick über die Clusteranalyse in Bezug auf die explorative Einteilung von Kompetenzniveaustufen mit Fokus auf den k-Means Algorithmus. Eine vertieftere (mathematische) Auseinandersetzung findet sich z. B. bei Wiedenbeck und Züll (2010).

<sup>14</sup> Hierarchische Verfahren werden in dieser Arbeit nicht eingesetzt und daher nicht näher spezifiziert. Eine Übersicht liefern Wiedenbeck und Züll (2010, S. 529 ff).



Milligan und Cooper (1985, S. 163 ff) stellen in einer Simulationsstudie 30 verschiedene statistische Indizes zur Bestimmung der Clusteranzahl gegenüber, die von Charrad, Ghazzali, Boiteau und Niknafs (2014) methodisch und anwendungsbezogen zugänglich gemacht wurden. Nach ihrem Verfahren werden alle Indizes berechnet und über die Mehrheitsregel eine Clusteranzahl vorgeschlagen, da die Indizes meist kein einheitliches Ergebnis liefern (Sedlmeier & Burkhardt, 2021, S. 315). Diese Vorgehensweise ist v. a. dann empfehlenswert, wenn sich die Anzahl der Cluster nicht aus der Theorie begründen lässt (ebd., S. 315). Es zeigte sich in der o. a. Simulationsstudie von Milligan und Cooper außerdem, dass das für metrische Daten geeignete *Calinski/Harabasz Kriterium* (Calinski & Harabasz, 1974) zu 90 % und damit am besten in der Lage war, die Gruppenstruktur korrekt zu definieren (vgl. Backhaus et al., 2015, S. 496).

Der k-Means Algorithmus arbeitet unter der Vorgabe einer Clusteranzahl nach dem Austauschverfahren (Backhaus et al., 2015, S. 477). Ausgehend von den jeweiligen Mittelwerten einer Startpartition und deren Fehlerquadratsumme (Varianzkriterium) versucht der Algorithmus iterativ durch Verschieben von Fällen zwischen den Gruppen das Varianzkriterium zu verringern (vgl. ebd., S. 477). Eine Verlagerung in eine andere Gruppe erfolgt dann, wenn der betrachtete Einzelfall zu einer maximalen Verringerung der Fehlerquadratsumme im Iterationszyklus führt (vgl. ebd., S. 477). Nach der jeweiligen Neuberechnung der Mittelwerte folgen so lange weitere Iterationszyklen bis keine Verbesserung des Varianzkriteriums mehr erzielt wird und die Clusterung abgeschlossen ist (vgl. ebd., S. 477).

## 5.8 Untersuchung von Zusammenhängen

Zusammenhänge zwischen mindestens intervallskalierten Variablen können anhand der Produkt-Moment-Korrelation untersucht werden (Sedlmeier & Renkewitz, 2018, S. 242 f). Die Produkt-Moment-Korrelation ermöglicht eine Aussage zu Stärke und Richtung eines linearen Zusammenhangs zwischen zwei Merkmalsausprägungen (vgl. ebd., S. 242 f) und berechnet sich gemäß Gleichung 3 als Quotient aus der Kovarianz und dem Produkt der Standardabweichungen  $sd_x$  und  $sd_y$  (entnommen und modif. aus Sedlmeier & Renkewitz, 2018, S. 220-221).

$$r_{xy} = \frac{\frac{1}{n} \cdot \sum_i^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{sd_x \cdot sd_y} = \frac{\text{cov}(x, y)}{sd_x \cdot sd_y} \quad \mathbf{3}$$

Die Spannweite des Korrelationskoeffizienten  $r$  rangiert zwischen  $-1 \leq r \leq +1$ , wobei  $-1$  einem perfekt negativen und  $+1$  einem perfekt positiven Zusammenhang entspricht (ebd., S. 243). Bei einem Wert von  $r = 0$  besteht kein Zusammenhang

zwischen Variablen. Zur Interpretation der Stärke des Zusammenhangs im Kontext der Sozialwissenschaften wird überwiegend die von Cohen (1988) vorgeschlagene Konvention verwendet (Sedlmeier & Renkewitz, 2018, S. 223). Gemäß der Konvention entspricht ein Wert von  $|r| \approx 0.1$  einem schwachen Zusammenhang (vgl. ebd., S. 223). Werte von  $|r| \approx 0.3$  und  $|r| \approx 0.5$  entsprechen einer mittleren bzw. einer hohen Korrelation (vgl. ebd., S. 223). Eine nennenswerte Einschränkung weisen Korrelationen bezüglich ihrer Aussagekraft zu Ursache-Wirkungs-Beziehungen auf (vgl. ebd., S. 243). Korrelationen können keinen Aufschluss darüber geben, ob die Ausgangsvariable die Zielvariable ursächlich beeinflusst (vgl. ebd., S. 243). Sofern eine Korrelation zwischen zwei Variablen belegt ist und theoriegeleitet eine Vermutung zur Wirkrichtung besteht, kann mittels linearer Regression überprüft werden, inwiefern die Zielvariable durch die Ausgangsvariable vorhergesagt werden kann (ebd., S. 246). Dadurch wird das Problem der Kausalbeziehung zwar nicht gelöst, es lassen sich jedoch auf Basis der Modellannahme Aussagen zu Ereignissen treffen, die in der Zukunft liegen (z. B. Schulabschlussnote als Prädiktor für spätere Studienleistungen) (vgl. ebd., S. 265).

Mittels linearer Regression kann der gerichtete Zusammenhang einer Ausgangsvariablen (Prädiktorvariable) mit einer Zielvariablen (Kriteriumsvariable) untersucht werden (Sedlmeier & Renkewitz, 2018, S. 246). Die allgemeine Form der Regressionsgeraden entspricht Gleichung 4 (entnommen aus Sedlmeier & Renkewitz, 2018, S. 286).

$$\hat{y} = bx + a \quad 4$$

Steigung und Achsenabschnitt in der Geradengleichung werden als Regressionskoeffizienten  $b$  und  $a$  bezeichnet (vgl. ebd., S. 286). Die Steigung der Geraden (Koeffizient  $b$ ) gibt an, wie stark sich die vorhergesagte Kriteriumsvariable  $\hat{y}$  durch einen Anstieg oder Abfall der Prädiktorvariablen um eine Einheit verändert (vgl. ebd., S. 286). Die Bestimmung der Regressionsgeraden erfolgt nach der Methode der kleinsten Quadrate<sup>15</sup> (vgl. ebd., S. 286). Zur besseren Vergleichbarkeit kann durch z-Standardisierung der Eingangsvariablen anstelle des Koeffizienten  $b$  ein standardisiertes Regressionsgewicht  $\beta$  erhalten werden (vgl. ebd., S. 286). Neben der Größe des Einflusses ist außerdem von Interesse, mit welcher Güte sich die Zielvariable durch die Ausgangsvariable vorhersagen

---

<sup>15</sup> „Die Gerade wird so gewählt, dass die Summe der quadrierten Abweichungen zwischen den vorhergesagten  $\hat{y}$ -Werten und den tatsächlichen  $y$ -Werten minimiert wird“ (Sedlmeier & Renkewitz, 2018, S. 286).

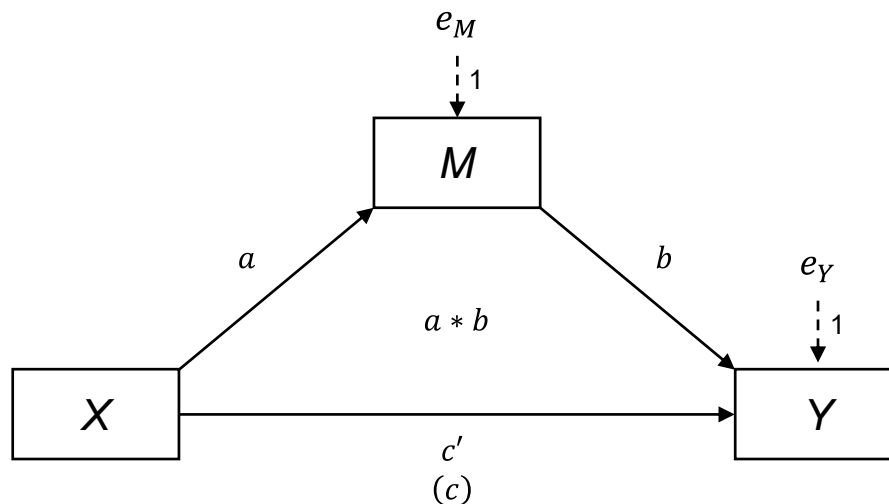
lässt (ebd., S. 247). Als Gütemaß der aufgeklärten Varianz durch die Regressionsanalyse wird der Determinationskoeffizient  $R^2$  betrachtet (vgl. ebd., S. 287).

Sobald mehrere Prädiktoren für die Vorhersage eines Kriteriums eingesetzt werden sollen, eignet sich die multiple Regressionsanalyse (vgl. ebd., S. 287) bzw. das allgemeine lineare Modell (vgl. ebd., S. 715). Damit geht die Erweiterung der Regressionsgleichung 4 einher und ergibt unter Einbezug weiterer Prädiktorvariablen sowie eines Fehlerterms  $e_i$  den wahren Wert der Kriteriumsvariablen  $y_i$  gemäß Gleichung 5 (entnommen aus Sedlmeier & Renkewitz, 2018, S. 715).

$$y_i = b_1 \cdot x_{1i} + b_2 \cdot x_{2i} + \dots + b_m \cdot x_{mi} + a + e_i \quad 5$$

Als Gütemaß wird ebenfalls der multiple Determinationskoeffizient  $R^2$  betrachtet, der den Anteil an aufgeklärter Varianz des Kriteriums durch sämtliche Prädiktoren angibt (vgl. ebd., S. 287). Auf Basis des allgemeinen linearen Modells eröffnen sich weitere Möglichkeiten der Zusammenhangsanalyse direkter (Moderationsanalyse) und indirekter Einflüsse (Mediationsanalyse) auf ein Kriterium (vgl. Hayes, 2022, S. 27; Sedlmeier & Renkewitz, 2018, S. 766). In dieser Arbeit soll der indirekte Einfluss fluider Intelligenz über das Vorwissen auf Fachwissen (s. Kapitel 4.1.4) untersucht werden. Im Folgenden wird daher die Mediationsanalyse ausführlicher beschrieben.

Das grundlegendste Mediatormodell enthält drei Variablen (Prädiktorvariable, Kriteriumsvariable und Mediatorvariable; Abbildung 9) und untersucht neben dem direkten Zusammenhang des Prädiktors  $X$  mit dem Kriterium  $Y$  den indirekten Einfluss des Prädiktors  $X$  auf das Kriterium  $Y$  über den Mediator  $M$  (vgl. Hayes, 2022, S. 79 ff). Die jeweiligen Pfadkoeffizienten werden im methodischen Ansatz von Hayes (2022) auf Basis linearer Regressionsmodelle berechnet (ebd., S. 80). Eine Mediationsanalyse ist erst dann sinnvoll, wenn ein bestätigter Gesamtzusammenhang von  $X$  und  $Y$  vorliegt (ebd., S. 81). Über diesen Zusammenhang gibt neben vorliegenden Korrelationen der Pfadkoeffizient  $c$  Aufschluss (ebd., S. 114). Sofern die Mediationsanalyse auf Basis eines bestehenden direkten Zusammenhangs zwischen Prädiktor- und Kriteriumsvariable sinnvoll erscheint, erfolgt die Aufnahme des Mediators in das Modell und es werden die Pfadkoeffizienten  $a$  und  $b$  ermittelt (ebd., S. 84ff).



**Abbildung 9:** Mediatormodell mit drei Variablen (angelehnt an Hayes, 2022, S. 85).

Der Pfad  $a$  beschreibt den Effekt des Prädiktors auf den Mediator (ebd., S. 86). Der vom Mediator auf die Kriteriumsvariable ausgehende Effekt wird durch den Pfad  $b$  dargestellt (ebd., S. 86). Der indirekte Einfluss des Prädiktors über den Mediator wird durch die Multiplikation der beiden einzelnen Pfadkoeffizienten  $a*b$  erhalten (vgl. ebd., S. 112). Der direkte Einfluss von  $X$  auf  $Y$  unter Einbezug von  $M$  wird durch den Pfadkoeffizienten  $c'$  repräsentiert (ebd., S. 85).  $e_M$  und  $e_Y$  entsprechen den Fehlern bei der Schätzung von Mediator- und Kriteriumsvariable (vgl. ebd., S. 84). Der regressionsbasierte Ansatz bietet zwei Vorteile. Erstens besteht die Möglichkeit über z-Standardisierung ebenfalls standardisierte Pfadkoeffizienten zu erhalten und zweitens kann die Varianzaufklärung des Modells mittels Determinationskoeffizient angegeben werden (Sedlmeier & Renkewitz, 2018, S. 773).

### 5.9 Zwischenfazit zu den methodischen und statistischen Grundlagen

Zunächst werden die Lehrpersonen als Bildungsexpert:innen in die Testentwicklung sowie die Analyse der Rahmenbedingungen einbezogen und erhalten im Gegenzug durch die entwickelten inhaltlich validierten Instrumente im Studienverlauf formative Rückmeldung zum Wissensstand der Lernenden. Die in der formativen Rückmeldung aufgezeigten Optimierungspotenziale können in der Schulpraxis in die Unterrichtsgestaltung und -durchführung einfließen. Von den erhaltenen Forschungsergebnissen profitieren im Sinne des Design-Based-Research Ansatzes (Anderson & Shattuck, 2012; Reinmann, 2005) somit die Bildungsforschung und die schulische Praxis gleichermaßen. Für die Lehrpersonenbefragung wird auf die Methode des teilstandardisierten Expert:inneninterviews zurückgegriffen. Im Zuge der Auswertung wird eine inhaltlich strukturierende Vor-

gehensweise gewählt, die sich an den Hauptkategorien aus dem Interviewleitfaden orientiert. Diese deduktiven Kategorien werden im Auswertungsverlauf ggf. induktiv ergänzt.

Die Abhandlungen der Aspekte zur Testgestaltung in Kapitel 5.3 sowie die Gegenüberstellung der Klassischen Testtheorie (Kapitel 5.5) mit der Probabilistischen Testtheorie (Kapitel 5.6) enthalten wichtige Implikationen zur Vorgehensweise bei der Testentwicklung im Rahmen dieser Arbeit. Auf Basis der beschriebenen Gütekriterien in Kapitel 5.4 werden Niveautests mit dichotomen Items entwickelt, um eine stringente Auswertung ohne Notwendigkeit der Differenzierung zwischen ausgelassenen und zeitlich nicht machbaren Items zu gewährleisten, die in der Testdurchführung aufwändig gekennzeichnet werden müssten. Für den Einsatz der Testinstrumente im Feld bedeutet diese Vorgehensweise, dass zeitlich der Zuschnitt des Testumfangs auf maximal 90 Minuten (zwei Schulstunden) erfolgt und die Anzahl der Aufgaben so festgelegt wird, dass diese von den Schüler:innen in der vorgegebenen Zeit gut bearbeitet werden können. Durch die gewählte Dichotomisierung vereinfacht sich der Zusammenhang zwischen Itemschwierigkeit und Itemvarianz und ermöglicht die nachfolgend dargelegte Vorgehensweise in Bezug auf die Aufgabenauswahl sowie eine Rasch-Modellierung der Testskalen.

Da erstens die Differenzierungsfähigkeit von Items bei sehr geringen sowie sehr hohen Lösungsquoten schlecht ausfällt, werden keine Items mit einer niedrigeren Lösungsquote als 20 % bzw. höheren Lösungsquote als 80 % in eine etwaige Skala einbezogen. Die im Rahmen dieser Arbeit zugrunde gelegte Anforderung an die Testskala orientiert sich außerdem an diesem Lösungsquotenbereich, um zweitens für den angedachten Einsatzzweck ausreichend zwischen Testteilnehmer:innen mit hohen bzw. niedrigen Merkmalsausprägungen des Fachwissens zu unterscheiden; vorausgesetzt diese Items weisen gleichzeitig eine hinreichend hohe Trennschärfe auf. Für dichotome Items in diesem Lösungsbereich kann drittens durch den direkten Zusammenhang mit der Lösungsquote eine ausreichende Differenzierung der Stichprobe bezüglich der Itemvarianz bei gleichzeitiger Abdeckung der extremeren Randbereiche angenommen werden. Aus dem in Kapitel 5.4.3 beschriebenen begünstigenden Zusammenhang hoher Itemvarianz in Bezug auf hohe Trennschärfen wird viertens angenommen, dass die einzelnen Items in der gewählten Spanne der Lösungsquoten ausreichend hohe Trennschärfen erzielen. Die Trennschärfen werden zudem für jedes einzelne Item untersucht. Der Grenzwert für den Ausschluss eines Items wird bei einer Trennschärfe von  $r < .20$  festgelegt.

Die Entwicklung der Testinstrumente zur Erfassung des Fachwissens erfolgt aufgrund der Pilotierungsstichprobe mit einem Stichprobenumfang von  $n < 100$  zunächst nach der Klassischen Testtheorie. In einer forschungsökonomischen Gegenüberstellung der vorgestellten Verfahren zur Reliabilitätsbestimmung in der Klassischen Testtheorie sind die beiden Ansätze Retest und Paralleltest als aufwändig einzuschätzen. Eine Unterschätzung der Reliabilität bei heterogenen Items ist zwar auszuschließen, liefert aber bei der Erfassung von Fachwissen keinen Mehrwert gegenüber des Zeit- oder Entwicklungsaufwands. Die interne Konsistenzprüfung ist der Splithalf-Methode vorzuziehen, weil jedes Item einzeln betrachtet wird und keine gleichen Testhälften gebildet werden müssen, was ggf. einen erhöhten Entwicklungsaufwand für Aufgaben mit sich bringen würde. Für die Entwicklung und Erprobung von Testinstrumenten zur Erfassung von Fachwissen an kleinen Pilotierungsstichproben ( $n < 100$ ) eignet sich damit das Verfahren der internen Konsistenzprüfung und wird in der Testentwicklung in dieser Arbeit zur Erstellung der deskriptiven Fachwissenstests angewandt.

Da die interne Konsistenzprüfung keine Aussagen zur Dimensionalität einer Skala treffen kann, ist es für weiterführende Analysen (z. B. Kompetenzniveau-modellierung) unerlässlich, das Analysepotenzial der Testinstrumente über ihren deskriptiven Charakter hinaus auszubauen. Dazu wird bei einer ausreichend großen Stichprobe ( $n > 100$ ; s. z. B. Chen et al., 2014) die Skalierung anhand des Rasch-Modells als probabilistische Methode vorgenommen. Neben der geschätzten Personenfähigkeit, die die Ausprägung der inhaltsbezogenen Kompetenzen repräsentiert, können über die Modelle Aussagen zu Wissensdimensionen von Vor- und Fachwissen getroffen werden.

Kompetenzniveau-modellierung setzt eine Segmentierung der Stichprobe in diskrete Gruppen voraus. Zur Gruppenfindung wird im Rahmen dieser Arbeit aus zwei Gründen die k-means Clusteranalyse eingesetzt. Zum einen handelt es sich um ein quantitatives Verfahren, das sich auf die vorliegenden Daten stützt und somit rein objektiv agiert. Zum anderen können verschiedene theoretisch begründbare Gruppierungsansätze verglichen werden. Andernfalls müsste die Niveaustufenfestlegung durch Expert:innen in qualitativen Verfahren erfolgen. Die Literatur gibt jedoch entscheidende Hinweise, dass die qualitative Vorgehensweise bei gleichzeitig höherem Aufwand keine nennenswert besseren Ergebnisse als das quantitative Verfahren der Clusteranalyse liefert (s. Kapitel 2.5).

Anschließend an die Kontrastierung durch Kompetenzniveaus interessiert die Überprüfung theoriebasierter Zusammenhänge zwischen fluider Intelligenz, Vorwissen und Fachwissen. Ausgehend von nachzuweisenden Zusammenhängen

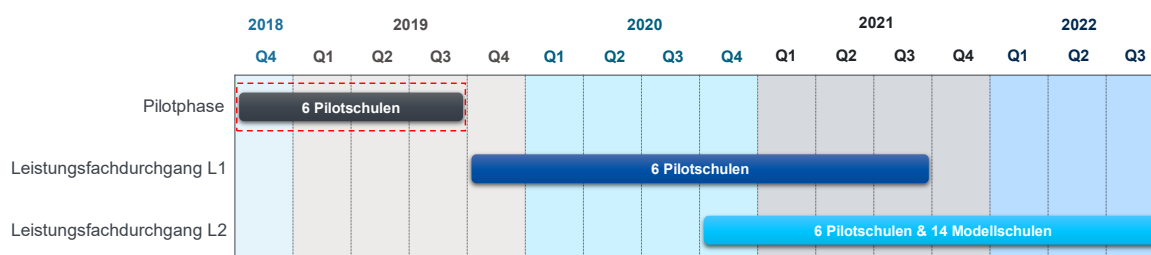
der Merkmale durch Korrelationsanalysen (H1, H2 und H3), eignet sich die regressionsbasierte Mediationsanalyse zur Überprüfung der Hypothese H4, die den indirekten Einfluss von fluiden Intelligenz über Vorwissen auf das Fachwissen der Lernenden im Leistungsfach unterstellt.





## 6 Inhaltliche und kontextuelle Rahmenbedingungen des Leistungsfachs NwT

Nach dem Angebots-Nutzungs-Modell (Seidel, 2014) sind für den Bildungsoutput mehrere Einflussfaktoren von Bedeutung. Darunter die inhaltlichen, formalen und kontextuellen Rahmenbedingungen des Unterrichts auf der Angebotsebene. Die Perspektive der Lehrpersonen gibt Aufschluss darüber, welche organisatorischen, methodisch-didaktischen und thematischen Schwerpunkte von den Lehrpersonen festgelegt werden.



**Abbildung 10:** Verortung der Studie zu den Rahmenbedingungen im Erhebungsablaufplan.

### 6.1 Untersuchungsdesign und Stichprobe

In einer Vollerhebung (Schuljahr 2018/19; Pilotphase) werden alle am Schulversuch teilnehmenden Lehrpersonen ( $n = 13$ , weiblich = 3, männlich = 10) zu inhaltlichen, kontextuellen und methodisch-didaktischen Rahmenbedingungen, zu ihrer Motivation und zu Qualifikationsmaßnahmen befragt. In Leitfadeninterviews werden sie mit teilstandardisierten Fragestellungen aufgefordert, zu den in Tabelle 4 aufgelisteten deduktiven Kategorien Stellung zu nehmen.

**Tabelle 4:** Kurzdefinition der deduktiven Kategorien des Interviewleitfadens.

Kategorie	Kurzbeschreibung
<b>Motivation und fachliche Voraussetzungen</b>	Erfassung der intrinsischen und extrinsischen Motive sowie geforderte fachliche Voraussetzungen, um den Kursstufenunterricht durchführen zu dürfen. Dabei wird v. a. die Notwendigkeit einer bestimmten Fächerkombination fokussiert.
<b>Unterrichtsorganisation</b>	Vorgegebene schulische Voraussetzungen und Rahmenorganisation. Besonders Teamlösungen und mögliche Lehrerwechsel zum Halbjahr/Schuljahr werden eruiert.
<b>Unterrichtsvorbereitung</b>	Vorgehensweise bei der Konzeption von Unterrichtseinheiten und bevorzugt verwendete Quellen. Des Weiteren interessiert die Korrespondenz und Kooperation schulintern und pilotschulübergreifend.

<b>Spiralcurriculare Kompetenzteilbereiche</b>	Kompetenzteilbereiche, die in der Mittelstufe NwT unterrichtet werden. Besonderes Augenmerk entfällt auf solche, die auf den Kursstufenunterricht vorbereiten und darauf, welche Implikationen daraus für den Unterricht in der Mittelstufe resultieren.
<b>Unterrichtsinhalte</b>	Fachliche Unterrichtsinhalte und ingenieurwissenschaftliche Bezüge der einzelnen Kompetenzteilbereiche im Kursstufenunterricht sollen genannt werden. Außerdem wird eine Einschätzung zur Aufteilung von theoretischen und praktischen Anteilen erfragt. In diesem Kontext werden ebenfalls die Dauer und Art der Projekte im Unterricht und mögliche Exkursionen beleuchtet.
<b>Unterrichtsmethoden</b>	Die Lehrpersonen beschreiben, welche Unterrichtsmethoden sie für geeignet halten und bevorzugt im Unterricht einsetzen wollen.
<b>Ausstattung der Schule und verfügbare Ressourcen</b>	Besonderer Fokus liegt auf den vorhandenen Geräten, Maschinen, Werkstoffen und dem Budget, das den Lehrpersonen der Pilotschulen zur Verfügung steht. Ein kurzer Umriss der räumlichen Situation und der Wunsch nach zusätzlichen Ressourcen und Optimierungsbedarfen soll vorgenommen werden.
<b>Kompetenzerwerb der Schüler:innen</b>	In Anlehnung an die definierten Kompetenzbereiche in der EPA-Technik (Autorengruppe der Kultusministerkonferenz, 2006) soll dargestellt werden, wie der Unterricht zum Erwerb ebendieser beiträgt und wie die Kompetenzen nachgewiesen werden sollen.
<b>Prüfungsformate</b>	Es werden Prüfungsformate für die Schüler:innen dargelegt und eine Gewichtung zwischen praktischen und schriftlichen Prüfungen vorgenommen. Ferner werden Aufgabenformate für schriftliche und praktische Prüfungen eruiert.
<b>Qualifikation der Lehrpersonen</b>	Lehrpersonen, die kein NwT-Studium absolviert haben, sollen ihre spezifische fachwissenschaftliche Ausbildung in NwT umreißen. Dabei sind besonders Zusatzqualifikationen, die spezifisch für die Kursstufe erworben wurden und deren zeitlicher Umfang von Interesse.
<b>Wahrgenommene Unterstützung</b>	Unterstützungsbedarfe (v. a. beim Fortbildungsangebot) und Optimierungspotenziale bei der Unterstützung durch verschiedene institutionelle Vertreter (u. a. Schulleitung, Ministerium für Kultus, Jugend und Sport Baden-Württemberg) sollen erläutert werden.

Die aus der Theorie hergeleiteten Kategorien decken die Angebotsebene in den Bereichen (1) Kontext des Bildungssystems, (2) Kontext der Schule, des Kollegiums und des Fachs, (3) Lehrerkompetenzen (in Form von Selbstauskunft) und (4) Lehrprozesse im Unterricht ab. Zusätzlich werden demographische Rahmenbedingungen über einen Kurzfragebogen im Vorfeld der Interviews zur Beschreibung der Stichprobe erhoben. Fokussiert wird dabei die individuelle Unterrichtserfahrung im Fach NwT und die studierte Fächerkombination der Lehrpersonen. Außerdem sollen die Kursgröße, Anzahl der unterrichteten Kurse und das entsprechende Geschlechterverhältnis angegeben werden. Bei der Transkription der Interviews wird der Dialekt des Interviewers und der Befragten geglättet und ggf. Sätze in eine korrekte grammatikalische Struktur gebracht, ohne dabei die Kernaussage zu verändern. Die Transkripte und Kurzfragebögen werden durch qualitativ inhaltsanalytische Methoden (Kuckartz & Rädiker, 2022; Mayring, 2022) erschlossen. Die insgesamt 1304 Codes im Kategoriensystem verteilen sich gemäß der Auflistung in Tabelle 5 auf die elf oben angeführten Kategorien.

**Tabelle 5:** Verteilung der Codes auf das Kategoriensystem mit zugehöriger Interraterreliabilität.

Kategorie	Codeanzahl	Cohens $\kappa$
Motivation und fachliche Voraussetzungen	68	.80
Unterrichtsorganisation	15	.97
Unterrichtsvorbereitung	65	.92
Spiralcurriculare Kompetenzteilbereiche	269	.76
Unterrichtsinhalte	222	.78
Unterrichtsmethoden	70	.85
Ausstattung der Schule und verfügbare Ressourcen	192	.85
Kompetenzerwerb der Schüler:innen	99	.64
Prüfungsformate	147	.82
Qualifikation der Lehrpersonen	68	.88
Wahrgenommene Unterstützung	63	.82
Kritik an zeitlichen Faktoren und Deputat	26	.91

Des Weiteren kann bei der Auswertung die Kategorie „Kritik an zeitlichen Faktoren und Deputat“ als Unterkategorie für die wahrgenommene Unterstützung induktiv erhalten werden. Zur Unterkategorie werden alle Erwähnungen einer ho-

hen zeitlichen Belastung im Vergleich zu anderen Kursstufenfächern und die damit zusammenhängende Entlastung über Stundendeputate zugeordnet. Die erzielten Interreliabilitätswerte bei zwei bewertenden Personen können als annehmbar bis gut bezeichnet werden (Wirtz & Caspar, 2002).

## **6.2 Untersuchungsergebnisse**

Zur Veranschaulichung der unterschiedlichen Positionen der befragten Lehrpersonen finden einzelne Zitate aus den Interviewtexten Verwendung. Zur hinreichenden Anonymisierung werden keine Informationen zu den interviewten Personen angegeben. Die einzelnen Lehrpersonen erhalten Nummerierungen, um einen möglichen Zusammenhang einzelner Zitate nachvollziehen zu können.

### **6.2.1 Demographische Rahmenbedingungen und individuelle Unterrichtserfahrung**

Die Pilotkurse der sechs Versuchsschulen werden von  $N=91$  (weiblich = 19, männlich = 72, Kursgröße max. = 30, Kursgröße min. = 6) Schüler:innen besucht. Aufgrund der unterschiedlichen Kursgröße resultiert eine Spanne beim Betreuungsverhältnis zwischen 1:2 bis 1:15. Der Median der Unterrichtserfahrung der Lehrpersonen beträgt 11 Jahre (*Min.* = 3 Jahre, *Max.* = 18 Jahre, *Modus* = 3 Jahre). Alle haben mindestens eine Naturwissenschaft, Geografie oder Mathematik fachwissenschaftlich studiert. Es sind außerdem zwei Lehrpersonen mit einem Studium der Technik bzw. NwT beteiligt.

### **6.2.2 Motivation und fachliche Voraussetzungen**

Zunächst werden die Beteiligten zu ihren intrinsischen und extrinsischen Motiven (in Anlehnung an Deci & Ryan, 1993) bezüglich der Durchführung des Pilotschulprojekts und v. a. der Organisation und Durchführung des Unterrichts im ersten Unterrichtsjahr als vierstündiges Kursstufenfach befragt. Intrinsische Motive sind als von der Lehrperson ausgehende und interessengeleitete Handlung zu verstehen (Deci & Ryan, 1993). Unter extrinsischen Motiven werden alle Handlungen im Bezugsfeld der NwT-Pilotphase zusammengefasst, die mit einer Aufforderung oder sonstiger Art von äußerer Einflussnahme (bspw. durch Vorgesetzte) in Zusammenhang stehen (ebd.).

Zwölf der befragten Lehrpersonen geben als intrinsische Motive an, selbst Spaß am Fach und Interesse an den Inhalten zu haben.

„Ich persönlich hätte das Fach auch sehr gerne belegt und ich finde auch die Inhalte für die Oberstufe sehr interessant. Deswegen war das für mich wichtig, dass das durchgeführt wird an der Schule.“ – Lp5

Als extrinsische Motivation wird von mehr als der Hälfte der Lehrpersonen ( $n = 7$ ) die logische Konsequenz der Fortführung eines Hauptfachs aus der Mittelstufe als abiturfähiger Oberstufenkurs genannt und als externes Selbstverständnis ihre personelle Besetzung als durchführende Lehrperson aufgeführt.

„Und es war auch logisch, dass im Land NwT als Leistungsfach kommen muss, weil ein Fach, das in der Mittelstufe ein Hauptfach ist, schlecht ist, wenn kein Leistungsfach in der Kursstufe weiterführt. [...] und es war klar: wenn es kommt, dass es mindestens einer von unserem Team hier machen muss.“ – Lp6

Eine besondere Fächerkombination ist über die generellen Vorgaben von Seiten der Verwaltungsebenen nicht vorgeschrieben, wenngleich knapp die Hälfte der Lehrpersonen ( $n = 5$ ) angibt, dass ein Fachstudium in einer Naturwissenschaft oder Geografie und gesammelte Unterrichtserfahrung aus Mittelstufe und zwei-stündigem Oberstufenkurs wünschenswert und hilfreich für den Unterricht im Leistungsfach ist.

### **6.2.3 Unterrichtsorganisation**

Für die Organisation des unterrichtlichen Rahmens ist die Vorgehensweise der Lehrpersonen in den einzelnen Versuchsschulen von Interesse. Alle geben an, den Unterricht in Form einer Teamlösung organisieren und durchführen zu wollen. Der Unterricht wird an fünf von sechs Pilotschulen von Lehrerteams aus zwei bis drei Lehrpersonen organisiert und zum größten Teil gemeinsam durchgeführt. Als Vorteil dieser Organisationsform wird dargestellt, dass der gehaltene Unterricht unmittelbar evaluiert werden kann.

„Jetzt bei der Erprobung sind wir immer mit drei Kollegen anwesend, um direkt zu sehen: Welche Inhalte werden vermittelt und um dann den Unterricht zu evaluieren.“ – Lp3

Die Befragten geben geschlossen an, dass ein Lehrpersonenwechsel zum Halbjahr oder Schuljahr nicht wünschenswert bzw. geplant ist und somit jedes Lehrerteam seinen Pilotkurs über die zwei Kursstufenschuljahre begleitet.

### **6.2.4 Unterrichtsvorbereitung**

In der Kategorie „Unterrichtsvorbereitung“ werden die Interviewten zur Vorgehensweise bei der Konzeption von Unterrichtseinheiten und zu den verwendeten Quellen befragt. Zehn Lehrpersonen bevorzugen eine schulinterne Konzeption im Team, wohingegen an zwei Schulen die Unterrichtskonzeption zunächst federführend von einer Lehrperson vorgenommen und die erarbeiteten Entwürfe wöchentlich im Team besprochen werden. Die bevorzugten Quellen zur Vorbereitung sind Fachbücher und wissenschaftliche Artikel. Zehn Lehrpersonen bestätigen, diese regelmäßig für ihre Unterrichtsvorbereitung zu nutzen. Außerdem

geben sieben Befragte an, auf Inhalte aus Fortbildungen zurückgreifen zu können. Drei Befragte berichten, dass sie die Unterrichtsvorbereitung teilweise in Kooperation mit externen Experten aus Industrie und Forschung durchführen.

### **6.2.5 Spiralcurriculare Kompetenzteilbereiche**

Für eine mögliche Implikation des Oberstufencurriculums auf den Unterricht der Mittelstufe wird untersucht, welche inhaltsbezogenen Kompetenzbereiche in der Mittelstufe unterrichtet und als relevantes Vorwissen für die Kursstufe charakterisiert werden. Zum Zeitpunkt der Erhebungsphase im Schuljahr 2018/19 werden die in die Kursstufe einmündenden Schüler:innen in der Mittelstufe nach dem Bildungsplan 2004 (MKJS, 2004) unterrichtet. Es lässt sich feststellen, dass an den Pilotschulen bereits eine Ausrichtung der inhaltsbezogenen Kompetenzen in Richtung des Oberstufencurriculums des Bildungsplanentwurfs (Autorengruppe Bildungsplan NwT, 2019b) und des Bildungsplans 2016 (MKJS, 2016) für das Profulfach der Mittelstufe erfolgt. Konkret werden von allen Lehrpersonen Kompetenzen im Bereich „*Konstruktion, technisches Zeichnen und CAD*“, „*Grundlagen der Elektronik*“ und „*Mikrocontroller*“ als wichtig für die Kursstufe eingestuft und deshalb bereits in der Mittelstufe gefördert.

„Ich denke tatsächlich der Mikrocontroller, das haben Sie ja gesehen an den Projekten heute und in der Präsentation [der Schüler:innen]. Das CAD kann man auch immer gut gebrauchen, die ganzen Grundlagen [...] natürlich zu Hebelwirkung und Kräften. Auch Elektronik werden die Schüler:innen schon benötigen, wenn es losgeht mit dem Mikrocontroller und den Bauteilen. Uns wird auch immer mehr dieser spiralcurriculare Aufbau bewusst.“ – Lp11

Die Hälfte der Interviewten nennt zusätzlich die Kompetenzteilbereiche „*Energietechnik*“, „*Produktentwicklung*“, „*Messtechnik*“, „*Statik*“ sowie „*Regelungs- und Automatisierungstechnik*“ und gibt an, diese ebenfalls in der Mittelstufe grundlegend zu unterrichten.

### **6.2.6 Unterrichtsinhalte in der Kursstufe**

In Bezug auf die inhaltsbezogenen Kompetenzen des Bildungsplanentwurfs ist es von Interesse, welche fachlichen Inhalte im Unterricht zur Kompetenzförderung umgesetzt und welche ingenieurwissenschaftlichen Bezüge hergestellt werden. Darüber hinaus sollen die Befragten das Verhältnis von praktischen zu theoretischen Unterrichtsinhalten einschätzen. Im Kontext der praktischen Inhalte ist es vor allem von Interesse, Projekte zu den einzelnen Kompetenzteilbereichen des Bildungsplans und deren Dauer sowie Durchführungsform zu eruieren. Von den Lehrpersonen werden praktisch durchgeführte Anteile am Unterricht zwischen 20 % - 50 % genannt. Der praktische Anteil entfällt zum größten Teil auf Projekte,

die von den Schüler:innen bearbeitet werden. Allerdings geben Befragte, bei denen der Theorieteil wesentlich größer ausfällt an, dass sie gerne mehr praktische Anteile implementieren wollen.

„Im Moment sind es vielleicht 70 % [theoretische Anteile] zu 30 % [praktischen Anteilen] und es soll auf jeden Fall 50 % zu 50 % werden.“ – Lp7

Bei der inhaltlichen Gestaltung werden u. a. aktuelle gesellschaftliche Themen und Zukunftstrends fokussiert. Folgendes Zitat bietet einen Einblick in konkrete Inhalte der Kompetenzteilbereiche „*Technikfolgenabschätzung*“ und „*Energieversorgung*“.

„Es gibt einen Überblick über die Gründe, Ziele, Maßnahmen, Finanzierung und Herausforderungen der Energiewende, einen Überblick zum aktuellen Stand der Stromerzeugung nach Energieträgern und deren Herkunft, zum Endenergieverbrauch nach Sektoren und zu Vor- und Nachteilen der Kraftwerkstypen. Es geht dabei um große Energiespeicher, Smart-Home, Smart-Grid, den NIMBY-Effekt, um Stromnetze, um den Wirkungsgrad [verschiedener Komponenten und Kraftwerke im Stromnetz] und um die Methode der technikethischen Fallanalyse.“ – Lp7

Die über die Unterrichtsinhalte hergestellten ingenieurwissenschaftlichen Bezüge werden von den Lehrpersonen größtenteils in die vier Domänen „*Informationstechnik*“, „*Elektrotechnik*“, „*Maschinenbau*“ und „*Bauingenieurwesen*“ untergliedert. Die Projektarbeit zu den Themen der Kursstufe wird von neun Befragten als arbeitsteilig durchzuführende Entwicklung eines Gesamtprodukts organisiert. Vier Lehrpersonen geben an, dass die Schüler:innen die Projekte arbeitsgleich durchführten und jeder ein eigenes Produkt entwickeln sollte. Für die durchschnittliche Dauer einer Projektdurchführung wird eine Zeitspanne von mindestens 12-16 Schulstunden angegeben. Diese Zeitspanne entspricht im vierstündigen Pilotkurs drei bis vier Schulwochen. Die Lehrpersonen können es sich vorstellen, wesentlich längere Projektlaufzeiten im Leistungskurs zu realisieren. Diese Projekte sollen dann von inhaltsbezogenen Qualifikationsphasen unterbrochen werden.

„Es wäre wahrscheinlich ein Projekt, das in mehreren Phasen abläuft und bei dem immer wieder eine Qualifikationsphase eingeschoben wird. Ich kann mir vorstellen, dass so ein Projekt über ein ganzes Jahr laufen könnte, mit den entsprechenden Qualifikationsphasen.“ – Lp3

### **6.2.7 Unterrichtsmethodik**

Im Rahmen der Frage zur Unterrichtsmethodik werden von den Befragten sowohl lehrpersonen- als auch schülerzentrierte Unterrichtsmethoden genannt. Als

häufigste lehrpersonenzentrierte Methode wird der Lehrervortrag ( $n = 9$ ), respektive ein Lehrer-Schüler-Gespräch, im Rahmen der Qualifizierungsphasen eingesetzt.

„Die Methodik sieht in der Regel so aus, dass es eine Qualifikationsphase gibt mit einem Lehrervortrag, bzw. häufiger ein Lehrer-Schüler-Gespräch, in dem Schüler:innen auch Fragen stellen.“ – Lp4

Als häufigste schülerbezogene Methoden werden Selbstlernphasen ( $n = 9$ ), Partner- bzw. Gruppenarbeiten ( $n = 9$ ), Praktika ( $n = 10$ ) und Projektarbeit ( $n = 7$ ) aufgeführt. Zusammenfassend lässt sich sowohl in praktischen als auch theoretischen Einheiten nach den Aussagen der Lehrpersonen gesamtheitlich eher eine Schülerzentrierung des Unterrichts feststellen.

### **6.2.8 Ausstattung der Schule und verfügbare Ressourcen**

In der Kategorie „Ausstattung der Schule und verfügbare Ressourcen“ wird das verfügbare Budget für den Fachbereich NwT an den jeweiligen Schulen, die räumliche Situation sowie verfügbare Geräte und Maschinen erfragt. Darüber hinaus sind Optimierungsbedarfe hinsichtlich zusätzlicher Ressourcen und räumlicher Möglichkeiten von Interesse. Das Budget, das dem Fachbereich NwT an den Versuchsschulen zur Verfügung steht, variiert zwischen 1000 € bis 8000 €. Ein Budget von 1000 € pro Jahr wird für die Anschaffung von Geräten, wie z. B. 3D-Drucker oder eine CNC-Fräse, als nicht ausreichend beschrieben.

„[...] insgesamt für NwT haben wir ein Budget von ca. 1000 € pro Jahr. Aber da muss natürlich auch einiges an Verbrauchsmaterial durch gehen. Solche Dinge wie die CNC-Fräse zum Beispiel ist aus Sondermitteln angeschafft worden, das bekommen wir sonst nicht. Oder einen dieser 3D-Drucker, den haben wir über einen Wettbewerb bekommen. Solche Möglichkeiten muss man da schon nutzen.“ – Lp8

Nach der Häufigkeit der Nennungen lassen sich als wichtigste Maschinen und Geräte im NwT-Unterricht der Kursstufe eine Grundausstattung an Mikrocontrollern, ein Laptop für jeden Lernenden inklusive eines entsprechenden Softwarepakets für CAD und Simulationen, mindestens ein 3D-Drucker, eine CNC-Fräse und entsprechende Handwerkzeuge sowie Bauteile für den elektrotechnischen Bereich identifizieren. Als weniger relevante Gerätschaften für die Kursstufe werden Tischbohrmaschinen und Dekupiersägen genannt, die in allen Schulen für den Mittelstufenunterricht vorhanden sind. Schulübergreifend wird eine ähnliche räumliche Situation beschrieben. Der Fachbereich NwT wird an den Versuchsschulen größtenteils eigenständig behandelt und ist mit separaten Sammlungsräumen ausgestattet. An allen Schulen ist mindestens ein Fachraum mit Werkbänken für die Schüler:innen vorhanden. Es sind pro Schule mindestens zwei



moderne Fachräume eingerichtet, die hauptsächlich oder ausschließlich dem Fachbereich NwT zugeordnet sind. Alle Schulen haben zusätzliche Werkstatt-räume für größere Maschinen, 3D-Drucker und CNC-Fräsen, um einen ungestör-ten Unterricht bei gleichzeitig laufender Fertigung durchführen zu können. Als Optimierungspotenzial nennen alle die Funktionalität und Stabilität der Internet-verbinding in den Fachräumen und der Schule allgemein. Es wird zudem ver-mehrt der Wunsch geäußert, die IT-Infrastruktur extern verwalten und warten zu lassen, um die Funktionalität zu verbessern und zu gewährleisten.

„Grundsätzlich ist ein massives Problem, mit dem wir kämpfen, das ganze Thema Computer und Netzwerkausstattung.“ – Lp8

„[...] was WLAN und Internet angeht, das muss alles ein Lehrer machen, der dafür wenige Stunden zur Verfügung hat, mehrere hundert Computer hier betreuen soll und zusätzlich noch einen Server. [...] da wäre es an der Zeit, dass man das in professionelle Hände gibt, dass von der Stadt eine professionelle IT-Infrastruktur gelegt und gepflegt wird.“ – Lp7

### 6.2.9 Kompetenzerwerb der Schüler:innen

Im Schulversuch ist es von Interesse, den Sach- und Methodenkompetenzerwerb anhand von Kompetenzfacetten aus den einheitlichen Prüfungsanforderungen in der Abiturprüfung (Autorengruppe der Kultusministerkonferenz, 2006) für das Fach Technik in Hinblick einer schulübergreifenden Standardisierung und in Bezug auf die Entwicklung geeigneter Testinstrumente zu untersuchen. Die Befrag-ten sollen erläutern, wie ihr Unterricht zum Sach- und Methodenkompetenzer-werb nach den EPA-Technik Kompetenzrichtlinien beiträgt und wie sie diese Kompetenzen nachweisen können. Die Lehrpersonen stellen den Bezug zur Durchführungsweise anhand der AQuAPRe (**A**usblick, **Q**ualifikation, **A**uftrag, **P**ro-jekt und **R**eflexion) Organisationsform (Landesinstitut für Schulentwicklung, 2017) für den Unterricht im Fach NwT her. Dabei soll der gesamte Prozess, vom Ausblick bis zur Reflexion, den Sachkompetenzgewinn fokussieren und gewähr-leisten.

„Wenn ich ein Produkt entwickle, dann muss ich zunächst einmal das technische Problem analysieren. [...] ich schaue mir an, wie das Problem [...] in Teilprobleme zerlegt werden kann und erstelle dann die Anforderungsliste. Dazu muss ich Wir-kungszusammenhänge ermitteln, Lösungen entwickeln und deren Wirksamkeit be-urteilen.“ – Lp7

Der Schwerpunkt des Methodenkompetenzerwerbs wird von acht Lehrpersonen in den praktischen Unterrichtsphasen, mit Bezug auf die Umsetzung einer Pro-duktentwicklung und deren fortlaufenden Dokumentation, in Praktika sowie der abschließenden Präsentation und Reflexion einer Projektarbeit gesehen.

„[...] wenn wir ein Projekt durchführen in der Kursstufe, dann gibt es da auch eine Dokumentation. Da erkennt man am Ende immer am Produkt ziemlich gut, wie es um die [Methoden]Kompetenzen steht: Umgang mit Technik, kritisches Beobachten, korrekte Fachsprache und Kommunikationsformen sowie Chancen und Risiken [der Technik] zu erkennen.“ – Lp3

Alle Befragten geben an, dass sie Sachkompetenz über das Schreiben von Klausuren und die Auswertung der Produktdokumentationen abprüfen können. Die Methodenkompetenz soll durch spezielle Aufgaben innerhalb der Klausuren erfasst werden, die beispielsweise die technische Kommunikation anhand der zeichnerischen Darstellung erfordern. Darüber hinaus dienen sowohl der praktische Teil der Projektdokumentation als auch die Funktionalität des Endprodukts und dessen technische Umsetzung als Nachweis für die Methodenkompetenz.

„Die [...] Klausur hat mit den Schüler:innen am PC stattgefunden. Sie haben da Simulationen durchgeführt, CAD-Zeichnungen erstellt und anhand vorgegebener Messdaten Rückschlüsse gezogen. Ansonsten sind die Methodenkompetenzen, glaube ich, hauptsächlich in der Arbeit in den Projekten nachzuweisen.“ – Lp8

Es wird zudem als Herausforderung identifiziert, die Methodenkompetenz im Rahmen einer schriftlichen Abiturprüfung zu erfassen.

„Da wurden sehr viele Methodenkompetenzen erfasst durch diese Klausur [am PC], was natürlich eine Herausforderung wird, das im Abitur so abzubilden. Da kann man sicherlich nicht an einem Rechner arbeiten.“ – Lp8

#### **6.2.10 Prüfungsformate**

Die Lehrpersonen werden zu den Prüfungsformaten und den Aufgabenarten für die Leistungsmessung im Kursstufenfach befragt. Des Weiteren sollen sie eine Gewichtung der schriftlichen gegenüber den praktischen Prüfungen vornehmen. Alle Befragten nennen als schriftliche Prüfungsform Klausuren. Ergänzend zur schriftlichen Benotung wird von allen Interviewten die Projektarbeit der Schüler:innen in Form einer praktischen Note bewertet. Die Lehrpersonen geben an, dass als Aufgabenstellung im Rahmen der praktischen Arbeiten eine Anforderungsliste erstellt wird, die von den Schüler:innen im Projekt zu erfüllen ist.

„Es wird eine Anforderungsliste geben, die im Rahmen der Problemlösung dann, inklusive Dokumentation, abgearbeitet werden muss.“ – Lp3

Fünf Lehrpersonen wünschen sich in diesem Kontext ein praxisbezogenes Prüfungsformat für die Abiturprüfung, bei dem die Schüler:innen ein Projekt selbstständig vorbereiten, das Endprodukt präsentieren und in einem Kolloquium abschließend zu fachlichen Details befragt werden.

„Eine Idee für mich ist, dass die Schüler:innen im Prinzip eine praktische Arbeit erstellen, aber nicht im Rahmen der Prüfung, sondern im Vorfeld. In der Prüfung werden sie dann zu dieser Produktentwicklung befragt.“ – Lp8

Die Hälfte der Befragten gibt an, dass praktische und schriftliche Anteile gleichwertig in die Gesamtnote einmünden. Die am zweithäufigsten genannte Gewichtungsumform entspricht einer dreifachen Wertung der schriftlichen Klausuren, der zweifachen Gewichtung der Projekte und einer einfachen Wertung der mündlichen Beiträge.

### **6.2.11 Qualifikation der Lehrpersonen**

Zwei Lehrpersonen im Schulversuch haben NwT bzw. Technik als Fachwissenschaft studiert und weitere zwei Befragte absolvierten das Zusatzstudium für NwT am KIT. Alle Interviewten besuchten darüber hinaus die angebotenen Weiterbildungen, die eigens für die Kompetenzteilbereiche der Kursstufe ausgearbeitet wurden. In Kompetenzteilbereiche, die eine Wissensvertiefung über das Fortbildungsangebot hinaus erfordern, arbeiten sich die Befragten größtenteils autodidaktisch oder im Austausch mit ihren Kolleg:innen ein.

„Ich habe mir sehr viel in Zusammenarbeit mit meinen Kollegen hier an der Schule über die verschiedenen Fächer erarbeitet.“ – Lp3

Als Kritikpunkt am Fortbildungsangebot wird die heterogene Organisation durch unterschiedliche Institutionen bemängelt.

„Von den einzelnen Veranstaltungen her sind die [Qualifizierungsmaßnahmen] bisher teilweise von einer Uni organisiert worden. Dann ist manchmal eine unterrichtliche Umsetzung versucht worden darzustellen, aber von Leuten, die wenig oder gar keine Unterrichtserfahrung haben. Das hilft nicht so gut weiter. [...] und teilweise haben die Veranstaltungen Kollegen durchgeführt, bei denen dann der fachwissenschaftliche Hintergrund nicht ganz so da ist.“ – Lp8

Für das Qualifikationsangebot wird deshalb der konkrete Wunsch geäußert, fachwissenschaftliche und fachdidaktische Konzepte stärker zu vernetzen. Dadurch soll gewährleistet werden, dass einerseits das Niveau der fachwissenschaftlichen und andererseits der fachdidaktischen Inhalte dem Anspruch der Kursstufe gerecht werden kann.

„[...] indem man die Kombination anbietet, dass eine Uni o. Ä. den fachwissenschaftlichen Hintergrund darstellt, auf das Niveau, was man braucht [...] heruntergebrochen und die Lehrpersonen aus den Pilotschulen, die dann Unterrichtserfahrung gesammelt haben, ihre Unterrichtseinheiten darstellen.“ – Lp8

### **6.2.12 Wahrgenommene Unterstützung, zeitliche Faktoren und Deputat**

Die wahrgenommene Unterstützung innerhalb der Versuchsschulen wird sowohl ausgehend vom Kollegium als auch ausgehend von der Schulleitung als gut beschrieben. Jeweils mehr als die Hälfte der Befragten üben Kritik an der Deputatzuweisung ( $n = 7$ ) und an zeitlichen Faktoren ( $n = 8$ ). In diesem Zusammenhang wird hauptsächlich der Zeitaufwand für die Unterrichtsorganisation, im Vergleich zum gewährten Deputat, bemängelt.

„[...] der Zeitaufwand, das neu einzuführen, sich da einzuarbeiten, ist enorm. Auch was ich an häuslicher Arbeit im Vorfeld einer Unterrichtsstunde investieren muss, ist enorm und das ist nicht zu vergleichen mit dem normalen klassischen Unterricht.“ – Lp3

„[...] der zeitliche Aufwand für diesen vierstündigen Kurs ist natürlich schon brutal. Ich unterrichte aktuell auch [ein anderes Fach] vierstündig, das ist locker Faktor vier [an Zeitaufwand für NwT], wenn es reicht.“ – Lp4

Es wird in diesem Kontext mehrmals angemerkt, dass ein größeres Zeitbudget zur Einführung des Kursstufenfachs notwendig erscheint.

„Wir fühlen uns jetzt natürlich schon ganz schön gedrängt. Das muss jetzt in kurzer Zeit so gut wie möglich stehen. Ich sehe es z. B. mit großen Bedenken, dass wir jetzt hergehen und eine Rückmeldung geben sollen über den Bildungsplan [...] und dass es dann möglicherweise der Bildungsplan sein wird, der dann für die nächsten 10 Jahre in Betrieb ist. [...] ich bin mir ganz sicher, dass sich einige Sachen, die jetzt so gemacht wurden, wieder ändern werden [im weiteren Verlauf der Schulversuchsphase].“ – Lp13

### **6.3 Zwischenfazit zur Interviewstudie**

Das Ziel der Interviewstudie ist es, die inhaltlichen und kontextuellen Rahmenbedingungen zu Eingang des Schulversuchs aus Sicht der Lehrpersonen zu untersuchen und systematisch zu betrachten, (1.) welche Rahmenbedingungen an den Versuchsschulen gegeben sind, (2.) wie der Unterricht organisiert und gestaltet wird, (3.) welche technischen und finanziellen Ressourcen für den Unterricht zur Verfügung stehen, (4.) wie die Förderung der Kompetenzen von Schüler:innen angestrebt wird und wie diese in Prüfungsformaten erfasst werden sowie (5.) welche Optimierungspotenziale in der frühen Schulversuchsphase in Bezug auf die Weiterbildungs- und Unterstützungsangebote eruiert werden können. Die Betrachtung der Angebotsebene dient u. a. als aktuelle Bestandsaufnahme zur Umsetzung von NwT in Mittel- und Kursstufe im Schuljahr 2018/19 und als Grundlage für die Testentwicklung. Die o. a. zentral relevanten Punkte und deren Implikationen für den Schulversuch werden im Folgenden zusammengefasst.

(1.) Die Ergebnisse der Interviewstudie belegen, dass die Lehrpersonen eine intrinsische Motivation in Bezug auf die Durchführung des Unterrichts in der Kursstufe aufweisen. Extrinsische Motive treten bei etwa der Hälfte der Interviewteilnehmer:innen hauptsächlich in Bezug auf die Fortsetzung der Mittelstufe auf. Die Lehrpersonen sehen sich dort aufgrund ihrer Expertise im Fach NwT als notwendige Besetzung. Ihre Unterrichtserfahrung erstreckt sich von minimal drei Jahren bis hin zu 18 Jahren. Die Kursgrößen und Betreuungsverhältnisse variieren zwischen den Versuchsschulen im Bereich von 1:2 bis 1:15.

(2.) Während der Pilotphase wird der Unterricht an den Schulen hauptsächlich im Team organisiert und durchgeführt. Eine Lehrperson setzt dabei meist das Unterrichtskonzept um, während die andere Lehrperson Optimierungspotenziale festhält. Im NwT-Unterricht der Mittel- und Kursstufe entfällt ein hoher Zeitanteil auf die Projektarbeit als Schüler:innenzentrierte und handlungsorientierte Unterrichtsmethode (vgl. Zendler, 2018, S. 22). Trotz der Vorgaben des zum Befragungszeitpunkt noch gültigen Bildungsplans in der Fassung von 2004 wird an den Pilotschulen bereits ein auf die Kursstufe zugeschnittenes Unterrichtskonzept in der Mittelstufe umgesetzt, da von allen Lehrpersonen die Kompetenzfacetten *Konstruktion, technischen Zeichnen und CAD, Grundlagen der Elektronik* und *Mikrocontroller* als wichtig für die Kursstufe eingeschätzt werden und die Lernenden bereits anschlussfähige Grundkompetenzen zu diesen Themen aus der Mittelstufe mitbringen sollen.

(3.) Trotz der Unterschiede beim Budget, das für das Fach NwT an den jeweiligen Schulen zur Verfügung steht, beschreiben die Lehrpersonen die technische Ausstattung ihrer Schule als ausreichend für eine Umsetzung von NwT in der Kursstufe. Ein Budget von 1000 € wird jedoch als zu niedrig eingeschätzt, um Geräteanschaffungen zu bestreiten. Die wichtigste Grundausstattung, die eine Schule für die Umsetzung von NwT in der Kursstufe im Klassensatz anschaffen sollte, sind Mikrocontroller und entsprechendes Zubehör, ein Laptop für jeden Lernenden inklusive eines entsprechenden Softwarepakets für CAD, Simulation und Datenauswertung sowie Werkzeuge für die Projektarbeit. Größere Maschinen, die in der Schule zur Verfügung stehen sollten, sind 3D-Drucker, eine CNC-Fräse, Heißdrahtschneider und Maschinen für die Holzbearbeitung. Als Problematik der infrastrukturellen Rahmenbedingungen konnte die Verfügbarkeit und Funktionalität der Internetverbindung an den Schulen identifiziert werden.

(4.) Die Förderung der in der EPA-Technik abgefassten Anforderungen in Bezug auf die Sach- und Methodenkompetenz bei Schüler:innen kann nach Aussagen der Interviewteilnehmer:innen grosso modo bestätigt werden. Eine Orientierung

an der EPA-Technik erscheint im Sinne einer übergeordneten Standardisierung in Hinblick auf die Abiturprüfung neben den inhaltlichen Vorgaben durch den Bildungsplan sinnvoll. Die Lehrpersonen prüfen die Sachkompetenz der Schüler:innen im Kontext der inhaltsbezogenen Kompetenzteilbereiche des Bildungsplans über Klausuren ab. Die Methodenkompetenz und damit verbundene prozessbezogene Kompetenzen des Bildungsplans werden mehrheitlich über die Projektdokumentation sowie die technische Umsetzung und Funktionalität des Endprodukts einer Projektarbeit bewertet. Alle Lehrpersonen geben an, dass sie die Möglichkeit nutzen, eine Klausur durch eine Projektarbeitsnote zu ersetzen.

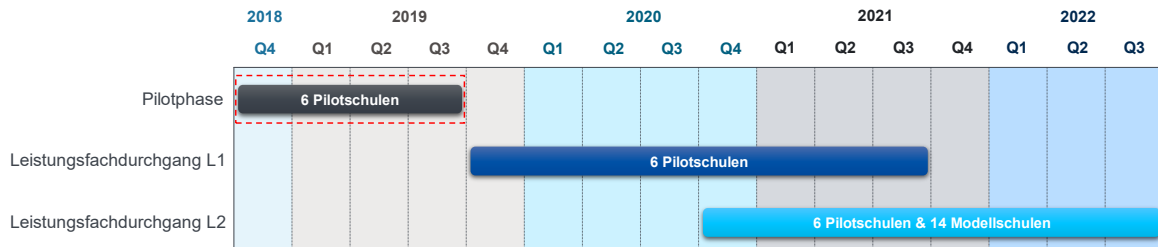
(5.) In Bezug auf die Qualifikation der Lehrpersonen für die Kursstufe geben zwei Personen an, NwT bzw. Technik als Fachwissenschaft studiert zu haben. Darüber hinaus nehmen alle Interviewten am eigens für die Kursstufe entwickelten Weiterbildungsprogramm teil. Optimierungswünsche werden in Bezug auf die Vernetzung von fachwissenschaftlichen und fachdidaktischen Inhalten geäußert.

#### **6.4 Implikation der Lehrpersonenbefragung für die Testentwicklung**

Mit den Ergebnissen der Interviewstudie verbinden sich darüber hinaus Implikationen für die Testentwicklung. Die Interviews bieten Einblick in die Qualifizierung der Lehrpersonen in den Kompetenzbereichen des Leistungsfachbildungsplans (s. Kapitel 2.4.2). Da immer zwei Pilotschulen als Tandem zwei Kompetenzteilbereiche vertieft erarbeiten, können speziell diese Lehrpersonen als Bildungsexpert:innen in die Testentwicklung miteinbezogen werden. In Kompetenzteilbereichen, in denen sie eine vertiefte Expertise erlangen, sollen sie die Aufgabenkomplexität sowie die curriculare und unterrichtliche Relevanz der Aufgaben einschätzen. Aus den Interviews geht zudem hervor, dass die Kompetenzbereiche des Mittelstufenbildungsplans eine stärker spiralcurriculare Anbindung an die Kursstufe erhalten sollen. Unter diesem Gesichtspunkt erscheint es notwendig, die Schüler:innen auf das erlangte Vorwissen aus der Mittelstufe zu testen, da von einem Einfluss auf die Fachwissensentwicklung auszugehen ist (z. B. Nickolaus et al., 2015; Reinfried & Tempelmann, 2014; Schmidt, 2018; Simonsmeier et al., 2022).

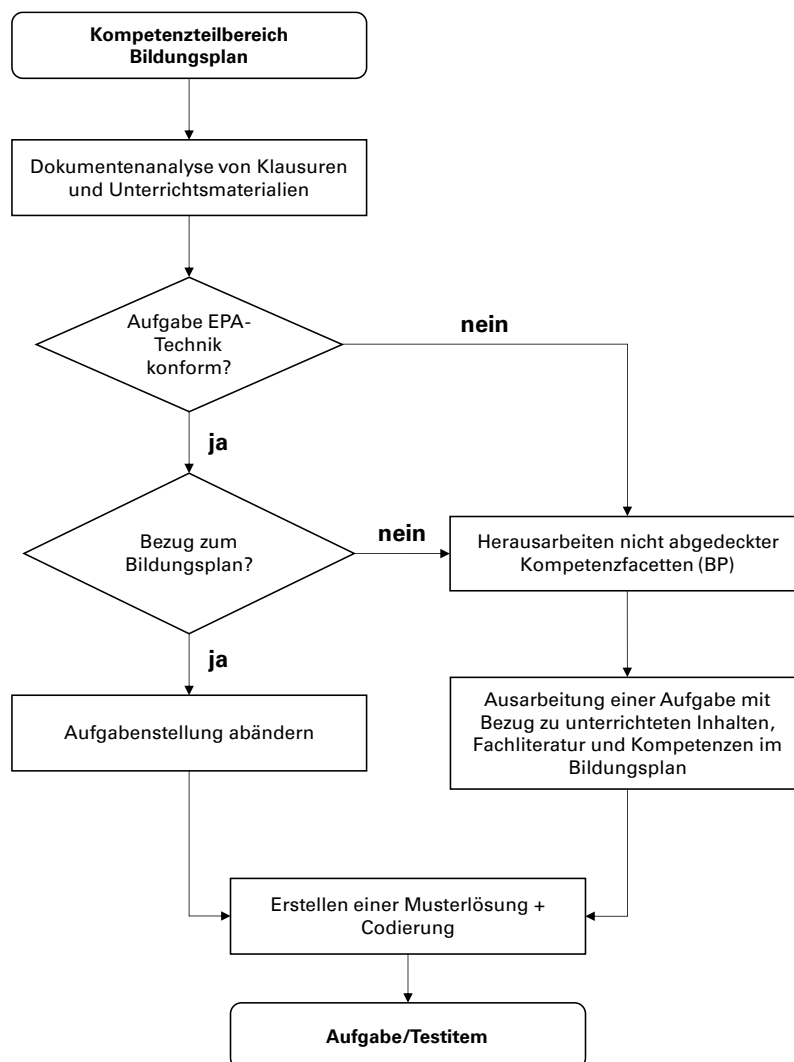
## 7 Testentwicklung zur Erfassung inhaltsbezogener Kompetenzen im Leistungsfach NwT

Im Rahmen der Pilotphase (s. Abbildung 11) erfolgt die Entwicklung valider und reliabler Testinstrumente zur Erfassung des Vor- und Fachwissens der Schüler:innen im Leistungsfach NwT.



**Abbildung 11:** Verortung der Testentwicklung und Pilotierungsstudie im Erhebungsablaufplan.

Dazu werden zunächst gemäß der Vorgehensweise in Abbildung 12 Aufgaben entwickelt, zu Testskalen zusammengefasst und deren Gütekriterien überprüft.



**Abbildung 12:** Vorgehensweise bei der Entwicklung von Testaufgaben.

Die Dateneingabe, Bereinigung und Auswertung erfolgt in *SPSS 25.0* (SPSS Inc., 2017). Für alle Auswertungen wird ein Signifikanzniveau von  $\alpha = .05$  zugrunde gelegt.

### **7.1 Testinstrument zur Erfassung des Vorwissens aus der Mittelstufe**

Das Vorwissen wird als mögliche Einflussvariable auf das Fachwissen erhoben. Darüber kann festgestellt werden, welchen Wissensstand die Schüler:innen aus der Mittelstufe als Voraussetzung für das Leistungsfach mitbringen und wie sich das Vorwissen auf die Fachwissensentwicklung im Leistungsfach auswirkt. Die im Rahmen dieser Arbeit entwickelten Testinstrumente kommen in zwei Leistungsfachdurchgängen L1 und L2 zum Einsatz.

#### **7.1.1 Bezug zu den Kompetenzteilbereichen des Profilfachs der Mittelstufe**

Die Erstellung der Testitems zur Erfassung des Vorwissens wurden auf den Bildungsplan des Profilfachs NwT in der Mittelstufe (MKJS, 2016) ausgelegt. Zur Eingrenzung des Umfangs wurden diejenigen inhaltsbezogenen Kompetenzen berücksichtigt, die Anknüpfungspunkte für das Leistungsfach bieten und somit potenziell eine spiralcurriculare Relevanz aufweisen (Tabelle 6). Es wurden insgesamt 40 Aufgaben ausgearbeitet, die sich auf die Kompetenzteilbereiche und deren Kompetenzfacetten verteilen. Aus vorhandenen Testinstrumenten im Rahmen der Evaluation des zweistündigen Schulversuchs durch das Institut für Erziehungswissenschaft der Universität Stuttgart können einige Aufgaben von Zinn et al. (2017) und Misic (2018) adaptiert werden. Im Folgenden werden aus der Literatur entnommene Aufgaben entsprechend gekennzeichnet.

#### **7.1.2 Inhaltliche Validierung der Testaufgaben zum Vorwissen**

Die Eingrenzung der Aufgaben für die Pilotierung mit den Schüler:innen erfolgt durch eine inhaltliche Validierung mittels Expert:innenbefragung der Lehrpersonen. Im Expertenrating sollen von insgesamt  $n = 12$  Lehrpersonen jeweils 20 der insgesamt 40 Aufgaben in den Kategorien *Komplexität*, *curriculare Relevanz* und *Praxisrelevanz* durch eine sechsstufige Likert-Skala bewertet werden, bei der 1 dem schlechtesten und 6 dem besten zu vergebendem Wert entspricht. Dazu werden zwei Testhefte erstellt, die von  $n = 6$  zufällig ausgewählten NwT-Lehrpersonen der Pilotschulen bearbeitet werden.



**Tabelle 6:** Inhaltliche Validierung der Aufgaben durch die Lehrpersonen. Die angegebenen Werte repräsentieren den Median der Stichprobe. Ausgegraute Aufgaben wurden nach der Expertenbefragung verworfen.

Item	Thema	Kompe- tenzfacette	Kom- ple- xität	Curr. Re- lev.	Pra- xis- re- lev.
A1 <sup>a</sup>	EVA-Prinzip	SP (4)	3.00	5.00	4.50
A2	Teilsysteme und Systemgrenzen	SP (2)	4.00	4.00	3.00
A3	Energie-, Stoff- und Informationsströme im Wohnhaus	SP (2)	3.00	4.50	4.00
A4 <sup>b</sup>	Temperaturunterschied Sommer/Winter	E (1)	3.00	4.00	4.00
A5	Energieübertragungskette Dampfmaschine	E (1)	2.00	5.00	6.00
A6	Strom aus Wind und Sonne	Evs (1),(2)	4.00	5.50	6.00
A7	Energiebedarf Agrarproduktion	E (6)	4.00	3.00	4.00
A8	Täglicher Stromverlauf Sommer/Winter	Evs (1)	3.00	4.00	3.00
A9 <sup>b</sup>	Effizienzsteigerung Sonnenkollektor	Evs (2)	4.00	2.50	3.00
A10 <sup>b</sup>	Definition des Wirkungsgrads und Dimensionierung einer Photovoltaikanlage	E(7); Evs (4)	4.00	5.00	5.50
A11	Getriebe und Übersetzung	BF (6)	2.50	5.50	5.00
A12	Hebel	BF (4)	3.00	6.00	5.00
A13	Bewegungsentstehung und -umwandlung	BF (3),(5)	3.00	4.00	3.50
A14 <sup>b</sup>	Materialeigenschaften Brückenbau	EigS (1),(2)	3.50	4.50	5.50
A15	Leitfähigkeit von Stoffen	EigS (1)	4.00	2.50	1.00
A16	Löslichkeit von Stoffen	EigS (1)	2.00	2.50	2.00
A17 <sup>b</sup>	Grundkonstruktionsweise Dreieck	StP (1)	4.00	4.50	5.00
A18 <sup>b</sup>	Trägerquerschnitte unter Biegung	StP (2)	4.00	5.00	5.00
A19 <sup>b</sup>	Brückenkonstruktion	StP (1)	3.50	2.50	4.50
A20	Ressourcenschonende Werkstoffe	Pe (3)	3.50	3.00	2.00
A21	Trennen, Fügen, Umformen	Pe (4)	3.00	5.00	3.50
A22 <sup>a</sup>	Ansichten	Pe (1)	4.00	5.50	5.00
A23 <sup>b</sup>	Verfahren in der Automatisierungstechnik	SsV (2)	4.00	3.00	1.50
A24	Kohlenstoffkreislauf	SsV (1)	3.50	3.50	2.00

A25 <sup>b</sup>	Windsystem	SsV (1)	3.00	3.50	1.00
A26 <sup>b</sup>	Sensorik bei Robotern	Ia (1),(5)	4.00	4.50	4.50
A27	Subjektives Erleben und Intensität des physikalischen Reizes	Ia (4)	4.50	4.00	1.00
A28 <sup>b</sup>	Erweiterung menschlicher Sinne durch Sensoren	Ia (5)	3.00	4.50	2.50
A29	Reproduzierbarkeit einer Messung	GAD (1)	4.00	5.50	4.00
A30	Messfehler	GAD (1)	4.50	5.50	5.00
A31 <sup>b</sup>	Schwingung und Ton	GAD (4)	3.50	4.00	2.50
A32 <sup>b</sup>	Analoge und digitale Signale	Iv (1)	4.00	5.50	4.50
A33 <sup>b</sup>	Steuerung	Iv (3)	5.00	5.50	5.00
A34 <sup>a</sup>	Regelung	Iv (1),(4)	4.00	4.50	5.50
A35	Bauteile und Schaltkomponenten	ES (3)	3.50	6.00	5.50
A36 <sup>a</sup>	Schaltpläne LED	ES (4)	4.00	6.00	6.00
A37 <sup>a,b</sup>	Widerstand, Spannung und Leistung	ES (1)	3.00	5.50	5.50
A38 <sup>a</sup>	Programmabschnitt Mikrocontroller	Iv (6)	4.00	5.50	6.00
A39 <sup>a</sup>	Analoge und digitale Signale am Mikrocontroller	Iv (1),(6)	4.00	5.00	4.00
A40 <sup>a</sup>	Fehlersuche Programmabschnitt	Iv (5)	5.00	6.00	6.00

<sup>a</sup> Aufgabe adaptiert aus Misic (2018); <sup>b</sup> Aufgabe adaptiert aus Zinn et al. (2017); SP = Systeme und Prozesse; E = Energie in Natur und Technik; Evs = Energieversorgungssysteme; BF = Bewegung und Fortbewegung; EigS = Eigenschaften von Stoffen; StP = Statische Prinzipien in Natur und Technik; Pe = Produktentwicklung; SsV = Stoffströme und Verfahren; Ia = Informationsaufnahme durch Sinne und Sensoren; GAD = Gewinnung und Auswertung von Daten; Iv = Informationsverarbeitung; ES = Elektronische Schaltungen

Die Auswahl der Aufgaben für die Pilotierung erfolgt mit einem gewählten Cut-Off Wert bei vier auf der Likert-Skala. Für den Cut-Off ist die repräsentierte hohe Zustimmung in den Kategorien „curriculare Relevanz“ und „Praxisrelevanz“ entscheidend. Sofern alle Aufgaben eines Kompetenzteilbereichs die Bedingungen erfüllen, werden die jeweils besten Aufgaben verwendet und die Aufgabe mit den schlechtesten Werten verworfen oder überarbeitet. Der Kompetenzteilbereich *Stoffströme und Verfahren* wird aufgrund der gering eingeschätzten Praxisrelevanz der Aufgaben und der fehlenden Bezugnahme der Expert:innen in der Interviewstudie (Kapitel 6.2.5) im Rahmen der Vorwissenserfassung ausgeschlossen. Durch diese Vorgehensweise erfolgt eine Auswahl von 24 Aufgaben zur Abbildung des Vorwissens aus der Mittelstufe, die den Schüler:innen im Schulversuch zur Pilotierung vorgelegt werden.

### 7.1.3 Testgüte des Vorwissenstests

Zur Ermittlung der Testgüte und der Item-Trennschärfen erfolgt die Auswertung für den Einsatz der pilotierten Tests nach der Methodik der Klassischen Testtheorie. Anhand der erhaltenen statistischen Kennwerte soll das Testinstrument für insgesamt zwei Erhebungsdurchgänge konsultiert werden. Aufgrund der geringen Stichprobengröße in der Pilotierung kann eine IRT-Skalierung zur finalen Erstellung und Verifikation des Testinstruments im Nachgang der Studien gerechnet werden.

Zur Erfassung der Testgüte nach der Klassischen Testtheorie werden die Lösungsquoten, Itemtrennschärfen und die interne Konsistenz der Skala (Cronbachs Alpha) als statistische Kennwerte betrachtet. An der Pilotierung haben  $n = 54$  ( $w = 25.93\%$ ,  $m = 68.51\%$ ,  $kA = 5.56\%$ ) Schüler:innen teilgenommen. Die erhaltenen Testergebnisse sind normalverteilt (Shapiro-Wilk-Test:  $W = 0.957$ ,  $p > .05$ ) mit einer mittleren Lösungsquote von  $40.74\%$  ( $Med = 37.84\%$ ,  $Min. = 8.11\%$ ,  $Max. = 75.68\%$ ,  $SD = 13.67\%$ ). Für die Abbildung des Vorwissens kommen ausschließlich Items mit einer Lösungsquote zwischen  $0.20$  ( $20\%$ ) und  $0.80$  ( $80\%$ ) und einer minimalen Trennschärfe von  $r \geq 0.20$  im Testinstrument zur Anwendung. Eine Übersicht über die Items und entsprechende Kennwerte liefert Tabelle 7.

**Tabelle 7:** Statistische Kennwerte der Testanalyse nach der Klassischen Testtheorie. Die finale Trennschärfe bezieht sich auf die Skala mit dem höchsten Cronbachs Alpha Wert.

Aufgabe (Item)	Lösungsquote	Trennschärfe	Finale Trennschärfe
VW1 (A1)	0.61	-0.34	-
VW2 (A40)	0.31	0.32	0.35
VW3 (A5)	0.43	0.26	0.29
VW4 (A38)	0.74	0.28	0.34
VW5a (A8)	0.63	0.34	0.39
VW5b (A8)	0.57	0.19 <sup>a</sup>	0.27
VW6 (A22)	0.43	0.30	0.33
VW7a (A11)	0.89	0.22	-
VW7b (A11)	0.85	0.22	-
VW8a (A34)	0.15 <sup>a</sup>	0.26	0.28
VW8b (A34)	0.33	0.33	0.37

VW9a (A14)	0.24	0.32	0.30
VW9b (A14)	0.30	0.37	0.44
VW10a (A32)	0.13	0.32	0.41
VW10b (A32)	0.57	0.28	0.27
VW11 (A18)	0.22	-0.15	-
VW12a (A29)	0.56	0.27	0.31
VW12b (A29)	0.19 <sup>a</sup>	0.47	0.40
VW13 (A3)	0.26	0.46	0.45
VW14 (A39)	0.76	0.27	0.30
VW15 (A6)	0.59	0.13	-
VW16a (A37)	0.54	0.37	0.38
VW16b (A37)	0.83	0.36	0.32
VW16c (A37)	0.35	0.24	0.26
VW17a (A10)	0.28	0.32	0.31
VW17b (A10)	0.28	0.35	0.37
VW17c (A10)	0.09	0.50	0.43
VW18 (A35)	0.63	0.39	0.38
VW19 (A12)	0.41	0.53	0.54
VW20 (A33)	0.11	0.43	0.45
VW21a (A17)	0.19	0.27	0.23
VW21b (A17)	0.09	0.14	-
VW22a (A30)	0.30	0.55	0.49
VW22b (A30)	0.24	0.57	0.54
VW23 (A22)	0.35	0.31	0.32
VW24a (A26)	0.43	0.08	-
VW24b (A26)	0.20	0.44	0.34

Die Nummerierung der Aufgaben orientiert sich an der Reihenfolge im Testheft A. Die vorherige Nummerierung aus der Expertenbefragung wird zur besseren Zuordnung in Klammern ergänzt. Ausgegraute Items weisen kritische Kennwerte in Bezug auf Lösungsquote oder Trennschärfe auf. <sup>a</sup> Items wurden durch ein nur geringes Unter- oder Überschreiten der Grenzwerte und aufgrund ihrer Wichtigkeit zur Abbildung der Kompetenzfacetten in der Skala beibehalten.

Für die Eliminierung der Items zur Skalenbildung werden zuerst Aufgaben entfernt, die sich nicht in Teilaufgaben gliedern, um die Vollständigkeit der Aufgaben bestmöglich beizubehalten. Zusätzlich verbleiben vor allem Aufgaben, die



Die finale Skala weist einen Cronbachs Alpha Wert von  $C_\alpha = .79$  auf und ist im Kontext eines Fachwissenstest als gut zu bewerten. Diejenigen Items, die nach der Pilotierung aufgrund der relevanten Kennwerte (Lösungsquote und Trennschärfe) als kritisch zu bewerten sind, werden überarbeitet und in veränderter Form in den weiteren Testungen erneut eingesetzt, um eine möglichst vollständige inhaltliche Abdeckung des Bildungsplans des Profulfachs durch das Testinstrument zu gewährleisten.

## **7.2 Testinstrument zur Erfassung des Fachwissens in der Kursstufe**

Die Erfassung des Fachwissens der Schüler:innen im Leistungsfach erfolgt anhand ausgewählter inhaltsbezogener Kompetenzfacetten für alle Kompetenzteilbereiche des Bildungsplanentwurfs für das Leistungsfach (Autorengruppe Bildungsplan NwT, 2019b). Aus eingereichten Unterrichtsmaterialien und unter Einbezug der Expert:innenmeinung aller unterrichtenden Lehrpersonen werden die zu pilotierenden Testinstrumente unter Berücksichtigung der Inhaltsvalidität erstellt. Die Vorgehensweise gleicht der zur Entwicklung des Vorwissenstests.

### **7.2.1 Bezug zu den Kompetenzteilbereichen des Leistungsfachs**

Die Aufgaben sollen das Fachwissen anhand der inhaltsbezogenen Kompetenzen des Bildungsplanentwurfs (Autorengruppe Bildungsplan NwT, 2019b) abbilden. Die inhaltsbezogenen Kompetenzen im Bildungsplan erstrecken sich über neun Kompetenzteilbereiche (Übersicht in Kapitel 2.4.2). Prozessbezogene Kompetenzen finden in diesem Testverfahren keine Berücksichtigung, um ein einheitliches Kompetenzkonstrukt zu erfassen, das sich ausschließlich auf die fachwissenschaftlichen Inhalte des Leistungsfachbildungsplans fokussiert.

### **7.2.2 Inhaltliche Validierung der Testaufgaben zum Fachwissen im Leistungsfach**

Für jeden Kompetenzteilbereich sind 20 bis 22 Aufgaben von den Expert:innen hinsichtlich *Komplexität*, *curricularer Relevanz* und *Praxisrelevanz* durch eine sechsstufige Likert-Skala zu bewerten. Die detaillierten Ergebnisse sind im Anhang (Tabelle 31 bis Tabelle 39) dargestellt. Aufgaben, die nach abschließender Bewertung der Expert:innen in einem der o. a. Bewertungsbereiche einen zu geringen Durchschnittswert erzielen, werden überarbeitet oder ggf. ersetzt, um mit mehr als 15 Aufgaben pro Kompetenzteilbereich in die Pilotierungsphase überzugehen. Dem zugrunde liegt ein Cut-Off bei einem mittleren Skalenwert von vier, der noch einer eher hohen Zustimmung zuzuordnen ist.

### **7.2.3 Testgüte des Fachwissenstests**

Die Testgüte wird über eine Pilotierungsstudie im Schuljahr 2018/19 anhand der Klassischen Testtheorie ermittelt. An der Pilotierung haben insgesamt  $n = 77$

Schüler:innen teilgenommen. Die Testinstrumente werden den Schulen nach behandelten Kompetenzteilbereichen zugeordnet. Ausführliche Informationen zu den einzelnen Items der Skalen sind im Anhang in Tabelle 40 bis Tabelle 48 dargestellt. Nachfolgend sind in Tabelle 8 nur die wichtigsten Skalenkennwerte der einzelnen Kompetenzteilbereiche des Leistungsfach zusammengefasst.

**Tabelle 8:** Übersicht der wichtigsten Testkennwerte der einzelnen Skalen zu den Kompetenzteilbereichen der Kursstufe.

<b>Thema (<math>n_{PI}</math>)</b>	<b>Anzahl Items</b>	<b>Cronbachs Alpha</b>	<b>Lösungsquoten</b>	<b>Trennschärfen</b>
Technikfolgenabschätzung (26)	12	0.70	Min. = 23 %; Max. = 69 %; SD = 23 %	Min. = 0.21; Max. = 0.82; MW = 0.44
Energieversorgung (25)	12	0.77	Min. = 19 %; Max. = 54 %; SD = 11 %	Min. = 0.29; Max. = 0.70; MW = 0.50
Elektrische Antriebstechnik (23)	13	0.62	Min. = 10 %; Max. = 52 %; SD = 22 %	Min. = 0.20; Max. = 0.78; MW = 0.46
Technische Mechanik (27)	16	0.79	Min. = 11 %; Max. = 70 %; SD = 22 %	Min. = 0.24; Max. = 0.71; MW = 0.48
Produktentwicklung (29)	11	0.75	Min. = 10 %; Max. = 86 %; SD = 24 %	Min. = 0.23; Max. = 0.73; MW = 0.47
Grundlagen der Elektronik (12)	17	0.84	Min. = 17 %; Max. = 67 %; SD = 27 %	Min. = 0.23; Max. = 0.79; MW = 0.56
Aufnahme von Signalen in der Mt. (19)	11	0.75	Min. = 16 %; Max. = 84 %; SD = 24 %	Min. = 0.30; Max. = 0.69; MW = 0.49
Regelungstechnik (13)	13	0.84	Min. = 15 %; Max. = 85 %; SD = 26 %	Min. = 0.31; Max. = 0.82; MW = 0.57
Datenkommunikation (28)	14	0.81	Min. = 18 %; Max. = 61 %; SD = 25 %	Min. = 0.29; Max. = 0.75; MW = 0.54

Insgesamt wurde für jede Skala ein Cronbachs Alpha Wert von mindestens  $C_a = 0.7$  angestrebt. *Elektrische Antriebstechnik* erreicht mit  $C_a = 0.62$  einen noch akzeptablen Wert. Außerdem verbleiben zusätzlich zu den Items mit einer Lösungsquote zwischen 20 % und 80 % solche in den Skalen, die für die umfangreiche Abdeckung eines Kompetenzbereichs als unabdingbar erachtet werden. In den Skalen wurden keine Items behalten, die weniger als 10 % oder mehr als 90 % Lösungsquote aufweisen, um eine ausreichende Abdeckung der Randbereiche zu gewährleisten. Zudem wurden die Daten auf Normalverteilung untersucht (Tabelle 9).

Tabelle 9: Übersicht über die Normalverteilung der einzelnen Skalen zu den Kompetenzteilbereichen der Kursstufe.

Thema	Normalverteilung (Shapiro-Wilk)
Technikfolgenabschätzung	$W = 0.96, p > .05, \text{Schiefe} = 0.269, \text{Kurtosis} = -0.931$
Energieversorgung	$W = 0.95, p > .05, \text{Schiefe} = 0.443, \text{Kurtosis} = -0.712$
Elektrische Antriebstechnik	$W = 0.95, p > .05, \text{Schiefe} = 0.353, \text{Kurtosis} = -0.099$
Technische Mechanik	$W = 0.93, p > .05, \text{Schiefe} = 0.297, \text{Kurtosis} = -1.191$
Produktentwicklung	$W = 0.96, p > .05, \text{Schiefe} = 0.296, \text{Kurtosis} = -0.739$
Grundlagen der Elektronik	$W = 0.95, p > .05, \text{Schiefe} = 0.093, \text{Kurtosis} = -1.494$
Aufnahme von Signalen in der Messtechnik	$W = 0.94, p > .05, \text{Schiefe} = -0.409, \text{Kurtosis} = -1.050$
Regelungstechnik	$W = 0.88, p > .05, \text{Schiefe} = 0.532, \text{Kurtosis} = -0.064$
Datenkommunikation	$W = 0.95, p > .05, \text{Schiefe} = 0.313, \text{Kurtosis} = -1.098$

Aus den pilotierten Skalen werden zwei Testhefte erstellt. Das erste Testheft umfasst die sieben Kompetenzteilbereiche *Technikfolgenabschätzung, Energieversorgung, Elektrische Antriebstechnik, Technische Mechanik, Produktentwicklung, Aufnahme von Signalen in der Messtechnik* und *Regelungstechnik*. Das zweite Testheft behandelt die Kompetenzteilbereiche *Grundlagen der Elektronik* und *Datenkommunikation*. In Abbildung 14 ist eine finale Aufgabe aus dem Bereich *Elektrische Antriebstechnik* exemplarisch dargestellt. Es handelt sich dabei um eine Aufgabe, die sich aus zwei Items mit einem halboffenen und offenen Antwortformat zusammensetzt.

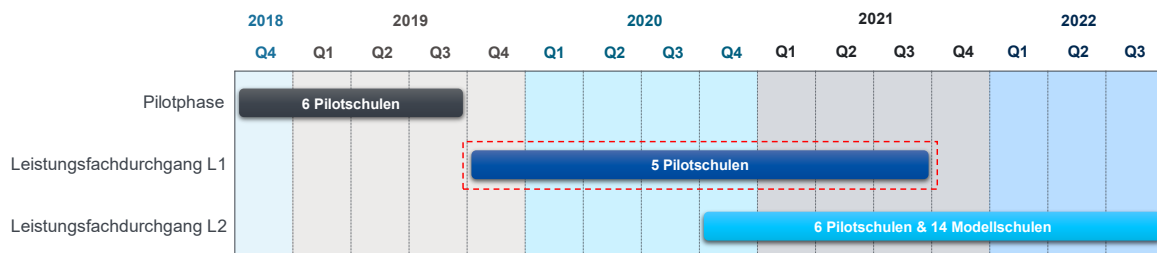






## 8 Fachwissen von Schüler:innen im Jahrgang 2019 – 2021 (L1)

Nach Abschluss der Pilotphase zur Entwicklung und Pilotierung der Testinstrumente erfolgt der erste Durchgang des fünfständigen Leistungsfachs NwT mit Abiturprüfung (L1). Die Befragung findet an den Schulen aus der Pilotphase statt. Über den Verlauf der Schuljahre 2019/20 und 2020/21 werden die Schüler:innen des Leistungskurses an diesen Schulen bis zu ihrem Abschluss begleitet (s. Abbildung 15). Eine der ursprünglich sechs Pilotschulen kann keinen Leistungskurs im Durchgang L1 anbieten.



**Abbildung 15:** Verortung der Studie zum Fachwissen im Leistungsfachdurchgang L1 im Erhebungsablaufplan.

### 8.1 Studiendesign

Zur Erfassung des Fachwissens erfolgt eine Befragung der Schüler:innen an vier Zeitpunkten über den Verlauf der Kursstufe hinweg. Zu Beginn der Kursstufe im Oktober des Schuljahres 2019/20 wird in der ersten Befragung das Vorwissen (Testentwicklung Kapitel 7.1) erhoben, mit dem die Lernenden aus der Mittelstufe in das Leistungsfach einmünden. Im weiteren Verlauf finden die Testungen nach einem dreiviertel Schuljahr und kurz vor der Abiturprüfung nach etwa drei Kursstufenhalbjahren statt, um das Fachwissen in den Kompetenzteilbereichen zu erfassen (Testentwicklung Kapitel 7.2). Die Testung deckt die folgenden Kompetenzteilbereiche des Bildungsplans ab (Autorengruppe Bildungsplan NwT, 2019b):

- *Technikfolgenabschätzung*
- *Energieversorgung*
- *Elektrische Antriebstechnik*
- *Technische Mechanik*
- *Produktentwicklung*
- *Aufnahme und Verarbeitung von Signalen in der Messtechnik (kurz: Messtechnik)*
- *Regelungstechnik*

Die Kompetenzteilbereiche *Grundlagen der Elektronik* und *Datenkommunikation* sind aus den Testungen vor der Abiturprüfung ausgeklammert. Ihre Erhebung ist am Ende des Leistungsfachdurchgangs L1 (Q3/2021) geplant. Diese testökonomische Maßnahme ist damit zu begründen, dass in den *Grundlagen der Elektronik* einige prozessbezogene Kompetenzfacetten angeführt sind, die nicht im Rahmen des schriftlichen Abiturs abgeprüft werden. *Datenkommunikation* wird aufgrund der gesamten Stoffmenge per Absprache durch alle Schulen nach der Abiturprüfung behandelt und kann damit erst zu diesem Zeitpunkt getestet werden.

## 8.2 Datenauswertung

Die Dateneingabe, Bereinigung und Berechnung der Testergebnisse erfolgt mit der Software *SPSS 25.0* (SPSS Inc., 2017). Alle weiteren Analysen werden in *R* (R Core Team, 2022) durchgeführt. In Erweiterung der Standardsoftware wird das *tidyverse* Softwarepaket (Wickham et al., 2019) verwendet. Abbildungen werden mit dem darin enthaltenen Package *ggplot2* (Wickham, 2016) erstellt. Für alle Auswertungen wird ein Signifikanzniveau von  $\alpha = .05$  zugrunde gelegt.

## 8.3 Beschreibung der Stichprobe

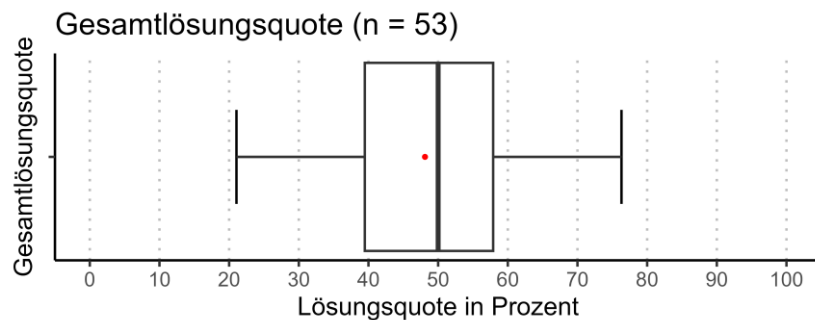
Die Befragung zum Schuljahresbeginn im Oktober 2019 erfolgt mit  $n_{\text{Vorwissen}, L1} = 53$  ( $w = 9\%$ ,  $m = 91\%$ ) und damit allen verfügbaren Schüler:innen im Leistungsfach NwT (Vollerhebung) an fünf der sechs Pilotschulen in Baden-Württemberg. An den folgenden Befragungen können im Juni 2020  $n_{\text{Fachwissen1}, L1} = 51$  ( $w = 6\%$ ,  $m = 96\%$ ) und im April bzw. Mai 2021 vor der Abiturprüfung  $n_{\text{Fachwissen2}, L1} = 41$  ( $w = 5\%$ ,  $m = 95\%$ ) Lernende des Leistungskurses erreicht werden. In der Abschlusserhebung im Juni 2021 standen  $n_{\text{Fachwissen3}, L1} = 42$  Schüler:innen ( $w = 7\%$ ,  $m = 93\%$ ) zur Verfügung.

## 8.4 Deskriptive Auswertung des Vorwissens

Die Auswertung des Vorwissens der Schüler:innen des Leistungsfachs bezieht sich auf die Daten der Testung zum Schuljahresbeginn 2019/20 (Oktober 2019). Es wurden keine weiteren Anpassungen der Aufgaben gegenüber der abschließenden Testversion aus der Entwicklungsphase vorgenommen. Die Schüler:innen erzielen eine mittlere Gesamtlösungsquote (Abbildung 16) von  $Med = 50.0\%$  ( $MW = 48.1\%$ ). Eine detaillierte Übersicht über die Lösungsquote der Gesamtstichprobe ist in Tabelle 10 aufgeführt. Die Testergebnisse der Schüler:innen liegen nach Shapiro-Wilk ( $W = 0.97$ ,  $p > .05$ ) normalverteilt vor. Mittels Varianzanalyse ( $F(4,48) = 1.38$ ,  $p > .05$ ) kann kein Unterschied beim mittleren Vorwissenstand der Lernenden zwischen den Schulen festgestellt werden.

Tabelle 10: Detaillierte Betrachtung der Lösungsquoten des Vorwissenstests der Studie 1 (Gesamtstichprobe).

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
21.05 %	39.47 %	48.11 %	50.00 %	57.89 %	76.32 %	13.28 %



**Abbildung 16:** Testergebnis des Vorwissenstests an den Pilotschulen im Leistungsfachdurchgang L1. Der rote Punkt in diesem und allen weiteren Boxplots entspricht dem Mittelwert.

## 8.5 Fachwissen im ersten Leistungsfachdurchgang (L1)

Das Fachwissen der Schüler:innen wird über zwei nachfolgend getrennt beschriebene Testzeitpunkte im Abstand eines dreiviertel Schuljahres erfasst. Dabei werden die Testpersonen in den in Kapitel 8.1 beschriebenen Kompetenzteilbereichen geprüft.

### 8.5.1 Deskriptive Auswertung des Fachwissens – erster Testzeitpunkt (L1)

Die erste Erhebung des Fachwissens der Testung erfolgt im Juni 2020 nach einem dreiviertel Schuljahr. Das Fachwissen liegt an Testzeitpunkt 1 mit einem Median von  $Med_{Ges} = 28.4\%$  ( $MW_{Ges} = 27.9\%$ ) erwartungsgemäß im unteren Bereich der Skala (s. Abbildung 17), da an den Schulen nur wenige Kompetenzteilbereiche vollständig abgeschlossen sind. Der Status der Kompetenzteilbereiche ist in Tabelle 11 dargestellt. Eine detaillierte Übersicht der Testergebnisse bietet Tabelle 12. Die Testergebnisse liegen nach Shapiro-Wilk ( $W = 0.99$ ,  $p > .05$ ) normalverteilt vor. Eine Varianzanalyse zwischen den Schulen wird zu diesem Testzeitpunkt aufgrund des o. a. heterogenen Themenfortschritts nicht vorgenommen.

Tabelle 11: Status der Kompetenzteilbereiche beim ersten Testzeitpunkt in L1.

Kompetenzteilbereich	Teilweise behandelt	Vollständig behandelt
Technikfolgenabschätzung	1	0
Energieversorgung	4	0
Elektrische Antriebstechnik	1	2
Technische Mechanik	1	4
Produktentwicklung	1	4
Messtechnik	1	1
Regelungstechnik	1	1

Tabelle 12: Ausführliche Betrachtung der Testergebnisse des Fachwissenstests der abiturrelevanten Kompetenzteilbereiche am Zeitpunkt der ersten Testung in L1.

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
12.50 %	22.73 %	27.94 %	28.41 %	32.95 %	45.46 %	7.59 %

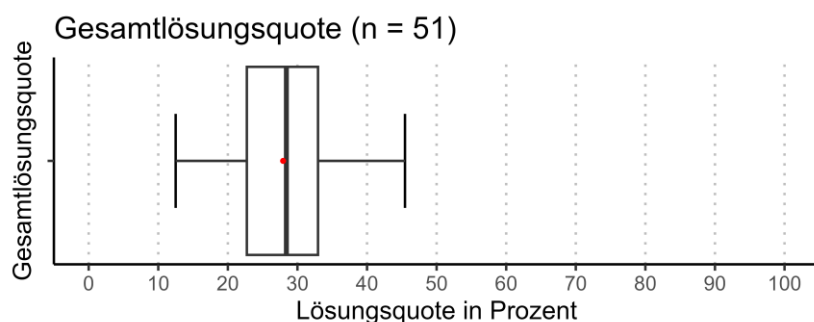


Abbildung 17: Gesamtlösungsquote des Fachwissenstests an den Pilotschulen zum ersten Testzeitpunkt im Leistungsfachdurchgang L1.

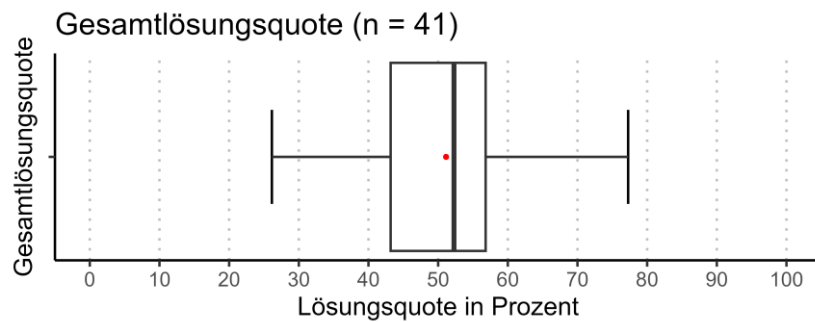
### 8.5.2 Deskriptive Auswertung des Fachwissens – zweiter Testzeitpunkt (L1)

In der zweiten Erhebung des Fachwissens vor der Abiturprüfung erreichen die Schüler:innen ein mittleres Ergebnis von  $Med = 52.27\%$  ( $MW = 51.14\%$ ). Eine Detaillierte Ergebnisübersicht liefert Tabelle 13.

Tabelle 13: Übersicht der Testergebnisse des Fachwissenstests der zweiten Testung der abiturrelevanten Kompetenzteilbereiche in L1.

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
26.14 %	43.18 %	51.14 %	52.27 %	56.82 %	77.27 %	10.43 %

Die Gesamtlösungsquote ist in Abbildung 18 aufgetragen.

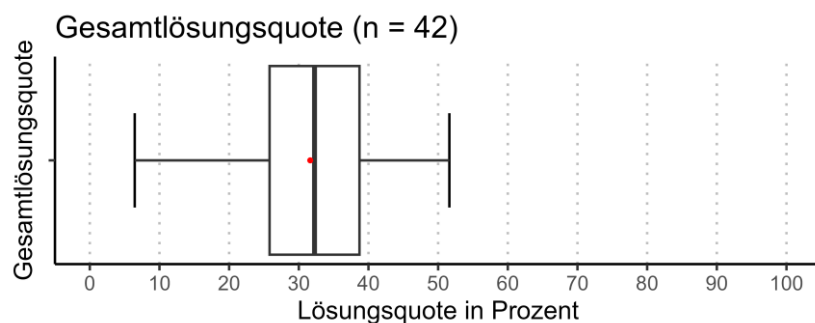


**Abbildung 18:** Gesamtlösungsquote des Fachwissenstests an den Pilotschulen des Leistungsfachdurchgangs L1 am zweiten Testzeitpunkt vor dem Abitur.

Das Gesamtergebnis des Fachwissenstests liegt nach Shapiro-Wilk ( $W= 0.98$ ,  $p > .05$ ) normalverteilt vor. Ein Vergleich der Schulen mittels ANOVA zeigt, dass keine signifikanten Unterschiede ( $F(4,36) = 1.10$ ,  $p > .05$ ) beim Fachwissen zwischen den Schulen am Messzeitpunkt vor der Abiturprüfung vorliegen.

### 8.5.3 Deskriptive Auswertung der Kompetenzteilbereiche *Grundlagen der Elektronik und Datenkommunikation* (L1)

In der Testung der verbliebenen Kompetenzteilbereiche *Grundlagen der Elektronik* und *Datenkommunikation* erreichen die Testteilnehmer:innen eine Gesamtlösungsquote (Abbildung 19) von  $Med = 32.26\%$  ( $MW = 31.64\%$ ).



**Abbildung 19:** Gesamtlösungsquote des Testteils *Grundlagen der Elektronik und Datenkommunikation* im Leistungsfachdurchgang L1.

Das Gesamtergebnis der beiden Kompetenzteilbereiche ist nach Shapiro-Wilk ( $W= 0.98$ ,  $p > .05$ ) normalverteilt. Eine detaillierte Übersicht über die Gesamtlösungsquote findet sich in Tabelle 14.

**Tabelle 14:** Übersicht der Testergebnisse der Kompetenzteilbereiche *Grundlagen der Elektronik* und *Datenkommunikation* in L1.

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
6.45 %	25.81 %	31.64 %	32.26 %	38.71 %	51.61 %	11.11 %

Nach Angabe der Lehrpersonen wurde *Datenkommunikation* nach dem Abitur nicht mit zufriedenstellendem Tiefgang unterrichtet. Im Schuljahr 2021 wurde das Abitur im Mai (und damit später im Vergleich zu anderen Schuljahren) durchgeführt, wodurch zeitliche Verzerrungen entstanden sind. Dadurch wird die Messung des Fachwissens im Kompetenzteilbereich *Grundlagen der Elektronik* und *Datenkommunikation* insgesamt beeinflusst. Aus diesem Grund erscheint zusätzlich eine gesonderte Betrachtung und Einordnung des Kompetenzteilbereichs *Grundlagen der Elektronik* sinnvoll.

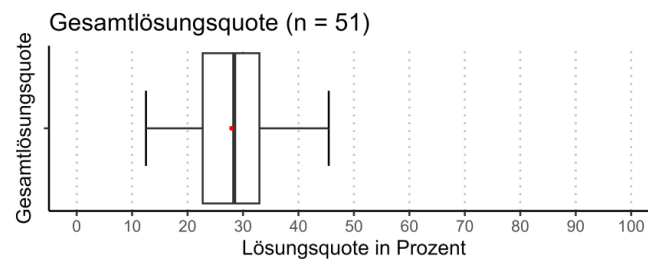
Unter Ausschluss von *Datenkommunikation* zeigt sich, dass die Schüler:innen im Kompetenzteilbereich *Grundlagen der Elektronik* eine mittlere Lösungsquote von *Med.* = 41.18 % (*MW* = 40.76 %) erreichen und im Bereich der zuvor erreichten Lösungsquote der Kompetenzteilbereiche für das Abitur rangieren. Es ist kein signifikanter Unterschied zwischen den Pilotschulen erkennbar ( $F(4,36) = 2.52$ ,  $p = .058$ ).

### 8.6 Zwischenfazit zu den Ergebnissen aus Studie L1

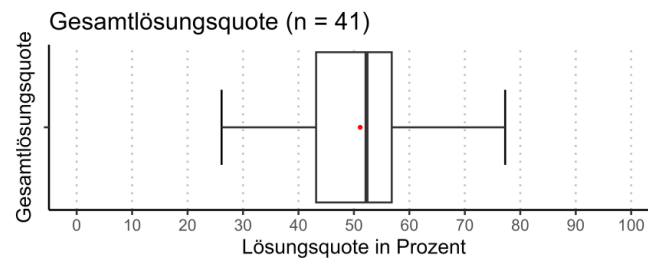
Die Ergebnisse belegen, dass die Schüler:innen im ersten Leistungsfachdurchgang L1 im Mittel mit weitestgehend gleichen mittleren Voraussetzungen in Form ihres Vorwissens (Abbildung 16) aus der Mittelstufe in die Kursstufe starten. Es ist ein Wissensfortschritt über den Verlauf der beiden Messzeitpunkte erkennbar (erste Messung: Abbildung 20a; zweite Messung: Abbildung 20b). Es zeigt sich schulübergreifend kein Unterschied beim mittleren Fachwissen am zweiten Messzeitpunkt vor dem Abitur. In den dem Abitur nachgelagerten Kompetenzteilbereichen wurde die Messung von *Datenkommunikation* durch bildungsorganisatorische Einflussfaktoren beeinträchtigt. Es erfolgt ein Ausschluss dieser Messskala und eine isolierte Betrachtung des verbleibenden Kompetenzteilbereichs *Grundlagen der Elektronik*. Die Schüler:innen erreichen in diesem Kompetenzteilbereich vergleichbare Leistungen wie in den jeweils vor dem schriftlichen Abitur abgeprüften Kompetenzteilbereichen.



**a**



**b**

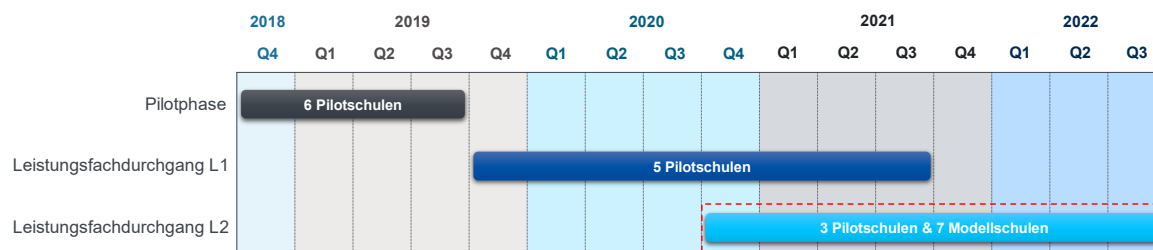


**Abbildung 20:** Entwicklung der Gesamtlösungsquote über den Verlauf zweier Testzeitpunkte im Leistungsfachdurchgang L1: a) nach einem dreiviertel Schuljahr und b) vor dem Abitur.



## 9 Fachwissen von Schüler:innen im Jahrgang 2020 – 2022 (L2)

An den Leistungsfachdurchgang L1 und der damit verbundenen Studie in Kapitel 8 schließt sich die Modellphase des Schulversuchs an. Im Leistungsfachdurchgang L2 ist neben den sechs Pilotschulen die Einmündung von 14 weiteren Modellschulen geplant.



**Abbildung 21:** Verortung der Studie zum Fachwissen im Leistungsfachdurchgang L2 im Erhebungsablaufplan.

Von den ursprünglich geplanten 14 Modellschulen können sieben das Leistungsfach in der Kursstufe durchführen. Drei der ursprünglichen sechs Pilotschulen können aufgrund geringer Anmeldezahlen keinen Kurs in L2 anbieten, wodurch sich eine insgesamt Schulzahl von zehn Schulen für den Leistungsfachdurchgang L2 ergibt (s. Abbildung 21).

### 9.1 Studiendesign

Das Studiendesign für den zweiten Leistungsfachdurchgang (L2) orientiert sich an dem der ersten Studie, um eine vergleichbare Stichprobe zu generieren. Dazu werden die Schüler:innen am Beginn der Kursstufe zu ihrem Vorwissen, nach einem dreiviertel Schuljahr und vor dem Abitur zu ihrem Fachwissen in den sieben bereits vorgestellten Kompetenzteilbereichen (Kapitel 8.1) und nachgelagert zu den Kompetenzteilbereichen *Grundlagen der Elektronik* sowie *Datenkommunikation* befragt.

Durch die größere Gesamtstichprobe der beiden Studien wird eine repräsentativere Gesamtanalyse des Schulversuchs angestrebt und in Kapitel 10 nochmals kombiniert betrachtet.

### 9.2 Datenauswertung

Die Dateneingabe, Bereinigung und Berechnung der Testergebnisse wurde analog zur ersten Befragungswelle mit *SPSS 28.0* (SPSS Inc., 2021) durchgeführt. Alle weiteren Analyseschritte wurden ebenfalls in *R* (R Core Team, 2022) unter Verwendung der bereits im vorherigen Abschnitt 8.2 angeführten Erweiterung

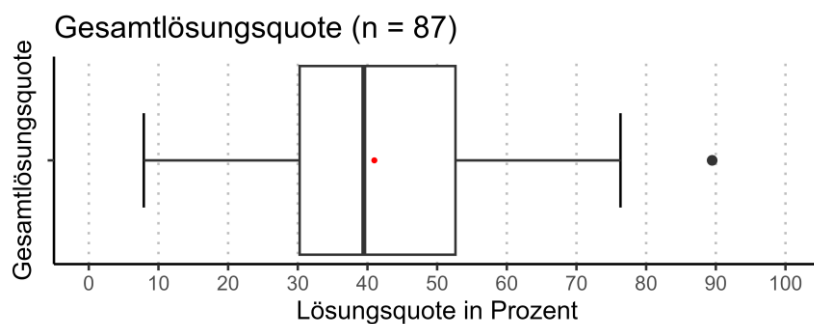
*tidyverse* (Wickham et al., 2019) und dem darin enthaltenen Paket *ggplot2* (Wickham, 2016) zur Erstellung von Diagrammen vorgenommen. Für alle Auswertungen wird ein Signifikanzniveau von  $\alpha = .05$  zugrunde gelegt.

### 9.3 Beschreibung der Stichprobe

Alle Testungen sind als Vollerhebungen anzusehen, da zu jedem Testzeitpunkt die maximal zur Verfügung stehende Anzahl an Proband:innen befragt wurde. An der Befragung zum Schuljahresbeginn 2020/21 im Oktober 2020 haben  $n_{Vorwissen, L2} = 87$  ( $w = 9.2\%$ ,  $m = 90.8\%$ ) Lernende teilgenommen. Für die Befragung zum Fachwissen am ersten Erhebungszeitpunkt nach einem dreiviertel Schuljahr im Mai 2021 konnten  $n_{Fachwissen1, L2} = 76$  ( $w = 7\%$ ,  $m = 93\%$ ) Schüler:innen akquiriert werden. Am Befragungszeitpunkt vor dem Abitur (April 2022) wurden  $n_{Fachwissen2, L2} = 74$  ( $w = 9.5\%$ ,  $m = 89.2\%$ ,  $k. A. = 1.4\%$ ) Lernende erreicht. Die Stichprobe der Abschlussbefragung im Mai/Juni 2022 umfasst  $n_{Fachwissen3, L2} = 60$  ( $w = 8.3\%$ ,  $m = 91.7\%$ ) Schüler:innen.

### 9.4 Deskriptive Auswertung des Vorwissens

Die Auswertung des Vorwissens der Schüler:innen des Leistungsfachs betrachtet die Daten der Testung zum Schuljahresbeginn 2020/21 (Oktober 2019) an den Pilot- und Modellschulen im Schulversuch. Es wurden keine Anpassungen der Aufgaben gegenüber der abschließenden Testversion aus der Entwicklungsphase und aus Studie 1 vorgenommen. Aufgrund der geringen Anzahl der Teilnehmerinnen in der Stichprobe wird auf eine geschlechtsspezifische Auswertung in Bezug auf deren Aussagekraft verzichtet.



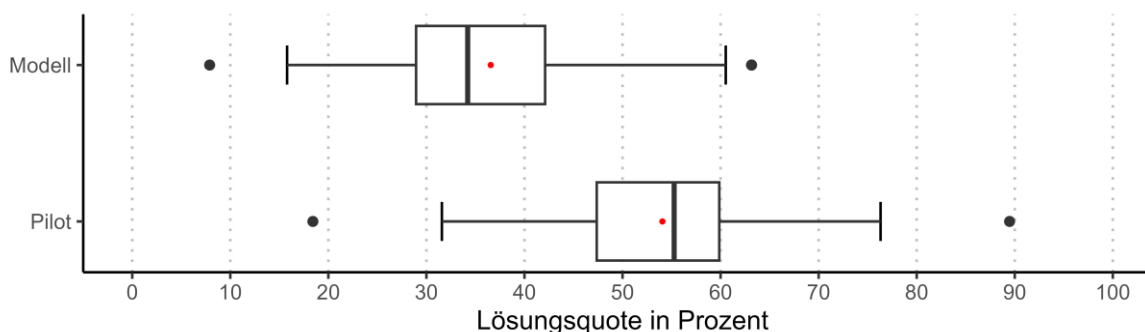
**Abbildung 22:** Testergebnis des Vorwissenstests an den Pilot- und Modellschulen im Leistungsfachdurchgang L2.

Die Lösungsquote der Gesamtstichprobe (Abbildung 22) beträgt  $Med = 39.5\%$  ( $MW = 41.0\%$ ). Eine detailliertere Übersicht der Teststatistik ist in Tabelle 15 dargestellt.

**Tabelle 15:** Ausführliche Betrachtung der Lösungsquoten des Vorwissenstests der Studie 2 (Gesamtstichprobe).

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
7.90 %	30.26 %	41.00 %	39.50 %	52.60 %	89.50 %	15.30 %

Die Schüler:innen an Pilotschulen erreichen ein mittleres Vorwissen von  $Med_{Pilot} = 55.26\%$  ( $MW = 54.07\%$ ). An den Modellschulen liegt ein mittleres Vorwissen von  $Med_{Modell} = 34.21\%$  ( $MW = 36.56\%$ ) vor. Aus den vorliegenden Unterschieden der mittleren Lösungsquoten zwischen Pilot- und Modellschulen beim Vorwissen (Abbildung 23) ist tendenziell ableitbar, dass an den Pilotschulen bereits eine stärkere Anpassung der Unterrichtsschwerpunkte in Richtung des Bildungsplans 2016 (MKJS, 2016) für die Mittelstufe und der Perspektive auf das Leistungsfach erfolgt ist. Die Daten liegen sowohl für die Gesamtstichprobe ( $W = 0.976$ ,  $p > .05$ ) als auch für die Gruppenunterteilung in Pilot- und Modellschulen (Pilot:  $W = 0.975$ ,  $p > .05$ ; Modell:  $W = 0.969$ ,  $p > .05$ ) normalverteilt nach Shapiro-Wilk vor.



**Abbildung 23:** Übersicht über die Lösungsquoten der Pilot- und Modellschulen im Vorwissenstest. Die Schüler:innen verteilen sich zu  $n_{Modell} = 65$  und  $n_{Pilot} = 22$ .

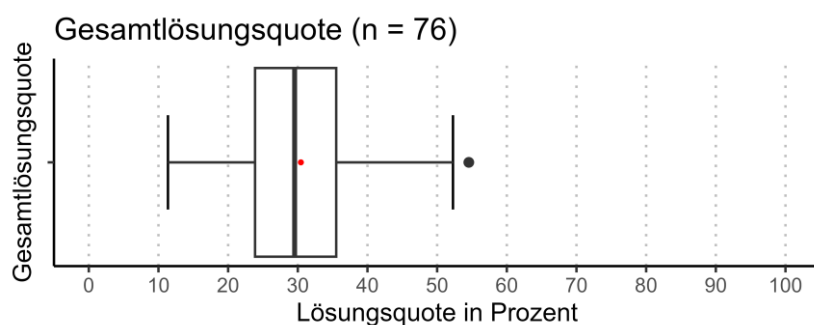
Die o. a. Vermutung zu den divergierenden mittleren Testergebnissen der Pilot- und Modellschulen wird per t-Test überprüft. Dieser bestätigt mit  $t(86) = 23.85$ ,  $p < .001$ , dass ein signifikanter Gruppenunterschied vorliegt. Innerhalb der Gruppen zeigen sich keine signifikanten Unterschiede (ANOVA; Pilot:  $F(2, 19) = 0.67$ ,  $p > .05$ ; Modell:  $F(6, 58) = 1.92$ ,  $p > .05$ ). Es kann somit festgestellt werden, dass sich bei einer signifikanten intergruppenspezifischen Divergenz keine intragruppenspezifischen Unterschiede bei den Pilot- und Modellschulen messen lassen und die Schüler:innen innerhalb ihrer Gruppen auf einem vergleichbaren mittleren Vorwissenstand starten.

## 9.5 Fachwissen im zweiten Leistungsfachdurchgang (L2)

Analog zum Leistungsfachdurchgang L1 wird die Entwicklung des Fachwissens der Lernenden an den Pilot- und Modellschulen über zwei nachfolgend getrennt beschriebene Testzeitpunkte im Abstand eines dreiviertel Schuljahres erfasst. Dazu wird die Stichprobe in L2 in den in Kapitel 8.1 beschriebenen, abiturrelevanten Kompetenzteilbereichen geprüft.

### 9.5.1 Deskriptive Auswertung des Fachwissens – erster Testzeitpunkt (L2)

Im Leistungsfachdurchgang L2 erfolgt die erste Testung des Fachwissens analog zu L1 nach einem dreiviertel Schuljahr. Mit einer mittleren Gesamtlösungsquote von  $Med = 29.55\%$  ( $MW = 30.44\%$ ) rangiert das Fachwissen in dieser Erhebungswelle erneut erwartungsgemäß im unteren Bereich der Skala (Abbildung 24).



**Abbildung 24:** Gesamtlösungsquote des Fachwissenstests an den Pilot- und Modellschulen zum ersten Testzeitpunkt im Leistungsfachdurchgang L2.

Tabelle 16 liefert eine detaillierte Übersicht über die Gesamtlösungsquote.

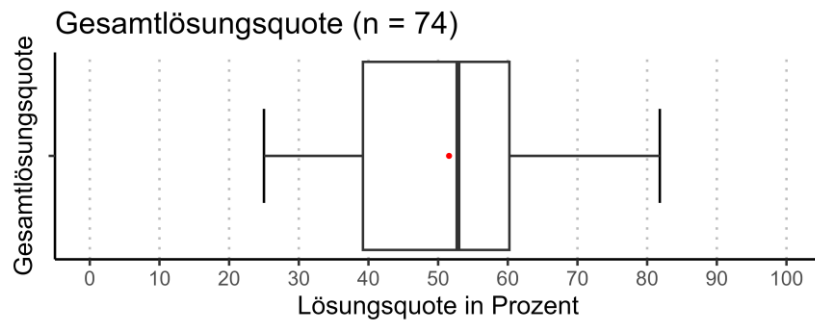
**Tabelle 16:** Aufstellung der Testergebnisse des Fachwissenstests der abiturrelevanten Kompetenzteilbereiche am Zeitpunkt der ersten Testung im Leistungsfachdurchgang L2.

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
11.36 %	23.86 %	30.44 %	29.55 %	35.51 %	54.55 %	9.75 %

Die Verteilung bereits abgeschlossen unterrichteter Kompetenzteilbereiche erweist sich in L2 ebenfalls als heterogen, weshalb keine weiteren Vergleichsanalysen für diesen Testzeitpunkt vorgenommen werden.

### 9.5.2 Deskriptive Auswertung des Fachwissens – zweiter Testzeitpunkt (L2)

Im Leistungsfachdurchgang L2 erreichen die Schüler:innen der Pilot- und Modellschulen in der zweiten Testung des Fachwissens vor dem Abitur eine mittlere Lösungsquote (Abbildung 25) von  $Med. = 52.84\%$  ( $MW = 51.57\%$ ).



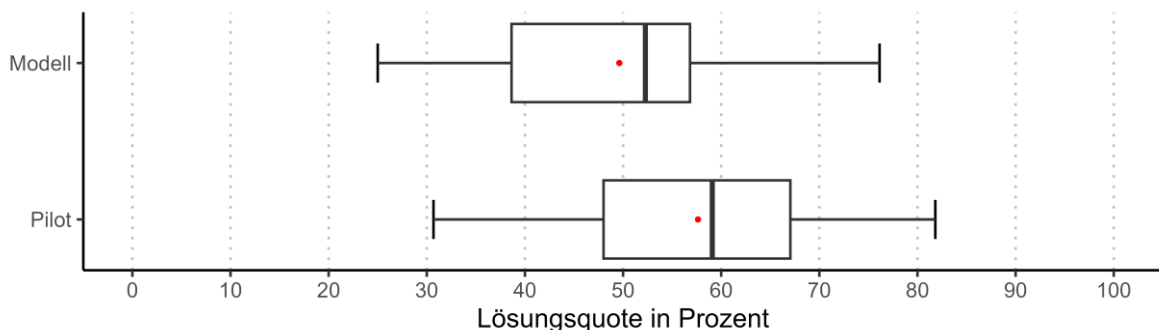
**Abbildung 25:** Gesamtlösungsquote des Fachwissenstests an den Pilot- und Modellschulen des Leistungsfachdurchgangs L2 am zweiten Testzeitpunkt vor dem Abitur.

Eine detaillierte Aufstellung der Lösungsquote ist Tabelle 17 zu entnehmen.

**Tabelle 17:** Aufstellung der Ergebnisse der zweiten Testung des Fachwissens der Schüler:innen zu den abiturrelevanten Kompetenzteilbereichen im Leistungsfachdurchgang L2.

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
25.00 %	39.20 %	51.57 %	52.84 %	60.23 %	81.82 %	13.33 %

Die Daten des Gesamtergebnisses liegen gemäß des Shapiro-Wilk-Tests normalverteilt vor ( $W = 0.98, p > .05$ ).

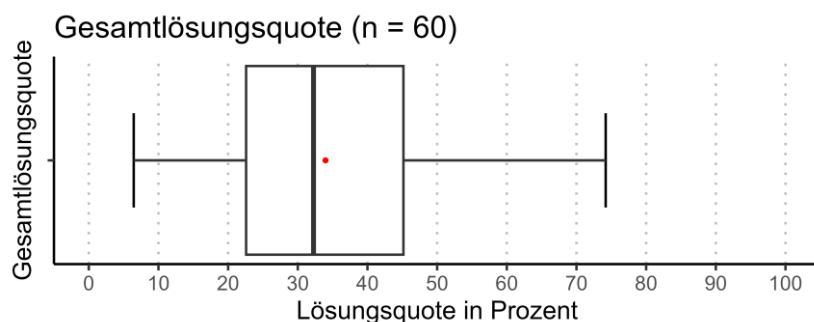


**Abbildung 26:** Übersicht über die Lösungsquoten der Pilot- und Modellschulen im Fachwissenstest des Leistungsfachdurchgangs L2 vor der Abiturprüfung. Die erfassten Ergebnisse verteilen sich zu  $n_{Modell} = 56$  auf die Modellschulen und  $n_{Pilot} = 18$  auf die Pilotschulen.

Innerhalb der Gruppen der Pilot- und Modellschulen kann die Normalverteilung der Daten mittels Shapiro-Wilk-Test ebenfalls bestätigt ( $W_{Pilot} = 0.97, p > .05$ ;  $W_{Modell} = 0.98, p > .05$ ) werden. Im Vergleich der Pilot- und Modellschulen (Abbildung 26) zeigt der t-Test einen signifikanten Gruppenunterschied ( $t(25) = 2.06, p = .0496$ ) an, der in Anbetracht des Signifikanzniveaus und der unterschiedlichen Stichprobengröße nochmals in der kombinierten Auswertung zu betrachten ist. Eine ANOVA kann darüber hinaus keinen signifikanten Unterschied zwischen den einzelnen Pilot- und Modellschulen belegen ( $F(8, 65) = 1.87, p > .05$ ).

### 9.5.3 Deskriptive Auswertung der Kompetenzteilbereiche *Grundlagen der Elektronik* und *Datenkommunikation* (L2)

In den beiden der Abiturprüfung nachgelagerten Kompetenzteilbereichen *Grundlagen der Elektronik* und *Datenkommunikation* erreichen die Schüler:innen eine Gesamtlösungsquote von  $Med = 32.26\%$  ( $MW = 33.98\%$ ).



**Abbildung 27:** Gesamtlösungsquote des Testteils *Grundlagen der Elektronik* und *Datenkommunikation* im Leistungsfachdurchgang L2.

Die Normalverteilung der Gesamtlösungsquote in den beiden Kompetenzteilbereichen kann mittels Shapiro-Wilk-Test ( $W = 0.97$ ,  $p > .05$ ) bestätigt werden. Einen detaillierten Überblick über die Ergebnisverteilung liefert Tabelle 18.

**Tabelle 18:** Übersicht der Ergebnisverteilung der Kompetenzteilbereiche *Grundlagen der Elektronik* und *Datenkommunikation* in L2.

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
6.45 %	22.58 %	33.98 %	32.26 %	45.16 %	74.19 %	15.38 %

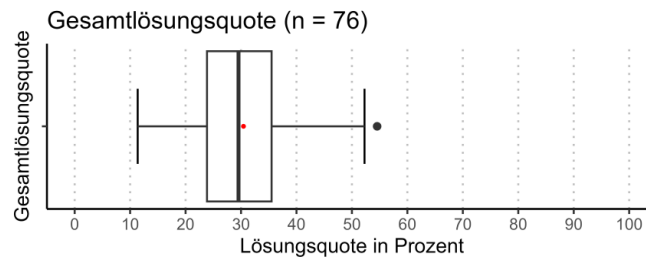
Der Kompetenzteilbereich *Datenkommunikation* konnte nach Angaben der Lehrpersonen in diesem Durchgang erneut nicht vollumfänglich unterrichtet werden. Unter Ausschluss von *Datenkommunikation* resultiert für *Grundlagen der Elektronik* eine mittlere Lösungsquote von  $Med = 41.18\%$  ( $MW = 41.96\%$ ) bei einer bestätigten Normalverteilung ( $W = 0.98$ ,  $p > .05$ ).

### 9.6 Zwischenfazit zu den Ergebnissen aus Studie L2

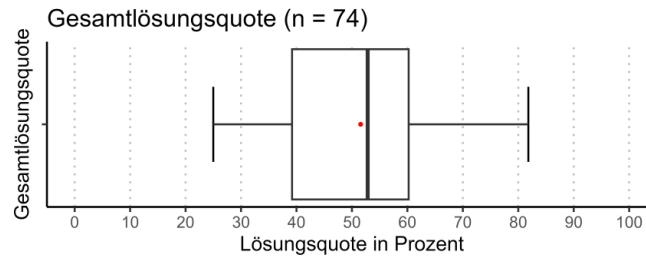
Beim Start in die Kursstufe besteht zwischen den Schüler:innen der Pilot- und Modellschulen im Leistungsfachdurchgang L2 ein signifikanter Unterschied bei deren mittlerem Vorwissen zugunsten der Lernenden an Pilotschulen. Die Ergebnisse der Erhebung des Fachwissens in L2 belegen einen Fortschritt der inhaltsbezogenen Kompetenzfacetten über den Zeitverlauf der Kursstufe bis zur Abiturprüfung (Abbildung 28).



**a**



**b**



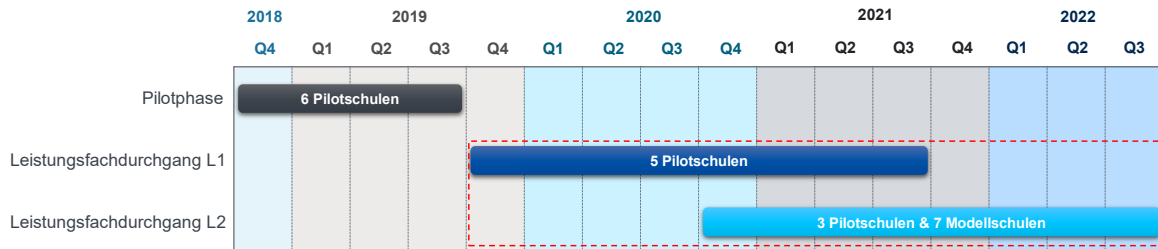
**Abbildung 28:** Vergleich der Gesamtlösungsquoten der beiden Testzeitpunkte nach einem drei- viertel Schuljahr (a) und vor dem Abitur (b) im Leistungsfachdurchgang L2.

Neben signifikanten Unterschieden des Vorwissens der Schüler:innen beim Start in die Kursstufe zeigen sich diese ebenfalls beim Fachwissen vor der Abiturprüfung. Der statistisch signifikante Befund im Gruppenvergleich erscheint auf Basis des Signifikanzniveaus jedoch wenig belastbar. Als Limitation diesbezüglich ist zusätzlich die unterschiedliche Stichprobengröße in den Gruppen anzuführen. Eine erneute Prüfung ermöglicht die kombinierte Betrachtung beider Stichproben in Kapitel 10. Der Kompetenzteilbereich *Datenkommunikation* konnte aufgrund zeitlicher Faktoren in diesem Leistungsfachdurchgang laut Angabe der Lehrpersonen erneut nicht vollumfänglich unterrichtet werden und wird deshalb aus der Auswertung ausgeschlossen. Für *Grundlagen der Elektronik* ergibt sich eine vergleichbare Lösungsquote mit den Kompetenzteilbereichen, die vor der Abiturprüfung erfasst wurden.



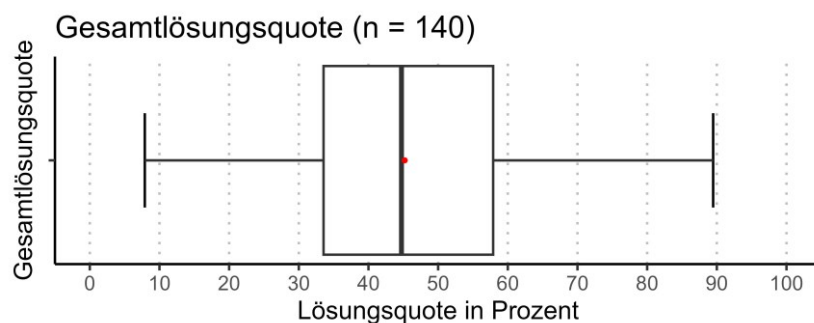
## 10 Betrachtung der kombinierten Stichprobe L1 & L2

Nach beiden Leistungsfachdurchgängen L1 und L2 wird die Stichprobe aufgrund des identischen Studiendesigns kombiniert (Abbildung 29) und eine Auswertung für die Gesamtstichprobe vorgenommen.



**Abbildung 29:** Verortung der kombinierten Betrachtung des Fachwissens der Leistungsfachdurchgänge L1 und L2 im Erhebungsablaufplan.

Daraus resultieren Gesamtstichproben von  $n_{VW, ges} = 140$  ( $w = 9.3 \%$ ,  $m = 90.7 \%$ ),  $n_{FW1, ges} = 127$  ( $w = 6.3 \%$ ,  $m = 93.7 \%$ ),  $n_{FW2, ges} = 115$  ( $w = 7.7 \%$ ,  $m = 91.3 \%$ ) und  $n_{GE-DK, ges} = 102$  ( $w = 7.8 \%$ ,  $m = 91.2 \%$ ) Schüler:innen für die jeweiligen Befragungszeitpunkte. Es fällt auf, dass in beiden Leistungsfachdurchgängen L1 und L2 nur wenige Mädchen das Leistungsfach belegen. Inwiefern sich mehr Mädchen für das Leistungsfach gewinnen lassen und welche Lehr-Lern-Kontexte dazu förderlich sind, kann Gegenstand zukünftiger Untersuchungen sein. Die Schüler:innen in der Gesamtstichprobe sind im Mittel  $MW_{Alter} = 16.35$  Jahre alt und weisen einen durchschnittlichen IQ von  $MW_{IQ} = 109.43$  auf.

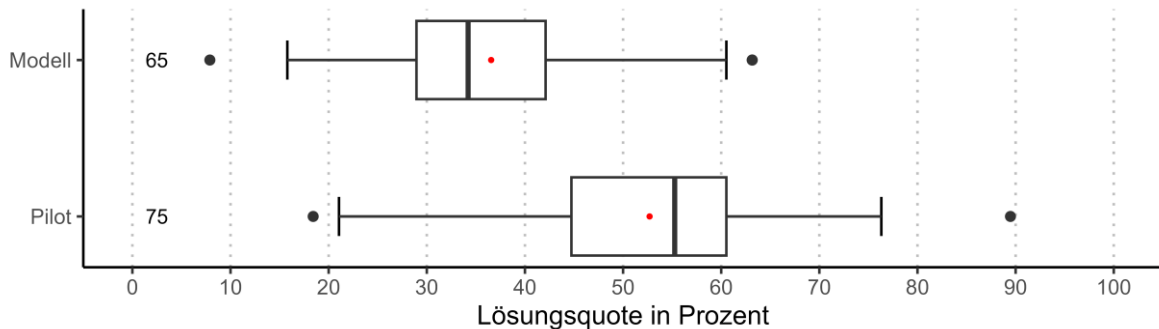


**Abbildung 30:** Gesamtlösungsquote des Vorwissenstests der kombinierten Stichprobe aus L1 und L2.

Die mittlere Lösungsquote des Vorwissens für die kombinierte Stichprobe (Abbildung 30) beträgt  $Med_{VW} = 44.74 \%$  ( $MW_{VW} = 45.21 \%$ ) bei einer bestätigten Normalverteilung der Daten (Shapiro-Wilk-Test:  $W = 0.99$ ,  $p > .05$ ).

Für den Vergleich der Pilot- und Modellschulen (Abbildung 31) wird bei bestätigter Normalverteilung des Testergebnisses in beiden Gruppen

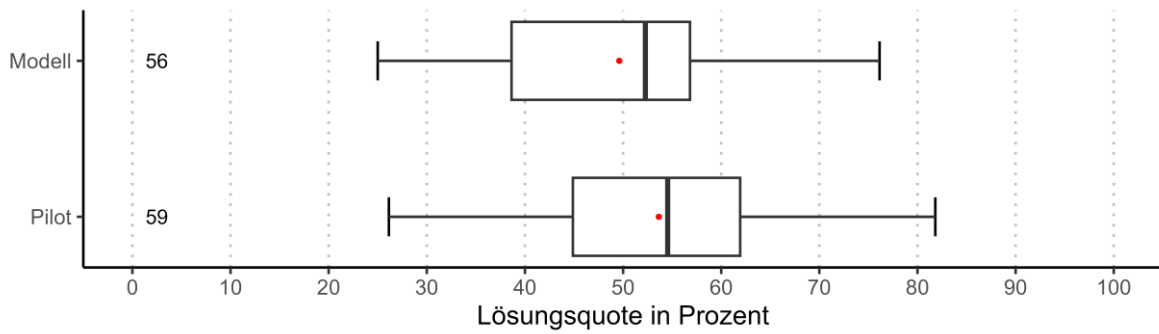
( $W_{Modell, ges, VW} = 0.97, p > .05$ ;  $W_{Pilot, ges, VW} = 0.97, p > .05$ ) ein t-Test berechnet. Dieser zeigt signifikante Gruppenunterschiede ( $t(138) = 7.26, p < .001$ ).



**Abbildung 31:** Vergleich der Lösungsquoten von Pilot- und Modellschulen im Vorwissenstest. Der Zahlenwert vor den Boxplots entspricht der jeweiligen Anzahl der Schüler:innen in den Gruppen.

Die Effektstärkeberechnung (*Cohens d*) erfolgt mit dem Package *rstatix* (Kassambara, 2023). Für den Gruppenunterschied zwischen Pilot- und Modellschulen beim mittleren Vorwissen der Lernenden kann ein großer Effekt von  $d = 1.23$  festgestellt werden. Innerhalb der Gruppen der Pilot- und Modellschulen sind keine signifikanten Unterschiede zwischen Lernenden an den einzelnen Schulen belegbar, wonach davon auszugehen ist, dass die Schüler:innen innerhalb ihrer jeweiligen Gruppe mit vergleichbaren mittleren Voraussetzungen in die Kursstufe einmünden ( $F_{Modell}(6, 58) = 1.92, p > .05$ ;  $F_{Pilot}(4, 70) = 0.098, p > .05$ ).

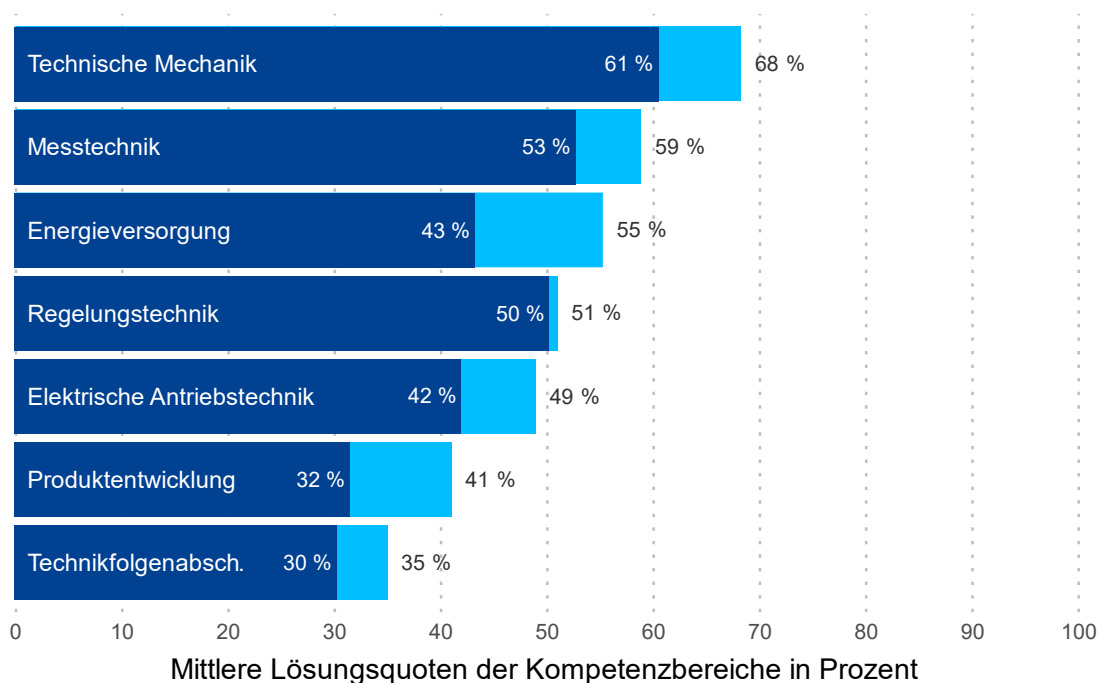
Eine Progression des Fachwissens der Gesamtstichprobe über die zwei Messzeitpunkte nach einem dreiviertel Schuljahr und vor der Abiturprüfung kann durch die Steigerung der Gesamtlösungsquote vom ersten ( $Med_{FW1, ges} = 29.55\%$ ,  $MW_{FW1, ges} = 29.44\%$ ) auf den zweiten Messzeitpunkt ( $Med_{FW2, ges} = 52.27\%$ ,  $MW_{FW2, ges} = 51.68\%$ ) festgestellt werden. Ein Vergleich des Fachwissens der Schüler:innen an den einzelnen Pilot- und Modellschulen (Abbildung 32) mittels ANOVA ( $F(11, 103) = 0.85, p > .05$ ) belegt, dass vor der Abiturprüfung schulübergreifend keine Unterschiede beim Mittelwert des Fachwissens vorliegen. Der Gruppenvergleich des mittleren Fachwissens der zusammengefassten Pilot- und Modellschulen ist ebenfalls nicht signifikant ( $t(113) = 1.76, p > .05$ ). Eine detaillierte Darstellung der Verteilung des Fachwissens der Gesamtstichprobe vor der Abiturprüfung liefert Tabelle 19. Der Fortschritt des Fachwissens zu den einzelnen Kompetenzteilbereichen ist in Abbildung 33 dargestellt. Bei der Feststellung des Fachwissens am ersten Testzeitpunkt (dunkelblau) wurden nur Schüler:innen berücksichtigt, die das Thema bereits vollständig im Unterricht behandelt haben.



**Abbildung 32:** Vergleich der Lösungsquoten von Pilot- und Modellschulen im Fachwissen vor der Abiturprüfung. Der Zahlenwert vor den Boxplots entspricht der jeweiligen Anzahl der Schüler:innen in den Gruppen.

**Tabelle 19:** Detaillierte Übersicht des Fachwissens der kombinierten Stichprobe der Pilot- und Modellschulen aus L1 und L2 vor der Abiturprüfung.

<i>Min. LQ</i>	<i>Q1</i>	<i>MW</i>	<i>Med</i>	<i>Q3</i>	<i>Max. LQ</i>	<i>SD</i>
25.00 %	42.05 %	51.68 %	52.27 %	57.96 %	81.82 %	12.36 %



**Abbildung 33:** Übersicht des Fachwissensfortschritts der Schüler:innen in den einzelnen Kompetenzteilbereichen des Bildungsplans zwischen den beiden Messzeitpunkten nach einem dreiviertel Schuljahr und vor der Abiturprüfung.

Die rein deskriptive Analyse lässt keinen direkten Rückschluss auf die tatsächliche Personenfähigkeit der Schüler:innen in den einzelnen Kompetenzteilbereichen zu, wenngleich sie erwartungskonform eine Progression zwischen den beiden Messzeitpunkten belegt.

Das Fachwissen in den Bereichen *Grundlagen der Elektronik* und *Datenkommunikation* wird im Folgenden aus den bereits diskutierten Gründen lediglich für *Grundlagen der Elektronik* betrachtet. Es zeigt sich, dass die Schüler:innen in diesem Thema ebenfalls ein schulübergreifendes Fachwissen erreichen, das sich im Mittel nicht signifikant unterscheidet ( $Med = 41.18 \%$ ,  $MW = 41.46 \%$ ;  $F(11, 90) = 1.23$ ,  $p > .05$ ).

### 10.1 Zwischenfazit zu den deskriptiven Befunden

In Bezug auf die Beantwortung der Forschungsfrage F2 zeigt sich, dass die Schüler:innen mit einem standortabhängigen Vorwissen in die Kursstufe starten. Lernende an Pilotschulen weisen im Mittel ein signifikant höheres Vorwissen auf als diejenigen an Modellschulen. Innerhalb ihrer jeweiligen Gruppe sind keine nennenswerten Unterschiede zwischen den Lernenden an Pilot- oder Modellschulen festzustellen. Diese Befundlage spiegelt möglicherweise die zeitlich längere Teilnahme der Pilotschulen am Schulversuch wider, kann durch die vorliegenden Daten jedoch nicht abschließend geklärt werden. Die Beobachtung der Fachwissensentwicklung und insbesondere der Fachwissensstand der Schüler:innen vor dem Abitur ermöglicht die Beantwortung der Forschungsfrage F3. Die Zielerreichung eines sich schulübergreifend nicht signifikant unterscheidenden mittleren Fachwissens sind aufgrund der heterogenen Startvoraussetzungen nicht als gesichert anzusehen. Die Ergebnisse belegen eine Progression des Fachwissens in allen Kompetenzteilbereichen über den Verlauf der Kursstufe an zwei Messzeitpunkten. Darüber hinaus liegen schulübergreifend keine signifikanten Unterschiede des mittleren Fachwissens an den Pilot- und Modellschulen vor der Abiturprüfung vor.

In den Leistungsfachdurchgängen L1 und L2 konnten keine ausreichend belastbaren Ergebnisse zum Kompetenzteilbereich *Datenkommunikation* generiert werden. In beiden Fällen führen die Lehrpersonen zeitliche Probleme bei der Durchführung des Kompetenzteilbereichs nach dem Abitur an. Um dahingehend Optimierungspotenziale auszuschöpfen, könnte eine bessere Vernetzung des Kompetenzteilbereichs gegenüber einer dem Abitur nachgelagerten epochalen Umsetzung angestrebt werden. Als weitere Limitation tritt auf, dass durch die erreichten mittleren Lösungsquoten in den einzelnen Kompetenzteilbereichen

auf Basis der deskriptiven Analyse kein direkter Rückschluss auf die tatsächliche Fähigkeit der Schüler:innen in diesen Kompetenzteilbereichen möglich ist. Es lässt sich auf deskriptiver Ebene lediglich eine Progression belegen. Um dieser Limitation zu begegnen, schließt sich Kapitel 11 an. Es erfolgt dort eine Skalierung auf Basis des Rasch-Modells, um Itemschwierigkeiten und Personenfähigkeiten zu schätzen.





## 11 Schätzen der Personenfähigkeiten

Das Analysepotenzial der Skalen zur Erfassung des Vor- und Fachwissens lässt sich mittels IRT-Skalierung anhand von Rasch-Modellen ausbauen. Durch die Schätzung der Personenfähigkeit werden Aussagen über die Ausprägung des Fachwissens der einzelnen Testpersonen und deren Vergleich ermöglicht. Mit der kombinierten Stichprobe aus den beiden durchgeführten Studien L1 und L2 von insgesamt  $n > 100$  (s. z. B. Chen et al., 2014) für alle relevanten Erhebungszeitpunkte kann die Modellierung statistisch aussagekräftig vorgenommen werden. Die weiterführende Analyse dient einerseits dem Zweck der Überprüfung der erhaltenen Skalen aus der Klassischen Testtheorie. Da sich andererseits über die Eindimensionalität eines erfassten Konstrukts mittels klassischer Testtheorie keine Aussage treffen lässt, erscheint eine abschließende Beschreibung des Vor- und Fachwissens in Bezug auf die Kompetenzbereiche der Bildungspläne (Autorengruppe Bildungsplan NwT, 2019b; MKJS, 2016) sinnvoll. Zudem können mittels der geschätzten Personenfähigkeiten Niveaustufen des Fachwissens abgeleitet werden. In den nachfolgenden Abschnitten wird die Auswertung mittels Rasch-Modell mit der kombinierten Stichprobe der Studien L1 (Kapitel 8) und L2 (Kapitel 9) vorgenommen.

Die IRT-Skalierung der Kompetenzbereiche des Vorwissens aus der Mittelstufe wird mit der kombinierten Stichprobe ( $n_{VV, ges} = 140$ ) aus L1 und L2 berechnet. Für die Skalierung des Fachwissenstests eignet sich der Befragungszeitpunkt vor der Abiturprüfung mit einer kombinierten Stichprobengröße von  $n_{FW2, ges} = 115$  Lernenden aus L1 und L2. An diesem Messzeitpunkt sollten die Schüler:innen in beiden Leistungsfachdurchgängen den höchsten Fachwissensstand erreichen sowie alle abiturrelevanten Kompetenzteilbereiche vollständig abbilden können und damit eine verlässliche Aussage hinsichtlich ihrer Personenfähigkeit ermöglichen. Die Skalierung der Skala *Grundlagen der Elektronik* erfolgt auf Basis der Datenerhebung nach der Abiturprüfung ( $n_{GE-DK, ges} = 102$ ). Für *Datenkommunikation* wird aufgrund der bereits diskutierten Erhebungsproblematik keine Skalierung vorgenommen. Für die Berechnung der Modelle und zugehöriger Kennwerte werden die Packages *eRm* (Mair, Hatzinger & Maier, 2021) und *TAM* (Robitzsch, Kiefer & Wu, 2022) verwendet. Die detaillierten Kennwerte der berechneten Rasch-Modelle können *Anhang 6: Ergebnisse der IRT-Skalierung* entnommen werden. Für alle Auswertungen wird ein Signifikanzniveau von  $\alpha = .05$  zugrunde gelegt.

### 11.1 Skalierung des Vorwissenstests

Die Ausgestaltung des Bildungsplans (MKJS, 2016) legt eine differenzierte Betrachtung des Vorwissens als mehrdimensionales Gesamtkonstrukt nahe. Dazu wird die Modellierung des Vorwissens an den normativen Kompetenzbereichen des Bildungsplans ausgerichtet. Zunächst wird überprüft, ob sich für jeden Kompetenzbereich ein eindimensionales Rasch-Modell berechnen lässt. Ein gültiges Rasch-Modell der Kompetenzbereiche *Energie und Mobilität* sowie *Stoffe und Produkte* kann zwar berechnet werden, die erhaltenen Reliabilitäten der Skalen weisen mit  $EAP Rel_{\text{Stoffe/Produkte}} = 0.49$  und  $EAP Rel_{\text{Energie/Mobilität}} = 0.47$  eine deutliche Messfehlerbehaftung auf (vgl. Adams, 2005, S. 170). Dieser Sachverhalt ist möglicherweise auf eine zu geringe Anzahl an Items für jeden Kompetenzteilbereich (Itemanzahl < 10) zurückzuführen oder deutet auf eine Vernetzung der Kompetenzbereiche hin, da sich mit der Kombination aus *Energie und Mobilität* sowie *Stoffe und Produkte* ein Modell mit zufriedenstellenden Parametern (Tabelle 20) berechnen lässt. Eine Auseinandersetzung mit diesem Sachverhalt kann in zukünftigen Studien im Rahmen einer Testrevision und weiteren Strukturaufklärung erfolgen. Für den Kompetenzbereich *Informationsaufnahme und -verarbeitung* resultiert ein Modell mit zufriedenstellenden Modellparametern (Tabelle 20).

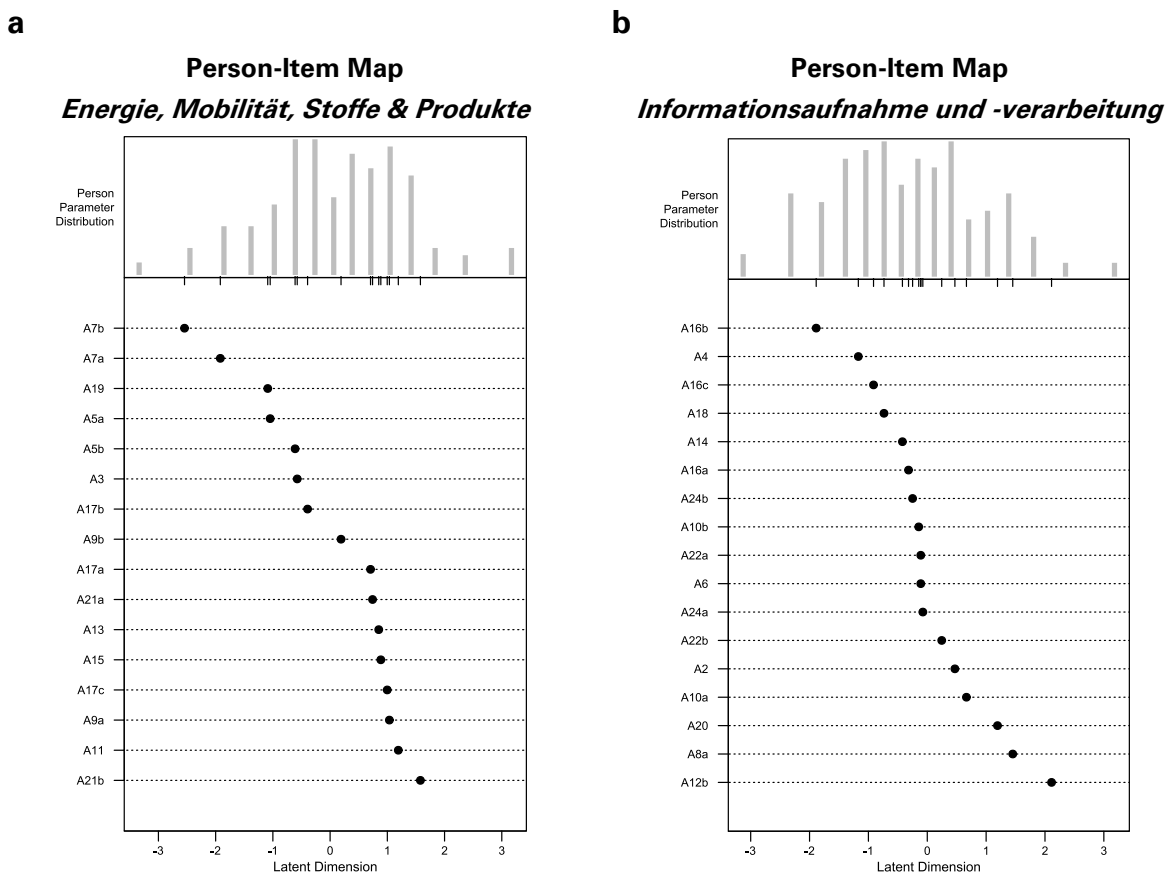
**Tabelle 20:** IRT-Skalierung des Vorwissens anhand der Dimensionen *Energie, Mobilität, Stoffe & Produkte* und *Informationsaufnahme und -verarbeitung*.

	<b>Energie, Mobilität, Stoffe &amp; Produkte</b>	<b>Informationsaufnahme und -verarbeitung</b>
<b>Andersen LRT</b>	$ALRT = 19.17, p > .05$	$ALRT = 16.88, p > .05$
<b>Wald-Test</b>	$p > .05$ für alle Items	$p > .05$ für alle Items <sup>c</sup>
<b>Martin Löff Test</b>	$MLT = 39.13, p > .05$	$MLT = 41.98, p > .05$
<b>MNSQ Infit<sup>a</sup></b>	0.85, 1.09	0.85, 1.12
<b>EAP Rel</b>	0.71	0.74
<b>WLE Rel</b>	0.68	0.71
<b>Personenfähigkeit <math>\hat{\theta}^b</math></b>	$-3.35 < 0.19 < 3.22$	$-3.13 < -0.44 < 3.18$
<b>Itemschwierigkeit <math>\hat{\varphi}^b</math></b>	$-2.54 < 0.19 < 1.58$	$-1.88 < -0.12 < 2.11$

<sup>a</sup> Bericht des minimalen und maximalen Werts. <sup>b</sup> Spanne der Personenfähigkeit und Itemschwierigkeit umfasst den Minimalwert, Median und Maximalwert. <sup>c</sup> Aufgabe A14 zeigt Auffälligkeiten beim Wald-Test, verhält sich jedoch gemäß MNSQ Infit modellkonform. Ein Ausschluss dieses Items verschlechtert zudem die Reliabilität der Skala. Aufgabe A14 wird aus diesem Grund nicht ausgeschlossen.

Mit den Items des Kompetenzbereichs *Systeme und Prozesse* lässt sich kein eigenständiges Rasch-Modell berechnen. Eine Aufteilung auf die übrigen Kompetenzbereiche führt ebenfalls zu einem Ausschluss der Items. Für diesen Kompetenzbereich sollten im Rahmen einer Testrevison weitere Aufgaben entwickelt und die bereits bestehenden überarbeitet werden.

Das erste Modell bildet damit eine Vorwissensdimension anhand der Kompetenzteilbereiche *Energie in Natur und Technik, Energieversorgungssysteme, Bewegung und Fortbewegung, Eigenschaften von Stoffen, Statische Prinzipien in Natur und Technik* sowie *Produktentwicklung* des Bildungsplans der Mittelstufe (MKJS, 2016) ab. Das zweite Modell repräsentiert eine weitere Vorwissensdimension über die Kompetenzteilbereiche *Informationsaufnahme durch Sinne und Sensoren, Gewinnung und Auswertung von Daten, Informationsverarbeitung* sowie *elektronische Schaltungen* des Mittelstufenbildungsplans (ebd.). Die Skala *Energie, Mobilität, Stoffe und Produkte* umfasst die Aufgaben *A3, A5a, A5b, A7a, A7b, A9a, A9b, A11, A13, A15, A17a, A17b, A17c, A19, A21a* und *A21b*. Aufgabe *A23* wird aufgrund fehlender Modellpassung ausgeschlossen.



**Abbildung 34:** Person-Item Maps der Dimensionen (a) *Energie, Mobilität, Stoffe & Produkte* und (b) *Informationsaufnahme und -verarbeitung*.

Das Rasch Modell für *Informationsaufnahme und -verarbeitung* ist für die Aufgaben *A2, A4, A6, A8a, A10a, A10b, A12b, A14, A16a, A16b, A16c, A18, A20, A22a, A22b, A24a* und *A24b* gültig. Die Aufgaben *A1a, A1b, A12a* und *A8b* werden aufgrund fehlender Modellpassung ausgeschlossen. Die Person-Item Maps in Abbildung 34 zeigen eine zufriedenstellende Verteilung der Aufgaben und der Personenfähigkeiten.

Der Mittelwert der geschätzten Itemschwierigkeit  $\hat{\phi}$  belegt, dass die Skala *Energie, Mobilität, Stoffe und Produkte* etwas schwerer ist als die Skala *Informationsaufnahme und -verarbeitung*. Zusammengefasst zeigt sich, dass das mit dem Testinstrument erfasste Vorwissen als zweidimensionales Konstrukt dargestellt werden kann.

### 11.2 Skalierung des Fachwissenstests

Die Testinstrumente zur Messung des Fachwissens bilden die neun inhaltsbezogenen Kompetenzteilbereiche des Leistungsfachbildungsplans ab (Autorengruppe Bildungsplan NwT, 2019b). Die Bezeichnung der Items orientiert sich an den finalen Skalen nach der Pilotierung und kann über Anhang 5: Pilotierungsergebnisse der Fachwissenstests (Leistungsfach) im Testheft zugeordnet werden.

Die Kompetenzteilbereiche des Leistungsfachs sind in analoger Weise zum Mittelstufenbildungsplan (MKJS, 2016) in Kompetenzbereichen aus zwei oder mehr Kompetenzteilbereichen zusammengefasst (Kapitel 2.4.2). Aus diesem Grund wird nicht von einem neundimensionalen Konstrukt aus abgrenzbaren Kompetenzteilbereichen ausgegangen. Vielmehr wird eine integrierte Förderung der Teilkompetenzen innerhalb abgrenzbarer Kompetenzbereiche vermutet. Als Grundannahme für die integrierte Förderung von Teilbereichen innerhalb eines Kompetenzbereichs sollten dazu die Lösungsquoten der jeweiligen Kompetenzteilbereiche signifikant zusammenhängen. Im Folgenden werden deshalb zunächst Korrelationen der Lösungsquoten der Kompetenzteilbereiche untersucht, um festzustellen, ob die Lösungsquoten den vermuteten Zusammenhang aufweisen (Tabelle 21). Die Korrelationen werden mit dem Package *correlation* (Makowski, Ben-Shachar, Patil & Lüdtke, 2020) berechnet. Die Korrelationsanalyse indiziert, dass die Lösungsquoten der Testskalen der Kompetenzteilbereiche u. a. zusammenhängen wie sie im Bildungsplan (Autorengruppe Bildungsplan NwT, 2019b) als Kompetenzbereiche subsummiert sind. Es bestehen darüber hinaus noch weitere Korrelationen, die dann berücksichtigt werden, wenn sich ein Kompetenzteilbereich inhaltlich und kontextuell nicht auf den ihm zugeordneten

Kompetenzbereich des Bildungsplans (ebd.) beschränkt (z. B. im Fall von *Technikfolgenabschätzung*).

**Tabelle 21:** Korrelative Zusammenhänge der prozentualen Lösungsquoten in den zu modellierenden Kompetenzteilbereichen.

	Ta	Ev	EA	TM	Pe	Mt	Rt
Ev	0.28	-					
EA	0.19	0.47***	-				
TM	0.32*	0.40***	0.41***	-			
Pe	0.21	0.39***	0.18	0.33**	-		
Mt	0.16	0.43***	0.52***	0.29*	0.13	-	
Rt	0.25	0.10	0.21	0.25	0.25	0.41***	-
GE	0.23	0.46***	0.57***	0.37***	0.33	0.35*	0.23

Ta = Technikfolgenabschätzung, Ev = Energieversorgung, EA = Elektrische Antriebstechnik, TM = Technische Mechanik, Pe = Produktentwicklung, Mt = Messtechnik, Rt = Regelungstechnik; GE = Grundlagen der Elektronik; \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ; Interpretation der Stärke des Zusammenhangs nach Cohen (1988); farblich hervorgehobene korrelierende Kompetenzteilbereiche werden kombiniert, ausgegraute signifikante Korrelationen werden vernachlässigt.

Im Bildungsplan finden sich Querverweise der Kompetenzfacetten der *Technikfolgenabschätzung* zu den Kompetenzbereichen *Energie und Antrieb* sowie *Technische Mechanik und Produktentwicklung* (Autorengruppe Bildungsplan NwT, 2019b, S. 16). Zur Modellierung wird *Technikfolgenabschätzung* aufgrund der einzig signifikanten Korrelation mit *Technische Mechanik* dem Kompetenzbereich *Technische Mechanik und Produktentwicklung* zugeteilt. Bei *Grundlagen der Elektronik* handelt es sich ebenfalls um einen Kompetenzteilbereich, der zwar formal dem Kompetenzbereich *Elektro- und Informationstechnik* zugeordnet ist, sich inhaltlich begründet jedoch nicht ausschließlich auf diesen Kompetenzbereich beschränkt (Autorengruppe Bildungsplan NwT, 2019b, S. 21) sondern mit mehreren Kompetenzteilbereichen kombiniert werden kann (z. B. *Elektrische Antriebstechnik, Energieversorgung, Messtechnik, Produktentwicklung*). Dadurch ist inhaltlich sowie auf Basis der erhaltenen Korrelationen keine eindeutige Zuordnung zu einem Kompetenzbereich möglich. Es wird daher eine Modellierung als eigenständiger Kompetenzteilbereich fokussiert. *Datenkommunikation* wird aufgrund bereits diskutierter Messungenauigkeiten nicht modelliert.

Durch die Korrelationsanalyse und die inhaltliche Begründung ergeben sich vier Dimensionen für die Abbildung des Fachwissens der Kursstufe, zu denen im Folgenden Rasch-Modelle berechnet werden:

- *Energieversorgung und Elektrische Antriebstechnik*
- *Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung*
- *Messtechnik und Regelungstechnik*
- *Grundlagen der Elektronik*

In Tabelle 22 werden die erhaltenen Modellparameter für die ersten drei Dimensionen dargestellt. Für alle Dimensionen lassen sich gültige Modelle bestätigen.

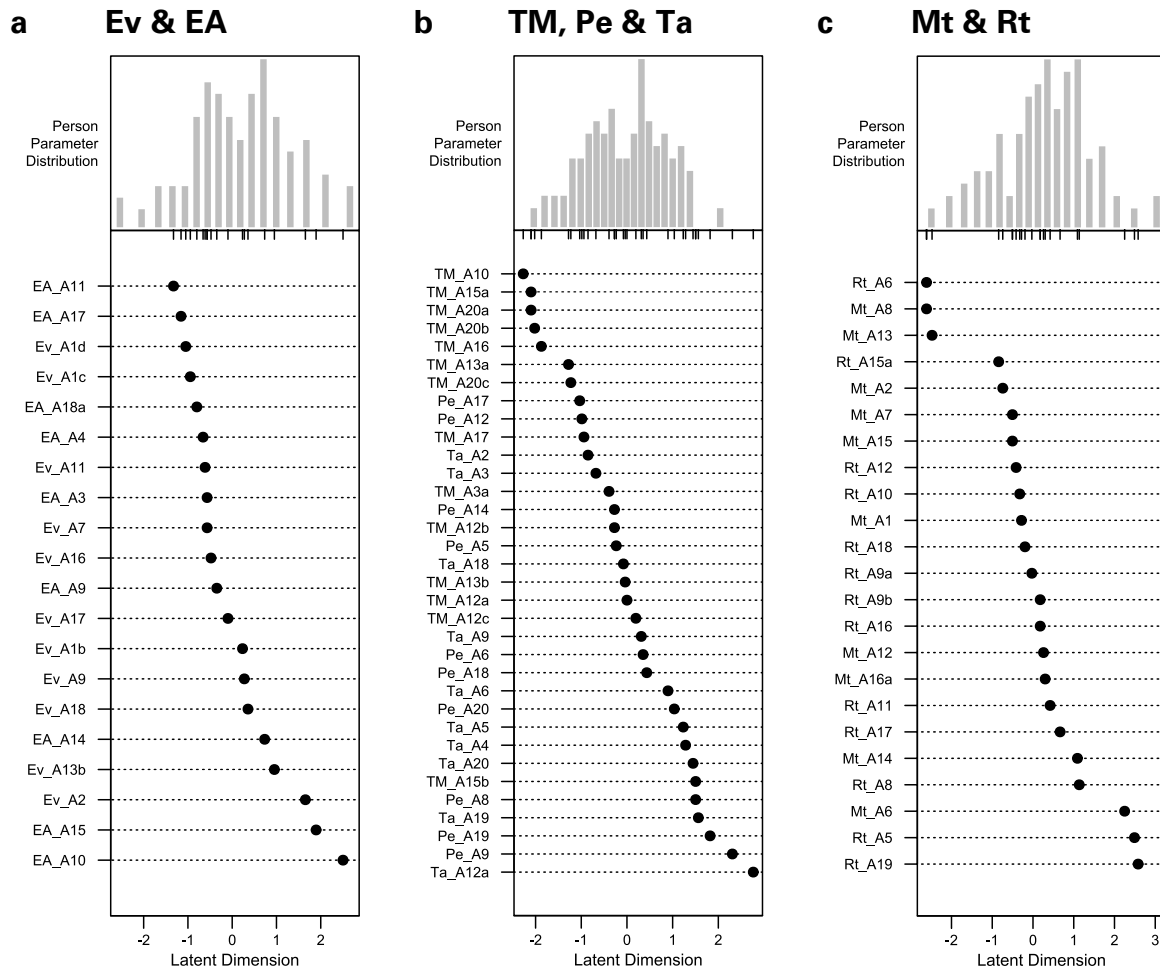
**Tabelle 22:** Modellparameter der Fachwissensdimensionen *Energieversorgung und Elektrische Antriebstechnik*; *Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung* sowie *Messtechnik und Regelungstechnik*.

	<b>Ev &amp; EA</b>	<b>TM, Pe &amp; Ta</b>	<b>Mt &amp; Rt</b>
<b>Andersen LRT</b>	$ALRT = 21.30, p > .05$	$ALRT = 22.82, p > .05$	$ALRT = 23.85, p > .05$
<b>Wald-Test</b>	$p > .05$ für alle Items <sup>c</sup>	$p > .05$ für alle Items	$p > .05$ für alle Items
<b>Martin Löf Test</b>	$MLT = 59.01, p > .05$	$MLT = 116.84, p > .05$	$MLT = 72.29, p > .05$
<b>MNSQ Infit<sup>a</sup></b>	0.81, 1.16	0.85, 1.09	0.77, 1.14 <sup>d</sup>
<b>EAP Rel</b>	0.75	0.75	0.76
<b>WLE Rel</b>	0.73	0.74	0.76
<b>Personenfähigkeit <math>\hat{\theta}^b</math></b>	-2.54 < 0.18 < 2.66	-2.04 < -0.01 < 2.04	-2.49 < 0.35 < 3.03
<b>Itemschwierigkeit <math>\hat{\varphi}^b</math></b>	-1.33 < -0.47 < 2.50	-2.27 < -0.04 < 2.76	-2.61 < 0.07 < 2.58

Ev & EA = *Energieversorgung und Elektrische Antriebstechnik*; TM, Pe & Ta = *Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung*; Mt & Rt = *Messtechnik und Regelungstechnik*; <sup>a</sup> Bericht des minimalen und maximalen Werts. <sup>b</sup> Spanne der Personenfähigkeit und Itemschwierigkeit umfasst den Minimalwert, Median und Maximalwert; <sup>c</sup> Aufgabe Ev\_A9 zeigt Auffälligkeiten beim Wald-Test, verhält sich aber gemäß MNSQ Infit modellkonform. Ein Ausschluss verschlechtert die Skalenreliabilität, das Item wird beibehalten; <sup>d</sup> Unterer MNSQ Infit Wert ist kleiner 0.8. Ein Ausschluss erfolgt nicht, da laut Wald-Test alle Items modellkonform sind.

Die Modellierung der Skala *Energieversorgung und Elektrische Antriebstechnik* ist unter Ausschluss der Items *EA\_A12*, *EA\_A13a*, *EA\_A13b*, *EA\_A18b* und *Ev\_A13a* möglich. Die Skala aus den übrigen Items und deren Aufgabenschwierigkeit ist in Abbildung 35a dargestellt. Das Rasch-Modell für *Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung* wird nach Ausschluss von *TM\_A3b*, *TM\_A9*, *Pe\_A10*, *Ta\_A12b* und *Ta\_A17* gültig. Eine Übersicht der finalen Items und deren Schwierigkeiten liefert Abbildung 35b. Nach Eliminierung von *Mt\_A5* ergibt sich für *Messtechnik und Regelungstechnik* ein gültiges

Rasch-Modell. Die finale Skala und deren Items sind in Abbildung 35c aufgetragen.



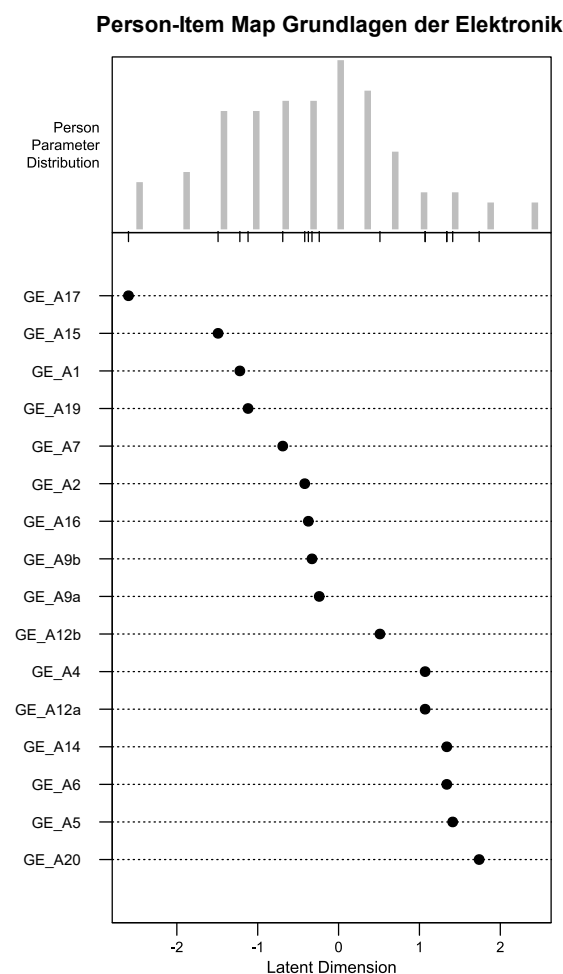
**Abbildung 35:** Finale Skalen der Dimensionen *Energieversorgung und Elektrische Antriebstechnik*; *Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung* sowie *Messtechnik und Regelungstechnik*.

Für *Grundlagen der Elektronik* ergeben sich unter Ausschluss von Aufgabe *GE\_A3* die in Tabelle 23 dargestellten Kennwerte. Ein Item weist einen geringfügig zu niedrigen MNSQ Infit auf. Es wird in der Skala behalten, da sich bei Ausschluss die Kennwerte des Modells verschlechtern. Die EAP- und WLE-Reliabilität bewegt sich an der unteren Akzeptanzgrenze. Die Werte sind noch als akzeptabel einzustufen. Einen Überblick über die Skala liefert Abbildung 36.

**Tabelle 23:** Kennwerte der IRT-Skalierung der Fachwissensdimension *Grundlagen der Elektronik*.

<b>Modellkennwerte <i>Grundlagen der Elektronik</i></b>	
<b>Andersen LRT</b>	$ALRT = 19.59, p > .05$
<b>Wald-Test</b>	$p > .05$ für alle Items
<b>Martin Löff Test</b>	$MLT = 44.59, p > .05$
<b>MNSQ Infit<sup>a</sup></b>	0.77, 1.17
<b>EAP Rel</b>	0.65
<b>WLE Rel</b>	0.63
<b>Personenfähigkeit <math>\hat{\theta}^b</math></b>	$-2.46 < -0.31 < 2.43$
<b>Itemschwierigkeit <math>\hat{\phi}^b</math></b>	$-2.60 < -0.24 < 1.74$

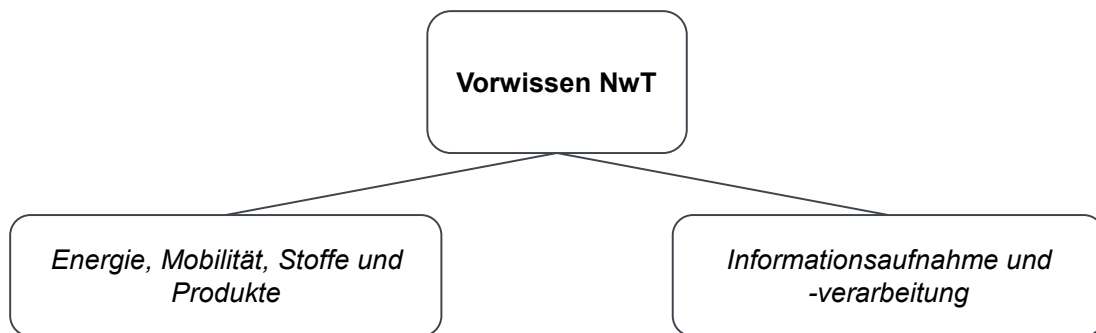
<sup>a</sup> Bericht des minimalen und maximalen Werts. <sup>b</sup> Spanne der Personenfähigkeit und Itemschwierigkeit umfasst den Minimalwert, Median und Maximalwert.

**Abbildung 36:** Person-Item Map des Kompetenzteilbereichs *Grundlagen der Elektronik*.



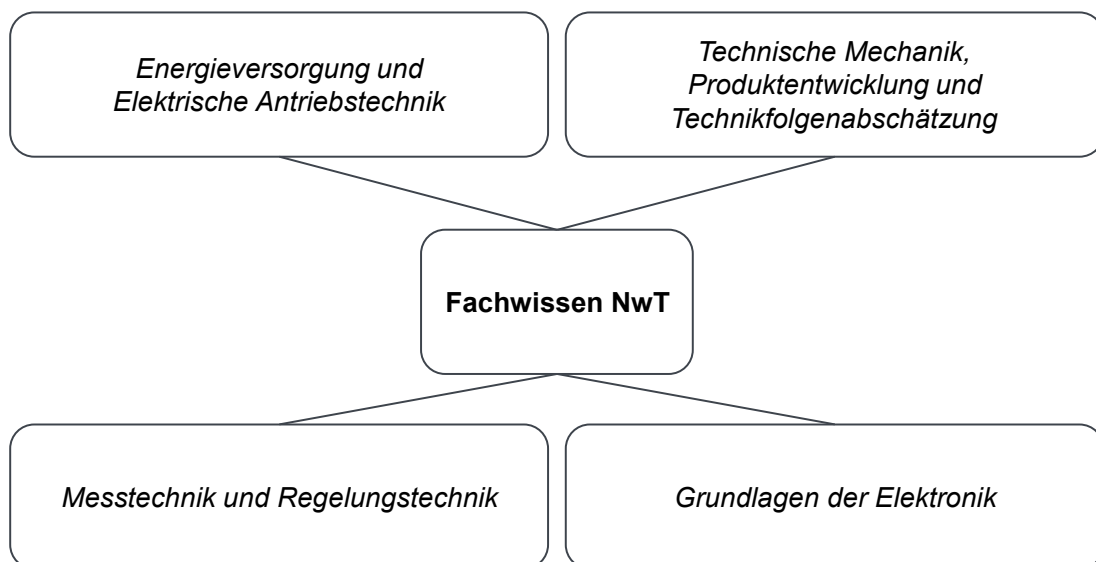
### 11.3 Zwischenfazit der IRT-Skalierung

Das Vor- und Fachwissen lassen sich jeweils als mehrdimensionale Konstrukte modellieren. Das Vorwissen NwT lässt sich anhand der beiden Dimensionen *Energie, Mobilität, Stoffe und Produkte* sowie *Informationsaufnahme und -verarbeitung* beschreiben (Abbildung 37). Für beide Vorwissensdimensionen resultieren gültige Rasch-Modelle mit zufriedenstellenden Modellparametern.



**Abbildung 37:** Vorwissen als zweidimensionales Konstrukt aus den Kompetenzbereichen *Energie, Mobilität, Stoffe und Produkte* sowie *Informationsaufnahme und -verarbeitung*.

Das Fachwissen lässt sich als vierdimensionales Konstrukt darstellen. Die einzelnen Dimensionen setzen sich aus den Kompetenzteilbereichen des Leistungsfachbildungsplans (Autorengruppe Bildungsplan NwT, 2019b) gemäß Abbildung 38 zusammen.



**Abbildung 38:** Fachwissen als vierdimensionales Konstrukt aus den Kompetenzteilbereichen *Energieversorgung und Elektrische Antriebstechnik; Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung; Messtechnik und Regelungstechnik* sowie *Grundlagen der Elektronik*.

Zusammenfassend konnten mit den generierten Daten auf Basis der Testinstrumente zur Erfassung des Vorwissens und des Fachwissens Skalen für unterschiedliche Wissensdimensionen modelliert und Personen- sowie Itemparameter geschätzt werden. Die geschätzten Personenfähigkeiten bilden den Ausgangspunkt für die nachfolgenden Analysen der Kapitel 12 und 13. Inwiefern sich die normative, über die inhaltliche Ebene der Kompetenzteilbereiche begründete Festlegung der einzelnen Dimensionen des Vor- und Fachwissens als haltbar erweist, kann in zukünftigen Studien überprüft werden.

## 12 Kompetenzniveaumodellierung

Zur Bildung der Niveauabstufungen im Rahmen der Kompetenzniveaumodellierung wird als Standard-Setting Verfahren (s. Kapitel 2.5) die Bildung von Kontrastgruppen mittels Clusteranalyse (s. Kapitel 5.7) umgesetzt. Diesbezüglich werden die R-Packages *cluster* (Maechler, Rousseeuw, Struyf, Hubert & Hornik, 2022), *factoextra* (Kassambara & Mundt, 2020), *stats* (R Core Team, 2022) und *NbClust* (Charrad et al., 2014) verwendet.

### 12.1 Standard-Setting mittels Clusteranalyse

Die Kontrastierung der Gruppen erfolgt anhand der geschätzten Personenfähigkeiten in den im vorherigen Kapitel 11.2 gebildeten Dimensionen *Energieversorgung und Elektrische Antriebstechnik; Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung* sowie *Messtechnik und Regelungstechnik*. Dazu müssen die vorliegenden Daten zunächst bereinigt werden und ggf. Fälle mit fehlenden Werten in einzelnen Skalen ausgeschlossen werden. Für die Modellierung wird der Datensatz des Fachwissens vor der Abiturprüfung verwendet. Nach Ausschluss von Personen mit fehlenden Werten verbleiben  $n_{Clus} = 113$  Personen im Datensatz für die Clusteranalyse.

Die Festlegung der Anzahl auf drei Cluster erfolgt in diesem Fall theoriegeleitet. Die Annahme von drei Clustern erscheint vertretbar, da sich die kognitive Aufgabenanforderung insgesamt an den Operatoren der EPA-Technik (Autorengruppe der Kultusministerkonferenz, 2006) orientiert und diesbezüglich von Florian et al. (2014) eine Schwierigkeitsabstufung empirisch belegt werden konnte. Die Dreiteilung ermöglicht, dass eine Niveauabstufung in Analogie zu den Anforderungsniveaus des Bildungsplans in Bezug auf praktische Implikationen interpretiert werden kann. Unter dieser Annahme lassen sich drei Cluster mit unterschiedlichen Personenfähigkeitenmittelwerten identifizieren.

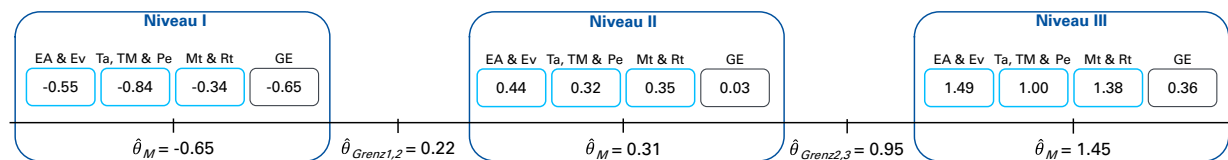
**Tabelle 24:** Verteilung der mittleren Personenfähigkeit innerhalb der Cluster.

	<b>Cluster 1 (<math>n_{c1} = 22</math>)</b>	<b>Cluster 2 (<math>n_{c2} = 51</math>)</b>	<b>Cluster 3 (<math>n_{c3} = 40</math>)</b>
<b>Min.</b>	0.95	0.22	-0.82
<b>Q1</b>	1.20	0.27	-0.74
<b>Median (MW)</b>	1.45 (1.30)	0.31 (0.30)	-0.65 (-0.63)
<b>Q3</b>	1.48	0.35	-0.53
<b>Max.</b>	1.52	0.38	-0.40

Die Stichprobe teilt sich auf zu je  $n_{C1} = 22$  in Cluster 1,  $n_{C2} = 51$  in Cluster 2 sowie  $n_{C3} = 40$  Schüler:innen in Cluster 3. Die Cluster charakterisieren sich nach mittlerer Personenfähigkeit (Mittelwert der Personenfähigkeiten der drei Dimensionen) gemäß den Kennwerten in Tabelle 24.

### 12.2 Kompetenzniveaustufen

Cluster 1 wird im Weiteren als *Niveau III*, Cluster 2 als *Niveau II* und Cluster 3 als *Niveau I* bezeichnet. Gleichzeitig ist im Kontext eines Leistungsfachs davon auszugehen, dass alle Schüler:innen weitestgehend auf einem hohen Niveau rangieren. Zur Einteilung der Kompetenzniveaustufen wird ausgehend von Niveau III jeweils das Minimum der mittleren Personenfähigkeit aus den drei Dimensionen des Modellierungsdatensatzes festgelegt. Die Mediane der geschätzten Personenfähigkeit in *Grundlagen der Elektronik* werden den Gruppen angefügt. Nach dieser Vorgehensweise ergeben sich die in Abbildung 39 dargestellten Kompetenzniveaus.



**Abbildung 39:** Kompetenzniveaumodell der geschätzten Personenfähigkeiten der Lernenden im Leistungsfach NwT. Für jedes Niveau sind die jeweiligen Medianwerte  $\hat{\theta}_M$  dargestellt. Die Bildung der Niveaus erfolgte mit den hellblau umrandeten Fachwissensdimensionen.

Erwartungskonform zeigt sich, dass sich Niveau I in *Grundlagen der Elektronik* signifikant von Niveau II und Niveau III unterscheidet. Ein signifikanter Unterschied zwischen Niveau II und Niveau III kann für diese Dimension nicht festgestellt werden, wenngleich sich der Mittelwert für die geschätzte Personenfähigkeit unterscheidet. Die Analyse der Gruppenunterschiede ist in Tabelle 25 dargestellt.

**Tabelle 25:** ANOVA der geschätzten Personenfähigkeit in der Fachwissensdimension *Grundlagen der Elektronik*.

Niveau I	Niveau II	Niveau III	ANOVA	Tukey-Kramer post-hoc
Median (MW)	Median (MW)	Median (MW)		
-0.65 (-0.91)	0.03 (-0.15)	0.36 (0.32)	$F(2, 88) = 11.72^{***}$	I-II*** I-III*** II-III

\*\*\*  $p < .001$

### **12.3 Zwischenfazit zur Kompetenzniveaumodellierung**

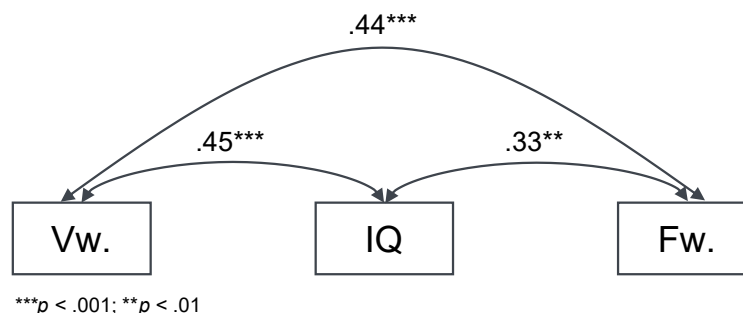
Als Zwischenfazit wird für die Kompetenzniveaumodellierung festgestellt, dass sich eine Dreiteilung der Stichprobe nach den Anforderungsniveaus der EPA-Technik (Autorengruppe der Kultusministerkonferenz, 2006) und des Bildungsplans (Autorengruppe Bildungsplan NwT, 2019b) abbilden lässt. Es können in Bezug auf die Fachwissensdimensionen *Energieversorgung und Elektrische Antriebstechnik; Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung* sowie *Messtechnik und Regelungstechnik* Kontrastgruppen anhand einer Clusteranalyse als Standard-Setting-Verfahren gebildet werden. Das Fachwissen der Schüler:innen differenziert sich in drei Niveaustufen. Niveau I unterscheidet sich zudem signifikant von Niveau II und Niveau III in Bezug auf die Dimension *Grundlagen der Elektronik*. Forschungsfrage F4 kann dahingehend beantwortet werden, als dass sich die drei unterstellten Kompetenzniveaus gemäß der Anforderungsbereiche abgrenzen lassen. Im nachstehenden Kapitel interessiert die Analyse der Personenfähigkeit der Schüler:innen in Abhängigkeit der kognitiven Voraussetzungen, um die erhaltenen Kompetenzniveaus differenzierter zu betrachten.



### 13 Zusammenhangsanalyse kognitiver Voraussetzungen mit dem Fachwissen im Leistungsfach

Die Forschungsliteratur liefert Hinweise auf einen Zusammenhang von Lernergebnissen mit kognitiven Voraussetzungen (u. a. Vorwissen und fluide Intelligenz). Im Folgenden wird dieser Zusammenhang untersucht, um eine differenziertere Beschreibung der in Kapitel 12 identifizierten Kompetenzniveaus zu ermöglichen und ggf. Hinweise für eine zielgerichtete Förderung leistungsschwächerer Schüler:innen zu erhalten. Als Merkmale werden das Vorwissen aus der Mittelstufe (s. Kapitel 11.1) sowie die fluide Intelligenz zur Beschreibung der Lernenden in den Kompetenzniveaus herangezogen. In der Zusammenhangsanalyse können nur Schüler:innen berücksichtigt werden, die an allen Messzeitpunkten teilgenommen haben. Aus diesem Grund reduziert sich die Stichprobe auf  $n_{Zsha} = 84$  Lernende.

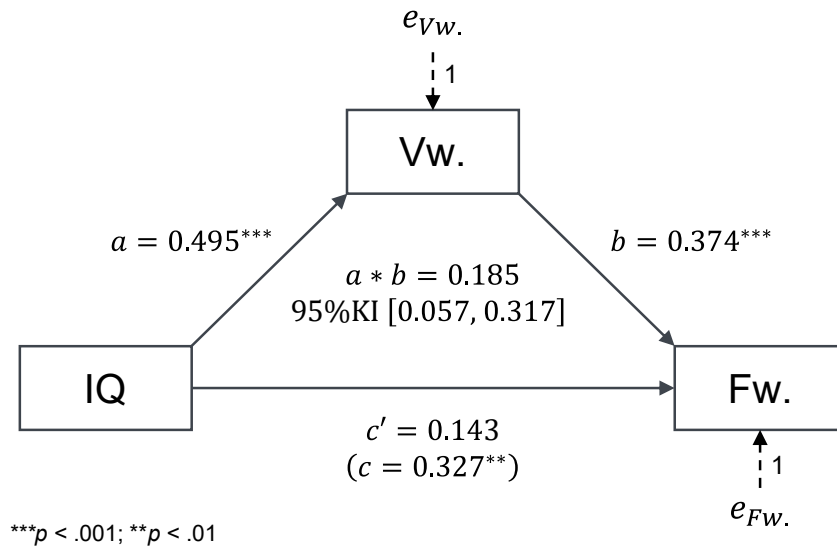
Bezüglich des Zusammenhangs der mittleren Personenfähigkeiten bei Vor- und Fachwissen mit der fluiden Intelligenz (IQ) werden nachfolgend die Hypothesen H1, H2 und H3 (Kapitel 4.1.4) überprüft. Dazu werden Mittelwerte der Personenfähigkeit aus den Kompetenzteilbereichen von Vor- und Fachwissen berechnet und einer Korrelationsanalyse unterzogen. Die Berechnung erfolgt mit *correlation* (Makowski et al., 2020). Für alle Auswertungen wird ein Signifikanzniveau von  $\alpha = .05$  zugrunde gelegt. Die Ergebnisse sind in Abbildung 40 dargestellt.



**Abbildung 40:** Korrelative Zusammenhänge zwischen fluider Intelligenz (IQ) sowie der mittleren Personenfähigkeiten des Vorwissens (Vw.) und Fachwissens (Fw.).

Es zeigen sich signifikante mittlere Korrelationen (nach Cohen, 1988) zwischen den drei Bereichen. Die Hypothesen H1, H2 und H3 lassen sich damit bestätigen und zeigen den u. a. von Hattie (2009, S. 41) dargestellten Zusammenhang der Konstrukte. Auf Basis der erhaltenen Korrelationen und bekannter Zusammenhänge der fluiden Intelligenz mit Vorwissen und Fachwissen (z. B. Nickolaus et al., 2015) wird eine Mediationsanalyse zur Ermittlung der konkreten Wirkungszusammenhänge zwischen den drei Konstrukten durchgeführt. Es soll überprüft

werden, ob die fluide Intelligenz die Entwicklung des Fachwissens vorhersagt und ob der direkte Zusammenhang durch das Vorwissen mediiert wird (H4; Kapitel 4.1.4). Die Durchführung der Mediationsanalyse erfolgt mit *PROCESS* (Hayes, 2022) in *R* (R Core Team, 2022). Die Berechnung der Konfidenzintervalle erfolgt mittels Bootstrapping in 10.000 Iterationen.



**Abbildung 41:** Mediationsanalyse des Einflusses der fluiden Intelligenz (IQ) auf Vorwissen (Vw.) und Fachwissen (Fw.). Alle Pfadkoeffizienten sind standardisiert. Das Modell klärt  $R^2 = 21\%$  der Varianz des Fachwissens auf. Die Interpretation der Effektstärken orientiert sich an Cohen (1988).

Es zeigt sich zunächst ein direkter mittlerer Effekt ( $c = 0.327$ ,  $p < .001$ ) der fluiden Intelligenz auf das Fachwissen. Mit der Aufnahme des Mediators *Vorwissen* in das Modell sagt die fluide Intelligenz das Vorwissen mit einem mittleren Effekt von  $a = 0.495$  ( $p < .001$ ) vorher. Das Vorwissen beeinflusst seinerseits das Fachwissen mit einem mittleren Effekt von  $b = 0.374$  ( $p < .001$ ). Es liegt eine vollständige Mediation ( $a * b = 0.185$ , 95 %KI [0.057, 0.317]) des Einflusses von fluider Intelligenz auf das Fachwissen über das Vorwissen vor. Der direkte Effekt nach Einbezug des Mediators ( $c' = 0.143$ ,  $p > .05$ ) ist nicht mehr signifikant. Es lässt sich damit konstatieren, dass die fluide Intelligenz keinen direkten Einfluss auf das Fachwissen ausübt, sondern eine höhere Intelligenz ein höheres Vorwissen begünstigt, mit dem wiederum ein höheres Fachwissen einhergeht. Die Ergebnisse der Mediationsanalyse sind in Abbildung 41 zusammengefasst. Insgesamt lassen sich damit  $R^2 = 21\%$  der Varianz des Fachwissens aufklären. Die Hypothese H4 gilt somit als bestätigt.

Auf Basis des Effekts der fluiden Intelligenz auf das Vorwissen und bekannter analoger Zusammenhänge aus der beruflichen Aus- und Weiterbildung (z. B.



Hedrich, 2021; Nickolaus et al., 2015, S. 20; Wyrwal, 2020) kann für den allgemeinbildenden Bereich davon ausgegangen werden, dass sich im Fach NwT in der Mittelstufe eine technische Grundfähigkeit entwickelt hat, die einen Einfluss auf das Fachwissen im Leistungsfach ausübt. Gleichzeitig deutet die Varianzaufklärung von  $R^2 = 21\%$  einerseits darauf hin, dass neben dem Vorwissen möglicherweise weitere bedeutsame Einflussfaktoren existieren (z. B. affektive und motivationale Lernendenmerkmale) die eine positive Fachwissensentwicklung begünstigen. Andererseits können die geringeren Effekte und die insgesamt kleinere Varianzaufklärung im Vergleich zu den Modellen von z. B. Nickolaus et al. (2015) möglicherweise auch bedeuten, dass die Anknüpfung der Mittel- an die Kursstufe im Sinne eines Spiralcurriculums noch optimiert werden könnte.

### 13.1 Korrelative Zusammenhänge der Kompetenzteilbereiche

Für eine differenziertere Beschreibung der Kompetenzniveaus und einen vertiefteren Einblick in die Zusammenhänge der Dimensionen des Vorwissens mit den Dimensionen des Fachwissens werden die Korrelationen des Vor- und Fachwissens sowie der fluiden Intelligenz (Tabelle 26) anhand des vorliegenden Gesamtdatensatzes untersucht.

**Tabelle 26:** Korrelative Zusammenhänge der kognitiven Auszeichnungsmerkmale (Personenfähigkeiten Vor- und Fachwissen sowie fluide Intelligenz).

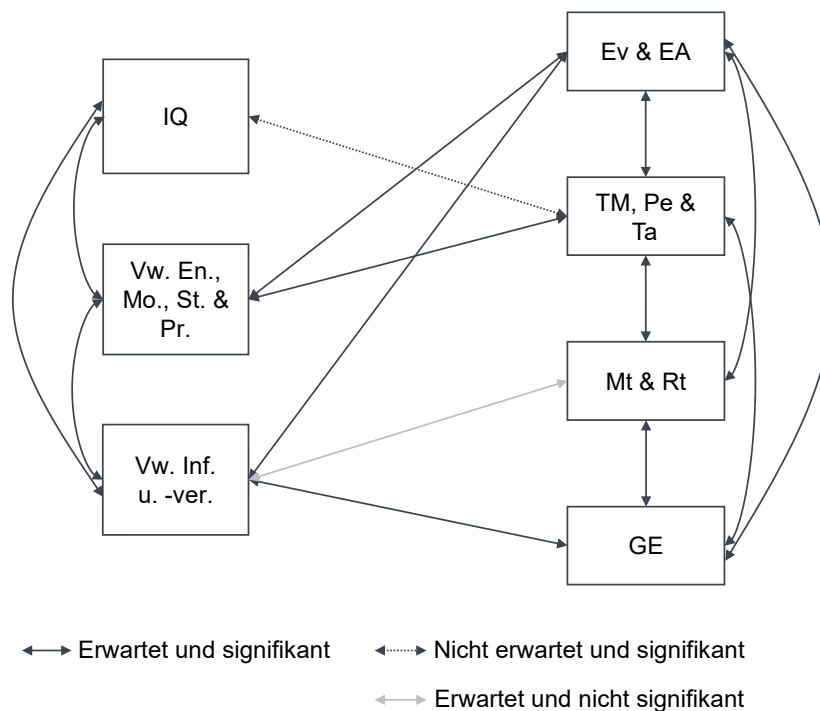
	IQ	Vw. Inf. u. -ver.	Vw. En., Mo., St. & Pr..	Ev & EA	TM, Pe & Ta	Mt & Rt
<b>Vw. Inf. u. -ver.</b>	0.34**	-				
<b>Vw. En., Mo., St. &amp; Pr.</b>	0.43***	0.46***	-			
<b>Ev &amp; EA</b>	0.23	0.53***	0.32**	-		
<b>TM, Pe &amp; Ta</b>	0.33**	0.24	0.40***	0.56***	-	
<b>Mt &amp; Rt</b>	0.09	0.22	0.12	0.38***	0.34**	-
<b>GE</b>	0.25	0.44***	0.23	0.56***	0.36**	0.28*

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ; Interpretation der Stärke des Zusammenhangs nach Cohen (1988). Vw. En., Mo., St. & Pr = Vorwissen *Energie, Mobilität, Stoffe & Produkte*; Vw. Inf. u. -ver. = Vorwissen *Informationsaufnahme und -verarbeitung*; Ev & EA = *Energieversorgung & Elektrische Antriebstechnik*; TM, Pe & Ta = *Technische Mechanik, Produktentwicklung & Technikfolgenabschätzung*; Mt & Rt = *Messtechnik & Regelungstechnik*; GE = *Grundlagen der Elektronik*.

Es kann erwartungskonform eine Korrelation der fluiden Intelligenz mit den beiden Kompetenzbereichen des Vorwissens festgestellt werden. Ebenfalls erwartungskonform sind fehlende Korrelationen (mit Ausnahme von *TM, Pe & Ta*) der

fluiden Intelligenz, die mit den Wissensdimensionen des Leistungsfachs über das Vorwissen in einem indirekten Zusammenhang steht (s. Mediationsanalyse Abbildung 41). Über den korrelativen Zusammenhang von *TM, Pe & Ta* und der fluiden Intelligenz kann nur gemutmaßt werden. Ein Erklärungsansatz könnte die kognitive Anforderung der Testskala darstellen, die möglicherweise u. a. die „Fähigkeit des Menschen Zusammenhänge zu erkennen und diese in neuartigen Situationen anwenden zu können“ (Weiß, 2019, S. 31) erfordert, die über den CFT 20-R repräsentiert wird. Zum Beispiel Geißel, Nickolaus, Stefanica, Härtig und Neumann (2013, S. 57) stellen diesbezüglich ebenfalls fest, dass der IQ in komplexeren Anforderungskontexten einen größeren Zusammenhang mit dem Fachwissen aufweist. Eine eingehendere Betrachtung dieses Sachverhalts kann Gegenstand zukünftiger Untersuchungen zur Kompetenzstruktur sein.

Für die Korrelation der Vorwissensdimensionen mit den Fachwissensdimensionen werden unterschiedliche Korrelationen erwartet, die auf der inhaltlichen Ebene der Vorwissens- und Fachwissensdimensionen begründet ist. Eine Übersicht über erwartete und tatsächliche Korrelationen gibt Abbildung 42.



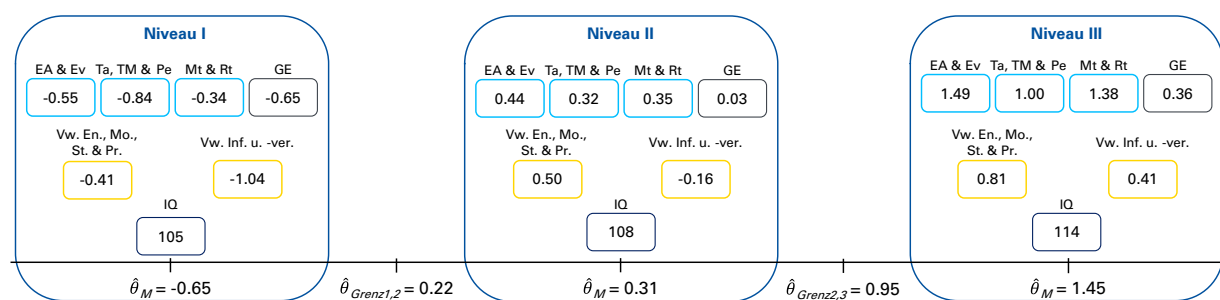
**Abbildung 42:** Darstellung der erwarteten und tatsächlich signifikanten Korrelationen zwischen den Vor- und Fachwissensdimensionen auf inhaltlicher Basis und der kognitiven Grundvoraussetzung. Die Korrelationswerte werden aus Übersichtsgründen nicht in die Abbildung integriert und sind in Tabelle 26 dargestellt.

Die Dimensionen des Vorwissens korrelieren überwiegend erwartungskonform mit den entsprechenden Dimensionen des Fachwissens im Leistungsfach und

deuten auf die inhaltliche Anknüpfung des Leistungsfachs an die Mittelstufe hin. Die Korrelation zwischen *Vorwissen Informationsaufnahme und -verarbeitung* und *Messtechnik und Regelungstechnik* ist als einzige erwartete Korrelation nicht signifikant und zeigt damit eine fehlende Anknüpfung an das Vorwissen auf. Die Tatsache, dass die Vorwissensdimensionen auf Basis einer inhaltlichen Begründung erwartungskonform mit den Fachwissensdimensionen korrelieren, untermauert neben der statistischen Modellierung zusätzlich die Annahme einer Mehrdimensionalität von Fach- und Vorwissen.

### 13.2 Beschreibung der Kompetenzniveaus unter Einbezug kognitiver Merkmale

Im Folgenden soll untersucht werden, ob sich die Niveaustufen auch in den Konstrukten *fluide Intelligenz* und *Vorwissen* wiederfinden. Falls sich hier ein signifikanter Gruppenunterschied bestätigen lässt, kann bereits anhand dieser beiden Konstrukte bei der Eingangsuntersuchung zu den Startvoraussetzungen abgeschätzt werden, welchem Niveau die Lernenden zuzuordnen sind und welche inhaltspezifische Förderung in Bezug auf die Zusammenhänge von Vor- und Fachwissen eine Wirkung entfalten kann. Gleichzeitig ist eine umfängliche Beschreibung der Kompetenzniveaus in Bezug auf die kognitiven Lernendenmerkmale (*fluide Intelligenz*, *Vorwissen* und *Fachwissen*) möglich. Die Kompetenzniveaus unter Einbezug der kognitiven Merkmale *fluide Intelligenz* und der Personenfähigkeiten in den Kompetenzdimensionen des Vorwissens sind in Abbildung 43 dargestellt. Insgesamt zeigt sich auf Basis der vorherigen Untersuchungen ein erwartungskonformes Bild ansteigender Personenfähigkeiten in den Facetten des Vorwissens sowie bei der fluiden Intelligenz über die Niveaus hinweg. Zur statistischen Betrachtung der Unterschiede in den Ausprägungen der Auszeichnungsmerkmale in den einzelnen Niveaustufen werden ANOVAs berechnet.



**Abbildung 43:** Beschreibung der Kompetenzniveaus durch kognitive Auszeichnungsmerkmale *Personenfähigkeit Vorwissen* und *fluide Intelligenz (IQ)*. Abgetragen sind die jeweiligen Medianwerte. Die Bildung der Niveaus erfolgte mit den hellblau umrandeten Fachwissensdimensionen.

Der Tukey-Kramer post-hoc-Test der Varianzanalyse zu den beiden Facetten des Vorwissens belegt einen signifikanten Unterschied zwischen Niveau I und Niveau III. Zwischen den Niveaus II und III lassen sich keine statistisch signifikanten Gruppenunterschiede bezüglich des Vorwissens belegen, wenngleich sich die Mediane der Personenfähigkeit unterscheiden. Für die fluide Intelligenz kann ebenfalls ein signifikanter Unterschied zwischen den Niveaus I und III gefunden werden; jedoch nicht zwischen Niveau I und II sowie Niveau II und III. Alle diesbezüglichen Ergebnisse sind in Tabelle 27 zusammengefasst.

**Tabelle 27:** ANOVAs der kognitiven Auszeichnungsmerkmale.

	Niveau I	Niveau II	Niveau III	ANOVA	Tukey-Kramer post-hoc
	Median (MW)	Median (MW)	Median (MW)		
<b>Vw. En., Mo., St. &amp; Pr.</b>	-0.41 (-0.22)	0.50 (0.30)	0.81 (0.80)	$F(2, 103) = 5.67^{**}$	I-II I-III** II-III
<b>Vw. Inf. u. -ver.</b>	-1.04 (-0.71)	-0.16 (-0.16)	0.41 (0.44)	$F(2, 102) = 7.18^{**}$	I-II I-III*** II-III
<b>IQ</b>	105.00 (104.97)	108.00 (109.22)	114.00 (114.53)	$F(2, 103) = 3.49^*$	I-II I-III* II-III

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

Die signifikanten Unterschiede zwischen Niveau I und III zeigen, dass Vorwissen und fluide Intelligenz sich zur Beschreibung der Niveaustufen des Fachwissens eignen und eine zumindest binäre Zuordnung der Lernenden zu Niveau I oder Niveau III ermöglichen. Insgesamt ist ein Anstieg der Personenfähigkeit des Vorwissens und der fluiden Intelligenz über die Niveaus erkennbar. Gleichzeitig haben sie durch die Varianzaufklärung von  $R^2 = 21\%$  beschränkt Vorhersagekraft für das spätere Fachwissen. Dies wird nochmals sichtbar, da sich die Niveaus I und II sowie II und III jeweils nicht signifikant in Bezug auf Vorwissen und fluide Intelligenz unterscheiden. Für die zukünftige Forschung resultiert daher das Desiderat der Beschreibung der Niveaus unter Einbezug weiterer Lernendenmerkmale, um die Varianz der Personenfähigkeiten der Fachwissensdimensionen weiter aufzuklären und damit den Lernerfolg genauer vorherzusagen.

### 13.3 Zwischenfazit zu den Zusammenhangsanalysen kognitiver Voraussetzungen mit dem Fachwissen im Leistungsfach

In Bezug zur Forschungsfrage F5 lässt sich feststellen: Ein Zusammenhang zwischen fluider Intelligenz und Vorwissen mit dem erreichten Fachwissen lässt sich theoriekonform neben dem beruflich-technischen Aus- und Weiterbildungsbe- reich (z. B. Hedrich, 2021; Nickolaus et al., 2015; Wyrwal, 2020) auch für den all- gemeinbildenden technischen Bereich und darin speziell für das Leistungsfach NwT belegen. Die Hypothesen H1 bis H3 können durch signifikant positive Zu- sammenhänge der Konstrukte fluider Intelligenz, Vorwissen und Fachwissen be- stätigt werden. Über die Korrelationen hinaus finden sich in der einschlägigen Literatur Hinweise auf Mediationseffekte des Vorwissens bei Untersuchungen des Zusammenhangs von fluider Intelligenz, Vorwissen und Fachwissen. Diese theoretische Herleitung wird anhand der Bestätigung von Hypothese H4 eben- falls für den allgemeinbildenden technischen Bereich nachgewiesen. Das berech- nete Modell klärt  $R^2 = 21\%$  der Varianz des Fachwissens durch die vollständige Mediation der fluiden Intelligenz über das Vorwissen auf. Es zeigen sich mittlere direkte Einflüsse der fluiden Intelligenz auf das Vorwissen und wiederum mittlere Einflüsse des Vorwissens auf das Fachwissen (s. Abbildung 41). Im Vergleich zu bisherigen Erkenntnissen aus dem beruflichen Bereich (z. B. Nickolaus et al., 2015) fällt die Varianzaufklärung des Fachwissens jedoch geringer aus. Diese Er- kenntnis lässt sich in zweierlei Hinsicht interpretieren. Erstens deutet der Befund darauf hin, dass neben fluider Intelligenz und Vorwissen weitere bedeutsame Einflussfaktoren auf das Fachwissen existieren. Zweitens erscheint es plausibel, dass zum jetzigen Stand im Schulversuch Optimierungspotenzial bei der An- knüpfung der Inhalte der Mittelstufe an die Kursstufe im Sinne eines Spiralcurri- culums besteht. Inwiefern sich die Varianzaufklärung durch den Einbezug weite- rer Einflussfaktoren und durch die weitere Entwicklung des Leistungsfachs stei- gert, stellt deshalb auch ein Desiderat für zukünftige Untersuchungen dar. For- schungsfrage F5 wird durch die theoriekonform nachgewiesenen Zusammen- hänge und die Bestätigung der Hypothesen H1 bis H4 als beantwortet angese- hen. Die kognitiven Merkmale *fluide Intelligenz* und *Vorwissen* eignen sich für die Einteilung von Lernenden in Kompetenzniveaus auf Basis ihrer Eingangsvor- aussetzungen, wenngleich sich die Merkmale lediglich zwischen Niveau I und III signifikant unterscheiden. Diese Erkenntnis unterstreicht das Desiderat nach der Suche weiterer Merkmale zur Beschreibung der Kompetenzniveaus. Die Eignung der kognitiven Auszeichnungsmerkmale *fluide Intelligenz* und *Vorwissen* zur bi- nären Zuordnung der Lernenden zu den Kompetenzniveaus beantwortet dahin- gehend Forschungsfrage F6.



## 14 Resümee

Es folgt eine Zusammenfassung der Erkenntnisse, begleitet durch eine diskursive Einbettung der Ergebnisse in den Forschungsstand. Unter Einbezug des theoretischen Hintergrunds werden Implikationen für die schulische Praxis und die Bildungsforschung im allgemeinbildenden naturwissenschaftlich-technischen Kontext unter kritischer Auseinandersetzung mit den Limitationen der Studien abgeleitet. Die Arbeit schließt mit einem Ausblick auf weitere Forschungsdesiderate.

### 14.1 Zusammenfassung

Ausgehend vom Desiderat eines Beschreibungswissens zu den inhaltsbezogenen Kompetenzen von Schüler:innen im Leistungsfach NwT geht die vorliegende Arbeit sechs Forschungsfragen (F1 bis F6) nach. Forschungsfrage F1 adressiert wie die valide und reliable Messung inhaltsbezogener Kompetenzen im Leistungsfach NwT erfolgen kann. Dazu interessieren verschiedene Aspekte der Rahmenbedingungen des Leistungsfachs, die Einschätzung der Lehrpersonen zu den entwickelten Testaufgaben bis hin zur Pilotierung der Testinstrumente mit Schüler:innen. Auf Basis der entwickelten Testinstrumente ist die Beschreibung der Eingangsvoraussetzungen der Lernenden bei Einmündung in die Kursstufe (F2) von Interesse und über welches Fachwissen als Ausprägung der inhaltsbezogenen Kompetenzen sie am Ende der gymnasialen Kursstufe verfügen (F3). Basierend auf der Kompetenzmessung widmet sich die Forschungsfrage F4 der Feststellung unterschiedlicher Kompetenzniveaus. Im Rahmen der Forschungsfrage F5 werden Zusammenhänge der fluiden Intelligenz und des Vorwissens mit dem Fachwissen betrachtet, um in Forschungsfrage F6 abschließend deren Eignung zur Einteilung der Lernenden in Kompetenzniveaus anhand ihrer Eingangsvoraussetzungen zu überprüfen. Die Beantwortung der Forschungsfragen erfolgt durch mehrere Studien, die im Zusammenhang mit dem Forschungsprojekt *NwT-Kursstufe* stehen.

Zu **F1**: Für die Entwicklung geeigneter Testinstrumente wird zunächst die Angebotsebene im Schulversuch *NwT-Kursstufe* betrachtet. Inhaltliche und kontextuelle Aspekte des Unterrichts werden durch Lehrpersoneninterviews eruiert. Die Interviewstudie ergibt, dass sich die Kursgrößen und damit das Betreuungsverhältnis an den Schulen (zw. 1:2 und 1:15) sowie die sächliche Ausstattung v. a. in Bezug auf das Budget (zw. 1.000 € und 8.000 €) heterogen darstellen. Bezüglich der sächlichen Ausstattung mit Geräten zur computergestützten Fertigung zeigen sich die Befragten größtenteils zufrieden. Die zumindest während der Anfangsphase meist im Team unterrichtenden Lehrpersonen wurden über

kursstufenspezifische Qualifikationsmaßnahmen weitergebildet und nehmen aus überwiegend intrinsischen Motiven am Schulversuch teil. Die befragten Lehrpersonen deuten an, dass vom Leistungsfach Implikationen für die Mittelstufe ausgehen und dass dort eine stärkere Verankerung des für die Kursstufe benötigten Vorwissens erfolgen muss. Die Kompetenzförderung der Schüler:innen erfolgt anhand der im Bildungsplan formulierten inhalts- und prozessbezogenen Kompetenzen und erfüllt laut den interviewten Lehrpersonen gleichzeitig die Anforderungen der EPA-Technik (Autorengruppe der Kultusministerkonferenz, 2006) in Bezug auf die Förderung von Sach- und Methodenkompetenz in einem fachspezifischen Kontext. Als geeignete Prüfungsformate im Leistungsfach werden Klausuren zur Erfassung inhaltsbezogener Kompetenzfacetten und Projektdokumentationen zur Prüfung prozessbezogener Kompetenzentwicklung eingesetzt. Die aus den Lehrpersoneninterviews gewonnen Einblicke ermöglichen eine zielgerichtete Entwicklung von Testinstrumenten für die Messung inhaltsbezogener Kompetenzen bei Schüler:innen im Leistungsfach. Einerseits können durch die konkrete Bezugnahme zu Kompetenzteilbereichen aus der Mittelstufe passgenaue Items für die Erfassung eines anknüpfungsfähigen Vorwissens entwickelt und andererseits den Lehrpersonen alle Aufgaben (Vorwissen und Fachwissen im Leistungsfach) als Bildungsexpert:innen zur Bewertung der inhaltlichen Validität in Bezug auf Praxis- und curriculare Relevanz sowie zur Einschätzung der Aufgabenkomplexität vorgelegt werden. Anhand dieser Einschätzung erfolgt die Zusammenstellung der Testhefte, deren Reliabilität in einer Pilotierungsstudie überprüft wird. Die Pilotierungsstudie ermöglicht mittels Klassischer Testtheorie die Bildung reliabler und valider Skalen für das kursstufenrelevante Vorwissen (MKJS, 2016) und die neun Kompetenzteilbereiche des Leistungsfachbildungsplans (Autorengruppe Bildungsplan NwT, 2019b).

Zu **F2** und **F3**: Die Erhebungen erfolgen in zwei Leistungsfachdurchgängen L1 und L2. Im Durchgang L1 sind fünf der ursprünglich sechs Pilotschulen vertreten. Im Durchgang L2 nehmen drei der ursprünglich sechs Pilotschulen und sieben neu hinzugekommene Modellschulen mit einem Leistungsfach am Schulversuch teil. Zu Beginn jedes Leistungsfachdurchgangs wird das Vorwissen der Schüler:innen erfasst. Die kombinierte Betrachtung der beiden Leistungsfachdurchgänge deutet zunächst heterogene Startvoraussetzungen bei den Schüler:innen an. Es zeigt sich ein signifikant höheres Vorwissen bei Lernenden an Pilotschulen im Vergleich zu den Schüler:innen an Modellschulen. An zwei



Messzeitpunkten<sup>16</sup> erfolgt die Erfassung der inhaltsbezogenen Kompetenzen in sieben abiturrelevanten Kompetenzteilbereichen<sup>17</sup> und schließt mit der Befragung zu zwei weiteren Kompetenzteilbereichen<sup>18</sup> nach dem Abitur ab. Trotz des unterschiedlichen Vorwissens ist kein signifikanter Unterschied des mittleren Fachwissens der Lernenden am Messzeitpunkt nach eineinhalb Schuljahren (jeweils kurz vor der Abiturprüfung) festzustellen. Durch die Kombination der Stichproben aus L1 und L2 ergibt sich darüber hinaus die Möglichkeit, das Analysepotenzial der Testskalen hinsichtlich der Schätzung von Personenfähigkeiten mittels IRT-Skalierung auszubauen.

Zu **F4**: Zur Kompetenzniveaumodellierung werden die Personenfähigkeiten der Schüler:innen mittels IRT-Skalierung geschätzt. Die Modellierung der einzelnen Kompetenzdimensionen orientiert sich primär an der Zuordnung der Kompetenzteilbereiche zu den jeweiligen inhaltsbezogenen Kompetenzbereichen des Leistungsfachbildungsplans (Autorengruppe Bildungsplan NwT, 2019b). Einige Kompetenzteilbereiche weisen jedoch auch inhaltsbezogene Quervernetzungen außerhalb des ihnen zugeordneten Kompetenzbereichs auf. Aus diesem Grund werden über eine Korrelationsanalyse zunächst Zusammenhänge zwischen den Lösungsquoten der einzelnen Kompetenzteilbereiche ermittelt, um die normative Annahme zusammenhängender Skalen zu überprüfen<sup>19</sup>. Ausgehend davon lassen sich Rasch-Modelle für die inhaltlich und auf Basis von korrelierenden Lösungsquoten der Kompetenzteilbereiche unterstellten Fachwissensdimensionen berechnen: *Energieversorgung und Elektrische Antriebstechnik; Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung; Messtechnik und Regelungstechnik* sowie *Grundlagen der Elektronik*. Insgesamt wird damit von einem vierdimensionalen Fachwissenskonstrukt ausgegangen. Auf Basis der drei Dimensionen *Energieversorgung und Elektrische Antriebstechnik; Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung* sowie *Messtechnik und Regelungstechnik* erfolgt mittels Clusteranalyse als Standard-Setting-Verfahren die Bildung dreier Kontrastgruppen. *Grundlagen der Elektronik* dient als Kontrollvariable. Die Kontrastierung lehnt sich an die durch Operatoren implizierten kognitiven Leistungsniveaus des Bil-

---

<sup>16</sup> Die erste Messung erfolgt nach einem dreiviertel Schuljahr. Die zweite Messung erfolgt unmittelbar vor dem Abitur nach ca. eineinhalb Schuljahren.

<sup>17</sup> Gegenstand der Erhebung sind die Teilbereiche: *Technikfolgenabschätzung, Energieversorgung, Elektrische Antriebstechnik, Technische Mechanik, Produktentwicklung, Aufnahme und Verarbeitung von Signalen in der Messtechnik* (kurz: *Messtechnik*) sowie *Regelungstechnik*.

<sup>18</sup> *Grundlagen der Elektronik* und *Datenkommunikation*

<sup>19</sup> *Datenkommunikation* wird aufgrund der nicht belastbaren Datengrundlage ausgeschlossen.

dungsplans (ebd.) bzw. der EPA-Technik (Autorengruppe der Kultusministerkonferenz, 2006) an und ergibt drei trennscharfe Cluster. Die Kontrastgruppen erhalten die Bezeichnungen *Niveau I*, *Niveau II* sowie *Niveau III*. Der mittlere Wert der Niveaustufen resultiert aus dem Mittelwert der Personenfähigkeiten aller Fachwissensdimensionen. Die Grenzen der Niveaus werden anhand des Minimalwerts der mittleren Personenfähigkeit der einzelnen Kontrastgruppen festgelegt.

Zu **F5**: Forschungsfrage F5 widmet sich den Zusammenhängen kognitiver Voraussetzungen (Vorwissen und fluide Intelligenz) mit dem Fachwissen der Schüler:innen. Dazu wird analog zu den Kompetenzbereichen des Fachwissens die Personenfähigkeit der Lernenden mittels IRT-Skalierung geschätzt. Es resultieren für die Beschreibung des Vorwissens zwei eindimensionale Skalen, die sich an den Kompetenzbereichen *Energie, Mobilität, Stoffe & Produkte* sowie *Informationsaufnahme und -verarbeitung* des Bildungsplans der Mittelstufe (MKJS, 2016) orientieren. Für die Untersuchung der Einflüsse kognitiver Voraussetzungen auf das Fachwissen wird ein Mittelwert der Personenfähigkeiten aus den Dimensionen des Vor- und Fachwissens gebildet. Es zeigt sich in Einklang mit dem theoretischen Hintergrund, dass signifikant positive Korrelationen zwischen Vor- und Fachwissen sowie der fluiden Intelligenz bestehen. Des Weiteren wird theoriekonform eine vollständige Mediation des Einflusses fluider Intelligenz über das Vorwissen auf das Fachwissen nachgewiesen. Mit dem Mediationsmodell können  $R^2 = 21\%$  der Varianz des Fachwissens aufgeklärt werden. In diesem Kontext interessiert auch der Zusammenhang von Vorwissens- und Fachwissensdimensionen. Die Korrelationsanalyse zeigt einen Zusammenhang zwischen der Vorwissensdimension *Energie, Mobilität, Stoffe und Produkte* mit den Fachwissensdimensionen *Energieversorgung und Elektrische Antriebstechnik* sowie *Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung*. Die Vorwissensdimension *Informationsaufnahme und -verarbeitung* korreliert mit *Energieversorgung und Elektrische Antriebstechnik* sowie *Grundlagen der Elektronik*. Eine erwartete Korrelation mit *Messtechnik und Regelungstechnik* ist nicht signifikant. Korrelationen der fluiden Intelligenz mit den Fachwissensdimensionen sind aufgrund der Mediation nicht zu erwarten. Trotzdem zeigt sich ein positiver Zusammenhang zwischen der fluiden Intelligenz und der Fachwissensdimension *Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung*.

Zu **F6**: Auf Basis der belegten Zusammenhänge der kognitiven Merkmale mit dem Fachwissen wird mittels Varianzanalysen aufgeklärt, ob sich die drei Niveaustufen ebenfalls bei den Konstrukten *fluide Intelligenz* und *Vorwissen* wiederfinden. Über eine möglichst frühe Zuordnung der Lernenden zu einem Kompetenzniveau kann eine zielgerichtete inhaltsorientierte Förderung von Lernenden im Leistungsfach NwT erfolgen. Es können signifikante Unterschiede der Vorwissensdimensionen und der fluiden Intelligenz zwischen den Niveaus I und III festgestellt werden. Dadurch wird zwar keine tertiäre, jedoch eine binäre Einordnung des Niveaus anhand des Vorwissens und der fluiden Intelligenz möglich.

## **14.2 Diskussion und Implikationen**

Ausgehend vom Forschungsstand zeigt sich durch die vorliegende Untersuchung, dass das Leistungsfach NwT als ein auf Standards basierendes, interdisziplinär ausgerichtetes und allgemeinbildendes naturwissenschaftlich-technisches Fach in der Lage ist, zielgerichtet (inhaltsbezogene) Kompetenzen zu fördern. Es leistet damit einen Beitrag in Bezug auf eine technische Mündigkeit der Lernenden sowie zur Förderung der wissenschaftspropädeutischen Anknüpfungsfähigkeit des allgemeinbildenden Gymnasiums an Technikwissenschaften. Dieser Beitrag wird u. a. durch die Revision des Bildungsplans 2016 sowie der Einführung von Basis- und Leistungsfach im Vergleich zum im Forschungsstand dargestellten Status des Fachs gestärkt. Wurden in den Abhandlung von Mokhonko et al. (2014) sowie Zinn und Latzel (2017) noch Dinge auf der Angebots-ebene wie bspw. fehlende Geräte zur computergestützten Fertigung (z. B. 3D-Drucker) oder keine eigenen Fachräume von den Lehrpersonen beklagt, zeigen sich zumindest die im Schulversuch vertretenen Lehrpersonen nach dem aktuellen Stand überwiegend zufrieden mit der räumlichen und sächlichen Ausstattung ihrer Schule in Bezug auf die Umsetzung des Leistungsfachs NwT. Gleichmaßen sollten Schulen, deren Ausstattung sich nicht auf diesem Stand befindet, entsprechend nachgerüstet werden, um flächendeckend homogene Bedingungen zu schaffen. Bezüglich der Qualifizierung der Lehrpersonen münden einerseits stetig Absolvent:innen mit einem grundständigen NwT-Studium in den Schuldienst ein. Auf der anderen Seite werden die praktizierenden Lehrer:innen gezielt für das Basis- und Leistungsfach in der Kursstufe weitergebildet. Die vorliegende Untersuchung leistet mit der aktuellen Bestandsaufnahme der Rahmenbedingungen des Fachs NwT in Mittel- und Kursstufe einen Beitrag zur Erweiterung des Wissensstands.

In Bezug auf die Förderung der inhaltsbezogenen Kompetenzen bei Schüler:innen im Leistungsfach erfolgt in der vorliegenden Arbeit eine Messung der Lernergebnisse in Form des Fachwissens zu den im aktuellen Bildungsplan (Autorengruppe Bildungsplan NwT, 2019b) formulierten Kompetenzbereichen und des Vorwissens aus der Mittelstufe. Die Ergebnisse belegen, dass sich bei den Lernenden an allen Schulen inhaltsbezogene Kompetenzen über den Verlauf des Leistungsfachs ausprägen. Die Lösungsquoten der Fachwissenstests unterscheiden sich im Mittel nicht signifikant zwischen Lernenden an Pilot- und Modellschulen. Bezüglich des Kompetenzteilbereichs *Datenkommunikation* kann für die Bildungspraxis der Vorschlag einer Anpassung der unterrichtlichen Umsetzung abgeleitet werden, da zur Ausprägung des Fachwissens in diesem Kompetenzteilbereich aufgrund der durch zeitliche Faktoren beeinträchtigten Behandlungsintensität keine Aussage getroffen werden kann. Durch die Kompetenzniveaumodellierung zeigen sich basierend auf unterschiedlichen kognitiven Startvoraussetzungen (fluide Intelligenz und Vorwissen) aber auch signifikante Unterschiede bei der Ausprägung des Fachwissens der Lernenden im Leistungsfach. Die zugrundeliegende Annahme eines Einflusses von fluider Intelligenz und Vorwissen stehen im Einklang mit bestehenden Modellen aus der beruflich-technischen Aus- und Weiterbildung und erweitern den Wissensstand um Ergebnisse aus einem allgemeinbildenden technischen Blickwinkel. Gleichzeitig wird durch diese Einflussfaktoren auf das Fachwissen eine geringere Varianzaufklärung erreicht, als durch die Theorie zu erwarten wäre. Diese Erkenntnis deutet möglicherweise auf eine noch zu optimierende Anknüpfung der Mittel- an die Kursstufe hin und wird auch von den Lehrpersonen in den Interviews zu den Rahmenbedingungen geäußert. Für die Bildungspraxis kann daraus der Vorschlag einer strukturellen Anpassung der Mittelstufe hinsichtlich der gezielten Förderung spiralcurricular anknüpfungsfähiger Kompetenzen abgeleitet werden. Hierzu bietet das vorliegende Testinstrument zur Erfassung des Vorwissens der Mittelstufe die Möglichkeit über Eingangsbefragungen bei zukünftigen Leistungsfachdurchgängen eine Bestandsaufnahme bezüglich der Heterogenität in der Zusammensetzung der Leistungskurse oder eine Oberstufenberatung zum Ende der Sekundarstufe I vorzunehmen. Es wäre zudem denkbar das Testinstrument im Rahmen einer landesweiten Bestandsaufnahme des Fachwissens der Schüler:innen im Fach NwT am Ende der Mittelstufe einzusetzen und die weitere Entwicklung des Fachs formativ zu begleiten.

Auf Basis der vorliegenden Befunde lässt sich für die naturwissenschaftlich-technische Allgemeinbildung folgendes Fazit ziehen: Das Schulfach NwT ist ausgehend von seiner Anlage als Profulfach der Mittelstufe über das Leistungsfach in

der Lage, technische Kompetenzen bei Lernenden zu fördern. Basierend auf einem anknüpfungsfähigen Vorwissen aus der Mittelstufe erlangen die Lernenden nachweisbar ein interdisziplinäres und an Technikwissenschaften orientiertes Fachwissen. Gleichzeitig konnte bisher kaum auf Erfahrungswerte bezüglich der Rahmenbedingungen und der Lernzielerreichung interdisziplinärer technischer Fächer am allgemeinbildenden Gymnasium zurückgegriffen werden. Aus diesem Grund bildet das im Rahmen dieser Arbeit generierte Beschreibungswissen zum einen die Grundlage für eine evidenzorientierte Weiterentwicklung des Fachs NwT und zum anderen einen Orientierungsrahmen für ähnliche Fächerangebote.

### **14.3 Limitationen**

Limitationen sind bezüglich der Operationalisierung inhaltsbezogener Kompetenzen der Schüler:innen über das Fachwissen, der Stichprobengröße in Bezug auf statistische Methoden und der Organisation der Messzeitpunkte zur Erfassung des Fachwissens zu nennen. Die Operationalisierung der inhaltsbezogenen Kompetenzen über das Fachwissen erscheint ausgehend vom grundlegenden Kompetenzbegriff sinnvoll. Ein möglicher Einfluss der prozessbezogenen Kompetenzen bleibt durch die Messung jedoch unberücksichtigt. Aus forschungspraktischer Sicht ist diese Reduktion hinzunehmen, da sich mit einem papierbasierten Testinstrument unter Berücksichtigung der Eindimensionalität von Testskalen zum einen nicht zwei unterschiedliche Konstrukte erfassen lassen und zum anderen eine Erfassung prozessbezogener Kompetenzen mittels papierbasierten Testinstrumenten nur schwer möglich ist. Es ist anzunehmen, dass die über das Fachwissen als Repräsentation der inhaltsbezogenen Kompetenzen vorgenommene Charakterisierung der Schüler:innen im Leistungsfach belastbar ist. Gleichwohl wurden aus wirtschaftlichen Gründen keine weiteren Kontrollvariablen wie z. B. mathematische Kompetenzen, Interesse oder Motivation erhoben. Trotz Vollerhebung stützen sich die statistischen Analysen auf eine insgesamt kleine Stichprobe. Dies schließt zum einen die Modellierung komplexerer Zusammenhänge aus und bedeutet für die vorgenommenen Modellierungen, dass bei eventuellem Ausbau der Stichprobe Modelle nochmals hinsichtlich ihrer Beständigkeit zu untersuchen sind. Als abschließende Limitation wird die praktische Umsetzung der Tests angeführt. Aufgrund der maximalen Verfügbarkeit von 90 min Testzeit (eine Doppelstunde) pro Testzeitpunkt erfolgte eine Aufteilung der Kompetenzteilbereiche in einen dem Abitur vorgelagerten Block und den einzelnen Kompetenzteilbereichen *Grundlagen der Elektronik* sowie *Datenkommunikation* nach dem Abitur. Die Erfassung aller Kompetenzteilbereiche an einem Testzeitpunkt wäre optimal, da dann eventuelle Vergessenseffekte oder

Veränderungen in der persönlichen Befindlichkeit (z. B. Motivationseffekte) zwischen zwei Messzeitpunkten ausgeschlossen werden können.

#### **14.4 Ausblick**

Insgesamt bietet die vorliegende Arbeit anschlussfähige Erkenntnisse für Forschung und Praxis, liefert aber auch Impulse für zukünftige Untersuchungen. Es resultieren Forschungsdesiderate erstens in Bezug auf prozessbezogene Kompetenzen. Zweitens zur Beschreibung der Kompetenzniveaus durch affektive und motivationale Lernendenmerkmale. Drittens bezüglich einer Betrachtung der Entwicklung der prädiktiven Kraft von fluider Intelligenz und Vorwissen auf Fachwissen im weiteren Verlauf des Schulversuchs als möglicher Indikator für die spiralcurriculare Entwicklung des Fachs sowie viertens zu Lehr-Lern-Kontexten, die dazu führen, dass ggf. ein ausgewogeneres Geschlechterverhältnis im Leistungsfach erreicht wird.

Die vorliegende Arbeit widmet sich ausschließlich den inhaltsbezogenen Kompetenzen von Schüler:innen im Leistungsfach NwT. Gleichzeitig ist zu unterstellen, dass die prozessbezogenen Kompetenzen im bevorzugt projektorientierten Unterricht bedeutsam für die technische Kompetenzentwicklung sind. Ihr Einfluss kann in zukünftigen Studien z. B. mittels Videoanalysen in Kombination mit den Fachwissenstests aus dieser Arbeit untersucht werden. Des Weiteren sind bisher affektive Lernendenmerkmale und ihr Einfluss auf die Lernergebnisse unberücksichtigt geblieben. Der Einbezug in die Betrachtung ergänzt das Gesamtbild zu möglichen Einflussfaktoren auf die Lernergebnisse und führt ggf. zu einer höheren Varianzaufklärung des Fachwissens. Außerdem ist mit der differenzierteren Strukturaufklärung eine Untersuchung der Entwicklung des Fachwissens unter Einbezug kognitiver, affektiver und motivationaler Voraussetzungen gewinnbringend für eine Identifikation möglicher leistungskritischer Momente im Verlauf der Kursstufe (z. B. abnehmende Motivation o. Ä.). Eine weitere Strukturaufklärung führt offene Fragen zur Entwicklung des Fachwissens im Fach NwT bei unterstützenden Bedingungen einer Klärung zu. Voraussetzung für eine belastbare Modellierung unter Einbezug vielfältiger Konstrukte ist jedoch die Vergrößerung der Stichprobe. Diese sollte im Rahmen eines regelmäßigen Bildungsmonitorings des Leistungsfachs NwT akquiriert werden. Gleichzeitig können über ein Bildungsmonitoring in den Schulversuch einmündende Schulen zukünftig formativ begleitet und Optimierungspotenziale evidenzorientiert ausgeschöpft werden. Positive Einflüsse formativer Unterrichtsevaluation sind mit großer Effektstärke in der einschlägigen Literatur belegt (z. B. Hattie, 2020, S. 215). Mit einer regelmäßigen Messung des Fachwissens über den zeitlichen Verlauf

des Schulversuchs hinweg wäre zudem zu untersuchen, ob sich die Varianzaufklärung des Fachwissens durch den Einfluss des Vorwissens den aus der Theorie erwartbaren Werten annähert und damit eine weiter zunehmende Vernetzung von Mittel- und Kursstufe belegt. Neben der Verknüpfung von Mittel- und Kursstufe zur spiralcurricularen Optimierung der Kompetenzentwicklung besteht ein weiteres Desiderat bezüglich der Identifikation von Faktoren und Lehr-Lern-Kontexten, die dazu führen, dass ein ausgewogenes Geschlechterverhältnis im Leistungsfach erreicht wird. Diesbezüglich wird erwartet, dass eine Erfassung förderlicher und hemmender Faktoren der Kurswahlentscheidung und geschlechterspezifisches Interesse an unterschiedlichen Themen im Rahmen der Mittel- und Kursstufe Einblicke liefert, die eine günstige kontextuelle Anpassung von Unterrichtsinhalten ermöglicht. Zukünftige Forschungsvorhaben sollten sich den aufgeführten Desideraten widmen, um die evidenzorientierte Entwicklung des Schulfachs Naturwissenschaft und Technik am Gymnasium weiterhin zu unterstützen.





## Literaturverzeichnis

- Acatech & Joachim-Herz-Stiftung. (2022). *MINT Nachwuchsbarometer 2022*. Zugriff am 16.09.2022. Verfügbar unter: <https://www.acatech.de/publikation/mint-nachwuchsbarometer-2022/>
- Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational Evaluation*, 31(2-3), 162–172. <https://doi.org/10.1016/j.stueduc.2005.05.008>
- Adams, R. J. & Wu, M. (Eds.). (2003). *Programme for International Student Assessment (PISA). PISA 2000 Technical Report (PISA)*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264199521-en>
- Adenstedt, V. (2018). How boys' and girls' technical interest differs: A research study. In M. J. de Vries, S. Fletcher, S. Kruse, P. Labudde, M. Lang, I. Mammes et al. (Eds.), *Research in technology education. International approaches* (Center of Excellence for Technology Education (CETE), Vol. 2, S. 21–44). Münster: Waxmann.
- Ames, A. J. & Penfield, R. D. (2015). An NCME Instructional Module on Item-Fit Statistics for Item Response Theory Models. *Educational Measurement: Issues and Practice*, 34(3), 39–48. <https://doi.org/10.1111/emip.12067>
- Anderson, T. & Shattuck, J. (2012). Design-Based Research. A Decade of Progress in Education Research? *Educational Researcher*, 41(1), 16–25. <https://doi.org/10.3102/0013189X11428813>
- Andrich, D. & Marais, I. (2019). *A Course in Rasch Measurement Theory. Measuring in the Educational, Social and Health Sciences* (Springer eBooks Education, 1st ed. 2019). Singapore: Springer. <https://doi.org/10.1007/978-981-13-7496-8>
- Autorengruppe Bildungsplan NwT. (2019a). *Erprobungsfassung des Bildungsplans Naturwissenschaft und Technik im Basisfach - Gymnasium*. Zugriff am 14.10.2022. Verfügbar unter: <http://www.bildungsplaene-bw.de/,Lde/LS/BP2016BW/ALLG/GYM/NWTBFO>
- Autorengruppe Bildungsplan NwT. (2019b). *Erprobungsfassung des Bildungsplans Naturwissenschaft und Technik im Leistungsfach - Gymnasium*. Zugriff am 14.10.2022. Verfügbar unter: <http://www.bildungsplaene-bw.de/,Lde/LS/BP2016BW/ALLG/GYM/NWTLFO>
- Autorengruppe der Kultusministerkonferenz. (2006). Beschlüsse der Kultusministerkonferenz. Einheitliche Prüfungsanforderungen in der Abiturprüfung Technik (EPA-Technik). Zugriff am 11.03.2023. Verfügbar unter: [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/1989/1989\\_12\\_01-EPA-Technik.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Technik.pdf)

- Avsec, S. & Jamšek, J. (2016). Technological literacy for students aged 6–18: a new method for holistic measuring of knowledge, capabilities, critical thinking and decision-making. *International Journal of Technology and Design Education, 26*(1), 43–60. <https://doi.org/10.1007/s10798-015-9299-y>
- Avsec, S. & Jamšek, J. (2018). A path model of factors affecting secondary school students' technological literacy. *International Journal of Technology and Design Education, 28*(1), 145–168. <https://doi.org/10.1007/s10798-016-9382-z>
- Backhaus, K., Erichson, B., Weiber, R. & Plinke, W. (2015). Clusteranalyse. In K. Backhaus, B. Erichson, W. Plinke & R. Weiber (Hrsg.), *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (14. Aufl. 2016, S. 453–516). Berlin, Heidelberg: Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-46076-4\\_9](https://doi.org/10.1007/978-3-662-46076-4_9)
- Bahr, T. & Zinn, B. (2023). Gender Differences in the New Interdisciplinary Subject Informatik, Mathematik, Physik (IMP) - Sticking with STEM? *Educ. Sci., im Druck*.
- Baumert, J., Bos, W. & Lehmann, R. (Hrsg.). (2000). *Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe* (TIMSS/III, Bd. 2). Opladen: Leske + Budrich.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9*(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Bernholt, S. (2010). *Kompetenzmodellierung in der Chemie. Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität* (Studien zum Physik- und Chemielernen, Bd. 98). Zugl.: Oldenburg, Univ., Diss., 2010. Berlin: Logos-Verl.
- Bogner, A. & Menz, W. (2002). Das theoriegenerierende Experteninterview. In A. Bogner, B. Littig & W. Menz (Hrsg.), *Das Experteninterview. Erkenntnisinteresse, Wissensformen, Interaktion* (S. 33–70). Wiesbaden: VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-322-93270-9\\_2](https://doi.org/10.1007/978-3-322-93270-9_2)
- Boone, W. J. (2016). Rasch Analysis for Instrument Development: Why, When, and How? *CBE Life Sciences Education, 15*(4). <https://doi.org/10.1187/cbe.16-04-0148>
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (Springer-Lehrbuch, 7., vollständig überarbeitete und erweiterte Auflage). Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-12770-0>
- Brändle, M. (2023). Das Leistungsfach Naturwissenschaft und Technik (NwT) in Baden-Württemberg. In M. Binder & C. Wiesmüller (Hrsg.), *Technikunterricht*

- konkret. (S. 88–111). Reutlingen, 23. - 24. September 2022. DGTB. Offenbach a. M.: BE.ER Konzept.
- Brühwiler, C. & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction, 21*(1), 95–108. <https://doi.org/10.1016/j.learninstruc.2009.11.004>
- Busch, M. (2016). *Empirische Studien zum fächerübergreifenden naturwissenschaftlichen Unterricht: Kompetenzförderung, Interessenentwicklung, Wahlmotive und Lehrerperspektive*. Dissertation, Jena, Friedrich-Schiller-Universität Jena, 2016.
- Calinski, T. & Harabasz, J. (1974). A Dendrite Method for Cluster Analysis. *Communications in Statistics - Theory and Methods, 3*(1), 1–27. <https://doi.org/10.1080/03610927408827101>
- Castillo, M. (2010). Technological Literacy: Design and Testing an Instrument to Measure Eighth Grade Achievement in Technology Education. In *2010 Annual Conference & Exposition Proceedings* (15.1196.1-15.1196.19). ASEE Conferences.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1–22. <https://doi.org/10.1037/h0046743>
- Charrad, M., Ghazzali, N., Boiteau, V. & Niknafs, A. (2014). NbClust : An R Package for Determining the Relevant Number of Clusters in a Data Set [Computer software].
- Chen, W.-H., Lenderking, W., Jin, Y., Wyrwich, K. W., Gelhorn, H. & Revicki, D. A. (2014). Is Rasch model analysis applicable in small sample size pilot studies for assessing item characteristics? An example using PROMIS pain behavior item bank data. *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 23*(2), 485–493. <https://doi.org/10.1007/s11136-013-0487-5>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hoboken: Taylor and Francis.
- Conrad, K.-J. (2019). *Grundlagen der Konstruktionslehre. Maschinenbau-Anwendungen und Orientierung auf Menschen; mit 104 Tabellen, zahlreichen Kenntnisfragen und Aufgabenstellungen mit Lösungen* (Hanser eLibrary, 7., aktualisierte und erweiterte Auflage). München: Hanser.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. <https://doi.org/10.1007/BF02310555>

- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–238.
- Deci, E. L. & Ryan, R. M. (Eds.). (2002). *Handbook of self-determination research* (1. publ). Rochester, NY: University of Rochester Press.
- DeMars, C. E. (2018). Classical Test Theory and Item Response Theory. In F. P. Irwing (Ed.), *The Wiley handbook of psychometric testing. A multidisciplinary reference on survey, scale, and test development* (S. 49-74). Hoboken: Wiley.
- Ditton, H. & Maaz, K. (2015). Sozioökonomischer Status und soziale Ungleichheit. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (S. 229–244). Wiesbaden: Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-531-19994-8\\_17](https://doi.org/10.1007/978-3-531-19994-8_17)
- Dörnyei, Z. (2001). *Teaching and researching motivation* (Applied linguistics in action). Harlow: Longman.
- Duismann, G. H. & Meschenmoser, H. (2009). Technisches Verständnis als arbeitsrelevante Basiskompetenz. Empirische Befunde zur Kompetenzdiagnostik technischer Grundbildung. In W. E. Theuerkauf & H. Meschenmoser (eds.), *Qualität technischer Bildung. Kompetenzmodelle und Kompetenzdiagnostik* (Schriften zu Arbeit - Beruf - Bildung, Bd. 3). Berlin: Machmit-Verlag.
- Ebrahiminejad, H. (2017). A Systematized Literature Review: Defining and Developing Engineering Competencies. In *2017 ASEE Annual Conference & Exposition Proceedings*. ASEE Conferences.
- European Commission. Directorate General for Education, Youth, Sport and Culture & ECORYS. (2020). *Supporting key competence development: learning approaches and environments in school education : input paper*. Publications Office. <https://doi.org/10.2766/8227>
- Fischer, F., Waibel, M. & Wecker, C. (2005). Nutzenorientierte Grundlagenforschung im Bildungsbereich. *Zeitschrift für Erziehungswissenschaft*, 8(3), 427–442. <https://doi.org/10.1007/s11618-005-0149-7>
- Fleischer, J., Koeppen, K., Kenk, M., Klieme, E. & Leutner, D. (2013). Kompetenzmodellierung. Struktur, Konzepte und Forschungszugänge des DFG-Schwerpunktprogramms. *Zeitschrift für Erziehungswissenschaft*, 16(1), 5–22. <https://doi.org/10.1007/s11618-013-0379-z>
- Fletcher, S. (2020). Technische Mündigkeit als wichtiger Bestandteil der Allgemeinbildung. Zwischen Ingenieurpädagogik, Lehrkräftebildung und betrieblicher Praxis. In F. Bünning, M. Dick, R. W. Jahn & A. Seltrecht (Hrsg.), *Zwischen Ingenieurpädagogik, Lehrkräftebildung und betrieblicher Praxis. Eine*

- Festschrift für Klaus Jenewein* (Berufsbildung, Arbeit und Innovation, Band 57, S. 127–142). Bielefeld: wbv.
- Fletcher, S. & Deutsch, J. (2016). Energiemündigkeit von Schülerinnen und Schülern am Ende der Sekundarstufe I - Konzeptionalisierung eines Modells zur Energiemündigkeit und Entwicklung eines darauf basierenden Testwerkzeugs. *Journal of Technical Education (JOTED)*, Bd. 4 Nr. 2 (2016): *Journal of Technical Education (JOTED)*. <https://doi.org/10.48513/JOTED.V4I2.81>
- Fletcher, S., Vries, M. J. de & Max, C. (2018). Die technische Mündigkeit von Schüler/-innen zum Ende der Sek. I im internationalen Vergleich. *Journal of Technical Education (JOTED)*, Bd. 6 Nr. 4 (2018): *Journal of Technical Education (JOTED)*. <https://doi.org/10.48513/JOTED.V6I4.154>
- Flick, U. (2021). *Qualitative Sozialforschung. Eine Einführung* (Rororo Rowohlts Enzyklopädie, Bd. 55694, 10. Auflage, Originalausgabe). Reinbek bei Hamburg: rowohlts enzyklopädie im Rowohlt Taschenbuch Verlag.
- Florian, C., Sandmann, A. & Schmiemann, P. (2014). Modellierung kognitiver Anforderungen schriftlicher Abituraufgaben im Fach Biologie. *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), 175–189.
- Förster, S. (2016). *Selbstbezogene Kognitionen und Motivationen im Grundschulalter* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 99, [1. Auflage]. Dissertation.
- Früh, W. (2004). *Inhaltsanalyse. Theorie und Praxis* (UTB Medien- und Kommunikationswissenschaft, Psychologie, Soziologie, Bd. 2501, Unveränd. Nachdr. der 5. Aufl. von 2001). Konstanz: UVK Verl.-Ges.
- Geißel, B., Nickolaus, R., Stefanica, F., Härtig, H. & Neumann, K. (2013). Die Relevanz mathematischer und naturwissenschaftlicher Kompetenzen für die fachliche Kompetenzentwicklung in gewerblich-technischen Berufen. In R. Nickolaus, J. Retelsdorf, E. Winther & O. Köller (Hrsg.), *Mathematisch-naturwissenschaftliche Kompetenzen in der beruflichen Erstausbildung. Stand der Forschung und Desiderata* (*Zeitschrift für Berufs- und Wirtschaftspädagogik – Beihefte*, Bd. 26, 1. Auflage, S. 39–66). Stuttgart: Franz Steiner Verlag.
- Gläser-Zikuda, M. (2015). III-7 Qualitative Auswertungsverfahren. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (S. 119–130). Wiesbaden: Springer Fachmedien Wiesbaden.  
[https://doi.org/10.1007/978-3-531-19992-4\\_9](https://doi.org/10.1007/978-3-531-19992-4_9)
- Göhner, M. & Krell, M. (2020). Qualitative Inhaltsanalyse in naturwissenschafts-didaktischer Forschung unter Berücksichtigung von Gütekriterien: Ein Review. *Zeitschrift für Didaktik der Naturwissenschaften*, 26(1), 207–225.  
<https://doi.org/10.1007/s40573-020-00111-0>

- Hahn, S. (2013). Wissenschaftspropädeutik in der gymnasialen Oberstufe. In D. Bosse, F. Eberle & B. Schneider-Taylor (Hrsg.), *Standardisierung in der gymnasialen Oberstufe* (SpringerLink Bücher, S. 161–174). Wiesbaden: Springer VS.
- Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (Springer-Lehrbuch, 2., aktualisierte und überarbeitete Auflage, S. 143-171). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Hartig, J., Frey, A., Nold, G. & Klieme, E. (2012). An Application of Explanatory Item Response Modeling for Model-Based Proficiency Scaling. *Educational and Psychological Measurement*, 72(4), 665–686.  
<https://doi.org/10.1177/0013164411430707>
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127–143). Berlin/Heidelberg: Springer-Verlag. [https://doi.org/10.1007/3-540-33020-8\\_9](https://doi.org/10.1007/3-540-33020-8_9)
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement* (Reprinted.). London: Routledge.
- Hattie, J. (2020). *Lernen sichtbar machen. Mit Index und Glossar* (5. unveränderte Auflage, erweiterte Auflage mit Index und Glossar). Baltmannsweiler: Schneider Verlag Hohengehren GmbH.
- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis. A regression-based approach* (Methodology in the social sciences, Third edition). New York, London: The Guilford Press.
- Hedrich, M. (2021). *Schulische, betriebliche und private Einflussfaktoren auf Fachwissen bei Elektronikern für Automatisierungstechnik am Ende der Ausbildung*. Dissertation. Universität Stuttgart. <https://doi.org/10.18419/OPUS-11714>
- Hess, B., Subhiyah, R. G. & Giordano, C. (2007). Convergence between cluster analysis and the Angoff method for setting minimum passing scores on credentialing examinations. *Evaluation & the Health Professions*, 30(4), 362–375.  
<https://doi.org/10.1177/0163278707307904>
- Hoffmann, C. (2021). *Studien zum fächerübergreifenden naturwissenschaftlichen Unterricht in der universitären Lehrerbildung: hochschuldidaktisches Veranstaltungskonzept, professionelle Überzeugungen, Natur der Naturwissenschaften*. Dissertation. Friedrich-Schiller-Universität Jena.  
<https://doi.org/10.22032/DBT.49162>

- International Technology and Engineering Educators Association. (2007). *Standards for technological Literacy: Content for the study of technology*. Reston, VA.
- International Technology and Engineering Educators Association. (2020). *Standards for Technological and Engineering Literacy. The Role of Technology and Engineering in STEM Education*. Executive Summary. Zugriff am 28.12.2022. Verfügbar unter: <https://www.iteea.org/189252.aspx>
- Jonkisz, E., Moosbrugger, H. & Brandt, H. (2012). Planung und Entwicklung von Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (Springer-Lehrbuch, 2., aktualisierte und überarbeitete Auflage, S. 27-74). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Kassambara, A. (2023). rstatix: Pipe-Friendly Framework for Basic Statistical Tests (Version 0.7.2) [Computer software]. Verfügbar unter: <https://CRAN.R-project.org/package=rstatix>
- Kassambara, A. & Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses (Version 1.0.7) [Computer software]. Verfügbar unter: <https://CRAN.R-project.org/package=factoextra>
- Kauertz, A. (2008). *Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben* (Studien zum Physik- und Chemielernen, Bd. 79). Zugl.: Duisburg, Essen, Univ., Diss., 2007. Berlin: Logos-Verl.
- Kauertz, A., Fischer, H. E., Mayer, J., Sumfleth, E. & Walpuski, M. (2010). Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 135–153.
- Kelava, A. & Moosbrugger, H. (2012). Deskriptivstatistische Evaluation von Items (Itemanalyse) und Testwertverteilungen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (Springer-Lehrbuch, 2., aktualisierte und überarbeitete Auflage, S. 75-102). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Kennedy, T. J. & Sundberg, C. W. (2020). 21st Century Skills. In B. Akpan & T. J. Kennedy (Hrsg.), *Science Education in Theory and Practice*. Cham: Springer International Publishing.
- Kleinknecht, M., Bohl, T., Maier, U. & Metz, K. (Hrsg.). (2013). *Lern- und Leistungsaufgaben im Unterricht. Fächerübergreifende Kriterien zur Auswahl und Analyse*. Bad Heilbrunn: Verlag Julius Klinkhardt.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. et al. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Verfügbar unter: [https://www.pedocs.de/frontdoor.php?source\\_opus=20901](https://www.pedocs.de/frontdoor.php?source_opus=20901)

- Koller, I., Alexandrowicz, R. & Hatzinger, R. (2012). *Das Rasch Modell in der Praxis* (1. Aufl.). Stuttgart: UTB GmbH; facultas.wuv.
- Köller, O. (2011). Standardsetzung im Bildungssystem. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (S. 179–192). Wiesbaden: VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-93015-2\\_15](https://doi.org/10.1007/978-3-531-93015-2_15)
- Krapp, A. & Prenzel, M. (1992). *Interesse, Lernen und Leistung. Neuere Ansätze der pädagogisch-psychologischen Interessenforschung*. Münster: Aschendorff.
- Krapp, A. & Prenzel, M. (2011). Research on Interest in Science: Theories, methods, and findings. *International Journal of Science Education*, 33(1), 27–50. <https://doi.org/10.1080/09500693.2010.518645>
- Krupczak, J. & Disney, K. (2013). Technological Literacy: Assessment and Measurement of Learning Gains. In *2013 ASEE Annual Conference & Exposition Proceedings* (23.1160.1-23.1160.11). ASEE Conferences.
- Kuckartz, U. & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung : Grundlagentexte Methoden* (Grundlagentexte Methoden, 5. Auflage). Weinheim, Basel: Beltz Juventa.
- Labudde, P. (2014). Fächerübergreifender naturwissenschaftlicher Unterricht – Mythen, Definitionen, Fakten. *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), 11–19. <https://doi.org/10.1007/s40573-014-0001-9>
- Labudde, P. & Schecker, H. (2021). Unterrichtskonzeptionen für fächerübergreifenden naturwissenschaftlichen Unterricht. In T. Wilhelm, H. Schecker, M. Hopf, R. Berger, J.-P. Burde, C. Haagen-Schützenhöfer et al. (Hrsg.), *Unterrichtskonzeptionen für den Physikunterricht. Ein Lehrbuch für Studium, Referendariat und Unterrichtspraxis* (Lehrbuch, S. 435–473). Berlin: Springer Spektrum. [https://doi.org/10.1007/978-3-662-63053-2\\_14](https://doi.org/10.1007/978-3-662-63053-2_14)
- Landesinstitut für Schulentwicklung. (2017). *Beispielcurriculum für das Fach NwT. Klassen 8 bis 10*, Stuttgart.
- Läpple, V. (2016). *Einführung in die Festigkeitslehre. Lehr- und Übungsbuch* (4., aktualisierte Auflage). Wiesbaden: Springer Vieweg. <https://doi.org/10.1007/978-3-658-10611-9>
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2022). cluster: Cluster Analysis Basics and Extensions (Version 2.1.4) [Computer software]. Verfügbar unter: <https://CRAN.R-project.org/package=cluster>
- Mahler, N., Schipolowski, S. & Weirich, S. (2019). Anlage, Durchführung und Auswertung des IQB-Bildungstrends 2018. In P. Stanat, S. Schipolowski, N.



- Mahler, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2018. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich* (1. Auflage, 99-124). Münster: Waxmann.
- Mair, P., Hatzinger, R. & Maier, M. J. (2021). eRm: Extended Rasch Modeling (Version 1.0-2) [Computer software]. Verfügbar unter: <https://cran.r-project.org/package=eRm>
- Makowski, D., Ben-Shachar, M. S., Patil, I. & Lüdtke, D. (2020). Methods and Algorithms for Correlation Analysis in R. *Journal of Open Source Software*, 5(51), 2306. <https://doi.org/10.21105/joss.02306>
- Mandl, H. & Friedrich, H. F. (Hrsg.). (1992). *Lern- und Denkstrategien. Analyse und Intervention*. Göttingen: Hogrefe Verl. für Psychologie.
- Mayring, P. (2022). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (13., überarbeitete Auflage). Weinheim, Basel: Beltz.
- Merk, S., Poindl, S. & Bohl, T. (2018). *Wie sollten Rückmeldungen von quantitativ erfasstem Schülerfeedback (nicht) gestaltet werden?* <https://doi.org/10.31219/osf.io/7r6kb>
- Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 2(3), 255–273. <https://doi.org/10.1080/10705519509540013>
- Milligan, G. W. & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50(2), 159–179. <https://doi.org/10.1007/BF02294245>
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg. (2016). *Allgemeinbildendes Gymnasium Baden-Württemberg. Bildungsplan Naturwissenschaft und Technik (NwT) - Profilmfach*. Stuttgart. Zugriff am 14.10.2022. Verfügbar unter: <http://www.bildungsplaene-bw.de/Lde/LS/BP2016BW/ALLG/GYM/NWT>
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg. (2022). *Leitfaden für die gymnasiale Oberstufe. Abitur 2024*. Zugriff am 02.11.2022. Verfügbar unter: [https://km-bw.de/site/pbs-bw-km-root/get/documents\\_E-1537580629/KULTUS.Dachmandant/KULTUS/KM-Homepage/Publikationen%202021/2021\\_Leitfaden%20fuer%20die%20gymnasiale%20Oberstufe\\_Abitur%202024.pdf](https://km-bw.de/site/pbs-bw-km-root/get/documents_E-1537580629/KULTUS.Dachmandant/KULTUS/KM-Homepage/Publikationen%202021/2021_Leitfaden%20fuer%20die%20gymnasiale%20Oberstufe_Abitur%202024.pdf)
- Misic, J. (2018). *Entwicklung und Pilotierung eines Testinstruments zur Erfassung fachlicher Kompetenzen in ausgewählten Bereichen der Kursstufe im Fach Naturwissenschaft und Technik*. Staatsexamensarbeit (unveröffentlicht). Universität Stuttgart, Stuttgart.

- MKJS. (2004). *Bildungsplan 2004. Allgemeinbildendes Gymnasium Baden-Württemberg*, Stuttgart.
- Mokhonko, S., Stefanica, F. & Nickolaus, R. (2014). NwT-Unterricht. Herausforderungen bei der Einführung eines neuen Faches im Spiegel einer aktuellen Bestandsaufnahme. *Journal of Technical Education (JOTED)*, 2(1).
- Moosbrugger, H. (2012). Klassische Testtheorie (KTT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (Springer-Lehrbuch, 2., aktualisierte und überarbeitete Auflage, S. 103-117). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Moosbrugger, H. & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (Springer-Lehrbuch, 2., aktualisierte und überarbeitete Auflage, S. 7-26). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- National Academy of Engineering and National Research Council. (2006). *Tech Tally: Approaches to Assessing Technological Literacy*. Washington, D.C.: National Academies Press. <https://doi.org/10.17226/11691>
- Neumann, K., Kauertz, A., Lau, A., Notarp, H. & Fischer, H. E. (2007). *Die Modellierung physikalischer Kompetenz und ihrer Entwicklung* (Bd. 13). Verfügbar unter: [https://www.researchgate.net/profile/hans-fischer/publication/281345574\\_die\\_modellierung\\_physikalischer\\_kompetenz\\_und\\_ihrer\\_entwicklung](https://www.researchgate.net/profile/hans-fischer/publication/281345574_die_modellierung_physikalischer_kompetenz_und_ihrer_entwicklung)
- Nickolaus, R., Abele, S. & Albus, A. (2015). Technisches Vorwissen als Prädiktor für die berufsfachliche Kompetenzentwicklung in gewerblich-technischen Berufen. In L. Windelband & S. Kruse (Hrsg.), *Technik im Spannungsfeld der Allgemeinen und Beruflichen Bildung. Erste Fachtagung der Technischen Bildung in Baden-Württemberg* (Schriftenreihe Didaktik in Forschung und Praxis, Bd. 77, S. 9–29). Hamburg: Kovač.
- Nickolaus, R., Abele, S., Gschwendtner, T., Nitzschke, A. & Greiff, S. (2012). Fachspezifische Problemlösefähigkeit in gewerblich-technischen Ausbildungsberufen-Modellierung, erreichte Niveaus und relevante Einflussfaktoren. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 108(2), 243–272.
- OECD. (2007). *PISA 2006. Naturwissenschaftliche Kompetenzen für die Welt von morgen*. Kurzzusammenfassung. Zugriff am 18.03.2023. Verfügbar unter: <https://www.oecd.org/pisa/39731064.pdf>
- OECD. (2013). *PISA 2012 Assessment and Analytical Framework*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264190511-en>

- OECD. (2016). *PISA 2015. Ergebnisse im Fokus*. Zugriff am 18.03.2023. Verfügbar unter: [https://www.oecd.org/berlin/themen/pisa-studie/PISA\\_2015\\_Zusammenfassung.pdf](https://www.oecd.org/berlin/themen/pisa-studie/PISA_2015_Zusammenfassung.pdf)
- OECD. (2019). *Trends Shaping Education 2019* (Trends Shaping Education). Paris: OECD Publishing. [https://doi.org/10.1787/trends\\_educ-2019-en](https://doi.org/10.1787/trends_educ-2019-en)
- OECD. (2020). *Lernkompass 2030. OECD-Projekt Future of Education and Skills 2030*. Rahmenkonzept des Lernens (Bertelsmann Stiftung, Deutsche Telekom Stiftung, Education Y e. V., Global Goals Curriculum e. V. & Siemens Stiftung, Hrsg.). Zugriff am 26.03.2023. Verfügbar unter: <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/oecd-lernkompass-2030-all#detail-content-7004-4>
- OECD. (2022). *Trends Shaping Education 2022* (Trends Shaping Education Ser., 1st ed.). Paris: Organization for Economic Cooperation & Development. <https://doi.org/10.1787/6ae8771a-en>
- Offermann, G. & Schäfer, A. (2012). Zur Ideen- und Entstehungsgeschichte des Faches Naturwissenschaften und Technik (NwT) in Baden-Württemberg. In U. Pfenning & O. Renn (Hrsg.), *Wissenschafts- und Technikbildung auf dem Prüfstand. Zum Fachkräftemangel und zur Attraktivität der MINT-Bildung und -Berufe im europäischen Vergleich* (1. Auflage, S. 183–199). Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG. Zugriff am 14.10.2022. Verfügbar unter: <https://www.nomos-elibrary.de/10.5771/9783845238289-183.pdf>
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (2013). *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Verfügbar unter: <https://elibrary.utb.de/doi/pdf/10.31244/9783830979906>
- Pant, H. A., Tiffin-Richards, S. P. & Köller, O. (2010). Standard-Setting für Kompetenztests im Large-Scale-Assessment. Projekt Standardsetting. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (Zeitschrift für Pädagogik, Beiheft; 56). Weinheim: Beltz. Verfügbar unter: [https://www.pedocs.de/frontdoor.php?source\\_opus=3406](https://www.pedocs.de/frontdoor.php?source_opus=3406)
- Parchmann, I. (2010). Kompetenzmodellierung in den Naturwissenschaften. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (Zeitschrift für Pädagogik, Beiheft; 56). Weinheim: Beltz. Verfügbar unter: [https://www.pedocs.de/volltexte/2010/3324/pdf/beiheft56\\_komplett\\_d\\_a.pdf#page=136](https://www.pedocs.de/volltexte/2010/3324/pdf/beiheft56_komplett_d_a.pdf#page=136)

- Petrillo, J., Cano, S. J., McLeod, L. D. & Coon, C. D. (2015). Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value in Health : the Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 18(1), 25–34.  
<https://doi.org/10.1016/j.jval.2014.10.005>
- Pittschellis, R. (2012). Der Beitrag von Festo für die didaktische Strukturierung der Technikbildung. In U. Pfenning & O. Renn (Hrsg.), *Wissenschafts- und Technikbildung auf dem Prüfstand. Zum Fachkräftemangel und zur Attraktivität der MINT-Bildung und -Berufe im europäischen Vergleich* (1. Auflage, S. 223–233). Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG. Zugriff am 14.10.2022. Verfügbar unter: <https://www.nomos-elibrary.de/10.5771/9783845238289-223.pdf>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, 88(5), 879–903.  
<https://doi.org/10.1037/0021-9010.88.5.879>
- R Core Team. (2022). R: A Language and Environment for Statistical Computing (Version 4.2.2) [Computer software]. Vienna, Austria. Verfügbar unter: <https://www.R-project.org/>
- Rädiker, S. & Kuckartz, U. (2019). Das Kategoriensystem gestalten. In S. Rädiker & U. Kuckartz (Hrsg.), *Analyse qualitativer Daten mit MAXQDA* (S. 95–110). Wiesbaden: Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-22095-2\\_8](https://doi.org/10.1007/978-3-658-22095-2_8)
- Reinfried, S. & Tempelmann, S. (2014). *Wie Vorwissen das Lernen beeinflusst – Eine Lernprozessstudie zur Wissenskonstruktion des Treibhauseffekt-Konzepts*. <https://doi.org/10.18452/23977>
- Reinmann, G. (2005). Innovation ohne Forschung? Ein Plädoyer für den Design-Based Research-Ansatz in der Lehr-Lernforschung. *Unterrichtswissenschaft*, 33(1), 52–69.
- Reiss, C., Sälzer, C., Schiepe-Tiska, A., Klieme, E. & Köller, O. ((2016). *PISA 2015. Eine Studie zwischen Kontinuität und Innovation*. Zusammenfassung. Münster: Waxmann. Zugriff am 18.03.2023. Verfügbar unter: [https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende\\_und\\_Zusammenfassungen/PISA\\_2015\\_Zusammenfassung\\_final.pdf](https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungen/PISA_2015_Zusammenfassung_final.pdf)
- Robitzsch, A., Kiefer, T. & Wu, M. (2022). TAM: Test Analysis Modules (Version 4.1-4) [Computer software]. Verfügbar unter: <https://CRAN.R-project.org/package=TAM>

- Rürup, M., Fuchs, H.-W. & Weishaupt, H. (2016). Bildungsberichterstattung – Bildungsmonitoring. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 411–437). Wiesbaden: Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-531-18942-0\\_15](https://doi.org/10.1007/978-3-531-18942-0_15)
- Sahin, A. & Anil, D. (2017). The Effects of Test Length and Sample Size on Item Parameters in Item Response Theory. *Educational Sciences: Theory & Practice*. <https://doi.org/10.12738/estp.2017.1.0270>
- Sälzer, C. (2016). *Studienbuch Schulleistungsstudien*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-45765-8>
- Sälzer, C. & Prenzel, M. (2013). PISA 2012–eine Einführung in die aktuelle Studie. In *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Waxmann.
- Schermelleh-Engel, K. & Werner, C. S. (2012). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (Springer-Lehrbuch, 2., aktualisierte und überarbeitete Auflage, S. 119-141). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Schlüter, K. (2015). *Ansätze zur Modellierung informatischer Anforderungen und Fähigkeiten von Schülern auf der Basis von Aufgabenmerkmalen*. Verfügbar unter: <https://core.ac.uk/download/pdf/33797233.pdf>
- Schmidt, H. K. (2018). "Man muss (wissen) entscheiden – schwer". *Die Bedeutung von Vorwissen beim Lernen mit bioethischen Dilemmata*. Verfügbar unter: <https://oops.uni-oldenburg.de/id/eprint/3935>
- Schrader, F.-W. & Helmke, A. (2008). Determinanten der Schulleistung. In M. K. W. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion* (S. 285–302). Wiesbaden: VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-91104-5\\_11](https://doi.org/10.1007/978-3-531-91104-5_11)
- Schreier, M. (2012). *Qualitative content analysis in practice*. Los Angeles, London, New Dehli, Singapore, Washington DC: SAGE.
- Sedlmeier, P. & Burkhardt, M. (2021). *Datenanalyse mit R. Beschreiben, Explorieren, Schätzen und Testen*. München: Pearson.
- Sedlmeier, P. & Renkewitz, F. (2018). *Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler* (3., aktualisierte und erweiterte Auflage). Hallbergmoos: Pearson.
- Seidel, T. (2014). *Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma*. Weinheim, Basel: BeltzJuventa.
- Seidel, T. & Reiss, K. (2014). Lerngelegenheiten im Unterricht. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie. Mit Online-Materialien* (6., vollständig überarbeitete Auflage, S. 253–276). Weinheim: Beltz.

- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15(2), 4–14.  
<https://doi.org/10.3102/0013189X015002004>
- Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L. & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist*, 57(1), 31–54. <https://doi.org/10.1080/00461520.2021.1939700>
- Sireci, S., Robin, F. & Patelis, T. (1999). Using Cluster Analysis to Facilitate Standard Setting. *Applied Measurement in Education*, 12, 301–325.  
[https://doi.org/10.1207/S15324818AME1203\\_5](https://doi.org/10.1207/S15324818AME1203_5)
- Spiegelhauer, J. (2017). *Bedeutung und Förderung funktionalen Denkens im Kontext des Unterrichts aus mathemathistorischer, fachdidaktischer und unterrichtspraktischer Perspektive*. Dissertation. Martin-Luther-Universität, Halle-Wittenberg.
- SPSS Inc. (2017). IBM SPSS Statistics for Windows (Version 25.0) [Computer software]. Armonk, NY: IBM Corp.
- SPSS Inc. (2021). IBM SPSS Statistics for Windows (Version 28.0) [Computer software]. Armonk, NY: IBM Corp.
- Stanat, P., Schipolowski, S., Mahler, N., Weirich, S. & Henschel, S. (2019). *IQB-Bildungstrend 2018: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich*. Verfügbar unter: <https://elibrary.utb.de/doi/pdf/10.31244/9783830990444>
- Stemmann, J. (2018). Problemlösen im Umgang mit technischen Geräten - eine allgemeine oder kontextspezifische Kompetenz? *Empirische Pädagogik*, 32, 46–62.
- Stemmann, J. & Lang, M. (2014). Theoretische Konzeption einer allgemeinen technischen Problemlösefähigkeit und Möglichkeiten ihrer Diagnose. *Journal of Technical Education (JOTED)*, Bd. 2 Nr. 1 (2014): *Journal of Technical Education (JOTED)*. <https://doi.org/10.48513/JOTED.V2I1.23>
- Stifterverband für die Deutsche Wissenschaft e. V. (2021). *Future Skills 2021. 21 Kompetenzen für eine Welt im Wandel*. Diskussionspapier Nr. 3. Verfügbar unter: [https://hochschulbildungsreport2020.de/2021/future\\_skills\\_2021](https://hochschulbildungsreport2020.de/2021/future_skills_2021)
- Strobl, C. (2015). *Das Rasch-Modell. Eine verständliche Einführung für Studium und Praxis* (Sozialwissenschaftliche Forschungsmethoden, Bd. 2, 3. Aufl.). Mering: Rainer Hampp Verlag.
- Stumpf, H. (1996). *Klassische Testtheorie*. Verfügbar unter: <https://madoc.bib.uni-mannheim.de/50962/1/hqm31-klassische-testtheorie.pdf>

- Templin, J. & Jiao, H. (2012). Applying Model-Based Approaches to Identify Performance Categories. In G. J. Cizek (Hrsg.), *Setting performance standards. Foundations, Methods and Innovations* (S. 379–397). Routledge.
- Theodoridis, S. & Koutroumbas, K. (Eds.). (2010). *Pattern recognition* (4th ed.). Burlington, MA: Academic Press.
- Theuerkauf, W. E. (2009). Voraussetzungen zur Erfassung von Kompetenzen und Standards der technischen Bildung in der allgemeinbildenden Schule. In W. E. Theuerkauf & H. Meschenmoser (eds.), *Qualität technischer Bildung. Kompetenzmodelle und Kompetenzdiagnostik* (Schriften zu Arbeit - Beruf - Bildung, Bd. 3, 78-92). Berlin: Machmit-Verlag.
- Thomas, C. L., Cassady, J. C. & Finch, W. H. (2018). Identifying Severity Standards on the Cognitive Test Anxiety Scale: Cut Score Determination Using Latent Class and Cluster Analysis. *Journal of Psychoeducational Assessment*, 36(5), 492–508. <https://doi.org/10.1177/0734282916686004>
- Trendtel, M., Pham, G. & Yanagida, T. (2016). Skalierung und Linking. In S. Breit & C. Schreiner (Hrsg.), *Large-Scale Assessment mit R. Methodische Grundlagen der österreichischen Bildungsstandardüberprüfung* (1. Auflage, S. 185–224). Wien: facultas.
- Tseng, F.-L., Chiou, J.-M. & Sung, Y.-T. (2017). Using Custer Analysis to validate the Angoff Standard Setting Method in Mixed-format Assessments, 2540–2546. <https://doi.org/10.1109/FSKD.2017.8393175>
- Van Laar, E., van Deursen, A. J., van Dijk, J. A. & Haan, J. de. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577–588. <https://doi.org/10.1016/j.chb.2017.03.010>
- Van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M. & Haan, J. de. (2020). Determinants of 21st-Century Skills and 21st-Century Digital Skills for Workers: A Systematic Literature Review. *SAGE Open*, 10(1), 215824401990017. <https://doi.org/10.1177/2158244019900176>
- VDMA. (2019). *Technikunterricht in Deutschland. Eine Analyse und Bewertung von Technik in den Curricula allgemeinbildender Schulen*. Zugriff am 24.04.2023. Verfügbar unter: [https://www.vdma.org/documents/34570/0/VDMA+Kompendium+-+Technikunterricht+in+Deutschland\\_1567435852469.pdf/f54af245-345d-4d7a-4ae9-cad4c1d38fdd?t=1612885653616](https://www.vdma.org/documents/34570/0/VDMA+Kompendium+-+Technikunterricht+in+Deutschland_1567435852469.pdf/f54af245-345d-4d7a-4ae9-cad4c1d38fdd?t=1612885653616)
- Violato, C., Marini, A. & Lee, C. (2003). A validity study of expert judgment procedures for setting cutoff scores on high-stakes credentialing examinations

- using cluster analysis. *Evaluation & the Health Professions*, 26(1), 59–72. <https://doi.org/10.1177/0163278702250082>
- Vries, M. J. de, Fletcher, S., Kruse, S., Labudde, P., Lang, M., Mammes, I. et al. (Eds.). (2018). *Research in technology education. International approaches* (Center of Excellence for Technology Education (CETE), Vol. 2). Münster: Waxmann.
- Walach, M. (2015). *Measuring the Influences That Affect Technological Literacy in Rhode Island High Schools*. University of Rhode Island.
- Walter, O. (2009). Herkunftsassoziierte Disparitäten im Lesen, der Mathematik und den Naturwissenschaften: ein Vergleich zwischen PISA 2000, PISA 2003 und PISA 2006. In M. Prenzel & J. Baumert (Hrsg.), *Vertiefende Analysen zu PISA 2006* (S. 149–168). Wiesbaden: VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-91815-0\\_8](https://doi.org/10.1007/978-3-531-91815-0_8)
- Weinert, F. E. (Hrsg.). (2001). *Leistungsmessungen in Schulen* (Beltz-Pädagogik). Weinheim und Basel: Beltz.
- Weiß, R. H. (2019). *CFT 20-R mit WS/ZF-R. Grundintelligenztest Skala 2 – Revision mit Wortschatztest und Zahlenfolgentest – Revision* (2. überarbeitete Auflage mit aktualisierten und erweiterten Normen). Göttingen: Hogrefe.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Verfügbar unter: <https://ggplot2.tidyverse.org>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R. et al. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Wiedenbeck, M. & Züll, C. (2010). Clusteranalyse. In H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (SpringerLink Bücher, S. 525–552). Wiesbaden: VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-92038-2\\_21](https://doi.org/10.1007/978-3-531-92038-2_21)
- Wilson, M. (2004). *Constructing Measures. An Item Response Modeling Approach*. New York: Taylor & Francis Group, LLC. <https://doi.org/10.4324/9781410611697>
- Wirtz, M. A. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Hogrefe.
- Woitkowski, D. & Riese, J. (2017). Kriterienorientierte Konstruktion eines Kompetenzniveau-modells im physikalischen Fachwissen. *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 39–52. <https://doi.org/10.1007/s40573-016-0054-z>



- Wyrwal, M. (2020). *Das berufsfachliche Wissen von Schülerinnen und Schülern in der Fachschule Bautechnik* (Sozialwissenschaften heute, Bd. 6). Dissertation.
- Zendler, A. (Hrsg.). (2018). *Unterrichtsmethoden für den Informatikunterricht*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-20675-8>
- Zieky, M. J. (2012). So Much has Changed. An Historical Overview of Setting Cut Scores. In G. J. Cizek (Hrsg.), *Setting performance standards. Foundations, Methods and Innovations* (S. 15–32). Routledge.
- Zieky, M. J. & Livingston, S. A. (1977). *Basic Skills Assessment. Manual for Setting Standards on the Basic Skills Assessment Tests*. Educational Testing Service, Basic Skills Assessment, Rosedale Road, Princeton, New Jersey 08541. Retrieved from <https://eric.ed.gov/?id=ed157930>
- Zinn, B. (2014). Technische Allgemeinbildung - Bedeutungsspektrum, Bildungsstandards und Forschungsperspektiven. *Journal of Technical Education (JOTED)*, 2(2). <https://doi.org/10.48513/JOTED.V2I2.37>
- Zinn, B. (2018a). Technikdidaktik in der Allgemeinbildung. In B. Zinn, R. Tenberg & D. Pittich (eds.), *Technikdidaktik. Eine interdisziplinäre Bestandsaufnahme* (S. 63–69). Stuttgart: Franz Steiner Verlag.
- Zinn, B. (2018b). Technischer Unterricht. In B. Zinn, R. Tenberg & D. Pittich (eds.), *Technikdidaktik. Eine interdisziplinäre Bestandsaufnahme* (S. 115–122). Stuttgart: Franz Steiner Verlag.
- Zinn, B. (2018c). Technisches Lernen am Gymnasium. In B. Zinn, R. Tenberg & D. Pittich (eds.), *Technikdidaktik. Eine interdisziplinäre Bestandsaufnahme* (S. 231 - 238). Stuttgart: Franz Steiner Verlag.
- Zinn, B. & Latzel, M. (2017). *Abschlussbericht zum Projekt "Evaluation des Schulversuchs NwT-K2"*. Institut für Erziehungswissenschaft der Universität Stuttgart: Abteilung Berufspädagogik mit Schwerpunkt Technikdidaktik (BPT).
- Zinn, B., Latzel, M. & Ariali, S. (2017). Entwicklung und Erprobung eines Instruments zur Erfassung technischen Wissens im Fach Naturwissenschaft und Technik. *Journal of Technical Education (JOTED)*, 5(1), 76–99.
- Zinn, B., Tenberg, R. & Pittich, D. (Eds.). (2018). *Technikdidaktik. Eine interdisziplinäre Bestandsaufnahme*. Stuttgart: Franz Steiner Verlag.

## Referierte Publikationen

Vereinzelte Textpassagen dieser Arbeit wurden vorab veröffentlicht. In wörtlicher oder sinngemäßer Form eingebettete Abschnitte sind der nachfolgenden Publikation entnommen. An den entsprechenden Stellen finden sich Verweise.

Brändle, M. (2023). Das Leistungsfach Naturwissenschaft und Technik (NwT) in Baden-Württemberg. In M. Binder & C. Wiesmüller (Hrsg.), *Technikunterricht - konkret*. (S. 88–111). Reutlingen, 23. - 24. September 2022. DGTB. Offenbach a. M.: BE.ER Konzept.

## Anhänge

### Anhang 1: Kategoriensystem der Interviewstudie zu inhaltlichen und kontextuellen Rahmenbedingungen

**Tabelle 28:** Kategoriensystem der Interviewstudie.

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Motivation/Be- weggründe	Extrinsisch		Angelehnt an Decy & Ryan, 1993	„Und dann passt halt auch die Kursstufe oben dann sehr gut zu dem bisherigen Konzept und die Möglichkeiten sind auch da, sie zu machen.“ → extrinsische Motivation in der Ausprägung <i>Internalisierung von Zielen</i> .
	Intrinsisch		Angelehnt an Decy & Ryan, 1993	„Ich persönlich habe Spaß an dem Fach.“ → intrinsische Motivation in der Ausprägung <i>Interne Prozessmotivation</i>

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Motivation/Beweggründe	Besondere Fächerkombination erforderlich/sinnvoll		Eine besondere Fächerkombination ist/war Voraussetzung, um als Lehrperson den Lehrauftrag in der Kursstufe NwT im vierstündigen Kurs zu übernehmen. In Form einer Ja/Nein-Frage formuliert beschreibt diese Unterkategorie die bejahende Antwort. Außerdem kann diese Antwort Codes enthalten, die auf ein wünschenswertes Fach abzielen.	„War bei Ihnen an der Schule eine besondere Fächerkombination notwendig, dass man NwT unterrichten darf in der Oberstufe?“ – Ja „Aber es ist auf jeden Fall zielführend, wenn eine Person Physik studiert hat.“
	Besondere Fächerkombination nicht erforderlich		Eine besondere Fächerkombination ist/war nicht Voraussetzung, um als Lehrperson den Lehrauftrag in der Kursstufe NwT im vierstündigen Kurs zu übernehmen. In Form einer Ja/Nein-Frage formuliert beschreibt diese Unterkategorie die verneinende Antwort.	„War bei Ihnen an der Schule eine besondere Fächerkombination notwendig, dass man NwT unterrichten darf in der Oberstufe?“ - Nein

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Ausstattung der Schule und verfügbare Ressourcen	Optimierungsbedarf		Die Befragten sehen Optimierungsbedarfe in unterschiedlichen Bereichen, die sie zur Umsetzung eines modernen Unterrichts auf hohem wissenschaftlichem und didaktischem Niveau benötigen.	„[...] das ist die Personalsituation.“
		Internet/Netzwerkinfrastruktur	Deckt netzwerkspezifische Optimierungsbedarfe ab. Dazu gehört: <ul style="list-style-type: none"> <li>• Zugriffsmöglichkeit auf das Internet</li> <li>• Stabilität der vorhandenen Netzwerkinfrastruktur</li> <li>• Wartungsintervalle von Servern</li> <li>• Netzwerksicherheit</li> </ul>	„[...] das Internet ist halt schlecht, weil wir da einen Flaschenhals beim Schuleingang haben.“
	Funktionalität der Ausstattung	Eine entsprechende Ausstattung ist zwar vorhanden, aber nicht oder nur teilweise einsatzbereit.	„Die Laptops sind noch nicht ganz stabil, da sind wir grad dran, aber da würde ich mir wünschen, dass das immer 100 % funktioniert.“  „Es gehen sehr viele Geräte kaputt, weil sie nicht sachgemäß gepflegt werden.“	

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel	
Ausstattung der Schule und verfügbare Ressourcen	Räumliche Situation		Beinhaltet Aussagen, die die räumliche Situation der Schule und die verfügbaren Räumlichkeiten für den NwT-Unterricht beschreiben. Dazu gehören auch Geräte und Einrichtungsgegenstände, die fest mit dem Raum verbunden sind (z. B. Beamer, Gas-/Wasseranschlüsse).	„Zwei Räume und ein Vorbereitungsraum und so ein Schüler-Projekt-raum.“	
	Budget		Beinhaltet Aussagen über das für den NwT-Unterricht verfügbare Budget.	„Im Jahr 6000 €.“	
	Werkstoffe		Definiert Werkstoffe, die für den NwT-Unterricht an der Schule zur Verfügung stehen und/oder verwendet werden.	„[...] viele Materialien wie Widerstände, Schrauben usw.“	
	Geräte/Maschinen	Mikrokontroller		Beschreibt die für den Unterricht verfügbaren Geräte und Maschinen. Die Facetten stellen eine Auswahl wichtiger Geräte für den Unterricht bereit. Wird bei Bedarf erweitert.	„Also Tischbohrmaschinen haben wir drei.“
		Laptop			
		Handwerkzeuge			
		Kappsäge			
		Kreissäge			
		Schleifmaschine			
		CNC-Fräs- oder Schneidemaschine			
3D-Drucker					
Dekupiersäge					
Standbohrmaschine					

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Ausstattung der Schule und verfügbare Ressourcen	Zusätzliche Ressourcen gewünscht		Zusätzliche Ressourcen werden benötigt, die für den Unterricht wichtig sind.	„[...] so insgesamt wäre es sicherlich hilfreich, wir hätten nochmal zwei bis drei 3D-Drucker.“
	Zusätzliche räumliche Möglichkeiten gewünscht		Zusätzliche Räume werden benötigt, um den Unterricht bestmöglich zu realisieren.	„Die Raumsituation ist prekär, weil wir keine Räume haben, die man jetzt nutzen kann.“
Qualifikation der Lehrpersonen	Aufbauende fachwissenschaftliche Ausbildung	Weiterbildungen	Fachwissenschaftliche Weiterbildungen, die für den NwT-Unterricht allgemein besucht wurden.	„[...] immer mal wieder auf Fortbildungen gegangen, wenn es um Solarzellen ging.“
		Zusatzausbildung im Referendariat	Die Lehrperson gibt an, dass Sie während des Referendariats an der Zusatzausbildung teilgenommen hat.	„Ich habe die Zusatzausbildung aus dem Referendariat.“
		Autodidaktisches Aneignen	Die Lehrperson gibt an, dass Sie sich die fachwissenschaftlichen Inhalte autodidaktisch angeeignet hat.	„Ich habe mir sehr viel in Zusammenarbeit mit meinen Kollegen hier in der Schule über die verschiedenen Fächer erarbeitet“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Qualifikation der Lehrpersonen		Industrieerfahrung/Handwerkliche Ausbildung	Die Lehrperson gibt an, dass sie einen Erfahrungsschatz aus der Industrie/aus dem Handwerk verfügt, der aus einer Ausbildung oder Beschäftigung stammt.	„Da war ich im Anschluss für zehn Jahre in der Industrie tätig als wissenschaftlicher Mitarbeiter bei AEG“
	Zusatzqualifikationen kursstufenspezifisch	Zeitlicher Umfang Zusatzqualifikationen	Erfasst den zeitlichen Aufwand der jeweils besuchten Weiterbildungsveranstaltungen	„eine pro Jahr“ o. Ä.
			Fachwissenschaftliche Weiterbildungen, die speziell für den Unterricht in der Kursstufe vierstündig besucht wurden.	„in einem Jahr 8-10 Tage. Machen wir 10“
Kompetenzerwerb der SuS	Keine konkrete Bezugnahme auf EPA-Technik		Während des Beantwortens der Fragestellung wird von der befragten Person nicht wie gefordert auf die EPA-Technik Bezug genommen, sondern auf den Bildungsplan o. Ä.	I: „Genau, dann kommen wir jetzt zu den Methodenkompetenzen. Also können Sie kurz skizzieren, wie der NwT-Unterricht den Methodenkompetenzerwerb gestaltet?“ B: „Ich denke jetzt an quasi so prozessbezogene Kompetenzen, wahrscheinlich bezeichnet es das eher als solches.“



Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Kompetenzerwerb der SuS	Nachweis: Methodenkompetenzerwerb		Beschreibung von Möglichkeiten zum Nachweis des Methodenkompetenzerwerbs der SuS im Unterricht der vierstündigen Kursstufe.	„Also sie [SuS] sollen halt eine technische Skizze machen z.B. und dann die Funktionsstruktur darstellen und dann einen Schaltplan, ein Flussdiagramm oder vielleicht einen Code.“
	Nachweis: Fachkompetenzerwerb		Beschreibung von Möglichkeiten zum Nachweis des Fachkompetenzerwerbs der SuS im Unterricht der vierstündigen Kursstufe.	„Was einem als erstes immer einfällt ist die Klausur.“
	Methodenkompetenzerwerb		Beschreibung von Möglichkeiten zur Schulung von Methodenkompetenzen bei den SuS.	„Ja und eben da wir projektartig arbeiten, werden halt diese Kompetenzen gefördert.“
	Fachkompetenzerwerb		Beschreibung von Möglichkeiten zur Schulung von Fachkompetenzen bei den SuS.	I: „Also Sie machen dann auch klassisch eine Qualifizierungsphase in dem Fall?“  B: „Genau.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Prüfungen	Gewichtung		Darstellung der Gewichtung schriftlicher/mündlicher/projektspezifischer Benotung.	„Meistens 50:50.“
	Arten von Prüfungen	Fachpraktische Arbeit	Praktische Arbeit (Projekt) mit Präsentation, die wie eine Klausur gewichtet wird.	„Früher hätte man gesagt eine GFS. Aber der kann die auch praktisch machen.“
		Protokoll/Laborjournal	Die Lehrperson lässt über den Unterrichts- und Versuchsablauf Protokoll führen. Wichtige Arbeitsschritte und Versuchsdurchführungen werden festgehalten und in einer propädeutischen Protokollform verschriftlicht.	„Dementsprechend spielt auch dieses Heft für die Notenfindung eine nicht unerhebliche Rolle.“
		Klausur	Klausur als Prüfungsformat zur Abfrage verschiedener fachbezogener Kompetenzen.	„Also Klausuren auf jeden Fall.“
	Projektdokumentation mit Präsentation	Projektdokumentation als Prüfungsart. Die SuS präsentieren darüber hinaus optional ihre Ergebnisse im Rahmen von Schulveranstaltungen oder als Prüfungsleistung.	„Und dann eben Projektdokumentation.“  „[...] die eigentliche Prüfung ist aber, dass wir einen NwT-Tag haben, da müssen sie vor der versammelten Schule die Projekte präsentieren.“	

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Prüfungen	Aufgabenarten für praktische Prüfung	Forschungsarbeit/Studie	Anlage einer praktischen Prüfung als Forschungsauftrag oder als Studie mit vollumfänglicher Dokumentation.	„[...] gute Themen finden, an denen die SuS in entsprechendem Rahmen forschen können.“
		Projektauftrag/Lastenheft/Pflichtenheft	Projektdurchführung mit festgelegten Kriterien, die das Produkt erfüllen muss.	I: „Welche Aufgabenarten werden Sie für die praktischen Prüfungen heranziehen? Also da kann z. B. ein Projektauftrag oder sowas gefordert sein?“ B: „[...] also einmal wenn es z.B. um ein Produkt geht, dann bietet sich das auf jeden Fall an.“
	Aufgabenarten für schriftliche Prüfung		Vom Interviewten genannte Aufgabenarten, die sich für eine schriftliche Prüfung eignen.	„[...] auch Aufgaben sein mit höherem Anforderungsgrad, um einen Transfer leisten zu müssen.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Kompetenzteilbereiche	Bereiten auf NwT-Unterricht in der Kursstufe vor		Kompetenzteilbereiche der Mittelstufe in NwT, die explizit auf den Unterricht im vierstündigen NwT-Unterricht der Kursstufe vorbereiten. Es sind dabei Kompetenzteilbereiche zu kategorisieren, die explizit im Kursstufenbildungsplan auftauchen und von der Lehrperson auf entsprechende Nachfrage genannt werden.	„Elektronik, da haben sie schon was gemacht.“
		Curriculare Anbindung	Anbindung vorhanden	Eine spiralcurriculare Anbindung von Inhalten aus der Mittelstufe ist vorhanden. Einzuordnen sind in diese Kategorie alle genannten Kompetenzteilbereiche, die in der Kursstufe nach dem Bildungsplan eine Anbindung erhalten.
		Anbindung nicht vorhanden	Eine spiralcurriculare Anbindung von Inhalten aus der Mittelstufe ist nicht vorhanden.	„zum Schuljahresende haben wir noch ein bisschen Medizintechnik drin.“
	Technische Kompetenzteilbereiche der Mittelstufe	Pneumatik	Lehrpersonen stellen einen Bezug zum Themengebiet Pneumatik her.	-
		Fliegen	Bezugnahme auf den Kompetenzteilbereich „Fliegen“, bei dem die Flügelprofile und Antrieb kennengelernt werden.	-

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Kompetenzteilbereiche	Mittelstufe (MKJS, 2016)	Energie	Kompetenzteilbereiche, die sich generell um die Energiethematik drehen. Dazu gehören erneuerbare Energien, Wirkungsgrad, Energieversorgung u. Ä.	-
		Bionik	Bionik als Themengebiet wird genannt.	„[...] , dass wir auf jeden Fall das Thema "Bionik" drin haben möchten.“
		Messtechnik/Datenmanagement	Bezugnahme auf Methoden der Messtechnik und des Datenmanagements. Dazu gehören bspw. das Aufnehmen von Messreihen und das Erstellen von Diagrammen.	„Wir haben auf jeden Fall Messen, Steuern und Regeln mit dem Mikrocontroller in Klasse 9.“
		Werkstoffe	SuS lernen Eigenschaften und Umgang mit verschiedenen Werkstoffen kennen.	-
		Regelungstechnik/Automatisierungstechnik	Lehrpersonen stellen einen Bezug zum Thema „Regelungstechnik“ her.	„[...] dann dementsprechend den Motor zu regeln.“
		Technikethik	Lehrpersonen stellen einen Bezug zur Technikethik her.	„[...] etwas, weil der Plastikmüll so ein großes Thema ist, dass die Schüler dann so eine Maschine entwickeln, die den Müll aus dem Meer holt.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Kompetenzteilbereiche	Mittelstufe (MKJS, 2016)	Grundlagen Elektronik/Sensorik	Aufbau grundlegender Schaltungen mit elektrischen Bauteilen.	„Vielleicht kommt da dieser Trübungssensor, das Photometer, zum Beispiel zum Einsatz.“
		Wirkungsgrad	Die SuS erfassen das Verhältnis von aufgewandter zu nutzbarer Energie.	„In neun haben wir grad die Windpumpe.“
		Getriebe	Die SuS haben Kenntnisse über Getriebearten und die damit zusammenhängenden Bewegungsgrößen wie Drehmoment, Kraft und Drehzahl.	„[...] ein bisschen so leicht in Richtung Getriebe gehend.“
		Konstruktion (Zeichnen/CAD)	Die SuS haben sich mit technischem Zeichnen und computergestütztem Programmieren beschäftigt und sind in der Lage, selbst (einfache) Bauteile in einem CAD-Programm zu konstruieren.	„[...] im anderen Jahrgang die Seifenblasenmaschine als ein Projekt.“
		Mikrocontroller	Lehrpersonen stellen einen Bezug zur Verwendung des Mikrocontrollers her.	„Mikrocontroller in Klasse 8.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
	Kurstufe (Autorengruppe Bildungsplan NwT, 2019b)		Umfasst alle Kompetenzteilbereiche, die von der interviewten Person in der vierstündigen Kursstufe unterrichtet werden.	„Also ich habe jetzt die <i>Technische Mechanik</i> und die <i>Produktentwicklung</i> unterrichtet und werde als nächstes <i>Elektronik, Messtechnik, Regelungstechnik</i> und <i>Datenkommunikation</i> unterrichten.“
Unterrichtsinhalte K4	Anteile praktischer Art		Beschreibt den Anteil der praktischen Unterrichtsinhalte des Unterrichts in NwT K4. Die interviewte Person soll gewichten und nennt ein Verhältnis.	„50:50“
	Anteile theoretischer Art		Beschreibt den Anteil der theoretischen Unterrichtsinhalte des Unterrichts in NwT K4. Die interviewte Person soll gewichten und nennt ein Verhältnis.	„50:50“
	Exkursionen		Die befragte Person äußert sich zur Thematik Exkursionen. Dabei werden alle Äußerungen berücksichtigt, selbst wenn die Person aussagt, dass keine Exkursionen geplant sind.	„Es sind aktuell noch keine angedacht.“ „Jetzt in der Kursstufe I fahren wir nach Hannover zur Industriemesse.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel	
Unterrichtsinhalte K4		Thema der Exkursion	Beinhaltet Aussagen über Exkursionen im Rahmen des vierstündigen Unterrichts in der Kursstufe und beschreibt, wenn möglich, deren thematische Ausrichtung.	-	
	Fachliche Unterrichtsinhalte der Kompetenzbereiche Oberstufe		Fachliche Unterrichtsinhalte, die im Rahmen des Unterrichts in NwT K4 zur Schulung der Kompetenzen herangezogen werden.	-	
	Ingenieurwissenschaftliche Bezüge	Informationstechnik	Die Unterrichtsinhalte weisen diverse ingenieurwissenschaftliche Bezüge auf. Eine Einteilung erfolgt über vier repräsentative technische Themenfelder.		„Jetzt kommt halt eben das, was die Elektrotechniker und die Informationstechniker letztendlich nachher machen werden.“
		Elektrotechnik			„In der Produktentwicklung ist es natürlich stark maschinenbauorientiert.“
		Maschinenbau			„[...] das ist die <i>Technische Mechanik</i> , das machen die Bauingenieure und die brauchen <i>Technische Mechanik</i> .“
	Bauingenieurwesen				



Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Unterrichtsinhalte K4	Projekte zu den einzelnen Kompetenzteilbereichen		Projekte, die von den SuS im Rahmen der Kompetenzteilbereiche der Kursstufe umgesetzt werden sollen.	„[...] über das Smart-Home [...].“
		Arbeitsteilige Durchführung	Die SuS arbeiten innerhalb der Projektdurchführung arbeitsteilig an einem Teilaspekt des Gesamtprojekts. Am Ende der Projektphase werden die einzelnen Komponenten zu einem Gesamtprodukt kombiniert.	„Die SuS differenzieren innerhalb des Projekts.“
		Arbeitsgleiche Durchführung	Alle SuS innerhalb des Projektteams arbeiten gemeinsam die gleichen Teilaspekte aus und jedes Teammitglied durchläuft alle Arbeitsschritte.	„Die SuS sollen alle dasselbe arbeiten.“
		Durchschnittliche Dauer	Beschreibt die durchschnittlich angesetzte oder wünschenswerte Zeitdauer für eine erfolgreiche Projektdurchführung.	„[...] ein Vierteljahr Projektphase.“
Unterrichtsmethodik	Differenzierung und Heterogenität		Die befragte Person nimmt Bezug auf ein unterschiedliches Niveau, das von Schüler:innen erreicht werden kann oder spricht explizit von einer differenzierten Aufgabenstellung.	„Drücken können sie sich nicht, aber sie erreichen unterschiedlichen Tiefgang.“
	Angewandte Unterrichtsmethoden	Referat	Die SuS halten im Unterricht Referate um ihren Mitschülerinnen und Mitschülern neue Themen und Unterrichtsinhalte zu vermitteln.	„Die SuS halten Referate.“
		Selbstlernphasen	SuS erarbeiten sich durch Material der Lehrperson einen thematischen Sachverhalt eigenständig.	„Oder auch mit einzelnen zu bearbeitenden Arbeitsaufträgen.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Unterrichtsmethodik	Angewandte Unterrichtsmethoden	Partnerarbeit/Gruppenarbeit	Die SuS erarbeiten sich mit einem Partner oder in Gruppen gemeinsam einen fachlichen Unterrichtsinhalt.	„[...] in der Projektphase arbeiten die SuS ja eh in Gruppen, meistens Vierergruppen, Dreiergruppen.“
		Rechercheauftrag	Die SuS erhalten in einem Arbeitsauftrag die Anweisung selbstständig und unter Verwendung von geeigneten Quellen Sachwissen zu recherchieren.	„[...] Rechercheaufträge, bei denen sich die SuS selbst was erarbeiten können.“
		Praktikum	Durch praktische Herangehensweise können SuS zuvor erlangtes theoretisches Wissen durch Anwendung verifizieren und Methodenkompetenzen erlernen.	„[...] dann Praktika.“
		Projekt	Das Projekt vereint die fachwissenschaftliche Qualifizierung mit der praktischen Anwendung. Die SuS sind in der Lage, ausgehend von ihrem Sachwissen, ein Produkt zu planen, zu entwickeln und zu konstruieren. Im schulischen Kontext gilt das Projekt als Unterrichtsmethode.	„[...] Projekt als Methode.“
		Lehrer-Schüler-Gespräch	Im Lehrer-Schüler-Gespräch erhalten die SuS in der Qualifizierungsphase fachwissenschaftlichen Input von der Lehrperson.	„Also Schüler-Lehrer-Gespräch auch mit Power-Point-Vorträgen.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Unterrichtsmethodik	Angewandte Unterrichtsmethoden	Lehrervortrag	Lehrpersonen nennen den Lehrervortrag als Unterrichtsmethode.	„[...] momentan lief sehr viel im Frontalunterricht mit Lehrervortrag.“
Unterrichtsorganisation	Alleinige Organisation		Die Lehrperson organisiert und hält ihren Unterricht im Alleingang ohne Unterstützung von Kollegen.	„[...], dass halt jetzt mein Partner in dem Team an eine andere Schule ist, mach ich's grad allein.“
	Lehrerwechsel pro Halbjahr		Pro Halbjahr der Kursstufe findet ein Lehrerwechsel statt. Dieser kann themenspezifische, aber auch organisatorische Gründe haben.	
	Lehrerwechsel pro Schuljahr		Pro Schuljahr der Kursstufe findet ein Lehrerwechsel statt. Dieser kann themenspezifische, aber auch organisatorische Gründe haben.	
	Teamlösung		Mehrere Lehrpersonen unterrichten einen vierstündigen Kurs im Team. Dabei sind jede Unterrichtsstunde mehrere betreuende Lehrpersonen im Klassenraum anwesend.	„[...] wir haben uns dann für eine Tandemlösung entschieden.“
Unterrichtsvorbereitung	Konzeption erfolgt allein		Die Konzeption von Unterrichtsinhalten und -materialien wird von einer Lehrperson allein vorgenommen.	„Im Wesentlichen erstmal jeder für sich selbst.“
	Konzeption erfolgt zusammen mit Kollegen schulintern		Die Konzeption von Unterrichtsinhalten und -materialien erfolgt zusammen mit anderen (im vierstündigen Kurs unterrichtenden) Lehrpersonen aus dem schulinternen Kollegium.	„[...], da wir ja jetzt im Tandem unterrichten machen wir die Konzeption natürlich im Team.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Unterrichtsvorbereitung	Konzeption erfolgt zusammen mit Kollegen schulübergreifend		Die Konzeption von Unterrichtsinhalten und -materialien erfolgt zusammen mit anderen (im vierstündigen Kurs unterrichtenden) Lehrpersonen aus einem schulübergreifenden Kollegium oder in enger Zusammenarbeit mit der Tandemschule.	I: „Sie haben es ja grade schon kurz angesprochen, die Konzeption der Unterrichtseinheiten erfolgt dann mit anderen Versuchsschulen zusammen in diesem Fall?“ B: „Ja.“
	Quellen für Konzeption UE	Expertenwissen	Expertenwissen, das z. B. durch ein Studium erlangt wird. Kann auch von einer externen Stelle (z. B. Beratung durch Ethiker:in) herangezogen werden.	„Also ich bin ja Bildungsplanmitglied [gemeint: Mitglied der Bildungsplankommission], habe von daher eine Quelle.“
		Zurückgreifen auf Unterrichtsinhalte K2	Als Quelle für die Konzeption der Unterrichtseinheiten wird genannt, dass auf Inhalte aus der zweistündigen Kursstufe zurückgegriffen wird.	„Da hatten wir natürlich einiges schon im Vorfeld oder viele Jahre vorher schon erprobt im Zweistündigen und schöpfen hauptsächlich daraus.“

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Unterrichtsvorbereitung	Quellen für Konzeption UE	Austausch zwischen Versuchsschulen	Als Wissensquelle für die Unterrichtsvorbereitung wird der Austausch zwischen den verschiedenen Versuchsschulen im vierstündigen Kursstufenversuch herangezogen.	„[...] ich kooperiere stark mit dem Tandem, wir haben ja dasselbe Thema am Anfang bearbeitet.“
		Weiterbildungsinhalte	Weiterbildungsinhalte, die für die Kursstufe relevant sind, werden als Wissensquelle herangezogen.	„Als Hilfe dient uns da die Fortbildungsreihe.“
		Studieninhalte	Inhalte/Skripte aus dem Fach- oder Aufbaustudium NwT werden für die Konzeption herangezogen.	„Und ich habe noch das Skript von der Uni.“
		Fachbücher/Wissenschaftliche Artikel	Spezielle themenspezifische Fachliteratur wird für die Konzeption der Inhalte herangezogen.	„Wir nutzen tatsächlich Literatur für die Uni bzw. Fachliteratur fürs Ingenieurstudium.“
		Internetrecherche	Informationen und fachwissenschaftliche Inhalte werden über eine Internetrecherche zusammengetragen und dienen als Grundlage der Konzeption. Recherchierte Quellen können Videos, Artikel oder fachwissenschaftliche Beiträge sein.	„[...] viel Internetrecherchen, wenn man dann das genau vorbereitet.“
	Vorgehensweise Konzeption von UE		Beschreibt die Vorgehensweise der Lehrperson(en) bei der Konzeption von Unterrichtselementen.	-

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Unterstützung	Gute Unterstützung	Arbeitsgruppe NwT	Die befragte Person fühlt sich von einer oder mehrerer dieser Parteien gut in der Umsetzung des Schulversuchs NwT K4 unterstützt.	„Also von [...] werde ich unterstützt.“
		Kollegium schulintern		
		Kollegium schulübergreifend		
		Kultusministerium		
		Schulleitung		
	Schlechte Unterstützung	Arbeitsgruppe NwT	Die befragte Person fühlt sich von einer oder mehrerer dieser Parteien schlecht in der Umsetzung des Schulversuchs NwT K4 unterstützt.	„Von [...] weiß ich gar nicht, ob es da eine Unterstützung gibt.“
		Kollegium schulintern		
		Kollegium schulübergreifend		
		Kultusministerium		
Schulleitung				
Unterstützungsbedarf		Die befragte Person nennt Unterstützungsbedarfe, die ihr die Umsetzung des Schulversuchs erleichtern würden.	„[...], wenn man sich in dieser Gruppe tatsächlich noch öfters trifft und dann gezielt fortbildet. Das wäre schon gut.“	

Kategorie	Unterkategorie	Ausprägungen/Facetten	Definition	Ankerbeispiel
Weitere Anmerkungen	Megatrends der Zukunft/Gegenwart	Bio-/Gentechnologie	Die befragte Person nimmt Bezug auf die Megatrends der Gegenwart/Zukunft und nennt bspw. Themenvorschläge oder Überbegriffe, die für den Bildungsplan relevant sein sollten.	„[...] sei es Digitalisierung, die aus meiner Sicht eben auch medial ganz zentral vertreten ist, die ja auch über die reine Technik hinaus enorme gesellschaftliche Reichweite hat.“
		Klimawandel/Nachhaltigkeit/Green Economy		
		Digitalisierung/Big Data/Datensicherheit		
		Automatisierung und Robotik		
		Urbanisierung		
Weitere Anmerkungen	Kritik an zeitlichen Faktoren		Die befragte Person übt Kritik an zeitlichen Faktoren, die sich auf den Arbeitsaufwand und/oder die Unterrichtsorganisation oder -durchführung beziehen.	„Der zeitliche Aufwand für diesen vierstündigen Kurs ist natürlich schon brutal.“
	Kritik an Deputatsangelegenheiten		Die befragte Person kritisiert die zur Verfügung gestellten Deputatsstunden.	„Also ich glaube, das wäre schon hart, wenn es jetzt gar nichts dafür gäbe. Ob wir dann trotz aller Motivation im nächsten Schuljahr weiter machen würden.“
Allgemeine Daten	Geschlecht der Lehrperson	männlich	Einteilung des genannten Geschlechts der Lehrperson auf dem Kurzfragebogen.	-
		weiblich		-
	Unterrichtseintritt NwT		Angabe der Lehrperson, seit welchem Jahr sie NwT unterrichtet.	-

<b>Kategorie</b>	<b>Unterkategorie</b>	<b>Ausprägungen/Facetten</b>	<b>Definition</b>	<b>Ankerbeispiel</b>
Allgemeine Daten	Fächerkombination		Angabe der studierten Fächerkombination der Lehrperson.	-
	Anzahl der Kurse		Anzahl der betreuten Kurse in der NwT-Kurstufe vierstündig.	-
	Anzahl der SuS	männlich	Angabe der Gesamtanzahl von SuS in den vierstündigen Kursen und deren Geschlechterverhältnis.	-
		weiblich		-



## Anhang 2: Codierungsrichtlinie der qualitativen Interviewauswertung

Um mehrfache Codierung zu vermeiden, wird immer die erste Nennung codiert.

**Bedingung:** Da es sich um ein Leitfadeninterview handelt, erfolgt eine Kategorisierung, sofern möglich, erst nach der entsprechenden Frage.

**(Erklärung:** Wird beispielsweise im Absatz unter der Fragestellung von Unterrichtsinhalten bereits Bezug auf Projekte genommen, ist es möglich, dass zu einem späteren Zeitpunkt nochmals explizit nach Projekten gefragt wird. → Eine Kategorisierung erfolgt erst bei der expliziten Nachfrage.)

Eine Codierung beginnt entweder am Anfang eines Satzes, nach einem „oder“, nach einem „und“ oder nach einem **Komma** im Satz. **Kommata**, „und“ und „oder“ werden nicht mitcodiert (Ausnahme: Aufzählung). Leerzeichen, Sprecherangaben und Satzzeichen am Satzende werden nicht codiert.

Wenn Gewichtungen genannt werden, wird die Gewichtung mit der entsprechenden Erläuterung codiert.

**Bsp.:** „Die Gewichtung von schriftlich zu mündlich ist 2:1“ → Codesegment ist der ganze Satz, nicht nur das Zahlenverhältnis.

Ein Codesegment endet erst, wenn der Satzteil mit der beinhaltenden Information abgeschlossen ist. Handelt es sich um mehrere Teilsätze oder eine Aufzählung, werden alle Teile in das Codesegment aufgenommen.

**Bsp.:** [...], weil der Plastikmüll so ein großes Thema ist, dass die Schüler dann so eine Maschine entwickeln, die den Müll aus dem Meer holt.

Codesegment „**Technikethik**“ über alle Informationen aus den Nebensätzen

Zwei Räume und ein Vorbereitungsraum und so ein Schüler-Projektraum.

Codesegment „**Räumliche Situation**“ deckt alle im Satz aufgezählten Elemente ab.

Ein Codesegment kann sich über mehrere Teilsätze oder Hauptsätze erstrecken, wenn sich kein Themenwechsel ergibt und der Kontext benötigt wird und wenn ein gesamter Abschnitt einem Codesegment zugeordnet werden muss.

**Bsp.:** Und in Klasse 10 machen wir gerade den Sonnenfolger, also diese nachgeführte Solarzelle. Dort geht's drum, halt mit Sensoren bisschen so die Helligkeit zu testen.

Codesegment „**Grundlagen Elektronik/Sensorik**“

Ja also eigentlich genauso wie in der Mittelstufe, wo man halt praktisch inhaltliche Dinge lernt und dann mit denen handelt, ist der Kompetenzerwerb sehr aussichtsreich. Also wie gesagt, durch das, dass man also nicht nur ein Praktikum macht und es dabei erlernt oder davon erfährt oder darüber informiert wird, wie auch immer man es bezeichnen möchte. Sondern, dass man

das dann im Projekt irgendwie einsetzt, handelnd einsetzt. Also das denke ich bringt eine andere Tiefe von Verständnis.

Codesegment „**Keine konkrete Bezugnahme auf EPA-Technik**“ trifft auf den ganzen Abschnitt zu, da die gestellte Frage nicht im gewünschten Kontext beantwortet wurde.

Begründungen für einen Sachverhalt werden entsprechend mit in das Codesegment aufgenommen.

**Bsp.:** Es hätte ne Teamlösung gegeben, aber durch das, dass halt jetzt mein Partner sozusagen in dem Team an eine andere Schule ist, mach ich's grad allein.

Codesegment „**Teamlösung**“ wird standardmäßig codiert.

Codesegment „**Alleinige Organisation**“ wird mit Begründung codiert, da der Kontext wichtig ist.

Bei Ja/Nein-Fragen wird nur die Antwort mit dem entsprechenden Code versehen.

**Bsp.:**

I: War bei Ihnen an der Schule irgendwie eine besondere Fächerkombination notwendig, dass man NwT unterrichten darf in der Oberstufe?

B: **Ne.**

Codesegment „**Besondere Fächerkombination nicht erforderlich**“ wird zugeordnet. Im gegenteiligen Fall einer Bejahung würde „**Besondere Fächerkombination erforderlich**“ zugeordnet werden.

Existiert für eine Ja/Nein-Frage noch keine Antwortkategorie, wird diese induktiv angelegt und mit einem entsprechenden Memo versehen. Wird Bezug zum Kurzfragebogen genommen, wird das entsprechende Element nur im Dokument mit den Kurzfragebögen codiert um Mehrfachnennungen zu vermeiden.

Wenn mehrere Kategorien für einen Sachverhalt in Frage kommen, wird dieser mit der Oberkategorie codiert.

**Bsp.:** [...] das ist „Technische Mechanik“

Trifft auf die Kategorien „Bauingenieurwesen“ und „Maschinenbau“ unter dem Codesegment „Ingenieurwissenschaftliche Bezüge“ zu und wird deshalb mit „Ingenieurwissenschaftliche Bezüge“ codiert.

Es wird möglichst immer auf Facettenebene codiert. In bestimmten Fällen (um z. B. einen Überblick über Themen von Projekten zu erhalten) kann mit der Unterkategorie codiert werden.

Projekte zu den einzelnen Kompetenzteilbereichen → [...] über das Smart-Home fällt in diese Unterkategorie, da für die verschiedenen Themen keine Facetten angelegt werden.

**Anhang 3: Kommentiertes Aufgabenverzeichnis****Tabelle 29:** Verzeichnis der Aufgaben des Vorwissenstests mit Herkunftsnachweis.

<b>Kompetenzteilbereich</b>	<b>Item</b>	<b>Thema</b>	<b>Herkunft</b>
<b>Systeme und Prozesse</b>	SP4_3 (A 1)	EVA-Prinzip	Misic, 2018, Aufgabe 3 Mikrocontroller
	SP2_4 (A 2)	Teilsysteme und Systemgrenzen	Eigenentwicklung
	SP2_5 (A 3)	Energie-, Stoff-, Informationsströme Wohnhaus	Eigenentwicklung
<b>Energie in Natur und Technik</b>	E1A_1 (A 4)	Sommer/Winter Temperaturunterschied	Zinn et al., 2017, Aufgabe 10
	E1_2 (A 5)	Energieübertragungskette Dampfmaschine	Eigenentwicklung
	EVS1/2_6 (A 6)	Strom aus Wind und Sonne	Eigenentwicklung
	E6_11 (A 7)	Energiebedarf Agrarproduktion	Eigenentwicklung
<b>Energieversorgungssysteme</b>	EVS1_7 (A 8)	Täglicher Stromverlauf Sommer/Winter	Eigenentwicklung
	EVS2_15 (A 9)	Effizienzsteigerung Sonnenkollektor	Zinn et al., 2017, Aufgabe 15
	EVS4_14 (A 10)	Wirkungsgrad Definition/Berechnung, Dimensionierung	Zinn et al., 2017, Aufgabe 14
<b>Bewegung und Fortbewegung</b>	BF6_10 (A 11)	Getriebe/Übersetzung	Eigenentwicklung
	BF4_13 (A 12)	Hebel	Eigenentwicklung
	BF3/5_9/12 (A 13)	Bewegungswandlung und -entstehung	Eigenentwicklung

<b>Eigenschaften von Stoffen</b>	EigS2_25 (A 14)	Materialeigenschaften Brückenbau	Zinn et al., 2017, Aufgabe 25
	EigS1_1 (A 15)	Leitfähigkeit von Stoffen	Eigenentwicklung
	EigS1_2 (A 16)	Löslichkeit eines Stoffes	Eigenentwicklung
<b>Statische Prinzipien in Natur und Technik</b>	SP1_23 (A 17)	Grundkonstruktion Dreieck	Zinn et al., 2017, Aufgabe 28
	SP2_28 (A 18)	Trägerquerschnitte	Zinn et al., 2017, Aufgabe 23
	SP1_26 (A 19)	Brückenkonstruktion	Zinn et al., 2017, Aufgabe 26
<b>Produktentwicklung</b>	PE3_6 (A 20)	Ressourcenschonende Werkstoffe	Eigenentwicklung
	PE4_7 (A 21)	Trennen, Fügen, Umformen	Eigenentwicklung
	PE1_5 (A 22)	Ansichten	Misic, 2018, Aufgabe 5 Technisches Zeichnen
<b>Stoffströme und Verfahren</b>	SSV2_4 (A 23)	Verfahren in der Automatisierungstechnik	Zinn et al., 2017, Aufgabe 4
	SSV1_24 (A 24)	Kohlenstoffkreislauf	Eigenentwicklung
	SSV1_7 (A 25)	Windsystem	Zinn et al., 2017, Aufgabe 7
<b>Informationsaufnahme durch Sinne und Sensoren</b>	IA1_31 (A 26)	Sensorik bei Robotern	Zinn et al., 2017, Aufgabe 31
	IA4_4 (A 27)	Subjektives Erleben und Intensität des phys. Reizes	Eigenentwicklung
	IA5_5 (A 28)	Erweiterung menschlicher Sinne durch Sensoren	Adaptiert nach Zinn et al. (2017), Aufgabe 9

<b>Gewinnung und Auswertung von Daten</b>	GAD1_8/10 (A 29)	Reproduzierbarkeit einer Messung	Eigenentwicklung
	GAD1_7 (A 30)	Messfehler	Eigenentwicklung
	GAD4_19 (A 31)	Schwingung/Ton	Zinn et al., 2017, Aufgabe 19
<b>Informationsverarbeitung</b>	IV1_32 (A 32)	Signal Analog, Digital	Zinn et al., 2017, Aufgabe 32
	IV3_30 (A 33)	Steuerung	Zinn et al., 2017, Aufgabe 30
	IV4_1/4 (A 34)	Regelung	Misic, 2018, Aufgabe 1+4 Regelungstechnik
<b>Elektronische Schaltungen</b>	ES3_12 (A 35)	Bauteile/Schaltkomponenten	Eigenentwicklung
	ES4_6 (A 36)	Schaltpläne LED	Misic, 2018, Aufgabe 6 Elektrotechnik
	ES1_3 + ES Basics (A 37)	Widerstand/Spannung/Leistung	Misic, 2018, Aufgabe 1 +2 Elektrotechnik; Zinn et al., 2017, Aufgabe 3
<b>Mikrocontroller</b>	MC_1 (A 38)	Programmabschnitt	Misic, 2018, Aufgabe 1 Mikrocontroller
	MC_2 (A 39)	Wissen zu Basics Mikrocontroller	Misic, 2018, Aufgabe 2 Mikrocontroller
	MC_6 (A 40)	Mikrocontroller Sketch/Fehlersuche	Misic, 2018, Aufgabe 6 Mikrocontroller

**Tabelle 30:** Verzeichnis der Aufgaben des Fachwissenstests mit Herkunftsnachweis.

<b>Kompetenzteilbereich</b>	<b>Item</b>	<b>Thema</b>	<b>Herkunft</b>
<b>Technikfolgenabschätzung</b>	TA_1	Nachhaltigkeitsbegriff	Eigenentwicklung
	TA_2	Zieldreieck der Energiewirtschaft	Eigenentwicklung
	TA_3	Wirkung von Produkten auf die Umwelt/Treibhausgase	Eigenentwicklung
	TA_4	Dreieck der Nachhaltigkeit	Eigenentwicklung
	TA_5	Partizipation an Entscheidungen von technikethischen Fragestellungen/Energieversorgung der Zukunft in Deutschland	Eigenentwicklung
	TA_6	Risikoanalyse	Eigenentwicklung
	TA_7	Technikethik/ Gründe für Energiewende	Eigenentwicklung in Zusammenarbeit mit Weisbrodt, U. & Hilgert, M., LFG Mannheim
	TA_8	NIMBY-Effekt	Eigenentwicklung
	TA_9	Ziele/Interessen von Akteuren im Themenbereich „Ausstieg aus Kernenergie“	Eigenentwicklung in Zusammenarbeit mit Weisbrodt, U. & Hilgert, M., LFG Mannheim
	TA_10	Sachanalyse	Eigenentwicklung in Zusammenarbeit mit Weisbrodt, U. & Hilgert, M., LFG Mannheim
	TA_11	Technikbewertung nach VDI 3780	Eigenentwicklung
	TA_12	Technikethische Fallanalyse	Eigenentwicklung in Zusammenarbeit mit Weisbrodt, U. & Hilgert, M., LFG Mannheim
	TA_13	Werteoktogonal	Eigenentwicklung
	TA_14	Vorrangregeln	Eigenentwicklung in Zusammenarbeit mit Weisbrodt, U. & Hilgert, M., LFG Mannheim

	TA_15	Ökobilanz/Aufbau	Eigenentwicklung
	TA_16	Etablierte Methoden zur Technikbewertung	Eigenentwicklung
	TA_17	Ökologischer Fußabdruck	Eigenentwicklung
	TA_18	Technikethische Konflikte/"Cradle to Cradle" Ansatz	Eigenentwicklung
	TA_19	Wertekonflikte „5G-Netzausbau“	Eigenentwicklung in Zusammenarbeit mit Weisbrodt, U. & Hilgert, M., LFG Mannheim
	TA_20	Technikethische Fragestellung und Spontanurteil	Eigenentwicklung
<b>Energieversorgung</b>	EV_1	Energieträger	Eigenentwicklung
	EV_2	Energiewandlung Vergleich Wärmekraftwerk vs. Photovoltaik	Eigenentwicklung
	EV_3	Überspannung	Eigenentwicklung
	EV_4	Dimensionierung Solarmodul Inselanlage	Eigenentwicklung
	EV_5	Abgasbehandlung in Wärmekraftwerken	Eigenentwicklung
	EV_6	Wirkungsgrad Energiewandlung/Kraftwerkstypen	Eigenentwicklung in Zusammenarbeit mit Weisbrodt, U. & Hilgert, M., LFG Mannheim
	EV_7	Wirkungsgrad Energiespeicher/Berechnung	Eigenentwicklung
	EV_8	Energiespeicher/Speicherprinzip	Eigenentwicklung
	EV_9	Wirkungsgrad/Steigerung des Wirkungsgrads von Wärmekraftwerken	Eigenentwicklung
	EV_10	Energiewandlung/Schadstoffe aus Primärenergieträgern	Eigenentwicklung
	EV_11	Brennstoffzelle	Eigenentwicklung
	EV_12	Definition Smart-Grid	Eigenentwicklung
	EV_13	Möglichkeiten der Energiespeicherung	Eigenentwicklung

	EV_14	Merit-Order-Effekt/Erklärung	Eigenentwicklung
	EV_15	Einflussfaktoren Höhe der EEG-Umlage	Eigenentwicklung
	EV_16	EEG-Umlage	Eigenentwicklung
	EV_17	Strompreisbildung an der Strombörse	Eigenentwicklung
	EV_18	Energiedichte/Energiespeicher	Eigenentwicklung
	EV_19	Strombedarfsdeckung/Kraftwerke	Eigenentwicklung in Zusammenarbeit mit Weisbrodt, U. & Hilgert, M., LFG Mannheim
	EV_20	Ökonomische Aspekte der Energiewende	Eigenentwicklung
<b>Elektrische Antriebs-technik</b>	EA_1	Reale Spannungsquelle	Eigenentwicklung
	EA_2	Bürstenloser Gleichstrommotor	Eigenentwicklung in Zusammenarbeit mit Bruckelt, S., Dr. Bernhard, K. & Berchtold, J., HSG Bad Wimpfen
	EA_3	Einfacher Gleichstrommotor/Aufbauskizze	Eigenentwicklung in Zusammenarbeit mit Bruckelt, S., Dr. Bernhard, K. & Berchtold, J., HSG Bad Wimpfen
	EA_4	Kennlinie Spannungsquellen/Lastwiderstand/Betriebspunkt	Eigenentwicklung in Zusammenarbeit mit Bruckelt, S., Dr. Bernhard, K. & Berchtold, J., HSG Bad Wimpfen
	EA_5	Motorstrom bei unterschiedlichen Drehmomenten	Eigenentwicklung in Zusammenarbeit mit Bruckelt, S., Dr. Bernhard, K. & Berchtold, J., HSG Bad Wimpfen
	EA_6	Erfassung Motorkennlinie/Schaltplan	Eigenentwicklung in Zusammenarbeit mit Bruckelt, S., Dr. Bernhard, K. & Berchtold, J., HSG Bad Wimpfen



EA_7	Motorbetrieb/Leistung	Eigenentwicklung in Zusammenarbeit mit Bruckelt, S., Dr. Bernhard, K. & Berchtold, J., HSG Bad Wimpfen
EA_8	Elektromotoren/elektrotechnische Grundlagen	Eigenentwicklung
EA_9	Elektromotor-Kennlinie/Betriebspunkte	Eigenentwicklung
EA_10	Elektromotor-Kennlinie/Dauerbetriebsbereich	Eigenentwicklung
EA_11	Elektromotor-Kennlinie/Start-Stopp-Betrieb DC-Motoren	Eigenentwicklung
EA_12	Elektromotor/Last-Betriebspunkte	Eigenentwicklung
EA_13	Servomotor/Aufbau	Eigenentwicklung in Zusammenarbeit mit Bruckelt, S., Dr. Bernhard, K. & Berchtold, J., HSG Bad Wimpfen
EA_14	Motorarten/Auswahl eines geeigneten Elektromotors	Eigenentwicklung
EA_15	Motorbetrieb/Notwendigkeit Motor-Shield	Eigenentwicklung in Zusammenarbeit mit Bruckelt, S., Dr. Bernhard, K. & Berchtold, J., HSG Bad Wimpfen
EA_16	Motorbetrieb/Bauteillösung induktiver Kickback	Eigenentwicklung
EA_17	Schrittmotor/Aufbau- und Funktionsprinzip	Eigenentwicklung
EA_18	Pulsweitenmodulation/PWM-Signale	Eigenentwicklung
EA_19	Vergleich Motor-Generatorbetrieb Gleichstrommaschine anhand von Ersatzschaltbildern	Eigenentwicklung
EA_20	Kennlinien Generatorbetrieb	Eigenentwicklung

<b>Technische Mechanik</b>	TM_1	Kräfteplan/resultierende Kraft	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_2	Kräftezerlegung/resultierende Kraft	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_3	Freischnitte	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_4	Spannungs-Dehnungs-Diagramm	Eigenentwicklung in Anlehnung an Läßle (2016, S. 6 f)
	TM_5	Definition E-Modul	Eigenentwicklung
	TM_6	Euler'sche Knickfälle	Eigenentwicklung
	TM_7	Hooke'sches Gesetz	Eigenentwicklung
	TM_8	Drehmoment	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_9	Einheiten TM und Festigkeitslehre	Eigenentwicklung
	TM_10	Kräfteysteme	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_11	Lager/Lagerreaktionen	Eigenentwicklung
	TM_12	Statische Bestimmtheit/Lagerung	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_13	Statische Bestimmtheit	Eigenentwicklung
	TM_14	Gleichgewicht/Drehmoment	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_15	FEM-Analyse/Doppel-T-Träger	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_16	Zweiwertige Gelenke/Kraftübertragung	Eigenentwicklung in Zusammenarbeit mit Zendler, S., IfE-BPT
	TM_17	Ebene Fachwerke/Nullstäbe	Eigenentwicklung

	TM_17 Alt: TM_18	Balken unter Biegebeanspruchung	Eigenentwicklung
	TM_18 Alt: TM_19	Werkstoff/Festigkeitsbedingungen/Zugspannung	Eigenentwicklung
	TM_19 Alt: TM_20	Werkstoffeigenschaften bei Belastung/Zug, Druck, Biegung	Eigenentwicklung
	TM_20 Alt: TM_21	Lagerungsarten	Eigenentwicklung
<b>Produktentwicklung</b>	PE_1	Grundregeln des Gestaltens	Eigenentwicklung
	PE_2	Wirtschaftliche Wertigkeit/Definition	Eigenentwicklung
	PE_3	Nutzwertanalyse	Eigenentwicklung
	PE_4	Technische und wirtschaftliche Wertigkeit/Diagramm	Eigenentwicklung
	PE_5	Analyse eines technischen Produkts/Funktionsbeschreibung	Eigenentwicklung
	PE_6	Konstruktionselemente/Hauptfunktionen	Eigenentwicklung
	PE_7	Konstruktionselemente/Zuordnung Verbindungsmöglichkeiten	Eigenentwicklung
	PE_8	Werkstoffeigenschaften/Werkstoffauswahl	Eigenentwicklung
	PE_9	Technische Systeme	Conrad, 2019, S. 448, Aufgabe 2.2 (modif.)
	PE_10	Gesamtfunktion/Black-Box	Eigenentwicklung
	PE_11	Methodisches Vorgehen Konstruieren	Eigenentwicklung
	PE_12	Flussdiagramm	Eigenentwicklung in Zusammenarbeit mit Bühler, M., Rotteck-Gymnasium Freiburg & Meyer, K., FAG Vaihingen

	PE_13	Skizzieren von technischen Lösungen/Pendelstütze Zugbrücke	Eigenentwicklung in Zusammenarbeit mit Bühler, M., Rotteck-Gymnasium Freiburg & Meyer, K., FAG Vaihingen
	PE_14	Anforderungsliste	Eigenentwicklung
	PE_15	Funktionsstruktur der Aufgabenstellung	Eigenentwicklung
	PE_16	Prototypisierung/Grundsätze	Eigenentwicklung
	PE_17	Normorientierte Konstruktionsweise/Ökonomischer Vorteil	Eigenentwicklung
	PE_18	Anforderungsliste/Kundenanforderungen und -wünsche	Eigenentwicklung
	PE_19	Fehlhandlungssichere Fertigung und Montage	Eigenentwicklung
	PE_20	Problemstellung/Ableitung Auftrag und Zielsetzung	Eigenentwicklung
<b>Grundlagen der Elektronik</b>	GE_1	Diodenkennlinie/Schaltplan entwerfen	Eigenentwicklung
	GE_2	Dioden/Allgemeine Daten	Eigenentwicklung
	GE_3	Diode/Widerstandskennlinie	Eigenentwicklung
	GE_4	NTC/PTC	Eigenentwicklung
	GE_5	Integrierte Schaltkreise/Definition	Eigenentwicklung
	GE_6	Operationsverstärker-Verwendungsmöglichkeiten	Eigenentwicklung
	GE_7	Realer vs. idealer OPV	Eigenentwicklung
	GE_8	Beschaltung OPV	Eigenentwicklung
	GE_9	Spannungsteilerschaltung	Eigenentwicklung
	GE_10	Formel für Verstärkung bei nichtinvertierender Verstärkung	Eigenentwicklung
	GE_11	Optimierung Schaltkreis/Schutz vor induktiver Last	Eigenentwicklung

	GE_12	Schaltung entwickeln/Rechts- Linkslauf	Eigenentwicklung
	GE_13	Freilaufdiode	Eigenentwicklung
	GE_14	Blockkondensator	Eigenentwicklung
	GE_15	Glättungskondensator	Eigenentwicklung
	GE_16	Technische Realisierung eines Leiterplattenentwurfs	Eigenentwicklung
	GE_17	Schaltkreis/Platinenfertigung	Eigenentwicklung
	GE_18	Kennlinie Fotowiderstand	Eigenentwicklung
	GE_19	Kennlinie eines Kaltleiters/Kurvenverlauf	Eigenentwicklung
	GE_20	OP-Datenblattanalyse	Eigenentwicklung
<b>Mess- technik</b>	MT_1	Spannungsteilerregel	Eigenentwicklung
	MT_2	Spannungsteiler bei NTC-Messschaltung	Eigenentwicklung
	MT_3	Brückenschaltung	Eigenentwicklung
	MT_4	Brückenschaltung Widerstandsbestimmung	Eigenentwicklung
	MT_5	Belasteter vs. Unbelasteter Spannungsteiler	Eigenentwicklung
	MT_6	Schaltung zur Bestimmung der Lichtintensität skizzieren	Eigenentwicklung
	MT_7	Funktion eines A/D-Wandlers	Eigenentwicklung
	MT_8	Schaltung mit PTC zur Wärmebestimmung	Eigenentwicklung
	MT_9	Invertierender Verstärker	Eigenentwicklung
	MT_10	Fotowiderstand auswählen	Eigenentwicklung
	MT_11	Brückenschaltung zeichnen	Eigenentwicklung
	MT_12	A/D-Wandler Eigenschaften	Eigenentwicklung
	MT_13	Abtastung	Eigenentwicklung
	MT_14	A/D-Wandlungsverfahren	Eigenentwicklung
	MT_15	A/D-Wandlung Blockschaltplan	Eigenentwicklung

	MT_16	Eigenschaften Spannungsfolger	Eigenentwicklung
	MT_17	Flash-Wandler	Eigenentwicklung
	MT_18	Instrumentenverstärker	Eigenentwicklung
	MT_19	Nicht invertierender Verstärker dimensionieren	Eigenentwicklung
	MT_20	Messgenauigkeit	Eigenentwicklung
	MT_21	Eigenschaften belasteter Spannungsteiler	Eigenentwicklung
	MT_22	Belasteter Spannungsteiler berechnen	Eigenentwicklung
<b>Rege- lungs- technik</b>	RT_1	Festwertregelung/Ziel	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_2	Unterschied Steuern vs. Regeln	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_3	Mechanische Regler/analoge Regler	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_4	Regelkreis/Elemente und Größen	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_5	Regelung/Sprungantwort	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_6	Sprungantwort/Definition	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_7	Steuerung vs. Regelung	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_8	Regelstrecken/Auswahl Regelstrecke	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_9	Ermittlung Sprungantwort	Eigenentwicklung

	RT_10	Regelstrecke/Fahren eines Autos unter Einfluss von Störgrößen	Misic, 2018, Aufgabe 5 Regelungstechnik
	RT_11	PI-Regler/Blockdiagramm	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_12	Zweipunktregler/Beispiel und Funktionsweise	Eigenentwicklung
	RT_13	P- und I-Regler/Vorteile vs. Nachteile	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
	RT_14	PI-Regler/Beispiel Kernreaktor	Eigenentwicklung
	RT_15	Zweipunktregelung/Diagramm + Ableitung Anwendungsmöglichkeit	Eigenentwicklung
	RT_16	Unstetige vs. Stetige Regelung	Eigenentwicklung
	RT_17	Biologische Regelvorgänge	Eigenentwicklung
	RT_18	Technische Regelvorgänge	Eigenentwicklung
	RT_19	Regelung vs. Steuerung/Vorteile einer Regelung	Eigenentwicklung in Zusammenarbeit mit Pfister, S. & Eisenmann, M., AEG Ulm
<b>Datenkommunikation</b>	DK_1	Bits/Byte	Eigenentwicklung
	DK_2	Technische Kommunikation	Eigenentwicklung in Zusammenarbeit mit Trittlter, F., Merkle, M. & Mangold, M., FSG Marbach
	DK_3	Binärkode	Eigenentwicklung
	DK_4	Erklärung „Baudrate“	Eigenentwicklung
	DK_5	OSI-3-Schicht-Modell	Eigenentwicklung
	DK_6	Netzwerktopologie Internet	Eigenentwicklung
	DK_7	Komponenten im Kommunikationsnetzwerk	Eigenentwicklung in Zusammenarbeit mit Trittlter, F., Merkle, M. & Mangold, M., FSG Marbach
	DK_8	ASCII-Codierung	Eigenentwicklung

---

DK_9	Reichweite Datenübertragungswege	Eigenentwicklung
DK_10	Signalmodulation	Eigenentwicklung
DK_11	Zusammenhang Übertragungsrate, Modulation, Störsicherheit	Eigenentwicklung
DK_12	Störfaktoren für Datenübertragung	Eigenentwicklung
DK_13	Flussdiagramm: Optisches Übertragungssystem	Eigenentwicklung
DK_14	Programmablaufplan kabelloser Chat	Eigenentwicklung
DK_15	Funktionsweise Ende-zu-Ende Verschlüsselung	Eigenentwicklung
DK_16	Vor-/Nachteile symmetrischer Verschlüsselung	Eigenentwicklung
DK_17	Arrays	Eigenentwicklung
DK_18	Bitübertragungsrate/Übertragungsfrequenz	Eigenentwicklung
DK_19	Beschreibung des Master-Slave-Prinzips	Eigenentwicklung
DK_20	Master-Slave – Code I2C Arduino	Eigenentwicklung

---



**Anhang 4: Ergebnisse der Expert:innenbefragung (Leistungsfach)**

**Tabelle 31:** Medianwerte der Expert:innenbefragung zu *Technikfolgenabschätzung* ( $n = 5$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>Ta_1</b>	2.00	5.00	5.00
<b>Ta_2</b>	2.00	3.00	4.00
<b>Ta_3</b>	4.00	5.00	4.00
<b>Ta_4</b>	1.00	4.00	4.00
<b>Ta_5</b>	5.00	6.00	6.00
<b>Ta_6</b>	4.00	4.00	3.00
<b>Ta_7</b>	3.00	5.00	5.00
<b>Ta_8</b>	3.00	5.00	6.00
<b>Ta_9</b>	5.00	5.00	5.00
<b>Ta_10</b>	2.00	5.00	5.00
<b>Ta_11</b>	2.00	3.50	2.50
<b>Ta_12</b>	3.00	5.00	5.00
<b>Ta_13</b>	4.00	5.00	5.00
<b>Ta_14</b>	3.50	5.00	5.00
<b>Ta_15</b>	3.00	2.00	1.00
<b>Ta_16</b>	2.50	2.50	2.00
<b>Ta_17</b>	5.00	2.00	1.00
<b>Ta_18</b>	6.00	2.50	1.50
<b>Ta_19</b>	4.00	5.00	5.00
<b>Ta_20</b>	6.00	5.00	5.00

**Tabelle 32:** Medianwerte der Expert:innenbefragung zu *Energieversorgung* ( $n = 5$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>Ev_1</b>	1.00	2.00	3.00
<b>Ev_2</b>	4.00	3.00	3.00
<b>Ev_3</b>	4.00	3.00	1.00
<b>Ev_4</b>	5.00	4.50	2.00
<b>Ev_5</b>	4.00	2.00	1.50
<b>Ev_6</b>	2.00	2.00	2.00
<b>Ev_7</b>	3.00	5.00	4.00
<b>Ev_8</b>	3.00	5.00	5.00
<b>Ev_9</b>	3.00	3.00	3.00
<b>Ev_10</b>	3.00	2.00	2.00
<b>Ev_11</b>	4.00	3.00	1.50
<b>Ev_12</b>	4.00	5.00	5.00
<b>Ev_13</b>	3.00	4.00	4.00
<b>Ev_14</b>	5.00	5.00	5.00
<b>Ev_15</b>	5.00	3.00	2.00
<b>Ev_16</b>	3.00	5.00	5.00
<b>Ev_17</b>	5.00	3.00	3.00
<b>Ev_18</b>	4.00	4.00	1.00
<b>Ev_19</b>	4.00	4.00	5.00
<b>Ev_20</b>	5.00	3.00	2.00

**Tabelle 33:** Medianwerte der Expert:innenbefragung zu *Elektrische Antriebstechnik* ( $n = 5$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>EA_1</b>	2.00	4.00	6.00
<b>EA_2</b>	4.00	5.00	5.00
<b>EA_3</b>	3.00	5.00	5.00
<b>EA_4</b>	2.00	6.00	5.00
<b>EA_5</b>	2.00	6.00	6.00
<b>EA_6</b>	4.00	6.00	6.00
<b>EA_7</b>	5.00	4.00	4.00
<b>EA_8</b>	5.00	5.00	5.00
<b>EA_9</b>	4.00	4.00	5.00
<b>EA_10</b>	4.00	4.50	4.50
<b>EA_11</b>	5.50	4.50	2.50
<b>EA_12</b>	5.00	5.00	5.00
<b>EA_13</b>	3.00	5.00	5.00
<b>EA_14</b>	3.00	5.00	6.00
<b>EA_15</b>	3.00	5.00	6.00
<b>EA_16</b>	5.00	4.00	3.00
<b>EA_17</b>	3.00	5.00	5.00
<b>EA_18</b>	3.00	5.00	5.00
<b>EA_19</b>	5.00	1.50	1.00
<b>EA_20</b>	5.00	1.00	1.00

**Tabelle 34:** Medianwerte der Expert:innenbefragung zu *Technische Mechanik* ( $n = 6$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>TM_1</b>	2.00	4.50	5.00
<b>TM_2</b>	2.50	5.00	5.00
<b>TM_3</b>	4.00	6.00	6.00
<b>TM_4</b>	5.00	5.00	5.00
<b>TM_5</b>	3.00	5.00	5.00
<b>TM_6</b>	5.00	5.00	4.00
<b>TM_7</b>	4.00	4.00	4.00
<b>TM_8</b>	4.00	3.00	3.00
<b>TM_9</b>	1.00	6.00	6.00
<b>TM_10</b>	4.00	4.00	4.50
<b>TM_11</b>	2.00	6.00	6.00
<b>TM_12</b>	3.50	6.00	6.00
<b>TM_13</b>	4.00	5.00	4.00
<b>TM_14</b>	5.50	5.50	5.50
<b>TM_15</b>	3.50	6.00	6.00
<b>TM_16</b>	2.00	4.00	5.00
<b>TM_17</b>	5.00	2.00	2.00
<b>TM_18</b>	4.00	5.00	5.00
<b>TM_19</b>	4.00	4.00	5.00
<b>TM_20</b>	3.00	5.50	5.50

**Tabelle 35:** Medianwerte der Expert:innenbefragung zu *Produktentwicklung* ( $n = 5$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>Pe_1</b>	2.00	6.00	6.00
<b>Pe_2</b>	4.00	4.00	4.00
<b>Pe_3</b>	3.00	4.00	4.00
<b>Pe_4</b>	3.00	5.00	5.00
<b>Pe_5</b>	3.00	3.00	3.00
<b>Pe_6</b>	3.00	5.00	5.00
<b>Pe_7</b>	3.00	4.00	4.00
<b>Pe_8</b>	5.00	3.00	3.00
<b>Pe_9</b>	3.00	5.00	4.00
<b>Pe_10</b>	4.00	5.00	5.00
<b>Pe_11</b>	3.00	5.00	5.00
<b>Pe_12</b>	3.00	6.00	6.00
<b>Pe_13</b>	5.00	5.00	5.00
<b>Pe_14</b>	2.00	3.00	2.00
<b>Pe_15</b>	3.00	3.00	3.00
<b>Pe_16</b>	2.50	4.50	4.50
<b>Pe_17</b>	3.00	5.00	5.00
<b>Pe_18</b>	4.00	4.00	4.00
<b>Pe_19</b>	5.00	6.00	6.00
<b>Pe_20</b>	5.00	5.00	5.00

**Tabelle 36:** Medianwerte der Expert:innenbefragung zu *Grundlagen der Elektronik* ( $n = 5$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>GE_1</b>	2.00	6.00	5.00
<b>GE_2</b>	3.00	4.00	4.00
<b>GE_3</b>	2.00	4.00	4.00
<b>GE_4</b>	3.00	2.00	3.00
<b>GE_5</b>	2.00	4.00	4.00
<b>GE_6</b>	5.00	2.00	1.00
<b>GE_7</b>	3.00	5.00	6.00
<b>GE_8</b>	4.50	4.50	5.50
<b>GE_9</b>	3.00	5.00	6.00
<b>GE_10</b>	3.00	6.00	6.00
<b>GE_11</b>	3.00	1.50	1.00
<b>GE_12</b>	2.00	2.00	2.00
<b>GE_13</b>	4.00	5.00	5.00
<b>GE_14</b>	3.00	3.00	3.00
<b>GE_15</b>	4.00	4.00	4.00
<b>GE_16</b>	2.00	4.00	5.50
<b>GE_17</b>	4.00	5.00	5.00
<b>GE_18</b>	4.00	5.00	4.00
<b>GE_19</b>	4.00	4.00	4.00
<b>GE_20</b>	4.00	5.00	5.00

**Tabelle 37:** Medianwerte der Expert:innenbefragung zu *Aufnahme und Verarbeitung von Signalen in der Messtechnik* ( $n = 5$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>Mt_1</b>	4.00	6.00	6.00
<b>Mt_2</b>	4.00	6.00	6.00
<b>Mt_3</b>	5.50	5.00	5.00
<b>Mt_4</b>	5.00	5.00	5.00
<b>Mt_5</b>	5.00	5.00	5.00
<b>Mt_6</b>	5.00	5.00	5.00
<b>Mt_7</b>	4.00	5.00	5.00
<b>Mt_8</b>	6.00	6.00	6.00
<b>Mt_9</b>	5.00	6.00	6.00
<b>Mt_10</b>	5.00	5.00	3.00
<b>Mt_11</b>	4.50	5.00	5.00
<b>Mt_12</b>	4.00	2.00	2.00
<b>Mt_13</b>	5.00	2.00	2.00
<b>Mt_14</b>	4.50	1.50	1.00
<b>Mt_15</b>	5.00	2.00	1.50
<b>Mt_16</b>	4.00	5.50	5.50
<b>Mt_17</b>	6.00	4.00	1.50
<b>Mt_18</b>	5.00	3.00	5.00
<b>Mt_19</b>	5.00	6.00	6.00
<b>Mt_20</b>	4.00	5.00	5.00
<b>Mt_21</b>	5.00	5.00	5.00
<b>Mt_22</b>	4.00	5.00	3.00

**Tabelle 38:** Medianwerte der Expert:innenbefragung zu *Regelungstechnik* ( $n = 5$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>Rt_1</b>	4.00	3.00	4.00
<b>Rt_2</b>	3.00	6.00	6.00
<b>Rt_3</b>	3.00	6.00	6.00
<b>Rt_4</b>	3.00	5.00	5.00
<b>Rt_5</b>	4.00	6.00	6.00
<b>Rt_6</b>	3.00	6.00	6.00
<b>Rt_7</b>	4.00	6.00	6.00
<b>Rt_8</b>	3.00	5.00	5.00
<b>Rt_9</b>	4.00	6.00	6.00
<b>Rt_10</b>	4.00	4.00	3.00
<b>Rt_11</b>	5.00	5.00	5.00
<b>Rt_12</b>	4.00	6.00	6.00
<b>Rt_13</b>	4.00	6.00	6.00
<b>Rt_14</b>	4.00	6.00	6.00
<b>Rt_15</b>	6.00	5.00	4.00
<b>Rt_16</b>	4.00	6.00	6.00
<b>Rt_17</b>	5.00	5.00	5.00
<b>Rt_18</b>	3.00	5.00	4.00
<b>Rt_19</b>	3.00	6.00	6.00
<b>Rt_20</b>	4.00	6.00	6.00



**Tabelle 39:** Medianwerte der Expert:innenbefragung zu *Datenkommunikation* ( $n = 6$ ).

<b>Item</b>	<b>Komplexität</b>	<b>Curriculare Relevanz</b>	<b>Praxisrelevanz</b>
<b>Dk_1</b>	3.00	5.00	5.00
<b>Dk_2</b>	4.00	3.50	3.00
<b>Dk_3</b>	2.50	4.50	4.50
<b>Dk_4</b>	3.50	3.00	3.00
<b>Dk_5</b>	3.00	4.50	4.00
<b>Dk_6</b>	4.00	4.50	4.00
<b>Dk_7</b>	3.50	5.00	5.00
<b>Dk_8</b>	4.00	5.00	5.50
<b>Dk_9</b>	3.50	5.00	4.50
<b>Dk_10</b>	4.50	5.50	4.50
<b>Dk_11</b>	6.00	5.00	3.50
<b>Dk_12</b>	3.50	5.00	5.00
<b>Dk_13</b>	5.00	3.00	3.00
<b>Dk_14</b>	3.00	2.00	2.50
<b>Dk_15</b>	4.50	5.00	5.50
<b>Dk_16</b>	3.00	5.00	5.00
<b>Dk_17</b>	3.50	5.00	5.00
<b>Dk_18</b>	4.00	5.00	6.00
<b>Dk_19</b>	3.50	4.50	4.50
<b>Dk_20</b>	4.50	5.00	5.00

**Anhang 5: Pilotierungsergebnisse der Fachwissenstests (Leistungsfach)**

Die Reihenfolge der Aufgaben wurde nach der Expertenbefragung neu abgestimmt, um Druckkosten und Papierverbrauch für die Pilotierung zu minimieren. Für alle weiteren Analysen gilt die neue Nummerierung. In Klammern ist die Aufgabenbezeichnung aus der Expertenbefragung zugeordnet.

**Tabelle 40:** Pilotierungsergebnisse der Skala *Technikfolgenabschätzung*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>c</sup></b>	<b>Kompetenz-facetten</b>
Ta_A1 (Ta_1)	0.65	-0.69	-	-
Ta_A2 (Ta_20)	0.54	0.22	0.21	(3)
Ta_A3 (Ta 3)	0.69	0.43	0.43	(2)
Ta_A4 (Ta 18)	0.23	0.66	0.67	(4)
Ta_A5 (Ta 5)	0.35	0.42	0.37	(1)
Ta_A6 <sup>a</sup>	0.54	0.39	0.34	(1)
Ta_A7 (Ta 7)	0.00	-	-	-
Ta_A8 (Ta 14)	0.19	-0.13	-	-
Ta_A9 (Ta 9)	0.46	0.34	0.40	(3)
Ta_A10 (Ta 12)	0.15	0.30	-	-
Ta_A11 <sup>a</sup>	0.08	0.44	-	-
Ta_A12a (Ta 19 <sup>b</sup> )	0.08	0.39	0.36	(3)
Ta_A12b (Ta 19 <sup>b</sup> )	0.39	0.53	0.45	(3)
Ta_A13 (Ta 4)	0.50	-0.58	-	-
Ta_A14 (Ta 17)	0.46	-0.16	-	-
Ta_A15 (Ta 6)	0.12	-	-	-
Ta_A16 (Ta 15)	0.00	-	-	-
Ta_A17 (Ta 8)	0.50	0.78	0.82	(3)
Ta_A18 (Ta 13)	0.19	0.25	0.31	(4)
Ta_A19 (Ta 10)	0.27	0.43	0.44	(2)
Ta_A20 <sup>a</sup>	0.39	0.46	0.49	(4)

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.

**Tabelle 41:** Pilotierungsergebnisse der Skala *Energieversorgung*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>c</sup></b>	<b>Kompetenz-facetten</b>
Ev_A1a (Ev 18 <sup>b</sup> )	0.31	-0.06	-	-
<b>Ev_A1b (Ev 18<sup>b</sup>)</b>	0.42	0.30	0.41	(2)
<b>Ev_A1c (Ev 18<sup>b</sup>)</b>	0.50	0.51	0.45	(2)
<b>Ev_A1d (Ev 18<sup>b</sup>)</b>	0.46	0.48	0.41	(2)
<b>Ev_A2<sup>a</sup></b>	0.27	0.59	0.60	(5)
Ev_A3 (Ev 3)	0.04	-	-	-
Ev_A4 <sup>a</sup>	0.23	0.08	-	-
Ev_A5 <sup>a</sup>	0.08	-	-	-
Ev_A6 (Ev 16)	0.15	-	-	-
<b>Ev_A7 (Ev 7)</b>	0.39	0.50	0.56	(2)
Ev_A8 (Ev 14)	0.04	-	-	-
<b>Ev_A9<sup>a</sup></b>	0.35	0.37	0.29	(1)
Ev_A10 (Ev 12)	0.35	-0.03	-	-
<b>Ev_A11<sup>a</sup></b>	0.31	0.14	0.29	(1)
Ev_A12 (Ev 19)	0.04	-	-	-
<b>Ev_A13a (Ev 4<sup>b</sup>)</b>	0.31	0.75	0.69	(4)
<b>Ev_A13b (Ev 4<sup>b</sup>)</b>	0.19	0.72	0.70	(4)
Ev_A14 <sup>a</sup>	0.08	-	-	-
Ev_A15 <sup>a</sup>	0.39	0.14	-	-
<b>Ev_A16<sup>a</sup></b>	0.39	0.37	0.42	(3)
<b>Ev_A17 (Ev 8)</b>	0.31	0.55	0.62	(2)
<b>Ev_A18 (Ev 13)</b>	0.23	0.67	0.51	(3)
Ev_A19 <sup>a</sup>	0.35	0.19	-	-
Ev_A20 (Ev 11)	0.08	-	-	-

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.

**Tabelle 42:** Pilotierungsergebnisse der Skala *Elektrische Antriebstechnik*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>d</sup></b>	<b>Kompetenz-facetten</b>
EA_A1 (EA_1)	0.25	-0.09	-	-
EA_A2 (EA_20)	0.06	-	-	-
<b>EA_A3 (EA_3)</b>	0.58	0.06	0.42	(1)
<b>EA_A4 (EA_18)</b>	0.22	0.28	0.70	(5)
EA_A5a (EA_A5 <sup>b</sup> )	0.11	-	-	-
EA_A5b (EA_A5 <sup>b</sup> )	0.17	-	-	-
EA_A6 (EA_6)	0.06	-	-	-
EA_A7 (EA_7)	0.06	-	-	-
EA_A8 (EA_14)	0.17	-	-	-
<b>EA_A9 (EA_9)</b>	0.31	0.42	0.37	-
<b>EA_A10 (EA_12)</b>	0.25	0.29	0.40	(1)
<b>EA_A11 (EA_2)</b>	0.75	0.15	0.19	(1)
<b>EA_A12 (EA_19)<sup>b</sup></b>	0.08	-	-	(3)
<b>EA_A13a (EA_4<sup>b</sup>)<sup>c</sup></b>	0.56	-0.03	-	(2)
<b>EA_A13b (EA_4<sup>b</sup>)<sup>c</sup></b>	0.33	0.21	-	(2)
<b>EA_A14 (EA_17)</b>	0.39	0.29	0.45	(1)
<b>EA_A15 (EA_6)</b>	0.33	0.71	0.40	(2)
EA_A16 (EA_15)	0.19	0.47	-	
<b>EA_A17 (EA_8)</b>	0.72	0.29	0.78	(1)
<b>EA_A18a (EA_13<sup>b</sup>)</b>	0.58	0.48	0.36	(1)
<b>EA_A18b (EA_13<sup>b</sup>)</b>	0.64	0.48	0.40	(1)
EA_A19 (EA_19)	0.64	0.06	-	-
EA_A20 (EA_20)	0.08	-	-	-

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> In Skala behalten aufgrund inhaltlicher Abdeckung der Kompetenzteilbereiche; <sup>d</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.

**Tabelle 43:** Pilotierungsergebnisse der Skala *Technische Mechanik*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>d</sup></b>	<b>Kompetenz-facetten</b>
TM_A1 (TM_1)	0.04	-	-	-
TM_A2 (TM_21)	0.04	-	-	-
<b>TM_A3a (TM_3<sup>b</sup>)<sup>c</sup></b>	0.11	-	-	(2)
<b>TM_A3b (TM_3<sup>b</sup>)<sup>c</sup></b>	0.11	-	-	(2)
TM_A4 <sup>a</sup>	0.04	-	-	-
TM_A5a (TM_5a)	0.00	-	-	-
TM_A5b (TM_5b)	0.04	-	-	-
TM_A6a (TM_16a)	0.00	-	-	-
TM_A6b (TM_16b)	0.07	-	-	-
TM_A7 (TM_A7)	0.00	-	-	-
TM_A8 (TM_A14)	0.00	-	-	-
<b>TM_A9 (TM_A9)</b>	0.30	0.34	0.22	(3),(5)
<b>TM_A10 (TM_A12)</b>	0.48	0.42	0.43	(1)
TM_A11 (TM_A2)	0.07	-	-	-
<b>TM_A12a (TM_A20<sup>b</sup>)</b>	0.33	0.46	0.56	(4)
<b>TM_A12b (TM_A20<sup>b</sup>)</b>	0.33	0.63	0.66	(4)
<b>TM_A12c (TM_A20<sup>b</sup>)</b>	0.52	0.47	0.59	(4)
<b>TM_A13a (TM_A4a)</b>	0.22	0.28	0.42	(3)
<b>TM_A13b (TM_A4b)</b>	0.04	0.18	0.19	(3)
TM_A14 (TM_A18)	0.78	-0.06	-	-
<b>TM_A15a (TM_A6<sup>b</sup>)</b>	0.67	0.32	0.42	(4)
<b>TM_A15b (TM_A6<sup>b</sup>)</b>	0.07	0.30	0.19	(4)
<b>TM_A16 (TM_A15)</b>	0.33	0.58	0.49	(5)
<b>TM_A17 (TM_A8)</b>	0.30	0.32	0.23	(1)
TM_A18 (TM_A13)	0.44	-0.30	-	-
TM_A19 (TM_A10)	0.70	-0.55	-	-
<b>TM_A20a (TM_A11<sup>b</sup>)</b>	0.41	0.70	0.76	(1)
<b>TM_A20b (TM_A11<sup>b</sup>)</b>	0.52	0.64	0.71	(1)
<b>TM_A20c (TM_A11<sup>b</sup>)</b>	0.37	0.66	0.68	(1)

(zu Tabelle 43)

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> In Skala behalten aufgrund inhaltlicher Abdeckung der Kompetenzteilbereiche; <sup>d</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.

**Tabelle 44:** Pilotierungsergebnisse der Skala *Produktentwicklung*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>c</sup></b>	<b>Kompetenz-facetten</b>
Pe_A1 (Pe_1)	0.50	-0.79	-	-
Pe_A2a (Pe_20 <sup>b</sup> )	0.77	-0.40	-	-
Pe_A2b (Pe_20 <sup>b</sup> )	0.65	-0.61	-	-
Pe_A3 (Pe_3)	0.21	-0.29	-	-
Pe_A4 (Pe_18)	0.38	0.27	-	-
<b>Pe_A5 (Pe_A5)</b>	0.29	0.23	0.63	(1)
<b>Pe_A6 (Pe_A16)</b>	0.44	0.20	0.73	(2)
Pe_A7 (Pe_A7)	0.56	-0.54	-	-
<b>Pe_A8<sup>a</sup></b>	0.27	0.45	0.53	(2)
<b>Pe_A9 (Pe_A9)</b>	0.21	0.25	0.22	(2)
<b>Pe_A10 (Pe_A12)</b>	0.24	0.49	0.57	(1)
Pe_A11 <sup>a</sup>	0.06	-	-	-
<b>Pe_A12 (Pe_A19)</b>	0.74	-0.19	0.35	(4)
Pe_A13 <sup>a</sup>	0.12	-	-	-
<b>Pe_A14 (Pe_A14)</b>	0.62	0.34	0.43	(3)
Pe_A15 (Pe_A6)	0.18	0.33	-	-
Pe_A16 <sup>a</sup>	0.21	0.45	-	-
<b>Pe_A17 (Pe_A8)</b>	0.56	0.30	0.37	(3)
<b>Pe_A18 (Pe_A13)</b>	0.35	0.62	0.63	(4)
<b>Pe_A19 (Pe_A10)</b>	0.35	-0.17	0.44	(1)
<b>Pe_A20 (Pe_A11)</b>	0.27	-0.10	0.31	(2)

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.

**Tabelle 45:** Pilotierungsergebnisse der Skala *Grundlagen der Elektronik*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>d</sup></b>	<b>Kompetenz-facetten</b>
<b>GE_A1 (GE_1)</b>	0.67	0.42	0.41	(2)
<b>GE_A2 (GE_20)</b>	0.33	0.47	0.46	(3)
<b>GE_A3 (GE_3)</b>	0.17	0.51	0.52	(2)
<b>GE_A4 (GE_18)</b>	0.50	0.21	0.26	(2)
<b>GE_A5 (GE_5)</b>	0.33	0.55	0.58	(3)
<b>GE_A6 (GE_16)</b>	0.50	0.81	0.79	(1)
<b>GE_A7 (GE_7)</b>	0.25	0.55	0.53	(3)
<b>GE_A8 (GE_14)</b>	0.00	-	-	-
<b>GE_A9a (GE_A9<sup>b</sup>)</b>	0.25	0.59	0.64	(1)
<b>GE_A9b (GE_A9<sup>b</sup>)</b>	0.25	0.73	0.78	(1)
<b>GE_A10<sup>a</sup></b>	0.17	-	-	-
<b>GE_A11 (GE_2)</b>	0.83	0.37	-	-
<b>GE_A12a (GE_19<sup>b</sup>)</b>	0.42	0.16	0.21	(2)
<b>GE_A12b (GE_19<sup>b</sup>)</b>	0.42	0.54	0.54	(2)
<b>GE_A13 (GE_4)</b>	0.42	0.27	-	-
<b>GE_A14 (GE_17)</b>	0.58	0.82	0.79	(1)
<b>GE_A15<sup>a</sup></b>	0.58	0.82	0.79	(4)
<b>GE_A16 (GE_15)<sup>c</sup></b>	0.00	-	-	(4)
<b>GE_A17<sup>a</sup></b>	0.58	0.67	0.67	(3)
<b>GE_A18 (GE_A18)</b>	0.08	-	-	-
<b>GE_A19 (GE_A10)</b>	0.42	0.42	0.40	(3)
<b>GE_A20<sup>a,c</sup></b>	0.17	-	-	(4)

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> In Skala behalten aufgrund inhaltlicher Abdeckung der Kompetenzteilbereiche; <sup>d</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.

**Tabelle 46:** Pilotierungsergebnisse der Skala *Aufnahme und Verarbeitung von Signalen in der Messtechnik*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>c</sup></b>	<b>Kompetenz-facetten</b>
<b>Mt_A1 (Mt_1)</b>	0.58	0.39	0.58	(2)
<b>Mt_A2<sup>a</sup></b>	0.84	0.21	0.30	(1)
Mt_A3 (Mt_21)	0.58	-0.11	-	-
Mt_A4 (Mt_5)	0.42	-0.12	-	-
<b>Mt_A5 (Mt_19)</b>	0.32	0.30	0.46	(4)
<b>Mt_A6 (Mt_7)</b>	0.16	0.37	0.41	(3)
<b>Mt_A7 (Mt_16)</b>	0.37	0.25	0.40	(4)
<b>Mt_A8<sup>a</sup></b>	0.90	0.21	0.40	(4)
Mt_A9a <sup>a</sup>	0.05	-	-	-
Mt_A9b <sup>a</sup>	0.19	-	-	-
Mt_A10 (Mt_2)	0.11	-	-	-
Mt_A11 (Mt_5)	0.00	-	-	-
<b>Mt_A12 (Mt_4)</b>	0.26	0.41	0.61	(1)
<b>Mt_A13 (Mt_20)</b>	0.84	0.59	0.69	(5)
<b>Mt_A14<sup>a</sup></b>	0.32	0.40	0.42	(5)
<b>Mt_A15<sup>a</sup></b>	0.53	0.45	0.56	(4)
<b>Mt_A16a (Mt_8<sup>b</sup>)</b>	0.42	0.31	0.60	(3)
Mt_A16b (Mt_8 <sup>b</sup> )	0.32	-0.28	-	-
Mt_A17a <sup>a</sup>	0.00	-	-	-
Mt_A17b <sup>a</sup>	0.11	-	-	-
Mt_A18 (Mt_10)	0.58	0.16	-	-
Mt_A19a <sup>a</sup>	0.05	-	-	-
Mt_A19b <sup>a</sup>	0.05	-	-	-

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.



**Tabelle 47:** Pilotierungsergebnisse der Skala *Regelungstechnik*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>c</sup></b>	<b>Kompetenz-facetten</b>
Rt_A1 (Rt_2)	0.00	-	-	-
Rt_A2a (Rt_4 <sup>b</sup> )	0.39	-0.11	-	-
Rt_A2b (Rt_4 <sup>b</sup> )	0.31	0.09	-	-
Rt_A3 (Rt_19)	0.54	0.68	-	-
Rt_A4 (Rt_6)	0.85	0.17	-	-
<b>Rt_A5 (Rt_17)</b>	0.23	0.69	0.72	(4)
<b>Rt_A6 (Rt_8)</b>	0.92	0.52	0.55	(1)
Rt_A7 (Rt_15)	0.23	0.04	-	-
<b>Rt_A8 (Rt_10)</b>	0.54	0.25	0.53	(2)
<b>Rt_A9a (Rt_13<sup>b</sup>)</b>	0.15	0.67	0.82	(3)
<b>Rt_A9b (Rt_13<sup>b</sup>)</b>	0.23	0.58	0.68	(3)
<b>Rt_A10 (Rt_3)</b>	0.85	0.36	0.51	(1)
<b>Rt_A11 (Rt_20)</b>	0.92	0.34	0.57	(1)
<b>Rt_A12 (Rt_5)</b>	0.39	0.29	0.56	(1)
Rt_A13 (Rt_18)	0.46	0.22	-	-
Rt_A14 (Rt_14)	0.00	-	-	-
<b>Rt_A15a (Rt_16<sup>b</sup>)</b>	0.54	0.48	0.57	(2)
Rt_A15b (Rt_16)	0.00	-	-	-
<b>Rt_A16 (Rt_9)</b>	0.54	0.53	0.57	(3)
<b>Rt_A17 (Rt_14)</b>	0.39	0.29	0.31	(4)
<b>Rt_A18 (Rt_11)</b>	0.39	0.48	0.57	(1)
<b>Rt_A19 (Rt_12)</b>	0.15	0.37	0.50	(1)

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.

**Tabelle 48:** Pilotierungsergebnisse der Skala *Datenkommunikation*.

<b>Aufgabe (Item)</b>	<b>Lösungsquote</b>	<b>Trennschärfe</b>	<b>Finale Trennschärfe<sup>c</sup></b>	<b>Kompetenz-facetten</b>
<b>Dk_A1 (Dk_1)</b>	0.50	0.46	0.54	(1)
<b>Dk_A2<sup>a</sup></b>	0.43	0.29	0.29	(5)
<b>Dk_A3a (Dk_3<sup>a</sup>)</b>	0.50	0.61	0.62	(1)
<b>Dk_A3b (Dk_3<sup>b</sup>)</b>	0.46	0.69	0.74	(1)
<b>Dk_A3c (Dk_3<sup>b</sup>)</b>	0.46	0.70	0.67	(1)
<b>Dk_A4<sup>a</sup></b>	0.04	-	-	-
<b>Dk_A5 (Dk_5)</b>	0.61	0.26	0.27	(1)
<b>Dk_A6 (Dk_16)</b>	0.25	0.11	0.26	(4)
<b>Dk_A7 (Dk_7)</b>	0.07	-	-	-
<b>Dk_A8<sup>a</sup></b>	0.04	-	-	-
<b>Dk_A9 (Dk_9)</b>	0.07	-0.47	-	-
<b>Dk_A10 (Dk_12)</b>	0.43	-0.07	-	-
<b>Dk_A11 (Dk_2)</b>	0.96	0.31	0.32	(1)
<b>Dk_A12<sup>a</sup></b>	0.18	-	-	-
<b>Dk_A13<sup>a</sup></b>	0.43	0.51	0.56	(5)
<b>Dk_A14 (Dk_17)</b>	0.93	-0.14	-	-
<b>Dk_A15 (Dk_6)</b>	0.04	-0.11	-	-
<b>Dk_A16 (Dk_15)</b>	0.36	0.47	0.50	(4)
<b>Dk_A17 (Dk_8)</b>	0.00	-	-	-
<b>Dk_A18 (Dk_13)</b>	0.04	0.76	0.74	(2)
<b>Dk_A19a (Dk_10<sup>b</sup>)</b>	0.29	0.60	0.55	(3)
<b>Dk_A19b (Dk_10<sup>b</sup>)</b>	0.21	0.76	0.72	(3)
<b>Dk_A20 (Dk_11)</b>	0.18	0.47	0.47	(3)

<sup>a</sup> Neuentwicklung nach Expert:innenbefragung; <sup>b</sup> Aufgabe in Teilaufgaben aufgeteilt; <sup>c</sup> Trennschärfe nach Ausschluss unpassender Items aus der Skala.

## Anhang 6: Ergebnisse der IRT-Skalierung

### Rasch-Modell *Vorwissen Energie, Mobilität, Stoffe & Produkte*

Results of RM estimation:

Call: RM(X = rasch\_data)

Conditional log-likelihood: -907.7357

Number of iterations: 20

Number of parameters: 15

Item (Category) Difficulty Parameters (eta): with 0.95 CI:

Item	Estimate	Std. Error	lower CI	upper CI
A17b	-0.394	0.184	-0.755	-0.032
A3	-0.575	0.188	-0.942	-0.207
A5a	-1.047	0.201	-1.441	-0.653
A5b	-0.612	0.188	-0.981	-0.242
A15	0.886	0.187	0.519	1.254
A17c	0.998	0.190	0.626	1.370
A7a	-1.918	0.247	-2.402	-1.435
A7b	-2.544	0.301	-3.134	-1.954
A19	-1.090	0.203	-1.487	-0.693
A9a	1.036	0.191	0.662	1.410
A9b	0.191	0.180	-0.162	0.544
A21a	0.742	0.185	0.380	1.104
A21b	1.578	0.210	1.166	1.990
A11	1.192	0.195	0.809	1.575
A13	0.850	0.187	0.484	1.216

Item Easiness Parameters (beta) with 0.95 CI:

Item	Estimate	Std. Error	lower CI	upper CI
beta A17a	-0.706	0.184	-1.067	-0.345
beta A17b	0.394	0.184	0.032	0.755
beta A3	0.575	0.188	0.207	0.942
beta A5a	1.047	0.201	0.653	1.441
beta A5b	0.612	0.188	0.242	0.981
beta A15	-0.886	0.187	-1.254	-0.519
beta A17c	-0.998	0.190	-1.370	-0.626
beta A7a	1.918	0.247	1.435	2.402
beta A7b	2.544	0.301	1.954	3.134
beta A19	1.090	0.203	0.693	1.487
beta A9a	-1.036	0.191	-1.410	-0.662
beta A9b	-0.191	0.180	-0.544	0.162
beta A21a	-0.742	0.185	-1.104	-0.380
beta A21b	-1.578	0.210	-1.990	-1.166
beta A11	-1.192	0.195	-1.575	-0.809
beta A13	-0.850	0.187	-1.216	-0.484

---

Separation Reliability: 0.6997

Observed Variance: 1.3116 (Squared Standard Deviation)  
 Mean Square Measurement Error: 0.3939 (Model Error Variance)

---

Anhang 6: Ergebnisse der IRT-Skalierung

---

Itemfit Statistics:

	Chisq	df	p-value	Outfit MSQ	Infit MSQ	Outfit t	Infit t
A17a	146.882	138	0.287	1.057	0.998	0.451	0.004
A17b	115.667	138	0.917	0.832	0.921	-1.301	-0.993
A3	194.156	138	0.001	1.397	1.081	2.499	0.961
A5a	133.321	138	0.597	0.959	0.938	-0.159	-0.582
A5b	125.384	138	0.771	0.902	0.951	-0.640	-0.560
A15	123.124	138	0.813	0.886	0.978	-0.741	-0.244
A17c	112.403	138	0.946	0.809	0.881	-1.231	-1.430
A7a	189.940	138	0.002	1.366	0.945	1.169	-0.305
A7b	220.075	138	0.000	1.583	1.076	1.289	0.418
A19	145.482	138	0.315	1.047	1.089	0.303	0.865
A9a	149.777	138	0.233	1.078	1.051	0.518	0.610
A9b	107.130	138	0.976	0.771	0.850	-2.078	-2.160
A21a	119.041	138	0.876	0.856	0.899	-1.035	-1.313
A21b	149.883	138	0.231	1.078	0.925	0.417	-0.647
A11	160.126	138	0.096	1.152	1.020	0.872	0.242
A13	109.580	138	0.964	0.788	0.884	-1.501	-1.470

Discrim

A17a	0.360
A17b	0.472
A3	0.252
A5a	0.404
A5b	0.441
A15	0.365
A17c	0.463
A7a	0.216
A7b	0.067
A19	0.233
A9a	0.294
A9b	0.548
A21a	0.492
A21b	0.343
A11	0.278
A13	0.474

Minimaler Itemfit: 0.84972  
 Maximaler Itemfit: 1.089224

---

Zusammenfassung Itemschwierigkeiten

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	-2.54405	-0.82931	0.19085	-0.04709	0.94218	1.57792

---

EAP und WLE

EAP Reliabilität: 0.7059561

WLE Reliabilität:

Object of class 'tam.wle'

Call: TAM::tam.wle(tamobj = rm.1dim)

WLEs for 140 observations and 1 dimension

WLE Reliability=0.684

Average error variance=0.394

WLE mean=0.002

WLE variance=1.248

---

Andersen LR-test:

LR-value: 19.173

Chi-square df: 15

p-value: 0.206

---

Subgruppeninvarianz auf Itemebene mittels Wald-Test

Zieldef.: Einzelne signifikante Items passen nicht zum Test

Wald test on item level (z-values):

	z-statistic	p-value
beta A17a	-1.036	0.300
beta A17b	-1.134	0.257
beta A3	0.819	0.413
beta A5a	0.938	0.348
beta A5b	-0.252	0.801
beta A15	0.523	0.601
beta A17c	-1.469	0.142
beta A7a	1.547	0.122
beta A7b	1.689	0.091
beta A19	1.545	0.122
beta A9a	-0.504	0.614
beta A9b	-1.299	0.194
beta A21a	-1.662	0.097
beta A21b	-0.486	0.627
beta A11	0.936	0.350
beta A13	-1.225	0.220

---

Martin-Loef-Test (split criterion: median)

LR-value: 39.131

Chi-square df: 63

p-value: 0.992

---

**Rasch-Modell Vorwissen Informationsaufnahme und -verarbeitung**

Results of RM estimation:

Call: RM(X = rasch\_data)

Conditional log-likelihood: -1012.674

Number of iterations: 9

Number of parameters: 16

Item (Category) Difficulty Parameters (eta): with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
A24b	-0.249	0.182	-0.605	0.107
A12b	2.112	0.275	1.573	2.651
A22a	-0.110	0.182	-0.468	0.247
A22b	0.245	0.186	-0.119	0.610
A2	0.469	0.190	0.096	0.843
A4	-1.171	0.191	-1.545	-0.798
A8a	1.452	0.228	1.006	1.899
A10a	0.665	0.195	0.283	1.048
A10b	-0.145	0.182	-0.502	0.212
A14	-0.422	0.182	-0.778	-0.066
A20	1.193	0.215	0.773	1.614
A6	-0.110	0.182	-0.468	0.247
A16a	-0.318	0.182	-0.674	0.037
A16b	-1.887	0.215	-2.309	-1.465
A16c	-0.913	0.186	-1.277	-0.549
A18	-0.735	0.184	-1.095	-0.376

Item Easiness Parameters (beta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
beta A24a	0.075	0.183	-0.282	0.433
beta A24b	0.249	0.182	-0.107	0.605
beta A12b	-2.112	0.275	-2.651	-1.573
beta A22a	0.110	0.182	-0.247	0.468
beta A22b	-0.245	0.186	-0.610	0.119
beta A2	-0.469	0.190	-0.843	-0.096
beta A4	1.171	0.191	0.798	1.545
beta A8a	-1.452	0.228	-1.899	-1.006
beta A10a	-0.665	0.195	-1.048	-0.283
beta A10b	0.145	0.182	-0.212	0.502
beta A14	0.422	0.182	0.066	0.778
beta A20	-1.193	0.215	-1.614	-0.773
beta A6	0.110	0.182	-0.247	0.468
beta A16a	0.318	0.182	-0.037	0.674
beta A16b	1.887	0.215	1.465	2.309
beta A16c	0.913	0.186	0.549	1.277
beta A18	0.735	0.184	0.376	1.095

---

Separation Reliability: 0.7305

Observed Variance: 1.3561 (Squared Standard Deviation)  
 Mean Square Measurement Error: 0.3654 (Model Error Variance)

---

## Anhang 6: Ergebnisse der IRT-Skalierung

---

### Itemfit Statistics:

	Chisq	df	p-value	Outfit MSQ	Infit MSQ	Outfit t	Infit t	Discrim
A24a	142.710	138	0.374	1.027	1.036	0.261	0.484	0.396
A24b	116.879	138	0.904	0.841	0.911	-1.405	-1.179	0.509
A12b	129.499	138	0.685	0.932	0.897	-0.057	-0.508	0.292
A22a	124.605	138	0.786	0.896	0.936	-0.864	-0.818	0.484
A22b	119.106	138	0.876	0.857	0.928	-1.077	-0.857	0.499
A2	173.502	138	0.022	1.248	1.117	1.584	1.301	0.277
A4	173.138	138	0.023	1.246	1.122	1.507	1.376	0.243
A8a	162.682	138	0.074	1.170	1.081	0.701	0.639	0.200
A10a	104.965	138	0.984	0.755	0.865	-1.547	-1.465	0.511
A10b	144.259	138	0.340	1.038	1.064	0.356	0.839	0.350
A14	110.065	138	0.962	0.792	0.847	-1.878	-2.100	0.559
A20	128.334	138	0.711	0.923	0.988	-0.271	-0.065	0.329
A6	152.519	138	0.188	1.097	1.059	0.833	0.775	0.346
A16a	127.782	138	0.723	0.919	0.877	-0.674	-1.663	0.515
A16b	185.183	138	0.005	1.332	1.035	1.351	0.335	0.189
A16c	135.208	138	0.551	0.973	1.028	-0.155	0.368	0.360
A18	110.141	138	0.961	0.792	0.846	-1.723	-2.051	0.552

Minimaler Itemfit: 0.8459349

Maximaler Itemfit: 1.122117

---

### Zusammenfassung Itemschwierigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-1.887191	-0.500400	-0.127701	0.004709	0.518350	2.111677

---

### EAP und WLE

EAP Reliabilität: 0.7449859

WLE Reliabilität:

Object of class 'tam.wle'

Call: TAM::tam.wle(tamobj = rm.1dim)

WLEs for 140 observations and 1 dimension

WLE Reliability=0.712

Average error variance=0.364

WLE mean=0.003

WLE variance=1.263

---

### Andersen LR-test:

LR-value: 16.875

Chi-square df: 16

p-value: 0.394

---

Wald test on item level (z-values):

	z-statistic	p-value
beta A24a	0.778	0.437
beta A24b	-0.523	0.601
beta A12b	0.306	0.760
beta A22a	-0.925	0.355
beta A22b	-0.795	0.427
beta A2	-0.685	0.493
beta A4	0.603	0.547
beta A8a	1.736	0.082
beta A10a	-1.243	0.214
beta A10b	0.036	0.971
beta A14	-2.258	0.024
beta A20	0.935	0.350
beta A6	0.976	0.329
beta A16a	-0.528	0.597
beta A16b	1.240	0.215
beta A16c	0.770	0.441
beta A18	-0.792	0.429

---

Martin-Loef-Test (split criterion: median)

LR-value: 41.984

Chi-square df: 71

p-value: 0.998



---

**Rasch-Modell *Energieversorgung und Elektrische Antriebstechnik***

Results of RM estimation:

Call: RM(X = rasch\_data)

Conditional log-likelihood: -1002.087

Number of iterations: 14

Number of parameters: 19

Item (Category) Difficulty Parameters (eta): with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
Ev_A1c	-0.946	0.219	-1.376	-0.516
Ev_A1d	-1.049	0.223	-1.487	-0.611
Ev_A7	-0.567	0.207	-0.974	-0.161
Ev_A2	1.653	0.235	1.194	2.113
Ev_A9	0.274	0.199	-0.115	0.664
Ev_A11	-0.612	0.209	-1.021	-0.204
Ev_A13b	0.953	0.208	0.545	1.361
Ev_A16	-0.478	0.205	-0.881	-0.076
Ev_A17	-0.096	0.200	-0.487	0.296
Ev_A18	0.356	0.199	-0.034	0.746
EA_A3	-0.567	0.207	-0.974	-0.161
EA_A4	-0.658	0.210	-1.069	-0.247
EA_A10	2.504	0.296	1.923	3.084
EA_A11	-1.326	0.237	-1.790	-0.862
EA_A9	-0.348	0.203	-0.746	0.049
EA_A14	0.733	0.203	0.335	1.132
EA_A15	1.895	0.248	1.409	2.382
EA_A18a	-0.799	0.214	-1.218	-0.379
EA_A17	-1.156	0.228	-1.603	-0.708

Item Easiness Parameters (beta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
beta Ev_A1b	-0.233	0.199	-0.622	0.156
beta Ev_A1c	0.946	0.219	0.516	1.376
beta Ev_A1d	1.049	0.223	0.611	1.487
beta Ev_A7	0.567	0.207	0.161	0.974
beta Ev_A2	-1.653	0.235	-2.113	-1.194
beta Ev_A9	-0.274	0.199	-0.664	0.115
beta Ev_A11	0.612	0.209	0.204	1.021
beta Ev_A13b	-0.953	0.208	-1.361	-0.545
beta Ev_A16	0.478	0.205	0.076	0.881
beta Ev_A17	0.096	0.200	-0.296	0.487
beta Ev_A18	-0.356	0.199	-0.746	0.034
beta EA_A3	0.567	0.207	0.161	0.974
beta EA_A4	0.658	0.210	0.247	1.069
beta EA_A10	-2.504	0.296	-3.084	-1.923
beta EA_A11	1.326	0.237	0.862	1.790
beta EA_A9	0.348	0.203	-0.049	0.746
beta EA_A14	-0.733	0.203	-1.132	-0.335
beta EA_A15	-1.895	0.248	-2.382	-1.409
beta EA_A18a	0.799	0.214	0.379	1.218
beta EA_A17	1.156	0.228	0.708	1.603

---

Anhang 6: Ergebnisse der IRT-Skalierung

---

Separation Reliability: 0.7325

Observed Variance: 1.1254 (Squared Standard Deviation)  
Mean Square Measurement Error: 0.301 (Model Error Variance)

---

Übersicht Personenfähigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-2.5367	-0.5535	0.1815	0.2077	0.9227	2.6599	1

Normalverteilung Personenfähigkeit

Shapiro-Wilk normality test

data: p.parm  
W = 0.98616, p-value = 0.2936

---

Itemfit Statistics:

	Chisq	df	p-value	Outfit MSQ	Infit MSQ	Outfit t	Infit t	Discrim
Ev_A1b	113.070	113	0.480	0.992	1.014	-0.034	0.209	0.354
Ev_A1c	116.027	113	0.404	1.018	0.963	0.157	-0.318	0.364
Ev_A1d	104.597	113	0.702	0.918	0.994	-0.356	-0.015	0.347
Ev_A7	124.749	113	0.212	1.094	0.955	0.659	-0.480	0.407
Ev_A2	118.090	113	0.353	1.036	1.031	0.232	0.287	0.240
Ev_A9	140.878	113	0.039	1.236	1.161	1.957	1.991	0.186
Ev_A11	99.499	113	0.814	0.873	0.881	-0.811	-1.322	0.480
Ev_A13b	92.484	113	0.921	0.811	0.883	-1.315	-1.294	0.494
Ev_A16	131.341	113	0.114	1.152	1.095	1.062	1.089	0.260
Ev_A17	89.347	113	0.951	0.784	0.854	-1.911	-1.926	0.543
Ev_A18	127.381	113	0.168	1.117	1.075	1.017	0.959	0.305
EA_A3	144.049	113	0.026	1.264	1.074	1.669	0.839	0.258
EA_A4	98.954	113	0.824	0.868	0.957	-0.820	-0.436	0.406
EA_A10	68.208	113	1.000	0.598	0.837	-1.157	-0.814	0.368
EA_A11	100.443	113	0.795	0.881	0.953	-0.443	-0.319	0.344
EA_A9	92.194	113	0.924	0.809	0.879	-1.492	-1.489	0.507
EA_A14	127.060	113	0.173	1.115	1.062	0.882	0.742	0.272
EA_A15	122.281	113	0.259	1.073	1.071	0.359	0.533	0.206
EA_A18a	80.814	113	0.990	0.709	0.814	-1.828	-1.965	0.550
EA_A17	107.409	113	0.631	0.942	1.001	-0.204	0.052	0.317

Minimaler Itemfit: 0.8140349

Maximaler Itemfit: 1.160546

---

Zusammenfassung Itemschwierigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-1.32592	-0.72839	-0.47842	-0.01227	0.54484	2.50359

---

EAP und WLE

EAP-Reliabilität: 0.7454778

WLE-Reliabilität:

Object of class 'tam.wle'

Call: TAM::tam.wle(tamobj = rm.1dim)

WLEs for 115 observations and 1 dimension

WLE Reliability=0.73

Average error variance=0.313

WLE mean=0.009

WLE variance=1.159

---

Andersen LR-test:

LR-value: 21.303

Chi-square df: 19

p-value: 0.32

---

Wald test on item level (z-values):

	z-statistic	p-value
beta Ev_A1b	0.220	0.826
beta Ev_A1c	0.844	0.399
beta Ev_A1d	0.642	0.521
beta Ev_A7	-0.833	0.405
beta Ev_A2	1.715	0.086
beta Ev_A9	2.063	0.039
beta Ev_A11	-1.150	0.250
beta Ev_A13b	-1.160	0.246
beta Ev_A16	0.673	0.501
beta Ev_A17	-1.649	0.099
beta Ev_A18	1.264	0.206
beta EA_A3	0.997	0.319
beta EA_A4	-0.544	0.587
beta EA_A10	-0.434	0.664
beta EA_A11	-0.284	0.777
beta EA_A9	-0.690	0.490
beta EA_A14	1.397	0.162
beta EA_A15	0.239	0.811
beta EA_A18a	-1.081	0.279
beta EA_A17	-0.683	0.495

---

Martin-Loef-Test (split criterion: median)

LR-value: 59.006

Chi-square df: 99

p-value: 1

**Rasch-Modell Technische Mechanik, Produktentwicklung und Technikfolgenabschätzung**

Results of RM estimation:

Call: RM(X = rasch\_data)

Conditional log-likelihood: -1764.235

Number of iterations: 16

Number of parameters: 33

Item (Category) Difficulty Parameters (eta): with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
TM_A10	-2.271	0.297	-2.852	-1.689
TM_A12a	0.000	0.196	-0.383	0.384
TM_A12b	-0.274	0.197	-0.660	0.112
TM_A12c	0.195	0.196	-0.190	0.580
TM_A13a	-1.279	0.224	-1.719	-0.839
TM_A13b	-0.039	0.196	-0.422	0.345
TM_A15a	-2.099	0.280	-2.648	-1.550
TM_A15b	1.504	0.238	1.037	1.970
TM_A16	-1.873	0.261	-2.385	-1.361
TM_A17	-0.942	0.211	-1.355	-0.528
TM_A20a	-2.099	0.280	-2.648	-1.550
TM_A20b	-2.020	0.273	-2.555	-1.485
TM_A20c	-1.228	0.222	-1.663	-0.793
Pe_A5	-0.234	0.197	-0.620	0.151
Pe_A6	0.353	0.198	-0.035	0.741
Pe_A9	2.307	0.305	1.709	2.905
Pe_A8	1.504	0.238	1.037	1.970
Pe_A12	-0.987	0.213	-1.404	-0.571
Pe_A14	-0.274	0.197	-0.660	0.112
Pe_A20	1.036	0.215	0.615	1.458
Pe_A19	1.820	0.260	1.311	2.329
Pe_A17	-1.034	0.214	-1.454	-0.614
Pe_A18	0.433	0.199	0.043	0.823
Ta_A2	-0.852	0.208	-1.260	-0.444
Ta_A3	-0.680	0.204	-1.079	-0.281
Ta_A4	1.284	0.226	0.841	1.726
Ta_A19	1.563	0.242	1.089	2.036
Ta_A5	1.232	0.223	0.795	1.670
Ta_A6	0.898	0.210	0.487	1.310
Ta_A9	0.314	0.197	-0.073	0.700
Ta_A18	-0.078	0.196	-0.462	0.306
Ta_A20	1.446	0.235	0.987	1.906
Ta_A12a	2.765	0.363	2.053	3.477

Item Easiness Parameters (beta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
beta TM_A3a	0.393	0.198	0.004	0.781
beta TM_A10	2.271	0.297	1.689	2.852
beta TM_A12a	0.000	0.196	-0.384	0.383
beta TM_A12b	0.274	0.197	-0.112	0.660
beta TM_A12c	-0.195	0.196	-0.580	0.190
beta TM_A13a	1.279	0.224	0.839	1.719
beta TM_A13b	0.039	0.196	-0.345	0.422
beta TM_A15a	2.099	0.280	1.550	2.648
beta TM_A15b	-1.504	0.238	-1.970	-1.037
beta TM_A16	1.873	0.261	1.361	2.385

Anhang 6: Ergebnisse der IRT-Skalierung

---

beta TM_A17	0.942	0.211	0.528	1.355
beta TM_A20a	2.099	0.280	1.550	2.648
beta TM_A20b	2.020	0.273	1.485	2.555
beta TM_A20c	1.228	0.222	0.793	1.663
beta Pe_A5	0.234	0.197	-0.151	0.620
beta Pe_A6	-0.353	0.198	-0.741	0.035
beta Pe_A9	-2.307	0.305	-2.905	-1.709
beta Pe_A8	-1.504	0.238	-1.970	-1.037
beta Pe_A12	0.987	0.213	0.571	1.404
beta Pe_A14	0.274	0.197	-0.112	0.660
beta Pe_A20	-1.036	0.215	-1.458	-0.615
beta Pe_A19	-1.820	0.260	-2.329	-1.311
beta Pe_A17	1.034	0.214	0.614	1.454
beta Pe_A18	-0.433	0.199	-0.823	-0.043
beta Ta_A2	0.852	0.208	0.444	1.260
beta Ta_A3	0.680	0.204	0.281	1.079
beta Ta_A4	-1.284	0.226	-1.726	-0.841
beta Ta_A19	-1.563	0.242	-2.036	-1.089
beta Ta_A5	-1.232	0.223	-1.670	-0.795
beta Ta_A6	-0.898	0.210	-1.310	-0.487
beta Ta_A9	-0.314	0.197	-0.700	0.073
beta Ta_A18	0.078	0.196	-0.306	0.462
beta Ta_A20	-1.446	0.235	-1.906	-0.987
beta Ta_A12a	-2.765	0.363	-3.477	-2.053

---

Separation Reliability: 0.7476

Observed Variance: 0.6918 (Squared Standard Deviation)  
Mean Square Measurement Error: 0.1746 (Model Error Variance)

---

Übersicht Personenfähigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-2.03928	-0.66556	-0.00884	-0.02443	0.56657	2.04111

Normalverteilung Personenfähigkeit

Shapiro-Wilk normality test

data: p.parm  
W = 0.98587, p-value = 0.272

---

Itemfit Statistics:

	Chisq	df	p-value	Outfit MSQ	Infit MSQ	Outfit t	Infit t	Discrim
TM_A3a	115.194	114	0.451	1.002	1.038	0.049	0.535	0.308
TM_A10	91.184	114	0.943	0.793	0.927	-0.612	-0.296	0.288
TM_A12a	99.827	114	0.825	0.868	0.901	-1.638	-1.483	0.461
TM_A12b	113.415	114	0.498	0.986	0.972	-0.130	-0.375	0.370
TM_A12c	96.597	114	0.879	0.840	0.882	-1.926	-1.780	0.488
TM_A13a	139.701	114	0.051	1.215	1.031	1.253	0.309	0.203
TM_A13b	111.857	114	0.539	0.973	0.980	-0.305	-0.277	0.358
TM_A15a	109.723	114	0.596	0.954	0.950	-0.071	-0.214	0.251
TM_A15b	101.293	114	0.797	0.881	0.933	-0.540	-0.477	0.323
TM_A16	139.161	114	0.055	1.210	0.955	0.881	-0.230	0.223

Anhang 6: Ergebnisse der IRT-Skalierung

---

TM_A17	130.380	114	0.140	1.134	1.079	1.012	0.855	0.179
TM_A20a	100.294	114	0.817	0.872	0.940	-0.380	-0.269	0.302
TM_A20b	105.128	114	0.712	0.914	0.944	-0.238	-0.261	0.301
TM_A20c	102.189	114	0.778	0.889	0.980	-0.660	-0.150	0.354
Pe_A5	102.159	114	0.779	0.888	0.915	-1.326	-1.231	0.445
Pe_A6	115.692	114	0.438	1.006	0.968	0.095	-0.430	0.341
Pe_A9	91.013	114	0.945	0.791	0.908	-0.568	-0.361	0.261
Pe_A8	139.171	114	0.055	1.210	1.086	1.033	0.684	0.088
Pe_A12	124.787	114	0.231	1.085	1.014	0.651	0.174	0.261
Pe_A14	113.695	114	0.490	0.989	0.975	-0.102	-0.330	0.362
Pe_A20	143.875	114	0.031	1.251	1.091	1.612	0.935	0.135
Pe_A19	111.227	114	0.556	0.967	0.941	-0.050	-0.319	0.283
Pe_A17	124.829	114	0.230	1.085	1.037	0.635	0.407	0.227
Pe_A18	122.104	114	0.285	1.062	1.062	0.651	0.862	0.254
Ta_A2	91.234	114	0.943	0.793	0.851	-1.752	-1.733	0.525
Ta_A3	111.901	114	0.538	0.973	0.988	-0.211	-0.114	0.335
Ta_A4	112.321	114	0.527	0.977	0.968	-0.074	-0.242	0.268
Ta_A19	120.669	114	0.317	1.049	1.035	0.299	0.302	0.171
Ta_A5	109.714	114	0.596	0.954	0.942	-0.217	-0.499	0.308
Ta_A6	110.590	114	0.573	0.962	0.976	-0.243	-0.240	0.315
Ta_A9	99.937	114	0.823	0.869	0.906	-1.470	-1.358	0.448
Ta_A18	115.633	114	0.440	1.006	1.005	0.094	0.100	0.333
Ta_A20	149.089	114	0.015	1.296	1.055	1.444	0.467	0.135
Ta_A12a	169.787	114	0.001	1.476	0.984	1.116	0.041	0.045

Minimaler Itemfit: 0.8509617

Maximaler Itemfit: 1.090877

---

Zusammenfassung Itemschwierigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-2.27058	-0.98744	-0.03882	0.01190	1.23220	2.76506

---

EAP-Reliabilität: 0.7454799

WLE-Reliabilität:

Object of class 'tam.wle'

Call: TAM::tam.wle(tamobj = rm.1dim)

WLEs for 115 observations and 1 dimension

WLE Reliability=0.738

Average error variance=0.174

WLE mean=-0.002

WLE variance=0.665

---

Andersen LR-test:

LR-value: 22.818

Chi-square df: 33

p-value: 0.908

---

Wald test on item level (z-values):

	z-statistic	p-value
beta TM_A3a	0.930	0.352
beta TM_A10	-1.250	0.211
beta TM_A12a	-0.949	0.343
beta TM_A12b	-0.417	0.677
beta TM_A12c	-0.787	0.431
beta TM_A13a	0.924	0.356
beta TM_A13b	-0.355	0.722
beta TM_A15a	0.426	0.670
beta TM_A15b	-0.973	0.330
beta TM_A16	0.515	0.607
beta TM_A17	-0.224	0.822
beta TM_A20a	-0.237	0.812
beta TM_A20b	0.247	0.805
beta TM_A20c	-0.211	0.833
beta Pe_A5	-0.600	0.549
beta Pe_A6	-0.060	0.952
beta Pe_A9	0.169	0.866
beta Pe_A8	1.134	0.257
beta Pe_A12	0.837	0.402
beta Pe_A14	-1.615	0.106
beta Pe_A20	1.457	0.145
beta Pe_A19	-1.374	0.169
beta Pe_A17	1.007	0.314
beta Pe_A18	1.106	0.269
beta Ta_A2	-0.978	0.328
beta Ta_A3	-0.327	0.743
beta Ta_A4	0.442	0.659
beta Ta_A19	1.317	0.188
beta Ta_A5	0.276	0.783
beta Ta_A6	0.055	0.956
beta Ta_A9	-0.240	0.810
beta Ta_A18	-0.160	0.873
beta Ta_A20	0.441	0.659
beta Ta_A12a	1.062	0.288

---

Martin-Loef-Test (split criterion: median)

LR-value: 116.814

Chi-square df: 288

p-value: 1

**Rasch-Modell Messtechnik und Regelungstechnik**

Results of RM estimation:  
 Call: RM(X = rasch\_data)

Conditional log-likelihood: -1104.134  
 Number of iterations: 17  
 Number of parameters: 22

Item (Category) Difficulty Parameters (eta): with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
Mt_A2	-0.742	0.219	-1.171	-0.313
Mt_A7	-0.505	0.211	-0.919	-0.091
Mt_A8	-2.612	0.367	-3.331	-1.894
Mt_A12	0.258	0.199	-0.132	0.649
Mt_A6	2.249	0.268	1.723	2.775
Mt_A13	-2.477	0.349	-3.161	-1.793
Mt_A14	1.091	0.209	0.681	1.500
Mt_A15	-0.505	0.211	-0.919	-0.091
Mt_A16a	0.299	0.199	-0.092	0.690
Rt_A6	-2.612	0.367	-3.331	-1.894
Rt_A5	2.491	0.289	1.924	3.058
Rt_A12	-0.414	0.209	-0.823	-0.005
Rt_A8	1.135	0.210	0.723	1.547
Rt_A9a	-0.028	0.202	-0.423	0.367
Rt_A9b	0.177	0.200	-0.214	0.569
Rt_A10	-0.326	0.207	-0.731	0.079
Rt_A11	0.420	0.199	0.030	0.811
Rt_A15a	-0.842	0.222	-1.278	-0.406
Rt_A16	0.177	0.200	-0.214	0.569
Rt_A17	0.665	0.201	0.271	1.059
Rt_A18	-0.196	0.204	-0.596	0.204
Rt_A19	2.580	0.298	1.996	3.164

Item Easiness Parameters (beta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
beta Mt_A1	0.282	0.206	-0.121	0.685
beta Mt_A2	0.742	0.219	0.313	1.171
beta Mt_A7	0.505	0.211	0.091	0.919
beta Mt_A8	2.612	0.367	1.894	3.331
beta Mt_A12	-0.258	0.199	-0.649	0.132
beta Mt_A6	-2.249	0.268	-2.775	-1.723
beta Mt_A13	2.477	0.349	1.793	3.161
beta Mt_A14	-1.091	0.209	-1.500	-0.681
beta Mt_A15	0.505	0.211	0.091	0.919
beta Mt_A16a	-0.299	0.199	-0.690	0.092
beta Rt_A6	2.612	0.367	1.894	3.331
beta Rt_A5	-2.491	0.289	-3.058	-1.924
beta Rt_A12	0.414	0.209	0.005	0.823
beta Rt_A8	-1.135	0.210	-1.547	-0.723
beta Rt_A9a	0.028	0.202	-0.367	0.423
beta Rt_A9b	-0.177	0.200	-0.569	0.214
beta Rt_A10	0.326	0.207	-0.079	0.731
beta Rt_A11	-0.420	0.199	-0.811	-0.030
beta Rt_A15a	0.842	0.222	0.406	1.278
beta Rt_A16	-0.177	0.200	-0.569	0.214
beta Rt_A17	-0.665	0.201	-1.059	-0.271
beta Rt_A18	0.196	0.204	-0.204	0.596
beta Rt_A19	-2.580	0.298	-3.164	-1.996



Anhang 6: Ergebnisse der IRT-Skalierung

---

Separation Reliability: 0.7447

Observed Variance: 1.0686 (Squared Standard Deviation)  
Mean Square Measurement Error: 0.2729 (Model Error Variance)

---

Übersicht Personenfähigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-2.4940	-0.3374	0.3536	0.2941	1.0334	3.0334	1

Normalverteilung Personenfähigkeit

Shapiro-Wilk normality test

data: p.parm  
W = 0.98367, p-value = 0.1806

---

Itemfit Statistics:

	Chisq	df	p-value	Outfit MSQ	Infit MSQ	Outfit t	Infit t	Discrim
Mt_A1	115.187	113	0.425	1.010	1.029	0.123	0.355	0.339
Mt_A2	129.741	113	0.134	1.138	1.122	0.828	1.126	0.227
Mt_A7	123.671	113	0.232	1.085	1.066	0.605	0.703	0.300
Mt_A8	93.775	113	0.906	0.823	0.932	-0.208	-0.162	0.255
Mt_A12	114.871	113	0.433	1.008	1.020	0.105	0.289	0.341
Mt_A6	97.028	113	0.858	0.851	0.893	-0.414	-0.609	0.292
Mt_A13	94.129	113	0.901	0.826	0.931	-0.240	-0.193	0.266
Mt_A14	102.705	113	0.746	0.901	0.919	-0.625	-0.934	0.386
Mt_A15	165.406	113	0.001	1.451	1.138	2.714	1.403	0.211
Mt_A16a	105.937	113	0.668	0.929	0.971	-0.611	-0.373	0.418
Rt_A6	122.647	113	0.252	1.076	0.920	0.315	-0.210	0.218
Rt_A5	80.062	113	0.992	0.702	0.920	-0.840	-0.360	0.313
Rt_A12	98.850	113	0.826	0.867	0.894	-0.939	-1.166	0.488
Rt_A8	111.213	113	0.530	0.976	1.016	-0.106	0.209	0.301
Rt_A9a	95.716	113	0.879	0.840	0.887	-1.402	-1.463	0.493
Rt_A9b	97.560	113	0.849	0.856	0.913	-1.313	-1.169	0.469
Rt_A10	129.168	113	0.142	1.133	0.942	0.993	-0.642	0.401
Rt_A11	140.997	113	0.038	1.237	1.122	1.984	1.638	0.223
Rt_A15a	103.636	113	0.725	0.909	0.901	-0.453	-0.873	0.437
Rt_A16	100.424	113	0.795	0.881	0.865	-1.067	-1.866	0.519
Rt_A17	105.680	113	0.675	0.927	0.958	-0.575	-0.546	0.375
Rt_A18	134.893	113	0.078	1.183	1.091	1.415	1.079	0.225
Rt_A19	61.907	113	1.000	0.543	0.774	-1.385	-1.148	0.390

Minimaler Itemfit: 0.7736593

Maximaler Itemfit: 1.137777

---

Zusammenfassung Itemschwierigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-2.61241	-0.50480	0.07449	0.01283	0.60373	2.58030

---

## Anhang 6: Ergebnisse der IRT-Skalierung

---

EAP-Reliabilität: 0.7631502

WLE-Reliabilität:

Object of class 'tam.wle'

Call: TAM::tam.wle(tamobj = rm.1dim)

WLEs for 115 observations and 1 dimension

WLE Reliability=0.76

Average error variance=0.288

WLE mean=-0.02

WLE variance=1.197

---

Andersen LR-test:

LR-value: 23.847

Chi-square df: 22

p-value: 0.355

---

Wald test on item level (z-values):

	z-statistic	p-value
beta Mt_A1	0.251	0.802
beta Mt_A2	1.798	0.072
beta Mt_A7	1.473	0.141
beta Mt_A8	-0.078	0.938
beta Mt_A12	-0.161	0.872
beta Mt_A6	0.339	0.735
beta Mt_A13	-0.230	0.818
beta Mt_A14	-0.964	0.335
beta Mt_A15	1.923	0.054
beta Mt_A16a	-0.335	0.738
beta Rt_A6	0.990	0.322
beta Rt_A5	-0.359	0.720
beta Rt_A12	-1.118	0.263
beta Rt_A8	0.933	0.351
beta Rt_A9a	-1.092	0.275
beta Rt_A9b	0.182	0.856
beta Rt_A10	-0.043	0.966
beta Rt_A11	1.546	0.122
beta Rt_A15a	-0.936	0.349
beta Rt_A16	-1.462	0.144
beta Rt_A17	0.006	0.995
beta Rt_A18	1.241	0.215
beta Rt_A19	-1.453	0.146

---

Martin-Loef-Test (split criterion: median)

LR-value: 72.292

Chi-square df: 131

p-value: 1

---

**Rasch-Modell Grundlagen der Elektronik**

Results of RM estimation:

Call: RM(X = rasch\_data)

Conditional log-likelihood: -650.1083

Number of iterations: 27

Number of parameters: 15

Item (Category) Difficulty Parameters (eta): with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
GEA2_dich	-0.418	0.208	-0.826	-0.011
GEA4_dich	1.070	0.243	0.594	1.547
GEA5_dich	1.411	0.264	0.893	1.929
GEA6_dich	1.338	0.259	0.830	1.847
GEA7_dich	-0.690	0.210	-1.103	-0.278
GEA16_dich	-0.374	0.208	-0.781	0.034
GEA9a_dich	-0.239	0.208	-0.646	0.168
GEA9b_dich	-0.329	0.208	-0.736	0.078
GEA17_dich	-2.598	0.314	-3.214	-1.982
GEA12a_dich	1.070	0.243	0.594	1.547
GEA12b_dich	0.512	0.220	0.081	0.943
GEA15_dich	-1.490	0.233	-1.947	-1.033
GEA14_dich	1.338	0.259	0.830	1.847
GEA19_dich	-1.119	0.219	-1.549	-0.689
GEA20_dich	1.738	0.290	1.170	2.307

Item Easiness Parameters (beta) with 0.95 CI:

	Estimate	Std. Error	lower CI	upper CI
beta GEA1_dich	1.221	0.223	0.785	1.657
beta GEA2_dich	0.418	0.208	0.011	0.826
beta GEA4_dich	-1.070	0.243	-1.547	-0.594
beta GEA5_dich	-1.411	0.264	-1.929	-0.893
beta GEA6_dich	-1.338	0.259	-1.847	-0.830
beta GEA7_dich	0.690	0.210	0.278	1.103
beta GEA16_dich	0.374	0.208	-0.034	0.781
beta GEA9a_dich	0.239	0.208	-0.168	0.646
beta GEA9b_dich	0.329	0.208	-0.078	0.736
beta GEA17_dich	2.598	0.314	1.982	3.214
beta GEA12a_dich	-1.070	0.243	-1.547	-0.594
beta GEA12b_dich	-0.512	0.220	-0.943	-0.081
beta GEA15_dich	1.490	0.233	1.033	1.947
beta GEA14_dich	-1.338	0.259	-1.847	-0.830
beta GEA19_dich	1.119	0.219	0.689	1.549
beta GEA20_dich	-1.738	0.290	-2.307	-1.170

---

Separation Reliability: 0.6332

Observed Variance: 1.0543 (Squared Standard Deviation)

Mean Square Measurement Error: 0.3867 (Model Error Variance)

---

Übersicht Personenfähigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-2.4607	-1.0178	-0.3093	-0.3115	0.3602	2.4274	1

## Anhang 6: Ergebnisse der IRT-Skalierung

---

Normalverteilung Personenfähigkeit

Shapiro-Wilk normality test

data: p.parm  
W = 0.97771, p-value = 0.08504

---

Itemfit Statistics:

	Chisq	df	p-value	Outfit MSQ	Infit MSQ	Outfit t	Infit t	Discrim
GEA1	98.194	100	0.532	0.972	1.008	-0.106	0.108	0.243
GEA2	119.979	100	0.085	1.188	1.134	1.540	1.646	0.167
GEA4	84.777	100	0.862	0.839	0.940	-0.695	-0.421	0.372
GEA5	66.160	100	0.996	0.655	0.849	-1.362	-0.944	0.485
GEA6	90.036	100	0.752	0.891	0.876	-0.347	-0.791	0.407
GEA7	87.915	100	0.801	0.870	0.919	-1.013	-0.973	0.399
GEA16	85.045	100	0.857	0.842	0.879	-1.385	-1.573	0.483
GEA9a	126.322	100	0.039	1.251	1.124	2.027	1.539	0.159
GEA9b	98.041	100	0.537	0.971	1.011	-0.214	0.164	0.305
GEA17	77.964	100	0.950	0.772	0.918	-0.507	-0.329	0.234
GEA12a	97.141	100	0.562	0.962	0.965	-0.101	-0.226	0.313
GEA12b	134.689	100	0.012	1.334	1.171	1.978	1.680	0.064
GEA15	67.680	100	0.995	0.670	0.778	-1.776	-2.036	0.495
GEA14	69.212	100	0.992	0.685	0.833	-1.278	-1.105	0.482
GEA19	101.290	100	0.445	1.003	1.010	0.072	0.138	0.286
GEA20	82.357	100	0.900	0.815	0.958	-0.491	-0.163	0.293

Minimaler Itemfit: 0.7782228

Maximaler Itemfit: 1.171184

---

Zusammenfassung Itemschwierigkeiten

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-2.59805	-0.55440	-0.23900	0.08139	1.20431	1.73841

---

Anhang 6: Ergebnisse der IRT-Skalierung

---

EAP-PV-Reliabilität: 0.650108

WLE-Reliabilität

Object of class 'tam.wle'

Call: TAM::tam.wle(tamobj = rm.1dim)

WLEs for 102 observations and 1 dimension

WLE Reliability=0.634

Average error variance=0.401

WLE mean=-0.004

WLE variance=1.094

---

Andersen LR-test:

LR-value: 19.592

Chi-square df: 15

p-value: 0.188

---

Wald test on item level (z-values):

	z-statistic	p-value
beta GEA1_dich	1.504	0.133
beta GEA2_dich	1.751	0.080
beta GEA4_dich	-0.849	0.396
beta GEA5_dich	-1.253	0.210
beta GEA6_dich	-0.800	0.423
beta GEA7_dich	0.637	0.524
beta GEA16_dich	-1.472	0.141
beta GEA9a_dich	1.842	0.066
beta GEA9b_dich	0.952	0.341
beta GEA17_dich	-0.580	0.562
beta GEA12a_dich	0.808	0.419
beta GEA12b_dich	1.789	0.074
beta GEA15_dich	-0.986	0.324
beta GEA14_dich	-0.800	0.423
beta GEA19_dich	0.689	0.491
beta GEA20_dich	0.040	0.968

---

Martin-Loef-Test (split criterion: median)

LR-value: 44.599

Chi-square df: 63

p-value: 0.962



### **Anhang 7: Weitere Materialien**

Auf Anfrage beim Autor sind folgende weitere Materialien erhältlich:

- Testhefte des Vorwissenstests
- Testhefte des Fachwissenstests
- Lösungsheft Vorwissenstest
- Lösungsheft Fachwissenstests





## Erklärung über die Eigenständigkeit der Dissertation

Ich versichere, dass ich die vorliegende Arbeit mit dem Titel **Inhaltsbezogene Kompetenzen von Schülerinnen und Schülern im Fach Naturwissenschaft und Technik in der gymnasialen Oberstufe** selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe; aus fremden Quellen entnommene Passagen und Gedanken sind als solche kenntlich gemacht.

Die Arbeit ist bislang keiner anderen Prüfungsbehörde vorgelegt und nicht an anderer Stelle veröffentlicht worden.

Ort, Datum, Unterschrift: \_\_\_\_\_

Marcus Brändle