

Technische Problemlösekompetenz im Alltag –
theoretische Entwicklung und empirische Prüfung des
Kompetenzkonstruktes

Problemlösen im Umgang mit technischen Geräten

Von der Fakultät für Ingenieurwissenschaften, Abteilung Bauwissenschaften,
Technologie und Didaktik der Technik
der Universität Duisburg-Essen
zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil.)

genehmigte Dissertation von

Dipl.-Ing. Jennifer Stemmann

aus Hattingen

Erstgutachter: Univ.-Prof. Dr. Martin Lang, Universität Duisburg-Essen

Zweitgutachter: Univ.-Prof. Dr. Stefan Fletcher, Universität Duisburg-Essen

Tag der mündlichen Prüfung: 14.12.2016

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen, all jenen Personen zu danken, die mich bei der Erstellung dieser Dissertation auf vielfältige Weise unterstützt haben.

Ein besonders großer Dank gilt meinem Doktorvater Prof. Dr. Martin Lang. Sein mir entgegengebrachtes Vertrauen sowie seine Haltung zur wissenschaftlichen Freiheit trugen maßgeblich zu einer produktiven und kreativitätsfördernden Arbeitsatmosphäre bei, in der meine Eigenständigkeit nie eingeschränkt wurde. Schon seit dem Beginn meiner Arbeit ermöglichte er mir die Teilnahme an wissenschaftlichen Diskursen und den Besuch zahlreicher Fortbildungen, die letztendlich zum Gelingen dieser Arbeit beigetragen haben. In Situationen, in denen ich mit scheinbar unüberwindbaren Schwierigkeiten zu ihm kam, halfen mir die zuversichtlichen und zukunftsorientierten Gespräche mit ihm und seine Kompetenz Probleme pragmatisch zu lösen.

Für die unkomplizierte und menschliche Art der Zusammenarbeit sowie für die Übernahme des Zweitgutachtens bedanke ich mich außerdem bei Herrn Prof. Dr. Stefan Fletcher. Die Gespräche mit ihm sowie seine kritischen Fragen zu Beginn meines Forschungsvorhabens haben zu einer gewinnbringenden Fokussierung dieser Arbeit geführt.

Bei Herrn Prof. Dr. Felix Walker bedanke ich mich zum einen für seine fachliche Expertise, die er mir zu jeder Zeit (!) zur Verfügung gestellt hat und zum anderen für seine langjährige kollegiale Unterstützung sowohl meiner Forschung als auch meiner wissenschaftlichen Entwicklung. Erwähnenswert und eines besonderen Dankes gebührend ist dabei die Tatsache, dass diese Unterstützung stets uneigennützig und selbst dann erfolgte, wenn er selbst keine Zeit hatte. Mit seiner Begeisterung für wissenschaftliche Forschung und seiner Genauigkeit bei der Bearbeitung von Forschungsfragen ist er für mich ein großes Vorbild.

Ein wesentlicher Teil meiner Arbeit war die Entwicklung eines computerbasierten Testinstruments, für die ich auf den CBA ItemBuilder zurückgegriffen habe. Diese Software wurde mir vom Arbeitsbereich Technologiebasiertes Assessment des Deutschen Institutes für Internationale Pädagogische Forschung (DIPF) zur Verfügung gestellt, dem ich hierfür danke. Carolin Hahnel danke ich außerdem für ihre Hilfe bei der Nutzung dieses Programmes. Herrn Prof. Dr. Samuel Greiff danke ich für das Zurverfügungstellen des Tests zur Erfassung allgemeiner Problemlösefähigkeit und Julia Rudolph für ihre Ideen und Anregungen bei der Auswertung der Daten. Dem IZfB danke ich für seine Unterstützung bei der Ausleihe von Laptops, dem Scannen von Testmaterial sowie der Möglichkeit meine Arbeit im Rahmen eines Doktorandenkolloquiums vorzustellen. Ohne die vielen Probandinnen und Probanden meiner Studie wäre diese Arbeit nicht möglich gewesen, ihnen gilt mein besonderer Dank.

Meinen studentischen Hilfskräften, insbesondere Emil Kaschner danke ich für die zahlreichen Stunden, die sie mit Programmierung, Grafikerstellung, Testdurchführung und Auswertung verbracht haben. Durch Emil wurden (nicht nur aus Zeitgründen) viele meiner Ideen erst umsetzbar. Danken möchte ich aber auch meinen wissenschaftlichen Kolleginnen und Kollegen für die belebende und freundschaftliche Stimmung am Arbeitsplatz. Namentlich ist zum einen Elmar Dammann zu nennen, der seine Dissertation zeitgleich zu Ende gebracht hat und vor allem in dem letzten Jahr ein Leidenspartner aber auch ein wichtiger Ansprechpartner

bei methodischen und statistischen Fragen war. Dem Büro C47 mit Marcel, Tatiana und Michaela danke ich für die offene Kommunikation und die kleinen und schnellen Hilfestellungen zwischendurch, Christian zudem für seine Übersetzungshilfe.

Bei meiner Familie, insbesondere meiner Schwester und bei meinen Freunden möchte ich mich dafür bedanken, dass sie so lange auf mich verzichteten und dafür auch noch Verständnis aufbringen mussten.

Meinem Ehemann Dennis danke ich dafür, dass er besonders in der arbeitsreichsten letzten Phase dieser Arbeit alle alltäglichen Arbeiten und Probleme von mir ferngehalten und seine eigenen Bedürfnisse für mich zurückgestellt hat. Vor allem aber danke ich ihm für seine Geduld, seine Toleranz, seinen Humor und seine Fähigkeit mich immer wieder in das *normale* Leben zurückzuholen. Ihm widme ich diese Arbeit.

Inhaltsverzeichnis

Abbildungsverzeichnis	5
Tabellenverzeichnis.....	9
1 Einleitung.....	13
Teil I: Stand der Forschung	
2 Probleme und Problemlösen.....	19
2.1 Problemverständnis	19
2.2 Klassifikation von Problemen (Taxonomien)	20
2.3 Schwierigkeitsbestimmende Problemmerkmale	26
2.4 Problemlösen.....	32
2.4.1 Wissenserwerb und Wissensanwendung im problemlösenden Umgang mit technischen Alltagsgeräten	34
2.4.2 Kognitionspsychologische Lerntheorien.....	39
3 Problemlösekompetenz.....	47
3.1 Kompetenzverständnis.....	47
3.2 Indikatoren von Problemlösekompetenz.....	52
4 Computerbasierte Kompetenzdiagnostik.....	60
4.1 Computersimulationen.....	61
4.2 Nutzen und Herausforderungen computerbasierter Tests	63
4.3 Finite Automaten als formale Modelle zur Erstellung von computersimulierten technischen Geräten.....	66
5 Einflussfaktoren auf den Prozess und die Leistung beim Problemlösen	72
5.1 Kontextübergreifende Einflussfaktoren	72
5.1.1 Intelligenz	72
5.1.2 Allgemeine Problemlösefähigkeit	80
5.1.3 Need for Cognition und Selbstwirksamkeitserwartungen.....	84
5.2 Kontextuelle Einflussfaktoren	88
5.2.1 Kontextbezug als Merkmal von Kompetenz	88
5.2.2 Erfahrungen.....	91
Teil II: Theoretisches Konstrukt	
6 Alltägliche technische Systeme als Problemkontext	97
6.1 Problemlösen in technischen Kontexten	97
6.2 Soziotechnische Systeme	99

6.3	Interaktionen in soziotechnischen Systemen.....	101
6.4	Merkmale, die die Interaktion mit technischen Systemen erschweren	105
6.4.1	Systemmerkmale	106
6.4.2	Situationsmerkmale.....	109
6.5	Indikatoren technischer Problemlösekompetenz	110
6.5.1	Systemexploration.....	111
6.5.2	Systemsteuerung	112
6.6	Formulierung eines theoretischen Kausalmodells	113
7	Forschungsfragen und Hypothesen.....	117
 Teil III: Methodik		
8	Operationalisierung technischer Problemlösekompetenz.....	121
8.1	Kontext – technische Alltagsgeräte	121
8.2	Interaktionsformen.....	122
8.3	Schwierigkeitsbestimmende Merkmale	123
8.4	Itemkonstruktion (im CBA-Item Builder).....	124
8.5	Beschreibung des Testmaterials	126
8.5.1	Beschreibung der einzelnen Items	126
8.5.2	Beschreibung des zusammengestellten Tests.....	148
8.5.3	Beschreibung des Scorings	151
8.6	Erprobung des Testinstrumentes	153
8.6.1	Erste Testerprobungsphase	154
8.6.2	Zweite Testerprobungsphase.....	158
8.7	Testinstruktion.....	160
9	Datenerhebung	161
9.1	Weitere Testinstrumente	161
9.1.1	CFT 20-R	161
9.1.2	Selbstwirksamkeitserwartungen.....	162
9.1.3	Need for Cognition.....	164
9.1.4	ICT-Literacy und gerätespezifisches Vorwissen.....	166
9.1.5	Allgemeine Problemlösefähigkeit	170
9.2	Testdesign.....	173
9.3	Stichprobe	174
9.4	Zeitlicher Ablauf.....	176

10	Validitätsbetrachtungen.....	178
10.1	Schwierigkeitsbestimmende Merkmale	182
10.1.1	Regressionsanalyse.....	183
10.1.2	Bestimmung der Merkmalsausprägungen	190
10.2	Dimensionalität technischer Problemlösekompetenz	198
10.3	Strukturmodell zur Verortung der Problemlösekompetenz im Umgang mit technischen Alltagsgeräten	203
 Teil IV: Forschungsergebnisse		
11	Deskriptive Befunde.....	211
11.1	Fluide Intelligenz.....	211
11.2	Need for Cognition.....	212
11.3	Selbstwirksamkeitserwartung	214
11.4	ICT-Literacy.....	216
11.5	Gerätespezifisches Vorwissen	217
11.6	Allgemeine Problemlösefähigkeit	219
11.7	Problemlösekompetenz im Umgang mit technischen Alltagsgeräten.....	222
11.8	Erste Zusammenhangsbetrachtungen.....	227
12	Güte des Messinstruments zur Erfassung des problemlösenden Umgangs mit technischen Alltagsgeräten.....	231
12.1	Messmodell der Systemexploration.....	231
12.2	Messmodell der Systemsteuerung	235
13	Schwierigkeitsbestimmende Merkmale	238
13.1	Empirische Itemschwierigkeiten	238
13.1.1	Explorationsphase	238
13.1.2	Steuerungsphase.....	240
13.2	Prüfung der Voraussetzungen für die multiple lineare Regression	243
13.2.1	Explorationsphase	244
13.2.2	Steuerungsphase.....	249
13.3	Regressionsergebnisse und Modellgüte.....	252
13.3.1	Explorationsphase	252
13.3.2	Steuerungsphase.....	255

14	Strukturanalyse technischer Problemlösekompetenz.....	261
15	Verortung technischer Problemlösekompetenz im nomologischen Netzwerk ..	267

Teil V: Diskussion

16	Problemlösender Umgang mit technischen Alltagsgeräten - ein eigenständiges Konstrukt?.....	277
16.1	Zusammenfassende Ergebnisbetrachtungen.....	277
16.2	Theoriebezogene Ergebnisinterpretation.....	281
16.3	Limitationen und Ausblick	292
17	Literaturverzeichnis	297

Abbildungsverzeichnis

Abbildung 2.1: Determinanten der Problemschwierigkeit dynamischer (komplexer) Probleme (Hussy, 1984b, S. 4)	29
Abbildung 2.2: Determinanten des Umgangs mit dynamischen Systemen (nach Funke, 1990)	30
Abbildung 2.3: Kategorien und Ebenen von Wissen zur Kontrolle und Steuerung technischer Systeme (Kluwe, 1997b, S. 20)	37
Abbildung 2.4: Modellarchitektur eines Produktionssystems (Opwis, 1988, S. 77)	40
Abbildung 2.5: Systemarchitektur von ACT-R (nach Anderson & Lebiere, 1998a, S. 11; Wallach & Lebiere, 1998, S. 95).....	41
Abbildung 4.1: Grad der Abstraktion von Modellen (Obermayer, 1964, S. 608; Übersetzungen und Ergänzungen (gestrichelte Linien) von Neber, 1987, S. 19).....	61
Abbildung 4.2: Überschneidung zwischen realem Gegenstandsbereich und Modell (vgl. Saam, 1997, S. 69).....	62
Abbildung 4.3: Zustandsdiagramm eines Automaten zum Bezahlen der Parkgebühr in einem Parkhaus (Balzert, 2005, S. 92); Bedingungen stehen in eckigen Klammern	68
Abbildung 4.4: Beispiel für das Betreten und Verlassen eines hierarchischen Zustands	69
Abbildung 4.5: Beispiel eines orthogonalen Systemzustandes.....	69
Abbildung 4.6: Notation eines einfachen Zustandsautomaten	70
Abbildung 5.1: Die Hierarchie der Bedürfnisse nach Maslow (1970) (Seel, 2000, S. 79)	85
Abbildung 6.1: Soziotechnisches System (nach Möller, 2003, S. 23)	101
Abbildung 6.2: Beispiele für den Übergang von realen zu virtuellen Bedienelementen (nach Baumann & Lanz, 1998)	104
Abbildung 6.3: Verortung von Problemlösekompetenz im Umgang mit technischen Alltagssystemen in einem nomologischen Netzwerk.....	114
Abbildung 6.4: Theoretisches Kausalmodell der Problemlösekompetenz im Umgang mit technischen Alltagsgeräten.....	116
Abbildung 8.1: simuliertes Interface eines Parkscheinautomaten	127
Abbildung 8.2: simuliertes Interface eines Receivers	128
Abbildung 8.3: simuliertes Interface eines Radioweckers	129
Abbildung 8.4: simuliertes Interface eines Fotodruckers	131
Abbildung 8.5: simuliertes Interface eines Kaffeeautomaten	132
Abbildung 8.6: simuliertes Interface eines Multifunktionsdruckers	134
Abbildung 8.7: simuliertes Interface einer Waschmaschine	135
Abbildung 8.8: simuliertes Interface eines Küchenradios	136
Abbildung 8.9: simuliertes Interface eines Navigationsgerätes.....	138

Abbildung 8.10: simuliertes Interface eines Festnetztelefons	139
Abbildung 8.11: simuliertes Interface einer Spülmaschine	140
Abbildung 8.12: simuliertes Interface einer Rollladensteuerung	141
Abbildung 8.13: simuliertes Interface einer Heizungssteuerung	142
Abbildung 8.14: simuliertes Interface eines Mobiltelefons	143
Abbildung 8.15: simuliertes Interface der Software eines Internetrouters.....	145
Abbildung 8.16: simuliertes Interface eines Fahrscheinautomaten.....	146
Abbildung 8.17: simuliertes Interface eines Fotoentwicklungsautomaten	147
Abbildung 8.18: simuliertes Interface eines Online-Banking Portals	148
Abbildung 8.19: gleichbleibendes Layout für alle Items des computerbasierten Tests zum Umgang mit technischen Geräten am Beispiel des Items Radiowecker	150
Abbildung 8.20: Beispiel für die Ermittlung der Explorationsvollständigkeit anhand der relativen Anzahl an explorierten Systemzuständen in einer Heizungssteuerung (die Zustände Start und Betriebsbereit werden nicht mitgezählt, da diese Zustände automatisch auch ohne Interaktionen erreicht werden).....	151
Abbildung 9.1: Ein Beispielitem des Untertests aus dem CFT 20-R (Grundintelligenztest Skala 2 Revision; Weiß (2006)); die Aufgabenstellung lautet die Figurenreihe logisch richtig fortzusetzen (Hornke, Amelang & Kersting, 2011, S. 21).....	162
Abbildung 9.2: Abfolge und Testzeiten der eingesetzten Testinstrumente und Skalen	174
Abbildung 9.3: Erhebungszeitraum.....	177
Abbildung 10.1: Zu prüfendes Strukturmodell zur Verortung technischer Problemlösekompetenz (die Messmodelle für die latenten Variablen sind hier nicht dargestellt).....	206
Abbildung 11.1: Häufigkeiten der erreichten IQ-Werte der $n = 147$ Probanden mit eingezeichneter Normalverteilungskurve.....	212
Abbildung 11.2: Häufigkeiten der Testwerte der Skala Need for Cognition der $n = 147$ Probanden mit eingezeichneter Normalverteilungskurve	214
Abbildung 11.3: Häufigkeit der beobachteten Testwerte für die Aufgabentypen Verifikation, Identifikation und Steuerung des allgemeinen Problemlösetests.....	221
Abbildung 11.4: Häufigkeiten der beobachteten Testwerte für die Exploration und die Steuerung technischer Geräte.....	226
Abbildung 12.1: kongenerisches Messmodell M2 für die Systemexploration mit geschätzten Faktorladungen und Messfehlervarianzen ($n = 136$)	234
Abbildung 12.2: kongenerisches Messmodell M2 für die Systemsteuerung mit geschätzten Faktorladungen und Messfehlervarianzen ($n = 141$)	237
Abbildung 13.1: Itemschwierigkeiten der Exploration technischer Geräte, sortiert nach ihrer empirischen Schwierigkeit (Explorationsvollständigkeit), von links nach rechts werden die Items entsprechend schwieriger zu explorieren ($n = 136$)	239

Abbildung 13.2: Itemschwierigkeiten der Steuerung technischer Geräte, berechnet aus der relativen Lösungshäufigkeit, sortiert nach ihrer empirischen Schwierigkeit, von links nach rechts werden die Items entsprechend schwieriger zu steuern ($n = 141$).....	241
Abbildung 13.3: Itemschwierigkeiten der Steuerung technischer Geräte berücksichtigt die für die Lösung benötigten Interaktionsschritte, sortiert nach ihrer empirischen Schwierigkeit, von links nach rechts werden die Items entsprechend schwieriger zu steuern ($n = 141$)	242
Abbildung 13.4: Vergleich der unterschiedlich berechneten empirischen Itemschwierigkeiten der Systemsteuerung ($n = 141$).....	243
Abbildung 13.5: Streudiagramme der als schwierigkeitsbestimmend angenommenen Systemmerkmale auf die empirische Schwierigkeit in der Systemexploration ($n = 15$)	245
Abbildung 13.6: Boxplot der als schwierigkeitsbestimmend angenommenen Vorwissenskompatibilität auf die Explorationsvollständigkeit ($n = 15$).....	246
Abbildung 13.7: Streudiagramme zur Beurteilung von Normalverteilung und Heteroskedastizität der Residuen für die verbleibenden als schwierigkeitsbestimmend angenommenen Merkmale der Systemexploration ($n = 15$)	248
Abbildung 13.8: Streudiagramme und Boxplot der als schwierigkeitsbestimmend angenommenen Merkmale auf die (dichotom ausgewertete) empirische Schwierigkeit der Systemsteuerung ($n = 15$).....	250
Abbildung 13.9: Residualplots zur Beurteilung von Normalverteilung und Heteroskedastizität der Residuen für die verbleibenden als schwierigkeitsbestimmend angenommenen Merkmale der Systemsteuerung (dichotome Ergebnisauswertung) ($n = 15$)	251
Abbildung 14.1: Modelle für die Struktur des problemlösenden Umgangs mit technischen Alltagsgeräten a) Systemexploration und Systemsteuerung erfassen dasselbe Merkmal; b) Systemexploration und Systemsteuerung sind zwei voneinander abgrenzbare Merkmale, die miteinander korrelieren; c) Systemexploration und Systemsteuerung sind zwei voneinander abgrenzbare Merkmale die miteinander korrelieren, die Residuen der Indikatoren sind teilweise korreliert.....	262
Abbildung 14.2: Ergebnis der Dimensionalitätsprüfung für das zweidimensionale Modell technischer Problemlösekompetenz mit korrelierenden Residuen ($n = 130$)	263
Abbildung 14.3: Modelle zur Prüfung diskriminanter Validität a) technische Systemexploration und Exploration im allgemeinen Problemlösen sind identisch, technische Systemsteuerung und Steuerung im allgemeinen Problemlösen sind identisch; b) technische Systemexploration, technische Systemsteuerung, Exploration und Steuerung im allgemeinen Problemlösen sind vier unterscheidbare latente Variablen	265
Abbildung 14.4: Ergebnis für das vierdimensionale Modell, das sowohl zwischen Exploration und Steuerung als auch zwischen technischem und allgemeinem Problemlösen differenziert. Alle eingezeichneten Korrelationen zwischen den latenten Variablen sind mit $p < .001$ signifikant ($n = 77$)	266

Abbildung 15.1: Strukturmodell zur Verortung technischer Problemlösekompetenz in einem Theoriemodell ($n = 77$); getrichelt dargestellte Effekte sind nicht signifikant, $*p < .05$; $**p < .01$; $***p < .001$; $R^2 =$ Kommunalität; Modellfit: $X^2/df = 1.21$; CFI = .96; RMSEA = .04 [.03, .05]; WRMR = .82 269

Abbildung 15.2: Einfluss fluider Intelligenz auf die Zusammenhänge zwischen Exploration und Steuerung im Umgang mit technischen Systemen und allgemeinen Problemlöseszenarien sowie auf den Zusammenhang zwischen technischem und allgemeinem Problemlösen ($n = 77$); getrichelt dargestellte Kovarianzen sind nicht signifikant; $*p < .05$; $**p < .01$; $***p < .001$; $R^2 =$ Kommunalität 273

Abbildung 15.3: Einfluss von Selbstwirksamkeitserwartungen auf die Zusammenhänge zwischen Exploration und Steuerung im Umgang mit technischen Systemen und allgemeinen Problemlöseszenarien sowie auf den Zusammenhang zwischen technischem und allgemeinem Problemlösen ($n = 77$); getrichelt dargestellte Effekte sind nicht signifikant; $*p < .05$; $**p < .01$; $***p < .001$; $R^2 =$ Kommunalität..... 274

Tabellenverzeichnis

Tabelle 2.1: Klassifikation von Problemen der Arbeits- und Ingenieurpsychologie (nach Krause, 1982a).....	21
Tabelle 2.2: Taxonomie der Wissensbegriffe (Wittmann, Süß & Oberauer, 1996, S. 5) mit Beispielen aus Walker (2013).....	35
Tabelle 6.1: zu operationalisierende schwierigkeitsbestimmende Merkmale	106
Tabelle 6.2: Ergebnis der Präzisierung schwierigkeitsbestimmender Merkmale für System- und Situationsmerkmale	110
Tabelle 8.1: Beispiele technischer Geräte aus dem Alltag.....	122
Tabelle 8.2: Arten der Transition.....	125
Tabelle 8.3: In der Steuerungsphase maximal zu erreichende itemabhängige Punktzahl sowie die Anzahl der hierfür nötigen Bearbeitungsschritte	152
Tabelle 8.4: Bewertung der Problemlöseleistung in der Steuerungsphase mit der Berücksichtigung der für die Lösung benötigten Anzahl an Bearbeitungsschritten bei gegebener Mindestanzahl an Bearbeitungsschritten	153
Tabelle 8.5: relative Lösungshäufigkeiten der einzelnen Items und deren Trennschärfe (Ergebnisse aus der zweiten Erprobung).....	158
Tabelle 9.1: Itemwortlaut der Skalen Ausdauer und Aufgeschlossenheit gegenüber Problemlösen (vgl. OECD, 2013b)	164
Tabelle 9.2: Itemwortlaut der Skala Need for Cognition (vgl. Preckel, 2014)	165
Tabelle 9.3: Itemwortlaut der Skala ICT – Nutzung (vgl. OECD, 2013a)	167
Tabelle 9.4: Itemwortlaut der Skala Erfahrungen im Umgang mit technischen Geräten	167
Tabelle 9.5: Beschreibung der Szenarien des Tests zur Erfassung allgemeiner Problemlösefähigkeit	171
Tabelle 9.6: Stichprobendaten.....	176
Tabelle 10.1: Blockweise Aufnahme unabhängiger Variablen in die Regressionsgleichung zur Bestimmung der Itemschwierigkeit	187
Tabelle 10.2: Ausprägungen der schwierigkeitsbestimmenden Systemmerkmale der Hardware in den einzelnen technischen Geräten	192
Tabelle 10.3: Ausprägungen der schwierigkeitsbestimmenden Systemmerkmale der Software in den einzelnen technischen Geräten.....	195
Tabelle 10.4: Ausprägungen der schwierigkeitsbestimmenden Systemmerkmale der Software in den einzelnen technischen Geräten.....	196
Tabelle 10.5: Ausprägungen der schwierigkeitsbestimmenden Situationsmerkmale der einzelnen technischen Geräte	197
Tabelle 11.1: Deskriptive Statistik und Subtest-Rohwert-Korrelationen der Ergebnisse des eingesetzten CFT 20-R	211

Tabelle 11.2: Deskriptive Statistik der, in genormte IQ-Werte transformierten Ergebnisse des eingesetzten CFT 20-R	212
Tabelle 11.3: relative Häufigkeiten der gewählten Kategorien aller Items der Skala Need for Cognition (n = 147).....	213
Tabelle 11.4: Deskriptive Statistik der Skala Need for Cognition.....	214
Tabelle 11.5: relative Häufigkeiten der gewählten Kategorien aller Items der Skalen Ausdauer und Aufgeschlossenheit gegenüber dem Problemlösen (n = 147)	215
Tabelle 11.6: Deskriptive Statistik der Skalen Ausdauer und Aufgeschlossenheit gegenüber Problemlösen	216
Tabelle 11.7: relative Häufigkeiten der gewählten Kategorien aller Items der Skala ICT-Nutzung (n = 147)	217
Tabelle 11.8: Deskriptive Statistik der Skala Nutzung von Informations- und Kommunikationstechnologien	217
Tabelle 11.9: relative Häufigkeiten der gewählten Kategorien aller Items der Skala gerätespezifisches Vorwissen (n = 147).....	218
Tabelle 11.10: Deskriptive Statistik der Skala gerätespezifisches Vorwissen	219
Tabelle 11.11: relative Lösungshäufigkeiten der unterschiedlichen Aufgaben für die Wissensabfrage nach der Exploration und für die Steuerungsaufgabe im allgemeinen Problemlösetest	220
Tabelle 11.12: Deskriptive Statistik für die unterschiedlichen Aufgaben der Wissensabfrage nach der Exploration sowie für die Steuerungsaufgabe im allgemeinen Problemlösetest	221
Tabelle 11.13: Interkorrelation zwischen dem Wissen nach der Exploration (Verifikation und Identifikation) und der Wissens-anwendung in der Steuerung im allgemeinen Problemlösetest	221
Tabelle 11.14: Explorationsvollständigkeit für den explorierenden Umgang und relative Lösungshäufigkeiten für den steuernden Umgang mit technischen Systemen in dem computerbasierten technischen Problemlösetest.....	223
Tabelle 11.15: deskriptive Statistiken des explorierenden und steuernden Umgangs mit den simulierten technischen Alltagsgeräten im technischen Problemlösekompetenztest..	225
Tabelle 11.16: bivariate Korrelationen (Spearman) zwischen dem gerätespezifischen Vorwissen und der Vollständigkeit der Systemexploration und der Steuerungsleistung für die 15 technischen Systeme	228
Tabelle 11.17: Skaleninterkorrelationen der manifesten Variablen Systemexploration und Systemsteuerung	229
Tabelle 11.18: Interkorrelationen zwischen allen manifesten Skalen- und Testwerten.....	230
Tabelle 12.1: Fitstatistiken für die Messmodelle der Systemexploration (n = 136)	232

Tabelle 12.2: standardisierte Faktorladungen, Itemleichtigkeiten, Messfehlervarianzen und Itemreliabilitäten des kongenerischen Messmodells M2 der Systemexploration (n = 136)	233
Tabelle 12.3: Fitstatistiken für die Messmodelle der Systemsteuerung (n = 141).....	235
Tabelle 12.4: standardisierte Faktorladungen, Thresholds, Messfehlervarianzen und Itemreliabilität des kongenerischen Messmodells M2 der Systemsteuerung (n = 141)	236
Tabelle 13.1: Bewertung der Steuerungsleistung.....	241
Tabelle 13.2: Interkorrelation schwierigkeitsbestimmender Systemmerkmale in der Phase der Exploration nach Entfernung der Variablen, die keinen (linearen) Einfluss auf die Explorationsvollständigkeit haben (n = 15).....	247
Tabelle 13.3: Partielle Korrelation der schwierigkeitsbestimmenden Merkmale der Komplexität auf Softwareebene unter Konstanthaltung der Variable Anzahl sichtbarer Zustände (n = 15)	247
Tabelle 13.4: Interkorrelation schwierigkeitsbestimmender Merkmale der Systemsteuerung (n = 15)	251
Tabelle 13.5: Regressionskoeffizienten und globale Modellgütemaße für die blockweise Analyse des Einflusses von Merkmalen auf die Itemschwierigkeit der Systemexploration (n = 15)	253
Tabelle 13.6: Regressionskoeffizienten und globale Modellgütemaße der Analyse schwierigkeitsbestimmender Merkmale der Systemexploration nach Entfernung von Variablen mit nichtsignifikanten Regressionskoeffizienten (n = 15)	254
Tabelle 13.7: Regressionskoeffizienten und globale Modellgütemaße für die blockweise Analyse des Einflusses von Merkmalen auf die Itemschwierigkeit der Systemsteuerung (dichotome Leistungsbewertung) (n = 15)	256
Tabelle 13.8: Regressionskoeffizienten und globale Modellgütemaße für die blockweise Analyse des Einflusses von Merkmalen auf die Itemschwierigkeit der Systemsteuerung (Leistungsbewertung berücksichtigt Anzahl der Interaktionen) (n = 15).....	257
Tabelle 13.9: Regressionskoeffizienten und globale Modellgütemaße der Analyse schwierigkeitsbestimmender Merkmale der Systemsteuerung (dichotome Leistungsbewertung) nach Entfernung von Variablen mit geringem Einfluss auf die Itemschwierigkeit (n = 15)	259
Tabelle 13.10: Regressionskoeffizienten und globale Modellgütemaße der Analyse schwierigkeitsbestimmender Merkmale der Systemsteuerung (Leistungsbewertung berücksichtigt Anzahl der Interaktionen) nach Entfernung von Variablen mit geringem Einfluss auf die Itemschwierigkeit (n = 15)	260
Tabelle 14.1: Fitstatistiken des Modellvergleichs für die Dimensionalitätsprüfung des problemlösenden Umgangs mit technischen Alltagsgeräten (n = 130).....	262
Tabelle 14.2: standardisierte Faktorladungen und Korrelationskoeffizienten der Residuen zwischen den Items der Systemexploration und der Systemsteuerung im technischen Problemlösetest (n =130)	264

Tabelle 14.3: Fitstatistiken des Modellvergleichs zur Prüfung diskriminanter Validität des problemlösenden Umgangs mit technischen Alltagsgeräten (n = 77).....	265
Tabelle 15.1: Modellgütemaße für die Messmodelle, der in dem Strukturmodell eingefügten latenten Variablen	268
Tabelle 15.2: Modellparamater des Teilmodells für den Einfluss von Selbstwirksamkeitserwartung auf allgemeines und technisches Problemlösen (Modellgütemaße: $\chi^2 = 874.48$; $p < .001$; $df = 727$; $\chi^2/df = 1.20$; CFI = .96; RMSEA = .04 [.03, .05]; WRMR = .82) (n = 77).....	271

1 Einleitung

In einer zunehmend technisierten Welt, in der Technologien in immer kürzeren Abständen entwickelt werden, nimmt das Lösen technischer Probleme eine bedeutende Stellung ein. Angesichts des sich aus dieser Entwicklung ergebenden Bedarfs an qualifiziertem Personal, werden von den Unternehmen bereits seit einiger Zeit Anstrengungen unternommen, die sich auf den betrieblichen Wissenserwerb (Lang, 2004) sowie das betriebliche Wissensmanagement (Nakhosteen, 2009) beziehen. Die Technologiefortschritte bleiben allerdings nicht auf den betrieblichen Kontext beschränkt, was sich unter anderem in der stetig steigenden Anzahl technischer Produkte im Alltag ausmachen lässt. Technische Artefakte ermöglichen das Erreichen von Handlungszielen, die ohne den Einsatz von Technik nicht oder nur mit einem erheblichen Mehraufwand zu verwirklichen wären (Greif, Mitrea & Werner, 2008). Mit diversen Weiterentwicklungen aus den Bereichen Mikrotechnologie und Digitaltechnik lassen sich zunehmend Informations- und Kommunikationstechniken in technische Geräte implementieren, die die Reichweite technischen Handelns weiter vergrößern.

Dem Vorteil eines Zugewinns an Handlungsmöglichkeiten steht jedoch der Nachteil gegenüber, dass Technik wiederum eine komplexere Umwelt schafft. In vielen Fällen erfolgt der Zugang zur und die Interaktion mit der Umwelt zuallererst über technische Geräte (ebd.). Beispielsweise kann der öffentliche Personennahverkehr in manchen Regionen nur genutzt werden, wenn es zuvor gelungen ist, die dafür erforderliche Fahrkarte an einem Selbstbedienungsautomaten zu erwerben, denn Fahrkartenschalter, an denen eine Bedienung stattfindet, gibt es immer weniger. Durch den hohen Diffusionsgrad von Technik in den Alltag wird vermehrt die Kompetenz im Umgang mit technischen Geräten vorausgesetzt, denn wer „technische Einrichtungen und Geräte nicht bedienen kann, bleibt vom Einsatz arbeitserleichternder, freizeitwerterhöhender oder mobilitätssteigernder Technikprodukte ausgeschlossen“ (Sackmann, Weymann & Hüttner, 1994).

Die Nutzung technischer Geräte wird häufig dann zu einem Problem, wenn technische Systeme neu sind, man noch keine Erfahrungen im Umgang mit ihnen sammeln konnte oder diese nicht erwartungsgemäß funktionieren. Die rasante Weiterentwicklung technischer Geräte sowie deren kürzere Nutzungszeiten führen zu einer ständigen Entwertung konkreten Bedienungswissens und einer permanenten Veränderung bewährter Handlungsroutinen, so dass sich der Mensch immer wieder mit unbekanntem technischen Systemen auseinandersetzen muss. Durch den gestiegenen Funktionsumfang und dem damit oft verschwindenden eindeutigen Zweckbezug technischer Geräte (Tully, 2003), ist ein alle Funktionen umfassendes Bedienungswissen ohnehin häufig gar nicht mehr denkbar. Weil sich Technikwissen aufgrund dessen nicht mehr entlang den Funktionen eines technischen Gerätes vermitteln lässt, verlieren Bedienungsanleitungen zunehmend an Bedeutung. Für die zusätzlichen Applikationen einiger Systeme (z. B. Computersoftware oder Apps für Smartphones) sind solche Anleitungen entweder überhaupt nicht oder nur mit Hilfe weiterer Technik (z. B. dem Internet) zu finden.

Der Erwerb von Bedienungswissen ist in vielen Situationen nur noch durch zielgerichtetes Ausprobieren (Explorieren) möglich, womit Ausprobieren ein wesentlicher Bestandteil der Techniksozialisation wird (ebd.). Voraussetzung für die selbstständige Aneignung von

Bedienungswissen sind allerdings einfach zu bedienende Systeme. Forderungen die Nutzungsphase bereits in die Entwicklung technischer Systeme einzubeziehen, werden von den Produktherstellern bislang jedoch noch nicht ausreichend berücksichtigt. Dabei besteht die Problematik in den starren Bedienkonzepten, die von einer vereinheitlichten Gruppe von Nutzern und Nutzungskontexten ausgehen und erwarten, dass sich diese der Technik anpassen (Zühlke, 2005). Gelingt eine solche Anpassung aufgrund mangelnder Erfahrungen oder falscher Vorannahmen nicht, führt die Nutzung technischer Geräte zu Hilflosigkeit und Frustration.

Die Folge derartiger negativer Erfahrungen kann eine Abnahme oder ein vollständiger Verlust des Interesses an Technik sein (Jakobs, 2005), womit die Konsequenz verbunden ist, dass eine weitere Auseinandersetzung mit technischen Sachverhalten und damit der Erwerb von Technikkompetenzen ausbleibt (Mammes, 2001). Ein Anhaltspunkt dafür findet sich beispielsweise in den Ergebnissen der von acatech und VDI im Auftrag gegebenen Studie *Nachwuchsbarometer Technikwissenschaften* (Pfenning & Renn, 2010). Hier zeigte sich, dass der Anteil an Schülerinnen und Schülern, die ein erhöhtes Interesse an Technik zeigen, in der Gruppe von Schülerinnen und Schülern, die überdurchschnittlich viele technische Geräte nutzen größer ist, als in den Gruppen von Schülerinnen und Schülern, die weniger technische Geräte nutzen.

Mit dem Ziel Schülerinnen und Schüler zu befähigen technische Probleme zu erkennen und zu lösen, soll der allgemeinbildende Technikunterricht dem Desinteresse und den fehlenden Technikkompetenzen, die aus den negativen Erfahrungen mit Technik resultieren, entgegenwirken. Dabei geht es, im Gegensatz zur gewerblich-technischen Bildung nicht darum, spezielles Fachwissen zu vermitteln, sondern um die Fähigkeit sich im technischen Alltag zurechtzufinden. Zwar wurde der Anspruch die Problemlösekompetenz im allgemeinbildenden Technikunterricht zu fördern in vielen Kernlehrplänen bereits festgehalten (z. B. Kernlehrplan Technik für die Sekundarstufe II in NRW, Kerncurriculum Technik für die Realschule in Niedersachsen), allerdings hat sich der Technikunterricht hierbei vor allem auf die Herstellung von Produkten fokussiert, weniger auf deren Nutzung (Sachs, 2005). Ein möglicher Grund hierfür wird u. a. in der vagen Spezifizierung eines Konstruktes Problemlösekompetenz im Umgang mit technischen Geräten und der damit verbundenen fehlenden Diagnosemöglichkeit gesehen (Stemmann & Lang, 2016). Denn das Ziel die Problemlösekompetenz einer Person zu fördern bedarf zunächst der Feststellung eines Förderbedarfs, um daraus entsprechende Maßnahmen ableiten und empirisch prüfen zu können. Die Entwicklung hierfür benötigter Messinstrumente setzt allerdings ein bereits formuliertes theoretisches Konstrukt sowie theoretische Annahmen darüber voraus, welche Anforderungen solche Problemstellungen an die problemlösende Person stellen (Jenßen, Dunekacke & Blömeke, 2015).

An diesem Punkt soll die vorliegende Arbeit anknüpfen, indem sie das Ziel hat, zunächst ein solches Konstrukt *Kompetenz im problemlösenden Umgang mit technischen Alltagsgeräten* auszuarbeiten und darauf aufbauend ein Testinstrument zu entwickeln und einer ersten empirischen Prüfung zu unterziehen. Zu diesem Zweck setzt sich die Arbeit aus fünf großen Teilen zusammen: dem Forschungsstand (Teil I), der Formulierung des theoretischen Konstruktes (Teil II), dem methodischen Design zur Überprüfung der Fragestellungen

(Teil III), den Forschungsergebnissen (Teil IV) sowie der anschließenden Interpretation dieser Ergebnisse (Teil V). Die Feingliederung der genannten Teile wird durch die folgenden Kapitel vorgenommen:

Eine erste Einführung in das Forschungsthema gibt das Kapitel 2, indem es zunächst den Begriff des Problems in seiner Bedeutung festlegt und anhand der Systematisierung und Charakterisierung verschiedener Problemtypen eine Einordnung von Problemen im Umgang mit technischen Systemen erlaubt. Für das unterschiedliche Problemlöseverhalten von Personen werden anschließend theoretische Erklärungen herangezogen, wobei ein erster Erklärungsansatz den Fokus auf potenziell schwierigkeitsbestimmende Merkmale und ein zweiter auf kognitionspsychologische Unterschiede seitens der problemlösenden Person legt.

Eine über kognitive Fähigkeiten hinausgehende Erklärung des Problemlöseprozesses macht die Einführung des Kompetenzbegriffs erforderlich. Die Festlegung des zugrundeliegenden Verständnisses von Kompetenz, bei dem auch der Situationskontext in die Betrachtungen einbezogen wird, erfolgt im Kapitel 3. Abschließend werden erste Möglichkeiten zur Operationalisierung von Problemlösekompetenzen vorgestellt. Daran anknüpfend beschreibt Kapitel 4 die Notwendigkeit den problemlösenden Umgang mit technischen Alltagsgeräten computerbasiert zu erfassen sowie die sich mit einer sorgfältig geplanten Computersimulation ergebenden Herausforderungen. Hierzu gehört die exakte Beschreibung der formalen Systemstruktur, für die verschiedene Möglichkeiten genannt und die in dieser Arbeit verwendeten finiten Automaten näher erläutert werden.

Neben den bereits vorgenommenen inhaltlichen Annäherungen an das zu untersuchende Konstrukt, erfolgt im Kapitel 5 auch eine strukturelle Annäherung, indem Zusammenhänge zu anderen Konstrukten in den Blick genommen werden. Als kontextübergreifende Einflussfaktoren werden Intelligenz, allgemeine Problemlösefähigkeit sowie Need for Cognition und problemlösebezogene Selbstwirksamkeitserwartungen und als kontextueller Einflussfaktor die bereits vorhandenen Erfahrungen im Umgang mit technischen Geräten analysiert.

Das Kompetenzkonstrukt *Problemlösen im Umgang mit technischen Alltagsgeräten* wird schließlich im Kapitel 6 unter Einbezug der bisherigen Betrachtungen aus dem Stand der Forschung ausgearbeitet. Aufbauend auf der Beschreibung technischer Kontexte wird das Konzept soziotechnischer Systeme vorgestellt, um den Umgang mit technischen Geräten besser operationalisieren zu können. Eine Annahme darüber, welche Merkmale diesen Umgang erschweren, wird differenziert nach System- und Situationsmerkmalen ausgeführt, bevor die Indikatoren, an denen sich ein kompetenter Umgang mit technischen Geräten zeigt, dargestellt werden. Den Abschluss des Kapitels bildet die Ausformulierung des theoretischen Kausalmodells, indem vorher beschriebene Zusammenhänge zu anderen Konstrukten in einem nomologischen Netzwerk dargestellt werden. Die Konkretisierung der Stellung des Konstruktes in diesem Theoriemodell dient der Prognose und Theorieprüfung im Rahmen der Konstruktvalidierung. Hierzu werden aus dem Theoriemodell in Kapitel 7 Hypothesen abgeleitet, die im weiteren Verlauf der Arbeit geprüft werden.

Für die Überprüfung der Hypothesen, die sich zum einen auf die Konstruktrepräsentation, die Konstruktstruktur sowie auf die Verortung in dem nomologischen Netzwerk beziehen,

sind Instrumente erforderlich, die eine Erfassung der Konstrukte möglich machen. Da bislang für den problemlösenden Umgang mit technischen Alltagsgeräten ein solches Messinstrument fehlte, wird dieses in Kapitel 8 auf Basis der bereits beschriebenen finiten Automaten zunächst entwickelt. Die zur Erfassung der anderen in dem nomologischen Netzwerk berücksichtigten Konstrukte eingesetzten Instrumente werden dann in Kapitel 9 beschrieben, ebenso wie die Rahmenbedingungen unter denen die Untersuchung durchgeführt wurde. Die statistischen Methoden, die benötigt werden, um die Überprüfung des angenommenen Theoriemodells anhand der formulierten Hypothesen vornehmen zu können, sind Inhalt des Kapitels 10, wobei sich die Reihenfolge ihrer Darstellung an der der aufgestellten Fragen und Hypothesen ausrichtet.

In dem sich an den Methodenteil anschließenden Ergebnisteil werden im Kapitel 11 zunächst die deskriptiven Befunde aller eingesetzten Testinstrumente berichtet und erste Zusammenhangsbetrachtungen vorgenommen, bevor Kapitel 12 die Güte des in dieser Arbeit entwickelten Testinstrumentes analysiert. Es folgen die Kapitel 13, 14 und 15, in denen die Prüfung der aufgestellten Hypothesen vorgenommen wird. Zunächst betrachtet Kapitel 13 die Konstruktrepräsentation, indem der Einfluss der als schwierigkeitsbestimmend angenommenen Merkmale auf die empirische Schwierigkeit analysiert wird. Kapitel 14 untersucht, inwieweit sich die angenommene Konstruktstruktur in den erhobenen Daten zeigt, indem die Dimensionalität des Konstruktes näher betrachtet wird. Abschließend erfolgt im Kapitel 15 die Überprüfung der kausalen Beziehungen zwischen den einzelnen Konstrukten in dem postulierten Theoriemodell.

Alle Befunde werden anschließend anhand der formulierten Fragestellungen im Kapitel 16 noch einmal zusammengefasst und unter Einbezug der betrachteten Theorie inhaltlich interpretiert. Erwartungswidrige Ergebnisse werden hier ebenfalls diskutiert. Zudem wird auf die, aufgrund vorhandener Limitierungen eingeschränkte Generalisierbarkeit der Befunde eingegangen und zuletzt auf die Bedeutung der Befunde für den allgemeinen Technikunterricht und die in ihm zu fördernde Problemlösekompetenz.

Teil I: Stand der Forschung

2 Probleme und Problemlösen

Das vorliegende Kapitel führt zunächst in das Forschungsthema ein, indem es grundlegende Begriffe erläutert. Für die Verortung der in dieser Arbeit verwendeten Probleme werden im Anschluss verschiedene Problemtypen vorgestellt und versucht diese systematisch zusammenzufassen. Es folgen Erklärungen für das unterschiedliche Problemlöseverhalten von Personen, wobei ein erster Erklärungsansatz den Fokus auf potenziell schwierighkeitsbestimmende Merkmale legt und ein zweiter auf kognitionspsychologische Unterschiede seitens der problemlösenden Personen.

2.1 Problemverständnis

Mit dem Problemlösen eng verbunden ist die Problementstehung, deren Ursache selten Inhalt wissenschaftlicher Forschung ist (vgl. Funke, 2003, S. 18). Das Vorhandensein von Bedürfnissen, Wünschen und Zielen motiviert Menschen zu handeln, um einen erwünschten Zustand zu erreichen (vgl. Ropohl, 1990, S. 115). Das zur Zielerreichung vollzogene Handeln ist dabei bewusst, planvoll und kann anhand von Handlungsergebnissen bewertet werden. Die Zielorientiertheit und das Bewusstsein der Aktivitäten bei der Zielverfolgung grenzt damit Handeln vom bloßen Verhalten ab (vgl. Betsch, Funke & Plessner, 2011, S. 140f.).

Oftmals lässt sich ein angestrebtes Ziel nicht ohne weiteres erreichen, weil die auf dem Weg dorthin liegenden Hindernisse die Zielerreichung zu einem Problem werden lassen (vgl. Funke, 2003, S. 18). Denn nach einer Definition von Duncker entsteht ein Problem genau dann, „[...] wenn ein Lebewesen ein Ziel hat und nicht ‚weiß‘, wie es dieses Ziel erreichen soll“ (Duncker, 1974, S. 1). Das Nichtwissen in der Definition von Duncker konkretisiert Dörner als das Fehlen von Mitteln, „[...] um den unerwünschten Zustand in den wünschenswerten Zustand zu überführen“ (Dörner, 1976, S. 10). Andere Autoren verwenden den Begriff der Barriere (Klix, 1971) oder den der Lücke in einem Handlungsplan (vgl. Funke, 2003, S. 25) zur Beschreibung des Problembegriffs.

Das Fehlen von sogenannten Operatoren, um Barrieren überwinden oder Lücken schließen zu können, grenzt in diesem Verständnis Probleme von Aufgaben ab. Bei einer Aufgabe sind die zur Bewältigung benötigten Methoden, Mittel und Regeln bekannt und die Lösung einer Aufgabe erfolgt mehr oder weniger routiniert (vgl. Dörner, 1976, S. 10; Funke, 2003, S. 25). Damit erfordert eine Aufgabe lediglich reproduktives Denken (Dörner, 1976, S. 10), während Probleme auch konstruktives Denken verlangen (Funke, 2003, S. 25). Ob es sich in einem konkreten Fall um eine Aufgabe oder um ein Problem handelt, ist gleichermaßen wie Ausgangs- und Zielzustand vom Subjekt abhängig, das entweder über die zur Zielerreichung benötigten Operatoren verfügt oder eben nicht (vgl. Hussy, 1984a, S. 114). Für Dörner hängt diese Unterscheidung von den Vorerfahrungen des Individuums ab (vgl. Dörner, 1976, S. 10). Was einem Novizen/einer Novizin kompliziert erscheint, stellt sich für den Experten/der Expertin oft als einfach dar (vgl. Gomez & Probst, 1995, S. 17f.).

Rollett (2008) merkt an, dass in konkreten inhaltsreichen (vorwissensinduzierenden) Situationen eine Unterscheidung zwischen Aufgabe und Problem zu Unklarheiten führt, denn die Expertise einer Person, für die jeder beliebige Wert auf einer kontinuierlichen Skala denkbar ist, kann sich auch nur auf Auszüge der Problemsituation beziehen (Rollett, 2008,

S. 20f.). Ist die fehlende Expertise einer Person eine Bedingung, um statt von einer Aufgabe von einem Problem zu sprechen, muss diese für jede Person in jeder Konstellation vorab bekannt sein. In diagnostischen Situationen, in denen die Problemlösekompetenz von Personen interessiert, werden Problemstellungen deshalb häufig vorwissensneutral konstruiert, um zu verhindern dass Probanden mit Vorwissen aus dem Bereich des Problems Vorteile haben (Greiff & Fischer, 2013a, S. 28). Die Konstruktion vorwissensneutraler Problemstellungen für die Diagnostik von Problemlösekompetenzen in spezifischen Kontexten birgt jedoch die Gefahr „[...] weniger authentische, d. h. weniger inhaltsvalide Aufgaben zu konstruieren [...]“ (Abele, Greiff, Gschwendtner et al., 2012, S. 366).

Um zur Erfassung der Problemlösekompetenz im Umgang mit alltäglichen technischen Geräten realitätsnahe Problemstellungen aus dem Alltag der Probanden konstruieren zu können, muss davon ausgegangen werden, dass Vorwissen seitens der Problemlöser vorhanden ist. In welchem Maße es für die Lösung der Probleme eingesetzt werden kann, ist vorab nicht abzusehen. Dementsprechend ist es hier nicht zweckmäßig zwischen den Begriffen Aufgabe und Problem zu unterscheiden, sondern in Anlehnung an Anderson (2007b) besser von Problemlösen als ein zielgerichtetes Verhalten zu sprechen, bei dem ein Ziel oftmals in Teilziele aufzuteilen ist, um Operatoren anzuwenden, die eine Zielerreichung ermöglichen (Anderson, 2007b, S. 292).

Unabhängig von dem Wissen und Können einer problemlösenden Person, unterscheiden sich Probleme in der Schwierigkeit ihrer Lösung, weshalb hier der Ansatz von Rollett übernommen wird, die Komplexität eines Problems (hier ein Merkmal der Schwierigkeit) über die objektive Aufgabenstruktur zu definieren und nicht in Abhängigkeit der Kompetenz einer Personen (vgl. Rollett, 2008, S. 22). Mit dem Ziel Probleme im Umgang mit technischen Alltagsgeräten objektiv beschreiben und in bestehende Problemklassifikationen einordnen zu können, erfolgt im nächsten Schritt die Betrachtung relevanter Problemtaxonomien und anschließend die der Problemmerkmale.

2.2 Klassifikation von Problemen (Taxonomien)

Der Wunsch Ordnung in die Vielzahl der denkbaren Problemsituationen zu bringen, führte zu der Suche nach Unterscheidungsmerkmalen und zu dem Versuch Probleme anhand dieser Merkmale in Typen einzuteilen (Funke, 2003, S. 29). Die Systematisierung von Problemen kann dabei auf Grundlage verschiedener Kriterien erfolgen, womit deutlich wird, dass es keine einheitliche, geschweige eine verbindliche Problemklassifikation gibt (vgl. Arbinger, 2015, S. 9). Probleme werden beispielsweise nach der:

- Art der Barrieren, die zwischen dem Anfangs- und dem Zielzustand liegen,
- Bedeutung von Wissen für die Problemlösung,
- Domäne, aus der sie stammen,
- Komplexität von Problemen differenziert (Arbinger, 2015, S. 9; Betsch, Funke & Plessner, 2011, S. 151f.).

Die Unterscheidung von Problemen anhand der Arten von Barrieren, die eine Überführung des Ausgangs- in den Zielzustand verhindern, nimmt Dörner vor, indem er Barrieren danach voneinander abgrenzt, wie viel Klarheit über die Ziele besteht und wie bekannt die Mittel zur

Zielerreichung sind (Dörner, 1976, S. 14). Sind sowohl der Zielzustand als auch die Anzahl der Operatoren bekannt, mit deren Hilfe sich eine Zustandsänderung herbeiführen lässt, liegt eine *Interpolationsbarriere* vor. Sie lässt sich beseitigen, indem die richtige Kombination aus den Operatoren in der richtigen Reihenfolge angewendet werden (Dörner, 1976, S. 12). Eine *synthetische Barriere* liegt vor, wenn zwar über den Zielzustand Klarheit besteht, aber die Operatoren zumindest teilweise unbekannt sind, also erst im Prozess des Problemlösens synthetisiert werden müssen. Eine dritte von Dörner beschriebene Barriere besteht, wenn die zur Zielerreichung benötigten Operatoren bekannt sind, aber bezüglich des Zielzustandes Unklarheiten vorliegen. Der von Dörner gewählte Name *dialektische Barriere* begründet sich in dem dialektischen Prozess, in dem die Problemlösung gefunden wird (Dörner, 1976, S. 13). Er zeichnet sich durch die Entwicklung, Prüfung, Revidierung und Verwerfung alternativer Vorstellungen über den Zielzustand aus, bis er in einer zufriedenstellenden Lösung endet (vgl. Arbinger, 2015, S. 11). In realen Problemsituationen treten diese Barrieren oft nicht in Reinform auf, sondern häufig kombiniert (vgl. Kotkamp, 1999, S. 16; Arbinger, 2015, S. 11).

Die Problemtypen von Dörner werden in der Literatur auch zur Unterscheidung von technischen Problemen herangezogen. Abele, Greiff, Gschwendtner et al. (2012) ordnen technische Probleme, aufgrund der von ihnen als transparent angenommenen Zielstellungen, Problemen mit Interpolations- und Synthesbarrieren zu (vgl. Abele, Greiff, Gschwendtner et al., 2012, S. 366). Krause (1982a), der Probleme aus dem Bereich der Arbeits- und Ingenieurpsychologie betrachtet, klassifiziert diese bei gleichzeitiger Betrachtung der Problemtypen nach Dörner und der Grundlage des Problemraumkonzeptes. Der Problemraum wird bei Krause repräsentiert durch einen Graphen, der alle möglichen Handlungsabläufe, die zur Problemlösung realisiert werden können, darstellt. Die Knotenpunkte des Graphen stellen die Zustände in dem Problemraum dar, während die Operatoren die Übergänge zwischen den Zuständen wiedergeben (Krause, 1982a, S. 18). Die Unterscheidung der Probleme erfolgt nach der Abgeschlossenheit von Zustands- und Operatormenge. Eine Zustandsmenge sieht er dann als abgeschlossen, wenn alle Zustände zur Problemlösung mit der Problemstellung gegeben oder durch sie ableitbar sind; gleiches gilt für die Operatormenge (Krause, 1982a, S. 20). Tabelle 2.1 gibt einen Ausschnitt der Klassifikationsmatrix von Problemen im Rahmen der Arbeits- und Ingenieurpsychologie wieder.

Tabelle 2.1: Klassifikation von Problemen der Arbeits- und Ingenieurpsychologie (nach Krause, 1982a)

Operatormenge	Zustandsmenge	
	Abgeschlossen	Nicht abgeschlossen
Abgeschlossen	Problemsituation I Steuerungsprobleme	Problemsituation III Konstruktionsprobleme
Nicht abgeschlossen	Problemsituation II Fertigungsprobleme	Problemsituation IV Allgemeine Entwurfsprobleme

Steuerungsprobleme sind ein Beispiel für die Problemsituation I, in der der Problemraum vollständig abgeschlossen ist. Sowohl die Operatoren als auch die Zustände sind durch die Aufgabenstellung gegeben oder können durch sie hergeleitet werden. Die gegebenen Operatoren müssen nur in der richtigen Kombination, analog zu den Problemen mit Interpolationsbarriere nach Dörner angewandt werden, um das Problem zu lösen (Krause,

1982a, S. 22). Die Problemsituation II, für die exemplarisch Fertigungsprobleme genannt werden, zeichnen sich durch das Fehlen von Operatoren aus. Diese müssen entweder erst entdeckt oder aus einer größeren Operatormenge zusammengestellt werden. Hierzu passen die von Dörner genannten Probleme mit Synthesbarriere (Krause, 1982a, S. 23). Konstruktionsprobleme repräsentieren die Problemsituation III, die durch einen nicht klaren Zielzustand und damit durch eine nicht abgeschlossene Zustandsmenge gekennzeichnet ist. Die Zustandsmenge muss also zunächst gesucht werden, während die Operatormenge durch den Prozess der klaren Begriffsdefinitionen als abgeschlossen betrachtet ist (Krause, 1982a, S. 24). Bei Entwurfsproblemen, die ein Beispiel der letzten Problemkategorie IV darstellen, ist weder die Zustandsmenge noch die Operatormenge abgeschlossen und gelten nach Krause als am schwierigsten zu bewältigen (Krause, 1982a, S. 24).

Der Verzicht auf die Dimension *Klarheit über den Zielzustand* aus der Taxonomie von Dörner und im Gegenzug die Hinzunahme der *Abgeschlossenheit der Zustandsmenge* für die Problemklassifizierung ist m. E. kritisch zu betrachten. Zwar argumentiert Krause, dass der Zielzustand ein Element der Zustandsmenge ist und wenn diese abgeschlossen ist, dann muss auch Klarheit über den Zielzustand im Sinne von Dörner bestehen (Krause, 1982a, S. 20f.). Der umgekehrte Schluss, dass bei bestehender Zielklarheit von einer abgeschlossenen Zustandsmenge auszugehen ist, ist jedoch nicht zulässig. Als problematisch sind außerdem die von ihm vorgenommenen Einordnungen von Problemtypen einzuschätzen. Ein Fertigungsproblem zeichnet sich nicht zwingend dadurch aus, dass die Zustandsmenge abgeschlossen ist, denn Werkstücke lassen sich oftmals mit verschiedenen Verfahren herstellen und je nach Wahl des entsprechenden Operators ist die Zustandsmenge eine andere. Die Problematik wird vor allem durch die Definition der Abgeschlossenheit der Zustands- und Operatormenge verursacht, nach der Abgeschlossenheit vorliegt, wenn alle Zustände und Operatoren in der Aufgabenstellung gegeben sind oder durch sie abgeleitet werden können. Im Fall der Steuerungsprobleme, die Krause für die Problemsituation I angibt, ist von einer vollständigen Abgeschlossenheit sowohl der Operator- als auch der Zustandsmenge wohl nur bei sehr einfachen Steuerungsaufgaben auszugehen, nicht aber bei Steuerungsproblemen, die intransparent oder sogar dynamisch sind. Schlussfolgernd lassen sich nur einzelne klar umrissene Probleme oder sehr eng gefasste Gruppen von Problemen diesen Klassen zuordnen, nicht aber ganze Problembereiche.

Neben der genannten Unterscheidung von Problemen anhand der vorliegenden Barriere, lassen sich Probleme außerdem mittels der Bedeutung des Wissens, das für die Bewältigung von Aufgaben und zur Lösung von Problemen erforderlich ist, voneinander abgrenzen. Wie bereits im Zusammenhang der Begriffsdefinition von Aufgabe und Problem angesprochen, gibt es Probleme für deren Lösungen bereits vorhandenes Wissen keine Rolle spielt. Hierzu werden den Probanden größtenteils künstliche Situationen vorgelegt, um die Möglichkeiten, bei der Bearbeitung der Problemstellungen auf bisherige Erfahrungen oder Wissen zurückzugreifen, zu minimieren (vgl. Arbinger, 2015, S. 2). Hinter dem Ziel, die beim Problemlösen stattfindenden kognitiven Prozesse unabhängig von Vorwissenseinflüssen erfassen zu können, steht die Vorstellung, dass es allgemeine Prozesse des Problemlösens gibt, die von der Domäne, aus der sie stammen unabhängig sind (vgl. Arbinger, 2015, S. 14).

Die Mehrzahl an Problemen, die im privaten, schulischen oder beruflichen Alltag zu bewältigen sind, zeichnen sich jedoch dadurch aus, dass sie nur mit entsprechendem Wissen zu lösen sind (Opwis, Beller, Spada et al., 2006, S. 219). Seitdem in zahlreichen Untersuchungen gezeigt werden konnte, dass sich Experten von Novizen nicht durch allgemeine Problemlösefähigkeiten, sondern durch ihr umfangreiches domänenspezifisches Wissen auszeichnen (Chi, Glaser & Rees, 1982, S. 71), klammert die Problemlöseforschung das Wissen einer Person nicht kategorisch aus, sondern berücksichtigt es als Strukturelement des Problemlösens (Arbinger, 2015, S. 2). Im Interesse der Erforschung realitätsnaher Probleme, werden diese in einen semantischen Kontext eingebettet, der bei den problemlösenden Personen Vorstellungen und Erwartungen auslöst und somit zu einer Beeinflussung des Problemlöseprozesses führt (vgl. Hesse, 1982, S. 64). Vorhandene Erfahrungen oder Wissen können dabei auf der einen Seite die Problemlöseleistung positiv beeinflussen (Beckmann, 1994, S. 63), indem neue Informationen besser in die vorhandene kognitive Struktur eingebunden werden (Krause, 1982b, S. 149). Auf der anderen Seite kann Vorwissen aber auch für den weiteren Wissenserwerb im Problemlöseprozess hinderlich sein, wenn etwa die neuen Informationen mit dem bestehenden Wissen unvereinbar sind (vgl. Wirth, 2004; Buchner & Funke, 1991).

Problemstellungen, die Wissen voraussetzen, welches wenig fachspezifisch¹ ist, erlauben den Ausgleich eventuell vorhandener Vorwissensdefizite noch im Problemlöseprozess. Je (fach-)spezifischer und umfangreicher hingegen das für die Problemlösung benötigte Wissen ist, desto schwieriger wird es, sich das Wissen erst während der Problembearbeitung anzueignen (Stemmann & Lang, 2014). Entsprechend eng verknüpft mit der Unterscheidung von Problemen hinsichtlich der Bedeutung des Vorwissens, ist die Differenzierung nach der Domäne, aus der die Probleme stammen (vgl. Arbinger, 2015, S. 12; Neber, 1987, S. 23). *Fachübergreifendes Problemlösen*, für das synonym der Begriff *allgemeines Problemlösen* verwendet wird (siehe z. B. Rüppel, Hinnersmann & Wiegand, 1987), benötigt Wissen, das keinem konkreten Fachbereich zugeordnet werden kann (vgl. Leutner, Funke, Klieme et al., 2005, S. 12). Solche Problemstellungen kommen aus dem weitgefassten Kontext von Privat-Persönlichem, Arbeit und Freizeit sowie Gesellschaft (vgl. Leutner, Klieme, Meyer et al., 2004, S. 148). Im Rahmen der PISA-Studie 2012 wurden solche Problemstellungen eingesetzt, um die fachübergreifende Problemlösefähigkeit von Schülerinnen und Schüler zu messen. Die Herausforderung einer exemplarisch zu nennenden Aufgabe bestand darin, den zeitlich günstigsten Weg in einem gegebenen Verkehrsnetz ausfindig zu machen (OECD, 2014a, S. 41). Demgegenüber stehen *fachspezifische Problemstellungen*, die von der problemlösenden Person umfangreiches Domänenwissen voraussetzen, das zur Problemlösung angewendet werden muss und ohne das eine Problemlösung nicht möglich ist. Sie entstammen zumeist aus beruflichen Kontexten, wie der Diagnose von Fehlern in der Automatisierungstechnik (Walker, Link & Nickolaus, 2015) oder der Diagnose von Störungen in einem Kraftfahrzeug (Abele, Walker & Nickolaus, 2014). Zwischen den hier beschriebenen Extremen von kontextfreien Problemen und den Problemen aus spezifischen Domänen sind alle graduellen Abstufungen denkbar (Arbinger, 2015, S. 12). Probleme im Umgang mit technischen Geräten, die aus dem Alltag mehr oder weniger bekannt sind, lassen sich dem

¹ Der Begriff *fachspezifisch* wird in dieser Arbeit synonym zu dem Begriff *domänenspezifisch* verwendet. Eine Auseinandersetzung mit den Begriffen *Domäne* und *Kontext* erfolgt in Kapitel 5.2.1.

spezifischen Kontext *Bedienung technischer Geräte* zuordnen. Im Gegensatz zu vollständig kontextfreien Problemsituation, kann seitens der problemlösenden Personen bereits Bedienungswissen vorhanden sein, das bei der Problembearbeitung hilft. Abgrenzend zu stark fachspezifischen Problemen, ist dieses Wissen aber nicht obligatorisch, da Personen fehlendes Wissen auch im Umgang mit den technischen Geräten explorativ erwerben können¹. Auf welchem Weg Wissen während des Problemlöseprozesses erworben und wie es dann zur Lösung konkreter Problemstellungen eingesetzt werden kann, wird in Kapitel 2.4 vorgestellt.

Eine weitere für diese Arbeit bedeutsame Differenzierung von Problemen ist die zwischen einfachem und komplexem Problemlösen. Die auf Dörner zurückgehende Bezeichnung komplexer Probleme beschreibt Situationen, in denen Personen mit computersimulierten Systemen interagieren müssen, um sie der Zielsetzung entsprechend zu steuern (vgl. Dörner & Kreuzig, 1983, S. 188). Zu den Problemen, die von Probanden in diesem Zusammenhang zu bewältigen sind, zählen beispielsweise: eine Stadt als Bürgermeister regieren, eine Schneiderwerkstatt führen, eine Mondfähre landen oder einem Entwicklungsland helfen (Müller, 1993, S. 1). Die Literatur nennt eine Reihe von Merkmalen, die komplexe Probleme von einfachen Problemen abgrenzen (Dörner, 1976, S. 18; Betsch, Funke & Plessner, 2011, S. 155). Zu den am häufigsten genannten Merkmalen, die derartige Probleme charakterisieren, zählen:

- Komplexität (Anzahl zu berücksichtigender Variablen),
- Vernetztheit (gegenseitige Variablenabhängigkeit),
- Dynamik (Veränderung von Variablen über die Zeit),
- Intransparenz (Zugänglichkeit von Informationen),
- Polytelie (Anzahl zu erreichender Ziele).²

Von den Anforderungen an die problemlösende Person, die sich aus diesen Merkmalen ergeben, wird eine größere Nähe zu den Anforderungen durch reale Probleme aus dem Alltag vermutet, als von den Anforderungen bei einfachen Problemen (vgl. Funke, 1986, S. 1). Das Verhalten der Personen im Umgang mit diesen komplexen Szenarien offenbarte allerdings große Schwierigkeiten (vgl. Schaub, 1993, S. 52) und führte sogar zu den Vermutungen, dass grundlegende Fähigkeiten für die Bewältigung der Anforderungen in komplexen Problemsituationen fehlen (vgl. Funke, 1986, S. 6; Dörner, 1981, S. 178). Das in den Studien beobachtete Scheitern der Versuchspersonen ist für Müller (1993) aufgrund der hohen Ausprägung von Komplexität vorhersehbar. Als weiteren Grund für den Misserfolg der Probanden gibt Müller die fehlenden Erfahrungen der Personen mit solchen Szenarien an, obwohl diese eigentlich Ausschnitte der Realität darstellen sollen (Müller, 1993, S. 4). Die von Müller auf die Komplexitätsausprägung bezogene Kritik wurde bereits von Hussy (1984b) angebracht. Ein Szenario, in dem gleichzeitig alle Merkmale der Problemkomplexität extrem ausgeprägt sind, schafft zwangsläufig Frustrationen (Hussy, 1984b, S. 11). Für ihn stellen die Merkmale komplexer Probleme Determinanten der Problemschwierigkeit dar, deren Ausprägungen er in seiner Studie systematisch variiert (Hussy, 1985).

¹ Eine genaue Beschreibung der hier betrachteten Problemtypen erfolgt im Kapitel 6.

² Kapitel 2.3 setzt sich mit diesen Merkmalen weiter auseinander.

Als weitere Kritikpunkte an den früheren Studien führt Funke (1984) die fehlenden Angaben zur Reliabilität, die fehlende Objektivität bei der Bewertung der Lösungsgüte (durch das Fehlen klar definierter Ziele) sowie die fehlende Operationalisierung der Komplexitätsmerkmale an (Funke, 1984). Mit dem Ziel die Schwächen dieser frühen Arbeiten zum komplexen Problemlösen überwinden zu wollen, entwickelt er Problemszenarien auf Basis formaler Modelle¹ (vgl. Funke, 1999, S. 195). Diese haben die präzise Gestaltung der Systemstruktur zum Ziel, die Einbettung in einen mehr oder weniger realitätsnahen Kontext erfolgt erst im Anschluss oder wird auch ganz weggelassen (Funke, 2003, S. 154f.). Komplexe Probleme, für die eine formale Struktur festgelegt wurde, lassen sich anhand dieser Struktur miteinander vergleichen.

Nicht behoben wurde bis zu diesem Zeitpunkt das Problem der mangelnden Reliabilität. Sowohl in Studien mit realitätsnahen Mikrowelten nach Dörner als auch in Studien mit formalisierten Problemszenarien bearbeiten die Probanden lediglich ein einzelnes Szenario über einen längeren Zeitraum. Auch die Bearbeitung von Teilsystemen oder die Beantwortung mehrerer Fragen zu dem bearbeiteten System können nicht als unabhängige Items betrachtet werden (vgl. Greiff & Funke, 2010, S. 218). Die mittels Retestungen gewonnenen Aussagen über die Reliabilität solcher One-Item-Tests sind aufgrund von Übungs- und Erinnerungseffekten als ebenso kritisch zu betrachten (Greiff, 2012, S. 32). Als Lösung schlägt Greiff einen Ansatz vor, in dem mehrere kleinere Systeme eingesetzt werden. Diese beruhen zwar auch auf formalen Modellen, sind aber so einfach wie möglich gestaltet. Um die in seinem Ansatz entscheidende Komplexitätsreduzierung zu betonen, nennt Greiff seinen Ansatz der minimal komplexen Szenarien *MicroDYN* (Greiff, 2012, S. 78).

Vor dem Hintergrund dieser stärker abstrahierten Problemszenarien merkt Arbinger kritisch an, dass nicht alle Probleme komplex sind, die sich so nennen (Arbinger, 2015, S. 133). Der Begriff komplexe Probleme ist für einige der präsentierten Szenarien unpassend. Die Szenarien, die auf Basis formaler Modelle konstruiert werden, weisen außerdem oftmals wenig Nähe zur Realität auf. Zwar gibt es häufig irgendeine Art semantischer Einbettung, diese stellen jedoch kaum naturgetreue Abbildungen dar, was von den Autoren auch gar nicht beabsichtigt wird. Müller kritisiert in diesem Zusammenhang jedoch, dass auf Basis der Ergebnisse, die in den Studien mit wenig realitätsnahen Szenarien gewonnen werden, allgemein auf Problemlösefähigkeit geschlossen wird (Müller, 1993, S. 4f.).

Die Verwendung der Begrifflichkeiten *einfache* und *komplexe Probleme* bzw. *einfaches* und *komplexes Problemlösen* für die Unterscheidung von Problemen ist also insofern schwierig, als dass nicht eindeutig bestimmt werden kann, wann ein Problem ein einfaches Problem und wann ein komplexes Problem ist. Weil die Natur des zu lösenden Problems durch diese Bezeichnung unklar bleibt (vgl. Betsch, Funke & Plessner, 2011, S. 154), schlägt Funke (2003) vor, besser von *statischen* und *dynamischen* Problemen zu sprechen (Funke, 2003, S. 107), wobei er selbst die Begriffe nicht konsequent verwendet (z. B. Funke, 2006a, 2010). Bezogen auf die Tätigkeit des Problemlösens wird kongruierend zwischen dem *analytischen* und dem *dynamischen Problemlösen* unterschieden.

¹ Simulationen auf Basis formaler Modelle stellen die Grundlage der, in dieser Arbeit eingesetzten Testinstrumente dar und werden in Kapitel 4.3 weiter präzisiert.

Analytisches Problemlösen ist durch seine transparente Ausgangslage gekennzeichnet, die sich, ebenso wie die Problemstellung selbst, während des Problemlöseprozesses nicht verändert (vgl. Leutner, Funke, Klieme et al., 2005, S. 17), also statisch ist. Eine transparente Ausgangslage zeichnet sich dadurch aus, dass alle relevanten Informationen gegeben bzw. durch schlussfolgerndes Denken abgeleitet werden können (vgl. ebd.). Im Unterschied zum analytischen Problemlösen muss die problemlösende Person beim dynamischen Problemlösen aktiv in die Problemsituation eingreifen, um fehlende relevante Informationen zu erhalten (vgl. Leutner, Klieme, Meyer et al., 2004, S. 162f.). Da sich der Problemzustand während des Eingreifens verändert, können Informationen nur im zeitlichen Verlauf generiert werden. Probleme lassen sich anhand des Merkmals der interaktiven Informationsbeschaffung eindeutiger zuordnen als es durch das Merkmal der Komplexität erfolgen kann, wobei die Begriffe *komplexes Problemlösen* und *dynamisches Problemlösen* in der Literatur synonym verwendet werden (vgl. Leutner, Fleischer, Wirth et al., 2012, S. 35).

Problematischer als die Bedeutungsgleichsetzung von komplexen und dynamischen Problemlösen erscheint die Verwendung der Synonyme analytisches Problemlösen für fachspezifische Probleme und dynamisches Problemlösen für fachübergreifende (allgemeine) Probleme (Greiff, Holt & Funke, 2013, S. 72). Wenn beispielsweise Abele, Greiff, Gschwendtner et al. (2012) schreiben, „[...] dass dynamisches Problemlösen eine zentrale Facette allgemeinen Problemlösens darstellt“ (Abele, Greiff, Gschwendtner et al., 2012, S. 368), dann ist die Aussage zwar nicht falsch, berücksichtigt aber auch nicht, dass dynamisches Problemlösen generell eine Facette von Problemlösen ist und nicht nur von allgemeinem Problemlösen. Gerade in beruflichen Situationen müssen Informationen durch aktives Eingreifen in die Problemsituation zunächst generiert werden, damit Problemlösen möglich wird.

Die Merkmale Komplexität, Vernetztheit, Dynamik, Intransparenz und Polytelie, die Dörner und andere zur Charakterisierung dynamischer Probleme verwenden, werden in Anlehnung an Hussy als Merkmale gesehen, die die Schwierigkeit eines Problems (unabhängig von der problemlösenden Person) beschreiben. Deswegen geht der nächste Abschnitt auf diese Merkmale nun weiter ein.

2.3 Schwierigkeitsbestimmende Problemmerkmale

Interindividuelle Unterschiede im Problemlöseverhalten von Personen lassen sich in der Regel durch die unterschiedlich ausgeprägten Problemlösekompetenzen erklären. Während kompetente Problemlöser und Problemlöserinnen ein Problem lösen können, scheitern Personen mit einer geringen Problemlösekompetenz an der Lösung desselben Problems. Darüber hinaus gibt es intraindividuelle Unterschiede bei der erfolgreichen Bearbeitung unterschiedlicher Probleme. So ist es denkbar, dass eine Person gewisse Probleme lösen kann, andere jedoch nicht. Probleme unterscheiden sich also nicht nur anhand der im vorherigen Abschnitt genannten Taxonomien, sondern auch in der Schwierigkeit ihrer Lösung. Schwierige Probleme lassen sich daran erkennen, dass sie nur von wenigen (besonders kompetenten) Personen gelöst werden und dass für ihre Lösung generell eine längere Bearbeitungszeit benötigt wird (Freunberger & Itzlinger-Bruneforth, 2013, S. 6).

Die Merkmale eines Problems oder einer Problemsituation, die einen Einfluss auf die Schwierigkeit bei der Bearbeitung haben, werden im Folgenden schwierigkeitsbestimmende Merkmale oder kurz Problemmerkmale genannt. Die Kenntnis über die Merkmale, die ein Problem schwierig machen, ist aus verschiedenen Gründen von Bedeutung. (1) Gibt es theoretisch abgeleitete Merkmale, deren Einflüsse auf die Schwierigkeit vorab formuliert wurden, dient ihre Überprüfung als Beitrag zur Konstruktvalidierung (Hartig & Frey, 2012, S. 44; Hartig, Frey & Jude, 2012, S. 161f.). (2) In bereits vorliegenden empirischen Daten können die als schwierigkeitsbestimmend angenommenen Merkmale dazu beitragen die Varianz theoriegeleitet zu erklären (Böhme, 2011, S. 82). (3) Für anstehende Testentwicklungen oder -revisionen lassen sich als relevant erwiesene schwierigkeitsbestimmende Merkmale gezielt variieren, um Items unterschiedlicher Schwierigkeit zu konstruieren (ebd.). (4) Aus individualdiagnostischer Perspektive dienen die schwierigkeitsbestimmenden Merkmale zur Definition von Grenzen in Kompetenzstufenmodellen (Hartig, 2007).

Mit dem Ziel das Kompetenzkonstrukt *Problemlösen im Umgang mit technischen Geräten* zu präzisieren, werden in der vorliegenden Arbeit zunächst potenziell schwierigkeitsbestimmende Merkmale aus der Literatur extrahiert und versucht deren Einfluss auf die Schwierigkeit im Umgang mit Problemen theoretisch zu begründen. Die Annahmen über ihren Schwierigkeitseinfluss werden anschließend empirisch geprüft. Merkmale, die ein Problem oder eine Aufgabe schwierig machen können, finden sich in der Literatur verschiedenster Fachbereiche. Viele der dort zu findenden Merkmale sind jedoch inhaltlich definiert und können demnach nicht auf die Schwierigkeit im problemlösenden Umgang mit technischen Geräten übertragen werden. Es kommt aus diesem Grund also nicht in Frage, in der Literatur nach konkreten Merkmalen wie etwa die *mathematischen Anforderungen in einem Kompetenztest zur Technischen Mechanik* (Dammann, Behrendt, Ștefănică et al., 2016) zu suchen, sondern zunächst nach Oberkategorien von Merkmalen, die dann für den Umgang mit technischen Systemen konkretisiert werden können. Als schwierigkeitsbestimmend können dann nur solche Merkmale statistisch untersucht werden, deren Ausprägung in der zu untersuchenden Gruppe von Aufgaben eine ausreichende Varianz aufweist. Merkmale, die zwar generell eine Schwierigkeit darstellen, wie die Interaktion mit dem Testmaterial bei technischen Geräten, aber bei allen Aufgaben gleich sind, führen zu keinem weiteren Erkenntnisgewinn.

Je nach Betrachtungsweise können Aufgabenmerkmale vielfältig systematisiert werden. Eine rationale, objektive Perspektive differenziert Aufgaben vor allem anhand formaler und inhaltlicher Aufgabenmerkmale, während aus psychologischer Perspektive die benötigten kognitiven Verarbeitungsprozesse als Merkmal einer Aufgabe in den Fokus treten (vgl. Kauertz, 2008, S. 19f.). Formale Aufgabenmerkmale beziehen sich im Wesentlichen auf die Oberflächenmerkmale wie das Antwortformat einer Aufgabe oder die Länge des Aufgabentextes (Prenzel, Häußler, Rost et al., 2002, S. 125) und lassen sich in der Regel abzählen (Kauertz, 2008, S. 21). Inhaltliche Merkmale verweisen auf die semantische Einbettung der Aufgabe, die sich beispielweise in ihrer Vorwissenskompatibilität, ihrer inhaltlichen Komplexität oder der Art des benötigten Wissens (z. B. deklaratives oder prozedurales Wissen) unterscheiden kann (Schumann & Eberle, 2011, S. 79f.). Während sich die Ausprägungen der formalen und inhaltlichen Aufgabenmerkmalen annähernd objektiv bestimmen lassen, können die während der Problembearbeitung stattfindenden kognitiven

Prozesse nur interpretativ geschätzt werden (vgl. Florian, Sandmann & Schmiemann, 2014, S. 178).

Als Grundlage zur Klassifizierung kognitiver Prozesse bei der Bearbeitung von Testaufgaben wird oftmals die zum Verständnis der Denkprozesse des Lernens entwickelte Taxonomie von Bloom verwendet, in der er Lernziele hierarchisch nach ihrer Komplexität wie folgt ordnet: Wissen, Verstehen, Anwendung, Analyse, Synthese und Bewertung (Bloom, 1976, S. 31; Anderson & Krathwohl, 2001, S. 31). Aufgaben, deren Lösung kognitive Prozesse höherer Komplexität erfordern, sollten demnach schwieriger sein, als solche, bei denen weniger komplexe Prozesse ausreichen (Hartig, Frey & Jude, 2012, S. 161; Neber, 1987, S. 72). Aus fachdidaktischer Sicht erscheinen kognitive Prozesse als schwierigkeitsbestimmende Merkmale geeignet, wenn ihre Ausprägungen beispielsweise in den Items eines Kompetenztests ausreichend breit streuen. So sollten demnach in einem Test sowohl Aufgaben vorliegen, in denen ausschließlich Wissen abgefragt wird als auch Items, die Begründungsleistungen oder Beurteilungen einfordern (vgl. Gschwendtner, 2008, S. 108). Neben dem Problem der Bewertung der tatsächlich für die Lösung einer konkreten Aufgabe notwendigen kognitiven Prozesse, kommt die heterogene Befundlage ihres Beitrages zur Aufklärung der Schwierigkeit als kritisch zu betrachtender Punkt hinzu. So können einzelne Studien den Einfluss der kognitiven Prozesse auf die Aufgabenschwierigkeit in Fachwissenstests zeigen (vgl. Gschwendtner, Geißel & Nickolaus, 2010, S. 266), in anderen Studien lassen sich die Schwierigkeitsabstufungen hingegen nicht konsistent nachweisen (Schabram, 2007).

Im Hinblick auf die Ermittlung von Schwierigkeitsursachen im interaktiven Umgang mit technischen Systemen, in dem mindestens *Anwenden* als kognitiver Prozess gefordert ist, erscheint eine nähere Betrachtung der Merkmale aus der Forschung zum dynamischen (komplexen) Problemlösen lohnenswert. Eine der ersten Arbeiten, in der die Größen, die die Schwierigkeit dynamischer Problemsituationen potenziell beeinflussen, systematisiert wurden, ist nach eigenen Aussagen, die von Hussy (1984b, S. 3). Prinzipiell klassiert er die Determinanten der Problemschwierigkeit in *Problemmerkmale* und *Personenmerkmale*. Die Problemmerkmale werden dabei noch weiter in *Problemumfang*, *Problemkomplexität* sowie die von Dörner unterschiedenen *Barrieretypen* gegliedert (Abbildung 2.1). Die meisten der ebenfalls von Dörner genannten Merkmale komplexer Probleme (*Variablenanzahl*, *Variablenvernetzung*, *Transparenz* und *Eigendynamik*) werden von Hussy der Kategorie Problemkomplexität zugeordnet (ebd.). Die in seinem Kategoriensystem aufgeführten und weiter nach *Fakten-* und *Operationswissen* differenzierten Personenmerkmale beziehen sich zum einen auf den Umfang des zur Problemlösung notwendigen Wissens, seiner Verfügbarkeit sowie seiner Organisation im Gedächtnis (Hussy, 1984a, S. 136f.).

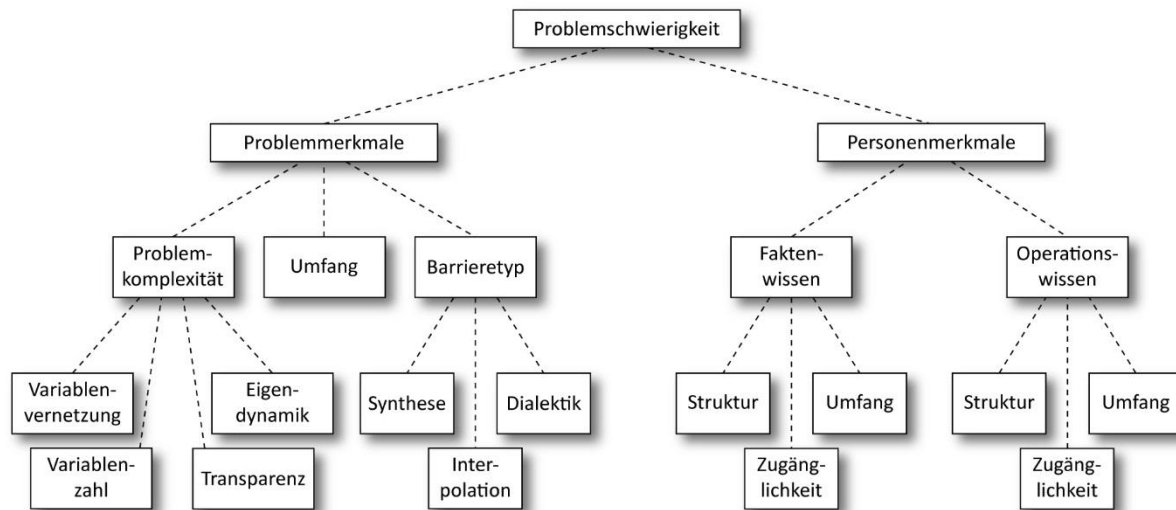


Abbildung 2.1: Determinanten der Problemschwierigkeit dynamischer (komplexer) Probleme (Hussy, 1984b, S. 4)

Im Rahmen seiner eigenen Untersuchung mit einem Problemszenario, in dem eine Mondlandung simuliert wird, variiert Hussy die Merkmale der Problemkomplexität (ausgenommen die Eigendynamik), um ihren Einfluss auf die Problemschwierigkeit zu prüfen (Hussy, 1984b, S. 12). Die Variation der Ausprägungen der Merkmale *Variablenanzahl*, *Variablenvernetzung* und *Transparenz* werden dabei jeweils in zwei Stufen vorgenommen (ebd., S. 13):

- Variablenanzahl (3 vs. 5 Variablen)
- Transparenz (transparente Version vs. implementierte Intransparenz)
- Variablenvernetzung (lineare Relationen vs. nichtlineare Relationen)

Mithilfe einer dreifaktoriellen Varianzanalyse zeigte Hussy dass die drei Merkmale insgesamt knapp die Hälfte (42 %) der Varianz in der Problemschwierigkeit erklären (ebd., S. 16). Die einzelnen Merkmale unterscheiden sich in ihrem Einfluss dabei nicht stark voneinander (Variablenanzahl und Variablenvernetzung tragen zu je 16 % und die Transparenz zu 10 % zur Varianzaufklärung bei) (ebd.). Wagener (2001, S. 33) kritisiert, dass sich die Gültigkeit der Taxonomie von Hussy nur auf sehr einfache dynamische Systeme beziehen lässt und sie keine Aussagen darüber enthält, wie die einzelnen Merkmale einzeln oder miteinander die Schwierigkeit beeinflussen.

Eine ähnliche Taxonomie findet sich bei Funke (1990, S. 145), der die Determinanten des Umgangs mit dynamischen Systemen einerseits in *Personenmerkmale* und andererseits in *System- und Situationsmerkmale* gruppiert (Abbildung 2.2). Anders als Hussy trennt Funke die Problemmerkmale außerdem in Merkmale, die das eigentliche System betreffen (*Systemmerkmale*) und solche der Situation, in der Personen mit diesem System umgehen (*Situationsmerkmale*). Unter der Rubrik *Personenmerkmale* fasst Funke alle Merkmale, die seitens einer Person in die Problemsituation eingebracht werden. Hierunter fallen kognitive Merkmale wie z. B. Wissen oder Intelligenz, emotionale und motivationale Merkmale sowie als relativ stabil geltende Persönlichkeitsmerkmale wie Selbstsicherheit oder Ängstlichkeit (ebd., S. 146). Unabhängig von der problemlösenden Person und damit objektiv bestimmbar

sind die Anforderungen, die durch die Situation und die Aufgabe entstehen. *Transparenz* als Merkmal der Situation kann beispielsweise durch die Offenlegung aller Variablen und ihren Relationen zueinander mit entsprechenden Grafiken realisiert werden. Eine Variation der Transparenz kann außerdem erfolgen, wenn die problemlösende Person direkt Eingriffe in das System vornimmt und nicht wie es in den Anfängen der Problemlöseforschung mit computersimulierten Szenarien üblich war, indirekt durch eine den Versuch leitende Person. Die *Aufgabenstellung* ist ein weiteres Merkmal der Problemsituation, die unabhängig von dem System, mit dem eine Person interagiert, geändert werden kann. So macht es einen Unterschied ob ein System nur exploriert oder auch nach konkreten Ziel- und eventuellen Zeitvorgaben gesteuert werden soll (ebd.).

Systemmerkmale, dienen der genauen Konstruktion und Beschreibung der verwendeten Systeme. Sie gliedert Funke in *formale* und *inhaltliche* Aspekte (ebd.). Als inhaltliche Aspekte werden von ihm die semantische Einbettung sowie damit verbunden, die Vorwissenskompatibilität genannt, auf die er aber nicht weiter eingeht, sondern sich in seinen nachfolgenden Beschreibungen auf die formalen Aspekte fokussiert (vgl. Wagener, 2001, S. 34). Zur Darstellung der formalen Systemstruktur bedient sich Funke der System- und später der kybernetischen Theorie. Auf Basis dieser ersten systemtheoretischen Betrachtungen identifiziert er zusätzlich zu den von Hübner (1989, S. 222) angeführten Merkmalen Stabilität, Beobachtbarkeit, Steuerbarkeit sowie Berechnung der optimalen Lösung, die formalen Systemmerkmale Zeitverzögerung, Wirkungen, Nebenwirkungen, Eigendynamik und Unabhängigkeit (Funke, 1990, S. 147ff.). Sowohl Hübner als auch Funke beschränken sich in ihren Ausführungen aber vor allem auf formale Systeme auf Basis linearer Gleichungssysteme, in denen die Verbindungen zwischen den Variablen meist kontinuierlicher und damit stetiger Natur sind.

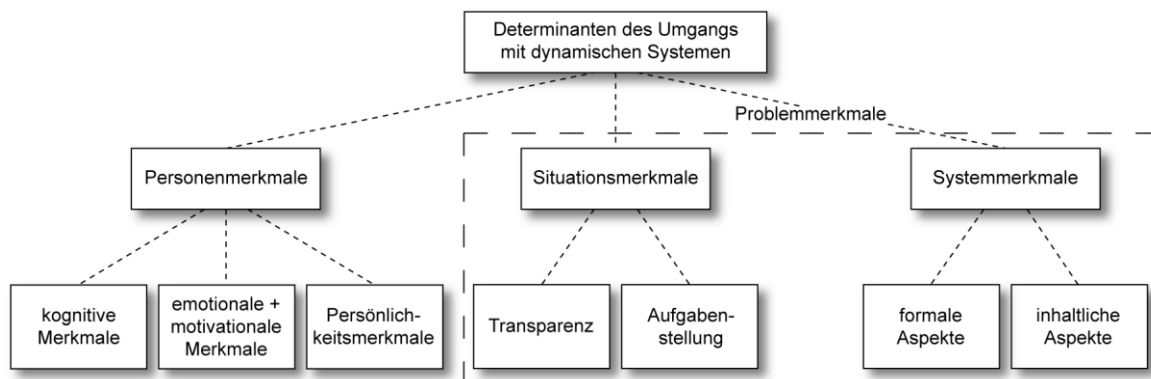


Abbildung 2.2: Determinanten des Umgangs mit dynamischen Systemen (nach Funke, 1990)

Gewinnbringend an Funkses Rahmentheorie zum dynamischen Problemlösen ist für die vorliegende Arbeit die differenzierte Betrachtung von Situations- und Systemmerkmalen sowie die Unterscheidung zwischen formalen und inhaltlichen Aspekten der Systemmerkmale. Die Frage, ob die von ihm bzw. Hübner konkret genannten Merkmale auch zur Beschreibung technischer Alltagsgeräte geeignet sind, wird in Kapitel 6.4 noch einmal aufgegriffen.

Da die Schwierigkeit eines Problems auch von dem Wissen und Können der Person abhängig ist, die das Problem lösen will, beziehen die Systematisierungen von Hussy und Funke

schwierigkeitsdeterminierende Merkmale auf Personenseite mit ein. Die als *subjektive Schwierigkeit* zu bezeichnende Komponente (siehe auch Schley, Eigenmann & Siegfried, 2015, S. 141), kann allerdings nicht zur Konstruktion unterschiedlich schwieriger Items herangezogen werden, da sie erst während der Problembearbeitung entsteht. In Anbetracht dessen, wird im Folgenden lediglich auf die objektiv zu bestimmenden Problemmerkmale eingegangen. Seit den Arbeiten von Dörner zum komplexen Problemlösen werden die folgenden Merkmale in etlichen Forschungsarbeiten (z. B. Sonntag & Schaper, 1997; Kluwe, 1997b; Schaub, 2001) zur Beschreibung von komplexen Problemen herangezogen: *Komplexität, Vernetztheit, Dynamik, Intransparenz und Polytelie*.

Komplexität

Der Komplexitätsbegriff wird zum einen als globaler Terminus verwendet, um dynamische Probleme von einfachen statischen Problemen abzugrenzen (siehe Abschnitt 2.2). Diesem Verständnis von Komplexität sind die noch folgenden Merkmale unterzuordnen (vgl. Hussy, 1985, S. 58; Schaub, 1993, S. 49). Zum anderen wird der Begriff aber auch mit der Anzahl an Komponenten und Variablen in einem System gleichgesetzt (vgl. Arbinger, 2015, S. 130; Kluwe, Schilde, Fischer et al., 1991, S. 292). Eine solche Simplifizierung auf rein quantitative Aspekte stößt jedoch auf Kritik (z. B. Funke, 2003, S. 128), sodass vorgeschlagen wird, den spezifischeren Begriff der Informationsquantität zu verwenden, wenn die Menge an Informationen in einem System betrachtet wird (Greiff, 2012, S. 84).

Komplexität im Sinne von Informationsquantität beeinflusst die Schwierigkeit im problemlösenden Umgang mit Systemen aufgrund der begrenzten Arbeitsgedächtniskapazität¹ des Menschen (Betsch, Funke & Plessner, 2011, S. 155). Diese wird nötig, um die Komplexität informationslastiger Probleme zu reduzieren (vgl. Kluwe, 1997b, S. 15). Je größer die Menge der zu berücksichtigenden Systemkomponenten und damit die zu verarbeitenden Informationen, desto schwieriger wird demnach der Umgang mit einem System.

Vernetztheit

Die Komponenten eines Systems können auf unterschiedliche Weise miteinander verknüpft sein. Die Beziehung zwischen den Variablen eines Systems wird durch das Merkmal Vernetztheit beschrieben. Vernetzte Variablen bewirken, dass ein Eingriff in ein System in Form von Variablenveränderungen zu einer Änderung an einer anderen Stelle in dem System führt, es also nicht möglich ist, Variablen isoliert voneinander zu beeinflussen (Dörner, 1976, S. 20). Vernetzungen sind die Ursache für die neben den intendierten Hauptwirkungen entstehenden Nebenwirkungen (Tauschek, 2006, S. 44). Von Vernetzung wird auch im Fall einer auf sich selbst rückgekoppelten Variable gesprochen (Funke, 2003, S. 129), die eine verstärkende oder abschwächende Wirkung auf ihre Veränderung haben kann. Um mit einem vernetzten System umgehen zu können, müssen die entsprechenden Haupt- und Nebenwirkungen prognostiziert und bei der Bedienhandlung berücksichtigt werden (Dörner, 1976, S. 20). Je stärker die Vernetzung in einem System, desto schwieriger wird eine korrekte Vorhersage.

¹ Eine detailliertere Begründung für den schwierigkeitsbestimmenden Einfluss der hier genannten Problemmerkmale liefert Kapitel 2.4.2.

Dynamik

Das Merkmal Dynamik beschreibt die Veränderung von Systemzuständen, die ohne das Eingreifen seitens der problemlösenden Person entstehen (Arbinger, 2015, S. 131). Der Einbezug der Zeitkomponente kann dabei kontinuierlich oder auch diskret sein. Dynamiken in zeitkontinuierlichen Systemen beschreiben fortlaufende Systemprozesse, die sowohl im ungestörten als auch im gestörten System stattfinden. Diskrete Dynamiken führen zu vermeintlich plötzlichen Veränderungen von Variablen und Zuständen zu bestimmten Zeitpunkten (vgl. auch Kluwe, Schilde, Fischer et al., 1991, S. 292). Diese Zeitpunkte können zufällig sein oder wiederum von anderen Variablen abhängen (Verweildauer in einem Systemzustand ohne eine Eingabe vorzunehmen). Der Umgang mit dynamischen Systemen ist deshalb schwierig, weil zeitliche Veränderungen der Systemzustände bedacht und künftige Entwicklungen prognostiziert werden müssen (Betsch, Funke & Plessner, 2011, S. 156).

Intransparenz

Systeme lassen sich nicht immer vollständig einsehen. Wenn Systemzustände, -variablen und/oder ihre Verknüpfungen zeitweise unbekannt sind oder es dauerhaft bleiben, spricht man von einem intransparenten System. Intransparenz führt zur Unsicherheit im Handeln der problemlösenden Person (Tauschek, 2006, S. 44). Die gezielte Interaktion mit einem intransparenten System verlangt durch geeignete Eingriffe die systematische Gewinnung von Informationen über das Verhalten des Systems (Kluwe, 1997b, S. 15; Funke, 2006b, S. 440), um Unsicherheiten zu beseitigen. Das Merkmal Intransparenz bzw. Transparenz wird oftmals zur Abgrenzung von dynamischen zu statischen Problemen bzw. von dynamischen zum analytischen Problemlösen verwendet (siehe Abschnitt 2.2), aber auch, um den Zusammenhang zwischen Problemlösen und Intelligenz zu erklären (Putz-Osterloh & Lüer, 1981, S. 330f.).

Polytelie

Müssen im Umgang mit komplexen Problemen mehrere Ziele gemeinsam erreicht werden, so liegt Polytelie (vom griechischen „poly telos = mehrere Ziele“ abgeleitet) vor (Arbinger, 2015, S. 131). Diese können sich möglicherweise teilweise oder ganz widersprechen, womit die Konsequenz verbunden ist, dass die problemlösende Person Prioritäten setzen und Ziele kombinieren und ausbalancieren muss (vgl. Schley, Eigenmann & Siegfried, 2015, S. 142).

Die weiteren (formalen) Merkmale, die in der Literatur genannt werden, können in der Regel den oben aufgeführten Merkmalen zugeordnet oder aus ihnen abgeleitet werden (Kotkamp, 1999, S. 28).

2.4 Problemlösen

Während es in den vorangegangenen Abschnitten vor allem um die Entstehung und Charakterisierung von Problemen und Problemtypen ging, widmet sich dieser Abschnitt nun dem Forschungsstand zum Lösen von Problemen. Der Fokus der Betrachtungen liegt hier vor allem auf dem Lösen von dynamischen (komplexen) Problemen, wie sie im Umgang mit technischen Systemen oftmals auftreten.

Problemlösen bezeichnet die Tätigkeit eines Individuums, die den Ausgangszustand eines Problems in seinen Zielzustand überführt (Rollett, 2008, S. 25). Dieses zielgerichtete Verhalten besteht in der Anwendung von geeigneten Operatoren, die, wenn sie nicht bekannt sind, zunächst entwickelt werden müssen und ggf. das Aufstellen von Teilzielen erfordern (Anderson, 2007b, S. 292). Bei der Entwicklung, Veränderung und Anwendung von Operatoren werden frühere Erfahrungen mit der gegenwärtigen Problemsituation verbunden (Ausubel, Novak & Hanesian, 1981, S. 649), ein Sachverhalt, der als problemlösendes Denken bezeichnet wird (Betsch, Funke & Plessner, 2011, S. 139). Denken wird aus kognitionspsychologischer Sicht als ein interner kognitiver Prozess definiert, der erst durch ein erzeugtes Verhalten zu beobachten ist (Mayer, 1979, S. 6). Problemlösendes Denken, das für die Autoren Betsch, Funke & Plessner (2011) die höchste Form geistiger Aktivität darstellt, steuert den Prozess des Erwerbs und der Anwendung von Wissen im Problemlöseprozess (vgl. Süß, 1996, S. 72).

Die Bedeutung, die dabei eine bereits vorhandene Wissensstruktur beim Problemlösen hat, wird unter anderem von den Autoren Ausubel, Novak & Hanesian (1981) besonders betont. Sie gehen davon aus, dass das Vorhandensein relevanter Vorkenntnisse das Problemlösen nicht nur erleichtert, sondern überhaupt erst ermöglicht (Ausubel, Novak & Hanesian, 1981, S. 655). Auch aus der Expertiseforschung ist bekannt, dass Experten sich vor allem durch ihr umfangreiches Wissen von Novizen unterscheiden und nicht durch generell hohe kognitive Fähigkeiten (Opwis, Beller, Spada et al., 2006, S. 225). Wissen und Problemlösen können nach Putz-Osterloh (1988) „[...] als einander ergänzende Determinanten für zielgerichtetes menschliches Verhalten betrachtet werden“ (Putz-Osterloh, 1988, S. 247). Fehlt das zur Zielerreichung benötigte Wissen, ist Problemlösen erforderlich, das wiederum den Erwerb und die Anwendung von Wissen verlangt. Welches spezifische Wissen für die Lösung eines Problems relevant wird, hängt dabei von der semantischen Einbettung des Problems ab (Süß, 1996, S. 72) und führt zu der Unterscheidung zwischen fachübergreifenden und fachspezifischen Problemen, die im Rahmen der Problemtaxonomie bereits diskutiert wurde.

Wissen über Problemlöseoperatoren setzt auch der problemlösende Umgang mit technischen Alltagsgeräten voraus (Schoppek, 1997, S. 271). Da nicht davon auszugehen ist, dass das benötigte Wissen bei allen Personen im gleichen Maß vorhanden ist, beinhaltet Problemlösen zunächst den Erwerb von Wissen. Das (dann) vorhandene Wissen ist zwar eine zwingende Voraussetzung um Probleme zu lösen, aber keine hinreichende Bedingung. Problemlösen verlangt auch die Wissensanwendung in Form von Handlungen, um eine Veränderung des Problemzustandes zu erreichen. Während erfolgreiches Problemlösen auf das Vorhandensein relevanten Wissens schließen lässt, ist der umgekehrte Schluss für erfolgloses Problemlösen nicht zulässig (Ausubel, Novak & Hanesian, 1981, S. 666).

Aufgrund der engen Verzahnung von Wissenserwerb und Wissensanwendung im Problemlöseprozess werden im Folgenden die zur Problemlösung benötigten Wissensarten besprochen und auf die Möglichkeiten des Wissenserwerbs sowie den Diskrepanzen zwischen vorhandenem und angewandtem Wissen eingegangen.

2.4.1 Wissenserwerb und Wissensanwendung im problemlösenden Umgang mit technischen Alltagsgeräten

Heutige technische Alltagsgeräte sind nicht mehr nur auf wenige Funktionen beschränkt und ihre Anwendung wird auch in Zukunft immer ergebnisoffener. In Folge kennen nur die wenigsten Nutzer alle Gerätefunktionen und Ausprobieren wird nach Tully ein wesentlicher Bestandteil der Techniksozialisation (Tully, 2003, S. 19). Ziel des Ausprobierens ist der Erwerb von Wissen, um individuelle Wissensstrukturen zu verändern und zu erweitern. Die anschließende Anwendung des erworbenen Wissens hingegen zielt auf das Herbeiführen von Veränderungen der Systemzustände. Die unterschiedliche Verortung der Zielsetzung von Wissenserwerb und -anwendung einmal innerhalb und einmal außerhalb einer Person, lässt eine getrennte Betrachtung dieser beiden Prozesskomponenten sinnvoll erscheinen (Wirth, 2004, S. 18). Ein weiteres Argument für die Trennung zwischen dem Erwerb und der Anwendung von Wissen im Problemlöseprozess liefert die *Theorie kognitiver Belastung* von Sweller, nach der eine erhöhte kognitive Belastung das Lernen während des Problemlösens behindern kann (Sweller, 1988, S. 283).

Dieser Effekt zeigte sich auch in einer Untersuchung von Vollmeyer & Burns (1999) zu einem minimalkomplexen System, dem BIOLOGY LAB. Zwar nahmen die Autoren eine Trennung zwischen der Wissenserwerbs- und der Wissensanwendungsphase vor, jedoch teilten sie der Experimentalgruppe bereits in der Explorationsphase (Wissenserwerbsphase) die Zielwerte für die anschließende Steuerungsphase (Wissensanwendungsphase) mit. Die Probanden aus der Experimentalgruppe erwarben tatsächlich weniger Wissen über das vorliegende System, da sie bereits früh das Erreichen der Zielwerte übten. Diese bereits bekannten Zielwerte erreichte die Experimentalgruppe in der Steuerungsphase trotz geringerem Wissen genauso gut wie die Kontrollgruppe, sie schafften es dann aber nicht, dieses sehr spezifische Wissen auf das Erreichen anderer Zielwerte zu übertragen (Vollmeyer & Burns, 1999, S. 107).

Problemlösen im Umgang mit unbekanntem technischen Systemen verlangt von der problemlösenden Person zum einen, solche Systeme sorgfältig zu explorieren, um Wissen zu erwerben und zum anderen Wissen anzuwenden, um sie der gewünschten Funktion entsprechend zu bedienen. Das in der Interaktion mit dynamischen Systemen erworbene Wissen lässt sich zunächst allgemein unterscheiden in *Strukturwissen* einerseits und *Eingriffswissen* andererseits (z. B. Preußler, 1998, S. 219; Schaub & Reimann, 1999, S. 175; Süß, 1996, S. 72). Das Wissen über die Struktur eines Systems umfasst das Wissen über die vorhandenen Systemvariablen und ihren Relationen (Putz-Osterloh, Bott & Houben, 1988, S. 241) sowie über die Systemzustände. Die meisten der vorliegenden Untersuchungen zum Wissenserwerb bei der Steuerung komplexer Systeme basieren auf dem Formalismus linearer Strukturgleichungen¹, für die Funke das Strukturwissen weiter in *Relations-*, *Richtungs-* und *Stärkewissen* differenziert (Funke, 1985, S. 456). Relationswissen ist das Wissen über den Zusammenhang zwischen zwei Variablen. Richtungswissen enthält darüber hinaus auch das Wissen über die Richtung dieser Wirkbeziehung und Stärkewissen außerdem Wissen über die Gewichtung dieser Beziehung (ebd.). Das Strukturwissen beinhaltet demnach die

¹ Die in dieser Arbeit verwendeten technischen Systeme basieren nicht auf dem Formalismus linearer Strukturgleichungen, sondern auf dem Formalismus finiter Automaten, die im Kapitel 4.3 vorgestellt werden.

unterschiedlich präzise Kenntnis darüber, wie das System prinzipiell funktioniert (vgl. Preußler, 1998, S. 219).

Neben dem kausalen Wissen über die Systemstruktur ist zur Steuerung eines komplexen Systems außerdem Wissen darüber erforderlich, wie in das System eingegriffen werden muss, um Variablen und Zustände konkret zu verändern (*Eingriffswissen*). Diese Form des Wissens besteht vor allem aus *Wenn-Dann-Regeln*, bei denen der Wenn-Teil die Bedingung beschreibt, unter der eine bestimmte Aktion (Dann-Teil) erlaubt ist (vgl. Opwis, Beller, Spada et al., 2006, S. 213). Wissen, das über die Systemstruktur erworben wird, ist in erster Linie *Sachwissen*, wohingegen sich das Eingriffswissen vor allem auf *Handlungswissen* bezieht (vgl. Schaub & Reimann, 1999, S. 175). Die Unterscheidung zwischen den Begriffen Sach- und Handlungswissen ist konfundiert mit der Unterscheidung zwischen *deklarativen* und *prozeduralen* Wissen und führt dazu, dass diese Wissensbegriffe synonym verwendet werden (z. B. bei Opwis, Beller, Spada et al., 2006, S. 213; Putz-Osterloh, 1988, S. 249).

Im Gegensatz zu der Unterscheidung zwischen Sach- und Handlungswissen wird die Differenzierung zwischen deklarativem und prozeduralem Wissen jedoch nicht anhand des Wissensinhaltes, sondern anhand ihrer Funktion innerhalb des kognitiven Systems vorgenommen (Schnotz, 1994, S. 36). Deklaratives Wissen ist das Wissen über Fakten, Zusammenhänge, Verfahren oder Prozesse und beschreibt damit die vorhandene kognitive Struktur über die Realität (ebd.). Prozedurales Wissen kennzeichnet hingegen diejenigen kognitiven Strukturen, die als Handlungen und Prozeduren direkt ausführbar sind (vgl. Süß, 1996, S. 63) und dazu dienen neue Informationen zu generieren (Putz-Osterloh, 1988, S. 249). Das Wissen *über* Handlungen, Verfahren und Prozeduren ist demnach dem deklarativen Wissen zuzuordnen und nicht dem prozeduralen Wissen. Erst die Fähigkeit zum Vollzug dieser Prozesse ist prozedurales Wissen (Schnotz, 1994, S. 36). Demzufolge ist es denkbar, dass eine Person den Ablauf einer Handlung genau beschreiben kann, also über deklaratives Wissen verfügt, jedoch nicht dazu in der Lage ist, diese Handlung selbst auszuführen (ebd.).

Die von Süß (1996) formulierte Taxonomie von Wissensbegriffen, die zur Steuerung komplexer Systeme und damit für problemlösendes Handeln relevant sein können, unterscheidet zwei Dimensionen. Eine Dimension stellt das nach den Wissensinhalten unterscheidbare Sach- und Handlungswissen dar, die andere Dimension das nach der Wissensform differenzierbare deklarative und prozedurale Wissen (Süß, 1996, S. 66) (Tabelle 2.2).

Tabelle 2.2: Taxonomie der Wissensbegriffe (Wittmann, Süß & Oberauer, 1996, S. 5) mit Beispielen aus Walker (2013)

	Sachwissen: Was ist der Fall?	Handlungswissen: Was ist zu tun?
Deklaratives Wissen: Fragen beantworten können	Benennen von Fakten, Sachverhalten und Begriffen eines Realitätsausschnitts.	Angeben und Beschreiben von situationsspezifischen Handlungsschritten (die zur Zielerreichung notwendig sind).
Prozedurales Wissen: Erfolgreich handeln können	Erkennen von situationspezifischen Merkmalen und Zusammenhängen.	Durchführen bzw. Ausführen einer kognitiven oder motorischen Operation/ Handlung bezogen auf einen bestimmten Kontext.

Das im explorierenden Umgang mit komplexen Systemen angeeignete Strukturwissen der Probanden wird in den meisten Untersuchungen als deklaratives Wissen erfasst, indem Probanden Fragen über das System beantworten (Berry & Broadbent, 1988) oder Variablenzusammenhänge in einem Kausaldiagramm einzeichnen sollen (Funke, 1985). Das Eingriffswissen wird in der Regel erfasst, indem Probanden die Systeme nach entsprechenden Zielvorgaben steuern sollen (z. B. Vollmeyer & Rheinberg, 1998; Strohschneider, 1990; Putz-Osterloh, Bott & Houben, 1988); bei dem erfassten Wissen handelt es sich also um prozedurales Wissen.

Während das spezifische Eingriffswissen allgemein als notwendige Voraussetzung für das Steuern dynamischer Systeme einzuschätzen ist, wird die Bedeutung des Strukturwissens für die Steuerung widersprüchlich bewertet (vgl. Schoppek, 1997, S. 270; Preußler, 1998, S. 219). Forscher, die davon ausgehen, dass deklaratives Wissen eine Voraussetzung für prozedurales Wissen darstellt, sind daran interessiert korrelative Beziehungen zwischen dem diagnostizierten Strukturwissen und der Steuerungsleistung zu ermitteln (z. B. Kluge, 2004, S. 131; Müller, 1993, S. 93f.; Kersting, 1999, S. 215). Während in vielen dieser Untersuchungen die Bedeutung von Strukturwissen für die Systemsteuerung zunächst bestätigt werden konnte, gibt es auch Arbeiten, die keinen oder nur einen geringen Zusammenhang feststellen konnten (z. B. Broadbent, FitzGerald & Broadbent, 1986; Berry & Broadbent, 1988; Putz-Osterloh, Bott & Houben, 1988). In diesen Studien zeigten Probanden trotz fehlendem Strukturwissen ein erfolgreiches Steuerungsverhalten¹.

Kluwe geht in seinen Annahmen sogar so weit, dass kein oder kaum Strukturwissen nötig ist, um ein System erfolgreich steuern zu können. Eine unvollständige oder stark vereinfachende Wissensstruktur reiche aus, um ein System kontrollieren oder steuern zu können (Kluwe, 1991, S. 315). In seinen Ausführungen nimmt er auf den Umgang mit technischen Systemen Bezug und unterscheidet innerhalb des Struktur- und des Steuerungswissens eine System- und eine Interfaceebene (Abbildung 2.3). Strukturwissen auf Systemebene (Anlagenwissen = *device knowledge*) bezieht sich auf das Wissen über die Komponenten eines technischen Gerätes oder einer Anlage und ihre strukturelle Organisation und Funktion innerhalb des Systems (vgl. Kluwe, 1997a, S. 66). Strukturwissen auf Interfaceebene (Displaywissen = *display knowledge*) ist Wissen über die Elemente der Benutzungsschnittstelle (Interface), die aus den Anzeigeelementen des Displays und den Bedienelementen besteht. Das auf Interfaceebene vorhandene Strukturwissen muss nicht mit dem Strukturwissen auf Systemebene verknüpft sein, d. h. eine Person kann die Funktion einer Taste des Interfaces kennen, aber nicht notwendigerweise, welche dahinterliegende Systemkomponente durch diese Taste angesprochen wird (Kluwe, 1997b, S. 20). Auch das Kontroll- und Steuerungswissen lässt sich auf diesen verschiedenen Ebenen betrachten. Das, die Systemebene betreffende Steuerungswissen nennt Kluwe Kausalwissen und „umfasst Kenntnisse über die internen

¹ Berry & Broadbent erklärten den fehlenden Zusammenhang zwischen Strukturwissen und Steuerungsleistung unter anderem mit dem Vorhandensein von impliziten Wissensbeständen. Das Nachzeichnen der anschließenden Diskussion um implizite Lernprozesse erscheint an dieser Stelle zu weit gehend, vor allem, weil der Nachweis impliziten Wissens nur indirekt möglich ist. „Implizites Wissen kann nur angenommen werden, wenn jede Form von deklarierbarem Wissen als Erklärung ausgeschlossen werden kann und auch keine andere Erklärung für eine erbrachte Leistung möglich ist“ (Süß, 1996, S. 71).

Ursache-Wirkungsgefüge des Systems“ (Kluwe, 1997b, S. 21). Kluwe vermutet, dass eine Person, die ein umfangreiches Kausalwissen über ein technisches Gerät hat, ihre Eingriffe in das System und den damit verbundenen Konsequenzen vermutlich besser versteht als eine Person, die nur über Eingriffswissen auf der Interfaceebene verfügt (ebd.). Letztere betrachtet das System eher als eine *Blackbox* und das Steuerungswissen ist Wissen über die In- und Outputs (IO-Wissen), das als Eingriffswissen bezeichnet wird (vgl. Kluwe & Haider, 1990, S. 185). Eine Person weiß in diesem Sinne, welche Inputs (z. B. Betätigen eines Knopfes) zu welchen Outputs führen (z. B. Kaffee fließt aus dem Auslauf eines Kaffeevollautomaten).

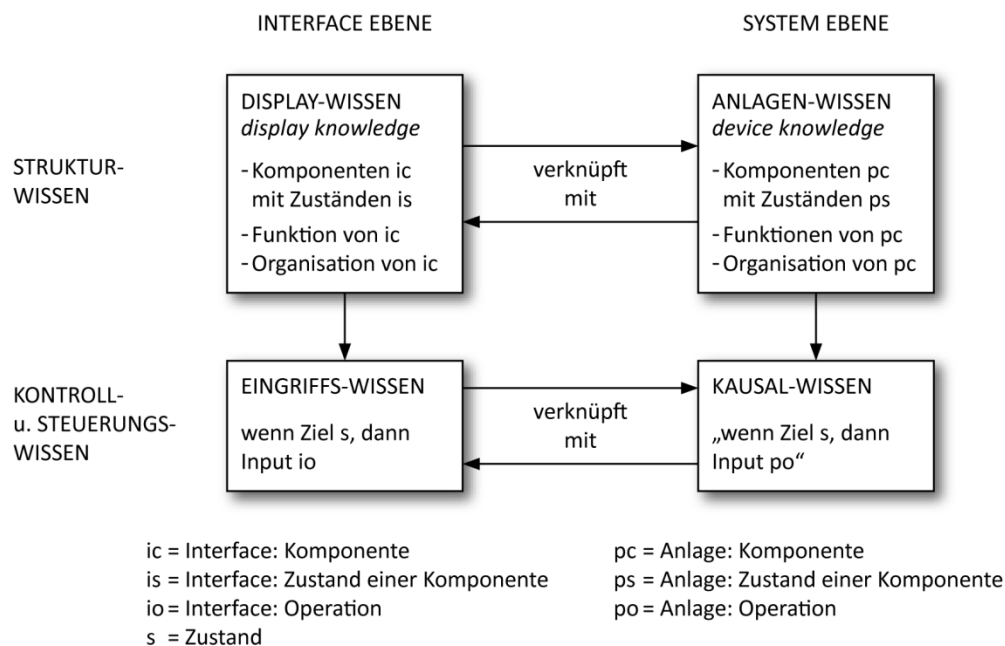


Abbildung 2.3: Kategorien und Ebenen von Wissen zur Kontrolle und Steuerung technischer Systeme (Kluwe, 1997b, S. 20)

In der von Kluwe vorgenommenen Unterscheidung zwischen Interface- und Systemebene für Wissen im Umgang mit technischen Systemen wird außerdem noch einmal der Unterschied zwischen eher fachspezifischen und eher fachübergreifenden Problemen deutlich. Während durchschnittliche Nutzer technischer Systeme vor allem auf der Interfaceebene interagieren, muss jemand, der diese Geräte fertigt, wartet oder repariert auch über umfangreiches Wissen auf der Systemebene verfügen. Die Vermittlung als auch die Diagnose von Wissen hat sich also daran zu orientieren, welches Wissen für die Erfüllung von welchen Aufgaben erforderlich ist (Kluwe, 1997b, S. 21). Die Abfrage umfangreichen Strukturwissens (von dem nur ein Teil benötigt wird, um ein System zu steuern) für die Prognose der Steuerungsleistung, erklärt für Haider (1992) die Unterschiede in den gefundenen Zusammenhangsmaßen (Haider, 1992, S. 98). Wird nur das für die Systemsteuerung relevante Strukturwissen abgefragt, dürfte der Zusammenhang wahrscheinlich höher ausfallen, führt aber dazu, dass genau das zur Steuerung der Systeme wesentliche Wissen vorab aktiviert wird.

Die in dieser Arbeit zu erfassende Kompetenz im problemlösenden Umgang mit technischen Alltagsgeräten, erfordert in der Regel kein Wissen auf Systemebene und damit auch kein fachspezifisches Wissen. Für die Interaktion mit technischen Geräten, die ausschließlich auf der Interfaceebene erfolgt, ist indessen prozedurales Wissen unverzichtbar. Der Nutzer/die

Nutzerin muss in der Lage sein, in das technische System einzugreifen, um entsprechende Zustandsänderungen herbeizuführen. Ob und in welchem Ausmaß für prozedurales Wissen auch deklaratives Wissen nötig ist, lässt sich nur bedingt durch die entsprechenden Wissensabfragen beantworten und ist mit dem Nachteil einer Beeinflussung der problemlösenden Person verbunden, indem durch Reflexion relevantes Wissen aktiviert wird (Süß, 1996, S. 71; Schoppek & Putz-Osterloh, 2003, S. 166) „und damit das Steuerungsverhalten und letztendlich die Steuerungsleistung beeinflusst würde“ (Tauschek, 2006, S. 57).

Das für die Steuerung technischer Systeme benötigte Eingriffswissen muss, wenn es nicht bereits als Vorwissen vorhanden ist, also zunächst erworben werden, ein Prozess, der in der Psychologie auch als *Lernen* bezeichnet wird (vgl. Fortmüller, 1997, S. 113; Renkl, 2015, S. 4). In diesem von Wirth (2004) als *aktiv* hervorzuhebenden Prozess werden Wissensstrukturen erweitert, verändert und neue Strukturen aufgebaut (vgl. Wirth, 2004, S. 16f.). Der Erfolg des Lernens ist dabei abhängig von dem Vorwissen, denn neues Wissen muss in die vorhandene Wissensstruktur eingebunden werden können (Krause, 1982b, S. 148). Erfolgreicher Wissenserwerb setzt zudem das Durchlaufen eines mehrstufigen Prozesses voraus, der die Generierung, Wahrnehmung sowie das Abrufen, Verstehen und Speichern von Informationen beinhaltet (Wirth, 2004, S. 17). Problemlösen als eine Form aktiven Lernens (vgl. Neber, 1987, S. 70; Opwis, Beller, Spada et al., 2006, S. 215) kann demzufolge als ein Informationsverarbeitungsprozess beschrieben werden (Dörner & van der Meer, 1995, S. I), in dem, wenn es sich um dynamisches Problemlösen handelt, eine Person Informationen (Problemlöseoperatoren) zunächst beschaffen muss, bevor sie diese verarbeiten und speichern kann.

Problemlöseoperatoren können prinzipiell auf verschiedene Arten erworben werden:

- Mitteilung durch andere (Instruktion),
- Beobachtung bei anderen (Analogiebildung),
- Entdecken (vgl. Anderson, 2007b, S. 296f.).

Übertragen auf den Umgang mit technischen Geräten, ermöglicht das Befragen von Experten, das Durchlesen von Bedienungsanleitungen (Instruktion), das Beobachten (Analogiebildung) sowie das eigene Ausprobieren (Entdecken) fehlende Informationen zu identifizieren (vgl. Wirth, 2004, S. 33). Vor dem Hintergrund dynamischer Systeme lässt sich durch Instruktion und Analogiebildung nur begrenzt (prozedurales) Eingriffswissen erwerben, so dass dem *Entdeckenden Lernen* im Umgang mit technischen Systemen eine besondere Bedeutung zugesprochen wird (vgl. Kluwe, 1997b, S. 22). Die zur Informationsbeschaffung erforderliche aktive Manipulation von technischen Systemen lässt sich nämlich nur bedingt durch rezeptives Lernen (bei dem die erforderlichen Informationen dem Lerner/der Lernerin bereits präsentiert werden) verdeutlichen (vgl. Renkl, 2015, S. 19; Fletcher, 2004, S. 51).

Damit die erzeugten Informationen zu Wissen weiterverarbeitet und gespeichert werden können, sind sie mit entsprechenden bereits vorhandenen Wissensbeständen zu verknüpfen (Seel, 2000, S. 207). Neu erworbenes Wissen, das in einer Situation angewandt werden soll, beansprucht zunächst noch erhebliche Aufmerksamkeitsressourcen. Erst die wiederholte Beschäftigung mit einem System führt zu routinierten Prozeduren, wodurch die Gedächtnisbelastung reduziert und die Ausführungen immer schneller werden (Opwis,

Beller, Spada et al., 2006, S. 215; Schoppek, 1997, S. 272; Seel, 2000, S. 207). Zusammenfassend erfolgt der Wissenserwerb im Umgang mit dynamischen Systemen durch die Identifikation von Informationen und deren Integration in die Wissensstruktur (Wirth, 2004, S. 22).

Mit dem in dieser Arbeit angestrebten Ziel, den problemlösenden Umgang mit technischen Alltagsgeräten beschreiben und erklären zu können, liegt der Fokus im nächsten Kapitel auf kognitionspsychologischen Theorien, die dem Erwerb und der Anwendung prozeduralen (Handlungs-)Wissens eine besondere Bedeutung beimessen. Die Theorie soll zum einen beschreiben, wie Wissenserwerb abläuft, zum anderen aber auch inter- und intraindividuelle Unterschiede in den Problemlöseleistungen erklären. Eine brauchbare psychologische Theorie des Problemlösens erklärt zudem, warum einige Probleme schwerer sind als andere (Betsch, Funke & Plessner, 2011, S. 174).

2.4.2 Kognitionspsychologische Lerntheorien

Das psychologische Verständnis von Lernen war bis Mitte des 20. Jahrhunderts im Behaviorismus ausschließlich mit beobachtbaren Verhaltensänderungen verbunden (Seel, 2000, S. 18; Opwis, Beller, Spada et al., 2006, S. 205). Demnach findet Lernen nur statt, wenn zuvor nicht beobachtbares Verhalten auftritt, wobei die nicht sichtbaren mentalen Prozesse als Black Box betrachtet und als Forschungsgegenstand ausgeschlossen wurden (ebd.). Erst allmählich setzten sich als Reaktion auf die Fortschritte der Informationstheorie und der Linguistik Theorieansätze durch, die sich der Erforschung der Prozesse innerhalb der Black Box widmen (Aebli, 1980, S. 56; Anderson, 2007b, S. 13). Der als *kognitive Wende* bezeichnete Paradigmenwechsel in der Psychologie führte dazu, dass sich die Psychologie verstärkt mit der Erklärung und der Vorhersage menschlichen Denkens und Handelns befasst. Entwickelt haben sich viele verschiedene Theorien, die Denken und Problemlösen aus unterschiedlichen Perspektiven betrachten. Es gibt also nicht nur *die* eine Theorie, sondern viele verschiedene (vgl. Arbinger, 2015, S. 31; Strohschneider, 1990, S. 15). Gemeinsam ist diesen Ansätzen hingegen, dass sie Problemlösen als Prozess des Wissenserwerbs und der Wissensanwendung beschreiben (Süß, 1996, S. 67) und damit als einen kognitiven Prozess.

Kognitive Prozesse wie *Wahrnehmen, Erinnern, Verarbeiten, Speichern, Repräsentieren* und *Nutzen* von Informationen (vgl. Norman & Rumelhart, 1978, S. 21; Hayes, 1995, S. 13) können auch als Informationsverarbeitungsprozesse interpretiert werden (Gruber, Mack & Ziegler, 1999, S. 7; Arbinger, 2015, S. 31; van der Meer & Döner, 1995 zitiert in Fortmüller, 1997, S. 5). Theoriemodelle, die sich des Informationsverarbeitungsansatzes bedienen, sehen die problemlösende Person als ein System, das mithilfe vorhandener Ein- und Ausgangskanäle Informationen aus der Umwelt aufnehmen und in ihr agieren kann (Opwis, Beller, Spada et al., 2006, S. 205). Kognitive Prozesse werden in diesem Verständnis als eine Reihe von Einzelschritten, in denen Informationen weiterverarbeitet werden, betrachtet (Anderson, 2007b, S. 13). Lernen und Denken als der Verarbeitungsprozess von Informationen benötigt analog zu einer Computerarchitektur eine *kognitive Architektur* (vgl. Artelt & Wirth, 2014, S. 168; Gruber, Mack & Ziegler, 1999, S. 10). Die Suche nach einer angemessenen kognitiven Architektur ist die Suche nach „Teilsystemen, nach deren Funktionen und Eigenschaften sowie nach dem Gefüge der Kommunikation zwischen den Teilsystemen ebenso wie zwischen dem Gesamtsystem und seiner Umwelt“ (Tack, 1995, S. 53). Im Folgenden wird der Entwurf einer solchen Architektur, der insbesondere den Erwerb prozeduralen Wissens in den Blick

nimmt, vorgestellt und auf die Repräsentationsformen dieses Wissen innerhalb des kognitiven Systems eingegangen.

2.4.2.1 Struktur des kognitiven Systems

Das für die Bedienung technischer Geräte zu erwerbende Eingriffswissen als prozedurales Wissen bezieht sich in der Hauptsache auf kognitive Fähigkeiten¹. Bei den zu erwerbenden Prozeduren handelt es sich um Folgen von Operatoren, die in Abhängigkeit des zu erreichenden Ziels Variablen und Zustände des Systems verändern können, wobei die Anwendung eines Operators an Bedingungen geknüpft ist. Solche *Wenn-Dann*-Regeln werden in der Psychologie als *Produktionen* bezeichnet (Seel, 2000, S. 207). Voraussetzung für prozedurales Lernen ist ein Gedächtnis, das diese Produktionen dauerhaft speichern kann (ebd., S. 208). In den ursprünglich für die Computermodellierung kognitiver Prozesse entwickelten *Produktionssystemen* ist ein solcher Produktionenspeicher implementiert (Opwis, 1988, S. 77). Neben dem *Produktionenspeicher* (Abbildung 2.4), der das prozedurale Wissen in Form einer Menge von Produktionen enthält, besteht ein Produktionssystem außerdem aus einem *Datenspeicher*, der das für eine Problemstellung relevante deklarative Wissen bereithält sowie einem *Interpreter*, der die Informationsverarbeitung kontrolliert und steuert (ebd.).

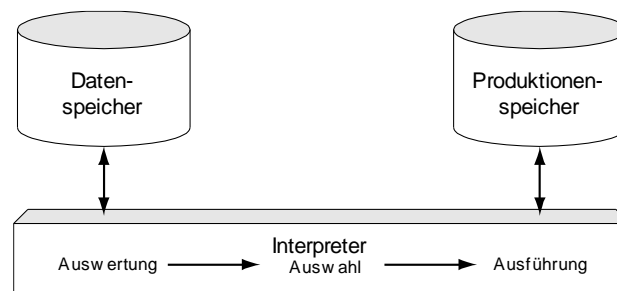


Abbildung 2.4: Modellarchitektur eines Produktionssystems (Opwis, 1988, S. 77)

Um das Zusammenspiel zwischen deklarativem und prozeduralem Wissen zu gewährleisten, muss der Interpreter den *Wenn*-Teil der Produktion unter Berücksichtigung des Datenspeichers auswerten und die aktuell anwendbare Produktion auswählen und auch ausführen. Er ist außerdem für die Bewertung der Ergebnisse verantwortlich, die zu Veränderungen der Produktionenspeicher und der Datenspeicher führen (Strohschneider, 1990, S. 31). Informationsverarbeitungsprozesse bestehen in dem Verständnis regelbasierter Systeme aus der sukzessiven Abarbeitung von Produktionen.

Als ein in der Kognitionspsychologie besonders bedeutendes Beispiel für ein regelbasiertes System gilt das ACT-Modell (ACT ist ein Akronym für Adaptive Control of Thought) von Anderson (1983). In seine Theorie des Erwerbs kognitiver Fertigkeiten, die Anderson im Laufe der Zeit kontinuierlich weiterentwickelt hat, flossen eine Vielzahl von Annahmen zur Gedächtniskapazität, Bearbeitungszeit beim Problemlösen sowie zur Wissensrepräsentation, -verarbeitung und -organisation ein (vgl. Seel, 2000, S. 215). Die Grundannahme der ACT-Theorie besteht in der prinzipiellen Trennung zweier Repräsentationsformen von Wissen: dem deklarativen und dem prozeduralen Wissen (Anderson & Lebiere, 1998b, S. 5).

¹ Der Erwerb von motorischen Fähigkeiten kann bei der Bedienung technischer Alltagsgeräte vernachlässigt werden.

Deklaratives Wissen ist das in sogenannten *Chunks* repräsentierte Faktenwissen, das prinzipiell verbalisierbar aber nicht direkt verhaltenssteuernd ist (vgl. Rüppel, Hinnermann & Wiegand, 1987, S. 176). Chunks bezeichnen dabei die im Gedächtnis gespeicherten, zu größeren unabhängigen Wissensseinheiten zusammengefassten Einzelinformationen. Im Vergleich dazu ist prozedurales Wissen das in Produktionssystemen repräsentierte direkt verhaltenssteuernde Wissen (ebd.). Beide Wissensformen können dauerhaft mental gespeichert werden (vgl. Fortmüller, 1997, S. 27), eine Annahme, die in der kognitiven Architektur von ACT durch zwei separate Gedächtnisse berücksichtigt wird (Abbildung 2.5).

In der Weiterentwicklung der kognitiven Architektur zu ACT-R¹ kommt ein drittes Gedächtnis, der *Zielstapel* (goal stack) hinzu, der die Aufgabe hat, die Hierarchie von Intentionen, die das Handeln leiten, zu enkodieren. Das unterste Ziel des Stapels ist dabei das allgemeinste Ziel und die darüber liegenden Ziele sind dem allgemeinen Ziel untergeordnete Teilziele (Anderson & Lebiere, 1998a, S. 10). Neue Ziele werden auf den Zielstapel gelegt (*push*) und erreichte beziehungsweise nicht mehr verfolgte Ziele werden vom Zielstapel entfernt (*pop*). ACT-R versucht immer das zuletzt auf den Stapel gelegte Ziel zu erreichen und richtet die Aufmerksamkeit auch nur auf ein *aktuelles Ziel*.

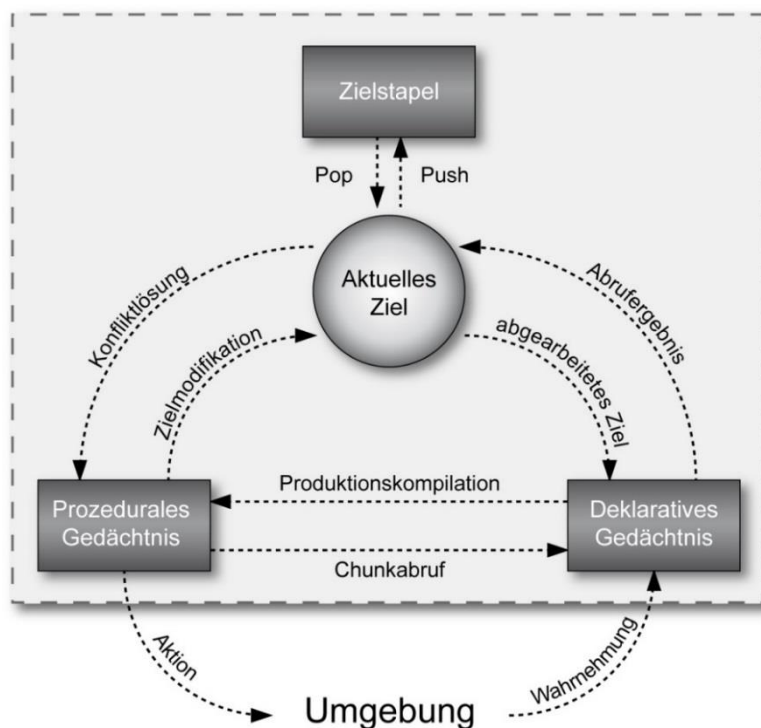


Abbildung 2.5: Systemarchitektur von ACT-R (nach Anderson & Lebiere, 1998a, S. 11; Wallach & Lebiere, 1998, S. 95)

Das aktuelle Ziel steuert die Produktionen aus dem prozeduralen Gedächtnis, indem im Fall einer Übereinstimmung zwischen einer Situation und dem Bedingungsteil einer Produktion, die im Aktionsteil der Produktion spezifizierte (kognitive) Handlung ausgeführt wird. Diese Produktionen können auch Leerstellen enthalten, so dass für die Handlungsausführung, je nach Grad der Prozeduralisierung des Wissens, fehlende Aspekte aus dem deklarativen

¹ Ein kurzer Überblick über die zeitliche und inhaltliche Entwicklung von ACT-R findet sich bei Anderson (2007a, S. 39ff.).

Gedächtnis repräsentiert, im verwendeten Wortlaut von ACT-R also Chunks abgerufen werden müssen (ebd., Wallach & Lebiere, 1998, S. 94).

Zu keinem Zeitpunkt können alle im deklarativen Gedächtnis vorhandenen Chunks aktiviert sein, sondern lediglich eine Teilmenge der Wissensstruktur (Fortmüller, 1997, S. 28). Diese Teilmenge wird in der Kognitionspsychologie als Arbeitsgedächtnis (*working memory*) bezeichnet (z. B. Greeno, 1974, S. 23; Fortmüller, 1997, S. 28). Bezüglich des Arbeitsgedächtnisses werden zwei alternative Modellvorstellungen diskutiert. Auf der einen Seite gibt es die Vorstellung, dass das Arbeitsgedächtnis als ein vom Langzeitgedächtnis abzugrenzender Speicher ist, der die gerade bewussten Informationen umfasst (vgl. Renkl, 2015, S. 8). Aufgrund des limitierten Umfangs der gleichzeitig fokussierbaren Informationen wird eine Begrenzung des Arbeitsspeichers angenommen. Die Konzeption eines separaten Arbeitsgedächtnisses liegt auch der kognitiven Architektur von ACT*, einer Vorversion von ACT-R zugrunde (Anderson, 1983, S. 19). Auf der anderen Seite wird das Arbeitsgedächtnis als der Teil des Wissens im Langzeitgedächtnis bezeichnet, auf den der Fokus durch das gerade aktuelle Ziel liegt. Eine klare Abgrenzung zwischen dem Arbeits- und dem Langzeitgedächtnis wird in dieser Modellvorstellung explizit ausgeschlossen. Diese Arbeitsgedächtniskonzeption wird in dem Kognitionsmodell ACT-R verwendet. Gemeinsam ist den alternativen Erklärungsansätzen die Annahme einer begrenzten Menge von Wissensinhalten die gleichzeitig fokussiert bzw. aktiviert werden kann (Fortmüller, 1997, S. 28).

Je mehr Leerstellen die Produktionen im prozeduralen Gedächtnis aufweisen, desto mehr Chunks müssen im deklarativen Gedächtnis aktiviert werden, wodurch es zu einer Belastung der Ressourcen des Arbeitsgedächtnisses kommt. Diese kognitive Belastung zeigt sich in der langsamen Bearbeitung von neuen Aufgaben beziehungsweise von Problemen (vgl. Haider & Frensch, 1997, S. 521). Erst wenn eine Umwandlung von deklarativem Wissen in prozedurales Wissen erfolgt, kann das Arbeitsgedächtnis entlastet werden (vgl. Spada & Mandl, 1988, S. 3). Der Prozess prozeduralen Lernens vor dem Hintergrund der vorgestellten kognitiven Architektur von ACT ist Inhalt des nächsten Abschnittes.

2.4.2.2 Prozedurales Lernen in kognitiven Modellen

Eine zentrale These in Andersons Lerntheorie ist, dass mit fortdauernder Übung, Teilschritte einer Aufgabenbearbeitung zu komplexen Prozeduren zusammengefasst werden, deren Aktivierung kein deklaratives Wissen mehr benötigt und die damit zu einer geringeren kognitiven Belastung führt. Der Prozess der Wissensprozeduralisierung erfolgt in der ACT*-Theorie in drei Phasen (Anderson, 1983, S. 217ff.): 1) die deklarative Phase (*The Declarative Stage*), 2) die Phase der Wissenskompilation (*Knowledge Compilation*) und 3) die Phase der Wissensoptimierung (*Tuning Productions*).

- 1) Die erste Phase beginnt zunächst mit dem Erwerb deklarativen Wissens, indem neue Informationen des Inhaltsbereichs wahrgenommen und im deklarativen Gedächtnis gespeichert werden. Für ein zielgerichtetes Verhalten in einer Problemsituation reicht die Repräsentation deklarativen Wissens jedoch nicht aus, vielmehr muss dieses Wissen in die Leerstellen allgemeiner, situationsunspezifischer Prozeduren eingesetzt und diese dann angewendet werden (vgl. Wirth, 2004, S. 32; Strohschneider, 1990,

S. 37; Rüppel, Hinnersmann & Wiegand, 1987, S. 177). Die Anwendung der Prozeduren in dieser Phase benötigt viele kognitive Ressourcen und ist daher zunächst noch langsam und fehlerbehaftet (vgl. Mandl, Friedrich & Hron, 1988, S. 139).

- 2) Durch das mehrfach wiederholte Anwenden (Üben) erfolgreicher Prozeduren kommt es zur zweiten Phase, der Wissenskompilation. Von den in der ersten Phase verwendeten allgemeinen Prozeduren werden spezielle Versionen erzeugt, bei denen das relevante deklarative Wissen als Bedingungsteil mit dem Aktionsteil der allgemeinen Prozedur verknüpft wird. Deklaratives Wissen wird in prozedurales Wissen überführt (Strohschneider, 1990, S. 37). Dadurch, dass kein Wissen mehr aus dem deklarativen Langzeitgedächtnis abgerufen und dann mit allgemeinen Prozeduren interpretiert werden muss, laufen die speziellen Prozeduren automatisiert und damit schneller ab (Rüppel, Hinnersmann & Wiegand, 1987, S. 177; Gruber, 1999, S. 62; Sembill, 1992, S. 91).
- 3) In der letzten Phase der Wissensoptimierung kommt es zu einer weiteren Feinabstimmung der Produktionen (vgl. Strohschneider, 1990, S. 38), die erklären, warum Experten komplexere Aufgaben schneller lösen können. Die Phase wird durch die Subprozesse *Generalisation*, *Diskrimination* und *Verstärkung* bestimmt (vgl. Kluwe & Haider, 1995, S. 265f.). *Generalisation* bezeichnet den Prozess, in dem die gemeinsamen Anteile verschiedener spezifischer Produktionen in einer neuen Produktion kombiniert werden (Strohschneider, 1990, S. 38). Diese neue Produktion lässt sich durch die Generalisation auf weitere Situationen anwenden. *Diskrimination* beschreibt einen gegenteiligen Prozess, in dem allgemeinere Prozeduren auf eine spezifische Situation eingeschränkt werden (vgl. Mandl, Friedrich & Hron, 1988, S. 142). Die *Verstärkung* von Prozeduren besteht darin, die Produktionen entsprechend der Häufigkeit ihrer erfolgreichen Anwendung mit einem Stärkewert zu versehen. Produktionen mit einem höheren Stärkewert kommen demnach schneller zur Anwendung als solche mit einem niedrigen Wert (ebd.).

Aus kognitionspsychologischer Perspektive kann prozedurales Lernen entsprechend auch als die graduelle Verbesserung kognitiver Fertigkeiten durch Übung angesehen werden (Mandl, Friedrich & Hron, 1988, S. 123). In Bezug auf den problemlösenden Umgang mit technischen Systemen kommt dem Wissen also eine besondere Bedeutung zu. Personen, die noch keine Erfahrungen mit einem bestimmten Gerät sammeln konnten, beginnen ihre Handhabung zunächst in der deklarativen Phase, während Personen, die über entsprechendes Vorwissen verfügen, bereits spezifische Prozeduren erwerben. Bei fortschreitender Beschäftigung mit einem System kommt es zur Prozeduralisierung des Wissens, ein Verlauf, der in Form abnehmender Bearbeitungszeiten zu beobachten ist (vgl. Schoppek, 1997, S. 272).

2.4.2.3 Kognitive Belastungen im Problemlöseprozess

Wie zuvor beschrieben, erfordert der Wissenserwerb im Prozess des Problemlösens die Verarbeitung neuer Informationen im kognitiven System. Hierfür muss bereits vorhandenes deklaratives und prozedurales Wissen aktiviert werden. Das Arbeitsgedächtnis, das den Teil des Wissens darstellt, der aktuell aktiviert ist, ist in seiner Kapazität jedoch begrenzt. Die Belastung bzw. die Nutzung der Ressourcen des Arbeitsgedächtnisses wird auch als *kognitive Belastung* bzw. *Cognitive Load* bezeichnet (Sweller, 1988, S. 265). Zur Erklärung des

Erfolges beim Wissenserwerb unter Berücksichtigung dieser Limitierung des menschlichen Informationsverarbeitungssystems wird bevorzugt die *Cognitive-Load-Theory* herangezogen (Renkl, Gruber, Weber et al., 2003, S. 4). Diese unterscheidet analog zu anderen kognitionspsychologischen Theorien zwischen einem Langzeitgedächtnis und einem Arbeitsgedächtnis, wobei das Arbeitsgedächtnis in der Anzahl der simultan zu verarbeitenden Informationen eingeschränkt ist. Im Gegensatz zu anderen Lerntheorien differenziert die *Cognitive-Load-Theory* jedoch verschiedene Formen der Arbeitsgedächtnisbelastung (Sweller, van Merriënboer & Paas, 1998, S. 259ff.): 1) den *Intrinsic Cognitive Load*, 2) den *Extraneous cognitive Load* und 3) den *Germane Cognitive Load*, die zusammen den gesamten Cognitive Load ergeben.

- 1) Der *Intrinsic Cognitive Load* beschreibt die Belastung, die durch den Umfang und die Schwierigkeit des zu lernenden Inhaltes, in Abhängigkeit bereits vorhandenen Wissens auf Seiten der Lerner verursacht wird (vgl. Gerjets & Scheiter, 2003, S. 34; Schmeck, 2010, S. 11). Je informationsreicher und wissensintensiver ein zu lösendes Problem ist, desto stärker fällt der *Intrinsic Cognitive Load* aus; je größer hingegen der Umfang des bereits vorhandenen Wissens, desto geringer wird er.
- 2) Der *Extraneous Cognitive Load* bezeichnet die durch die Lernvorgaben entstehende Belastung, die nicht durch den eigentlichen Lerninhalt bestimmt wird. Informationen, die beispielsweise redundant sind, erhöhen den *Extraneous Cognitive Load*, weil Lernende, diese relevanten Informationen von den entsprechend redundanten unterscheiden müssen (vgl. Schmeck, 2010, S. 12). Das gleichzeitige Lösen von Problemen (als das Anwenden von Wissen) im Wissenserwerbsprozess sollte aufgrund einer möglichen Überlastung des Arbeitsgedächtnisses deshalb vermieden werden (Sweller, van Merriënboer & Paas, 1998, S. 263).
- 3) *Germane Cognitive Load* ist der Teil der kognitiven Belastung, der sich aus der Auseinandersetzung mit dem eigentlichen Lernmaterial ergibt und dazu führt, dass es zur Bildung von Wissenschunks oder sogar zur Wissensprozeduralisierung kommen kann (Gerjets & Scheiter, 2003, S. 34). Dieser auch als lernerbezogene Belastung bezeichnete *Germane Cognitive Load* sollte möglichst den höchsten Anteil an dem gesamten Cognitive Load haben. Erreicht werden kann das durch Lernvorgaben, die dem *Extraneous Cognitive Load* möglichst gering halten, wenn der *Intrinsic Cognitive Load* durch das eigentliche Lernmaterial nicht verringert werden kann (Schmeck, 2010, S. 12f.).

Das bereits in Abschnitt 2.4.1 vorgebrachte Argument, den Prozess des Wissenserwerbs von der entsprechenden Wissensanwendung zu trennen, wird mit der *Cognitive-Load-Theory* theoretisch begründet. Hat der Umgang mit technischen Geräten zunächst den Erwerb von Eingriffs- oder auch Strukturwissen auf Interfaceebene zum Ziel, so sollten Personen zur Vermeidung eines *Cognitive Overloads* nicht gleichzeitig konkrete Steuerungsziele verfolgen. Cox & Young (2000) unterscheiden bezüglich des explorierenden Lernens von interaktiven Geräten zwischen dem Erwerb von *Device-Knowledge* und *Task-Knowledge*. Wenn eine Person den Umgang mit einem neuen Gerät erlernen möchte, so versucht sie zunächst herauszufinden, was die Betätigung eines jeden Bedienelementes bewirkt, um sich so die Gesamtfunktion des Gerätes zu erschließen (vgl. Cox & Young, 2000, S. 70). *Task-Knowledge* beschreibt hingegen das Wissen, das benötigt wird, um mithilfe des jeweiligen Gerätes eine

Aufgabe bewältigen bzw. ein Ziel erreichen zu können. Ein technisches Gerät bedienen zu können, heißt sowohl über Device-Knowledge als auch über Task-Knowledge zu verfügen; z. B. „Was passiert, wenn ich diese Taste drücke?“ und „Wie kann ich an diesem Gerät die Uhrzeit einstellen?“ (ebd.).

Kongruierend mit den in der Cognitive-Load-Theory angenommenen Ursachen für die Arbeitsgedächtnisbelastung unterscheiden auch Frensch & Funke (1995) zur Beschreibung komplexen Problemlösens Merkmale, die das eigentliche komplexe System betreffen; solche, die den Problemkontext beschreiben sowie Merkmale auf Personenebene. Sowohl die Merkmale des Systems, als auch die Merkmale der Situation stellen im Gegensatz zu den Personenmerkmalen die objektiv bestimmbar Anforderungen an das Arbeitsgedächtnis dar. Eine Abgrenzung zwischen den eigentlichen Problemmerkmalen und den Situationsmerkmalen wird durch die Beantwortung der Frage, welche Systemmerkmale unabhängig von der Problemsituation betrachtet werden können, möglich (Funke, 1990, S. 146f.). Die Anzahl der Bedienelemente eines technischen Gerätes ist ein Merkmal des Systems und in der Regel unabhängig davon, in welcher Situation das entsprechende System genutzt wird.

In Abschnitt 2.3 wurden bereits aus der Literatur Merkmale extrahiert, die einen Einfluss auf die Schwierigkeit im Umgang mit interaktiven Systemen haben können: *Komplexität*, *Vernetztheit*, *Dynamik* und *Intransparenz*. Vor dem hier beschriebenen Forschungsstand zum kognitiven System und zur kognitiven Belastung wird nun zusammenfassend begründet, warum die genannten Merkmale als schwierigkeitsbestimmend anzunehmen sind. Komplexität als quantitatives Maß wirkt deshalb schwierigkeitsbestimmend, weil die Anzahl der Produktionsschritte zum Erwerb von Wissen und dem Anwenden von Wissen erhöht ist. „If the task being performed is complex (e. g., many pieces of information are relevant for the goal), the source activation must be spread thinly, and each relevant node will receive less source activation. This makes the relevant information less distinct and less easily accessed. As a result, performance suffers“ (Daily, Lovett & Reder, 2001, S. 319). Je komplexer ein System, desto mehr Informationen müssen wahrgenommen und zunächst im deklarativen Gedächtnis an vorhandene Chunks angebunden werden. Die Anzahl simultan zu verarbeitender Informationen ist aufgrund des Arbeitsgedächtnisses jedoch beschränkt. Beziehungen zwischen den Variablen in einem System werden durch das Merkmal Vernetztheit beschrieben. Änderungen an vernetzten Variablen führen zu Änderungen an anderen Stellen in dem System. Diese zusätzlichen Veränderungen müssen wiederum wahrgenommen werden, wodurch die Repräsentation größerer Bereiche des deklarativen Wissens erforderlich ist. Neben den Systemvariablen, die als zu verarbeitende Informationen die Arbeitsgedächtnisressourcen belasten, stellen die vorhandenen Relationen zwischen ihnen weitere Anforderungen dar. Wenn die Wahrnehmung von Veränderungen durch das ausgeprägte Merkmal der Intransparenz behindert ist, kann keine Aktivierung erforderlichen deklarativen Wissen und damit auch keine Einbindung in den vorhandenen Leerstellen der Produktionen erfolgen. In der Folge bleiben die ausgeführten Aktionen fehlerhaft und die Steuerung des entsprechenden Systems ist erschwert. Das Merkmal Dynamik fasst mögliche zeitverzögerte Reaktionen eines Systems und Eigendynamiken zusammen. Letztere beschreiben Zustandsänderungen von Geräten, die ohne eine Interaktion seitens des Nutzers erfolgen können. Dynamische Systeme belasten das Arbeitsgedächtnis des kognitiven

Systems, weil ständig Zwischenergebnisse abrufbar bleiben müssen, also Chunks länger aktiviert bleiben, obwohl bereits neues Wissen repräsentiert werden muss. Die Wahrnehmung des Zusammenhangs zwischen einer getätigten Aktion und der Reaktion des Systems wird erschwert.

3 Problemlösekompetenz

Voraussetzung für die Erfassung von Kompetenzen und hier im Speziellen von Problemlösekompetenz, ist die Klärung des zugrundeliegenden Verständnisses von Kompetenz (vgl. Kaufhold, 2006, S. 14). Der nächste Abschnitt setzt sich deshalb zunächst mit unterschiedlichen Kompetenzkonzepten und -konstrukten¹ und der Bewertung dieser für die vorliegende Arbeit auseinander. Der Handlungsbezug als Merkmal von Kompetenzen erfordert im weiteren Verlauf zudem die Betrachtung von Situationskontexten, da sie das Handeln von Personen beeinflussen können. Der zweite Teil dieses Kapitels setzt sich mit der Frage nach der Operationalisierung von Problemlösekompetenz auseinander und geht auf die Möglichkeiten ergebnisorientierter und prozessorientierter Erfassung des Erwerbs und der Anwendung von Wissen ein.

3.1 Kompetenzverständnis

Kompetenz wird ganz allgemein als eine Befähigung zur Bewältigung von Anforderungen, die in einer Situation an eine Person gestellt werden, verstanden (vgl. Jung, 2010, S. 1; Kaiser, 1998, S. 199). Auch wenn es keine einheitliche Definition des Begriffes Kompetenz gibt, so ist den meisten Kompetenzkonzepten zu entnehmen, dass es sich im oben genannten Sinne um ein Potenzial handelt, das erfolgreiches Handeln in einem Bereich überhaupt erst ermöglicht (vgl. Gillen & Kaufhold, 2005, S. 367; Jung, 2010, S. 19). Ist eine Person den Anforderungen in einer Situation nicht gewachsen, wird sie als inkompetent bezeichnet (vgl. Klieme, Maag-Merki & Hartig, 2007, S. 5). Ob eine Passung zwischen den Ansprüchen der Situation einerseits und den individuellen Voraussetzungen andererseits vorliegt (vgl. Wollersheim, 1993, S. 89), zeigt sich erst in der Handlungsausführung, also in der Bewältigung konkreter Handlungssituationen (vgl. Weiß, 1999, S. 458). Für die Erfassung von Kompetenzen sind deshalb Handlungskontexte zu schaffen, in denen Personen ihre Kompetenz zeigen können (ebd.). Die eben angedeutete Heterogenität im Verständnis von Kompetenz resultiert m. E. zum einen aus den unterschiedlichen Annahmen darüber, welche individuellen Voraussetzungen nötig sind, um eine Situation zu bewältigen und zum anderen, ob Kompetenz als latentes Konstrukt zu definieren ist (vgl. Bernien, 1997, S. 30), auf das nur durch die beobachtbare Handlung geschlossen werden kann oder im Gegensatz dazu, ob Kompetenz gleich der Handlungsausführung selbst und damit direkt erfassbar ist. Weil sich aus der inhaltlichen Definition von Kompetenz Konsequenzen für die Kompetenzerfassung ergeben (vgl. Hartig, 2006, S. 2), setzt sich der Abschnitt zunächst mit diesen Definitionen und ihren Vor- und Nachteilen für die vorliegende Arbeit auseinander. Abschließend wird das Verständnis von Kompetenz für die hier zu betrachtende Problemlösekompetenz im Umgang mit technischen Geräten angegeben.

Eine der frühesten Beschreibungen des Kompetenzkonstruktes stammt aus der Linguistik von Noam Chomsky (Vonken, 2005, S. 19). Ausgehend von der Kritik an Skinners

¹ Während sowohl Konzepte als auch Konstrukte etwas (basierend auf theoretischen Annahmen) Konstruiertes beschreiben (vgl. Seidel, Prenzel & Krapp, 2014, S. 30), wird der Begriff Konzept verwendet, wenn die theoretischen Annahmen und Vorstellungen noch sehr vage sind, der Ausdruck Konstrukt hingegen bei schon klar definierten theoretischen Begriffen (vgl. Schnell, Hill & Esser, 2013, S. 118).

Veröffentlichung von 1957, in der er Sprache als das Ergebnis behavioristischer Reiz-Reaktions-Schemata bezeichnet, postuliert er in der sprechenden Person ein System von Regeln, aus denen die Person Sätze erzeugen kann (vgl. Aebli, 1980, S. 58). Dieses System von Regeln, über das eine sprechende Person verfügt, nennt Chomsky Kompetenz (ebd.) und grenzt sie von der Anwendung des Regelsystems bei der Erzeugung von Sprache ab, die er Performanz nennt (vgl. Chomsky, 1970, S. 14). Auch wenn Chomsky diese Definition von Kompetenz nur auf den Erwerb und die Verwendung von Muttersprache bezogen hat und der Erwerb von Muttersprache schwer mit dem Kompetenzerwerb in anderen Bereichen zu vergleichen ist, hat die grundlegende Unterscheidung von Kompetenz und Performanz auch in anderen Fachrichtungen Eingang gefunden.

Der Entwicklungspsychologe Flammer, der sich der Entwicklung und Veränderung von Kompetenz widmet, grenzt Kompetenz ebenfalls von ihrer Anwendung im tatsächlichen Verhalten¹ (Performanz) ab. Für ihn setzt Performanz Kompetenz zwingend voraus, Performanz ergibt sich aber nicht notwendigerweise aus Kompetenz (Flammer, 2009, S. 31). Ähnlich formulieren es auch die Autoren Grob & Maag Merki (2001); sie beschreiben Kompetenzen als Potenziale für Performanz, die notwendige, aber nicht hinreichende Voraussetzungen für das Bewältigen von Situationsanforderungen darstellen (Grob & Maag Merki, 2001, S. 60). Das heißt, auch wenn Kompetenz nicht durch ein entsprechendes Verhalten gezeigt wird, kann sie dennoch vorhanden sein (Flammer, 2009, S. 22), sie wird aber nicht sichtbar. Nach diesem Verständnis handelt es sich bei Kompetenz um ein latentes Konstrukt, dessen Erfassung mit Herausforderungen behaftet ist.

Aufgrund der fehlenden Möglichkeit Kompetenzen direkt beobachten und messen zu können, lassen sie sich ausschließlich anhand des gezeigten Verhaltens einer Person in einer Situation erschließen (vgl. Weiß, 1999, S. 449; Kaufhold, 2006, S. 22) und auch nur, wenn dieses Verhalten mit einer gewissen Regelmäßigkeit auftritt (Flammer, 2009, S. 22). Das Sichtbarwerden der Kompetenz in einer Situation hängt dabei von den Anforderungen ab, die an die Person gestellt werden. Ein vorhandenes Potenzial, das nicht abgerufen wird, weil die Anforderungen in einer Situation das nicht erfordern, kann demnach auch nicht erfasst werden (Kaufhold, 2006, S. 23) und stellt ein noch ungelöstes Problem dar (Bernien, 1997, S. 30).

Trotz der bei einigen Autoren übereinstimmenden Auffassung von Kompetenz als Potenzial von Performanz, gibt es bei ihnen kein einheitliches Verständnis darüber, welche individuellen Voraussetzungen unter dem Begriff Potenzial subsumiert werden. Während Bernien zur Kompetenz Wissensbestände sowie Fähigkeiten und Fertigkeiten zählt (Bernien, 1997, S. 24 f.), stellen Kompetenzen für Rupp, Leucht & Hartung lediglich „Aufschlüsselungen größer definierter Fähigkeiten“ dar (Rupp, Leucht & Hartung, 2006, S. 198). Hubig verwendet den Kompetenzbegriff sogar als Synonym für „Fähigkeiten, die ein habitualisiertes Können ausmachen“ (Hubig, 2012, S. 34). Auch für Rindermann stellen Kompetenzen lediglich Fähigkeiten dar, obwohl er zu bedenken gibt, dass z. B. Motivationen und Einstellungen genauso einen Einfluss auf die Bewältigung von Anforderungen haben können (Rindermann,

¹ Flammer (2009) spricht in seiner Begriffserklärung von Verhalten. Er unterscheidet aber nicht zwischen Verhalten und Handeln, so dass hier auch Handeln gemeint sein kann. In dieser Arbeit ist ausschließlich Handeln gemeint.

2012, S. 53). Die von Rindermann angesprochenen Aspekte, die neben den Fähigkeiten einen Einfluss auf das Handeln in einer konkreten Situation ausüben können, werden von anderen Autoren (Weinert, 2001b; Jung, 2010; Erpenbeck & Rosenstiel, 2007) in das Kompetenzkonstrukt integriert. Sie verknüpfen damit in dem Kompetenzbegriff die Möglichkeit zu Handeln im Sinne von *Können* mit der Bereitschaft die Handlung auch tatsächlich auszuführen im Sinne von *Wollen* (Gillen & Kaufhold, 2005, S. 367). Die Einbeziehung weiterer potenzieller Einflussfaktoren erfolgt allerdings keineswegs einheitlich (vgl. auch Seeber & Nickolaus, 2010, S. 251). Häufig werden Werte, Einstellungen, Bedürfnisse, Emotionen und motivationale Aspekte bzw. deren Verknüpfungen mit Wissen, Fertigkeiten und Fähigkeiten in das Konstrukt aufgenommen (vgl. Frei, Duell & Baitsch, 1984, S. 30ff.; Jung, 2010, S. 2; Gnahs, 2010, S. 21; Weinert, 2001c, S. 27f.). Dieses Verständnis von Kompetenz prägte vor allem die Diskussion im Bereich der Berufs- und Wirtschaftspädagogik, die sich auf den Erwerb und die Erfassung beruflicher Handlungskompetenz fokussiert (vgl. Seeber & Nickolaus, 2010, S. 249). Handlungskompetenz „[...] wird hier verstanden als die Bereitschaft und Fähigkeit des Einzelnen, sich in gesellschaftlichen, beruflichen und privaten Situationen sachgerecht, durchdacht sowie individuell und sozial verantwortlich zu verhalten“ (KMK 1996/2000 zitiert in Breuer, 2006, S. 199). Mit dem Begriff der Handlungskompetenz wird so versucht, die Trennung von Kompetenz und Performanz aufzugreifen und beide Konstrukte zu verknüpfen (Gillen & Kaufhold, 2005, S. 367). Aber auch in der pädagogischen Psychologie hat sich ein sehr weit gefasster Kompetenzbegriff durchgesetzt. Eine vielzitierte Definition, auf die sich auch die in Bildungsstandards formulierten Kompetenzen beziehen (Zeitler, Köller & Tesch, 2010, S. 24), findet sich bei Weinert (2001c), der Kompetenz 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, 2001c, S. 27f.) definiert. Klieme & Hartig (2007, S. 16) geben zu bedenken, dass ein Kompetenzkonstrukt, das alle Faktoren berücksichtigt, die zur Bewältigung von Anforderungen in konkreten Situationen erforderlich sind, unter Kompetenz das versteht, was bei Chomsky und seinen Nachfolgern Performanz ist, nämlich die tatsächliche Handlungsausführung.

Diese als exemplarisch zu betrachtende Darlegung unterschiedlicher Verständnisse von Kompetenz zeigt, dass es (wie oftmals bei konstruierten Begriffen) keine übereinstimmende Definition gibt. Die in Abhängigkeit der Wissenschaftsdisziplinen verwendeten Kompetenzbegriffe können sich in ihrer Bedeutung auch gegenseitig ausschließen (vgl. Klieme, Maag-Merki & Hartig, 2007, S. 5). Der Begriff der Kompetenz hat also nur innerhalb seiner spezifischen Theorie, in der er konstruiert wird, eine definierte Bedeutung (Erpenbeck & Rosenstiel, 2007, S. XX) und man kann nicht von einer *richtigen* oder *gültigen* Definition ausgehen. Die Verständigung innerhalb einer Disziplin gelingt jedoch nur bei Einnahme desselben theoretischen Standpunktes. Liegen innerhalb einer Disziplin unterschiedliche Perspektiven auf Kompetenz und Kompetenzerfassung vor, sollten diese auch transparent werden. Wenig förderlich für einen Dialog ist das Hin- und Herspringen zwischen den unterschiedlichen Sichtweisen (Bernien, 1997, S. 26). In diesem Zusammenhang erscheint die von Gut-Glanzmann vorgeschlagene Unterscheidung zwischen einem *Standard-*, einem

Modell- und einem *Assessmentdiskurs* hilfreich, wenn über Kompetenzen gesprochen wird (Gut-Glanzmann, 2012, S. 13).

Der Standarddiskurs verleiht dem normativen Aspekt des Kompetenzbegriffes besonderes Gewicht. Ausgehend von Vorstellungen über Bildungsideale werden Bildungs- und Leistungsstandards entwickelt, die festlegen, welche Kompetenzen bis zu einem bestimmten Zeitpunkt erworben werden sollen (Klieme, Avenarius, Blum et al., 2007, S. 9). Im Zentrum des Modelldiskurses, der das Bindeglied zwischen Standard- und Assessmentdiskurs ist, stehen Kompetenzmodelle. Ihre Funktion ist die Beschreibung der Struktur und Entwicklung von Kompetenzen sowie die diesbezügliche Interpretation empirischer Ergebnisse (Gut-Glanzmann, 2012, S. 19). Letztere sind Inhalt des Assessmentdiskurses, der die objektive, reliable und valide Erfassung von Kompetenzen zum Ziel hat. Während im Standard- und Modelldiskurs motivationale und volitionale Aspekte als wertvoll für den Kompetenzbegriff beschrieben werden, spielen sie im Assessmentdiskurs eine untergeordnete Rolle (ebd., S. 21).

Ausgehend von theoretischen Annahmen über das Konstrukt *technische Problemlösekompetenz* wird in dieser Arbeit ein Bezug zu konkreten Problemstellungen hergestellt. Auf Grundlage der bewältigten Anforderungen bei der Lösung dieser Probleme wird versucht, die Ausprägung der Problemlösekompetenz von Testpersonen abzuschätzen und mithilfe weiterer Skalen Rückschlüsse auf die Position des Konstruktes innerhalb eines nomologischen Netzwerkes zu ziehen. Damit lässt sich der hier verwendete Kompetenzbegriff in den Assessmentdiskurs einordnen, wobei der Modelldiskurs (in dem die Ergebnisse des Assessmentdiskurses einfließen) nicht völlig ausgeblendet werden darf. Vor dem Hintergrund des Assessmentdiskurses sind also die Vor- und Nachteile des Einbezuges von Faktoren wie Einstellungen, Werte, Motivation etc. zu diskutieren.

Ein wissenschaftstheoretisch eng gefasstes Konstrukt führt zu aussagekräftigeren Erkenntnissen als ein sehr weit gefasster Begriff (vgl. Hartig, 2006, S. 4), weil sich aus ihm wenige aber einfacher zu falsifizierende Aussagen ableiten lassen. Des Weiteren lässt sich ein Konstrukt schwieriger operationalisieren und damit in konkrete Testaufgaben umsetzen, je komplexer seine inhaltliche Definition ist (Klieme & Hartig, 2007, S. 25). Hinsichtlich des Kompetenzkonstrukts merken Seeber & Nickolaus (2010, S. 251) zudem an, dass die methodischen Möglichkeiten für die Erfassung eines komplex definierten Kompetenzkonstrukts derzeit kaum ausreichen. Dieser Auffassung ist auch Kaufhold (2006, S. 31) und schlägt in Anlehnung an Bernien (1997, S. 18) vor, die Komplexität von Kompetenz zu reduzieren und sich auf einzelne Bestandteile von Kompetenz zu fokussieren. Infolgedessen empfehlen einige Autoren, besser mehrere eng definierte Konstrukte zu unterscheiden und zu erfassen, als diese in einem sehr weit definierten Konstrukt zusammenzuführen (Klieme & Hartig, 2007, S. 25; Klieme, Maag-Merki & Hartig, 2007). Hartig (2006, S. 4) sieht darin zusätzlich den Vorteil die wechselseitige Beziehung dieser Konstrukte untersuchen zu können.

Vor dem Hintergrund, dass die Testbearbeitung und damit die Testleistung in Abhängigkeit vorhandener Erwartungen an Vor- und Nachteile variiert, werden diese durch die Realisierung von *High-Stakes*-Testsituationen annähernd konstant gehalten (Asseburg, 2011, S. 46). Im Gegensatz zu so genannten *Low-Stakes*-Testsituationen, in denen das Abschneiden im Test keine Konsequenzen für die Testperson hat, sind Personen in *High-Stakes*-

Testsituationen aufgrund vorhandener Auswirkungen ihres Testergebnisses hochmotiviert (Wise, 2009). Die Möglichkeit durch ein gutes Testergebnis einen Vorteil zu erlangen, führt zu einer höheren Leistungsbereitschaft und damit dazu, dass die gezeigte Leistung in der Regel bei allen Testpersonen eher der maximalen Leistung entspricht (Sundre & Kitsantas, 2004).

Sind keine persönlichen Konsequenzen aufgrund der erreichten Testergebnisse zu erwarten, so spielt die intrinsische Motivation für das Zustandekommen einer Testleistung eine größere Rolle, die stark von der Testsituation abhängt. Problematisch ist hierbei, dass sich die einzelnen Testaufgaben nicht anhand ihres Motivationsgrades in eine für alle Personen gleiche Rangfolge bringen lassen, um zwischen unterschiedlich kompetenten Personen zu differenzieren (Brandstätter, 2004, S. 261). So kann eine Situation für eine Person eine Herausforderung darstellen und dieselbe Situation für eine andere Person als Bedrohung (ebd.). „Verzerrungen der Testergebnisse aufgrund einer unterschiedlich hohen Motivation zur Testbearbeitung sind unter High-Stakes-Testbedingungen weniger wahrscheinlich“ (Asseburg, 2011, S. 46). In Testsituationen, in denen die Bearbeitungszeit sehr lang ist und die Testleistung sehr stark von der Motivation abhängt, sind materielle oder immaterielle Anreize fast unerlässlich, um zu aussagekräftigen Testergebnissen zu gelangen. Wird Motivation als einflussnehmender Faktor in das Kompetenzkonstrukt aufgenommen, das ein Test erfassen soll, so verbietet sich hingegen die Möglichkeit einer gezielten Beeinflussung von Motivation in High-Stakes-Testsituationen.

Um eine gezielte Beeinflussung geht es auch, wenn Kompetenzen gefördert werden sollen. Ein Testinstrument, das der Gewinnung von Daten dient, auf deren Grundlage Fördermaßnahmen ergriffen werden, muss unterscheiden können, welche der in dem Kompetenzkonstrukt aufgenommenen Dispositionen defizitär sind. So kann eine Person über eine hohe Problemlösefähigkeit verfügen, sich das Lösen eines Problems jedoch selbst nicht zutrauen. Die Förderung muss in diesem Fall also den Fokus auf die Selbstwirksamkeit legen. Für eine derart präzise Differenzierung stellt bei einem sehr weit gefassten Kompetenzkonstrukt die Operationalisierung in ein Testinstrument ein weiteres Problem dar.

Trotz der genannten Vorteile, die ein sehr eng gefasster Kompetenzbegriff bietet, sollte berücksichtigt werden, dass sich eine zu eng gefasste Kompetenz nicht auf andere ähnliche Situationen übertragen lässt. Fälschlicherweise werden so isolierte Fähigkeiten oder Wissen mit dem Kompetenzbegriff versehen (Klieme, Maag-Merki & Hartig, 2007, S. 8). Führt man sich noch einmal den anfangs angesprochenen Handlungsbezug als Merkmal von Kompetenz vor Augen, wird deutlich, dass neben kognitiven Faktoren auch affektive Faktoren wie Haltungen, Einstellungen und Werte erfolgreiches Handeln determinieren. Affektive Dispositionen bestimmen, wie eine Situation wahrgenommen und individuell interpretiert wird (Kaufhold, 2006, S. 23) und auf Grundlage dieser Wahrnehmung eine Person in einer Situation handelt. Es sollten jedoch nur solche Merkmale zu dem Kompetenzbegriff gezählt werden, die als weitgehend stabil gelten, weil Kompetenz ebenfalls ein Konstrukt ist, das zwar veränderbar ist, jedoch nicht täglich variiert (Schott & Ghanbari, 2009, S. 15). Hier wird der Empfehlung von Hartig (2006, S. 4) gefolgt, aktuelle Gefühlslagen oder Motivationen von dem Kompetenzkonstrukt zu trennen: „Ein weiteres Argument für eine explizite Trennung von Kompetenz und Motivation ist, dass Motivation in der Regel als eine Größe betrachtet wird, die im Zeitverlauf variieren kann – eine Schüler kann an einem Tag wenig, am nächsten Tag

hochmotiviert zur Schule kommen. Die Kompetenzen, die er erwirbt, werden hingegen als vergleichsweise stabil betrachtet“. Ähnlich formulieren es auch die Autoren Blömeke & Zlatkin-Troitschanskaia (2015, S. 9): „Kompetenzen werden über die Zeit und Einzelsituationen hinweg als relativ stabile (trait) Dispositionen betrachtet, die von dynamischen (state) Komponenten beeinflusst werden können“.

Kompetenzen sind demnach als Dispositionen zu betrachten, deren Ausprägung in ähnlichen Situationen kaum variiert. Die Anforderungsbereiche, auf die das Kompetenzkonstrukt generalisierbar ist, sind jedoch begrenzt; Kompetenz ist demnach kontextspezifisch (Klieme, Maag-Merki & Hartig, 2007, S. 6; Hartig, 2006, S. 3; Weinert, 2001a, S. 63; Bernien, 1997, S. 25). Der Handlungs- und Situationsbezug als Merkmal von Kompetenz grenzt das Konstrukt von einem grundlegend kognitiven Fähigkeitskonstrukt wie Intelligenz ab, das anhand weitgehend dekontextualisierten Denkleistungen gemessen wird (Wilhelm & Nickolaus, 2013, S. 25). Wie groß der Bereich sein darf, auf den Kompetenz generalisiert werden kann, wird in der Literatur kontrovers diskutiert. So genannte Schlüsselkompetenzen (auch allgemeine Kompetenzen, Grundkompetenzen) lassen sich aufgrund ihrer fehlenden Bindung an spezifische Kontexte nicht so klar von Intelligenz abgrenzen (Hartig & Klieme, 2006, S. 130).

Die Kompetenz zum Lösen von Problemen wird im schulischen Bereich oft als fachübergreifende (Leutner, Funke, Klieme et al., 2005, S. 12) oder cross-curriculare (vgl. Greiff & Fischer, 2013a, S. 27) Kompetenz bezeichnet und analog zum Begriff der Schlüsselkompetenz als eine situationsübergreifend generalisierbare Kompetenz aufgefasst. Die Annahme eines Konstrukts allgemeiner Problemlösekompetenz und das Ziel eine solche zu erfassen, führt zur Konstruktion vermeintlich vorwissensneutraler Problemstellungen, um zu verhindern, dass Probanden mit domänenspezifischem Vorwissen einen Vorteil haben (ebd., S. 28). Aber schon Wirth (2004) bezweifelt, dass derart vorwissensfreie Systeme überhaupt existieren können, da bereits die verwendeten Symbole, die räumliche Anordnung von Informationen oder besonders die Untersuchungsumgebung selbst Vorwissen aktivieren können. Deswegen ist es für die Kompetenzerfassung so wichtig, die Situation exakt zu beschreiben, um die erfasste Kompetenz in den Kontext einzuordnen (Kaufhold, 2006, S. 24).

3.2 Indikatoren von Problemlösekompetenz

Die Frage nach der Operationalisierung eines Konstruktes wie der Problemlösekompetenz ist die Frage danach, worin sich erfolgreiches und erfolgloses Problemlösen unterscheidet. Dieser Frage ist indirekt der Abschnitt 2.4 *Problemlösen* nachgegangen. Dabei konnte die enge Verzahnung zwischen dem Erwerb und der Anwendung von Wissen aufgezeigt und vor dem Hintergrund kognitionspsychologischer Theorien begründet werden. Für die Übersetzung in ein geeignetes Messinstrument sollten sich diese Unterschiede in dem Verhalten zeigen, so dass mithilfe passender Indikatoren Problemlösekompetenz erfasst werden kann.

Prinzipiell lassen sich Indikatoren danach unterscheiden, ob sie ein Ergebnis oder einen Prozess anzeigen. Nach der entsprechenden Vorgehensweise bei der Problemlöseanalyse unterscheidet Hussy (1984a, S. 117) einen *resultatorientierten* und einen *prozessorientierten Ansatz*. Der resultatorientierte Ansatz zeichnet sich vor allem durch das Interesse an dem Ergebnis eines Problemlöseprozesses aus. Mit ihm können insbesondere interindividuelle Unterschiede in der Lösungsgüte konstatiert werden (ebd., S. 118). Vor dem Hintergrund der

Problemschwierigkeit, die sich auf bestimmte Problemmerkmale zurückführen lassen, interessieren aber auch intraindividuelle Unterschiede der Güte von Problemlösungen. Im prozessorientierten Ansatz wird hingegen betrachtet, wie es zu dem beobachteten Ergebnis der Problemlösung gekommen ist. Dieser im zeitlichen Verlauf der Problemlösegeschichte jüngere Ansatz (vgl. Arbinger, 2015, S. 135), fokussiert auf die Vorgehensweise des Problemlösens.

Für Ausubel, Novak & Hanesian (1981) hat der prozessorientierte Ansatz zum Ziel, die Unterschiede zwischen erfolgreichen und erfolglosen Problemlösern erklären und sogar vorhersagen zu können (vgl. Ausubel, Novak & Hanesian, 1981, S. 663f.). Hierzu werden die Problemlöseprozesse von Gruppen unterschiedlich erfolgreicher Problemlöser miteinander verglichen und aus diesem Vergleich Prozessindikatoren abgeleitet. Gut-Glanzmann, der sich in seiner Dissertation der Modellierung und Messung experimenteller Kompetenz widmet, betrachtet Prozessindikatoren außerdem unter einem anderen Aspekt. Für ihn sollte das Messinstrument auch die verschiedenen Möglichkeiten der Kompetenzentwicklung berücksichtigen (Gut-Glanzmann, 2012, S. 15f.). Eine Kompetenzzunahme zeigt sich zum einen darin, dass umfangreichere und komplexere Aufgaben und zum anderen, dass Aufgaben qualitativ besser bewältigt werden (ebd.). Aus dieser Sichtweise können Ergebnisindikatoren nicht durch Prozessindikatoren ersetzt werden, sondern müssen diese ergänzen, um ein differenzierteres Bild der Kompetenz einer Person zu erhalten.

Für die Erfassung der *Kompetenz im problemlösenden Umgang mit technischen Geräten* sind geeignete Indikatoren nicht nur danach auszuwählen, ob sie das Ergebnis oder den Prozess des Problemlösens indizieren, sondern auch danach, ob sie (aufgrund der vorzunehmenden Trennung der Prozesse) auf den Erwerb oder die Anwendung prozeduralen Wissen hinweisen. Mit dieser Intention werden zunächst potenzielle Indikatoren aus der Problemlöseforschung extrahiert und ihre Eignung für diese Arbeit diskutiert.

Ergebnisindikatoren für den Wissenserwerb

In den Untersuchungen zum dynamischen Problemlösen, die zwischen dem Erwerb und der Nutzung von Wissen trennen, wird das in der Lernphase erworbene Wissen oftmals getrennt von der Wissensanwendung erfasst. Die Methode und die sich aus ihr ergebenden Ergebnisindikatoren sind dabei abhängig von der Art des erworbenen Wissens. Für das Steuern von Systemen, die auf Basis linearer Strukturgleichungen konstruiert sind, ist Wissen darüber erforderlich, in welchem Zusammenhang die Variablen des zu steuernden Systems zueinander stehen. Wissen darüber, dass es einen Zusammenhang zwischen zwei Variablen gibt, bezeichnet Funke (1985, S. 456) als *Relationswissen*. Kann die problemlösende Person weiterhin angeben, ob eine Variable positiv oder negativ auf eine andere wirkt, so besitzt sie außerdem entsprechendes *Richtungswissen*. Das qualitativ weitreichendste Wissen stellt das *Stärkewissen* dar, mit dem eine Person sogar einen numerischen Wert für die Stärke des Zusammenhangs zwischen zwei Variablen nennen kann (ebd.). Ob eine Person über dieses Zusammenhangswissen verfügt, wird transparent, wenn die Zusammenhangsstruktur in einem sogenannten Kausaldiagramm eingezeichnet wird (vgl. Beckmann, 1994, S. 69; Blech & Funke, 2006, S. 186). Über die Güte der Kausaldiagramme wird ein Index gebildet, der als Ergebnisindikator für den Wissenserwerb steht (genauer bei Müller, 1993, S. 41ff.). Die Erfassung des in der separaten Phase erworbenen Wissens mit Kausaldiagrammen ist aber

nur sinnvoll, wenn die Variablen in dem zu steuernden System eine entsprechende Zusammenhangsstruktur aufweisen. Für die technischen Systeme, die in dieser Arbeit zu entwickeln sind, kommt eine solche Wissensabfrage also nicht in Betracht.

Zwei weitere Möglichkeiten, das während der Explorationsphase erworbene Wissen zu erfassen sind *Identifikationsaufgaben* und *Verifikationsaufgaben* (vgl. Buchner, 1999, S. 209). Die den Testpersonen in Identifikationsaufgaben gestellten Fragen unterscheiden sich anhand drei Basistypen (Buchner & Funke, 1993, S. 92):

- 1) *Prognostic question*: Bei gegebenen Ausgangszustand und einer Intervention soll der daraus resultierende neue Zustand angegeben werden (z. B. Berry & Broadbent, 1988),
- 2) *Interpolation question*: Sowohl der Ausgangszustand als auch der neue Zustand sind bekannt und die Angabe der Intervention, die erforderlich ist, um den Ausgangszustand in den neuen Zustand zu überführen, wird verlangt,
- 3) *Retrognostic question*: Eine Intervention und der daraus resultierende Zustand sind gegeben, der ursprüngliche Ausgangszustand ist gesucht (z. B. bei Neubert, Kretzschmar, Wüstenberg et al., 2015, S. 185).

In Verifikationsaufgaben wird den Probanden im Anschluss an die Systemexploration eine zufällige Auswahl an Systemzuständen präsentiert, über die sie entscheiden müssen, ob es sich um tatsächlich aufgetretene oder nicht aufgetretene Zustände handelt (vgl. Blech & Funke, 2006, S. 188; Buchner & Funke, 1993; Buchner, 1999).

Prozessindikatoren für den Wissenserwerb

Bei den oben beschriebenen Identifikations- und Verifikationsaufgaben sowie den Kausaldiagrammen handelt es sich um statische Verfahren der Wissensdiagnostik, mit denen sich nur Ergebnisse aber keine Prozesse des Wissenserwerbs erfassen lassen (Wirth, 2004, S. 65). Einen Zugang zu den im Laufe des Wissenserwerbs stattfindenden kognitiven Prozessen bieten hingegen verbale Daten, die mithilfe von Verbalisierungsmethoden vor, während oder nach der untersuchten Lernsequenz entstehen (Funke & Spering, 2006, S. 674). Zu den Methoden, die synchron zu dem laufenden Lernprozess eingesetzt werden, gehören *Protokollnotizen* sowie die Methode des *Lauten Denkens* (ebd.). Letztere kann als „prominenteste Methode kognitionswissenschaftlicher Prozeßdiagnostik angesehen werden“ (Wallach, 1998, S. 164). Bei der Methode des Lauten Denkens wird eine Person gebeten, alles laut auszusprechen, was sie gerade denkt, während sie beispielsweise eine Aufgabe bearbeitet oder ein Problem löst (vgl. Kluwe, 1988, S. 362). Lautes Denken stellt in diesem Sinne den Sekundärprozess dar, der den Primärprozess (etwa die Aufgabenbearbeitung) begleitet (Wallach, 1998, S. 165). Unterschieden werden dabei zwei Formen der Methode: Lautes Denken, bei der die Handlung durch den Versuchsleiter/die Versuchsleiterin unterbrochen wird und kontinuierliches Lautes Denken (Funke & Spering, 2006, S. 675). Wenn bestimmte Voraussetzungen¹ erfüllt sind, ist davon auszugehen dass sich mit der Methode Daten über das aktivierte Wissen einer Person und über dessen Veränderungen gewinnen lassen (Kluwe, 1988, S. 362). Hierfür müssen die verbalen Informationen zunächst protokolliert und dann

¹ Die zu verbalisierenden kognitiven Prozesse und die Handlung müssen in enger zeitlicher Relation zueinander stehen. Außerdem muss die handelnde Person prinzipiell einen Zugang zu ihren Kognitionen haben (Funke & Spering, 2006, S. 675).

weiter aufbereitet werden, ein meist sehr zeitaufwändiger Prozess (vgl. Sandmann, 2014, S. 186).

Inwieweit es zu einer Modifikation des Primärprozesses durch die Methode des Lauten Denkens kommt, ist eine in der Literatur kontrovers diskutierte Frage, zumal die Befunde zu einer möglichen Reaktivität¹ des Lauten Denkens nicht einheitlich sind (vgl. Erdfelder, 1994, S. 61f.; Knoblich & Rhenius, 1995, S. 421). Leistungssteigerungen, die infolge der eingesetzten Methode beispielsweise von Merz (1969) und Kuhl (1983) berichtet wurden, führt Wallach (1998, S. 166) vor dem Hintergrund der ACT-R Theorie auf die Aktivitätserhöhung der Chunks, die in den Verbalisierungsprodukten genutzt werden, zurück. Über negative Effekte des Lauten Denkens auf die Problemlöseleistung berichten hingegen Schooler, Ohlsson & Brooks (1993, S. 174) sowie Kluwe & Reimann (1983). Eine Verlängerung der Bearbeitungszeit in der Phase der Wissensanwendung aufgrund des Lauten Denkens findet sich bei Wallach (1998, S. 218). Eine ebenfalls auf informationstheoretischen Annahmen basierende Begründung für diese negativen Effekte auf den Primärprozess liefern Ericsson & Simon (1980, S. 227f.), indem sie drei verschiedene Level von Verbalisierungen unterscheiden. Inhalte des Arbeitsgedächtnisses, die bereits in sprachlicher Form vorliegen, können direkt artikuliert werden und haben somit keinen Einfluss auf den Primärprozess (Level I Verbalisierung). Liegen die Inhalte des Arbeitsgedächtnisses z. B. in bildhafter Form vor, müssen sie zunächst rekodiert werden, damit eine Verbalisierung möglich wird (Level II Verbalisierung). Die Notwendigkeit der Rekodierung führt dabei zu einer Verlangsamung des Primärprozesses und folglich zu einer längeren Bearbeitungszeit, nicht aber zu einer Verschlechterung der Leistung. Leistungsbeeinflussungen treten dann auf, wenn die Instruktion zum Lauten Denken spezifische Verbalisierungen verlangt, die dazu führen, dass die Aufmerksamkeit von Probanden auf Aspekte gerichtet werden, die sie sonst weniger beachtet hätten (Level III Verbalisierung) (Wallach, 1998, S. 167).

Problematisch hält Wirth (2004, S. 70) die Methode des Lauten Denkens, wenn Prozesse verbalisiert werden sollen, die so stark automatisiert sind, dass sie sich fast vollständig dem Bewusstsein entziehen. Die Bewusstmachung eines eigentlich unbewussten kognitiven Prozesses führt dann zwangsläufig zu einer Veränderung desselben. Gelingt es jedoch nicht, die automatisierten Prozesse zu verbalisieren, führt dies fälschlicherweise oftmals dazu, das Vorhandensein dieser Prozesse anzuzweifeln (ebd.).

Im Hinblick auf die mögliche Reaktivität der Methode des Lauten Denkens und den mit ihrem Einsatz verbundenen Aufwand erscheint eine Fokussierung auf Prozessindikatoren, die aus Logfiles entnommen werden können, unproblematischer (vgl. Wirth, 2004, S. 83). Logfiles stellen digitale, automatisch aufgezeichnete, chronologische Protokolle von Ereignissen in einem computerbasierten System dar (Schmitz & Yanenko, 2014, S. 847). Im Umgang mit komplexen Systemen handelt es sich bei den Ereignissen um vorab als relevant definierte Systemeingriffe wie etwa Klicks (ebd.), Zeicheneingaben oder Drag and Drop-Aktionen. Weil die Aufzeichnung von Logfiles stattfindet, ohne dass die Testperson davon etwas bemerkt,

¹ Reaktivität bezeichnet die Veränderung von Verhalten aufgrund der Aufzeichnung bestimmter Aktivitäten (Bortz & Döring, 2006, S. 325).

liefern die aus ihnen entnommenen Prozessindikatoren unverzerrte Daten über den Wissensprozess (Wirth, 2004, S. 84).

Die von einer Person in der Wissenserwerbsphase vorgenommenen Eingriffe in ein System lassen sich beispielsweise danach analysieren, ob sie einer systematischen Abfolge im Sinne einer Strategie entsprechen (Funke, 2004, S. 103). Personen können sich in ihren Strategien darin unterscheiden, ob sie die in technischen Geräten oftmals vorhandenen Softwaremenüs zunächst in der Breite oder zunächst in der Tiefe erkunden, beziehungsweise wie hoch der Anteil an systematischen Eingriffen ist (vgl. Funke, 2001, S. 83). Für die Bewertung von Eingriffsfolgen reicht es jedoch nicht aus, sich nur die Sequenz der erfolgten Interventionen anzusehen. Aus dem dreimaligen Drücken der Plus-Taste an einem Radio, lässt sich noch keine Systematik erkennen, da diese Taste in unterschiedlichen Systemzuständen unterschiedliche Funktionen haben kann (im Radiobetrieb die Erhöhung der Lautstärke, in den Systemeinstellungen die Wahl des nächsten Menüpunktes). Erst wenn der jeweilige Ist-Zustand, die Intervention und der resultierende Zustand bekannt sind, ist eine prozessorientierte Bewertung der Eingriffsfolgen möglich (Funke, 2004, S. 103). Da hierbei jede einzelne getätigte Aktion bewertet werden muss, ist diese Form der Prozessanalyse nur für wenige und wenig komplexe Systeme geeignet.

Naumann, Goldhammer, Rölke et al. (2014, S. 194) stellen für den Umgang mit Informations- und Kommunikationstechnologien einen einfacheren Prozessindikator vor: die *Anzahl der Interaktionen* mit einem technologischen System. Mit dem Ziel den Erfolg beim technologiebasierten Problemlösen anhand der Anzahl an Interaktionen vorherzusagen zu können (ebd.), werden Daten des PIAAC-Feldtests (OECD, 2013c) aus Deutschland und Kanada analysiert. Die Hypothese, dass es einen positiven Zusammenhang zwischen der Interaktionsanzahl und der Lösungswahrscheinlichkeit gibt, basiert vermutlich auf der Annahme, dass die Interaktionen mit den präsentierten Umgebungen zum Erwerb von Wissen führt, das zur Problemlösung benötigt wird. Diese Annahme wird jedoch nicht explizit von den Autoren formuliert. Der Zusammenhang zwischen dem Prozessindikator Interaktionsanzahl und der Lösungswahrscheinlichkeit wird außerdem nicht als eine lineare Beziehung angenommen, sondern als negativ U-förmig (Naumann, Goldhammer, Rölke et al., 2014, S. 195). Die Autoren vermuten, dass ab einer gewissen Anzahl an Interaktionen auf eine Desorientierung der Testpersonen zu schließen ist und damit die Lösungswahrscheinlichkeit abnimmt (ebd.). Hierbei ist zu berücksichtigen, dass die in der PIAAC-Studie eingesetzten Items zum technologiebasierten Problemlösen *keine* Trennung zwischen einer Wissenserwerbs- und einer Wissensanwendungsphase vornehmen. Vorhandene Wissensdefizite müssen also parallel zur Problembearbeitung ausgeglichen werden und erfordern folgerichtig die aktive Informationsbeschaffung in dem technologischen System, ein Sachverhalt, der sich in der Anzahl der Interaktionen widerspiegelt.

Als alleiniger Indikator für den Erwerb von Wissen in einer, von der Wissensanwendung separierten Phase, erscheint der Autorin die Interaktionszahl aufgrund mangelnder theoretischer Rechtfertigung nicht geeignet. Wenn eine Person immer wieder die gleiche Taste drückt, obwohl diese zu keinem neuen Systemzustand führt und auch keine neuen Informationen generiert werden, wird sie weniger Eingriffswissen erwerben können, als eine

Person, die mit einer geringeren Anzahl von Eingriffen mehr Systemzustände erkundet und damit potenziell mehr Wissen erwerben kann.

Wirth (2004, S. 86), der sich der Selbstregulation von Lernprozessen in einem dynamischen System widmet, bewertet jeden einzelnen Systemeingriff danach, ob dieser einen Beitrag zur Identifikation von Informationen oder zur Integration der bereits identifizierten Information in die vorhandene Wissensstruktur leistet. Interaktionen, die der Informationsidentifikation dienen, zeichnen sich dadurch aus, dass mit ihnen neue, bis dahin nicht wahrnehmbare Informationen generiert werden (ebd., S. 22). Das erstmalige Drücken einer bestimmten Taste an einem neuen Gerät produziert Informationen, die erst zu dem Zeitpunkt durch die nutzende Person wahrnehmbar sind. Die Wahrnehmung neuer Informationen erfolgt durch das Aktivieren von Chunks im deklarativen Gedächtnis, wodurch Ressourcen des Arbeitsgedächtnisses belastet werden. Wird derselbe Eingriff wiederholt, generiert er keine neuen Informationen, sondern führt dazu, dass die dann mehrfach wahrgenommene Information in vorhandenes Wissen integriert wird (ebd., S. 87). Je häufiger ein Eingriff wiederholt wird, desto geringer ist durch die beginnende Prozeduralisierung der kognitive Aufwand für den Abruf des erworbenen Wissens. Um für jede Person den charakteristischen Verlauf der Lernprozessregulation erstellen zu können, wird der gesamte Wissenserwerbsprozess in mehrere gleich große Messzeiträume unterteilt und für jeden Messzeitraum, das Verhältnis von integrierenden zu identifizierenden Eingriffen gebildet und schließlich anhand der für jede Person zur Verfügung stehenden Eingriffsalternativen relativiert. Da jede getätigte Interaktion einer Person mit einem System bewertet werden muss, ist der Aufwand ähnlich hoch wie in der von Funke vorgeschlagenen Prozessanalyse der verwendeten Strategien.

Für das in dieser Arbeit zu entwickelnde Testinstrument zur Erfassung der Problemlösekompetenz im Umgang mit technischen Geräten, ist eine größere Anzahl an interaktiven Systemen zu konstruieren, die sich idealerweise in ihrer Schwierigkeit unterscheiden. Eine anschließende Logfileanalyse, deren Auflösungsgrad bis auf Interaktionsebene reicht, kann in diesem Rahmen nicht geleistet werden. Damit aber eine Prozessanalyse nicht auf den Prozessindikator Anzahl der Interaktionen begrenzt bleibt, wird die Vollständigkeit der Systemexploration als ein Maß für die Informationsidentifikation betrachtet. Wird die Explorationsvollständigkeit (die im Abschnitt 6.5.1 detailliert beschrieben wird) als Indikator für den Wissenserwerb bereits in der Erstellung der simulierten technischen Geräte berücksichtigt, so kann sie als Information in Logfiles aufgezeichnet werden.

Ergebnisindikatoren für die Wissensanwendung

In der Phase der Wissensanwendung müssen Testpersonen in der Regel das vorab erkundete dynamische System nach mehr oder weniger klar definierten Zielen steuern. Der Erfolg bei der Erreichung der vorgegebenen Ziele stellt den wichtigsten resultatorientierten Indikator für die Wissensanwendung dar. Die Bewertung des Erfolges kann dabei auf unterschiedlichen Wegen erfolgen und ist grundsätzlich von der Zielvorgabe abhängig (vgl. Kersting, 1999, S. 70). Wenn es, wie in vielen bereits untersuchten Problemsituationen darum geht, eine konkrete Zielvariable zu verändern, indem diese entweder möglichst genau einen vorgegebenen Wert oder etwas unpräziser, einen möglichst hohen oder niedrigen Wert

erreichen soll, so dient der Wert für die Variable beziehungsweise seine Differenz zum Zielwert als Steuerungsergebnis. In einer Studie mit dem komplexen Szenario *Tailorshop*, sollen Probanden beispielsweise als Manager einer Textilfabrik deren Kapital erhöhen. Nach zwölf simulierten Monaten ergab sich die Steuerungsleistung aus dem aus Einnahmen und Ausgaben berechneten Gewinn (Barth & Funke, 2010, S. 1262). Hier stellt der Indikator für den Erfolg der Wissensanwendung ein intervallskaliertes Maß dar.

In vielen Problemstellungen kann die Ausprägung des zu erreichendes Zielwertes nicht als intervallskaliertes sondern lediglich als nominalskaliertes Wert angegeben werden, also in der Form *Ziel erreicht* oder *Ziel nicht erreicht* (z. B. Arbinger, 2015, S. 135). Hat eine Person die Eingabe einer konkreten Adresse in ein Navigationsgerät zum Ziel, so kann der Erfolg nur danach bewertet werden, ob die Adresse vollständig und ohne Fehler eingegeben wurde, also ob das Ziel erreicht wurde oder nicht. Wurde beispielsweise im Straßennamen ein Buchstabe vergessen einzugeben (statt Stanzweg z. B. Tanzweg), kann das in der Praxis unangenehme Folgen haben. Werden einer Person mehrere Problemstellungen, hier konkret zu steuernde Systeme vorgelegt, können die richtig gelösten Probleme zusammengezählt werden.

Prozessindikatoren für die Wissensanwendung

Die Kompetenz einer Person im problemlösenden Umgang mit dynamischen Systemen zeigt sich vor allem darin, dass sie vorgegebene Systeme zielgerichtet steuern kann. Die Steuerungsleistung, als ergebnisorientierter Indikator ist für die Beurteilung dieser Kompetenz unverzichtbar und kann nicht durch prozessorientierte Indikatoren ersetzt werden. Wollte man Leistungsmaße nämlich durch Verhaltensmaße ersetzen (was aufgrund des damit verbundenen Aufwandes schon nicht sinnvoll ist), so muss zum einen ein geeignetes Verhaltensmaß gefunden werden und zum anderen muss dieses Maß hoch mit der Steuerungsleistung korrelieren (Kersting, 1999, S. 81).

Kompetente Personen unterscheiden sich von weniger kompetenten Personen jedoch nicht nur darin, dass sie schwierigere Aufgaben lösen können, sondern auch darin, dass sie Aufgaben qualitativ besser lösen (Gut-Glanzmann, 2012). Die Beurteilung der Qualität einer Problemlösung kann bei den hier betrachteten Systemsteuerungen nicht anhand des Ergebnisindikators *Steuerungsleistung* beurteilt werden. Hierzu ist wiederum der Prozess der Problembearbeitung näher zu betrachten. Als Prozessinformation können zum Beispiel die Bearbeitungszeit (Hussy, 1984a, S. 123), die Nutzung eines optimalen Pfades an Eingriffen in das System oder die Effektivität bestimmt werden (vgl. Funke, 2004, S. 103; Opwis, Beller, Spada et al., 2006, S. 222).

Zusammenfassung

Einführend wurde in diesem Abschnitt darauf hingewiesen, dass die in dieser Arbeit zu wählenden Indikatoren den Erwerb und die Anwendung prozeduralen Wissens anzeigen müssen. Die Bedienung technischer Geräte basiert in der Hauptsache auf Handlungswissen in der Form von Wenn-Dann-Produktionen. Eine resultatorientierte Wissensabfrage mithilfe von Kausaldiagrammen oder Identifikations- und Verifikationsaufgaben eignet sich vor allem für die Diagnose erworbenen deklarativen Wissens. Indikatoren für den Erwerb prozeduralen Wissen sollten sich hingegen an den Prozess des Wissenserwerbs orientieren. Die

Aufzeichnung der Prozess- bzw. Verhaltensmaße sollte die problemlösende Person jedoch nicht in ihrem Handeln beeinflussen, sondern von der Person unbemerkt im Hintergrund erfolgen. Eine Herausforderung stellt die Wahl eines geeigneten Indikators dar, denn dieser sollte zum einen theoretisch begründet und zum anderen auch mit dem tatsächlichen Problemlöseerfolg korreliert sein. Für die Prüfung des Zusammenhangs zwischen dem Erwerb und der Anwendung des Wissens wird die Bildung eines Index für den Indikator nötig. Der Aufwand für die Bildung eines oder mehrerer Messwerte fällt in einigen Studien jedoch unverhältnismäßig hoch aus. Als Alternative wird im Abschnitt 6.5.1 die Explorationsvollständigkeit als einen Indikator, der automatisch ausgewertet werden kann vorgestellt.

Der Problemlöseerfolg als der Erfolg bei der Anwendung des erworbenen Wissen lässt sich hingegen am sichersten mit einem resultatorientierten Indikator beschreiben: der Steuerungsleistung. Diese gibt in Bezug zu einer vorgegebenen Problemstellung an, ob eine Person das Problem gelöst hat oder nicht. Personen die über eine hohe Problemlösekompetenz verfügen, werden in der Lage sein, schwierigere Probleme zu lösen als Personen mit einer geringeren Kompetenz. Der Unterschied in der Kompetenz zeigt sich außerdem durch die Qualität der bewältigten Aufgaben. Diese lässt sich mithilfe prozessorientierter Indikatoren wie die Bearbeitungszeit erfassen.

4 Computerbasierte Kompetenzdiagnostik

Ziel der vorliegenden Arbeit ist die Entwicklung und Validierung eines Testinstruments zur Erfassung der Kompetenz im problemlösenden Umgang mit technischen Alltagsgeräten. Ein technisches Gerät nutzen und bedienen zu können, bedeutet Änderungen an dem Gerät vorzunehmen. Hierzu sind zum einen gezielte Eingriffe in das System nötig und zum anderen müssen die Reaktionen wahrgenommen und interpretiert werden, um zielsicher weitere Eingriffskombinationen zu tätigen. Für die Konstruktion eines Tests, der diese Kompetenz erfassen soll, stellt die Interaktion, die konstituierend für die Problemlösekompetenz mit technischen Geräten ist, eine Herausforderung dar. Während sich schulisches und berufliches Fachwissen, das deklarativ repräsentiert ist, häufig mit Papier-Bleistift-Verfahren abfragen lässt, eignet sich dieses Verfahren nicht, um kompetentes Handeln in dynamischen Situationen diagnostizieren zu können. Um zu erfahren, wie Personen in unbekanntem Situationen agieren, bietet sich beispielsweise die Beobachtung der Akteure im Feld und die damit mögliche differenzierte Untersuchung ihrer Handlungsstrategien an (Badke-Schaub & Tisdale, 1995, S. 48). Neben dem Vorteil einer hohen Realitätsnähe erscheinen der Aufwand des Verfahrens sowie die nicht standardisierbaren und damit auch nicht vergleichbaren Situationen hingegen problematisch (vgl. Jude & Wirth, 2007, S. 50). Alternativ zur Untersuchung im natürlichen Umfeld der Probanden kommen auch Leistungstests mit realen Gegenständen in Laborstudien in Betracht. Der Einfluss umweltbedingter Störvariablen könnte so zwar weitgehend kontrolliert werden (Döring & Bortz, 2016b, S. 206), jedoch müssten auch die eingesetzten technischen Geräte bestimmte Bedingungen erfüllen. Bereits existierende Produkte sind nach ihrer Ausprägung von als schwierigkeitsbestimmend angenommenen Merkmalen auszuwählen, um Testpersonen unterschiedlich schwierig zu bedienende Geräte vorlegen zu können. Analog zu den, in beruflichen Bereichen eingesetzten Tests in Form von Arbeitsproben¹ eignen sich Laborstudien mit realen technischen Geräten zudem nur für kleine Stichproben, denn der Erwerb einer ausreichenden Anzahl an technischen Geräten zur ökonomischen Erfassung technischer Problemlösekompetenz wäre mit hohen Kosten verbunden (vgl. Abele, Gschwendtner & Nickolaus, 2009b, S. 24).

Für eine sowohl hoch standardisierte, reliable als auch praktikable und für große Stichproben zugleich kostengünstige Erfassung von Handlungen bieten sich computerbasierte Messinstrumente an (Gschwendtner, Abele & Nickolaus, 2009, S. 557f.). Mit ihnen lassen sich dynamische Situationen konstruieren, die eine wiederholbare und kontrollierbare Untersuchung von Personen (Badke-Schaub & Tisdale, 1995, S. 48) im Labor und im Feld erlauben (Funke & Reuschenbach, 2011, S. 601). Für die Diagnostik mit computerbasierten Tests hat sich der Begriff *Computerbasiertes Assessment (CBA)* durchgesetzt (Jurecka & Hartig, 2007, S. 38). Im Gegensatz zum übergeordneten Begriff des *Technologiebasierten Assessments (TBA)*, unter dem alle informationstechnologischen Verfahren (wie z. B. Mobiltelefone, Eye-Tracking-Brillen, Computer etc.) zur psychologischen und pädagogischen Diagnostik zusammengefasst werden können (ebd., S. 37), beschreibt computerbasiertes Assessment die Verwendung von Computern zu diagnostischen Zwecken (ebd., S. 38). Das Testmaterial wird

¹ Arbeitsproben stellen standardisierte Aufgaben dar, die inhaltlich valide und erkennbar äquivalent erfolgsrelevantes berufliches Handeln abbilden (Funke, 1993, S. 120).

in der Regel über den Bildschirm präsentiert, die Interaktionen der getesteten Person erfolgen entweder über denselben oder über andere Eingabehardware wie Maus oder Tastatur (ebd.).

Neben den bereits oben angedeuteten Vorteilen computerbasierter Tests ergeben sich auch Herausforderungen, die sich zum einen auf die Erstellung und zum anderen auf den Einsatz computerbasierter Instrumente beziehen. Nachdem im folgenden Abschnitt zunächst Simulationen als eine spezielle Form computerbasierter Tests vorgestellt werden, setzt sich der anschließende Abschnitt mit beiden Aspekten, dem Nutzen und der Schwierigkeit von Computersimulationen auseinander. Die technischen Herausforderungen aufgreifend wendet sich der letzte Abschnitt den Möglichkeiten der Modellierung konkreter Simulationen von technischen Geräten zu. Hierbei wird der hinter vielen technischen Geräten stehende Formalismus *finiten Automaten* vorgestellt, der auch für die Programmierung der in dieser Arbeit einzusetzenden Simulationen verwendet wird.

4.1 Computersimulationen

Während zu Beginn des Computereinsatzes in der Diagnostik vor allem die Umwandlung bereits existierender Papier-Bleistift-Verfahren in eine elektronische Form im Vordergrund standen (Funke & Reuschenbach, 2011, S. 600), entwickelten sich seitdem Testmaterialien, die ausschließlich mit einem Computer bearbeitet werden können (Jude & Wirth, 2007, S. 49). Hierzu zählen die mit dem Computer simulierten Problemlöseszenarien, die seit den Anfängen komplexer Problemlöseforschung zum Einsatz kommen. Computersimulationen als Spezialfall von Modellen, bilden Realitätsausschnitte ab und können gewünschte Merkmale und das Verhalten der Realitätsausschnitte auf einem Computer reproduzieren (Saam, 1997, S. 15). Modelle geben jedoch nicht alle Merkmale und Variablen der Realität wieder, sondern stellen ein verkürztes und verändertes Abbild der Realität dar (ebd.), das in Computersimulationen umgesetzt wird. Die in Studien zum komplexen Problemlösen verwendeten Simulationen unterscheiden sich unter anderem darin, wie stark sie die Realität abstrahieren (Neber, 1987, S. 19). Um Modelle hinsichtlich ihres Abstraktionsgrades besser differenzieren zu können, ordnet Obermayer (1964, S. 608) verschiedene Realitätsmodelle innerhalb eines Spektrums, dessen eines Ende die reale Welt und dessen anderes Ende mathematische Modelle bilden (Abbildung 4.1). Je weiter man sich auf der Skala von der realen Welt nach rechts bewegt, desto größer ist der Abstraktionsgrad des entsprechenden Modells und entsprechend höher ist seine Generalisierbarkeit. Gleichzeitig sinkt damit die Menge an Details aus der Realität und die Validität der aus den eingesetzten Modellen gewonnenen Ergebnisse.



Abbildung 4.1: Grad der Abstraktion von Modellen (Obermayer, 1964, S. 608; Übersetzungen und Ergänzungen (gestrichelte Linien) von Neber, 1987, S. 19)

Je weniger gemeinsame Merkmale die in einer Untersuchungssituation verwendeten Simulationen und der Realitätsausschnitt aufweisen, desto weniger gut lassen sich Rückschlüsse von der Simulation zurück auf die Realität ziehen (vgl. Müller, 1993, S. 5). Der Anteil der Überschneidung, der bei Stachowiak (1973) als *partielle Isomorphie* und bei Gray (2002, S. 210f.) als *Correspondence* bezeichnet wird (Abbildung 4.2), bestimmt dabei den Grad der Abstraktion. Die Aspekte aus dem Realitätsbereich, die im Modell nicht repräsentiert sind, werden als Präeritonsklasse bezeichnet und diejenigen Aspekte, die durch die Modellierung entstanden sind und nicht in der Realität vorkommen, werden als Abundanzklasse bezeichnet (vgl. Saam, 1997, S. 69).

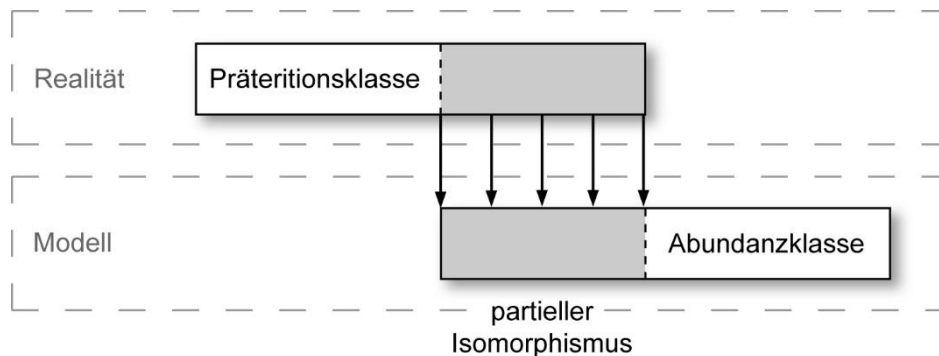


Abbildung 4.2: Überschneidung zwischen realem Gegenstandsbereich und Modell (vgl. Saam, 1997, S. 69)

Neben dem Grad der Abstraktion unterscheiden sich Simulationen außerdem darin, wie detailliert und genau sie die zu modellierenden Aspekte aus der Realität darstellen können (Neber, 1987, S. 19). Dieses als *Fidelity* (Gray, 2002) bezeichnete Merkmal von Computersimulationen kann sich zum einen auf die Ähnlichkeit des Aussehens (*display realism*) und zum anderen auf die Ähnlichkeit des Verhaltens (*response realism*) zwischen Realitätsausschnitt und Modell beziehen (Funke & Reuschenbach, 2011, S. 602). So kritisiert Müller (1993, S. 4), dass ein großer Teil der als *Realitätsausschnitt* oder als *realitätsnah* bezeichneten Problemlöseszenarien kaum naturgetreue Abbildungen des simulierten Gegenstandsbereichs darstellt, obwohl die Möglichkeiten durch die Einbindung multimedialer Inhalte wie (animierte) Grafiken, Videos und Audios (Goldhammer, Kröhne, Keßel et al., 2014, S. 12) durchaus gegeben sind. Die als *response realism* bezeichnete Realitätsnähe bezieht sich unter anderem auf die Eingriffsmöglichkeiten (Tastatur, Maus, Joystick, Touchdisplay etc.) in die Simulation sowie die daraus folgenden Reaktionen (Funke & Reuschenbach, 2011, S. 602). Während auf Objekte in der Realität direkt eingewirkt werden kann, erfolgen die Eingriffe in Computersimulationen zumeist durch das Anklicken von Buttons oder über das Drag and Drop von Objekten auf dem Bildschirm (ebd.). Durch diese Eingaben werden Variablen in der Simulation verändert, die dann durch einen, der die Simulation bedienenden Person meist unbekanntem Rechenalgorithmus durchlaufen und den Zustand der Simulation verändern (Strauß & Kleinmann, 1995, S. 108). Als Grund für die Verwendung der teilweise holzschnittartigen Realitätsabbilder (Buerschaper, 2000, S. 147) wird angegeben, dass es für diagnostische Zwecke eben nicht um die Gestaltung von *high fidelity Simulationen* geht, sondern um die Umsetzung von Anforderungen, die Menschen im Alltag bewältigen sollten (Badke-Schaub & Tisdale, 1995, S. 48).

Mit Computersimulationen ist es also möglich, reale technische Geräte aus dem Alltag mehr oder weniger detailgetreu abzubilden, Eingriffsmöglichkeiten zu generieren sowie das entsprechende Systemverhalten nachzubilden. Für die Entwicklung eines Testinstrument zur Erfassung der Kompetenz im problemlösenden Umgang mit technischen Alltagsgeräten sind solche Simulationen unverzichtbar, da sich die Interaktivität und Dynamik in solchen Handlungssituationen nicht mit traditionellen Tests umsetzen lassen. Ob das in einem computerbasierten Test beobachtete Verhalten mit dem Verhalten im Umgang mit realen Gegenständen übereinstimmt, ist ein Aspekt, der bestimmt, inwieweit aus den Ergebnissen eines solchen Tests Rückschlüsse auf das kompetente Verhalten im Alltag gezogen werden dürfen. Den Herausforderungen, die sich aus der Äquivalenzproblematik für das zu operationalisierende Konstrukt ergeben, widmet sich der folgende Abschnitt.

4.2 Nutzen und Herausforderungen computerbasierter Tests

Die Äquivalenzproblematik bezieht sich auf die Frage, inwieweit sich durch einen Wechsel des Testverfahrens Veränderungen an dem Testverhalten aufgrund des erfassten Konstrukt ergeben (Richter, Naumann & Noller, 1999). Soll beispielsweise ein bislang in Papierform vorliegender Test nun auch in computergestützter Form angewendet werden, stellt sich die Frage nach der Vergleichbarkeit der Ergebnisse dieser beiden Testversionen. Unterscheiden sich die Testergebnisse infolge unterschiedlichen Testverhaltens voneinander, so besteht zwischen den beiden Testformen keine Äquivalenz und die diagnostischen Aussagen aus beiden Tests sind nicht miteinander vergleichbar (Schaper, 2009, S. 23). Kreuzpointner (2010, S. 47) differenziert zwischen einer psychometrischen und einer experimentellen Äquivalenz. Als psychometrisch äquivalent bezeichnet er zwei Tests, die parallel sind, deren Mittelwerte, Varianzen, Verteilungen bzw. Randordnungen also gleich sind und die mit anderen Variablen ähnlich hoch korrelieren. Unter experimenteller Äquivalenz fasst er Faktoren wie die emotionalen, wahrnehmungs- und einstellungsbezogenen Reaktionen auf die jeweiligen Testformen zusammen (ebd., S. 48). Richter, Naumann & Noller (1999) erwähnen außerdem eine erfahrungsbezogene und populationsspezifische Äquivalenz. Im Fall der Übertragung eines Papier-Bleistift-Tests auf den Computer können beispielsweise Erfahrungen mit Computern, die in Abhängigkeit von bestimmten Gruppen oder Individuen variieren, die Äquivalenz der Testverfahren beeinflussen. Während eine Verringerung der Äquivalenz zwischen papierbasierten Fragebögen oder Multiple-Choice-Tests und ihren computerbasierten Pendanten häufig als Validitätsproblem des computerbasierten Tests interpretiert werden, sprechen sie in Kompetenzbereichen, die über rein kognitive Wissensbereiche hinausgehen, für eine Validitätssteigerung (Jurecka & Hartig, 2007, S. 45).

Es sollte prinzipiell nicht darum gehen, ein Testmedium adäquat durch ein anderes zu ersetzen, sondern darum, mit welchem Testverfahren ein zu messendes Konstrukt am besten operationalisiert werden kann (Frey & Hartig, 2013, S. 56). Eignen sich Multiple-Choice-Aufgaben vor allem zur Erfassung deklarativen Sach- und Handlungswissen, ist ihr Einsatz in Tests, die prozedurales Handlungswissen diagnostizieren sollen, eher unzweckmäßig (vgl. Goldhammer, Kröhne, Keßel et al., 2014, S. 12). Kompetenzkonstrukte (wie dynamisches Problemlösen), in denen tatsächliches Handeln in dynamischen Situationen von Bedeutung ist (Jude & Wirth, 2007, S. 49f.), sollten als Antwortformat die Interaktion mit dem Testmaterial zulassen. Von computerbasierten Simulationen, in denen Testpersonen ähnlich handeln

können wie in realen Situationen, darf eine höhere prognostische Validität erwartet werden, als von Papier-Bleistift-Tests (Frey & Hartig, 2013, S. 55; Jude & Wirth, 2007, S. 49). Berücksichtigen entsprechende Simulationen außerdem die Situationsgebundenheit von Kompetenzen, führt das auch zu einer Erhöhung der Testvalidität (Schaper, 2009, S. 28).

Jurecka & Hartig (2007, S. 43) merken an, dass sich die Äquivalenzfrage nur stellt, wenn ein Testverfahren durch ein anderes ersetzt oder dauerhaft zwei Testformen parallel verwendet werden sollen. Für die Autoren stellt sich die Frage demnach nicht, wenn ein computerbasierter Test neu entwickelt wird und kein entsprechendes Verfahren in Papierform existiert und auch nicht vorgesehen ist. Demgegenüber argumentieren Jude & Wirth (2007, S. 56) dass Computersimulationen nicht nur mit bereits vorhandenen traditionellen Tests verglichen werden müssen, sondern die in ihnen simulierten Situationen auch mit entsprechenden realen Situationen, um aus dem Testverhalten Rückschlüsse auf die tatsächliche Kompetenz ziehen zu können.

Ob ein computerbasierter Test die gleiche Fähigkeit misst, wie sie in realen Anforderungssituationen benötigt wird, ist bislang nur vereinzelt Gegenstand wissenschaftlicher Forschung. Im naturwissenschaftlichen und technischen Unterricht ist der Erwerb experimenteller Kompetenzen von großer Bedeutung (Schreiber, Theyßen & Schecker, 2014, S. 162; Walker, 2013, S. 15). Die hierfür benötigten Realexperimente sind jedoch zum einen mit einem hohen zeitlichen und monetären Aufwand verbunden (Schreiber, Theyßen & Schecker, 2014, S. 162) und zum anderen mit Unsicherheiten behaftet (Baumann, Simon, Wonisch et al., 2013, S. 306). Ein nicht exakt nach Anleitung durchgeführter Versuch kann zur einer Abweichung der Ergebnisse oder zum Scheitern des kompletten Versuchs führen, was zur Enttäuschung und dann zu einem Interessensverlust seitens der Schülerinnen und Schüler führen kann (ebd.). In einer von Schreiber, Theyßen & Schecker (2014) durchgeführten Studie wurde untersucht, ob ein Test mit einem Realexperiment durch einen Test mit einem Simulationsbaukasten ersetzt werden kann. Für die Beantwortung ihrer Frage betrachten die Autoren die prozessorientierten Schülerleistungen bei der Durchführung zweier Experimente (Realexperiment vs. Simulation) und konnten keine hinreichend hohen Korrelationen finden (ebd., S. 170). Baumann, Simon, Wonisch et al. (2013) verglichen Experiment und Simulation im Biologieunterricht hinsichtlich des durch die jeweilige Unterrichtsmethode erzielten Wissenszuwachs miteinander. Neben dem Ergebnis, dass sowohl die Anwendung der computersimulierten Lernumgebung als auch das eigenständige Experimentieren zu einem vergleichbaren nachhaltigen Wissenszuwachs bei den Schülerinnen und Schüler führt, berichten sie weitere positive Erfahrung im Umgang mit der Computersimulation. Beispielsweise konnten die Schülerinnen und Schüler eigenständig mit der Simulation umgehen, wohingegen in den Realexperimenten Hilfestellungen nötig sind (ebd., S. 309). Damit die Anforderungen zwischen einer realen und einer simulierten Situation überhaupt vergleichbar sind, muss die Simulation der Wirklichkeit möglichst gut entsprechen, in dem Aussehen (durch Grafiken) und Verhalten (Zulassen fehlerhaften Verhaltens) möglichst viel Ähnlichkeiten zur Realität haben (ebd.). Diese Aussage stimmt mit der von Funke & Reuschenbach (2011, S. 602) überein, nach der die Passung zwischen Simulation und realer Anforderung die Höhe der Validität bestimmt.

Mit dem Ziel für internationale Vergleichsstudien eine Alternative zu den aufwändigen und kostenintensiven Arbeitsproben in der Berufsbildung zu finden, wurden in diesem Bereich entsprechend detailgetreue Simulationen entwickelt (z. B. bei Gschwendtner, Abele & Nickolaus, 2009; Abele, Gschwendtner & Nickolaus, 2009b; Walker, Link & Nickolaus, 2015). Die Eignung der Computersimulationen für eine valide Kompetenzmessung wurde für mehrere technische Berufe untersucht, indem reale Arbeitsproben mit den entsprechenden simulierten Situationen verglichen wurden. Für den Beruf des Kfz-Mechatronikers/der Kfz-Mechatronikerin wurden verschiedene Fehlerdiagnoseaufgaben sowohl an einem realen Fahrzeug als auch in der entsprechenden Simulation konstruiert (Abele, Gschwendtner & Nickolaus, 2009a, S. 252). In dieser Studie konnte zwischen der Diagnoseleistung in der Realität und der in der Simulation ein sehr hoher Zusammenhang ($r = .94$) gezeigt werden, so dass für die Forscher keine Anzeichen dafürsprechen, dass „die Arbeit mit der Simulation andere Fähigkeiten erfordert als die Arbeit am Fahrzeug“ (ebd., S. 253). Für den Beruf des Elektrikers/der Elektrikerin für Automatisierungstechnik sind zur Erfassung der Fehlerdiagnosefähigkeit ebenfalls verschiedene Fehler in einer realen und einer simulierten Automatisierungsanlage implementiert worden (Walker, Link & Nickolaus, 2015). Auch hier korrelierten die Diagnoseleistungen in den realen und den simulierten Anforderungskontexten sehr hoch miteinander ($r = .99$), sodass davon ausgegangen wird, dass beide Instrumente dieselbe Fähigkeit erfassen (Walker, Link & Nickolaus, 2015, S. 232f.).

Zusammenfassend ist davon auszugehen, dass mit einer sorgfältig geplanten Simulation (hohe *correspondence* nach Gray (2002)) die gleichen Kompetenzen erfasst werden können, wie in realen Anforderungssituationen (vgl. Gschwendtner, Abele & Nickolaus, 2009, S. 574). Nur unter dieser erfüllten Validitätsbedingung können die sich aus dem Einsatz von Computersimulationen ergebenden Vorteile für die Kompetenzdiagnostik genutzt werden. Einer der Vorteile ist der durch eine standardisierte Testvorgabe, -bearbeitung und -auswertung verringerte Einfluss durch Testleitung und Testumgebung und damit eine Erhöhung der Objektivität der Messung (Jude & Wirth, 2007, S. 56). Eine Verringerung von Mess- und Auswertungsfehlern, zum Beispiel durch die Übertragung der Testdaten in ein computergestütztes Format, führt außerdem zu einer Verbesserung der Reliabilität der Messung (Jurecka & Hartig, 2007, S. 44). In Studien, in denen das Konstrukt der Problemlösekompetenz untersucht werden soll, sind aufgrund der Operationalisierung des Kompetenzkonstrukts nicht nur ergebnisorientierte Indikatoren von Bedeutung, sondern auch Prozessindikatoren. Erfolgt die Testung mit einem Computer, kann der Input, der von den Probanden vorgenommen wird, in sogenannten Logfiles aufgezeichnet werden. Logfiles bezeichnen automatisch erstellte Dateien, in denen bestimmte Ereignisse elektronisch aufgezeichnet werden (Schmitz & Yanenko, 2014, S. 847). Ereignisse können Klickaktionen, Drag-and-Drop-Vorgänge, Texteingaben sowie die zugehörigen Reaktions- und Bearbeitungszeiten sein, die ebenfalls einen diagnostischen Wert haben (Funke & Reuschenbach, 2011, S. 606). Die hier allerdings anfallende große Datenmenge kann für die auswertende Person auch eine Herausforderung darstellen (ebd.). Schwierigkeiten ergeben sich aber nicht erst mit der Analyse der entstandenen Prozessdaten, sondern bereits bei der Erstellung von Computersimulationen. Um auch sehr realitätsnahe Simulationen exakt beschreiben, entwickeln und anschließend technisch umsetzen zu können, bedarf es eines formalen Modells, das im nächsten Abschnitt näher beschrieben wird.

4.3 Finite Automaten als formale Modelle zur Erstellung von computersimulierten technischen Geräten

Mit dem Ziel eine bestmögliche Annäherung an realistische Problemstellungen zu erreichen, wurden in früheren Studien zur Erforschung dynamischen Problemlösens hochkomplexe Szenarien eingesetzt. Sie bestehen aus einer Anzahl stark vernetzter Variablen, von denen einige zufällig variieren und deren Beziehungen zueinander teilweise nicht linear sind (Funke, 1983b, S. 284f.). Kritisiert werden diese Systeme aufgrund der fehlenden Möglichkeit eine optimale Eingriffsfolge für ihre gezielte Steuerung anzugeben. Ohne das Vorhandensein einer idealen *Musterlösung* lässt sich das Eingriffsverhalten der problemlösenden Person jedoch nicht objektiv bewerten (Strauß & Kleinmann, 1995, S. 112f.). Die für die Diagnostik notwendige Kenntnis der Anforderungen, die das computersimulierte Szenario an die problemlösende Person stellt, macht eine exakte Beschreibung der formalen Systemstruktur erforderlich (ebd., S. 106). Die Betrachtung beziehungsweise die gezielte Variation von Simulationen hinsichtlich formaler Merkmale wie Variablen und Vernetzungen erleichtert zum einen die Konstruktion von Systemen beliebiger Schwierigkeit und zum anderen deren Vergleichbarkeit (Funke, 2004, S. 92). Jedoch lassen sich die computersimulierten Szenarien anhand dieser formalen Merkmale nicht vollständig beschreiben (Strauß & Kleinmann, 1995, S. 107), es ist außerdem notwendig das Verhalten und die Eigenschaften der Systeme auf einer benutzerrelevanten Ebene zu charakterisieren (Funke & Gerdes, 1993, S. 44).

Da es sich bei den komplexen Szenarien um modellierte Systeme handelt, kann für deren nähere Beschreibung die Systemtheorie beziehungsweise die Kybernetik herangezogen werden (Tauschek, 2006, S. 60). Wallach schlägt für die Modellierung technischer Systeme die folgenden Formalismen vor:

- Algebraische Gleichungen,
- Finite Automaten,
- Differenzgleichungen,
- Differentialgleichungen (Wallach, 1997, S. 142).

Davon sind in der späteren Problemlöseforschung vor allem lineare Differenzgleichungen (lineare Gleichungssysteme) und finite Automaten verwendet worden, die trotz ihrer einfachen formalen (mathematischen) Struktur die entsprechenden Systeme auf einem hohen Auflösungsgrad beschreiben können (Strauß & Kleinmann, 1995, S. 113f.). Problemlöseszenarien, in denen Regelungs- und Steuerungsprozesse implementiert werden sollen, lassen sich formal als lineare Gleichungen darstellen. Hierbei werden kontinuierliche Eingangssignale so verändert, dass sie je nach Art des Zusammenhangs ein ebenfalls kontinuierliches Ausgangssignal erzeugen. Als technisches Beispiel sei hier ein einfaches Heizungssystem (ohne Störgrößen) genannt, bei dem die Weite der Thermostatöffnung die Höhe der Temperatur bestimmt (Stemann & Lang, 2014, S. 92). Die auf Basis linearer Gleichungssysteme konstruierten Systeme können theoretisch unendlich viele Zustände annehmen; je nachdem wie groß das Eingangssignal gewählt wird, ergeben sich unterschiedlich große Ausgangssignale (Funke, 2004, S. 95f.). Schwierigkeiten im Umgang mit solchen Systemen resultieren aus der Anzahl der (sich beeinflussenden) Variablen, den Gewichtungen ihrer Zusammenhänge sowie eventuell vorhandenen Nebeneffekten oder

Eigendynamiken (Greiff, 2012, S. 33f.). Lineare Differenzgleichungen (ebenso wie die von Wallach genannten Algebraischen Gleichungen und Differentialgleichungen) als Formalismus für die Erstellung von Simulationen, setzen die Verwendung kontinuierlicher Ein- und Ausgangsvariablen voraus (Strauß & Kleinmann, 1995, S. 113f.). Viele technische Systeme wie Bankautomaten, Kaffeemaschinen oder Rollladensteuerungen werden hingegen durch diskrete Eingabehandlungen bedient, für deren formale Konstruktion die Theorie finiter Automaten geeigneter erscheint (Funke & Gerdes, 1993, S. 44). Mit ihr lassen sich solche Systeme definieren sowie ihre Eigenschaften losgelöst von der tatsächlichen physikalischen Umsetzung beschreiben (ebd., S. 45).

Ein *finiter (endlicher) Automat* ist ein Modell zur Beschreibung eines Systems, das sich zu jedem gegebenen Zeitpunkt, in einem der, aus einer *endlichen* Menge von Zuständen befindet (Hopcroft, Motwani & Ullman, 2011, S. 25). Ein *Zustand* ist dabei definiert als die Zeitspanne, in der ein System auf ein Ereignis wartet. Als Reaktion auf Ereignisse durchläuft ein System der Reihe nach verschiedene Zustände (Balzert, 2005, S. 87). Finite Automaten werden nach *deterministischen* und *nichtdeterministischen* finiten Automaten unterschieden. Ein Automat ist deterministisch, wenn sich der Automat zu jedem Zeitpunkt in nur einem Zustand befindet und nichtdeterministisch, wenn der Automat zeitgleich mehrere Zustände besitzt (Hopcroft, Motwani & Ullman, 2011, S. 62). In einen Zustand gelangt ein System durch ein entsprechendes Ereignis, das selbst von vernachlässigbarer Dauer ist, so dass ein System sich immer in einem Zustand befindet und nie im Übergang zwischen zwei Zuständen (Balzert, 2005, S. 87; Lunze, 2009, S. 42). In Abhängigkeit der Ereignisfolgen kann ein System zwischen Zuständen hin- und herspringen; der nachfolgende Zustand ist damit das Resultat aus dem vorherigen Zustand und dem erfolgten Ereignis, so dass der aktuelle Systemzustand implizit Informationen über die bisherigen Eingaben beinhaltet (Balzert, 2005, S. 87).

Weil Ereignisse die Ursache für Veränderungen von Zuständen innerhalb eines Systems sind, werden finite Automaten auch als *ereignisdiskrete* Systeme bezeichnet (vgl. Lunze, 2009, S. 42). Ereignisse können danach unterschieden werden, ob es sich bei ihnen um externe Eingaben oder interne Ereignisse handelt. Externe Eingaben entsprechen in technischen Systemen Benutzereingaben (beispielsweise das Drücken einer Taste) (Funke, 2004, S. 93), die zum Aufrufen entsprechender Operationen führen (*Aufrufereignis*) (Noack, 2001, S. 128). Zu den internen Ereignissen zählen Ereignisse, die durch das Ablauf einer Zeitspanne oder das Eintreten eines konkreten Zeitpunktes (*Zeitereignis*) erfolgen, sowie Ereignisse, die durch eine Änderung bestimmter Werte ausgelöst werden (*Änderungsereignis*) (ebd.).

Für die exakte Beschreibung computersimulierter Systeme innerhalb des Rahmenmodells finiter Automaten dient die sogenannte *Zustandsübergangsfunktion* (kurz *Übergangsfunktion* oder *Transitionsfunktion*) (Hopcroft, Motwani & Ullman, 2011, S. 72). Die Übergangsfunktion gibt für einen gegebenen Systemzustand die möglichen Zustandsübergänge mit den entsprechenden Ein- und Ausgaben an (Funke & Buchner, 1992, S. 28).

Zusammenfassend definieren Hopcroft, Motwani & Ullman (2011, S. 71f.) einen deterministischen finiten Automaten bestehend aus:

- einer endlichen Menge von Zuständen,
- einer endlichen Menge von Eingabesymbolen,
- einer Übergangsfunktion, der ein Zustand und ein Eingabesymbol als Argumente übergeben werden und die einen Zustand zurückgibt,
- einem Startzustand und
- einer Menge finaler oder akzeptierter Zustände.

Die von Hopcroft, Motwani & Ullman beschriebenen endlichen Automaten werden heute als Grundform ergebnisdiskreter Modelle aufgefasst, die in vielfältiger Weise erweitert wurden (Lunze, 2009, S. 147). Eine der grundlegendsten Erweiterungen wurden von Harel (1987) in seinem Konzept der *statecharts* (bekannt als Harel-Zustandsautomat) veröffentlicht. Eine der Erweiterungen betrifft die Einführung von Bedingungen für den Zustandsübergang. Befindet sich ein System in einem bestimmten Zustand und tritt dann ein entsprechendes Ereignis ein, wird zunächst die Bedingung überprüft (Abbildung 4.3). Ist die Bedingung erfüllt, erfolgt die Transition, ist sie nicht erfüllt, bleibt das System in dem Zustand und das Ereignis geht verloren. Ist die Bedingung erst zu einem späteren Zeitpunkt erfüllt, muss das entsprechende Ereignis erneut eintreten, um einen Zustandsübergang zu erreichen (Seidl, Brandsteidl, Huemer et al., 2012, S. 93). Diese Bedingungen basieren auf der Einführung von Variablen, die bei Transitionen verändert werden können (Lunze, 2009, S. 148), indem bei der Transition außerdem Aktionen ausgeführt werden (Noack, 2001, S. 129). Hierzu werden zwei Arten von Transitionen unterschieden: Befindet sich ein System in einem Zustand und tritt ein Ereignis ein, das dazu führt, dass das System in einen anderen Zustand übergeht, spricht man von äußerer Transition. Möglich ist aber auch, dass man als Zielzustand den Ausgangszustand definiert und als Reaktion auf ein entsprechendes Ereignis nur eine Aktion ausführt, die zum Beispiel eine Variable ändert. Hier spricht man von innerer Transition oder Selbsttransition (Seidl, Brandsteidl, Huemer et al., 2012, S. 100; Balzert, 2005, S. 91).

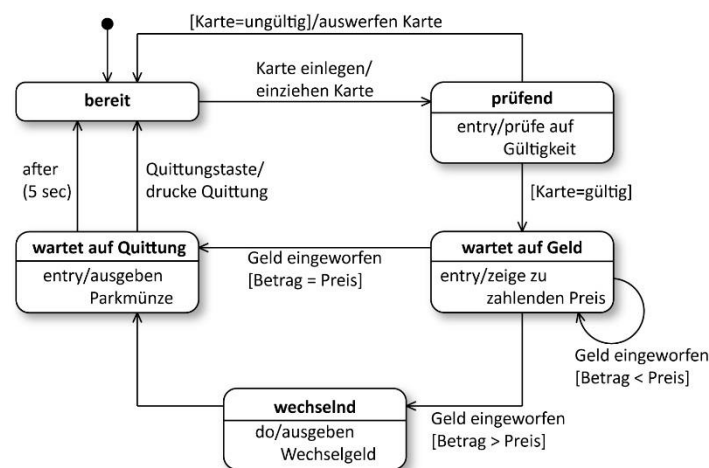


Abbildung 4.3: Zustandsdiagramm eines Automaten zum Bezahlen der Parkgebühr in einem Parkhaus (Balzert, 2005, S. 92); Bedingungen stehen in eckigen Klammern

Weitere konzeptionelle Erweiterungen endlicher Automaten betreffen die Komplexität der Zustände. So sieht Harel (1987, S. 234) *hierarchische* und *orthogonale Zustände* vor. Hierarchische Zustände beinhalten weitere sogenannte Subzustände. Wenn sich das System in einem hierarchischen Zustand befindet, befindet es sich immer nur in einem der Subzustände. Die in einem Oberzustand enthaltenen Subzustände dienen dabei der Verfeinerung von Beschreibungen des Systemverhaltens. Als Beispiel sei das technische System Radiowecker genannt, dessen Funktion *Alarm stellen* als Oberzustand modelliert werden kann, in dem die Subzustände *Stunde stellen*, *Minute stellen* und *Alarm einschalten* enthalten sind (Abbildung 4.4).

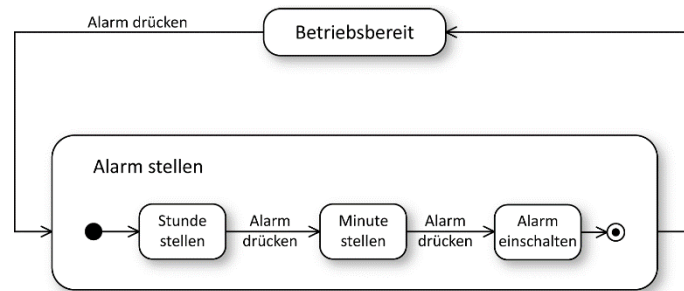


Abbildung 4.4: Beispiel für das Betreten und Verlassen eines hierarchischen Zustands

Orthogonale Zustände sind Zustände die in zwei oder mehrere Regionen aufgeteilt werden, von denen zu einem Zeitpunkt jeweils ein Zustand *jeder* Region aktiv ist (Seidl, Brandsteidl, Huemer et al., 2012, S. 105). Die Zustandswechsel innerhalb einer Region können dabei unabhängig von Systemveränderungen in anderen Regionen sein. Als Beispiel dient hier wieder das System Radiowecker. Durch die Modellierung eines Oberzustandes, der aus den Regionen *Display*, *Alarmstatus* und *Licht* besteht, können Zustandswechsel in den jeweiligen Regionen unabhängig von den Zuständen der jeweils anderen Regionen stattfinden (Abbildung 4.5). Der Zustand *Licht ein* hat dann keine Auswirkungen auf den Zustand *Alarm aus* (Harel, 1987, S. 248).

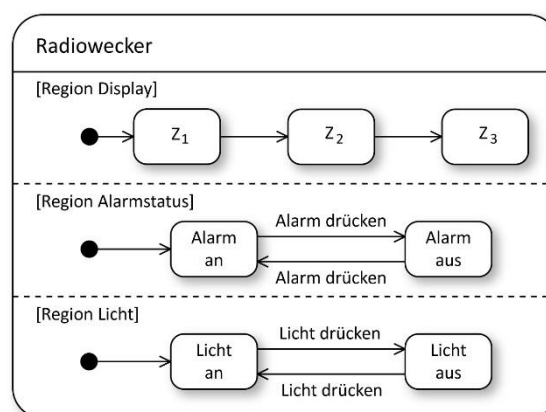


Abbildung 4.5: Beispiel eines orthogonalen Systemzustandes

Um das Verhalten eines technischen Systems auf der Basis des (erweiterten) Modells finiter Automaten exakt zu beschreiben, muss für jeden Zustand die Übergangsfunktion ermittelt werden, die beschreibt mit welchem Ereignis, unter welcher Bedingung und mit welcher Aktion der Automat in welchen Zustand wechselt. Die Spezifikation eines endlichen

Automaten mit diesen Übergangsfunktionen ist zwar nötig, allerdings schwer lesbar (Hopcroft, Motwani & Ullman, 2011, S. 74) und eignet sich nur bedingt als Kommunikationsmittel im Erstellungs-, Anwendungs- und Auswertungsprozess von computersimulierten Testinstrumenten. Neben dieser textuellen Notation von Zustandsübergängen stellen Zustandsübergangsdiagramme eine weitere Möglichkeit zur Visualisierung von Automaten dar. Die in ihnen verwendeten Zeichen und Symbole werden in Modellierungssprachen vereinheitlicht, die klare Regeln zur strukturierten Konstruktion von Modellen zur Verfügung stellen (Seidl, Brandsteidl, Huemer et al., 2012, S. 2). Die Sprache UML (Unified Modeling Language) ist die heutzutage am weitesten verbreitete, universell einsetzbare Modellierungssprache (Heinrich & Mairon, 2008, S. 11), mit der Zustandsübergänge auch grafisch beschreibbar sind.

Basierend auf dem von Harel (1987) verallgemeinerten Konzept finiter Automaten, stellen *Zustandsdiagramme* in der UML eine einheitliche Notation zur Darstellung von Zuständen und ihren Übergängen, als Reaktion auf externe und interne Ereignisse zur Verfügung (Brügge & Dutoit, 2004, S. 89). Zustandsdiagramme kommen immer dann zum Einsatz, wenn das Verhalten des zu betrachtenden Systems modelliert werden soll, die UML spricht dann auch von einem Verhaltenszustandsautomat (behavioral state machine) (Balzert, 2005, S. 92f.). Ein Zustandsdiagramm ist ein Graph, in dem die Zustände als Knoten und die Zustandsübergänge als Kanten definiert sind (Hopcroft, Motwani & Ullman, 2011, S. 74). In dem Diagramm werden die Zustände mit Rechtecken dargestellt, die den Namen des jeweiligen Zustands enthalten. Optional können in dem Rechteck außerdem Aktionen genannt werden, die beim Eintreffen in den Zustand (entry), während seiner Verweildauer (do) und beim Verlassen des Zustandes (exit) ausgeführt werden. Die Zustandsübergänge werden durch einen Pfeil symbolisiert, dessen Pfeilspitze die Richtung anzeigt, in die der Zustandsübergang erfolgt (Seidl, Brandsteidl, Huemer et al., 2012, S. 93). Beschriftet wird der Pfeil mit dem Ereignis, das den Zustandsübergang auslöst (bei mehreren Ereignissen eine Liste), den eventuell vorhandenen Bedingungen (in eckigen Klammern [...]) an die der Zustandsübergang gebunden ist, sowie etwaigen Aktionen (nach einem schrägen Strich /), die während des Übergangs ausgeführt werden (ebd.). Jeder Automat besitzt einen Anfangszustand (initial state) als Startpunkt für seine Ausführungen, von dem aus eine Transition in den ersten Zustand erfolgt. Der Anfangszustand wird mit einem ausgefüllten Kreis veranschaulicht. Ein Automat kann außerdem einen Endzustand haben, der durch einen ausgefüllten Kreis mit einem umschließenden Ring symbolisiert wird. Aus ihm führen keine Transitionen mehr hinaus (Balzert, 2005, S. 90f.). Abbildung 4.6 fasst die in der UML vereinheitlichten Regeln für die Darstellung von Zustandsübergangsdiagrammen zusammen.

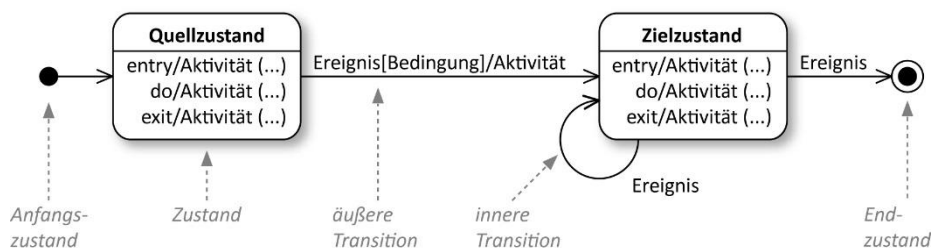


Abbildung 4.6: Notation eines einfachen Zustandsautomaten

Die Verwendung der Automatentheorie zur formalisierten Erstellung und Beschreibung von computersimulierten technischen Systemen bietet für deren Einsatz in einem Kompetenztest mehrere Vorteile. Für die Konstruktion von Computersimulationen erleichtern Zustandsübergangsdiagramme die strukturierte Erstellung eines Modells, aus dem dann manuell ein ausführbarer Code abgeleitet werden kann. Schon im Modellierungsprozess lassen sich schwierigkeitsbestimmende Merkmale technischer Systeme berücksichtigen, wie etwa die Komplexität eines technischen Systems durch die Variation der Anzahl von Zuständen und Ereignissen sowie Eigendynamiken, indem autonom ausgelöste Ereignisse integriert werden. Ein weiterer Vorzug betrifft die Möglichkeit bei gegebener Problemstellung eine optimale Eingriffsfolge für die Lösung angeben zu können. Damit (und der Möglichkeit Verhaltensdaten aufzeichnen zu können) lässt sich das Eingriffsverhalten der problemlösenden Person objektiv bewerten, womit dem Anspruch nach prozessorientierten Indikatoren bei Bedarf nachgekommen werden könnte.

Ein für diese Arbeit wichtiger Aspekt bezieht sich auf die theoretische Einbettung des mithilfe finiter Automaten darstellbaren Problemlöseprozesses. Die Zustandsübergänge eines endlichen Automaten lassen sich mit den Produktionsregeln, den kleinsten Einheiten des prozeduralen Wissens aus der ACT-R Theorie von Anderson vergleichen (Funke & Gerdes, 1993, S. 45). Die Produktionsregeln bestehen aus einem *Wenn*- und einem *Dann*-Teil¹. Die *Wenn*-Komponente besteht aus dem Quellzustand, in dem sich der Automat aktuell befindet und dem Zielzustand, in den der Automat überführt werden soll (Beispiel: *Wenn* der Backofen ausgeschaltet ist (Quellzustand) und er eingeschaltet werden soll (Zielzustand)). In der *Dann*-Komponente, werden die Eingaben spezifiziert, die den entsprechenden Zustandsübergang auslösen (Beispiel: ..., *dann* muss der Power-Button gedrückt werden) (ebd.). Der problemlösende Umgang mit technischen Systemen erfordert das Erlernen solcher Prozeduren (Eingriffswissen), um das Verhalten eines technischen Gerätes für dessen Steuerung vorhersagen zu können.

Der Systematisierung von Simulationen in realitätsnahe Szenarien und solche, die auf Basis von formalen Modellen erstellt wurden (Funke & Reuschenbach, 2011, S. 608), muss an dieser Stelle widersprochen werden. Die Verwendung finiter Automaten als formales Modell muss nicht zwingend zu Lasten der Realitätsnähe gehen. Sowohl das Verhalten, als auch das Aussehen computersimulierter technischer Systeme können hoch mit den realen technischen Alltagsgeräten korrespondieren, womit der Vorteil verbunden ist, dass die Fähigkeit im Umgang mit dem Computermodellen dieselbe ist, wie im Umgang mit den realen Geräten. Nachteilig ist der höhere Aufwand für die Programmierung und die grafische Umsetzung von realitätsnahen Simulationen.

¹ Produktionen als Bestandteil des prozeduralen Wissens kognitiver Systeme werden in Abschnitt 2.4.2.1 näher erläutert.

5 Einflussfaktoren auf den Prozess und die Leistung beim Problemlösen

Für das Zustandekommen einer Problemlöseleistung als Ergebnis eines Problemlöseprozesses sind, neben der erforderlichen Problemlösekompetenz, die im Kapitel 3 betrachtet wurde, weitere Einflussfaktoren verantwortlich. Demnach empfiehlt es sich diese Einflussfaktoren in die Betrachtung des Problemlöseprozesses und der Problemlöseleistung einzubeziehen (Betsch, Funke & Plessner, 2011, S. 140). Der Fokus liegt im Folgenden auf den Konstrukten, die einen Einfluss auf das Problemlöseverhalten haben und damit eine Erklärung für individuelle Unterschiede liefern können. Hinweise darauf, welche Konstrukte das Handeln in konkreten Problemsituationen beeinflussen, finden sich beispielsweise bei Frensch & Funke (1995). In ihrem theoretischen Rahmenmodell für komplexes Problemlösen unterscheiden sie subjektive Faktoren nach kognitiven und nichtkognitiven Variablen. Zu den kognitiven Variablen zählen sie das Hintergrundwissen, Selbstreflektions- und Selbstregulationsstrategien, kognitive Stile und allgemeine Intelligenz. Als nichtkognitive Variablen, die auf Seiten der problemlösenden Person eine Situation beeinflussen können, nennen die Autoren Selbstvertrauen, Ausdauer, Motivation und Freude beim Problemlösen (Frensch & Funke, 1995, S. 20f.). Ausubel, Novak & Hanesian zählen zu den intrapersonalen Faktoren außerdem die Aufgeschlossenheit, intellektuelle Neugier, Flexibilität, Interesse und Erfahrung einer Person (Ausubel, Novak & Hanesian, 1981, S. 666).

Während Intelligenz, kognitive Stile und Einstellungen gegenüber dem Nachdenken als Eigenschaften einer Person unabhängig von dem Kontext, in dem die Person handelt, sind, wird das Vorwissen, die Erfahrung und die Motivation einer Person sehr wohl von der konkreten Problemsituation bestimmt. Die Unterteilung in Einflussfaktoren, die unabhängig von der Situation und solche die kontextsensitiv sind, werden in diesem Kapitel aufgegriffen. Zunächst werden diejenigen Merkmale vorgestellt, die situationsübergreifend das Verhalten von Personen beeinflussen. Intelligenz stellt dabei die stärkste Determinante dar, weshalb ihr Einfluss auf den problemlösenden Umgang ausführlich beschrieben wird. Die Existenz einer allgemeinen Problemlösefähigkeit als ein Konstrukt, das problemübergreifend zusätzlich zur Intelligenz interindividuelle Unterschiede im Problemlöseverhalten erklären kann, wird in der Literatur kontrovers diskutiert. Der Forschungsstand zur allgemeinen Problemlösefähigkeit als eigenständiges Konstrukt wird wiedergegeben und die Theorieeinbindung allgemeiner Problemlösefähigkeit diskutiert. Der zweite Teil des Kapitels kommt dann auf die Faktoren zu sprechen, die in Abhängigkeit von der jeweiligen Situation variieren können, wie etwa das in einer Situation relevante Wissen, die Motivation oder das Interesse an der Lösung des entsprechenden Problems.

5.1 Kontextübergreifende Einflussfaktoren

5.1.1 Intelligenz

Zu den bereits am intensivsten erforschten Merkmalen der Psychologie gehört die Intelligenz (Rost, 2009, S. 1) und auch der Zusammenhang zwischen Intelligenz und Problemlösen ist bereits seit Langem Gegenstand vieler Untersuchungen (z. B. Dörner, 1984; Dörner & Kreuzig, 1983; Funke, 1983b; Hörmann & Thomas, 1989; Hussy, 1989, 1990, 1991; Kersting, 1999; Leutner, 2002; Putz-Osterloh & Lüer, 1981; Strohschneider, 1991b; Süß, 1999). Trotz

umfangreicher Forschungsaktivitäten konnte man sich bislang auf keinen einheitlichen Intelligenzbegriff einigen, weshalb es eine unüberschaubare Anzahl an Definitionen gibt (Rost, 2009, S. 2; Meili & Steingrüber, 1978, S. 51). In den meisten Definitionen wird der Begriff der Intelligenz jedoch mit globalen kognitiven Fähigkeiten wie die zum abstrakten und schlussfolgernden Denken, sich an neue Situationen anzupassen und neue Anforderungen zu bewältigen (Fähigkeit zum Problemlösen) bzw. ganz allgemein als die Fähigkeit zum Lernen in Verbindung gebracht (vgl. z. B. Mack, 1999, S. 122; Rost, 2009; Wirth, 2004; Seel, 2003). Die von Stern (1920) gegebene Definition lautet: „Intelligenz ist die allgemeine Fähigkeit eines Individuums, sein Denken bewußt auf neue Forderungen einzustellen; sie ist allgemeine geistige Anpassungsfähigkeit an neue Aufgaben und Bedingungen des Lebens“ (Stern, 1920, S. 2f). Die intellektuellen Voraussetzungen, Wissen erwerben zu können werden dabei als unabhängig von bereits gemachten Erfahrungen und bestehenden Kultureinflüssen angenommen (Mack, 1999, S. 124). Wissen kann dennoch einen Hinweis auf die Ausprägung von Intelligenz geben, denn die Fähigkeit zu lernen kann (unter der Bedingung vorhandener Lerngelegenheiten) zu einer entsprechenden Wissenszunahme führen (Meili & Steingrüber, 1978, S. 51f.).

Zentrales Anliegen der Intelligenzforschung war es lange Zeit Intelligenz in ihrer Struktur zu erforschen (Schweizer, 2006a, S. 4). Der Auffassung, dass sich Intelligenz auf eine Maßzahl „g“ (generell intelligence) reduzieren lässt, stehen Annahmen gegenüber, die verschiedene Facetten von Intelligenz differenzieren (Süß & Beauducel, 2011, S. 98). Die Unterscheidung verschiedener Formen von Intelligenz basiert auf der Beobachtung individueller Leistungsunterschiede im Lösen bestimmter Aufgaben. Auf Basis von Korrelationsrechnungen, die Übereinstimmungen im Lösungsverhalten aufdecken, hat Spearman seine *Zwei-Faktoren-Theorie* aufgebaut, in der er neben einem gemeinsamen Faktor „g“ einen weiteren Faktor „s“ postuliert, der die Wirkung spezifischer Testbedingungen repräsentiert (vgl. Seidel & Krapp, 2014, S. 178; Meili & Steingrüber, 1978, S. 55). Im Gegensatz zu Spearmans Generalfaktor „g“ geht Thurstone (1938) von neun – später sieben – Primärfaktoren aus: *induktives Denken, deduktives Denken, Wahrnehmungsgeschwindigkeit, Merkfähigkeit, Sprachverständnis, Wortflüssigkeit, Rechenfähigkeit, arithmetisches Schlussfolgern und räumliches Vorstellungsvermögen* (Süß & Beauducel, 2011, S. 107).

Da im Laufe der Zeit immer wieder neue Faktoren definiert wurden (Guilford beschreibt 120 Faktoren), denen zum Teil eine Abhängigkeit nachgewiesen werden konnte (vgl. Myers, 2014, S. 401), bemühte sich Cattell (1971) Ordnung in die Faktorenvielzahl zu bringen (vgl. Meili & Steingrüber, 1978, S. 61). Kern seines hierarchischen Intelligenzmodells ist die Differenzierung zwischen *fluiden* und *kristalliner Intelligenz* als Teilkomponenten (Faktoren zweiter Ordnung) der allgemeinen Intelligenz. Fluide Intelligenz g_f stellt dabei die Fähigkeit Anforderungen in neuen Situationen ohne Vorwissen bewältigen zu können dar. Sie ist überwiegend genetisch determiniert (Gruber & Stamouli, 2015, S. 31; Schweizer, 2006a, S. 5) und bestimmt die in dieser Arbeit im Fokus stehende Kompetenz zum Lösen von Problemen (vgl. Holling, Preckel & Vock, 2004, S. 21). Kristalline Intelligenz g_c beschreibt die kognitive Fähigkeit bereits erworbenes (kristallisiertes) Wissen anzuwenden und vertraute Informationen zu verarbeiten (Burmeister, 2009, S. 58). Im Rahmen seiner *Investment-Theorie* nimmt Cattell (1971) außerdem an, dass fluide Intelligenz die Voraussetzung für die Bildung kristalliner Intelligenz ist. Eine Person investiert ihre fluide Intelligenz seit ihrer Geburt in das Sammeln von Lernerfahrungen

und damit in den Aufbau von kristalliner Intelligenz (Rüppel, Hinnermann & Wiegand, 1987, S. 178). Individuelle Unterschiede in der fluiden Intelligenz stellen damit individuelle Unterschiede in den grundlegenden Voraussetzungen zum Wissenserwerb dar. Fluide Intelligenz ist für den Lernzuwachs in neuen Gebieten demnach wichtiger als kristalline Intelligenz (Rost, 2009, S. 52). Letztere ist stark von dem jeweiligen kulturellen Kontext und den Lerngelegenheiten geprägt (Artelt & Wirth, 2014, S. 178).

Die häufig in der Literatur zu findende Vorstellung, nach der fluide Intelligenz ganz oder zumindest zu einem großen Teil mit der biologischen kognitiven Struktur in Verbindung gebracht wird (z. B. Guthke & Beckmann, 2004, S. 144f.), lässt sich über die Beziehung zur Arbeitsgedächtniskapazität auch theoretisch erklären. So umfasst fluide Intelligenz, als die Fähigkeit zum Lernen, Prozesse des Bewusstwerdens von Informationen (Aufmerksamkeit) und des Haltens unterschiedlicher Informationsaspekte im Arbeitsgedächtnis (Rost, 2009, S. 56). Durch die Begrenzung der Arbeitsgedächtniskapazität ist die Menge an Informationen, die vorübergehend gehalten und wieder repräsentiert werden können, jedoch begrenzt¹. Für das aus der Kognitionspsychologie stammende Konzept des Arbeitsgedächtnisses eignet sich demnach die Metapher des Flaschenhalses, der die Leistungsfähigkeit bei der Bewältigung von kognitiven Aufgaben limitiert (vgl. Süß, 1999, S. 223). Untersuchungen, die sich dem Zusammenhang zwischen Arbeitsgedächtniskapazität und Intelligenz widmen, bekräftigen diese Vermutung.

Ausgehend von der Betrachtungsweise der kognitiven Struktur als informationsverarbeitendes System, das dem Arbeitsgedächtnis neben dem deklarativem und dem prozeduralem Gedächtnis eine wichtige Bedeutung für die kognitiven, motorischen und wahrnehmenden Prozesse zuspricht, kommen die Autoren Kyllonen & Christal (1990, S. 390) zu der Schlussfolgerung, dass sich intraindividuelle Leistungsunterschiede bei der Bewältigung kognitiver Aufgaben auf Unterschiede in den Komponenten des kognitiven Systems zurückführen lassen. Ihre Hypothese, nach der Kapazitätsdifferenzen im Arbeitsgedächtnis vorhandene Leistungsunterschiede in Intelligenztests erklären können, prüfen die Autoren in vier Studien. Dabei konnten sie Korrelationen zwischen den latent modellierten Variablen Intelligenz und Arbeitsgedächtnis von $r = .80$ bis $.88$ finden (Kyllonen & Christal, 1990, S. 426). Engle, Tuholski, Laughlin et al. (1999) können die Annahme eines engen Zusammenhanges zwischen Arbeitsgedächtnis und fluider Intelligenz in ihrer Studie ebenfalls bestätigen. Neben Skalen zur Messung fluider Intelligenz und Arbeitsgedächtniskapazität setzten sie außerdem Skalen zur Erfassung der Kapazität des Kurzzeitgedächtnisses ein. Als bedeutsam für fluide Intelligenz zeigte sich jedoch lediglich die Arbeitsgedächtniskapazität.

Die, im Abschnitt 2.4.2.1 beschriebene ACT-R-Theorie wurde nicht allein zur Erklärung des kognitiven Systems entwickelt, sondern auch mit dem Ziel die Prozesse innerhalb des kognitiven Systems zu simulieren (Anderson & Lebiere, 1998a, S. 4f.). Für die Programmierung der Simulation sind Algorithmen erforderlich, die Parameter zur Beschreibung des kognitiven Systems enthalten. Eine Größe, die die Höhe der Aufmerksamkeit auf vorhandene Zielelemente beschreibt und damit abhängig von der

¹ Die Bedeutung des Arbeitsgedächtnisses als Teil des kognitiven Systems wird in Abschnitt 2.4.2.1 näher erläutert.

Arbeitsgedächtniskapazität ist, ist der *Quellaktivierungs-Parameter (Source Activations W)* (Anderson, Lebiere & Lovett, 1998, S. 71; Schoppek & Putz-Osterloh, 2003, S. 167). Bei mehreren gleichzeitig zu betrachtenden Elementen wird die Aufmerksamkeit und damit der *Quellaktivierungsparameter W* gleichmäßig auf alle Elemente verteilt, so dass bei einer hohen Anzahl zu berücksichtigender Elemente, jedem Element nur wenig Aufmerksamkeit zuteilwird (ebd.). Bei gleichzeitiger Konstanthaltung aller weiteren Parameter, führt ein niedriger *Quellaktivierungsparameter W* im modellierten kognitiven System zu einer geringeren Leistung bei der Bewältigung kognitiver Anforderungen. Dass ein solches Modell dem tatsächlichen Verhalten von Personen entspricht, zeigen die Autoren Daily, Lovett & Reder (2001). In ihrer Studie, in der sie den Modellparameter *W* individuell, basierend auf dem Ergebnis einer das Arbeitsgedächtnis belastenden Aufgabe variieren, konnte das ACT-R-Modell die Leistungen der Personen in einer anderen Aufgabe gut vorhersagen (Daily, Lovett & Reder, 2001, S. 343ff.). Taatgen, Gray & Schunn (2002) gehen davon aus, dass sich die kognitive Architektur verschiedener Personen hinsichtlich *generelle Intelligenz, Wahrnehmungsgeschwindigkeit* und *psychomotorische Fähigkeiten* unterscheiden. Ausgehend von der Arbeit von Kyllonen & Christal (1990), in der ein enger Zusammenhang zwischen dem Arbeitsgedächtnis und Intelligenz konstatiert wurde, modellieren sie generelle Intelligenz in ACT-R, in dem sie den *W*-Parameter variieren, der eher mit der Arbeitsgedächtniskapazität assoziiert ist.

Die in vielen Studien berichteten hohen Zusammenhänge zwischen Intelligenz (allgemeine oder fluide Intelligenz) und Arbeitsgedächtnis und der daraus gezogenen Schlussfolgerung, beide Konstrukte erfassen dasselbe (Burmeister, 2009, S. 61), werden von Ackerman, Beier & Boyle (2005) kritisiert. In ihrer Metaanalyse, in der sie die in 86 Studien berichteten Korrelationen der latent modellierten Variablen an ihrer Reliabilität relativierten, kommen die Autoren nur auf eine durchschnittliche Korrelationen von $r = .48$. Ihre Ergebnisse zeigen, dass Arbeitsgedächtniskapazität und Intelligenz nicht dasselbe seien, so die Argumentation der Autoren (Ackerman, Beier & Boyle, 2005, S. 51). Die anschließende Diskussion (Kane, Hambrick & Conway, 2005; Oberauer, Schulze, Wilhelm et al., 2005; Beier & Ackerman, 2005) zeigt, dass sich die Konstrukte Arbeitsgedächtnis und Intelligenz zwar sehr ähneln, ob es sich jedoch um ein einziges Konstrukt handelt bzw. wodurch der hohe Zusammenhang zustande kommt, ist Gegenstand weiterer Untersuchungen (z. B. Colom, Abad, Quiroga et al., 2008; Bühner, Kröner & Ziegler, 2008).

Ebenso kontrovers wie über die Verbindung zwischen Intelligenz und Arbeitsgedächtniskapazität wird über den Zusammenhang zwischen Intelligenz und Problemlöseleistung diskutiert (z. B. Dörner, 1986; Hussy, 1989; Strohschneider, 1991b; Wüstenberg, Greiff & Funke, 2012). Obwohl der Einfluss kognitiver Grundfähigkeiten auf die Fähigkeit zum Lösen komplexer Probleme weitaus früher im Blickpunkt des Forschungsinteresses gerückt ist, erweisen sich die späteren Impulse aus der Intelligenz-Arbeitsgedächtnis-Debatte als durchaus gewinnbringend für die Erklärung der Beziehung zwischen Intelligenz und Problemlösen (z. B. bei Wittmann, Süß & Oberauer, 1996; Schweizer, Wüstenberg & Greiff, 2013). Wenn individuelle Unterschiede in der Kapazität des Arbeitsgedächtnisses Unterschiede in Intelligenztestleistungen vorhersagen, gleichzeitig die Arbeitsgedächtniskapazität einen Einfluss auf den Wissenserwerb beim Problemlösen hat, so sollten auch die Intelligenztestleistungen mit den Problemlöseleistungen korrelieren.

Aus der Perspektive der Intelligenzforschung, in der zumindest fluide Intelligenz definiert wird, als die Fähigkeit, sich an neue Situationen anzupassen und neue Anforderungen zu bewältigen (Cattell, 1963), kommt die Annahme, dass Intelligenztests auch allgemeine Problemlösefähigkeit erfassen sollen (Beckmann, 1994, S. 48). „Die den Intelligenztests zugeschriebene Bedeutsamkeit wäre unverdient, wenn es mit ihnen nicht möglich wäre, wesentliche Charakteristika des Denkens und Problemlösens zu prognostizieren“ (Dörner & Kreuzig, 1983, S. 185). Während Ergebnisse aus traditionellen Problemlösetests wie der *Turm von Hanoi* sehr hoch mit den Leistungen aus Intelligenztests zusammenhängen (Süß, 1996, S. 94f.), sehen die Befunde aus komplexen Problemlöseszenarien inhomogener aus¹. Die zu Beginn der komplexen Problemlöseforschung berichteten niedrigen oder teilweise auch fehlenden Korrelationen von traditionellen Intelligenztestleistungen und Steuerungsleistungen (Dörner & Kreuzig, 1983, S. 187; Putz-Osterloh, 1981; Putz-Osterloh & Lüer, 1981) ließen die Vermutung zu, dass Intelligenztests keine geeigneten Instrumente zur Prognose von Leistungen im Umgang mit komplexen Problemen sind. Ihre Validität wurde deshalb angezweifelt (z. B. Dörner, 1986, S. 292).

Schon früh wurden die Studien, aus denen entsprechende Resultate stammen, aufgrund verschiedener methodischer Aspekte kritisiert (vgl. Funke, 1983b, S. 283; Hörmann & Thomas, 1989, S. 24). In einigen der ersten Untersuchungen erfolgt die Interaktion mit dem komplexen System über eine den Versuch leitende Person. Die Probanden müssen hierbei entsprechende Fragen formulieren, um Informationen über das System zu erhalten (vgl. Beckmann, 1994, S. 53), ein Aspekt, der die Objektivität des Testes in Frage stellt (Kluwe, Schilde, Fischer et al., 1991, S. 297). Ein weiterer Kritikpunkt betrifft die unklare Zieldefinition der unterschiedlichen eingesetzten Problemszenarien. Sehr offene Problemstellungen führen dazu, dass Probanden eigene Ziele definieren, deren Erreichen nicht ohne entsprechende Informationen bewertet werden können (Strohschneider, 1991b, S. 360). Aber auch bei klar definierten Zielen, ist die Operationalisierung der Lösungsgüte in einigen Studien unzureichend. Funke (1983b) kritisiert beispielsweise die subjektive Wahl von Kriterien, anhand derer die Leistung der Probanden bewertet wird. So werden post-hoc Befunde aus Verhaltensdaten sowie aus Protokollen Lauten Denkens für die Bewertung der Problemlöseleistungen herangezogen, die selbst untereinander korrelieren (Funke, 1983b, S. 287). Nur in wenigen Studien, in denen komplexe Problemszenarien zum Einsatz kommen, werden außerdem Reliabilitätskoeffizienten genannt. In den Fällen, in denen Angaben zur Reliabilität gemacht werden, gehen die Koeffizienten aus der Korrelation zwischen den Problemlösegütemaßen aus der wiederholten Testbearbeitung hervor (Test-Retest-Reliabilität) (vgl. Kluwe, Schilde, Fischer et al., 1991, S. 299), Lern- und Behaltenseffekte können hierbei jedoch nicht ausgeschlossen werden. Eine valide Zusammenhangsbetrachtung zweier Variablen wie Intelligenz und Problemlösekompetenz ist jedoch nur unter der Bedingung ihrer reliablen Erfassung möglich (vgl. Beckmann, 1994, S. 59).

Problemlöseszenarien, die ohne die oben genannten Mängel entwickelt wurden, weisen dennoch nur einen mittelhohen Zusammenhang zu Intelligenztests auf (maximal $r = .49$ bei

¹ Eine ausführliche Zusammenschau der heterogenen Befundlage zum Zusammenhang zwischen Intelligenz und Problemlösen findet sich bei Kluwe, Schilde, Fischer et al. (1991), Süß (1996) und Burmeister (2009).

Süß, Kersting & Oberauer (1991); maximal $r = .50$ bei Kersting (1999)). Die Diskussion um den Zusammenhang zwischen Intelligenz und Problemlösefähigkeit ist deshalb noch immer aktuell, da diese Ergebnisse aus unterschiedlichen Perspektiven mit unterschiedlichen Interessen verschieden interpretiert werden können. Ein Teil der Forscher (z. B. Rüppel, Hinnersmann & Wiegand, 1987; Kersting, 1999) geht davon aus, dass fluide Intelligenz sehr wohl ein geeigneter Indikator für eine allgemeine Problemlösefähigkeit ist, denn auch wenn eine einheitliche Begriffsdefinition von Intelligenz fehlt, Problemlösefähigkeit ist ein gemeinsames Element vieler Definitionsansätze. Entgegen den theoretisch beschriebenen Anforderungen, die durch Intelligenz zu bewältigen sein sollten, erfolgt die Operationalisierung des Konstruktes in Intelligenztestaufgaben derart, dass diese Anforderungen fehlen (Kersting, 1999, S. 119). Bei den Testitems handelt es sich in der Regel um sehr kleine, nicht dynamische, nicht vernetzte und nicht transparente Problemstellungen, in denen die Art der Informationsverarbeitung nahezu vollständig vorgeschrieben wird (Kersting, 1999, S. 119; Rüppel, Hinnersmann & Wiegand, 1987, S. 182). Der Forderung Intelligenztestaufgaben zu revidieren und die beim Problemlösen hinzukommenden Anforderungen, die das spezifische Wissen stärker berücksichtigen, in die Intelligenzdiagnostik mit einzubeziehen (Rüppel, Hinnersmann & Wiegand, 1987, S. 182) wird bislang nur vereinzelt nachgekommen. Wenn Aufgaben zum komplexen Problemlösen als eine Erweiterung konventioneller Intelligenztest betrachtet werden, so sollten ihre Ergebnisse mit Intelligenz ähnlich hoch korrelieren, wie Aufgaben zu verschiedenen Dimensionen von Intelligenz (z. B. räumliches Vorstellungsvermögen, verbale Intelligenz) (Kersting, 1999, S. 120).

Mit dem Ziel die Lernfähigkeit als Bestandteil des Intelligenzkonstruktes in entsprechenden Tests zu erfassen, wurden sogenannte Lerntests entwickelt, die im Gegensatz zu traditionellen Intelligenztests nicht nur Kenntnisse über den aktuellen Intelligenzstatus liefern, sondern darüber hinaus auch Informationen über die Reaktionen der Probanden auf die im Test eingebauten Lernanregungen (Beckmann, 1994, S. 1). Für die Konstruktvalidierung der Lerntest, die als Alternative zu herkömmlichen Intelligenztests gesehen werden (Guthke & Beckmann, 2004, S. 137; Beckmann & Guthke, 1995, S. 186), werden komplexe dynamische Problemlöseszenarien herangezogen. Beckmann (1994) prüfte, wie gut die Lerntests im Gegensatz zu herkömmlichen Intelligenztest den Lernerfolg bei der Bearbeitung eines dynamischen Szenarios (Simulation eines Kirschbaumes und einer Maschine auf Basis linearer Gleichungssysteme) vorhersagen (Beckmann, 1994, S. 63f.). Dabei erwiesen sich die Ergebnisse aus den Lerntests als signifikant bessere Prädiktoren sowohl für die Exploration als auch für die Steuerung des dynamischen Problemszenarios als die Statusintelligenzleistungen (Beckmann, 1994, S. 133; Guthke & Beckmann, 2004, S. 155).

Kröner (2001), der die Auffassung von Süß (1996) teilt, nach der sich mit computersimulierten Problemlöseszenarien lediglich Intelligenz und simulationsspezifisches Wissen diagnostizieren lassen, entwickelt das komplexe, dynamische Szenario *MultiFlux*. Zur konvergenten Validierung korreliert er die Ergebnisse aus der Bearbeitung des Systems *MultiFlux* mit denen aus verschiedenen Intelligenztests. Die Erwartung, dass der Gesamtzusammenhang zwischen *MultiFlux*-Score und der Intelligenzdimension *Verarbeitungskapazität* aus dem BIS-Test von Jäger (1997) ähnlich hoch ist, wie der

Zusammenhang zwischen anderen Intelligenztests untereinander, kann mit $r = .65$ als erfüllt gelten (Kröner, 2001, S. 105).

Mit der ursprünglich gleichen Intention wie Kröner, nämlich ein innovatives Instrument zur Intelligenzdiagnostik zu entwickeln und zu evaluieren, wurde *The Genetics Lab* erstellt (Hazotte, Mayer, Djaghloul et al., 2011, S. 296). Hierbei handelt es sich um ein auf Basis linearer Strukturgleichungen simuliertes Problemlöseszenario, in dem Probanden den Einfluss der Gene von fiktiven Kreaturen auf deren körperlichen Eigenschaften untersuchen müssen. Um einen differenzierteren Eindruck von der Relation zwischen der Fähigkeit zum dynamischen Problemlösen und Intelligenz zu bekommen, prüfen Sonnleitner, Keller, Martin et al. (2013) verschiedene Strukturmodelle, in denen Problemlösen und Intelligenz jeweils einmal als Facettenmodell und je einmal als hierarchisches Modell konzeptualisiert werden. Basierend auf den Daten ihrer Studie kommen die Autoren zu dem Schluss, dass dynamisches Problemlösen keinen zusätzlichen Wert aufgrund seiner inkrementellen Validität im Vergleich zu traditionellen Intelligenztests aufweisen kann (ebd., S. 303). Sie räumen auch ein, dass es überzeugende Zeichen dafür gibt, dass Problemlösefähigkeit ein Konstrukt ist, dass zwar stark mit Intelligenz zusammenhängt, dennoch von Intelligenz unabhängig ist (Sonnleitner, Keller, Martin et al., 2013, S. 303).

An diesem Punkt setzt eine andere Forschergruppe, die komplexe Systeme als Problemstellungen in Untersuchungen einsetzen, an. Ihrer gemeinsamen Auffassung nach ist Problemlösefähigkeit ein von Intelligenz beeinflusstes aber eigenständiges Konstrukt (vgl. Greiff, 2012, S. 232; Wüstenberg, Greiff & Funke, 2012, S. 2; Greiff & Fischer, 2013a, S. 31). Um sich von der traditionellen Problemlöseforschung, aber auch von der Intelligenzforschung abzusetzen, argumentieren sie, dass die Problemstellungen Anforderungen an die Problemlöser stellen, die bislang in der Intelligenzforschung nicht betrachtet werden (Kluwe, Schilde, Fischer et al., 1991, S. 292f.). Die wesentlichsten Unterschiede zwischen Intelligenztest- und komplexen Problemlöseaufgaben bestehen darin, dass nicht alle benötigten Informationen in der Ausgangssituation gegeben sind und in der Folge diese zunächst interaktiv erworben werden müssen, wofür das vom System gegebene Feedback genutzt werden muss (Wüstenberg, Greiff & Funke, 2012, S. 2). Damit sprechen die Autoren die bereits von Kersting (1999, S. 119) vorgebrachte Kritik an, nach der diese Anforderungen zwar sehr wohl theoretisch im Intelligenzkonstrukt vorgesehen sind, die Operationalisierung in Intelligenztestaufgaben jedoch ausgeblieben ist.

Abgesehen von den substantiellen theoretischen und empirischen Überlappungen, die es zwischen fluider Intelligenz und allgemeiner Problemlösefähigkeit gibt, sollte eine darüber hinaus gehende Erklärungskraft im Sinne inkrementeller Validität des Problemlösekonstruktes auf externe Kriterien wie den Erfolg in Schule, Hochschule oder im Beruf auch theoretisch begründet sein. Hierzu mutmaßt Greiff (2012, S. 41), dass Problemlösen noch „Aspekte wie allgemeines prozedurales Wissen oder Erfahrungswerte beinhaltet“. In einer Untersuchung zum fachspezifischen Problemlösen bei Elektronikern und Kfz-Mechatronikern zeigte sich, dass Intelligenz nur dann einen Einfluss auf das Lösen der gegebenen technischen Probleme ausübte, wenn das Fachwissen nicht mit einbezogen wurde (Abele, Greiff, Gschwendtner et al., 2012, S. 385). Die Enge des Zusammenhangs zwischen Intelligenz und Problemlöseleistung scheint vermutlich mit dem (problemspezifischen)

Wissen zusammenzuhängen, das bereits vorhanden oder in dem Problemlöseprozess erworben werden muss.

Der semantische Kontext, in dem Problemstellungen eingebettet sind, löst Vorstellungen, Erwartungen und Hypothesen bei der problemlösenden Person aus und aktiviert bereits vorhandenes Wissen, wodurch ein Einfluss auf die weitere Vorgehensweise bei der Problembearbeitung ausgeübt wird (Hesse, 1982, S. 64). Die auf das Wissen zurückzuführende Varianz in der Problemlöseleistung kann dabei nicht durch den Intelligenztest erfasst werden, so dass der Zusammenhang zwischen Intelligenz und Problemlösefähigkeit geringer ausfällt (Kersting, 1999, S. 122). Erklärt werden kann das durch die Erleichterung des Wissenserwerbs, wenn auf bereits vorhandenes Wissen zurückgegriffen werden kann (Renkl, 1996, S. 178f.). Durch ein umfangreiches und für die Problemlösung nützliches Wissen können Intelligenzdefizite ausgeglichen werden, womit Intelligenz als Prädiktor der Problemlöseleistung an Bedeutung verliert (Wirth, 2004, S. 39). Ist jedoch überhaupt kein Vorwissen vorhanden, so dürfte der Zusammenhang zwischen Intelligenz und Problemlöseleistungen ebenfalls gering ausfallen, da kein Wissen verfügbar ist, an dem neues Wissen im Rahmen der Informationsverarbeitung anknüpfen kann (ebd.). Bei mittlerem Vorwissen kann dieses für den weiteren Lernprozess genutzt werden, so dass hier der Zusammenhang zwischen Intelligenz und Problemlöseleistung maximal wird (Elshout, 1987; Raaheim, 1988). Dieser als *Elshout-Raaheim-Hypothese* bezeichnete umgekehrt U-förmige Zusammenhang zwischen Intelligenz und Problemlösen konnte von Leutner (2002) bestätigt werden.

Neben dem problemspezifischen Vorwissen wird auch die Problemschwierigkeit als Moderatorvariable für die Stärke des Zusammenhangs zwischen Intelligenz und Problemlöseleistung diskutiert (Beckmann, 1994, S. 57). Je intransparenter das Ziel (Strohschneider, 1991a) sowie die in einer Problemsituation wirkenden Variablen (Hussy, 1984b) sind, desto geringer ist der Zusammenhang zur Intelligenz. Mit dem Wissen über die beeinflussenden Merkmale und den methodischen Schwächen einiger Studien erscheinen die empirischen Befunde über den Zusammenhang zwischen Intelligenz und Problemlösen nicht mehr so heterogen (Kersting, 1999, S. 128).

In konkreten Problemsituationen, in denen der jeweilige Kontext eine große Rolle spielt, lässt sich die Fähigkeit zum Problemlösen eben nicht nur auf die kognitiven Grundfähigkeiten reduzieren. Während es sich bei Intelligenztestleistungen um die Ergebnisse weitgehend dekontextualisierter Denkprozesse handelt (Wilhelm & Nickolaus, 2013), wird Problemlösen neben Intelligenz und Wissen auch von Werten, Motivation und Emotionen beeinflusst (Strohschneider, 1991b, S. 355). In diesem Zusammenhang erscheint es für diese Arbeit sinnvoll, die Fähigkeit im problemlösenden Umgang mit technischen Geräten als Problemlösekompetenz zu bezeichnen, da angenommen werden kann, dass der (technische) Kontext einen erheblichen Einfluss auf das Problemlöseverhalten haben wird. Für die Stärke des Zusammenhangs zur Intelligenz wird angenommen, dass diese geringer ausfällt, als für stark dekontextualisierte Problemlösesituationen.

5.1.2 Allgemeine Problemlösefähigkeit

Wie in Abschnitt 2.2 *Klassifikation von Problemen* bereits erwähnt, lassen sich Probleme danach unterscheiden, wie spezifisch das Wissen, das zu ihrer Lösung notwendig ist, ist. Das Lösen von Problemen, für die bereits vorhandenes Wissen keine Rolle spielt, wird in der Literatur als fachübergreifendes, allgemeines, komplexes oder auch dynamisches Problemlösen bezeichnet. Die Begriffe Komplexität und Dynamik werden in dieser Arbeit als Merkmale einer Problemsituation verwendet, die unabhängig von dem Kontext sind, in dem die Probleme entstehen beziehungsweise präsentiert werden. Die Bezeichnung allgemeines bzw. fachübergreifendes Problemlösen erscheint geeigneter, da hier die besonderen Annahmen zur Geltung kommen, die mit dem Begriff verbunden sind. Der Begriff *allgemein* bedeutet „allen gemeinsam, überall verbreitet“ und wird synonym für den Begriff *generell* verwendet „für die meisten oder alle Fälle derselben Art geltend, zutreffend“ (Duden - Das Bedeutungswörterbuch, 4. Aufl. Mannheim 2010 [CD-ROM]). Eine allgemeine Problemlösefähigkeit beschreibt demnach eine Fähigkeit, die zur Lösung aller bzw. einer großen Anzahl von Problemen eingesetzt werden kann. Das setzt voraus, dass es generalisierbare Prozesse des Problemlösens gibt, die unabhängig von dem Problem und seinem Kontext sind (vgl. Arbinger, 2015, S. 14).

Dieses Verständnis findet sich auch in der bei PISA (OECD, 2003) verwendeten Definition von Problemlösekompetenz als fachübergreifende Kompetenz wieder. Ähnlich wie im Konzept der Schlüsselqualifikationen (Mertens, 1970) werden fachübergreifende Kompetenzen als *übergeordnete Bildungselemente* oder *situationsübergreifend transferierbare Qualifikationen* aufgefasst (Leutner, Funke, Klieme et al., 2005, S. 12). Dementsprechend wird Problemlösekompetenz in PISA 2003 definiert als die „Fähigkeit einer Person kognitive Prozesse zu nutzen, um sich mit solchen realen, fächerübergreifenden Problemstellungen auseinanderzusetzen und sie zu lösen, bei denen der Lösungsweg nicht unmittelbar erkennbar ist und die zur Lösung nutzbaren Wissensbereiche nicht einem einzelnen Fachgebiet der Mathematik, der Naturwissenschaften oder des Lesens entstammen“ (OECD, 2003, S. 156 deutsche Übersetzung von Leutner, Klieme, Meyer et al., 2004, S. 148). Dynamisches Problemlösen, als eine zentrale Facette allgemeinen bzw. fachübergreifenden Problemlösens (Abele, Greiff, Gschwendtner et al., 2012, S. 368), wird von Buchner beschrieben als die erfolgreiche Interaktion mit einer Problemsituation, die dynamisch ist (Systemänderungen erfolgen aufgrund von Systemeingriffen oder als Funktion der Zeit) und in der nicht alle Informationen zur Systemsteuerung gegeben sind, sondern nur durch eine erfolgreiche Exploration und Integration der Informationen im Problemlöseprozess gewonnen werden können (vgl. Frensch & Funke, 1995, S. 14).

In den Aufgaben zum allgemeinen Problemlösen geht es den Autoren im Gegensatz zu domänenspezifischen Problemlöseaufgaben nicht um die Anwendung von spezifischen, fachbezogenen Wissen und Strategien, sondern um Denk- und Problemlösestrategien allgemeinerer Art, wie induktives und deduktives Denken, kombinatorisches Denken, Denken in Analogien, Kontrolle und Reflexion des eigenen Denkprozesses etc. (Leutner, Klieme, Meyer et al., 2004, S. 148). Für die Erfassung solcher allgemeiner Strategien erscheinen vorwissensneutrale Problemstellungen brauchbarer (Greiff & Fischer, 2013a, S. 28).

Bereits in der jüngeren Problemlöseforschung finden sich Ansätze solche universell einsetzbaren Strategien ausfindig zu machen und damit jegliche Art des Problemlösens beschreiben bzw. erklären zu können (vgl. Funke, 2006a, S. 518). In dem *General Problem Solver*-Ansatz (GPS) von Newell & Simon (1972) führt das Aneignen von allgemeinen Problemlösemethoden zum Entstehen von Expertise, sodass Experten domänenübergreifende Problemlösefähigkeiten zugesprochen werden (Gruber, 1999, S. 48f.). Diese generellen Problemlösemechanismen lassen sich auf verschiedene Problemtypen anwenden, indem sich die problemlösende Person mittels allgemeiner Suchstrategien durch einen Problemraum bewegt. Die von den Autoren zur Theorie des *Human Problem Solving* ausgearbeiteten Vorstellungen ließen sich jedoch nur auf wohldefinierte, wenig komplexe Probleme übertragen (ebd.) und erwiesen sich als zu einfach, um domänenspezifisches Problemlösen verstehen und beschreiben zu können (vgl. OECD PISA Deutschland, 2003, S. 7; Kluwe, Schilde, Fischer et al., 1991, S. 306f.).

Um die weiterhin offene Frage zu beantworten, ob ein eigenständiges Konstrukt allgemeiner Problemlösefähigkeit postuliert werden darf, werden in einigen Studien Versuche unternommen eine gemeinsame Varianz aus den Leistungen bei der Bearbeitung verschiedener Problemlösetests zu extrahieren. Die Arbeiten von Süß, Kersting & Oberauer (1991; 1993) zeigten zunächst, dass ein großer Teil (bis zu 35 %) der Varianz in der Problemlöseleistung durch die Prädiktoren Intelligenz und bereichsspezifisches Wissen aufgeklärt werden kann, wobei in diesen Arbeiten kein Vergleich zwischen den Problemszenarien unterschiedlicher Autoren vorgenommen wurde. Wenn es neben den bestehenden Konstrukten Intelligenz und Wissen ein weiteres Fähigkeitskonstrukt Problemlösefähigkeit gibt, sollte dieses über die Zeit relativ stabil sein, so die Vermutung von Süß (1996, S. 194). In seiner Studie überprüfte er diese Hypothese, indem er aus der Varianz in den Steuerungsleistungen zwischen einer Erstuntersuchung und einer Wiederholungsuntersuchung die Varianz der Intelligenz und des Vorwissen auspartialisierte. Die Interkorrelation der Steuerungsleistungen der zwei Messzeitpunkte sank dabei nach Auspartialisierung der Varianz der beiden Prädiktoren von $r = .46$ auf $r = .13$ und war nicht signifikant, so dass Süß nicht für die Annahme einer eigenständigen Problemlösefähigkeit plädiert (ebd.). In einer Studie von Wittmann, Süß & Oberauer (1996) sowie in einer Untersuchung von Kersting (1999), in denen Probanden unterschiedliche komplexe Problemszenarien bearbeiteten, zeigten sich ähnliche Befunde. Nach statistischer Kontrolle der Prädiktoren Intelligenz und Wissen konnten keine statistisch bedeutsamen Zusammenhänge zwischen den einzelnen Problemlöseleistungen nachgewiesen werden (ebd., S. 21).

Ausgehend von der Beobachtung, dass eine allgemeine Problemlösefähigkeit über Intelligenz hinausgehende inkrementelle Validität auf die Vorhersage von Noten in naturwissenschaftlichen Fächern aufweist (Greiff & Fischer, 2013b) und dass die Leistungen in unterschiedlichen Problemlöseszenarien miteinander korrelieren (Danner, Hagemann, Schankin et al., 2011), nehmen Greiff, Fischer, Wüstenberg et al. (2013) an, dass die Validität allgemeiner Problemlösefähigkeit über verschiedene Messinstrumente hinaus generalisierbar ist. Für drei Messinstrumente, von denen zwei auf Basis linearer Strukturgleichungssysteme und eines auf Basis finiter Automaten konstruiert sind, überprüfen sie konvergente Validität, indem sie die Instrumente in einem latenten Multitrait-Multimethod-Ansatz (MTMM)

kombinieren. Hierbei wird allgemeines Problemlösen auf latenter Ebene anhand der zwei Dimensionen *Wissenserwerb* und *Steuerungsleistung* modelliert, von denen angenommen wird, dass sie das Verhalten in den jeweiligen Phasen der Problemlöseszenarien (unabhängig von den drei Testinstrumenten aus denen die Items stammen) beeinflussen. Die von den Autoren berichteten guten Modellfitwerte werden dahingehend interpretiert, dass die Messinstrumente für allgemeines Problemlösen konvergieren (Greiff, Fischer, Wüstenberg et al., 2013, S. 592). Das von den Autoren aufgestellte Modell wird anschließend in Relation zu den Variablen Intelligenz und Schulerfolg (in naturwissenschaftlichen Fächern) gesetzt, um Informationen zur diskriminanten Validität zu erhalten. Unabhängig von dem Problemlöseinstrument konnten im Durchschnitt 25 % der Varianz durch Intelligenz aufgeklärt werden. Zusätzlich zur Intelligenz konnten mit der Leistung in der Wissenserwerbsphase zwischen 3 und 5 % der Varianz in den Noten der Naturwissenschaftlichen Fächer erklärt werden, Problemlösen hat demnach inkrementelle Validität gegenüber Intelligenz.

In derselben Studie wurde außerdem ein weiteres allgemeines Problemszenario von den Probanden bearbeitet (Greiff, Stadler, Sonnleitner et al., 2015). Hierbei handelt es sich um eine simulierte *Schneiderwerkstatt (Tailorshop)*, eine Mikrowelt, die bereits in vielen früheren Untersuchungen (bspw. von Funke, 1983b und Putz-Osterloh & Lür, 1981) eingesetzt wurde und weitaus komplexer ist, als die Szenarien der Instrumente *MicroDYN*, *Genetics Lab* und *MicroFIN*. *Tailorshop* besteht aus 24 Variablen, von denen 21 von den Probanden eingesehen und 12 von ihnen manipuliert werden können (vgl. Danner, Hagemann, Holt et al., 2011, S. 225). Aufgrund ihrer größeren Ähnlichkeit korrelieren die Leistungen aus den zuvor berichteten Problemlösetestinstrumenten höher untereinander als zwischen ihnen und der Mikrowelt *Tailorshop*. Die Vermutung, dass der höhere Zusammenhang zwischen den drei Messinstrumenten auf Basis formaler Strukturen durch einen größeren Intelligenzanteil an der Leistung zustande kommt, konnte nicht bestätigt werden. Nach statistischer Kontrolle der Intelligenz verringerten sich die Korrelationen zwischen den Messinstrumenten zwar, sie blieben dennoch statistisch bedeutsam und auch die gefundene Differenz in der Korrelation zwischen den Testinstrumenten blieb bestehen (Greiff, Stadler, Sonnleitner et al., 2015, S. 110). Die Autoren schließen aus ihren Befunden, dass es sich bei den Messinstrumenten auf Basis formaler Strukturen um validere Instrumente handelt, als bei den klassischen Mikrowelten (ebd.).

Wenn, wie die Befunde der Forschergruppe um Greiff annehmen lassen, sich die durch Intelligenz und Vorwissen nicht erklärbare Varianz in den Problemlöseleistungen aus unterschiedlichen Problemszenarien auf etwas Gemeinsames zurückführen lassen, so stellt sich die Frage, wie diese Gemeinsamkeit theoretisch beschrieben werden kann. Zu dieser Frage nehmen Greiff & Fischer (2013a) indirekt Stellung, indem sie schreiben, dass Problemlösen neben entsprechendem Wissen auch „koordinierte Prozesse des Repräsentierens, Suchens, Planens und Entscheidens“ (Greiff & Fischer, 2013a, S. 27) erfordert. Ähnlich formulieren es Strohschneider & Schaub (1991), sie nennen diese Prozesse Heuristiken und machen die unterschiedliche heuristische Kompetenz ihrer Probanden für den fehlenden Zusammenhang zwischen Intelligenz und dem Problemlöseverhalten verantwortlich. Unter heuristischer Kompetenz verstehen sie in Anlehnung an Stäudel (1987) die über eine „Situation hinweg generalisierte Einschätzung der eigenen Fähigkeit, neuartige Situationen

bewältigen zu können“ (Stäudel, 1987, S. 54). Bei Heuristiken, die auch als Strategien bezeichnet werden (Buerschaper, 2000, S. 158), handelt es sich um unvollständige Prinzipien oder Regeln, deren Anwendungen nicht mit Sicherheit zur Problemlösung führen, aber durch ihre grobe Festlegung von Ordnungen, von Zwischenzielen oder Operatoren für den, der sie anwendet, einen Vorteil bieten (Frackmann & Tärre, 2009, S. 44). Den Kern des heuristischen Wissens stellen Wenn-Dann-Regeln dar (von der Weth, 1994, S. 73), die, wenn es sich um allgemeine, unspezifische Regeln handelt, Leerstellen enthalten. Heuristiken sind also das, was bei Anderson (1983) als allgemeine Problemlöseprozeduren bzw. Produktionen bezeichnet wird¹. Im Sinne des von Anderson beschriebenen Prozesses der Wissensprozeduralisierung gehen Strohschneider & Schaub (1991) davon aus, dass sich die Erfahrungen im Umgang mit verschiedenartigen Problemen im Laufe der Zeit „zu hochgeneralisiertem Problemwissen verdichten, dessen Effekte auch beim Bearbeiten neuartiger Anforderungen nachweisbar sind“ (Strohschneider & Schaub, 1991, S. 325).

Aus fachdidaktischer Perspektive stellt sich in diesem Zusammenhang die Frage, ob es überhaupt solche allgemeinen Strategien gibt und wenn die Antwort positiv ausfällt, wie diese vermittelt werden können, so dass ihr Einsatz bei unterschiedlichen (z. B. fach- bzw. domänenspezifischen) Problemen einen Nutzen bringen kann (vgl. Frackmann & Tärre, 2009, S. 33). Ausubel, Novak & Hanesian (1981, S. 648) merken an, dass es kaum möglich sei, Problemlösestrategien auf unterschiedliche Arten von Problemen zu transferieren und die Möglichkeit Problemlösen einzuüben sei demnach begrenzt. Aber auch, wenn es keine Strategien gibt, die universell auf verschiedene Problemsituationen anwendbar sind (Kluwe, Schilde, Fischer et al., 1991, S. 307), so vermuten die Autoren Frackmann & Tärre (2009, S. 34) die Existenz etwaiger Prozesse, die über die Bereichsspezifität vieler Probleme hinausgehen und verweisen an anderer Stelle darauf, dass die Trainierbarkeit dieser Strategien Inhalt beispielsweise beruflicher Bildung sein sollten (Frackmann & Tärre, 2009, S. 33). Während Fleischer, Wirth, Rumann et al. (2010, S. 239) der allgemeinen Problemlösefähigkeit für den beruflichen Erfolg sogar per se eine zentrale Rolle beimessen, wird die Annahme, nach der allgemeine Problemlösefähigkeit die Kompetenz im Lösen von fachspezifischen (technischen) Problemen beeinflusst, von Abele, Greiff, Gschwendtner et al. (2012) genauer untersucht. In ihrer Untersuchung stellen die Forscher sich zwei Fragen: Zum einen ob allgemeine, dynamische Problemlösefähigkeit bei der Vorhersage technischer Problemlöseleistungen inkrementelle Validität gegenüber Intelligenz aufweist und zum anderen inwieweit sich der Zusammenhang zwischen allgemeinem und fachspezifischem Problemlösen bei Integration des für das fachspezifische Problemlösen benötigten Fachwissens ändert (Abele, Greiff, Gschwendtner et al., 2012, S. 370). Die Beantwortung ihrer Hypothese basiert auf Daten einer Studie, in der Kfz-Mechatroniker und Elektroniker sowohl Items zum allgemeinen Problemlösen (MicroDYN) als auch für ihre Domäne spezifische Problemstellungen bearbeiteten. Lediglich für die Stichprobe der Elektroniker konnte allgemeine Problemlösefähigkeit einen Anteil von 7 % zusätzliche Varianz der technischen Problemlöseleistung über Intelligenz hinaus aufklären (ebd., S. 383). Wurde außerdem das Fachwissen in das Vorhersagemodell integriert, wies allgemeine Problemlösefähigkeit ebenso

¹ Kapitel 3.4.2.1 und 3.4.2.2 gehen auf diese Produktionen als Teil des kognitiven Systems und auf ihre Bedeutung für prozedurales Lernen ein.

wie fluide Intelligenz keine signifikant inkrementelle Validität auf. Fachwissen kann demnach als der stärkste Prädiktor für das Lösen fachspezifischer Probleme angesehen werden.

In einer anderen Studie von Molnár, Greiff & Csapó (2013), die den Einfluss allgemeinen komplexen Problemlösens auf domänenspezifisches Problemlösen untersucht, stellt die Domäne das Fach Mathematik dar. Ziel der Studie ist es den Zusammenhang zwischen Intelligenz, allgemeinem und fachspezifischem Problemlösen und seine Veränderung über die Zeit darzustellen. Hierzu bearbeiten die an der Studie teilnehmenden, 9-17 Jahre alten ungarischen Schülerinnen und Schüler der Klassen 5, 7 und 9 Aufgaben eines Intelligenz-, eines allgemeinen Problemlöse- und eines fachspezifischen Problemlösetests. Die Höhe der Zusammenhänge zwischen den drei Konstrukten über alle Klassen hinweg liegt zwischen $r = .35$ und $.44$ und ist auch nach statistischer Kontrolle je einer der Variablen relativ stabil. Schaut man sich die Zusammenhänge differenziert nach Klassen an, fällt auf, dass zum einen der Zusammenhang zwischen Intelligenz und allgemeinem Problemlösen in allen Klassen am stärksten ist und, dass der Zusammenhang zwischen allgemeinem und domänenspezifischem Problemlösen im Lauf der Zeit höher wird. Die wachsende Korrelation zwischen domänenspezifischem und allgemeinem Problemlösen erklären die Autoren durch die Vermutung, dass sich die Strategien, die in beiden Problemsituationen zum Einsatz kommen, im Laufe der Zeit immer ähnlicher werden. Domänenspezifisches Problemlösen basiert auf der Anwendung von Wissen, wohingegen allgemeines Problemlösen die Voraussetzung für das Gewinnen und Anwenden neuen Wissens ist, das bereits im domänenspezifischen Problemlösen repräsentiert ist (Molnár, Greiff & Csapó, 2013, S. 43).

5.1.3 Need for Cognition und Selbstwirksamkeitserwartungen

Neben den bereits erwähnten kognitiven Faktoren wie Wissen und Fähigkeiten haben affektive und insbesondere motivationale Konstrukte ebenfalls einen Einfluss auf das Lernen und Verhalten einer Person in einer Situation (Bandura, 1977). Ausubel, Novak & Hanesian (1981), die erfolgreiche und erfolglose Problemlöser miteinander vergleichen, stellen fest, dass die Einstellungen erfolgreicher Problemlöser gegenüber dem Nachdenken positiver sind und dass erfolgreiche Problemlöser gegenüber erfolglosen Problemlösern mehr Selbstvertrauen in die eigene Fähigkeit, Probleme zu lösen zeigen und sich durch Komplexität nicht so leicht entmutigen lassen (Ausubel, Novak & Hanesian, 1981, S. 664). Das Verhalten einer Person in einer Problemsituation ergibt sich demnach aus dem Wechselspiel zwischen den Anforderungen der Situation, den kognitiven Fähigkeiten sowie der aktualisierten Motivation einer Person (vgl. Hesse, Spies & Lüer, 1983, S. 404). Letztere setzt sich zusammen aus überdauernden Motiven und den situativen Anregungen (ebd.). Überdauernde individuelle Motivdispositionen werden auch als implizite Motive bezeichnet (Heckhausen & Heckhausen, 2006, S. 4), die mit Bedürfnissen, Wünschen und Erwartungen umschrieben werden können (vgl. Seel, 2000, S. 78). Diese lassen sich nach Maslow (1970) in eine hierarchische Ordnung bringen (Abbildung 5.1).



Abbildung 5.1: Die Hierarchie der Bedürfnisse nach Maslow (1970) (Seel, 2000, S. 79)

Während Defizite in den primären Bedürfnissen der untersten beiden Ebenen wie Hunger, Durst, Schlaf etc. von fast allen Personen als Defizitmotive wahrgenommen werden und dazu führen, dass eine Person hoch motiviert ist, diese defizitären Bedürfnisse zu befriedigen, unterscheiden sich die *höheren* sekundären Bedürfnisse verschiedener Personen hinsichtlich ihrer Hierarchien voneinander (Seel, 2000, S. 80). Während eine sehr leistungsorientierte Person ein stärkeres Bedürfnis nach Wertschätzung und Anerkennung durch andere hat, dafür aber ein geringeres Bedürfnis zu lieben verspürt, können die genannten Bedürfnisse für eine andere Person in umgekehrter Richtung Bedeutung haben. Die individuelle Ausprägung eines konkreten Motivs wird aber erst dann im Verhalten einer Person sichtbar, wenn sich diese in einer motivpassenden Situation befindet (Rheinberg, 2006, S. 70). In einer Testsituation werden in der Regel keine Bedürfnisse nach Sicherheit oder Liebe wirksam, sondern eher Bedürfnisse nach Selbstwertschätzung oder kognitive Bedürfnisse. Bei Auftreten der entsprechenden Situation ist eine habituelle Motivation (ausgelöst durch überdauernde Motive) durch ihr wiederholtes Auftreten gekennzeichnet (Schiefele & Schaffner, 2015, S. 154).

Von Interesse sind in dieser Arbeit vor allem diejenigen Situationen, in denen Personen unbekanntem technischen Geräten gegenüberstehen und die zunächst vorhandenen Bedienungsprobleme durch aktiven explorierenden Wissenserwerb überbrücken müssen. Die für das Lösen von Problemen notwendige Motivation resultiert dabei vor allem aus Bedürfnissen nach Selbstvertrauen, dem Gefühl eigener Kompetenz sowie dem Bedürfnis, das eigene Potenzial auszuschöpfen, der Neugier und des Verstehens. Da sich die Personen, deren Umgang mit technischen Geräten im Rahmen dieser Arbeit untersucht wird, zudem in einer Testsituation befinden, interessieren auch diejenigen Bedürfnisse und Motive, die das Verhalten in solchen Testsituationen beeinflussen.

Motiviertes Verhalten bei der Bearbeitung eines Tests stellt ein Spezialfall der Leistungsmotivation dar (Asseburg, 2011, S. 22) und kann durch unterschiedliche Motive hervorgebracht werden. Motive einer Person, sich mit Gütemaßstäben (die durch Tests repräsentiert werden) auseinanderzusetzen, können der Spaß an der Herausforderung, die Bestätigung über die eigenen Fähigkeiten oder die Abwendung von drohendem Misserfolg sein (Langens, 2009, S. 217f.). Wenn die angestrebten Absichten innerhalb der Person liegen, wird die aus ihnen in einer Situation hervorgebrachte Motivation als intrinsische Motivation bezeichnet (Schiefele & Schaffner, 2015, S. 155). Intrinsisch motiviertes Verhalten erfordert keine positiven (Versprechungen) oder negativen Anreize (Drohungen) und beinhaltet vor allem Neugier, Exploration, Spontanität und Interesse (Deci & Ryan, 1993, S. 225). Im Gegensatz dazu tritt extrinsisch motiviertes Verhalten nicht spontan auf, sondern erst nach Aufforderung, in der Erwartung positiver oder der Vermeidung negativer Konsequenzen (vgl. ebd.).

Die Basis des Konstruktes der intrinsischen Motivation stellt das Motiv der *Neugier* dar (Seel, 2000, S. 81), das sich im unterschiedlichen Explorationsverhalten von Menschen zeigt (vgl. Schiefele & Schaffner, 2015, S. 157). Mit Neugier wird das Bestreben, durch aktives oder passives Explorieren neue Erkenntnisse zu erlangen, bezeichnet (vgl. Lengning, 2009, S. 252). Wie alle Motive bedarf Neugier einer entsprechenden Situation um wirksam zu werden. Berlyne (1974) unterscheidet in Abhängigkeit der Intensität der gebotenen Anreize zwischen *diversivem* und *spezifischem* Neugieverhalten. Bietet eine Situation keine oder nur sehr wenige Anreize, wird sie von einer Person als unangenehm erlebt und sie empfindet Unruhe beziehungsweise Langeweile. Mit dem Ziel, die durch die reizarme Situation verursachte innerliche Spannung abzubauen, werden stimulierende Aktivitäten ausgeführt, die als *diversives* Neugieverhalten bezeichnet werden (vgl. Rothermund & Eder, 2011, S. 35). *Spezifisches* Neugieverhalten wird durch Situationen ausgelöst, in denen Reize neu, komplex, zweideutig sind und damit zu einer subjektiven Ungewissheit führen, die ähnlich wie fehlende Reize zu innerlichen Spannungen führen. Die Reduzierung der erlebten Dissonanz wird durch spezifisches Neugieverhalten erreicht (ebd.). Je stärker das Motiv der Neugier und damit das Bedürfnis nach Erkenntnisgewinn, desto eher erlebt eine Person kognitive Konflikte, die sie beseitigen möchte.

Indem sie die Vorstellungen von Maslow und anderen Theoretikern über die Existenz verschiedener Bedürfnisse aufgreifen, konzeptualisieren Cohen, Stotland & Wolfe (1955) ein Persönlichkeitsmerkmal, das sie *Need for Cognition* nennen. Darunter verstehen sie die Tendenz einer Person, kognitive Unstimmigkeiten und Doppeldeutigkeiten zu vermeiden sowie eigene Erfahrungen bedeutungsvoll zu organisieren, um ein Verstehen der Situation zu ermöglichen. Cacioppo & Petty (1982), die sich mit der Erfassung des individuellen *Need for Cognition* befassen, sehen in dem Konstrukt die Tendenz einer Person sich mit Denken zu beschäftigen und Spaß an dieser Tätigkeit zu haben (ebd., S. 116). Die definitorische Nähe zum Konstrukt der Neugier konnte in einer Studie von Olson, Camp & Fuller (1984) empirisch bestätigt werden. Die mittlere Korrelation (berechnet aus acht unterschiedlichen Skalen zur Erfassung von Neugier) zwischen Neugier und *Need for Cognition* beträgt in dieser Untersuchung $r = .57$. Es konnte lediglich ein Zusammenhang zu dem oben beschriebenen spezifischen Neugieverhalten gefunden werden, nicht zu diversiver Neugier.

Dieses Ergebnis stützt die Annahme, dass *Need for Cognition* ein Motiv ist, das sich insbesondere in der Motivation zur Bewältigung kognitiver Anforderungen und dem Lösen von Problemen zeigt (vgl. Cacioppo, Petty, Feinstein et al., 1996, S. 198), wobei der Fokus auf intrinsischer Motivation liegt (Beißert, Köhler, Rempel et al., 2014, S. 5). So sind Personen mit einem hohen *Need for Cognition* eher bereit sich kognitiv aufwändig mit einem Problem auseinanderzusetzen, als Personen mit einem geringeren *Need for Cognition* (Pechtl, 2009, S. 5). Pechtl fasst die Merkmale, die eine Person mit hohem *Need for Cognition* beschreiben, wie folgt zusammen:

- Kognitive Aktivitäten werden als angenehm, unterhaltsam und positiv empfunden,
- neue kognitive Herausforderungen werden gesucht und kognitive Überraschungen und unübliche Gedankengänge als positiv eingeschätzt,
- kognitive Investitionen werden gerne getätigt,
- es besteht der Wunsch, Sachverhalte und Probleme kognitiv zu durchdringen sowie Zusammenhänge und Ursache-Wirkungsbeziehungen selbst zu erkennen,
- es werden umfangreichere Informationsaktivitäten getätigt und Medien gewählt, die eine aktivere Informationsnutzung erfordern (Pechtl, 2009, S. 8).

Für die Sichtbarkeit des Einflusses, den *Need for Cognition* als repräsentatives Konstrukt für intrinsische Motivation auf den problemlösenden Umgang mit technischen Geräten haben kann, erscheint der letztgenannte Punkt in der obigen Merkmalsliste besonders relevant. Allgemein zeigt sich motiviertes Verhalten in einer (Problem-)Situation darin, was eine Person tut (Vollmeyer, 2009, S. 336), mit welcher Ausdauer diese Tätigkeit ausgeführt wird und in welcher Intensität (wie sehr sich eine Person bei einer Tätigkeit konzentriert bzw. anstrengt) (Schiefele & Schaffner, 2015, S. 154). Motivation in einer Testsituation zeigt sich beispielsweise in der Bereitschaft, Testaufgaben auch unter Anstrengung bestmöglich zu bearbeiten und diese Anstrengung über die gesamte Testdauer hinweg aufrechtzuerhalten (Asseburg, 2011, S. 22). In Situationen, die den Erwerb von Wissen, also die Suche und Verarbeitung von Informationen verlangen, wird Motivation in dem Umfang an Informationssuche sowie in einer deutlich ausdauernden Bearbeitung sichtbar (Hesse, Spies & Lüer, 1983, S. 420; Vollmeyer, 2009, S. 336).

Eine zurückhaltende Informationssuche (als Ergebnis mangelnder Motivation) lässt sich auch auf eine geringe Selbstwirksamkeitserwartung zurückführen (vgl. Naumann, Goldhammer, Rölke et al., 2014, S. 195). Unter Selbstwirksamkeitserwartungen wird seit Bandura (1977) die subjektive Überzeugung verstanden, aufgrund der eigenen Kompetenz Anforderungen in Situationen erfolgreich bewältigen zu können, um ein gewünschtes Ziel zu erreichen (Beierlein, Kemper, Kovaleva et al., 2013, S. 251; Jerusalem & Mittag, 1999, S. 224). Sie bezieht sich auf die Einschätzung der eigenen Fähigkeiten, Handlungen erfolgreich planen und ausführen zu können in dem Sinne, ob sich eine Person in der Lage sieht, das zur Erreichung eines Ergebnisses erforderliche Verhalten zu zeigen (Mielke, 1984, S. 64).

Die Erwartung an die Wirksamkeit des eigenen Verhaltens basiert größtenteils auf direkte eigene Erfahrungen (vgl. Gebauer, 2013, S. 60). Die Quelle einer geringen Selbstwirksamkeit kann zum einen in bisherigen Misserfolgen aufgrund fehlenden Wissens, Fertigkeiten oder Expertise liegen (Rothermund & Eder, 2011, S. 85). Zum anderen kann eine erfolgreiche

Handlungsausführung auch als ein zufälliges Ergebnis interpretiert werden, bei dem die eigene Fähigkeit keine Rolle gespielt hat (vgl. Mielke, 1984, S. 75). Erst wenn das Zustandekommen eines Erfolgserlebnisses auf die eigene Fähigkeit zurückgeführt wird, lässt sich die Überzeugung der Selbstwirksamkeit erhöhen (ebd.). Eine hohe Selbstwirksamkeit erhöht wiederum die Anstrengung und Ausdauer bei der Bewältigung neuer Aufgaben mit der Konsequenz einer größeren Erfolgswahrscheinlichkeit. Positive Erfolgserlebnisse begünstigen die weitere Stabilisierung der Selbstwirksamkeitserwartungen (Jerusalem & Mittag, 1999, S. 224).

Wirth (2004) betont, dass die Erwartungen an die Wirksamkeit eigenen Handelns vor allem von den spezifischen Anforderungen in einer Situation und deren wahrgenommenen Schwierigkeiten abhängig sind. Seiner Einschätzung zufolge sind Selbstwirksamkeitserwartungen aufgabenbezogen und demnach keine generellen Personeneigenschaften (Wirth, 2004, S. 51). Eine Person, deren Selbstwirksamkeitserwartung hinsichtlich ihrer fremdsprachlichen Fähigkeiten hoch ist, hat demnach nicht zwingend auch hohe Selbstwirksamkeitserwartungen hinsichtlich mathematischer Anforderungen. Andere Autoren (z. B. Rothermund & Eder, 2011, S. 86) ziehen außerdem in Betracht, dass sich im Laufe der Zeit auch generalisierte Selbstwirksamkeitserwartungen herausbilden können. Sie entstehen infolge verschiedener Erfahrungen und sind konzeptualisiert als situationsübergreifende und relativ stabile Erwartungen an das eigene Können (Rothermund & Eder, 2011, S. 86). Solch allgemeine Selbstwirksamkeitserwartungen steuern als Teil der Persönlichkeit das Handeln in neuen und unvertrauten Situationen (ebd.) (Beierlein, Kemper, Kovaleva et al., 2013, S. 254).

Da die Skala zur Erfassung des *Need for Cognition* nicht nur eine Differenzierung zwischen Personen hinsichtlich ihres Spaßes am Denken vornimmt, sondern außerdem bezüglich ihrer positiven Einschätzung der eigenen kognitiven Fähigkeiten (Bless, Fellhauer, Bohner et al., 1991, S. 4), kann von einem Zusammenhang zwischen *Need for Cognition* und der allgemeinen (auf die Bewältigung neuer Anforderungen bezogene) Selbstwirksamkeit ausgegangen werden. Beide Konstrukte stellen Motive beziehungsweise Bedürfnisse im Sinne von Maslow (siehe Abbildung 5.1) dar und führen in einer entsprechenden Situation zu motiviertem Verhalten. Über Motivation wiederum ist bekannt, dass sie zu systematischerem Vorgehen sowie zur Reduzierung der Anstrengung beim Lernen führt und damit den Erwerb und die Anwendung von Wissen beeinflusst (Rheinberg, Vollmeyer & Lehnik, 2000, S. 83; Vollmeyer & Rheinberg, 1998, S. 22; Schoppek & Putz-Osterloh, 2003, S. 170). In Bezug auf das Problemlösen beobachteten Ausubel, Novak & Hanesian (1981), dass sich Problemlöser mit mehr Selbstvertrauen in die eigene Fähigkeit nicht so schnell entmutigen lassen und somit erfolgreicher sind.

5.2 Kontextuelle Einflussfaktoren

5.2.1 Kontextbezug als Merkmal von Kompetenz

Die in bereits vorhandenen Problemlösetheorien verwendeten Kompetenzbegriffe sind oftmals domänenübergreifend definiert, so dass sich die Frage stellt, ob eine solche weit gefasste Problemlösekompetenz eine ausreichende Vorhersagekraft für das Problemlöseverhalten in konkreten technischen Situationen hat, oder ob Problemlösen immer

abhängig von dem Kontext ist, in dem es sich vollzieht. Vor dem Hintergrund dieser Arbeit erscheint es also zunächst notwendig zu klären, inwieweit sich die Kompetenz zum Lösen von Problemen im Umgang mit technischen Geräten von der Kompetenz zum allgemeinen Problemlösen unterscheidet. Hierzu ist zunächst herauszuarbeiten, in welchem Maß sich typische Probleme der beiden Kategorien voneinander abgrenzen. Ein wesentliches Unterscheidungsmerkmal ist hierbei der spezifische Kontext, in dem die Problemlöseprozesse stattfinden.

Kompetenzen, die sich in der Bewältigung von Anforderungen zeigen und sich „[...] sowohl auf Handlungsvollzüge als auch auf die ihnen zugrundeliegenden mentalen Prozesse und Kapazitäten [...]“ (Klieme & Hartig, 2007, S. 13) beziehen, lassen sich nur in Situationen, in denen diese Handlungsperformanz gezeigt werden kann, bestimmen. Eine sinnvolle Interpretation der so erfassten Daten ist nur unter Berücksichtigung der Situation, in denen sie entstanden sind, möglich (vgl. Kaufhold, 2006, S. 82). Denn die Ausführung von Handlungen oder das Verhalten von Personen ist nicht ausschließlich eine Funktion ihrer Eigenschaften oder Fähigkeiten, sondern wird auch durch situative Faktoren beeinflusst (vgl. Amelang & Schmidt-Atzert, 2006, S. 16). Damit Personen handeln, ist eine Interpretation der Situation sowie die Reflexion der eigenen Handlung erforderlich. Hier und bei der Wahrnehmung und Bedeutungszuschreibung einer Situation bezieht die handelnde Person eigene Erfahrungen und Handlungsgewohnheiten mit ein (Kaufhold, 2006, S. 81). Handeln wird also nicht nur durch Persönlichkeitseigenschaften und andere individuelle Dispositionen, sondern auch durch die (subjektive Wahrnehmung und Deutung einer) Situation bestimmt (vgl. Berg & Calderone, 1994, S. 113). Dementsprechend ist nach Kaufhold (2006, S. 118) der Situationsbezug als ein Merkmal von Kompetenz anzusehen, der wenn es um die Erfassung von Kompetenz geht, nachvollziehbar beschrieben werden muss. Als notwendige zentrale Elemente zur Beschreibung einer Situation nennt Kaufhold die in der Situation beteiligten Akteure, die Rahmenbedingungen sowie die Anforderungen an eine Person durch die zu beschreibende Situation (ebd., S. 120).

Bevor sich Kapitel 6 dem speziellen Problemkontext technischer Geräte widmet, werden eine Definition des Begriffes Kontext eingeführt und allgemein die Einflüsse des Kontextes auf Fähigkeiten, Motivation und Selbstwirksamkeit beschrieben. In Anbetracht einer mangelnden Übereinstimmung bezüglich des Begriffes Kontext und einer Vielzahl von Definitionen (vgl. Brézillon, 1999, S. 47) erscheint diese Aufgabe schwierig. Für die Verwendung des Kontextbegriffes im Zusammenhang mit Kompetenzkonstrukten eignet sich die Begriffsbestimmung von Hartig (2006): „Mit Kontext ist gemeint, auf welchen Bereich von Situationen und Anforderungen sich ein spezifisches Kompetenzkonstrukt bezieht“ (Hartig, 2006, S. 4). Für ihn hat der Kontext eine wesentliche Aufgabe für die Definition des Kompetenzkonstrukts, nämlich die Abgrenzung des spezifischen Kompetenzkonstrukts gegenüber bereits existierenden Konstrukten wie Intelligenz (ebd., S. 4), für die vorliegende Arbeit aber auch die Abgrenzung gegenüber sehr weit gefassten Kompetenzdefinitionen. Auf je mehr Situationen und Anforderungen sich das Kompetenzkonstrukt bezieht, desto weiter fasst man den Kontext und damit das Konstrukt und je schwieriger wird die Ableitung eines konkreten Messinstrumentes. Die in einem zu definierenden Konstrukt enthaltenen Situationen sollten sich ähnlich und die in ihnen zu bewältigenden Anforderungen vergleichbar sein (ebd., S. 5). Ein bezüglich der Kompetenz zum Lösen von Problemen

ebenfalls geläufiger Begriff ist der der *Domäne*. Während Abele (2016) den Begriff der *Domäne* vage mit *bestimmter Umweltausschnitt* umschreibt, definiert Gelman eine Wissensdomäne als eine Reihe zusammenhängender Prinzipien, den Regeln ihrer Anwendung sowie den Objekten, die sie als Domäne gemeinsam haben (Gelman, 1998, S. 558). Eine exakte Abgrenzung einzelner Domänen zueinander wird als schwierig, wenn nicht als unmöglich betrachtet, so dass Domänen in der Regel relativ weit gefasst werden (Renkl, 1996, S. 175).

Analog zu dem Begriff der Domäne versteht man unter Kontext in Bezug auf naturwissenschaftliches Lernen „eine Verflechtung von fachlichen Inhalten und Zugängen [...]“ (van Vorst, Dorschu, Fechner et al., 2015, S. 30), für deren Charakterisierung Gilbert (2006, S. 964) vier Attribute vorstellt:

1. Ein soziales, zeitliches und räumliches Bezugssystem (Setting), in dem mentale Begegnungen mit den „focal event“ stattfinden,
2. Ein Handlungsrahmen dieser Begegnung, der typischerweise ein zu lösendes Problem darstellt,
3. Die Fachsprache als „talk“, die zur Lösung des Problems benutzt werden muss,
4. Die Beziehung zu außer-situationalem Hintergrundwissen (Fachwissen), das gegeben sein muss, um im Setting angemessen agieren zu können (Kölbach, 2011, S. 21f.).

Betrachtet man die von Gilbert (2006, S. 964) formulierten vier Attribute, durch die ein Lernkontext charakterisiert werden kann, erscheint eine Trennung zwischen den Begriffen *Domäne* und *Kontext* schwierig. Da der Begriff Domäne laut Duden ein „Spezialgebiet, also ein Gebiet, auf dem sich jmd. besonders betätigt“ (Duden - Das Bedeutungswörterbuch, 4. Aufl. Mannheim) bezeichnet, wird von domänenspezifischem Problemlösen in dieser Arbeit nur an den Stellen gesprochen, an denen es sich um stark fachspezifische Problemstellungen handelt, für die das benötigte Wissen nur über eine länger andauernde Beschäftigung (z. B. durch eine Berufsausbildung) mit dem Gegenstand zu erwerben ist. Der problemlösende Umgang mit technischen Alltagsgeräten wird hingegen als *kontextspezifisches* Problemlösen bezeichnet.

Vor dem Hintergrund des bereits in Abschnitt 5.1.2 thematisierten möglichen Einflusses allgemeinen Problemlösens auf das Lösen domänenspezifischer Probleme, liegt der Fokus in diesem Kapitel auf die Bedeutung des Kontextes für den Prozess und die Leistung beim Problemlösen. Die in der Realität vorliegenden komplexen Probleme sind immer in einen Kontext eingebettet (Hesse, 1982, S. 62) und an dem gescheiterten Versuch diese mit generellen Heuristiken und Strategien beschreiben und erklären zu können, kamen bereits Newell & Simon (1972) zu dem Schluss, „dass die Domänenspezifität ein wesentliches Charakteristikum von Problemen darstellt [...]“ (Rollett, 2008, S. 14). Für diese Vorstellung sprechen auch die Ergebnisse der PISA-Studie 2003, in der für Schülerinnen und Schüler in Deutschland abweichende allgemeine und mathematikspezifische Problemlösefähigkeit festgestellt wurden. Die allgemeine Problemlösefähigkeit wird dabei post hoc als Potenzial betrachtet, das sich in dem guten Abschneiden des Tests zum allgemeinen Problemlösen zeigte, aber nicht genutzt werden konnte, um stark fachspezifische Probleme zu lösen, was sich in dem schlechteren Abschneiden im Mathematiktest zeigte (Fleischer, Wirth & Leutner, 2014, S. 217f.; Buchwald, 2015, S. 12). Andersherum weisen im Bereich des kontextuellen Lernens Aufgabenstellungen, die in einem für die Schülerinnen und Schüler alltagsnahen Kontext eingebettet sind, häufig höhere Lösungsquoten auf, als kontextfreie Aufgaben (Van den

Heuvel-Panhuizen, 2005, S. 7). Das sich hinsichtlich der Möglichkeit des Transfers von Fähigkeiten zwischen verschiedenen Kontexten zeigende heterogene Bild ist Anlass für die Autoren Barnett & Ceci (2002) eine Taxonomie des Transfers von Fähigkeiten vorzuschlagen. Der Erfolg, mit dem kognitive Operatoren, die in einem konkreten Kontext beherrscht werden, in einen anderen Kontext übertragen werden, hängt demnach davon ab, wie stark sich Ursprungs- und Zielkontext voneinander unterscheiden. Ein wesentliches Unterscheidungsmerkmal stellt neben zeitlichen, räumlichen und weiteren Aspekten, die Wissensdomäne dar (Barnett & Ceci, 2002, S. 621). Je ähnlicher sich die Kontexte bezüglich der zu verwendenden Wissensinhalte sind, desto besser kann ein Transfer erfolgen. So sollte die Kompetenz im Umgang mit neuen technischen Geräten, zwischen dem Kontext *Bedienung einer Waschmaschine* und *Bedienung einer Spülmaschine* relativ gut zu transferieren sein. Die Übertragung einer solchen Kompetenz auf ein technisches Gerät im industriellen Kontext, wie die *Steuerung einer 5-Achsen Fräsmaschine* dürfte aufgrund der geringen Überschneidung der erforderlichen Wissensinhalte demnach problematischer sein.

Selbst bei sich ähnelnden Kontexten, gelingt der Transfer häufig jedoch nicht, wie beispielsweise Studien bezüglich der Fehlerdiagnosestrategien von Probanden in technischen Systemen zeigen (Nickolaus, Abele, Gschwendtner et al., 2012, S. 245). Das kann daran liegen, dass jede kontextualisierte Problemsituation bei der problemlösenden Person eine Reihe von Vorstellungen, Erwartungen und Bewertungen auslöst, die die weitere Vorgehensweise der Problembearbeitung beeinflusst (Hesse, 1982, S. 65). Im Sinne der kognitiven Lerntheorie nach Anderson, werden also je nach Bekanntheit des entsprechenden Kontextes Chunks im kognitiven System aktiviert, die für Handlungen zur Verfügung stehen. Aus diesem Grund wird die gleiche Situation von verschiedenen Personen oftmals anders gesehen und gedeutet (Esser, 1996, S. 3). Diese Vorerfahrungen können für den Problemlöseprozess sowohl förderlich (Hesse, 1982, S. 86) aber auch hinderlich sein (Funke, 1983a). Neben der Aktivierung von Wissen legen unterschiedliche Kontexte außerdem unterschiedliche Lösungsstrategien nahe (Fleischer, Wirth & Leutner, 2014, S. 219) und haben einen Einfluss auf die Motivation einer Person, sich mit dem vorliegenden Problem auseinanderzusetzen.

Nach diesen zunächst allgemeinen theoretischen Überlegungen zum Einfluss des Kontextes, setzt sich der folgende Abschnitt mit dem konkreten Einflussbereich des Vorwissens auseinander. Hierbei interessiert vor allem inwieweit bereits erworbene Erfahrungen auf den probierenden Umgang mit technischen Geräten einwirken und ob sich dieser Einfluss auch in der Leistung konkreter Systemsteuerungen zeigt.

5.2.2 Erfahrungen

Der Erfolg des während des Problemlöseprozesses stattfindenden Wissenserwerbs ist davon abhängig, wie gut neue Inhalte verstanden, behalten und angewendet werden können. Möglich wird die Verarbeitung und Speicherung neuer Informationen allerdings erst durch eine Verknüpfung an bereits vorhandene Wissensstrukturen (Krause & Stark, 2006, S. 41), die hierfür zunächst aktiviert werden müssen. Lernen stellt demnach einen ständigen Wechsel zwischen dem Rückgriff auf Bekanntes und der Bewältigung neuer Situationen dar (Gruber & Stamouli, 2015, S. 33). Gerade bei Problemen, die in einem spezifischen Kontext eingebettet sind, kann im Gegensatz zu kontextfreien Problemstellungen, ein solcher Rückgriff auf eventuell vorhandenes Vorwissen erfolgen, um damit gezielt neue Informationen zu

generieren und dann auch zu verstehen und in Wissen weiterzuverarbeiten (Burmeister, 2009, S. 69).

Unter Vorwissen, ein Begriff der bereits impliziert, dass eine Wissensveränderung stattfindet (Krause & Stark, 2006, S. 38), ist all das problembezogene Wissen zu verstehen, das einer Person *vor* Beginn des Bearbeitungsprozesses eines Problems zur Verfügung steht und das sie abrufen kann (Wirth, 2004, S. 42; Süß, Kersting & Oberauer, 1993, S. 193). Dieses Wissen kann deklarativ oder prozedural repräsentiert sein (Renkl, 1996, S. 175), implizit oder explizit vorliegen (Krause & Stark, 2006, S. 38f.) und auch weniger strukturierte Wissensaspekte umfassen, wie etwa Alltagserfahrungen (ebd., S. 39), die von Mack (1999, S. 144) als subjektive Informationen bezeichnet werden. Jedes Wissen ist zu einem bestimmten Zeitpunkt als Vorwissen zu verstehen, das zur Bewältigung neuer Situationen transferiert wird (ebd.), jedoch kann nicht abgrenzend definiert werden, welches Wissen in einer konkreten Situation als Vorwissen dient (Wirth, 2004, S. 42). Hilfreich für eine zweckmäßige Beschreibung des Vorwissens sind die von Krause & Stark (2006, S. 39) formulierten Dimensionen, anhand derer Vorwissen kategorisiert werden kann: *Inhalt, Bewusstheit, Repräsentation, Wissenschaftlichkeit, Umfang* und *Handlungsrelevanz*. Die Inhaltsdimension stellt die Frage, worauf sich das Vorwissen bezieht. Sowohl fachspezifisches als auch fachübergreifendes Wissen können als Vorwissen für das Problemlösen relevant werden, wobei sich fachspezifisches Wissen auf die konkreten Informationsinhalte bezieht und fachübergreifendes Wissen auf Strategien und Heuristiken, die problemübergreifend angewendet werden können (Baker & O'Neil, jr., 2002, S. 616). Die Bewusstheitsdimension bezieht sich auf die Verbalisierbarkeit und damit die Bewusstheit des Wissens und unterscheidet zwischen explizitem und implizitem Wissen (vgl. Krause & Stark, 2006, S. 39). Eng mit der Bewusstheit des Vorwissens verbunden ist die Form der Repräsentation des Wissens. Hierbei geht es um die Frage, ob das benötigte Vorwissen in deklarativer oder prozeduraler Form abgespeichert und wie gut es vernetzt ist (ebd., S. 40). Inwiefern das Vorwissen wahr bzw. mit wissenschaftlichen Erkenntnissen vereinbar ist, berücksichtigt die Dimension Wissenschaftlichkeit. Der Umfang bezieht sich auf die Weite der relevanten Wissensstruktur, die für neue Informationen Anknüpfungspunkte bietet (ebd.). Die Ausprägung der genannten Dimensionen des Vorwissens bestimmen, wie gut das Vorwissen in einer konkreten Lernsituation genutzt werden kann und damit die Ausprägung der Dimension Handlungsrelevanz (ebd., S. 41).

Dass relevantes Vorwissen eine zentrale Rolle für den weiteren Wissenserwerbsprozess neben Intelligenz spielt, ist in der Lernforschung unbestritten (Krause & Stark, 2006, S. 42). Unterschiede im Vorwissen innerhalb einer Domäne führen zu unterschiedlichen Lernverläufen und damit zu unterschiedlichen Lernergebnissen (Mack, 1999, S. 142). Im Umgang mit komplexen Problemen, die in einem semantischen Kontext eingebunden sind, stellt Vorwissen damit eine Determinante der Problemlöseleistung dar (Burmeister, 2009, S. 69; Schoppek & Putz-Osterloh, 2003, S. 166). Vorhandenes Wissen lenkt die Aufmerksamkeit auf bestimmte Bereiche und beeinflusst folglich, welche Informationen wahrgenommen werden (vgl. Süß, Kersting & Oberauer, 1993, S. 193). Neben den zum Teil unbewusst stattfindenden Aufmerksamkeitsprozessen induziert Vorwissen auch bewusst Vorannahmen und Hypothesen über den vorliegenden Problemkontext (Beckmann & Goode, 2014; Wittmann, Süß & Oberauer, 1996, S. 4). Auch wenn es sich dabei noch um vage Vorstellungen und Kenntnisse handelt, beeinflussen sie die Anwendung spezifischer

Strategien im weiteren Bearbeitungsprozess (Scherer, 2014, S. 182; Molnár, Greiff & Csapó, 2013). Vorwissen bewirkt demnach entweder eine Erleichterung oder eine Erschwernis des weiteren Wissenserwerbs (Mack, 1999, S. 146).

Personen, die überhaupt keine Erfahrungen und Kenntnisse im Umgang mit einem bestimmten technischen System haben, sind nicht in der Lage Hypothesen über das Systemverhalten aufzustellen, die sich dann gezielt prüfen lassen. Die trial-and-error Strategie ist in diesen Fällen die einzige mögliche Vorgehensweise, um Wissen über das unbekannte System zu erwerben (OECD PISA Deutschland, 2003, S. 7). Da die trial-and-error Strategie durch eine fehlende Systematik der Handlung gekennzeichnet ist, ist sie wenig effektiv und auch wenig effizient (Van Der Linden, Sonntag, Frese et al., 2001, S. 190). Unsystematische Explorationsstrategien sind in der Folge fehleranfällig und führen zu einem geringeren Wissenserwerb als systematische Strategien, die eine Person einsetzen kann, wenn sie wenigstens über eine geringe Menge an Vorwissen verfügt (ebd.). Durch häufiger auftretende Fehler und eventuell wahrnehmbare Widersprüche erzeugen unsystematische Strategien bei der Erkundung komplexer Systeme negative Kognitionen und Emotionen (ebd.) und sorgen damit für ein Absinken der aktuellen Motivation. Der Vorteil bereits vorhandenen Wissens für den Umgang mit unbekanntem technischen Systemen liegt in der bereits zu Beginn der Systemexploration besseren Orientierung und des gezielteren Informationsbedarfs, wodurch systematischere Explorationsstrategien verwendet werden (Hesse, 1982, S. 85).

Vorhandenes Vorwissen wirkt sich aber nur dann vorteilhaft auf den weiteren Wissenserwerbsprozess und damit auf die Systemsteuerung aus, wenn das Vorwissen nicht im Widerspruch zur implementierten Systemstruktur steht (vgl. Funke, 1992b, S. 135; Beckmann, 1994, S. 63). Sind die generierten Informationen nicht mit dem bestehenden Wissen vereinbar, kann der weitere Wissenserwerb sogar behindert werden. Um den kognitiven Konflikt, der aus der Diskrepanz zwischen dem vorhandenen Wissen und den neuen Informationen entsteht, zu lösen, müssen bestehende Wissensstrukturen modifiziert werden (Akkommodation) (Krause & Stark, 2006, S. 42). Oftmals werden hingegen die neuen Informationen so ausgelegt, dass sie zu dem bestehenden Wissen passen (Assimilation) (ebd.). Wissen über die Funktion eines Objektes führt beispielsweise dazu, dass andere neue Verwendungsmöglichkeiten desselben Objektes nicht erkannt werden (Bertholet & Spada, 2004, S. 72). Dieser Effekt ist als *funktionale Gebundenheit* oder als *funktionale Fixierung* bekannt (Anderson, 2007b, S. 318f.). Ein weiterer Grund für die trotz vorhandenem Vorwissen erfolglose Problembearbeitung kann die Bevorzugung bestimmter Operatoren und Vorgehensweisen aufgrund früherer Erfahrungen sein (Ausubel, Novak & Hanesian, 1981, S. 656; Bertholet & Spada, 2004, S. 72). Solche sogenannten *Einstellungseffekte* verzerren den Problemlöseprozess (Anderson, 2007b, S. 321), wenn sie nicht korrigiert werden. Weil die aus ungeprüften Annahmen gezogenen Schlussfolgerungen zu einer Verschlechterung des Wissenserwerbs führen (Beckmann & Goode, 2014; Funke, 1983a, S. 21) spricht Wirth (2004, S. 45) in diesen Fällen von *negativem* Vorwissen.

Unter der Voraussetzung eines erfolgreichen Wissenserwerbs im probierenden Umgang mit einem dynamischen System, wird erwartet, dass das anfängliche Vorwissen an Bedeutung verliert, weil es durch spezifisches Systemwissen ergänzt oder verändert wurde (Wirth, 2004, S. 85). Süß, Kersting & Oberauer (1993, S. 193) bestätigen, dass mit zunehmender

Systemerfahrung der Einfluss des Vorwissens geringer wird und für die nachfolgende Steuerungsleistung das unmittelbar vorher erhobene Wissen der beste Prädiktor ist. Die Autoren fanden zudem heraus, dass die interindividuellen Differenzen des Wissens relativ stabil blieben, obwohl zwischen den beiden Messzeitpunkten eine Lernphase von etwa einer Stunde lag (ebd.). Dieser Befund spricht für die Annahme, dass das Vorwissen einen entscheidenden Einfluss auf den weiteren Wissenserwerb und damit auch indirekt auf die Problemlöseleistung hat. Vor diesem Hintergrund betont Wallach (1997, S. 144) die Bedeutung der Kontrolle des Vorwissens, wenn Problemstellungen in einem semantischen Kontext dargeboten werden. Zur Vorwissenskontrolle schlägt er vor, entweder eine wissenshomogenisierte Stichprobe einzusetzen oder die Möglichkeit zum Wissenserwerb zu geben (ebd.).

Mit zunehmender und differenzierterer Wissensbasis sollte außerdem der Einfluss von Intelligenz und anderen allgemeinen Fähigkeiten abnehmen. Diese kognitiven Grundfähigkeiten sind für den Ausgleich fehlender problemspezifischer Bearbeitungsstrategien relevant. Sie haben keine Bedeutung mehr, wenn diese spezifischen Prozeduren bekannt sind (Friege & Lind, 2003, S. 65). Je spezifischer die Wissensvoraussetzungen für das Lösen eines Problems, desto geringer wird der Zusammenhang zur Intelligenz ausfallen (OECD PISA Deutschland, 2003, S. 8). Die Leistungen in der Bearbeitung abstrakter, vorwissensfreier Problemstellungen sollte hingegen sehr hoch mit Intelligenz korrelieren (Beckmann, 1994, S. 55). Burmeister (2009, S. 76) merkt hingegen an, dass der Zusammenhang von Vorwissen und Steuerungsleistung bei kontextspezifischen Problemen lediglich eine Scheinkorrelation darstellt, weil sich sowohl das Vorwissen als auch die Steuerungsleistung auf eine Drittvariable, bei der es sich um Intelligenz handeln könnte, zurückführen ließen. Intelligenz als die Fähigkeit zu Lernen beeinflusst den Erwerb von Wissen in allen Bereichen und erklärt damit sowohl das bereits vorhandene Vorwissen als auch das während der Systemexploration erworbene Wissen (ebd.).

Teil II: Theoretisches Konstrukt

6 Alltägliche technische Systeme als Problemkontext

Nachdem im Teil I dieser Arbeit der bisherige Forschungsstand zum Problemlöseprozess, zur Problemlösekompetenz und zu den weiteren Einflussfaktoren auf den Problemlöseprozess und die Problemlöseleistung dargelegt wurde, wird nun das zu untersuchende Konstrukt technischer Problemlösekompetenz näher beschrieben. Die Spezifizierung des Konstruktes ist im Sinne wissenschaftlicher Erkenntnisgewinnung erforderlich für die Operationalisierung von Variablen, die wiederum für die anschließende Entwicklung eines Messinstrumentes notwendig sind. Da es sich bei dem in dieser Arbeit zu erfassenden Konstrukt um eine Kompetenz handelt, ist die bereits an mehreren Stellen dieser Arbeit geforderte exakte Beschreibung des *Kontextes* Inhalt der Konstruktspezifizierung. Demnach führt dieses Kapitel zunächst in den technischen Kontext, in dem Problemlösen stattfindet ein und verbindet dann Kontext- und Personenmerkmale im Rahmen der Theorie *soziotechnischer Systeme* (Ropohl, 1999) miteinander. Abschließend werden die im Forschungsstand beschriebenen Merkmale, die die Schwierigkeit von Problemen und Aufgaben bestimmen, aufgegriffen und für den Kontext soziotechnischer Systeme konkretisiert.

6.1 Problemlösen in technischen Kontexten

Die Kompetenz im Umgang mit dynamischen Systemen wurde im Laufe der Problemlöseforschung zu einem großen Teil in vorwissensfreien Kontexten untersucht. Vor allem in der Allgemeinen Psychologie werden hierzu abstrakte Szenarien eingesetzt, die auf den Anspruch naturgetreuer Abbildungen weitgehend verzichten (Müller, 1993, S. 31). Kritisiert werden solche semantikarmen dynamischen Szenarien, da sie aufgrund ihrer Realitätsferne die Komplexität der Problemsituation zu stark reduzieren (Neber, 1987, S. 22). Problemlösen, das hingegen in konkreten realen Situationen stattfindet, greift zusätzlich auf Informationen in Form von Wissen aus dem deklarativen und prozeduralen Gedächtnis zurück. Um die Anforderungen, die in *technischen Kontexten* an die Problemlöser gestellt werden, besser nachvollziehen und gezielt beeinflussen zu können, folgt an dieser Stelle die Beschreibung des Kontextes Technik.

Technische Problemlösekontexte umfassen zunächst einmal alle Kontexte, in denen *Technik* eine Rolle in der Problemsituation spielt. In einer von Ropohl (1990, S. 111) verfassten Definition des Technikbegriffes beinhaltet Technik dabei die folgenden drei Mengen:

- „die Menge der nutzenorientierten, künstlichen, gegenständlichen Gebilde (Artefakte bzw. Sachsysteme),
- die Menge menschlicher Handlungen und Einrichtungen, in denen Artefakte entstehen und
- die Menge menschlicher Handlungen, in denen Artefakte verwendet werden“.

Diese drei Mengen lassen sich zwei Kategorien zuordnen: technische Gegenstände bilden dabei die erste Kategorie und technische Handlungen die zweite Kategorie. Unter technischem Handeln wird alles menschliche Handeln im Umgang mit Technik verstanden (Banse, 1996, S. 109). Probleme entstehen allerdings nicht durch ein technisches Artefakt selbst, sondern erst durch den Umgang mit ihm, also durch *technisches Handeln*. Erst mit dem Vorhandensein von Zielen ergeben sich eventuelle Barrieren, die im Prozess des

Problemlösens zunächst überwunden werden müssen, um das Ziel zu erreichen. Ein technisches Produkt wie beispielsweise ein Navigationsgerät kann also erst dann zu einem Problem werden, wenn es gefertigt, eingebaut, verwendet, aktualisiert oder entsorgt werden muss. Unabhängig von dem technischen Artefakt, lassen sich Situationen, in denen mit Technik umgegangen wird, nach den spezifischen Handlungen, die typisch für sie sind, unterscheiden (Graube, 2009, S. 72). In der entsprechenden Abfolge beschreiben diese Situationen den Lebenslauf eines Produktes von der *Idee* über dessen *Entwicklung, Herstellung, Vertrieb, Nutzung* bis zu seiner *Entsorgung* (Pahl, Beitz, Feldhusen et al., 2007, S. 98f.). Probleme, die während der Konstruktion eines Produktes entstehen, stellen demnach andere Anforderungen an die problemlösende Person, als Probleme, die im Fertigungsprozess aufkommen. Während das spezifische technische Handeln der meisten Phasen des Produktlebenslaufes vorwiegend in industriellen oder handwerklichen Bereichen stattfindet, nimmt die Nutzung technischer Produkte auch im privaten Bereich eine große Stellung ein. Technische Gegenstände bieten dabei die Möglichkeit Handlungsziele zu erreichen, die ohne Technik nicht oder nur mit erheblichem Mehraufwand zu realisieren wären (Greif, Mitrea & Werner, 2008, S. 56). Dem Vorteil eines Zugewinns an Handlungsmöglichkeiten steht jedoch der Nachteil gegenüber, dass Technik wiederum eine komplexere Umwelt schafft, die weiteres technisches Handeln nötig macht (ebd., S. 57). Die Kompetenz im Umgang mit Technik ist demnach kein abstraktes Wissen, denn wer „technische Einrichtungen und Geräte nicht bedienen kann, bleibt vom Einsatz arbeitserleichternder, freizeitwerterhöhender oder mobilitätssteigernder Technikprodukte ausgeschlossen“ (Sackmann, Weymann & Hüttner, 1994, S. 59).

Um den unterschiedlich kompetenten Kunden möglichst einfach zu bedienende Technikgeräte anbieten zu können, fordern ganzheitliche Produktmanagementansätze, auch die Nutzung als Phase des Produktlebenslaufes bereits in die Entwicklung der Geräte einzubeziehen. Bisher wird diese Forderung nur unzureichend berücksichtigt. In der Regel wird von einer vereinheitlichten Gruppe von Nutzern und Nutzungskontexten ausgegangen (Greif, Mitrea & Werner, 2008, S. 59), von denen erwartet wird, dass sie sich der Technik anpassen (Zühlke, 2005, S. 29). Somit beeinflussen die im Alltag kaum noch wahrgenommenen technischen Geräte Handlungs- und Erfahrungsweisen sowie deren symbolische Verarbeitung (Hörning, 1997, S. 220; Rammert, 2006, S. 4). Erschwert wird der Aneignungsprozess außerdem durch die, aufgrund der Fortschritte in der Mikroelektronik, immer kürzeren Abstände, in denen Technologien und technische Produkte hervorgebracht werden (Zühlke, 2005, S. 29), sowie durch die immer kürzer werdenden Phasen des Lebenslaufes eines Produktes. In vielen der erzeugten Geräte ersetzen digitale Schnittstellen mechanische Schalter und Knöpfe (Tully, 2003, S. 37). Demzufolge sind heutige Geräte leistungsfähiger und komplexer geworden und in Zukunft werden sie multimodaler, das bedeutet, sie werden mehr Sinneskanäle der Nutzer ansprechen (Zühlke, 2005, S. 30). Der Umgang mit technischen Geräten nimmt damit eine völlig neue Form der Kommunikation an (Baier, 2003, S. 5).

Die immer schnellere Entwicklung technischer Geräte sowie deren kürzere Nutzungszeiten führen zu einer ständigen Entwertung konkreten Eingriffswissens und zu einer ständigen Veränderung bewährter Handlungsrouitinen. Aber auch die immer größer werdende Menge an Funktionen, die in den Geräten implementiert wird, lässt den Umgang mit ihnen zu einem

Problem werden. Denn durch die vielen verschiedenen Optionen, ein technisches Gerät zu nutzen, verschwindet dessen eindeutiger Zweckbezug (Tully, 2003, S. 92). Mit einem heute erhältlichen Mobiltelefon lässt sich beispielsweise nicht nur telefonieren oder Kurznachrichten senden, sondern auch e-mailen, Musik hören, im Internet surfen und mit entsprechenden Applikationen außerdem die Haustechnik steuern oder andere technische Geräte fernbedienen. Die Verwendung eines technischen Gerätes ist nicht mehr von vornherein festgelegt, Handlungsergebnisse sind zunächst einmal offen, weshalb Tully (2003, S. 64) diese neue Technik als *ergebnisoffene Technik* bezeichnet. Zur ergebnisoffenen Technik zählt er sowohl Informations- und Kommunikationstechniken als auch Unterhaltungselektronik, Haushalts- und Küchentechnik, Fahrzeuge und Umwelttechnik (Tully, 2003, S. 57), wobei technische Geräte, aufgrund der Vielfalt ihrer Funktionen nicht mehr eindeutig einem dieser Bereiche zugeordnet werden können. Aufgelöst wird aber nicht nur die Zweckbindung, sondern auch der Zusammenschluss von Aufgabe und Ort. Ergebnisoffene Technik erhält ihren Zweck folglich erst in der Anwendung durch den Nutzer/die Nutzerin. Dementsprechend liegt die Initiative auf der Seite der Benutzer, deren Unsicherheit im Umgang mit Technik nur durch eine entsprechende Beherrschung der Technik und durch das Anwenden von Eingriffswissen vermieden werden kann (ebd., S. 61).

In Bezug auf technische Alltagsgeräte meint die Anwendung von Eingriffswissen die sachgemäße Bedienung technischer Gegenstände. In Anbetracht der gestiegenen Komplexität technischer Systeme erscheint eine umfassende Kenntnis aller Funktionen aber kaum denkbar. Weil klare Zweckbezüge immer weniger zu unterstellen sind, lässt sich Technikwissen auch nicht mehr entlang den Funktionen eines technischen Gerätes vermitteln, womit *Ausprobieren* ein wesentlicher Bestandteil der Techniksozialisation wird (ebd.). Digitale Benutzungsschnittstellen sorgen dafür, dass Ausprobieren (bzw. *Explorieren*; ein Begriff der synonym verwendet wird), die Interaktion zwischen Mensch und technischem Artefakt verlangt (Schachtner, 1997, S. 16). Aufbauend auf der Erläuterung des Systembegriffs im nächsten Abschnitt, wird das Konzept soziotechnischer Systeme vorgestellt, um die Interaktionen mit technischen Geräten sinnvoll operationalisieren zu können.

6.2 Soziotechnische Systeme

Technische Artefakte können als vom Menschen geschaffene Systeme bezeichnet bzw. beschrieben werden (Pahl, Beitz, Feldhusen et al., 2007, S. 17). Ein (technisches) System setzt entsprechend seiner Funktion(en) Eingangsgrößen in Ausgangsgrößen um. Als Ein- und Ausgangsgrößen, die auch als Systeminputs und -outputs bezeichnet werden, kommen sowohl Stoffe, Energien als auch Informationen in Frage (Vahling, 2000, S. 338f.). Die Funktionen eines Systems bestimmen den Zusammenhang zwischen den Systeminputs und -outputs und damit die innere Struktur eines Systems. Die Systemstruktur wird beschrieben durch eine Menge an Objekten und Elementen, die so miteinander in Beziehung stehen und aufeinander einwirken, dass sie das Verhalten des Gesamtsystems beeinflussen (Graube, 2009, S. 57f.). Ein System kann je nach Komplexität in Teilsysteme untergliedert werden, die als Subsysteme bezeichnet werden. Sie erfüllen unterschiedliche Teilfunktionen innerhalb des Systems. Sowohl ihre Elemente als auch ihre Funktionen stehen dabei in Relation zueinander (Vahling, 2000, S. 338f.). Kann ein Objekt aufgrund nicht mehr existierender Elemente und Relationen einzelne oder alle Systemfunktionen nicht mehr erfüllen, verliert ein Objekt seine

Systemeigenschaft (Tauschek, 2006, S. 135). Ein technisches System ist außerdem durch seine Abgrenzung gegenüber der Umwelt gekennzeichnet. Durch die Systemgrenze wird festgelegt, welche Elemente noch Teil des Systems und welche Elemente Teil der Umgebung sind (Pahl, Beitz, Feldhusen et al., 2007, S. 40). Systeme stellen jedoch keine isolierte Menge von Elementen dar, ihre Grenze ist offen für Wechselbeziehungen mit der Umwelt, indem die Ein- und Ausgangsgrößen die Systemgrenzen überschreiten (Graube, 2009, S. 58). Auf einer hohen Abstraktionsebene lassen sich Systeme durch folgende Merkmale zusammenfassend beschreiben:

- Abgrenzung gegenüber Umgebung und anderen Systemen,
- Eingangs- und Ausgangsgrößen,
- Funktion(en) zur Beschreibung des Zusammenhangs zwischen Eingangs-, Ausgangs- bzw. Zustandsgrößen,
- Struktur als Aussage über das Anordnungs- und Beziehungsgefüge der Subsysteme (Ropohl, 1999, S. 77; Wagener & Haupt, 2000, S. 64).

Will man die Struktur eines technischen Systems auf einem höheren Auflösungslevel betrachten, so bieten sich nach Pahl, Beitz, Feldhusen et al. (2007) sowie Ropohl (1999) und Graube (2009) unterschiedliche Sichtweisen auf die Systemstruktur an. Die *Wirkstruktur* beschreibt die Kombination von Wirkprinzipien, die zur Erreichung der Gesamtfunktion verknüpft werden. Als Wirkprinzipien gelten dabei die physikalischen, chemischen oder biologischen Effekte (Hebeleffekt, Keileffekt, Reibungseffekt, etc.), die zusammen die nicht weiter zerlegbaren Elementarfunktionen realisieren (ebd.). Die *Baustruktur* legt den Fokus ihrer Betrachtung hingegen auf die einzelnen Bauteile, Baugruppen und ihren Verbindungen, mit denen die Wirkprinzipien ermöglicht werden können (Pahl, Beitz, Feldhusen et al., 2007, S. 56). Basis für die Konkretisierung eines technischen Systems innerhalb der Wirk- und Baustruktur ist die Definition seiner Aufgabe, aus der die Funktion(en) des Systems abgeleitet werden. Die Verknüpfung der Teilfunktionen, in die die Gesamtfunktion zerlegbar ist, wird durch die *Funktionsstruktur* beschrieben (Graube, 2009, S. 76f.). In diesem, von Ropohl (1999, S. 75ff.) als *funktionales Konzept* bezeichneten Systemaspekt, wird das technische System als eine *black box* beschrieben, in der die von außen zu beobachtenden Zusammenhänge von Funktionen und Eigenschaften dargestellt werden. Im Gegensatz zur Wirk- oder Baustruktur beschreibt die Funktionsstruktur nicht, wie ein technisches Gerät aufgebaut ist, sondern lediglich sein Verhalten. Die auf Ein- und Ausgangsgrößen sowie Systemzustände ausgerichtete Betrachtungsweise liegt auch dem alltäglichen Umgang mit technischen Systemen zu Grunde. Um mit einem technischen System, wie einem Fahrscheinautomaten agieren zu können, muss der Nutzer/die Nutzerin nicht wissen, aus welchen Einzelteilen das Gerät aufgebaut ist und welche Funktion jedes einzelne Teil hat. Ausgehend von der eigenen Handlung erwartet die bedienende Person eine bestimmte Reaktion seitens des Systems, auf deren Basis sie die nächste Aktion ausführt.

Diese Wechselwirkungen zwischen dem handelnden System *Mensch* auf der einen Seite und dem *technischem System* auf der anderen Seite werden in dem übergeordneten *soziotechnischen System* näher betrachtet (Eckert, 2000, S. 128) (Abbildung 6.1). Durch die Wahrnehmung technischer Artefakte als Teilsysteme soziotechnischer Systeme rückt der Mensch dabei in den Vordergrund (Wagener & Haupt, 2000, S. 64). Er wirkt mit dem Ziel der

Funktionserfüllung auf das technische System ein und erhält von dem System eine, entsprechend der inneren Struktur mögliche Rückmeldung. Ist er in der Lage, die erhaltene Rückmeldung richtig zu interpretieren, kann er zielgerichtet weitere Benutzereingaben vornehmen (Pahl, Beitz, Feldhusen et al., 2007, S. 56). Auch hoch automatisierte Systeme, die zu ihrer Funktionserfüllung nur wenige Benutzereingaben benötigen, werden als soziotechnische Systeme betrachtet, weil sie von Menschen für Menschen entwickelt und betrieben werden. Demnach dürfen keine technischen Systeme isoliert betrachtet werden, sondern immer als Teil eines soziotechnischen Systems, zu dem außerdem der Mensch und die weiteren die Kommunikation beeinflussenden Faktoren gehören (Grote, 2005, S. 65f.). Zu den weiteren Einflussfaktoren zählen beispielsweise die auszuführende Aufgabe sowie konkrete Umgebungseinflüsse wie Lärm, Zeitdruck oder vorhandene Hilfestellungen.

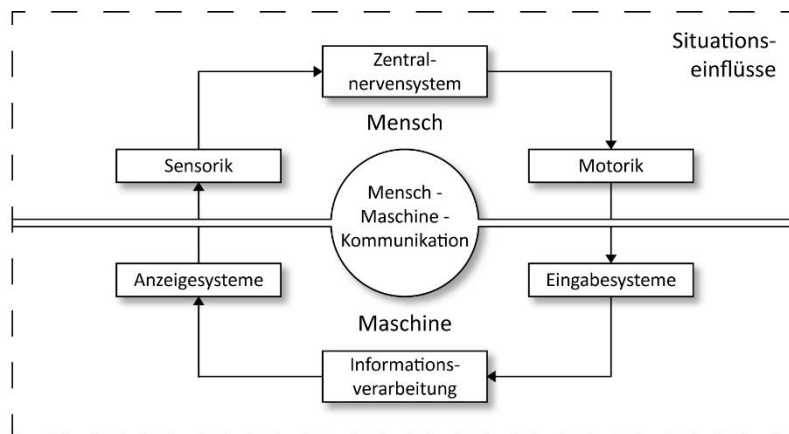


Abbildung 6.1: Soziotechnisches System (nach Möller, 2003, S. 23)

6.3 Interaktionen in soziotechnischen Systemen

Weil menschliches Handeln und technische Abläufe in soziotechnischen Systemen eng miteinander verbunden sind (Rammert & Schulz-Schaeffer, 2002, S. 7), wird in diesen Fällen oftmals von Kommunikation oder Interaktion zwischen Mensch und technischem System gesprochen (Hohmann, 2002, S. 19). Die Begriffe, die in diesem Zusammenhang häufig Verwendung finden, sind *Mensch-Maschine-Kommunikation (-Interaktion)* oder *Mensch-Computer-Kommunikation (-Interaktion)*. Heinecke (2012, S. 4) merkt an, dass die Unterscheidung zwischen Maschine und Computer zunehmend schwieriger wird, da in technischen Geräten immer mehr Komponenten enthalten sind, die früher fast ausschließlich in Computeranwendungen zu finden waren. Deshalb sei es zweckmäßig allgemeiner von *Mensch-System-Interaktion* zu sprechen.

Inwieweit der menschliche Umgang mit technischen Artefakten als Kommunikation oder Interaktion bezeichnet werden sollte, ist eine Fragestellung der Techniksoziologie. Die Verwendung des Kommunikationsbegriffs im Hinblick auf technische Systeme hat zur Konsequenz, dass diese nicht mehr nur als Objekte für menschliche Zwecke betrachtet werden, sondern oftmals als Teilnehmer sozialer Situationen (Braun-Thürmann, 2002, S. 10; Rammert, 2006, S. 6). Kommunikation beschreibt einen Vorgang, bei dem zwei oder mehrere Kommunikationspartner einander Nachrichten übermitteln. Die Übergabe von Nachrichten erfolgt mittels Signalen, indem ein Kommunikationspartner ein Signal sendet und ein anderer Partner dieses Signal empfängt (Rupprecht, 2014, S. 1). Nach Shannon's Informationstheorie

(Shannon & Weaver, 1963) soll das zum Zweck der Kommunikation erzeugte Signal dem Empfänger eine Information liefern. Der Informationsgehalt ist jedoch von der Neuigkeit abhängig, die das Signal für den empfangenen Kommunikationspartner hat.

Von Kommunikation wird auch dann gesprochen, wenn es keinen Wechsel zwischen Empfänger und Sender gibt, die Kommunikation also nur in eine Richtung erfolgt. Bedingung für eine Kommunikation ist jedoch, dass die gesendete Nachricht von dem Empfänger auch interpretiert werden kann (Hohmann, 2002, S. 26). Im Idealfall erfolgt eine störungsfreie Kommunikation, wenn der Sender die mitzuteilende Nachricht einem spezifischen Signal zuordnet, aus dem der Empfänger dieselbe Nachricht entnehmen kann (Rupprecht, 2014, S. 9). Als Kommunikationspartner kommen hierfür allerdings nicht nur Menschen, sondern auch Tiere, Maschinen und andere Objekte in Frage (Rupprecht, 2014, S. 1). Die Besonderheit einer Kommunikation zwischen technischen Systemen untereinander beziehungsweise zwischen technischen Geräten und ihren Nutzern liegt in dem Umstand, dass die von dem System empfangenen Signale (z. B. binäre Codes) nur auf Basis vorheriger Vereinbarungen interpretiert werden können (Hohmann, 2002, S. 25).

Eine Spezifizierung von Kommunikation stellt die Interaktion dar (ebd., S. 29), bei der die Kommunikationspartner außerdem interagieren („in | ter | agie | ren <lat.-nlat.>: sich, agierend u. aufeinander reagierend, wechselseitig in seinem Verhalten beeinflussen (von Menschen, auch z. B. von Computersystemen, Medien usw. u. deren Benutzern)“ (Duden - Das Fremdwörterbuch, 9. Aufl. Mannheim 2007 [CD-ROM])). Als interaktiv wird ein technisches System demnach bezeichnet, wenn sein Verhalten durch eine Person oder ein Artefakt beeinflusst werden kann (Heinecke, 2012, S. 3). Anders als bei Menschen ist das Verhalten eines technischen Systems nicht die Folge eines Sozialisierungsprozesses, sondern ein Produkt technischen Handelns (Braun-Thürmann, 2002, S. 15). Eine durchaus komplexe Form der Interaktivität technischer Systeme lässt sich beispielsweise durch den Dialogbetrieb realisieren, bei dem der Begriff *Dialog* die wechselseitige Aktivitäten zwischen System und Nutzer/Nutzerin ausdrückt (Heinecke, 2012, S. 3). Dabei reagiert das technische System auf die Eingaben der nutzenden Person durch eine entsprechende Rückmeldung. Nachdem die Eingabe verarbeitet wurde, kann die Person wieder eine Aktion ausführen. Ein Dialog ist immer nur dann möglich, wenn beide Kommunikationspartner die gleiche Sprache sprechen. Kann keine direkte Kommunikation stattfinden, müssen sogenannte Zwischeninstanzen vorgesehen werden, über die die Kommunikation abläuft (Rupprecht, 2014, S. 12).

Eine solche Zwischeninstanz beschreibt die sogenannte *Benutzungsschnittstelle* (das *Interface* = Schnittstelle). Über die Benutzungsschnittstelle kommuniziert der Benutzer/die Benutzerin mit dem technischen System (Stary, 1996, S. 20) (Kapitel 2.4.1). Sie umfasst alle Bestandteile eines interaktiven Systems, mit denen in das System eingegriffen werden kann und über die das System Informationen zur Verfügung stellt (Sarodnick & Brau, 2011, S. 21). Generell lassen sich Systembestandteile in Hardware- und Softwarekomponenten einteilen (Stary, 1996, S. 14; Haase, 2005) und die Komponenten einer Benutzungsschnittstelle außerdem in Bedien- und Anzeigeelemente (Baumann & Lanz, 1998, S. 28). Über die Bedienelemente erfolgen in der Regel die Eingaben der nutzenden Person, um eine Veränderung an dem System vorzunehmen und über die Anzeigeelemente erfolgen in der Regel die Rückmeldungen des Systems, indem Informationen angezeigt werden (ebd.). Mit der Betätigung eines

Bedienelementes wird in den meisten Fällen eine (diskrete) Variable verändert und/oder ein Ereignis ausgelöst, das den aktuellen Zustand eines Systems in einen anderen Zustand überführt. Als traditionelle Bedienelemente kommen zum Beispiel Tasten, Druckknöpfe, Wippschalter, Drehknöpfe, Stellscheiben, Schieber sowie Tasten mit Mehrfachfunktion in Frage (Baumann & Lanz, 1998, S. 37ff.). Diese werden jedoch durch den zunehmenden Einbau von Computersystemen in den technischen Geräten verdrängt. Die beispielsweise in Kühlschränken, Fernsehern, Telefonanlagen, Uhren und Waschmaschinen eingebauten Computersysteme werden benötigt, um die Geräte:

- intelligenter zu machen,
- sie an die Präferenzen und Bedürfnisse der Benutzer anzupassen,
- ihnen ein Gedächtnis zu geben,
- sie in höherem Maße autonom zu machen und
- sie mit anderen Geräten zu vernetzen (Herczeg, 2006, S. 67).

Reale Bedienelemente werden dabei von virtuellen Bedienelementen ersetzt. Diese werden von einem Rechner erzeugt und auf einem Bildschirm dargestellt (Baumann & Lanz, 1998, S. 61). Ein realer Taster lässt sich etwa durch eine Schaltfläche ersetzen, die bei einem Klick das gleiche Ereignis auslöst wie das Drücken einer Taste. Als Ersatz für einen Drehknopf bietet sich ein Radio-Button-Menü oder eine Auswahlliste an, aus denen jeweils eine Option ausgewählt werden kann (Dahm, 2006, S. 219ff.). Beispiele für den Ersatz realer durch virtuelle Bedienelemente gibt Abbildung 6.2. Für eine größere Menge von Funktionen bzw. Bedienelemente bieten sich permanente oder flüchtige Menüs als virtuelle Bedienelemente an, da sie sich mit Menübäumen gut strukturieren lassen und nur bei Bedarf sichtbar gemacht werden (Baumann & Lanz, 1998, S. 63ff.).

Die Art und Weise, mit der eine Person mit einem technischen System über die Benutzungsschnittstelle interagiert, unterscheidet sich anhand des Ziels, das erreicht werden soll. Intendiert eine Person mit einem technischen System eine konkrete Aufgabe zu erledigen (z. B. einen Telefonanruf zu tätigen), muss das technische Gerät (hier das Telefon) zielgerichtet bedient werden. Um die zur Zielerreichung erforderlichen Eingriffe zu tätigen, müssen diese der Person bekannt sein. Die zielgerichtete Steuerung eines technischen Systems stellt eine Form der Interaktion dar. Eine andere Form der Interaktion ergibt sich, wenn einer Person ein ihr (teils) unbekanntes technisches System vorliegt. In diesem Fall stellt die Bedienung des Systems für die Erledigung konkreter Aufgaben ein Problem dar, weil das benötigte Wissen hierfür fehlt. Die Interaktion mit dem technischen System erfolgt hier mit dem Ziel, zunächst das fehlende Eingriffswissen zu erwerben.



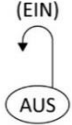
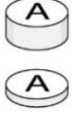
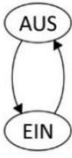
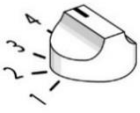
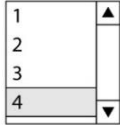
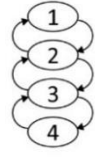
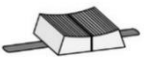

Bedienelement	mechanisch	virtuell	Zustandsdiagramm
Taste			
Knopf		<input type="checkbox"/> A <input checked="" type="checkbox"/> A	
Drehknopf			
Schieber			

Abbildung 6.2: Beispiele für den Übergang von realen zu virtuellen Bedienelementen (nach Baumann & Lanz, 1998)

Die gestiegene Komplexität heutiger technischer Systeme lässt die Erwartungen an vorhandene Funktionen unklarer werden. So ist bei dessen Kauf das Verkaufspersonal ebenso wenig mit den Funktionen vertraut, wie die Käufer. Bedienungsanleitungen als Informationsquelle haben in vielen Bereichen ebenfalls an Bedeutung verloren, wenn zunächst allgemein die Nutzungsmöglichkeiten der Systeme kennengelernt werden wollen. Viele Produkthersteller sind bereits dazu übergegangen, ihren Produkten keine Bedienungsanleitung mehr in Papierform beizulegen. Bedienungsanleitungen sind oftmals in digitaler Version auf einer CD beigelegt oder müssen von der Herstellerwebsite im Internet bezogen werden. Für zusätzliche Applikationen (z. B. Software für den Computer oder Apps für Smartphones) sind solche Anleitungen nur schwer oder überhaupt nicht zu finden. Der Erwerb von Wissen ist somit zum größten Teil durch eigene zufällige oder selbstregulierte Entdeckungen möglich. Wenn diese Entdeckungen nicht zufällig erfolgen, sondern zielgerichtet, spricht man auch von Ausprobieren oder Explorieren.

Um ein technisches Gerät bedienen zu können ist vor allem Eingriffswissen erforderlich, also prozedurales Wissen in Form von Wenn-Dann-Regeln. Hierzu muss eine Person über die Benutzungsschnittstelle Eingaben tätigen, die intern eine Zustandsänderung herbeiführen oder auch nicht, und hierüber eine Rückmeldung an die Benutzungsschnittstelle geben. Die Informationen müssen von der bedienenden Person wahrgenommen und im deklarativen Gedächtnis gespeichert werden. Je ausführlicher eine Person ein technisches System exploriert, desto mehr Zustandsübergänge finden statt und je mehr Informationen werden generiert. Je öfter eine Person dieselben Eingaben in demselben Zustand ausführt, desto eher erfolgt eine Wissenskompilation, bei der das deklarative Wissen in prozedurales Wissen überführt wird.

Dieses Wissen ist erforderlich, um in einer konkreten Situation eine Aufgabe mit dem technischen System zu erledigen. Dazu muss das System mit einer gezielten Folge von Eingriffen in einen bestimmten Systemzustand gebracht werden. Je nach Aufgabenstellung sind das unterschiedliche und unterschiedlich viele Eingriffe in das System. An einem Backofen mit digitaler Benutzungsschnittstelle lässt sich beispielsweise die Funktion *Backen mit Ober-/Unterhitze* aufrufen sowie die Einstellung der Temperatur mit wenigen Eingriffen vornehmen. Soll jedoch eine Uhrzeit programmiert werden, die den Backvorgang beginnen lässt, sind völlig andere Eingriffe nötig. Um eine Folge von Eingriffen auf einem direkten und fehlerfreien Weg durchführen zu können, muss das Verhalten des technischen Systems vorhersagbar sein. Hierzu müssen die Wenn-Dann-Regeln (die Zustandsübergangsfunktionen) bekannt sein.

Die Interaktion mit technischen Systemen unterscheidet sich also danach, welches Ziel mit der Interaktion erreicht werden soll. Steht der Erwerb von Eingriffswissen im Vordergrund, so wird eine Person sehr viele Eingriffe in das System vornehmen, um so viele Informationen über die Zustandsübergänge zu erhalten. Besteht das Ziel der Interaktion hingegen darin, das System in einen klar definierten Systemzustand zu bringen, so müssen die zu tätigen Eingriffe in einer der Aufgabenstellung entsprechenden Reihenfolge ausgeführt werden.

6.4 Merkmale, die die Interaktion mit technischen Systemen erschweren

Nachdem in den vorherigen Abschnitten technische Systeme als solche sowie die Möglichkeiten mit ihnen zu kommunizieren erläutert wurden, fehlt für die Entwicklung eines Testinstrumentes bzw. den in ihm enthaltenen Items, außerdem die Beschreibung der Merkmale, die die Schwierigkeit im Umgang mit technischen Systemen bestimmen. Für die Testentwicklung ist die Kenntnis über solche schwierigkeitsbestimmenden Merkmale insofern von Bedeutung, als dass sie dann gezielt variiert werden können, um unterschiedlich schwierige Items zu generieren. In dem Theorieteil dieser Arbeit wurden bereits potenziell schwierigkeitsbestimmende Merkmale aus der Literatur extrahiert und anhand der kognitionspsychologischen Lerntheorie von Anderson theoretisch begründet.

Die von Funke (1990) zur systematischeren Differenzierung von Problemmerkmalen vorgenommene Unterscheidung in Personen-, System- und Situationsmerkmale wird für diese Arbeit übernommen. Zu den Merkmalen einer Person gehören alle Eigenschaften und Kenntnisse, die eine Person in die Problemsituation mitbringt (ebd.). Schwierigkeitsrelevante Faktoren wie das Selbstvertrauen der Probanden, ein Problem lösen zu können, die Motivation oder das Vorwissen können bei der Testerstellung jedoch nicht berücksichtigt werden, da diese vorab nicht bekannt sind (Schumann & Eberle, 2011).

Zu den eigentlichen Systemmerkmalen werden die bereits von Dörner & Kreuzig (1983) zur Beschreibung komplexer Probleme formulierten Merkmale *Komplexität*, *Intransparenz*, *Vernetztheit* und *Eigendynamik* gezählt, wobei die Autoren nicht zwischen Situations- und Systemmerkmalen unterschieden haben. Situationsmerkmale beschreiben nach Funke (1990) die konkrete Problemstellung unabhängig von dem gegebenen dynamischen System. Für die Entwicklung von technischen Systemen, die sich hinsichtlich der Schwierigkeit mit ihnen zu interagieren unterscheiden, müssen die zunächst grob definierten Merkmale Komplexität, Vernetztheit, Dynamik, Intransparenz durch konkrete Variablen präzisiert werden. Dabei sind

jedoch die unterschiedlichen Interaktionsformen zu berücksichtigen. Soll ein technisches System zunächst ausprobiert werden, mit dem Ziel Eingriffswissen zu erwerben, so sind es ausschließlich die Merkmale des Systems, die die Interaktion erschweren können. Soll das technische Gerät gezielt gesteuert werden, bestimmt die konkrete Aufgabe ebenfalls die Interaktionsschwierigkeit. Die konkrete Aufgabenstellung ist unabhängig von dem jeweils vorliegenden System und stellt ein Merkmal der Situation dar. In den nächsten beiden Abschnitten sollen differenziert nach dem technischen System und der konkreten Aufgabenstellung die genannten Merkmale konkretisiert werden, so dass Tabelle 6.1 mit Inhalt gefüllt wird.

Tabelle 6.1: zu operationalisierende schwierigkeitsbestimmende Merkmale

	Komplexität	Vernetztheit	Dynamik	Intransparenz
Systemmerkmale	?	?	?	?
Situationsmerkmale	?	?	?	?

6.4.1 Systemmerkmale

Im alltäglichen Umgang mit technischen Systemen erfolgt die Interaktion über deren Benutzungsschnittstellen (Interface). Merkmale, die eben diese Interaktion schwierig machen, sind deshalb auch vorrangig auf der Interfaceebene auszumachen. Die Aufspaltung des Interface in die Komponenten Hardware und Software bietet sich auch zur Präzisierung von Systemmerkmalen an, denn das Merkmal Vernetztheit muss in der Hardware anders realisiert werden, als in der Software. Zunächst werden konkrete Merkmale der Systemhardware besprochen und anschließend die der Software technischer Systeme.

Hardware

Komplexität als schwierigkeitsbestimmendes Merkmal meint in dieser Arbeit den quantitativen Umfang des technischen Systems, also die Anzahl an Komponenten und Variablen in einem System. Bezogen auf die Hardware der Benutzungsschnittstelle kann die Menge der zu berücksichtigenden Systemkomponenten durch die Anzahl der Bedienelemente realisiert werden. Die Variable Anzahl der Bedienelemente umfasst dabei alle Tasten, Knöpfe, Schalter und Regler der Hardware, die betätigt werden können, um mit den technischen Systemen zu interagieren. Mit jedem zusätzlichen Bedienelement ist auch eine zusätzliche Information verbunden, die wahrgenommen und zunächst im deklarativen Gedächtnis repräsentiert werden muss, weshalb von einem Einfluss auf die Schwierigkeit im explorierenden Umgang mit einem technischen System ausgegangen werden muss.

Das Merkmal der Vernetztheit beschreibt, inwieweit Systemkomponenten und Variablen miteinander verknüpft sind. Ein stark vernetztes System lässt sich auf Hardwareebene konstruieren, indem nur wenige Bedienelemente für viele Funktionen zur Verfügung stehen. Um alle Bedienhandlungen ausführen zu können, müssen Bedienelemente mit mehreren Funktionen belegt sein. Diese können dann erreicht werden, indem Bedienelemente mehrfach gedrückt, gedrückt gehalten oder in unterschiedlichen Systemzuständen (bspw.

ausgeschalteter Systemzustand) betätigt werden. Mehrfachbelegungen führen dazu, dass die Reaktion von einem Systemzustand abhängt. Bedienfehler resultieren dann häufig daraus, dass sich das System nicht in dem Zustand befindet, den der Benutzer erwartet (Preim & Dachzelt, 2010, S. 145).

Die Veränderungen von Systemzuständen, die ohne das Eingreifen seitens der bedienenden Person entstehen, werden mit dem Merkmal Dynamik beschrieben. Autonome Veränderungen an Hardwarekomponenten lassen sich in einer Computersimulation umsetzen, indem die Bedienelemente der Hardware (etwa Knöpfe, Tasten oder Regler) von allein betätigt werden. Da ein solches Systemverhalten in der Realität nicht (oder vereinzelt in Störungsfällen) auftritt, soll bei der Itemkonstruktion hardwareseitig auf das Merkmal Dynamik verzichtet werden.

Das Merkmal Transparenz beschreibt wie deutlich und erkennbar Systemzustände, -variablen sowie erforderliche Ereignisse sind. Eine fehlende Transparenz führt zur Unsicherheit im Handeln und zur falschen Interpretation von Ein- und Ausgabegrößen. Variieren lässt sich das Merkmal für die Hardware eines technischen Systems, indem Bedienelemente hinsichtlich der Verständlichkeit ihrer Funktionen und hinsichtlich der Sichtbarkeit ihrer Betätigung modifiziert werden. Eine Variationsquelle ergibt sich durch die Gestaltung der Bedienelemente selbst. Mit ihrem Aussehen geben Bedienelemente visuelle Hinweise zu ihrer Bedienung. So vermitteln Knöpfe den Eindruck, dass sie gedrückt werden können und Hebel, dass sie hin- und herzubewegen sind. Diese wahrgenommenen Eigenschaften eines technischen Geräts werden als *Affordance* bezeichnet (Preim & Dachzelt, 2010, S. 136). Bedienelemente, die eine Bedienung vermitteln, die nicht möglich ist (man spricht von falscher *Affordance*), können dazu führen, dass das technische Gerät beschädigt wird (ebd., S. 138). So kann beispielsweise ein Hebel, der zwar aussieht, als könne man ihn in eine Richtung bewegen, aber sich lediglich ziehen lässt abgebrochen werden. Durch ihr spezifisches Aussehen können Bedienelemente ein technisches Gerät selbsterklärender und damit transparenter machen. Je größer die Anzahl der Bedienelementarten ist, desto einfacher wird demnach die Handhabung eines Gerätes. Eine weitere Möglichkeit der Transparenzvariation ist die Beschriftung der Bedienelemente. Anhand bestimmter Termini und Symbole, die auf oder in unmittelbarer Nähe zu den Bedienelementen angebracht sind, können die bedienenden Personen Rückschlüsse auf die hinter den Bedienelementen liegenden Funktionen ziehen (Preim, 1999, S. 57). Um die Sichtbarkeit der erfolgten Bedienhandlungen mithilfe der Hardware zu erhöhen oder zu verringern, kann eine Rückkopplung durch das entsprechend betätigte Bedienelement erfolgen. Rückkopplung bezeichnet die (unmittelbare) Rückmeldung des Systems auf die erfolgte Bedienhandlung (Preim & Dachzelt, 2010, S. 137), um den Benutzer/die Benutzerin über die erfolgreiche Interaktion zu informieren. Die Rückmeldung von Bedienelementen der Hardware kann durch die Bauteile selbst erfolgen. So nimmt beispielsweise ein Kippschalter bei seiner Bedienung eine neue Stellung ein und eine Taste lässt durch ihren Druckpunkt die erfolgreiche Interaktion erkennen (Baumann & Lanz, 1998, S. 17). Neben haptischen Rückmeldungen ermöglichen auch akustische und optische Signale eine erfolgte Eingabe. Die Rückmeldung kann hier etwa durch einen Signalton oder durch das Einschalten einer Tastenbeleuchtung dargestellt werden.

Software

Für die Umsetzung von Komplexität als Maß für die Quantität der Informationen, die bei der Systemexploration generiert und verarbeitet werden müssen, gibt es auf Softwareebene weitaus mehr Möglichkeiten, als auf Hardwareebene. Liegt die Intention im interaktiven Umgang mit technischen Geräten darin, Wissen über vorhandene Eingriffsmöglichkeiten zu erwerben, bestimmt die Menge an Eingriffsmöglichkeiten die Schwierigkeit. Mit der Anzahl der Zustände, die das technische System annehmen kann, kann die Menge der Eingriffsmöglichkeiten variiert werden. Die Anzahl der Variablen, die als Zustandsgrößen das System beschreiben und mit denen Zustandsänderungen herbeigeführt werden können, beschreiben ebenso wie die Zustandsanzahl den quantitativen Umfang und damit die Komplexität eines technischen Systems. Die Möglichkeit mittels der Software Funktionen herbeizuführen oder Variablen zu verändern, kann durch den Einsatz virtueller Bedienelemente umgesetzt werden. In technischen Systemen, die eine große Anzahl an virtuellen Bedienelementen vorsehen, werden diese aus Übersichts- und Platzgründen oftmals zu Gruppen zusammengefasst und in Menüs strukturiert (Preim, 1999, S. 59). Menüeinträge, die inhaltlich zusammengehören, stehen dabei auf einer Navigationsebene (ebd., S. 60). Die Menübreite beschreibt dabei die maximale Anzahl der Menüpunkte auf einer Navigationsebene. Die Menütiefe gibt die maximale Anzahl der Ebenen an, durch die die Nutzer in einem Menü navigieren können. Menüs sind demnach hierarchisch aufgebaut, die Form der Hierarchie bestimmt dabei die Schwierigkeit in der Interaktion mit einem Gerät. Es wird davon ausgegangen, dass die Navigation in einer breiten Hierarchie einfacher ist als in einer schmalen und tiefen Hierarchie (Preim, 1999, S. 102).

Um das Merkmal Vernetztheit für die Software technischer Systeme zu präzisieren, kann man sich zum einen die Vernetztheit der Menüs anschauen und zum anderen die Vernetztheit von Variablen. In einem stark vernetzten Menü sind Menüpunkte aus einer Navigationsebene mit anderen Menüpunkten aus einer anderen Navigationsebene verknüpft. Je größer die Anzahl dieser Verknüpfungen und damit die Vernetztheit eines Menüs, desto mehr verliert das Menü an Struktur und Übersicht, wodurch eine Orientierungslosigkeit bei der Navigation eintreten kann. Lassen sich vorgenommene Interaktionen mit einem technischen System, wie beispielsweise die Navigation durch ein Menü oder die Veränderung von Variablen direkt wieder rückgängig machen, so handelt es sich um eine reversible Bedienhandlung. Muss die nutzende Person erst durch alle Menüpunkte navigieren, um wieder an den gewünschten Ausgangspunkt zu gelangen fehlt die Reversibilität und die Bedienung des entsprechenden Gerätes wird erschwert. Kann eine Variable nur verändert werden, wenn eine andere Variable einen bestimmten Wert aufweist, so liegt eine Variablenbedingung vor. Je größer die Anzahl solcher Bedingungen, desto stärker ist die Vernetzung des technischen Systems und damit die Schwierigkeit es zu bedienen.

Die dynamische Eigenschaft eines technischen Systems wird durch das Übertragungsverhalten zwischen Eingangs- und Ausgangsgrößen charakterisiert (Johannsen, 1993, S. 34). Die Merkmalskategorie Dynamik kann einzig mit Variablen auf Seiten der Software realisiert werden. Dabei lassen sich zeitverzögerte Reaktionen eines technischen Systems sowie Eigendynamiken konkretisieren. Reaktionen des technischen Systems auf eine von der nutzenden Person getätigte Bedienhandlung, die erst nach Ablauf einer gewissen Zeit

erfolgen, führen dazu, dass sie diese nicht als Antwort auf das ausgelöste Ereignis wahrnimmt, sondern eventuell als Eigendynamik oder als Wirkung auf bereits nachfolgende Bedienaktionen. Eigendynamiken beschreiben Zustandsänderungen von Geräten, die ohne eine Interaktion seitens des Nutzers/der Nutzerin erfolgen können. Eine Software ließe sich beispielsweise so programmieren, dass ein System nach einer festgelegten Zeit in den Startzustand übergeht, unabhängig von den Eingaben des Nutzers/der Nutzerin.

Zur Konkretisierung des Merkmals Transparenz für die Softwarekomponente des Interfaces können analog zur Konkretisierung für die Hardwarekomponente, die Verständlichkeit der virtuellen Bedienelemente sowie die Rückmeldung der Software bei ausgeführten Bedienhandlungen herangezogen werden. Zwar haben virtuelle Bedienelemente im Gegensatz zu realen Bedienelementen keine reale Affordance, Icons, Buttons oder Scrollbars sollten aber eine wahrnehmbare Affordance haben. Das heißt, sie sollten so gestaltet sein, dass eine bestimmte Form der Bedienung nahe gelegt wird. Ein Button sollte also so gestaltet sein, dass er angeklickt wird und eine Scrollbar so, dass sie bewegt wird. „Dabei ist die wahrgenommene Affordance teilweise schwer von Konvention zu trennen [...] also etwas, was erlernt worden ist“ (Preim & Dachsel, 2010, S. 138).

6.4.2 Situationsmerkmale

Möchte eine Person ein technisches System mit dem Ziel bedienen, eine konkrete Aufgabe zu erledigen, lässt sich die Schwierigkeit in der Bedienung des Systems nicht nur auf die Systemmerkmale zurückführen. Das Interaktionsverhalten ändert sich entsprechend dem zu erreichenden Ziel, das unabhängig von dem vorliegenden System sein kann. So macht es für die nutzende Person einen Unterschied, ob sie an einer ihr unbekanntem Spülmaschine lediglich ein Programm auswählen möchte (ein Problem, dessen Lösung wenige Bedienhandlungen erfordert), oder ob sie die Angabe über den Härtegrad des Wassers ändern möchte (eine Problemstellung, bei der eventuell mehrfachbelegte Bedienelemente zu berücksichtigen sind). Die Aufgabenstellung ist daher ein Element der Situation, für die die schwierigkeitsbestimmenden Merkmale analog zum System konkretisiert werden müssen.

Wie in dem oben angeführten Beispiel für unterschiedliche Ziele bei der Bedienung einer Spülmaschine bereits angedeutet, lassen sich Aufgabenstellungen hinsichtlich ihrer Komplexität variieren, indem für sie eine unterschiedlich große Anzahl an Eingriffen nötig ist. Je mehr Bedienhandlungen nötig sind, um ein System zielgerichtet zu steuern, umso wahrscheinlicher ist das Auftreten fehlerhafter Eingriffe. Die Komplexität der Aufgabenstellung kann außerdem durch die Anzahl der Variablen, die verändert werden müssen, angepasst werden. Je mehr Variablen modifiziert werden müssen, umso schwieriger wird das Erreichen des Steuerungsziels.

Die in dem System vorhandene Vernetzung, die sich durch eine Mehrfachbelegung von Bedienelementen oder durch Variablenbedingungen zeigt, wirkt nur dann schwierigkeitsbestimmend, wenn sie für die Aufgabenstellung auch berücksichtigt werden muss. Lässt sich eine bestimmte Funktion nur aufrufen, wenn ein bestimmtes Bedienelement für drei Sekunden gedrückt gehalten wird, dann hat diese Mehrfachbelegung nur einen Einfluss auf die Steuerungsschwierigkeit, wenn diese Funktion aufgerufen werden muss; gleiches gilt für eventuell vorhandene Variablenbedingungen in dem technischen System.

Soll die Transparenz in einer konkreten Situation unabhängig von der Transparenz des technischen Systems erhöht werden, können Hilfestellungen zur Verfügung gestellt werden, wie etwa Bedienungsanleitungen oder FAQ-Seiten (engl. frequently asked questions »häufig gestellte Fragen«) auf Herstellerwebseiten. Das Merkmal Dynamik kann für die konkrete Aufgabenstellung als Situationsmerkmal nicht präzisiert werden, da sich das Steuerungsziel während der Bedienung eines technischen Systems zwar prinzipiell ändern könnte, für die Umsetzung dessen in einem Testinstrument wird für die Arbeit kein diagnostischer Mehrwert gesehen. Tabelle 6.2 fasst die Ergebnisse der Konkretisierung der Merkmale Komplexität, Vernetztheit, Dynamik und Intransparenz nach den beiden Interaktionsformen zusammen.

Tabelle 6.2: Ergebnis der Präzisierung schwierigkeitsbestimmender Merkmale für System- und Situationsmerkmale

	Komplexität	Vernetztheit	Dynamik	Transparenz
Systemmerkmale				
<i>Hardware</i>	Anzahl Bedienelemente	Mehrfachbelegte Bedienelemente	nicht vorhanden	Beschriftung, Rückkopplung
<i>Software</i>	Menütiefe und –breite, Funktions- und Variablenanzahl	Menüvernetzung, Reversibilität, Variablenbedingungen	Zeitverzögerungen, Eigendynamiken	Menüverständlichkeit, Rückkopplung
Situationsmerkmale				
<i>Problemstellung</i>	Anzahl Lösungsschritte, Anzahl einzustellender Variablen	Zu berücksichtigende Mehrfachbelegungen und Bedingungen	nicht vorhanden	Hilfsmittel

6.5 Indikatoren technischer Problemlösekompetenz

Kompetenzen zeigen sich in der Bewältigung von Anforderungen, die in einer Situation an eine Person gestellt werden. Für die Konzeptualisierung des in dieser Arbeit interessierenden Konstrukts *Problemlösekompetenz im Umgang mit technischen Alltagsgeräten* haben die vorherigen Abschnitte die besondere Situation sowie die in ihr entstehenden Anforderungen beschrieben. In diesem Abschnitt sollen die Möglichkeiten aufgezeigt werden, ein Testinstrument so zu entwickeln, dass es Personen auffordert, sich ihrer Problemlösekompetenz entsprechend zu verhalten, um aus dem Verhalten dann entsprechende Skalen zu bilden.

Um mit einem technischen Alltagsgerät versiert umgehen zu können, wird lediglich *Eingriffswissen* in Form von Prozeduren (Wenn-Dann-Regeln) benötigt. Für die Bedienung ist es hingegen nicht erforderlich, über Kenntnisse der Wirk- oder Baustruktur eines Systems zu verfügen. Fehlt allerdings auch das Eingriffswissen, weil einer Person ein vorliegendes technisches Gerät nicht bekannt ist, muss sie diese Wissenslücke zunächst schließen. Stehen keine Bedienungsanleitungen oder andere Hilfsmittel zur Verfügung, lassen sich

Wissensdefizite nur durch den probierenden Umgang mit dem technischen Gerät ausgleichen. Hierzu ist im soziotechnischen Sinn die Kommunikation bzw. die Interaktion mit dem technischen System notwendig. Durch gezielte Eingriffe in das System erhält die Person Rückmeldungen, die sie wahrnehmen und interpretieren muss, um weitere Handlungen vornehmen zu können. Mithilfe der erfolgten Rückmeldungen kann eine Person Eingriffswissen (Prozeduren) erwerben, das sie dann für die gezielte Steuerung des Systems einsetzen kann. Je nachdem, ob eine Person den Wissenserwerb oder die Wissensanwendung zum Ziel hat, unterscheidet sich ihre Interaktion mit dem technischen System nach den oben genannten Formen *Exploration* und *Steuerung*. Ein zu entwickelndes Testinstrument sollte demnach auch in der Lage sein, diese Interaktionsformen voneinander zu unterscheiden.

Voraussetzung hierfür sind Items, anhand derer sich das interaktive Verhalten von Personen mit technischen Systemen beobachten lässt. Die Eignung von Computersimulationen für dynamisches Problemlösen und damit für den Umgang mit technischen Systemen wurde bereits in Kapitel 4 hervorgehoben. Ein computerbasierter Test stellt eine kostengünstige Alternative zu einem Test mit realen technischen Geräten dar. Bedingung dafür, dass aus dem Verhalten der Probanden in dem computerbasierten Test Rückschlüsse auf das Verhalten der Probanden im Umgang mit realen technischen Geräten gezogen werden darf, ist, dass die Simulationen der technischen Systeme sowohl im Aussehen als auch im Verhalten möglichst realitätsnah gestaltet sind. Um den Aufwand, der durch die detaillierte Umsetzung in der Programmierung und der grafischen Darstellung entsteht, verhältnismäßig gering zu halten, sind lediglich die Hard- und Softwarekomponenten der Benutzungsschnittstellen zu simulieren, die für die alltägliche Nutzung relevant sind. Bauteile innerhalb des Gerätegehäuses und Anschlüsse auf den Gehäuserückseiten, die eher selten genutzt werden, sollen nicht nachgebildet werden.

Die zu simulierenden Alltagsgeräte sollen sich dabei hinsichtlich der Schwierigkeit sie zu bedienen unterscheiden. Für den *explorierenden* Umgang müssen die oben als potenziell schwierigkeitsbestimmend genannten *Systemmerkmale* in den Geräten unterschiedlich stark ausgeprägt sein. Für den *steuernden* Umgang ist die Ausprägung der *Situationsmerkmale* entsprechend zu variieren. Um aus dem Verhalten, das Personen im Umgang mit den technischen Systemen zeigen, ein Maß für ihre Kompetenz abzuleiten, ist entsprechend dem Interaktionsziel eine Skalierung vorzunehmen. Darunter wird die Verrechnungsvorschrift eines Tests verstanden. Diese sollte dazu führen, dass sich Kompetenzunterschiede zwischen Probanden auch in entsprechenden Abständen zwischen den Punktwerten des Tests abbilden (vgl. Pospeschill, 2010, S. 29). Im Folgenden werden getrennt nach den Interaktionsformen Systemexploration und Systemsteuerung geeignete Skalierungen besprochen.

6.5.1 Systemexploration

Wenn das Ziel einer Person darin besteht, Wissen über die Bedienung eines ihr noch unbekanntes technischen Gerätes zu erwerben, so muss die Interaktion mit dem System danach ausgerichtet sein. Die Eingriffe sollten dabei so vorgenommen werden, dass möglichst viele Informationen über vorhandene Systemzustände und Variablen generiert werden. Je mehr Zustandsübergänge im explorierenden Umgang mit dem technischen System kennengelernt werden, desto größer sollte das erworbene Eingriffswissen ausfallen. Als Indikatoren kommen computersimulierte technische Systeme in Frage, mit denen die

Probanden derart interagieren sollen, dass sie möglichst viel Eingriffswissen erwerben. Die technische Problemlösekompetenz einer Person sollte sich darin zeigen, dass sie in der Lage ist, derart mit dem System umzugehen, dass sie möglichst viele Informationen generieren und so Wissen erwerben kann. Für die Bildung eines numerischen Messwertes, der dem beobachteten Verhalten zugeordnet werden kann, gibt es verschiedene Möglichkeiten, die sich anhand eines resultatorientierten und eines prozessorientierten Ansatzes differenzieren lassen und bereits in Kapitel 3.2 vorgestellt wurden. Ein resultatorientierter Testwert entsteht, wenn den Probanden nach der Systemexploration Aufgaben vorgelegt werden, in denen sie Systemzustände identifizieren oder verifizieren müssen und die Richtiglösungen anschließend zu einem Wert summiert werden. Schwachpunkte dieses Ansatzes sind jedoch die vorhandene Ratewahrscheinlichkeit bei Unsicherheit oder Nichtwissen sowie die zusätzliche Testzeit, die durch die weiteren zu beantwortenden Items benötigt wird. Je nach Komplexität des explorierten Systems kann die Anzahl der Aufgaben zur Wissensdiagnose die erforderliche Testzeit bedeutend erhöhen und es bestünde die Gefahr, dass der Test nicht mehr ökonomisch ist.

Denkbar ist auch die Bildung eines Messwertes für die *Kompetenz im explorierenden Umgang mit technischen Systemen* über den prozessorientierten Ansatz, bei dem nicht das Ergebnis, sondern der Prozess der Exploration bewertet wird. Die Prozessanalyse muss dabei den Fokus auf die Menge der generierten Informationen legen und nicht auf die Explorationszeit oder die Anzahl der Interaktionen in dieser Zeit, denn anhand der beiden letztgenannten Variablen kann nicht auf das Verhalten in der Exploration geschlossen werden. Wenn die in einem Test einzusetzenden technischen Systeme mithilfe dem in der UML (Unified Modeling Language) verallgemeinerten Konzept finiter Automaten¹ modelliert werden, so ist die aus einem System generierbare Informationsmenge eindeutig festgelegt, durch die Anzahl der Zustände in einem System, die Anzahl veränderbarer Variablen sowie die Ereignisse, die Zustandsübergänge oder Variablenänderungen herbeiführen. Das Verhältnis zwischen der, von einer Person während des Explorierens, generierten Informationsmenge und der Gesamtzahl an Informationen, die aus einem System maximal generiert werden können, stellt ein Maß für das Explorationsverhalten dar. Dieses Verhältnis sollte für die Exploration von Systemen, deren schwierigkeitsbestimmende Merkmale schwach ausgeprägt sind, höher sein, als für die Exploration von Systemen, bei denen die Ausprägung der Merkmale besonders stark ist. Bei Personen, die über eine hohe Kompetenz im Umgang mit technischen Alltagsgeräten verfügen, sollte das Verhältnis zwischen generierten und vorhandenen Informationen ebenfalls höher sein, als bei Personen mit einer geringen Kompetenz. Die Verhältnisse, die im Folgenden als Explorationsvollständigkeits bezeichnet werden, sind intervallskaliert und können für jede Testperson aufsummiert werden und stellen damit ein Testwert dar, der als numerische Zahl die Kompetenzausprägung der Person widerspiegelt.

6.5.2 Systemsteuerung

Soll mit der Bedienung eines technischen Systems ein konkretes Anwendungsziel erreicht werden, kann die Menge an Informationen, die jemand in der Interaktion generiert, kein Maß für kompetentes Verhalten in dieser Situation sein. Für eine optimale Zielerreichung ist

¹ Finite Automaten als formale Modelle zur Erstellung von computersimulierten technischen Geräten sowie die hierfür verwendete UML werden in Kapitel 4.3 näher beschrieben.

vielmehr eine feste Folge von Eingriffen nötig, die die nutzende Person vorab kennengelernt haben und an die sie sich erinnern muss. Der Fokus bei der Steuerung eines technischen Systems liegt auf der Wissensanwendung und nicht wie bei der Exploration auf dem Erwerb von Wissen. Kompetenzunterschiede spiegeln sich in dem Erfolg der Wissensanwendung wider, sodass der Erfolg bei der Erreichung der vorgegebenen Ziele ein wichtiges resultatorientiertes Maß für die Systemsteuerung darstellt. Die Bewertung des Erfolges kann dabei auf unterschiedlichen Wegen erfolgen und ist grundsätzlich von der Zielvorgabe abhängig. Zielvorgaben sind im Umgang mit technischen Alltagsgeräten in der Regel sehr präzise formuliert, beispielsweise soll ein konkreter Systemzustand erreicht bzw. eine oder mehrere Systemvariable(n) verändert werden. Eine Zielvorgabe im Umgang mit einem Backofen kann sein: Stelle die Heizart des Backofens auf Ober-/Unterhitze und die Temperatur auf 210°C ein. Der Erfolg bei der Steuerung des technischen Systems lässt sich anhand des Wertes der Systemvariablen überprüfen. Weist die Systemvariable Heizart den Wert 2 = Ober-/Unterhitze und die Systemvariable Temperatur den Wert 210 = 210°C auf, so wurde die Aufgabe erfolgreich bewältigt. Die Zielerreichung kann als geeignetes Maß für die Kompetenz bei der Steuerung eines technischen Systems gesehen werden, da die Wahrscheinlichkeit bei der Systemsteuerung erfolgreich zu sein, mit zunehmender Kompetenz ansteigt. Kompetentes Verhalten sollte sich nicht nur in der Quantität erfolgreich bewältigter Aufgaben zeigen, sondern zudem auch in der Qualität des Bewältigungsprozesses. Als kompetent zu bezeichnende Nutzer steuern ein technisches System zielgerichteter, das heißt auf direktem Weg. Dieser direkte Weg zeichnet sich dadurch aus, dass die Anzahl der benötigten Eingriffe geringer ist, als für alle anderen Wege. Auf Basis der in Zustandsdiagrammen modellierten Systeme lässt sich vorab die optimale, sprich kürzeste Eingriffsfolge angeben. Je mehr eine Person von dem direkten Weg abweicht, desto größer wird die Anzahl an Interaktionen, die sie für die Steuerung eines Systems benötigt.

6.6 Formulierung eines theoretischen Kausalmodells

In den vorherigen Abschnitten des Kapitels 6 wurden bereits erste Schritte zur Spezifizierung des Konstruktes durchgeführt. Zum einen erfolgte die Beschreibung des Kontextes, in dem die Kompetenz zum Problemlösen im Umgang mit technischen Systemen erforderlich ist und zum anderen die Erarbeitung unterschiedlicher Konstruktdimensionen, anhand derer sich kompetentes Problemlöseverhalten unterscheidet. Für die weitere Bestimmung und Eingrenzung des Konstrukts sind außerdem Anknüpfungspunkte und Überlappungen zu anderen Konstrukten zu betrachten. Eine solche Betrachtung soll in diesem Abschnitt vorgenommen werden.

Ganz allgemein beschreibt ein Kompetenzkonstrukt das Potenzial zur Bewältigung von spezifischen Anforderungen in einer konkreten Situation. Auf die Problemlösekompetenz im Umgang mit technischen Alltagsgeräten bezogen, beschreibt das Konstrukt das Potenzial Anforderungen, die sich aus der Interaktion mit technischen Alltagsgeräten ergeben, bewältigen zu können, wobei das Ziel der Interaktion einmal im Wissenserwerb (Exploration) und einmal in der Wissensanwendung (Steuerung) liegen kann. Unter dem Begriff Potenzial werden alle individuellen Voraussetzungen subsummiert, die zur Bewältigung der Anforderungen erforderlich sind. Dazu gehören generelle und spezifische (kognitive) Fähigkeiten, Einstellungen, Erfahrungen und motivationale Aspekte. Weil eine Kompetenz

keine Einzelfähigkeit darstellt, ist die Überschneidung zu anderen Konstrukten bereits theoretisch vorgesehen. Es wird jedoch nicht möglich sein, alle Einzelfaktoren anzugeben, die für die Bewältigung einer Situation erforderlich sind, so dass es immer einen, nicht durch andere Konstrukte erklärbaren Anteil gibt. Für die Einordnung der zu messenden Problemlösekompetenz in ein nomologisches Netzwerk sollten aber mindestens diejenigen Konstrukte herangezogen werden, von denen auf Grundlage theoretischer und empirischer Vorarbeiten ein Einfluss angenommen werden kann. Hierzu enthält der Theorieteil dieser Arbeit bereits Informationen über Konstrukte, die mit dem zu messenden Konstrukt unterschiedlich große Überschneidungen aufweisen. Abbildung 6.3 gibt eine vereinfachte Darstellung über den Zusammenhang zwischen technischer Problemlösekompetenz und den Konstrukten Intelligenz, allgemeines (komplexes) Problemlösen, Selbstwirksamkeitserwartungen, Need for Cognition und Vorwissen.

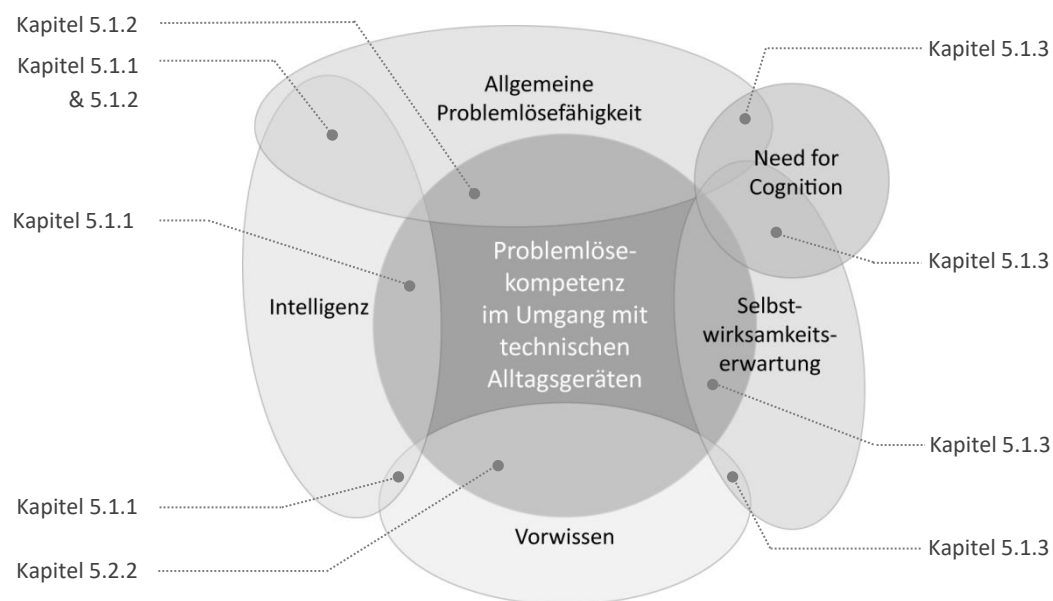


Abbildung 6.3: Verortung von Problemlösekompetenz im Umgang mit technischen Alltagssystemen in einem nomologischen Netzwerk

Um das Konstrukt der technischen Problemlösekompetenz weiter spezifizieren zu können, sollen die als unterschiedlich große Überlappungen in dem nomologischen Netzwerk dargestellten Zusammenhänge präzisiert und in einem Kausalmodell formalisiert werden. Hierzu werden die wichtigsten Aspekte aus dem Forschungsstand noch einmal zusammengefasst und die Richtung der angenommenen Kausalbeziehungen formuliert.

Fluide Intelligenz, als grundlegende Voraussetzung zum Wissenserwerb, bestimmt wie gut Anforderungen in neuen Situationen ohne Vorwissen bewältigt werden können. Damit gilt fluide Intelligenz als ein starker Prädiktor für den Erfolg beim Lösen von Problemen, aber auch als Prädiktor für bereits erworbenes Wissen (Kapitel 5.1.1). In Problemsituationen, die durch ihren semantischen Kontext bereits vorhandenes Wissen bei der problemlösenden Person aktivieren, fällt der Zusammenhang zwischen Intelligenz und Problemlöseleistungen geringer aus (geringere Überschneidung in Abbildung 6.3). Deshalb wird hier angenommen, dass sich für den Zusammenhang zwischen Intelligenz und allgemeiner (komplexer) Problemlösefähigkeit ein größeres Maß ergibt (größere Überschneidung in Abbildung 6.3), als

für den Zusammenhang zwischen Intelligenz und der Problemlösekompetenz im Umgang mit technischen Geräten. Betrachtet man Problemlösen auf der Ebene der Interaktionsziele differenziert nach Wissenserwerb und Wissensanwendung, so sollte fluide Intelligenz einen größeren Einfluss auf den Erwerb von Wissen haben als auf die Anwendung. Das gilt sowohl für allgemeines Problemlösen als auch für technisches Problemlösen.

Eine weitere kausale Beziehung wird zwischen allgemeiner Problemlösefähigkeit und technischer Problemlösekompetenz angenommen. Eine allgemeine Problemlösefähigkeit beschreibt über Problemsituationen hinweg generalisierbare Prozesse, Heuristiken oder Strategien. Hierzu zählen Prozesse des Repräsentierens, Suchens, Planes und Entscheidens, die sich aus Erfahrungen im Umgang mit verschiedenartigen Problemen im Laufe der Zeit zu allgemeinen Denk- und Problemlösestrategien entwickelt haben (Kapitel 5.1.2). Da angenommen wird, dass diese Fähigkeit zur Lösung aller bzw. einer großen Anzahl an Problemen eingesetzt werden kann, sollte diese Fähigkeit auch in Problemen im Umgang mit technischen Alltagsgeräten sichtbar werden.

Neben den eher allgemeinen kognitiven Fähigkeiten spielt im Umgang mit technischen Systemen außerdem vorhandenes Vorwissen eine Rolle. Zwar können auch ohne eine bereits vorhandene Wissensstruktur neue Informationen generiert werden, sie zu verarbeiten und zu speichern ist allerdings erst durch eine Verknüpfung mit bereits vorhandenem Wissen möglich. Dabei lenkt vorhandenes Wissen (teils unbewusst) die Aufmerksamkeit auf bestimmte Bereiche und beeinflusst folglich, welche Informationen wahrgenommen werden (Kapitel 5.2.2). Anhand bestehender Vorstellungen und Kenntnisse ist es der problemlösenden Person möglich, systematische Explorationsstrategien anzuwenden, die zu einem höheren Wissenserwerb führen, als fehleranfällige unsystematische Strategien, von denen Personen ohne Vorwissen Gebrauch machen. Ein positiver Zusammenhang zwischen dem Vorwissen und der Wissenserwerbsphase im Problemlöseprozess ist nur dann anzunehmen, wenn die zu explorierenden technischen Systeme nicht im Widerspruch zum Vorwissen stehen. Da im Laufe des Problemlöseprozesses neues spezifisches Systemwissen erworben wird, verliert das anfängliche Vorwissen zunehmend an Bedeutung und für die Leistung in der Systemsteuerung ist das unmittelbar vorher erworbene Wissen der beste Prädiktor. Anzunehmen ist deshalb nur ein geringer Zusammenhang zwischen Vorwissen und der Systemsteuerung.

Zu den beeinflussenden Faktoren einer Kompetenz zählen neben den bereits betrachteten kognitiven Fähigkeiten auch überdauernde motivationale und affektive Aspekte. Das Konstrukt Need for Cognition, als die Tendenz einer Person sich mit dem Denken zu beschäftigen und Spaß daran zu haben, sollte sich in der Motivation zum Lösen von Problemen zeigen, denn Problemlösen ist eine Tätigkeit, die zum größten Teil Denken erfordert (Kapitel 5.1.3). Das intrinsisch motivierte Verhalten von Personen mit einem hohen Need for Cognition sollte sich durch eine ausdauernde und intensive Bearbeitung der gegebenen Problemstellungen beobachten lassen. Es wird dabei ein positiver Einfluss von Need for Cognition sowohl auf allgemeines als auch auf technisches Problemlösen erwartet.

Angenommen werden kann außerdem ein (ungerichteter) Zusammenhang zwischen Need for Cognition und den Selbstwirksamkeitserwartungen, denn das Konstrukt Need for Cognition kann nur durch die subjektive Einschätzung der Personen erfasst werden (Kapitel 5.1.3).

Eigene Einschätzungen werden dabei von Selbstwirksamkeitserwartungen beeinflusst. Diese beziehen sich auf die Einschätzung der eigenen Fähigkeit Anforderungen in neuen Situationen erfolgreich bewältigen zu können. Eine bezüglich des Problemlösens hohe Selbstwirksamkeitserwartung führt in der entsprechenden Situation ebenso wie ein hohes Need for Cognition zu motiviertem Verhalten und damit zu einer höheren Leistung im allgemeinen und technischen Problemlösen (Kapitel 5.1.3). Da die Quelle einer geringen Selbstwirksamkeit unter anderem in bisherigen Misserfolgen aufgrund fehlenden Wissens liegen kann, ist ein Zusammenhang zwischen Selbstwirksamkeitserwartungen und dem Vorwissen denkbar.

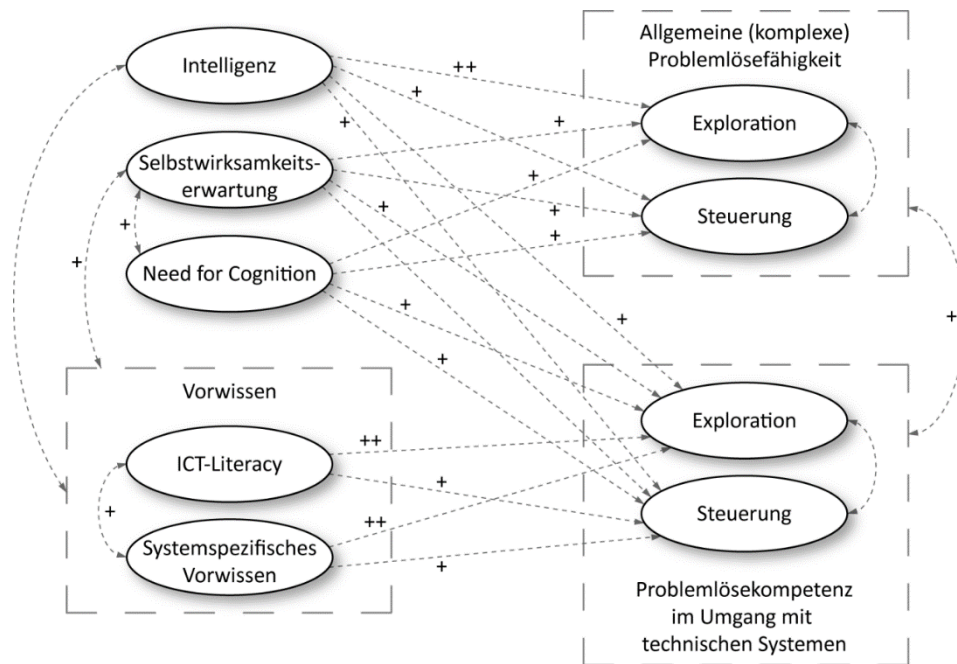


Abbildung 6.4: Theoretisches Kausalmodell der Problemlösekompetenz im Umgang mit technischen Alltagsgeräten

Diese Wirkungsbeziehungen zwischen den einzelnen Konstrukten bilden das in Abbildung 6.4 dargestellte theoretische Kausalmodell. Die in dem Modell eingezeichneten Kausalbeziehungen werden in dem nächsten Kapitel in Form von Hypothesen formuliert, mit denen sich die Gültigkeit der angenommenen Theoriestructur überprüfen lässt.

7 Forschungsfragen und Hypothesen

Ziel der vorliegenden Arbeit ist die Konzeptualisierung und Präzisierung des hypothetischen Konstrukts einer *Problemlösekompetenz im Umgang mit technischen Alltagsgeräten* sowie dessen empirische Überprüfung. In dem vorherigen Abschnitt wurde dieses Konstrukt bereits beschrieben und als mögliche Interaktionsformen der explorierende und der steuernde Umgang mit technischen Systemen ausgemacht. Theoretische Annahmen, die bereits getroffen wurden, beziehen sich zum einen auf die Anforderungen in den jeweiligen Problemsituationen sowie auf die Verortung des Konstruktes in einem nomologischen Netzwerk mit anderen Konstrukten. Die empirische Überprüfung des postulierten Kompetenzkonstruktes basiert auf der Prüfung der getroffenen Annahmen und ist Aufgabe der Konstruktvalidierung. Hierzu werden aus dem formulierten theoretischen Konstrukt Hypothesen abgeleitet, die in dem weiteren Fortgang der Arbeit geprüft werden. Die Aufstellung der Hypothesen erfolgt dabei getrennt nach dem zu validierenden Konstruktaspekt¹, der jeweils in den folgenden drei Fragestellungen fokussiert wird.

Fragestellung 1: Beeinflussen die angenommenen schwierigkeitsbestimmenden System- und Situationsmerkmale tatsächlich die Schwierigkeit im Umgang mit technischen Systemen?

H_{1.1}: Je größer die Komplexität eines technischen Systems, desto schwieriger ist es zu explorieren.

H_{1.2}: Je stärker die Vernetzung eines technischen Systems, desto schwieriger ist es zu explorieren.

H_{1.3}: Je mehr Eigendynamiken in einem technischen System vorhanden sind, desto schwieriger ist es zu explorieren.

H_{1.4}: Je transparenter ein technisches System ist, desto einfacher ist es zu explorieren.

H_{1.5}: Je komplexer die Aufgabenstellung desto schwieriger ist die Steuerung eines technischen Systems.

H_{1.6}: Je mehr Systemvernetzungen für eine Aufgabenstellung berücksichtigt werden müssen, desto schwieriger ist die Steuerung eines technischen Systems.

Fragestellung 2: Lässt sich die logisch-inhaltliche Analyse des Konstruktes (Konstruktstruktur) empirisch bestätigen?

H_{2.1}: Explorieren und Steuern sind zwei voneinander abgrenzbare Problemlösebereiche im Umgang mit technischen Alltagsgeräten.

H_{2.2}: Je höher der Wert für die Explorationsleistung, desto höher der Wert für die Steuerungsleistung im Umgang mit technischen Systemen.

H_{2.3}: Allgemeines Problemlösen und Problemlösen im Umgang mit technischen Systemen sind zwei voneinander abgrenzbare Problembereiche.

¹ Eine genauere Erläuterung der drei Validitätsaspekte wird als Einführung in die methodische Beschreibung der Validitätsprüfung in Kapitel 10 gegeben.

Fragestellung 3: Lässt sich die Verortung des Konstruktes in dem nomologischen Netzwerk empirisch bestätigen?

H_{3.1}: Je intelligenter eine Person, desto höher sind ihre Leistungen im explorierenden Umgang mit technischen Systemen.

H_{3.2}: Je intelligenter eine Person, desto höher sind ihre Leistungen im steuernden Umgang mit technischen Systemen.

H_{3.3}: Je intelligenter eine Person, desto höher sind ihre Leistungen im explorierenden Umgang mit allgemeinen Problemlöseszenarien.

H_{3.4}: Je intelligenter eine Person, desto höher sind ihre Leistungen im steuernden Umgang mit allgemeinen Problemlöseszenarien.

H_{3.5}: Die Intelligenz einer Person hat einen größeren Einfluss auf den explorierenden und steuernden Umgang mit allgemeinen Problemlöseszenarien als auf den explorierenden und steuernden Umgang mit technischen Alltagsgeräten.

H_{3.6}: Je intelligenter eine Person, desto höher ist ihr Vorwissen.

H_{3.7}: Je höher die Selbstwirksamkeitserwartungen einer Person, desto höher sind ihre Leistungen im explorierenden Umgang mit technischen Systemen.

H_{3.8}: Je höher die Selbstwirksamkeitserwartungen einer Person, desto höher sind ihre Leistungen im steuernden Umgang mit technischen Systemen.

H_{3.9}: Je höher die Selbstwirksamkeitserwartungen einer Person, desto höher sind ihre Leistungen im explorierenden Umgang mit allgemeinen Problemlöseszenarien.

H_{3.10}: Je höher die Selbstwirksamkeitserwartungen einer Person, desto höher sind ihre Leistungen im steuernden Umgang mit allgemeinen Problemlöseszenarien.

H_{3.11}: Die Selbstwirksamkeitserwartungen einer Person und ihr Need for Cognition sind zusammenhängende Konstrukte.

H_{3.12}: Je höher das Need for Cognition einer Person, desto höher sind ihre Leistungen im explorierenden Umgang mit technischen Systemen.

H_{3.13}: Je höher das Need for Cognition einer Person, desto höher sind ihre Leistungen im steuernden Umgang mit technischen Systemen.

H_{3.14}: Je höher das Need for Cognition einer Person, desto höher sind ihre Leistungen im explorierenden Umgang mit allgemeinen Problemlöseszenarien.

H_{3.15}: Je höher das Need for Cognition einer Person, desto höher sind ihre Leistungen im steuernden Umgang mit allgemeinen Problemlöseszenarien.

H_{3.16}: Je höher das Vorwissen einer Person, desto höher sind ihre Leistungen im explorierenden Umgang mit technischen Systemen.

H_{3.17}: Je höher das Vorwissen einer Person, desto höher sind ihre Leistungen im steuernden Umgang mit technischen Systemen.

H_{3.18}: Ein hohes Vorwissen geht mit einer hohen Selbstwirksamkeitserwartung einher.

Teil III: Methodik

8 Operationalisierung technischer Problemlösekompetenz

Das theoretisch in Abschnitt 6 beschriebene Konstrukt der *Problemlösekompetenz im Umgang mit technischen Alltagsgeräten* kann nicht direkt beobachtet werden, sodass in diesem Abschnitt die Operationalisierung in ein Messmodell vorgenommen wird. Aufgabe der Operationalisierung ist die Festlegung von Indikatoren, anhand derer die Ausprägung des Konstruktes bei den Untersuchungspersonen festgestellt werden kann. Die Indikatoren selbst müssen dabei direkt zu beobachten sein und bilden die Items des Tests. Da das zu messende Konstrukt als Ursache für das Antwortverhalten einer Person auf die Indikatoren angenommen wird, müssen für das Messmodell reflektive Indikatoren gebildet werden, also Indikatoren in denen sich das Konstrukt widerspiegelt. Es eignen sich hierfür Aufgaben, die inhaltsähnliche Anforderungen an die Probanden stellen, die sich jedoch in ihrer Schwierigkeit unterscheiden sollten. Zu berücksichtigen sind bei der Operationalisierung eines Konstruktes, dass zum einen alle wesentlichen Aspekte des theoretischen Konstruktes in den Indikatoren repräsentiert und zum anderen keine konstruktirrelevanten Aspekte enthalten sind.

Im Folgenden wird die Entwicklung von Items für das zu messende Konstrukt der Problemlösekompetenz im Umgang mit technischen Alltagsgeräten beschrieben. Hierzu ist zunächst eine repräsentative Auswahl technischer Geräte aus dem Alltag zu erstellen, mit denen eine Interaktion möglich ist. Die technischen Geräte sollen sowohl ausprobiert als auch gezielt gesteuert werden können, wobei die Schwierigkeit in dem interagierenden Umgang mit den Geräten variieren muss. Dementsprechend müssen sich die zu entwickelnden Indikatoren in den bereits in Kapitel 6.4 operationalisierten System- und Situationsmerkmalen unterscheiden. Die konkrete Itemerstellung erfolgt mit dem *CBA ItemBuilder*, einer Software, die nachfolgend kurz vorgestellt wird. Um potenzielle Fehler in den entwickelten Items und Probleme bei der Testdurchführung vorab erkennen und abstellen sowie eventuelle Anpassungen vornehmen zu können, wird das Testinstrument an Gelegenheitsstichproben erprobt. Das Ergebnis dieser Erprobung sowie sich daraus ergebene Konsequenzen werden am Ende des Abschnittes berichtet.

8.1 Kontext – technische Alltagsgeräte

Der zu entwickelnde Test soll technische Geräte aus dem Alltag enthalten, die von den Probanden exploriert und gesteuert werden können. Technische Geräte, mit denen der Mensch interagieren kann beziehungsweise sogar interagieren muss, um sie nutzen zu können, finden sich im Alltag in vielfältiger Form. Technische Geräte, die aus einem gemeinsamen Bereich stammen, weisen Ähnlichkeiten hinsichtlich ihres Aussehens, ihrer Komplexität, ihres Systemverhaltens oder ihres Zwecks auf. Die folgenden Kategorien beschreiben solche Bereiche: *Selbstbedienungsautomaten*, *Hausgerätetechnik*, *Haustechnik*, *Multimedengeräte* und *Informations- und Kommunikationstechnik*. Tabelle 8.1 gibt für alle der oben genannten Bereiche Beispiele an.

Tabelle 8.1: Beispiele technischer Geräte aus dem Alltag

Selbstbedienungs- automaten	Hausgeräte- technik	Haustechnik	Multimedia- geräte	Informations- und Kommunikations- technik
<ul style="list-style-type: none"> • Briefmarkenautomat • Fahrscheinautomat • Parkscheinautomat • Fotoentwicklungs- automat • Geldautomat • Packstation • Münzwechselautomat • Getränkeautomat 	<ul style="list-style-type: none"> • Spülmaschine • Waschmaschine • Backofen • Küchenradio • Kaffeeautomat • Brotback- automat • Staubsauger- roboter • Radiowecker 	<ul style="list-style-type: none"> • Rollladen- steuerung • Heizungs- steuerung • Markisen- steuerung • Licht- steuerung 	<ul style="list-style-type: none"> • DVD-Recorder • Receiver • MP3-Player • Beamer • Multi- funktions- drucker 	<ul style="list-style-type: none"> • Internetrouter • Telefon • Navigations- gerät • Online-Banking • Mobiltelefon

Für den technischen Problemlösetest ist zunächst eine Anzahl von 18 unterschiedlichen technischen Systemen vorgesehen. Damit die zu konstruierenden Items eine möglichst repräsentative Auswahl der Menge an technischen Systemen darstellen, die im Alltag vorzufinden sind, werden aus jedem der fünf genannten Bereiche Items entwickelt. Dabei richtet sich die Anzahl der pro Bereich zu entwickelnden Items danach, wie präsent der jeweilige Bereich in dem Alltag der meisten Personen ist. Nach eigener Einschätzung sind viele der technischen Geräte aus dem Bereich der Hausgerätetechnik in den meisten Haushalten vorhanden, während technische Systeme der Haustechnik weniger weit verbreitet sind. Demnach ist es zweckmäßig eine größere Anzahl an Items aus dem Bereich der Hausgerätetechnik in den Test zu implementieren, als aus dem Bereich der Haustechnik. Diese Überlegungen spiegeln sich in der folgenden Auswahl an technischen Systemen wider, die für den technischen Problemlösetest als Items ausgearbeitet werden:

Selbstbedienungsautomaten:	<i>Fahrscheinautomat, Parkscheinautomat, Fotoentwicklungsautomat</i>
Hausgerätetechnik:	<i>Spülmaschine, Waschmaschine, Küchenradio, Kaffeeautomat, Radiowecker</i>
Haustechnik:	<i>Rollladensteuerung, Heizungssteuerung</i>
Multimediageräte:	<i>Receiver, Multifunktionsdrucker, Fotodrucker</i>
Informations- und Kommunikationstechnik:	<i>Internetrouter, Telefon, Navigationsgerät, Online-Banking, Mobiltelefon</i>

8.2 Interaktionsformen

Weil sich die Art und Weise wie Personen mit einem technischen System über die Benutzungsschnittstelle interagieren, anhand des Ziels, das sie erreichen wollen, unterscheidet, müssen diese Unterschiede in dem Testinstrument abgebildet werden. Steht der Erwerb von Eingriffswissen im Vordergrund, so wird eine Person sehr viele Eingriffe in das System vornehmen, um so viele Informationen über die Zustandsübergänge zu erhalten. Besteht das Ziel der Interaktion hingegen darin, das System in einen klar definierten Systemzustand zu bringen, müssen die zu tätigen Eingriffe in einer der Aufgabenstellung entsprechenden Reihenfolge ausgeführt werden. Für jedes der in dem Test präsentierten technischen Systeme sind die Interaktionsformen *Explorieren* und *Steuern* vorzusehen. Weil für

das zielgerichtete Steuern eines technischen Systems Eingriffswissen erforderlich ist, muss die Phase des Explorierens, in der dieses Wissen erworben werden soll, *vor* der Steuerungsphase abgeschlossen sein.

Eine hohe Problemlösekompetenz in der Handhabung technischer Systeme zeigt sich im explorierenden Umgang darin, dass die Systeme sehr ausführlich exploriert, also viele Informationen generiert werden. Das Verhalten einer Person in dieser Phase ist also danach zu bewerten, wie ausführlich die Person das jeweilige technische System exploriert. Die Aufgabenstellungen der Explorationsphase müssen hierzu so formuliert sein, dass das geforderte Handlungsziel für die Probanden transparent ist, sprich die Probanden müssen dazu aufgefordert werden, die technischen Systeme sorgfältig zu erkunden, um möglichst viel Eingriffswissen zu erwerben. Damit eine Person ein technisches Gerät in der anschließenden Steuerungsphase zielgerichtet bedienen kann, müssen die Aufgabenstellungen ein jeweils zu erreichendes Ziel nennen. Die Kompetenz einer Person im steuernden Umgang mit den technischen Systemen zeigt sich darin, wie erfolgreich sie bei der Steuerung ist. Der Erfolg bei der Systemsteuerung lässt sich danach beurteilen, ob eine Person das vorgegebene Ziel erreicht hat und wie viele Eingriffe sie für die Zielerreichung vornehmen musste. Die Zielvorgaben sind dabei so präzise zu formulieren, dass eine objektive Bewertung möglich ist.

8.3 Schwierigkeitsbestimmende Merkmale

Damit die zu entwickelnden technischen Geräte unterschiedlich schwierig zu bedienen sind, müssen die als schwierigkeitsbestimmend angenommenen System- und Situationsmerkmale in den einzelnen Systemen bzw. in den jeweiligen Aufgabenstellungen unterschiedlich stark ausgeprägt sein. Für den Konstruktionsprozess der einzelnen technischen Systeme ist entsprechend eine Variation der Systemkomponenten von Hard- und Software vorzusehen. Die Hardware der Geräte soll sich hinsichtlich der *Anzahl an Bedienelementen*, der *Anzahl an mehrfachbelegten Bedienelementen*, dem *Vorhandensein von Rückkopplungen* auf Bedienhandlungen sowie der *Verständlichkeit von Elementbeschriftungen* unterscheiden. Als Merkmale der Software technischer Systeme können die *Menütiefe*, die *Menübreite*, die *Anzahl an Funktionen*, die *Anzahl an Variablen*, der *Vernetzungsgrad der Menüs*, die Ausprägung der *Reversibilität*, die *Anzahl der Bedingungen* um Variablen ändern zu können, die *Anzahl an Eigendynamiken* und *Zeitverzögerungen* sowie die *Verständlichkeit von Menübeschriftungen* und *Rückmeldungen auf Bedienhandlungen* variiert werden, um unterschiedlich schwierig zu bedienende technische Systeme zu erhalten.

Die praktische Umsetzung der Systementwicklung erfolgt jedoch nicht ausschließlich auf Basis der Variationen der oben genannten Systemmerkmale, da auf diese Weise wenig realitätsnahe Systeme entstehen würden. Als Vorbild für den tatsächlichen Konstruktionsprozess dienen reale technische Systeme, deren Aussehen und Verhalten in die Entwicklung der Items miteinfließen. Die Ausprägung der oben genannten Merkmale wird berücksichtigt, indem beispielsweise ein Teil der Funktionen oder Variablen der realen technischen Systeme weggelassen, Beschriftungen verändert oder Eigendynamiken implementiert werden. Eine ausführliche Darstellung der Merkmalsausprägungen in den einzelnen Items erfolgt in Kapitel 10.1.2.

8.4 Itemkonstruktion (im CBA-Item Builder)

Nach den theoretischen Überlegungen zu der Auswahl und der Erstellung der Items, erfolgt die konkrete Umsetzung in einen computerbasierten Test. Für diese Aufgabe wird der *CBA ItemBuilder* (zunächst Version 4.14, später Version 5.02) eingesetzt, eine Software, die von dem Deutschen Institut für Internationale Pädagogische Forschung (DIPF) zusammen mit der Firma Nagarro AG (früher Softcon AG) entwickelt wurde (Rölke, 2012, S. 344). Ziel der Entwicklung war ein Autorenwerkzeug, für dessen Anwendung der Endnutzer über keine speziellen IT- oder Programmierkenntnisse verfügen muss (Greiff, Wüstenberg, Holt et al., 2013, S. 411). Anhand seiner grafischen Benutzeroberfläche lassen sich mit dem *CBA ItemBuilder* unterschiedliche teils sehr komplexe Aufgabentypen erstellen. Dazu gehören klassische Aufgabenformate wie Markierungsfelder (Mehrfachauswahl), Texteingabefelder, Optionsfelder (Multiple Choice) oder Auswahllisten (Drop-Down Menüs), aber auch neuere Formate wie das Markieren von Text, Drag and Drop Eingaben oder Image Maps (klickbare Bereiche in Bildern) (vgl. Hahnel & Hacker, 2013; Rölke, 2012).

Das in dieser Arbeit für die Überführung realer technischer Systeme in Computersimulationen verwendete Modell finiter Automaten wurde in Kooperation mit anderen Forschern ebenfalls als Funktion in den *CBA ItemBuilder* implementiert. Wie in Abschnitt 4.3 genauer beschrieben, besteht ein finiter (endlicher) Automat mindestens aus einem Startzustand, einer endlichen Menge von Systemzuständen, einer endlichen Menge von Eingaben und einer Übergangsfunktion. Die in dem verallgemeinerten Konzept finiter Automaten vorgesehenen Erweiterungen wie Zustände, die selbst Unterzustände enthalten, sowie Bedingungen für Zustandsübergänge und ausführbare Aktionen können im *CBA ItemBuilder* ebenfalls modelliert werden.

Die Erstellung der einzelnen Items erfolgt dabei in mehreren Schritten. Zu Beginn ist das prinzipielle Verhalten der technischen Systeme in einem noch eher skizzenhaften Zustandsdiagramm zu beschreiben. In dem *CBA ItemBuilder* bietet das *State Machine Diagram* hierfür die Möglichkeit der grafischen Darstellung. Das dort erstellte Zustandsdiagramm dient lediglich der Übersicht, Zustandsübergangsfunktionen (*rules*) werden aus ihm nicht automatisch generiert. Hierfür müssen im nächsten Schritt zunächst alle Systemzustände (*states*) angelegt, benannt und eventuell in eine Hierarchie gebracht werden, sowie der jeweilige Startzustand und mögliche Endzustände definiert werden. Sollen außerdem Bedingungen (*conditions*) für Zustandsübergänge (*transitions*) genutzt werden, so müssen außerdem vorab entsprechende Variablen erstellt, benannt und gegebenenfalls schon Werte festgelegt werden. Sowohl die Zustände als auch die Variablen sind im *CBA ItemBuilder* in der *State Machine* zu definieren. Nach der Festlegung der einzelnen Systemzustände erfolgt die Konkretisierung der Zustandsübergänge in den *State Machine Rules*. Wenn die Übergänge zwischen den Zuständen erst durch das Eingreifen der Probanden ausgelöst werden sollen, sind zunächst entsprechende Ereignisse (*events*) zu definieren. Die Zustandsübergänge sowie die Aktivitäten (*action*),

die beim Ein- oder Austreten ausgelöst werden sollen, sind dabei nach einer durch den *CBA ItemBuilder* vorgegebenen Syntax zu beschreiben.

Tabelle 8.2: Arten der Transition

Transition	Syntax
Start Transition	STARTZUSTAND -> NORMALZUSTAND {EVENT: CONDITION ACTION}
Äußere Transition	NORMALZUSTAND => NORMAL_ODER_ENDZUSTAND {EVENT: CONDITION ACTION}
Innere Transition	NORMALZUSTAND internal {EVENT: CONDITION ACTION}
Eingangsaktivität	NORMAL_STATE entry {ACTION}
Ausgangsaktivität	NORMAL_STATE exit {ACTION}

Die Elemente der Syntax, die in Tabelle 8.2 in Großbuchstaben genannt werden, sind durch entsprechende Benennungen oder Ausdrücke zu ersetzen. `CONDITION` beschreibt die Bedingung, die erfüllt sein muss, damit das genannte (optionale) Event den Quellzustand (Start- oder Normalzustand) in den Zielzustand (Normal- oder Endzustand) überführt. Conditions werden in Form logischer Ausdrücke formuliert, die auf den Vergleich mit einem numerischen Wert basieren (z. B. `VARIABLE>5`). Mit der (optionalen) Angabe einer `ACTION` kann beispielsweise ein Event ausgelöst (`raise(EVENT)`), eine Variable zurückgesetzt (`reset(VARIABLES)`) oder auf einen bestimmten numerischen Wert gesetzt (`set(VARIABLE, VALUE)`), ein MediaPlayer gestartet (`setMediaPlayer(ID, OPERATION)`) oder dessen Lautstärke verändert (`setMediaPlayerVolume(ID, INTEGER)`) werden.

Um aus der syntaxbasierten Beschreibung des finiten Automaten eine Benutzungsschnittstelle, mit der die Probanden interagieren können, zu erstellen, muss dieses Interface zunächst *gebaut* werden. Dazu sind Grafiken für die Bedienelemente, der Anzeigeelemente sowie des Hintergrundes anzufertigen, die denen eines realen technischen Gerätes möglichst ähneln. Aus der Palette des *CBA ItemBuilders* werden die vorher erstellten Elemente auf einer Szenarioseite so positioniert, dass sie das zu simulierende technische System abbilden. Die Bedienelemente werden mit den vorher definierten Events verknüpft. Damit wird erreicht, dass bei Betätigung eines Bedienelements, das mit ihm verknüpfte Event ausgelöst wird und der passende in den *Rules* hinterlegte Zustandsübergang erfolgt. Die Anzeigeelemente (Displays, Leuchten) werden mit Variablen verknüpft, deren Werte wiederum in einer *ValueMap* mit entsprechenden Grafiken, Text, Audios oder Videos verbunden werden. Wenn eine Variable beispielsweise durch die, in einem Zustandsübergang angegebene Action einen bestimmten Wert zugewiesen bekommt und für diesen Wert in der *ValueMap* ein bestimmtes Bild vorgesehen ist, so wird dieses Bild in dem Display, das mit dieser Variable und der *ValueMap* verbunden ist, angezeigt.

Der *CBA ItemBuilder* bietet ferner die Möglichkeit einer automatischen Bewertung der, von den Probanden vorgenommenen Itembearbeitung (*Scoring*). Pro erstelltem finiten Automaten können zu diesem Zweck mehrere Aufgaben (*tasks*) angelegt werden. Damit lassen sich die Exploration und die Steuerung eines technischen Systems in dieser Arbeit getrennt voneinander bewerten. Für eine automatische Bewertung des Testverhaltens sind in jeder Task die Richtiglösungen (*Hits*) sowie optional die Falschlösungen (*Misses*) anzugeben. Als

Hit oder Miss eines finiten Automaten kann beispielsweise der letzte besuchte Systemzustand, die Ausführung aller Events, die Ausführung aller Events in einem Zustand, der Wert einer Variable oder eine Kombination dieser Bedingungen bestimmt werden.

Zur Verrechnung der einzelnen Hits und Misses innerhalb eines Items zu einem Ergebnis gibt es in dem *CBA ItemBuilder* verschiedene Optionen. Um die Steuerung eines technischen Systems derart bewerten zu können, dass nur bei Erreichen des Steuerungsziels das Item *richtig* gelöst ist, muss das Steuerungsziel in den Hits abgebildet sein. Nur wenn alle Hits erreicht werden, ist ein Item als richtig gelöst zu bewerten. Das Ergebnis kommt in diesem Fall aus einer Mindestanzahl an Hits zustande. In den Aufgabenstellungen zur Exploration der technischen Systeme werden keine konkreten Steuerungsziele vorgegeben, sondern dazu aufgefordert, das System ausführlich zu erkunden, um es im Anschluss zielgerichtet steuern zu können. Demnach kann auch nicht das Erreichen eines Steuerungsziels bewertet werden, sondern das Ausmaß der Informationsgenerierung. Für jede neu generierte Information (erreichter Systemzustand, veränderte Variable) wird ein Hit definiert. Die Summe der erreichten Hits bezogen auf die maximal zu erreichenden Hits bildet das Ergebnis für die Explorationsphase. Sowohl das Erreichen als auch das Nichterreichen der Hits (und der Misses) werden in Logfile-Dateien geschrieben, aus denen sie nach der Testdurchführung ausgelesen und weiter analysiert werden können.

Im Anschluss an die Konstruktion der einzelnen finiten Automaten und der Implementierung der dazugehörigen Aufgabenstellungen müssen diese zu einem Gesamttest zusammengestellt und den Probanden zur Verfügung gestellt werden. Die Testzusammenstellung und Auslieferung erfolgt dabei nicht im *CBA ItemBuilder*, sondern benötigt eine passende Laufumgebung, wie *TBA Tools*, *TAO* oder die *CBA Execution Environment* (Hahnel & Hacker, 2013). Für den im Rahmen dieser Arbeit durchzuführenden Computertest wird die, ebenfalls vom DIPF zur Verfügung gestellte Execution Environment (EE) verwendet. Mit dem darin integrierten *JBoss® Server* sowie einem Browser wird 1) der Test administriert, 2) der Test durchgeführt und 3) die Testergebnisse abgerufen (ebd.). Die Testergebnisse werden in Form von *XML-Dateien* heruntergeladen. Diese können in einem Texteditor geöffnet und gelesen werden. Für statistische Auswertungen müssen die darin enthaltenen Hits entweder manuell in ein Tabellenkalkulations- oder Statistikprogramm übertragen werden oder es ist ein entsprechendes Programm zu schreiben, das diese Übertragung automatisch ausführt.

8.5 Beschreibung des Testmaterials

8.5.1 Beschreibung der einzelnen Items

Parkscheinautomat

Um das eigene Auto an einem kostenpflichtigen Parkplatz ohne Schranke abstellen zu können, muss in der Regel ein Parkschein erworben werden. Hierfür sind an den meisten Parkplätzen Automaten aufgestellt, an denen man ein solches Ticket selbstständig kaufen kann. Parkscheinautomaten umfassen meist nicht viele Funktionen. Für den Erwerb eines Parkscheins ist die Dauer anzugeben, für die man das Auto auf dem Parkplatz abstellen möchte sowie einen sich aus der Parkdauer ergebenden Betrag zu zahlen, indem Münzen in den vorhandenen Geldschlitz geworfen werden. Stimmt der bezahlte Geldbetrag mit dem zu

zahlenden Betrag überein, kann das Ticket angefordert werden. Das gedruckte Ticket wird von dem Automaten durch einen Schlitz ausgegeben und der Automat ist bereit für den nächsten Ticketkauf. In der erstellten Computersimulation kann über die Tasten plus und minus die Parkdauer geändert werden (Abbildung 8.1). Über die rote Taste (mit einem Kreuz symbolisiert) kann der aktuelle Vorgang abgebrochen werden und der Parkscheinautomat gibt bereits eingeworfenes Geld wieder zurück und setzt die Variablen für die Parkdauer, den zu zahlenden Betrag sowie dem eingeworfenen Betrag auf ihren Ursprungswert zurück. Über die grüne Taste (mit einem Häkchen symbolisiert) kann der Parkschein angefordert werden, der bei erfolgreicher Bedienung des Parkscheinautomaten über den unteren Schlitz ausgegeben wird. Ein zu zahlender Betrag kann mit den rechts neben dem Parkscheinautomaten abgebildeten Geldmünzen über den Geldschlitz eingeworfen werden. Dazu muss zunächst eine Münze durch Anklicken gewählt werden und anschließend durch Anklicken des Geldschlitzes eingeworfen werden.

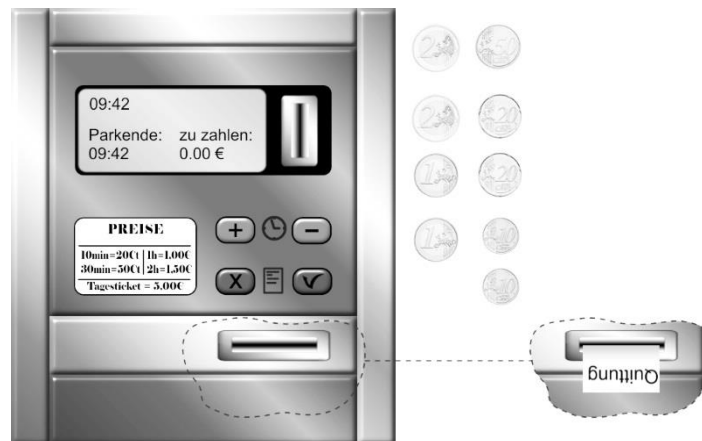


Abbildung 8.1: simuliertes Interface eines Parkscheinautomaten

Die Aufgabenstellungen für die Exploration und die Steuerung des Parkscheinautomaten lauten wie folgt:

Exploration: Sie haben Ihr Auto auf einem kostenpflichtigen Parkplatz abgestellt. Für das Abstellen Ihres Fahrzeugs müssen Sie ein Parkticket aus dem Parkscheinautomaten ziehen. Erkunden Sie, wie sich aus dem Parkscheinautomat ein Parkticket entnehmen lässt.

Steuerung: Kaufen Sie an dem Parkscheinautomat ein Ticket, mit dem Sie bis 15:30 Uhr auf dem kostenpflichtigen Parkplatz mit Ihrem Auto parken können.

Receiver

Für den Empfang und die Dekodierung digital übertragener Fernsehdienste wird ein Digitalreceiver benötigt, der oftmals bereits in den Fernseher integriert ist. Neben dieser Grundfunktion bietet ein Digitalreceiver die Möglichkeit Fernsehprogramme zu suchen und bereits vorhandenen Programmen hinzuzufügen. Programme können außerdem nach den Vorlieben der nutzenden Person in eine bestimmte Reihenfolge gebracht und als Favoritenliste abgespeichert werden. Damit sich Fernsehsendungen digital aufzeichnen lassen, ist es zudem erforderlich Angaben über Zeitzone und Sommerzeit in der Software ändern zu können. Änderungen können üblicherweise auch an Bild- und Toneinstellungen durchgeführt werden.

Die Bedienelemente, mit denen eine Person die gewünschten Einstellungsänderungen an der Software eines Receivers vornimmt, befinden sich auf einer Fernbedienung.

In der erzeugten Computersimulation sind sowohl die Fernbedienung mit den entsprechenden Bedienelementen sowie das Fernsehgerät, das als Anzeigeelement für den Receiver dient, zu sehen (Abbildung 8.2). Zu Beginn ist das Hauptmenü der Receiversoftware auf dem Bildschirm abgebildet. Mit den Pfeiltasten *oben* und *unten* lässt sich einer der Menüpunkte auswählen und durch Drücken der *OK*-Taste auf der Fernbedienung gelangt der Proband in das entsprechende Untermenü. In den Untermenüs können entweder Informationen generiert, Änderungen an Einstellungen vorgenommen oder eine tiefer liegende Menüebene gewählt werden. Aus einem Untermenü erreicht der Proband eine höhere Menüebene durch das Anklicken der *exit*-Taste auf der Fernbedienung. Einstellungsoptionen können mit den Pfeiltasten *rechts* und *links* gewählt werden. Die *gelbe* und die *rote* Taste, die sich rechts und links neben der Pfeiltaste nach unten befinden, haben nur für die Bearbeitung von Favoritenlisten eine Bedeutung. Nachdem Programme in eine gewünschte Reihenfolge innerhalb einer Favoritenliste gebracht wurden, lassen sich einzelne Einträge der Liste mit der roten Taste löschen und die Löschung aller Einträge einer Favoritenliste wird durch das Anklicken der gelben Taste herbeigeführt. Die mit den Ziffern von 0-9 beschrifteten Bedienelemente der Fernbedienung für den Receiver können genutzt werden, um den im Gerät voreingestellten Jugendschutz-PIN zu ändern.

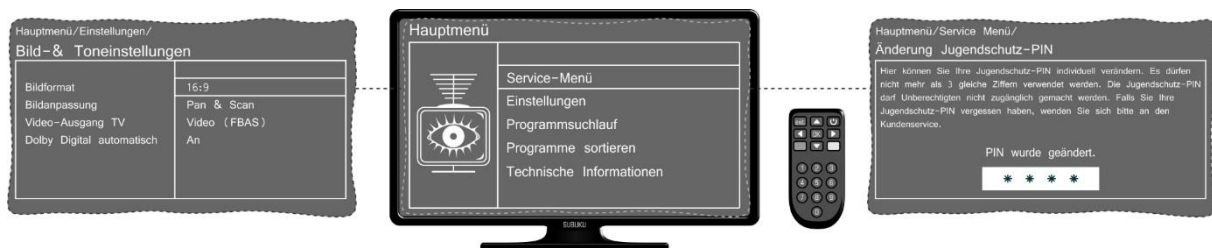


Abbildung 8.2: simuliertes Interface eines Receivers

Die Aufgabenstellungen für die Exploration und die Steuerung des Receivers lauten wie folgt:

Exploration: Der abgebildete Fernsehbildschirm zeigt die Softwareoberfläche eines integrierten Receivers. Kundschaften Sie die vielfältigen Einstellungsmöglichkeiten der Software aus. Die Bedienung dieser kann ausschließlich über die Fernbedienung erfolgen.

Steuerung: Erstellen Sie eine Favoritenliste, die die folgenden TV-Programme in dieser Reihenfolge enthält: ARD, ZDF, WDR, RTL, SAT1, Pro7

Radiowecker

Um zu einer zuvor einzustellenden Uhrzeit aus dem Schlaf geweckt oder an einen bestimmten Termin erinnert zu werden, werden Wecker benutzt. Radiowecker unterscheiden sich von herkömmlichen Weckern durch einen integrierten Radioempfänger, der zum Wecken aktiviert wird. Zum Einschlafen verfügen Radiowecker oft zusätzlich über eine sogenannte Schlaffunktion (Sleeptimer), durch die das Radio nach einer gewünschten Zeit automatisch ausgeschaltet wird. Die benutzende Person kann die aktuelle Uhrzeit, die Zeit, zu der das

Radio eingeschaltet wird, den Radiosender, die Lautstärke des empfangenen Radiosenders sowie die aktuelle Funktion des Radioweckers einstellen.

Das simulierte Interface eines Radioweckers (Abbildung 8.3) sieht hierfür verschiedene Bedienelemente vor. Oben links befindet sich der Funktionswahlschalter, der je nach Stellung zwischen dem Radiobetrieb (*ON*), dem ausgeschalteten Zustand (*OFF*) und dem Alarmbetrieb (*AUTO*) unterscheidet. Die Stellung des Schalters kann in der Simulation geändert werden, indem die gewünschte Stellung des Schalters angeklickt wird. Wenn der Schalter auf *ON* steht, hören die Probanden den jeweils eingestellten Radiosender. Soll dieser Sender geändert werden, müssen die Pfeiltasten *rechts* und *links*, die sich ebenfalls oben befinden und mit *STATION* beschriftet sind, angeklickt werden. In der Simulation können die Probanden zwischen elf verschiedenen Radiosendern wählen. Die Lautstärke kann durch Anklicken der *Plus* und *Minus* Tasten oben rechts (mit *VOLUME* beschriftet) angepasst werden. Sowohl der Sender als auch die Lautstärke können nur im Radiobetrieb (Funktionsschalter auf *ON* stehend) gewählt werden. Soll die Uhrzeit verändert werden, müssen die Probanden im *OFF*-Modus des Gerätes zunächst auf die Taste *TIME SET* drücken und können dann über die Taste *MIN* die Minuten und über die Taste *HOURL* die Stunden einstellen. Die Änderungen werden gespeichert, wenn anschließend noch einmal die Taste *TIME SET* angeklickt wird. Zur Veränderung der Alarmzeit müssen die Probanden analog mit der Taste *ALARM SET* verfahren. Damit der Wecker den gewählten Sender zu der voreingestellten Alarmzeit aktiviert, muss der Funktionsschalter auf *AUTO* gestellt werden; in dem Display erscheint dann ein kleiner Punkt. Soll die Sleep-Funktion eingeschaltet werden, muss der Funktionsschalter zunächst auf *OFF* gestellt werden und anschließend die Taste *SLEEP* gedrückt werden. Der eingestellte Radiosender wird abgespielt und im Display wird 00:59 als Hinweis angezeigt, dass sich das Radio in 59 Minuten automatisch abschaltet. Die Sleep Funktion kann vorzeitig durch Drücken der *SLEEP*-Taste beendet werden.



Abbildung 8.3: simulierte Interface eines Radioweckers

Die Aufgabenstellungen für die Exploration und die Steuerung des Radioweckers lauten wie folgt:

Exploration: Bringen Sie in Erfahrung, wie Einstellungen an dem abgebildeten Radiowecker vorgenommen werden können, damit er zu vorgegebenen Zeiten einen entsprechenden Radiosender abspielt.

Steuerung: Stellen Sie die Uhrzeit des Radioweckers auf 09:49 Uhr ein und die Weckzeit auf 07:00 Uhr. Stellen Sie den Radiowecker dann so ein, dass er sich zu dieser Zeit dann auch automatisch einschaltet.

Fotodrucker

Als Fotodrucker werden Drucker bezeichnet, die zur Ausgabe digitaler Bilder für den Heimgebrauch konzipiert wurden. Die meisten dieser Geräte können unabhängig von einem PC betrieben werden. Über einem Kartenslot zur Aufnahme von Speicherkarten kann der Fotodrucker auf die digitalen Bilder zugreifen und sie auf einem druckereigenen Display in verschiedenen Ansichten anzeigen. Die nutzende Person kann die gewünschten Bilder auswählen, Farbe, Schärfe und Helligkeit der Fotos anpassen sowie das Format und Rahmenoptionen für den Druck wählen. Die auf der Speicherkarte vorhandenen Bilder können dabei einzeln oder zusammen ausgedruckt werden.

In der entwickelten Computersimulation (Abbildung 8.4) muss dazu die Speicherkarte, die im Ausgangszustand auf der rechten Seite neben dem Fotodrucker abgebildet ist, in den für sie vorgesehenen Speicherkartenplatz gebracht werden. Hierzu müssen die Probanden auf die Speicherkarte klicken und sie mit gedrückter Maustaste auf den Speicherkartenplatz, der sich auf der linken Seite des Fotodruckers befindet, ziehen und dort die Maustaste loslassen. Bevor die Bilder ausgewählt, bearbeitet oder gedruckt werden können, ist das Gerät zunächst über den *Power*-Knopf, der sich auf der Geräteoberseite links unten befindet, einzuschalten. Nach dem Einschalten erscheint zunächst die Meldung, dass das Gerät bereit ist und nach einer kleinen Zeitverzögerung erscheint das erste Foto, das auf der Speicherkarte zu finden ist. In diesem Systemzustand können die Probanden die Anzahl der von diesem Foto zu druckenden Exemplare schrittweise durch das (mehrmalige) Drücken der *oberen* Taste des *Bedienkreuzes* erhöhen. Die Anzahl kann durch das Drücken der *unteren* Taste des *Bedienkreuzes* wieder verringert werden, bis null Exemplare gewählt sind. Das nächste Bild kann durch die *rechte* Taste und das vorherige Bild durch die *linke* Taste des *Bedienkreuzes* aufgerufen werden. Sollen mehrere Bilder gleichzeitig angezeigt werden, so kann mit der *Minus*-Taste (*ZOOM OUT*) die Anzeige der einzelnen Bilder verringert werden, so dass beim einmaligen Drücken der *Minus*-Taste vier Bilder gleichzeitig und beim zweimaligen Drücken der *Minus*-Taste neun Bilder gleichzeitig angezeigt werden. Eine Vergrößerung der Anzeige erfolgt analog mit der *Plus*-Taste (*ZOOM IN*). Zeigt das Display mehrere Bilder gleichzeitig an, so kann der Proband mithilfe der Tasten des *Bedienkreuzes* lediglich ein Foto auswählen. Die Änderung der Anzahl an Druckexemplaren ist nur in der Vollbildanzeige möglich, ebenso wie die Änderung der Bildparameter. Für letztere muss in der Vollbildanzeige die Taste *MENU* gedrückt werden. In dem Menü kann der Farbmodus, die Schärfe sowie die Helligkeit der Bilder geändert werden, indem der jeweilige Menüpunkt durch die *obere* und *untere* Taste des *Bedienkreuzes* gewählt wird und anschließend die jeweilige Option durch die *rechte* und *linke* Taste des *Bedienkreuzes* ausgesucht wird. Durch nochmaliges Drücken der Menütaste kommt der Proband in ein zweites Menü, in dem ein Rahmen, die Druckformate und die Option *Alles drucken* geändert werden kann. Drücken der Menütaste führt hier zu einem Wechsel zwischen den beiden Menüs. Sollen die getätigten Änderungen übernommen werden, so müssen die Probanden die Taste *OK* in der Mitte des *Bedienkreuzes* betätigen. Ist keine Änderungsübernahme gewünscht, so wird die Taste *CANCEL* gedrückt. In beiden Fällen wird das jeweils angezeigte Menü verlassen und die Vollbildanzeige des vorher gewählten Fotos erscheint im Display. Sollen Fotos gedruckt werden, so muss die Funktion über die Taste *PRINT* aufgerufen werden. Die ausgedruckten Bilder erscheinen bei erfolgreicher Bedienung auf der Papierkassette des Druckers, die sich unterhalb des Interfaces befindet.

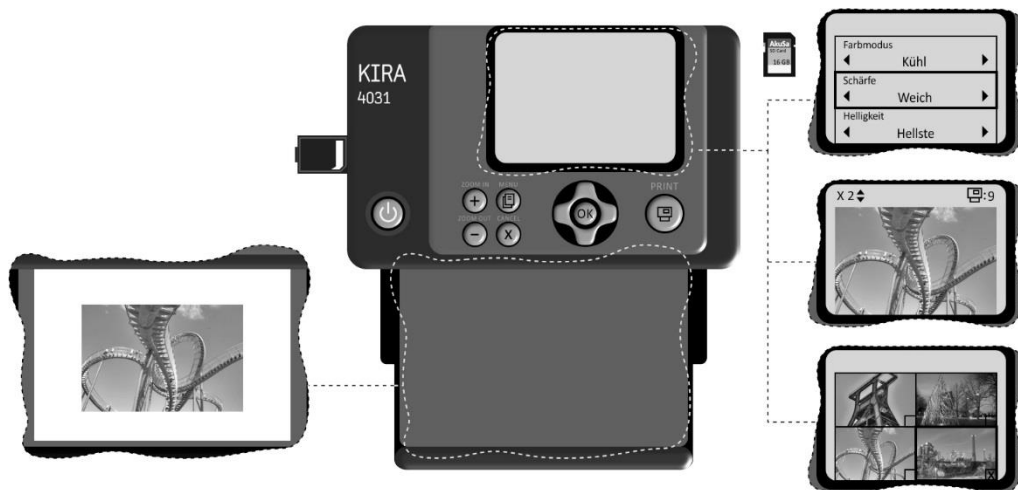


Abbildung 8.4: simuliertes Interface eines Fotodruckers

Die Aufgabenstellungen für die Exploration und die Steuerung des Fotodruckers lauten wie folgt:

Exploration: Mit dem abgebildeten Fotodrucker können Fotos direkt von einem Speichermedium aus gedruckt werden. Ihre Aufgabe ist es zunächst, vorhandene Einstellungsmöglichkeiten und Funktionen des Fotodruckers ausfindig zu machen.

Steuerung: Drucken Sie das erste Foto im Layout 20x30 mm und das zweite Foto ohne eine besondere Layoutvorgabe aus.

Kaffeevollautomat

Die zur vollautomatischen Kaffeezubereitung angebotenen Kaffeevollautomaten für den Heimbedarf unterscheiden sich in ihrer Bedienung von normalen Kaffeemaschinen. Während die Zubereitung von Kaffee mit einer herkömmlichen Kaffeemaschine jedes Mal eine exakte Portionierung von Wasser und Kaffeepulver bedarf, haben Kaffeevollautomaten jeweils ein Vorratsbehälter für Wasser und Kaffeebohnen, aus denen sich der Automat, die entsprechend benötigte Menge selbst entnimmt. Bei der Kaffeezubereitung zerkleinert ein Mahlwerk die Bohnen direkt vor dem Brühvorgang und befördert die gemahlene und ausgepresste Bohnenrückstände nach der Kaffeezubereitung in einen Auffangbehälter, der regelmäßig geleert und gereinigt werden muss. Ebenfalls in regelmäßigen Abständen muss der Vorrat an Wasser und Kaffeebohnen aufgefüllt werden. Neben den benutzerdefinierten Einstellungen zur Tassengröße (Menge an Kaffee) und der verwendeten Menge an Kaffeepulver (Stärke des Kaffees) ist in einem großen Teil der Kaffeevollautomaten zudem eine Entnahmestelle für heißes Wasser zur Zubereitung von Tee oder anderen Heißgetränken integriert.

Damit die Probanden in einer Simulation mit einem Kaffeevollautomaten interagieren können, müssen weitere Gegenstände Inhalt der Simulation sein, die die Probanden zum Ausprobieren der Kaffeemaschine benötigen. Dazu gehören Tassen, Wasser und Kaffeebohnen zum Auffüllen der Vorratsbehälter sowie eine Entsorgungsmöglichkeit für den zu entleerenden Auffangbehälter der Bohnenrückstände (Abbildung 8.5). Die Tassen auf dem Regal des Küchenszenarios können per Drag and Drop unter dem Kaffeeauslauf, den Heißwasserauslauf sowie auf die Abstellflächen rechts neben dem Kaffeeautomaten gestellt

werden. Der Automat wird über die *Power*-Taste eingeschaltet. In dem *Display* erscheinen auf der linken Seite Informationen über die Menge der Kaffeebohnen und des Wassers in den jeweiligen Vorratsbehältern sowie die Menge an Rückständen im Auffangbehälter. In der Mitte des Displays wird der aktuelle Systemzustand des Automaten angezeigt. Die rechte Seite des Displays informiert über die gewählte Tassengröße und die eingestellte Menge des verwendeten Kaffeepulvers und damit über die Kaffeestärke. Diese Einstellungen lassen sich über die beiden Schieberegler ändern. Die Position des linken Schiebereglers bestimmt die *Stärke des Kaffees* und kann durch Anklicken und Ziehen mit der Maus verändert werden. Die Position des rechten Schiebereglers bestimmt die *Tassengröße* bzw. die Menge des zubereiteten Kaffees. Um eine Tasse Kaffee zuzubereiten wird eine Tasse unter den Kaffeeauslauf gestellt und die Taste für *1 Tasse Kaffee* gedrückt. Sollen gleichzeitig zwei Tassen zubereitet werden, müssen zwei Tassen unter den Kaffeeauslauf gestellt und die Taste für *2 Tassen* gedrückt werden. Gefüllte Tassen können ebenfalls per Drag and Drop auf die Fläche neben dem Kaffeeautomaten abgestellt werden. Heißes Wasser lässt sich zubereiten, indem eine Tasse unter dem Wasserauslauf gestellt und die Taste oberhalb des Auslaufes für *heißes Wasser* gedrückt wird. Ist zu wenig Wasser oder sind zu wenige Kaffeebohnen in den oberen Behältern der Kaffeemaschine, erhalten die Probanden den Hinweis über das Display diese aufzufüllen. Für das Nachgießen des Wasserbehälters ist der mit Wasser gefüllte Messbecher von dem Küchenregal per Drag and Drop in den linken oberen Behälter des Kaffeeautomaten zu ziehen. Für das Nachfüllen des Kaffeebohnenbehälters ist analog dazu die Kaffeedose auf dem Regal in den rechten Behälter des Kaffeeautomaten zu ziehen. Falls der Auffangbehälter für Kaffeebohnenreste voll ist, müssen die Probanden den Behälter, der sich an der Kaffeemaschine rechts neben dem Kaffeeauslauf befindet, in den Mülleimer ziehen, um ihn zu entleeren.

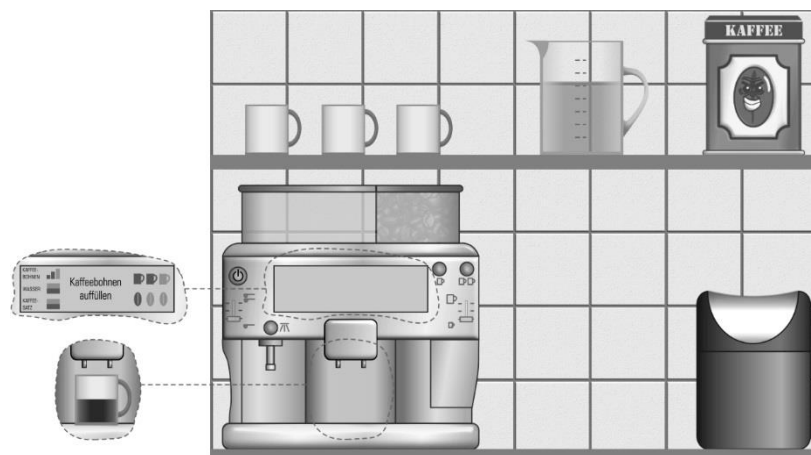


Abbildung 8.5: simuliertes Interface eines Kaffeeautomaten

Die Aufgabenstellungen für die Exploration und die Steuerung der Kaffeemaschine lauten wie folgt:

Exploration: Zu sehen ist ein Kaffeevollautomat, der sowohl Kaffee als auch Heißwasser zubereitet. Die Gegenstände, die sich außerdem in der Küche befinden, können per Drag and Drop verschiedene Funktionen erfüllen. Finden Sie heraus, wie in dem abgebildeten Szenario zu agieren ist.

Steuerung: Bereiten Sie eine Tasse sehr starken Kaffee und zwei Tassen Kaffee mittlerer Stärke zu.

Multifunktionsdrucker

Ein Multifunktionsdrucker ist die Bezeichnung für ein Gerät, das die Funktionen von mehreren, ansonsten getrennt arbeitenden Geräten vereint. Zu den Geräten, die durch die meisten Multifunktionsdrucker ersetzt werden, gehören ein Drucker, ein Scanner sowie ein Kopierer. Multifunktionsdrucker, die unabhängig von einem PC arbeiten, finden sich oftmals in Bibliotheken, großen Supermärkten, Einkaufszentren sowie in Schreibwarengeschäften. Die Bedienelemente der Hardware eines Multifunktionsdruckers stehen meist allen Hauptfunktionen (Drucken, Kopieren, Scannen) zur Verfügung. Über Funktionstasten lässt sich dann bestimmen welche Funktion mit einem Bedienelement gerade angesprochen wird. Um gezielte Einstellungen, innerhalb der jeweiligen Hauptfunktionen Drucken, Kopieren, Scannen etc. vornehmen zu können, enthält ein Multifunktionsdrucker eine Vielzahl von Bedienelementen innerhalb der Software, mit der, bei neueren Geräten über ein Touchdisplay interagiert wird. Zu den häufigsten Einstellungen, die in solchen Geräten vorgenommen werden können, gehören das Papierformat, die Farbe, eine eventuelle Dokumentenvergrößerung, der Duplexdruck, Qualitätseinstellungen, die Textausrichtung und vieles mehr. Für das Drucken und Scannen an öffentlichen Multifunktionsdrucker wird außerdem ein Speichermedium benötigt, auf dem das zu druckende Dokument gespeichert ist und auf dem das zu scannende Dokument gespeichert wird.

In der Simulation der Benutzungsschnittstelle eines solchen Multifunktionsdruckers (Abbildung 8.6) können die Probanden über die Taste *KOPIE* die Einstellungen für das Kopieren von Dokumenten vornehmen, über die Taste *SCAN* die Einstellungen für das Scannen von Dokumenten sowie das anschließende Versenden des Dokumentes per E-Mail und über die Taste *BOX* die Einstellungen für das Speichern gescannter Dokumente auf dem angeschlossenen USB-Stick. Je nach gewählter Funktion erscheinen in dem Gerätedisplay andere Menüs mit den entsprechenden Menüebenen in Form von Registerkarten. Sowohl die Reiter der Registerkarten als auch die einzelnen Menüeinträge können durch direktes Anklicken ausgewählt werden. Die jeweils gewählte Option wird dabei farblich unterlegt. Ist die Hauptfunktion *KOPIE* markiert, kann auf der oberen Menüebene der Kopiervorgang durch Drücken der *START*-Taste des Bedienpanels ausgelöst werden. Ist die Hauptfunktion *SCAN* aktiviert, lässt sich der Scanvorgang durch Drücken der *START*-Taste erst starten, wenn eine E-Mail-Adresse in den Scaneinstellungen angegeben wurde. Hierzu erscheint in dem Display eine Tastatur, mit der die Eingabe vorgenommen werden kann. Soll ein Dokument nach dem Scan nicht an eine E-Mail-Adresse gesendet, sondern auf dem angeschlossenen USB-Stick gespeichert werden, muss die Hauptfunktion *BOX* aktiviert sein. Neben der Möglichkeit Anpassungen an dem Scan (wie Auflösung, Scanformat oder Bildanpassungen) vorzunehmen, können die Probanden vor dem Scan optional einen Dokumentnamen vergeben. Der Scan und die Speicherung des Dokumentes auf dem USB-Stick werden durch Drücken der *START*-Taste ausgelöst.



Abbildung 8.6: simuliertes Interface eines Multifunktionsdruckers

Die Aufgabenstellungen für die Exploration und die Steuerung des Multifunktionsdruckers lauten wie folgt:

Exploration: Zu sehen ist ein Multifunktionsdrucker, mit dem es möglich ist, Dokumente zu kopieren, zu scannen und per E-Mail zu versenden. Finden Sie heraus, wie diese Funktionen genutzt werden können.

Steuerung: Scannen Sie das auf der Scanaufgabe liegende (nicht sichtbare) Dokument und speichern Sie es auf dem angeschlossenen USB-Stick.

Waschmaschine

Die in Haushalten am meisten verbreitete Waschmaschinenbauart sind Trommelwaschmaschinen, die von oben oder von vorne mit waschbaren Textilien beladen werden. Neuere Modelle bieten die Möglichkeit das Waschergebnis durch verschiedene Einstellungsmöglichkeiten zu beeinflussen. Je nach Verschmutzungsgrad, Art der Verschmutzung und den zu waschenden Materialien können beispielsweise die Waschwassertemperatur, die Waschwassermenge, die Waschdauer und sogar die einzelnen Phasen eines Waschvorgangs (Einweichen, Vorwäsche, Hauptwäsche, Spülen, Weichspülen/Stärken, Endschleudern) eingestellt werden. Dazu werden den Nutzern entsprechende Waschprogramme angeboten, aus denen sie wählen und die sie durch weitere Optionen an die eigenen Bedürfnisse anpassen können.

In der Simulation der Waschmaschine (Abbildung 8.7) können die Probanden zwischen diversen Waschprogrammen wählen, indem sie sich durch Drücken der hoch- und runterpfeiltasten in der Auswahl, die sich auf der rechten Seite der Benutzungsschnittstelle befindet, bewegen. Ist in dieser Auswahl die Option *weitere Programme* ausgewählt, so kann zwischen weiteren vier Waschprogrammen durch Drücken der *Rechtspfeil*- und der *Linkspfeil*-Taste gewählt werden. Die Entscheidung für eines dieser Spezialprogramme muss durch das Drücken der *OK*-Taste bestätigt werden. Für die meisten dieser Programme lassen sich Anpassungen bezüglich der Waschwassertemperatur und der maximalen Schleuderdrehzahl vornehmen, wobei die Größe des Auswahlbereiches von dem entsprechenden Programm abhängt. Die Waschwassertemperatur lässt sich im Ausgangszustand durch (mehrmaliges) Drücken der *OK*-Taste ändern, die Drehzahl durch (mehrmaliges) Drücken der *Rechtspfeil*-Taste. Der Waschvorgang kann in manchen Waschprogrammen weiter durch die Angabe der

vorhandenen Flecken optimiert werden. Hierzu muss zunächst die Taste *Flecken* gedrückt werden und mit der *Rechtspfeil*- oder *Linkspfeil*-Taste die entsprechende Fleckenart ausgewählt und die Auswahl dann mit der *OK*-Taste bestätigt werden. Durch Drücken der Taste *weitere Extras* können zusätzliche Optionen wie *Wasser plus*, *Vorwäsche*, *Extra leise* oder *Einweichen* wiederum mit der *Rechtspfeil*- oder *Linkspfeil*-Taste gewählt und mit *OK* bestätigt werden. Unabhängig von dem Waschprogramm und den gewählten weiteren Extras kann jeder Waschvorgang durch Drücken der Taste *kurz* verkürzt werden. Wenn sich die Waschmaschine im Systemzustand *Betriebsbereit* befindet, werden beim Drücken der Taste *Eco Feedback* Informationen über die benötigte Wasser- und Energiemenge des aktuell eingestellten Waschprogramms mit den zusätzlichen Einstellungen angezeigt. Der Waschvorgang kann entweder direkt mit der *Start/Stop*-Taste gestartet werden oder es lässt sich eine Dauer einstellen, nach deren Ablauf der Waschvorgang automatisch gestartet wird. Dazu ist die *Linkspfeil*-Taste im Betriebsbereit-Zustand (mehrmals) zu drücken, bis die gewünschte Dauer im Display steht.

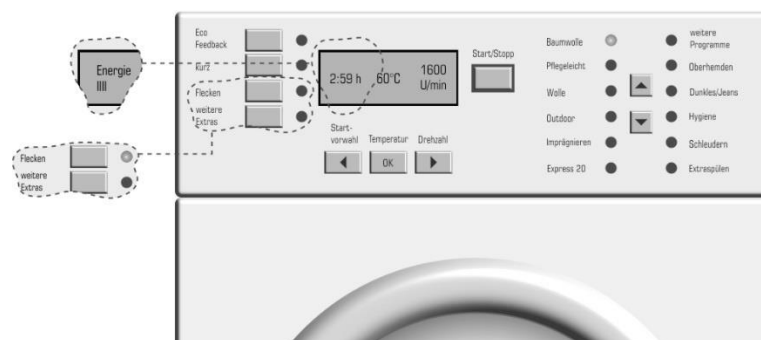


Abbildung 8.7: simuliertes Interface einer Waschmaschine

Die Aufgabenstellungen für die Exploration und die Steuerung der Waschmaschine lauten wie folgt:

Exploration: *Sie machen Urlaub in einer Ferienwohnung, in der Sie auch eine Waschmaschine nutzen können. Diese bietet neben den Waschprogrammen auch diverse Zusatzfunktionen. Bringen Sie in Erfahrung, wie sich welche Einstellungen vornehmen lassen.*

Steuerung: *Sie haben die Waschmaschine mit Wäsche befüllt, die mit Rückständen von Kirschen verunreinigt ist. Es handelt sich hierbei um Baumwollwäsche, die nur bei einer Temperatur von 40°C gewaschen werden darf. Stellen Sie die Waschmaschine passend für Ihre Wäsche und die vorhandene Verunreinigung (Obstflecken) ein und starten Sie die Maschine.*

Küchenradio

Ein Küchenradio ist ein Radio, das speziell für den Gebrauch in der Küche konzipiert wird. Zwar lässt sich jedes Radio und Abspielgerät für Musik in der Küche nutzen, doch bieten spezielle Küchenradios einige Vorteile. Dazu gehört die flache Bauhöhe der Radios sowie die damit verbundene Möglichkeit, sie unter Regalen oder Küchenhängeschränken montieren zu können. Auf diese Weise beanspruchen Küchenradios keine eigenen Stellflächen in der Küche. Weiterhin enthalten viele der erhältlichen Küchenradios nützliche Funktionen wie einen Küchentimer (Eieruhr), Speicherung von Radiosendern oder diverse Weck- und Alarmfunktionen.

Die Simulation der Benutzungsschnittstelle eines Küchenradios ist in Abbildung 8.8 zu sehen. Zu Beginn der Interaktion befindet sich das Gerät im Stand-by-Betrieb. Um einen Radiosender zu hören, müssen die Probanden das Gerät durch Drücken der *Power*-Taste in den Radiobetrieb überführen. Der voreingestellte Radiosender wird abgespielt und kann über die Tasten *Up* und *Down* verändert werden. Drücken der Tasten *Vol +* und *Vol -* führen zu einer Änderung der Abspiellautstärke. Um einen Radiosender mit den *Speichertasten 1-6* abzuspeichern, ist im Radiobetrieb die Taste *Memory Setting* zu drücken. Im Display erscheint der Text *Mem.* Nun muss die Taste der sechs Speichertasten gedrückt werden, auf die der Sender abgespeichert werden soll. Um die Eingabe abzuschließen, ist hiernach noch einmal die Taste *Memory Setting* zu betätigen. Der in dem Küchenradio integrierte Küchentimer lässt sich nur im Stand-by-Betrieb des Gerätes durch Anklicken der Taste *Timer* stellen. Mit den Tasten *Up* und *Down* kann die gewünschte Zeit eingestellt werden. Erneutes Drücken der Taste *Timer* startet den Küchentimer. Nach Ablauf der eingestellten Zeit wird der zuletzt gehörte Radiosender eingeschaltet. Der Küchentimer kann durch erneutes Drücken der Taste *Timer* wieder ausgeschaltet werden. Mit dem Küchenradio können außerdem zwei Weckzeiten programmiert werden. Ebenso wie der Timer lassen sich diese nur im Stand-by-Betrieb erstellen. Für die Erstellung der ersten Weckzeit wird die Taste *Alarm 1* (identisch mit der Taste *Vol +*) gedrückt. Die Eingabe der Stunde und der Minute der Weckzeit kann über die Tasten *Up* und *Down* erfolgen. Der Wechsel zwischen den Stunden und den Minuten ermöglicht die Taste *Rechtspfeil*. Wird die Taste *Alarm 1* mehrmals betätigt, lässt sich zwischen den akustischen Hinweisen *Signaltöne*, *Radio* oder *Abgeschaltet* auswählen, für die das jeweilige Symbol in dem Display angezeigt wird. Um die Wochentage, an denen das Wecksignal ertönen soll, einzustellen, wird die Taste *Alarm Mode* (identisch mit der Taste *Timer*) mehrmals angeklickt. Die Änderungen werden mit der Taste *Mode Setting* übernommen und das Gerät wird wieder in den Stand-by-Betrieb gesetzt. Für die Erstellung der zweiten Weckzeit wird die Taste *Alarm 2* gedrückt und dann wie für das Erstellen der ersten Weckzeit verfahren. Weitere Anpassungen wie das Einstellen der aktuellen Uhrzeit, das aktuelle Datum, das Ein- und Ausschalten des Dimmers, die Weckmelodie sowie den Wechsel zwischen einem 12- oder 24-Stundenmodus werden im Stand-by-Betrieb durch das Drücken der Taste *Mode Setting* vorgenommen. Dabei wird durch das mehrmalige Drücken dieser Taste zwischen den genannten Einstellungsoptionen gewählt und mit den Pfeiltasten entsprechende Änderungen herbeigeführt.



Abbildung 8.8: simuliertes Interface eines Küchenradios

Die Aufgabenstellungen für die Exploration und die Steuerung des Küchenradios lauten wie folgt:

Exploration: *Das abgebildete Küchenradio bietet neben dem Empfang von Radiosendern noch weitere Funktionen. Erkunden Sie das Küchenradio und finden Sie heraus, über welche Funktionen es verfügt.*

Steuerung: Mit den Tasten 1 bis 6 des abgebildeten Küchenradios können Radiosender abgespeichert werden. Diese können dann durch Drücken der jeweiligen Speichertaste abgerufen werden. Die Aufgabe ist es nun, mit der Taste 1 den Sender mit der Frequenz von 98,5 kHz abzuspeichern.

Kraftfahrzeug-Navigationsgerät

Navigationssysteme, die eine Zielführung zu einem gewählten Ort oder eine Route unter Beachtung gewünschter Kriterien ermöglichen, gibt es als transportable Stand-alone-Geräte, also Software-Erweiterungen für Laptops und Smartphones und als Kraftfahrzeug-Navigationsysteme, die in der Regel durch die Integration in Autoradios fest im Auto verbaut sind. Der damit verbundene Vorteil ist die automatische Absenkung der Lautstärke von Radio- oder Musikwiedergaben, wenn die Autolautsprecher für Richtungsansagen des Navigationssystems verwendet werden. Individuelle Wünsche bei der Verkehrsführung und der Navigationsanzeige können ebenso wie Sonderziele oder die Heimatadresse von dem Navigationssystem berücksichtigt werden.

Die Simulation eines Navigationsgeräts, das in einem Autoradio integriert ist, ist demnach komplex, so dass sich die Beschreibung des simulierten Systems auf einige wenige Hauptfunktionen beschränkt. Die Interaktion mit dem Navigationsgerät erfolgt ausschließlich über die in Abbildung 8.9 abgebildeten Tasten, das Display kann nicht für Bedienhandlungen genutzt werden, sondern dient lediglich zur Anzeige der Systemreaktionen. Wenn das Gerät über die *Power*-Taste eingeschaltet wird, wird die Funktion wieder aktiviert, die zum Zeitpunkt des Ausschaltens des Gerätes aktiviert war, beim ersten Einschalten des Systems ist es die Radiofunktion. Mit den Funktionstasten *PHONE*, *MEDIA*, *RADIO*, *TRAFFIC*, *SETUP* und *NAV* kann zwischen den einzelnen Hauptfunktionen gewechselt werden. Die durch die jeweilige Wahl einer Hauptfunktion dargebotenen Menüeinträge können mit dem Pfeiltasten erreicht und durch die *OK*-Taste gewählt werden. Mit den Tasten *Plus* und *Minus* lässt sich die Wiedergabelautstärke von Radio und anderen Musikmedien anpassen, unabhängig in welchem Betriebszustand sich das Gerät aktuell befindet. Das Drücken der Funktionstaste *SETUP* ermöglicht in Abhängigkeit des aktuellen Betriebszustandes verschiedene Einstellungen vorzunehmen. Im Radiobetrieb können beispielsweise Einstellungen zum Verkehrsfunk oder zur Übertragung von Zusatzfunktionen geändert werden; im Navigationsbetrieb hingegen Routenoptionen angepasst, Sonderziele aktiviert oder eine Heimatadresse hinterlegt werden. Innerhalb einiger Hauptfunktionen werden unten links und unten rechts im Display zusätzliche Auswahloptionen angeboten, die mit den sich darunter befindlichen Bedienelementen angewählt werden können. Aus den unteren Menüebenen gelangen die Probanden durch die *Zurück*-Taste wieder auf die oberste Menüebene innerhalb einer Funktion.



Abbildung 8.9: simuliertes Interface eines Navigationsgerätes

Die Aufgabenstellungen für die Exploration und die Steuerung des Navigationsgerätes lauten wie folgt:

Exploration: Das in Ihrem neuen Auto eingebaute Radio hat auch ein Navigationssystem. Versuchen Sie möglichst vollständig das gesamte Gerät zu erkunden, um alle vorhandenen Funktionen bedienen zu können.

Steuerung: Speichern Sie die folgende Adresse als Heimatadresse ab, um zu dieser über nur einen Tastendruck schnell navigieren zu können. Stadt: Essen, Strasse: Hauptstrasse, Hausnummer: 2.

Festnetztelefon

Der Begriff Festnetztelefon wird zur Unterscheidung von einem Mobiltelefon verwendet. Die für private Haushalte erhältlichen Festnetztelefone unterscheiden sich in ihrem Aussehen und in ihren Funktionen kaum mehr von Mobiltelefonen, im Gegensatz zu den in Büros verwendeten Geräten. Diese sind meist mit weiteren Endgeräten an eine virtuelle Telefonanlage angeschlossen und erlauben so kostenlose interne Gespräche. Weitere Funktionen, die sich auf das Telefonieren beziehen, sind die Einrichtung von Umleitungen, die individuelle Belegung von Tasten mit Rufnummern, die Speicherung von Wahlwiederholungen sowie die Abfrage des Anrufbeantworters oder die Rufnummernunterdrückung.

Die Interaktion mit dem Festnetztelefon (Abbildung 8.10) kann in der entwickelten Simulation ohne das Abheben des abgebildeten Telefonhörers erfolgen (ist in der Simulation auch nicht möglich). Soll eine Telefonnummer gewählt werden, muss zunächst die Taste *Lautsprecher* gedrückt werden. Die Eingabe einer Telefonnummer wird mit dem Tasten 0-9 vorgenommen und der Anruf mit der Taste ✓ ausgeführt. Zum Beenden des Anrufes muss die Taste *Lautsprecher* erneut gedrückt werden. Eventuell vorhandene Nachrichten auf dem Anrufbeantworter lassen sich mit der Taste *Briefkasten* abfragen und die Abfrage über die Taste *Lautsprecher* wieder beenden. Eine zuvor abgespeicherte Wahlwiederholung kann über die Taste *Wahlwiederholung* angerufen werden, zum Beenden des Anrufes ist wieder die Taste *Lautsprecher* zu drücken. Um eine Wahlwiederholung abzuspeichern oder Tasten mit Rufnummern zu belegen, muss die Taste *Service Menü* angeklickt werden. Auf einer Menüebene können sich die Probanden mit der *Rechtspfeil*- und der *Linkspfeil*-Taste bewegen und eine Auswahl mit der Taste ✓ bestätigen. Jede Menüebene enthält einen Eintrag *zurück*, dessen Bestätigung die Probanden zurück auf die nächsthöhere Menüebene führen. Zur Belegung der freien grauen Tasten auf der rechten Seite des Telefons, muss im *Service Menü*

zunächst die Option *Zielwahl* und dann die zu belegende Taste gewählt werden. Nach Eingabe der zu hinterlegenden Rufnummer wird diese durch ✓ bestätigt.

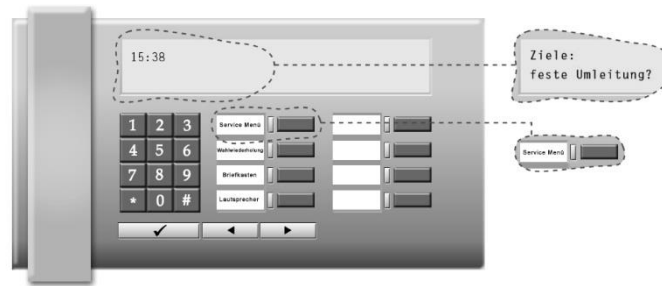


Abbildung 8.10: simuliertes Interface eines Festnetztelefons

Die Aufgabenstellungen für die Exploration und die Steuerung des Festnetztelefons lauten wie folgt:

Exploration: Dies ist Ihr Telefon an Ihrem neuen Arbeitsplatz. Es bietet die Möglichkeit Rufnummern oder Funktionen auf den 8 rechteckigen grauen Tasten zu hinterlegen. Der Hersteller hat die 4 Tasten auf der linken Seite mit bestimmten Funktionen vordefiniert. Erkunden Sie Ihr neues Telefon.

Steuerung: Da Sie bestimmte Rufnummern sehr häufig wählen werden, möchten Sie sie auf den belegbaren Tasten abspeichern. Hinterlegen Sie folgende Rufnummer Ihrer Chefin: 01577/1234567 auf die erste frei belegbare Taste auf der rechten Seite, so dass, wenn Sie sie betätigen, direkt Ihre Chefin angerufen wird.

Spülmaschine

Für die maschinelle Reinigung von Töpfen, Geschirr, Küchenutensilien und Besteck bieten Geschirrspülmaschinen unterschiedliche Programme und Zusatzfunktionen an, aus denen der Nutzer in Abhängigkeit der Geschirr- und Verschmutzungsart wählen kann. Die Programmauswahl hat einen Einfluss auf die Spülwassermenge, Spülwassertemperatur, den Sprühdruk sowie die Dauer der einzelnen Phasen des Spülvorgangs. Ein Spülvorgang besteht aus den Phasen Vorspülen, Reinigen, Zwischenspülen, Klarspülen und Trocknen. Neben Leitungswasser benötigt eine Spülmaschine außerdem Reinigungsmittel (Pulver, Tabs oder ähnliches), Spülmaschinensalz zur Wasserenthärtung, sowie Klarspülmittel. Da sich die für einen Spülvorgang benötigte Salz- und Klarspülmittelmenge nach dem Wasserhärtewert richtet, sollten die Salzzugabemenge und die Klarspülmenge einzustellen sein.

Die Wahl eines Spülprogramms erfolgt in der Simulation der Spülmaschine (Abbildung 8.11) im eingeschalteten Zustand (Drücken der *Power*-Taste) durch die Betätigung eines der sechs rechteckigen *Programmtasten*. In dem Display leuchtet die Dauer des entsprechenden Spülprogramms auf. Mit den, neben den Programmtasten angebrachten runden *Zusatzfunktionstasten* können je nach gewähltem Programm weitere Optionen eingestellt werden. Die Wahl der Funktion *Zeitsparen* verkürzt die Laufzeit des Spülprogramms, die Wahl der Funktion *Hygiene* erhöht die Temperatur während des Reinigungsvorgangs. Um den Spülvorgang sofort zu starten, muss die Taste *Start* gedrückt werden. Soll der Spülvorgang auf einen späteren Zeitpunkt verschoben werden, kann mit den Pfeiltasten der Zeitpunkt in Stundenschritten bis zu 24 Stunden verschoben werden. Die Anpassung der pro Spülvorgang

zuzugebenden Menge an Spezi­alsalz kann eingestellt werden, indem die erste Programmtaste auf der linken Seite im ausgeschalteten Zustand der Geschirrspülmaschine betätigt wird. Das Display zeigt den zunächst eingestellten Wert von *H04* an. Dieser kann durch (mehrmaliges) Drücken der *Start*-Taste verändert werden. Um getätigte Änderungen zu übernehmen, muss die erste *Programmtaste* noch einmal betätigt werden. Analog dazu kann mit der mittleren *Programmtaste* auf der linken Seite des Displays die zuzugebende Klarspülermenge verändert werden. Die Zeit, nach der sich der Geschirrspüler nach Programmende automatisch abschaltet, kann auf die gleiche Weise mit der unteren, linken Programmtaste und die Ein- und Ausschaltung der Intensivtrocknung mit der oberen, rechten Programmtaste verändert werden.



Abbildung 8.11: simuliertes Interface einer Spülmaschine

Die Aufgabenstellungen für die Exploration und die Steuerung der Geschirrspülmaschine lauten wie folgt:

Exploration: Über das hier sichtbare Bedienfeld einer Spülmaschine können Programme und Zusatzfunktionen gewählt sowie grundsätzliche programmunabhängige Einstellungen vorgenommen werden. Erkunden Sie das abgebildete Gerät möglichst genau.

Steuerung: Für gute Spülergebnisse benötigt die Geschirrspülmaschine kalkarmes Wasser. Leitungswasser muss mithilfe von Spezi­alsalz enthärtet werden. Die Einstellung und damit die benötigte Salzmenge ist abhängig vom Härtegrad des Leitungswassers und ist bei der abgebildeten Geschirrspülmaschine von *H00* bis *H07* einstellbar. Stellen Sie einen Wert von *H06* für sehr hartes Leitungswasser ein.

Rollladensteuerung

Neben der Möglichkeit der manuellen Bedienung von Rollläden per Gurt oder Kurbel stellt der Motorantrieb eine weitere Bedienungsvariante dar. Die Steuerung erfolgt bei motorbetriebenen Rollläden über eine Funksteuerung, mit der die Rollläden per Tastendruck geöffnet oder geschlossen werden. Viele Funksteuerungen bieten außerdem die Möglichkeit mithilfe von Zeit- und Zufallsschaltungen die Rollladensteuerung zu automatisieren.

In der Simulation der Rollladensteuerung lässt sich der Rollladen durch Drücken der großen Pfeiltaste nach unten der Funksteuerung schließen. In dem, in Abbildung 8.12 links neben der Funksteuerung abgebildeten Fenster bewegt sich daraufhin der Rollladen nach unten. Die Abwärtsbewegung kann durch Drücken der *Stop*-Taste (in der Mitte des Bedienkreuzes) gestoppt werden. Das Öffnen des Rollladens über die große Pfeiltaste nach oben kann nur im geschlossenen oder gestoppten Zustand erfolgen. Wird die Menü­taste *M* betätigt, wird die aktuell eingestellte Uhrzeit und der Wochentag ausgeblendet und die Probanden gelangen in das Einstellungsmenü der Funksteuerung. Die Menü­einträge sind aus Platzgründen mit

Ziffern bezeichnet. Die Bewegung auf einer Menüebene erfolgt mit den *kleinen Pfeiltasten*. Um in das jeweilige Untermenü zu gelangen, muss der jeweilige Menüeintrag mit der Taste *OK* bestätigt werden. Die zusätzlich eingeblendeten Symbole dienen hier als Hinweis auf die Funktionen, die sich durch den jeweiligen Menüeintrag aufrufen lassen. Änderungen an den Einstellungen lassen sich durch die *kleinen Pfeiltasten* vornehmen und durch Drücken der Taste *OK* übernehmen.

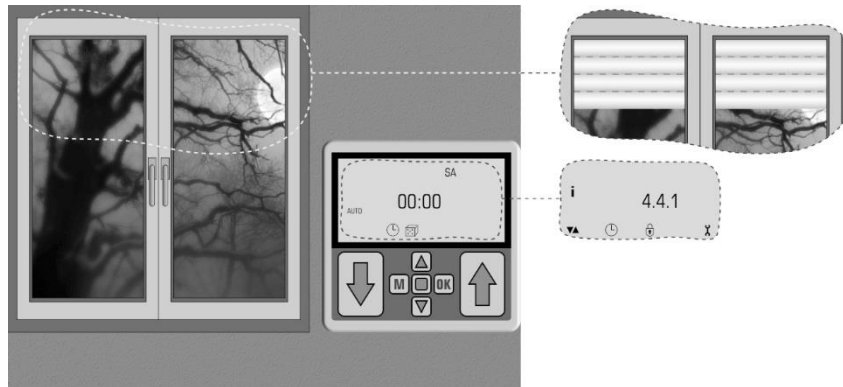


Abbildung 8.12: simuliertes Interface einer Rollladensteuerung

Die Aufgabenstellungen für die Exploration und die Steuerung der Rollladensteuerung lauten wie folgt:

Exploration: Zu sehen ist eine Rollladensteuerung, mit der Rollläden automatisch und manuell bedient werden können. Lernen Sie die Bedienung dieser Steuerung kennen.

Steuerung: Schalten Sie in dem Bedienpanel der Rollladensteuerung die Displaysperre ein, so dass ein versehentliches Verstellen der Einstellungen nicht möglich ist.

Heizungssteuerung

Heizkörper, die mit Funk-Thermostaten ausgestattet sind, lassen sich über eine Wandsteuerung entweder über Tastendruck oder vollautomatisch steuern. Zeitprogramme, die individuell angepasst werden können, steuern die Heizung so, dass nur zu bestimmten Zeiten das Thermostat geöffnet und damit geheizt wird. In der Regel kann zwischen einem Automatik- und einen Manuell-Modus hin und hergewechselt werden.

Der Wechsel zwischen diesen beiden Modi ist in der erstellten Simulation durch Drücken der Taste *FUNKTION* möglich (Abbildung 8.13). In dem Display erkennt man den aktuellen Modus durch den Text *MANU* (für Manuell) und *AUTO* (für Automatik). Befindet sich das Gerät im manuellen Betrieb, kann mithilfe der *Pfeiltasten* die gewünschte Raumtemperatur von 6°C bis 30°C eingestellt oder auf *OFF* gestellt werden. Um ein Heizprogramm erstellen zu können, ist die Taste *PROG* zu betätigen. Im ersten Schritt kann durch die Pfeiltasten der zu programmierende Wochentag ausgewählt werden. Erneutes Drücken der Taste *PROG* führt zu dem nächsten Schritt, in dem ebenfalls durch Drücken der Pfeiltasten schrittweise die Uhrzeit (in vollen Stunden) eingestellt wird, zu der auf die, an anderer Stelle zu definierende Tagestemperatur geheizt wird. Weiteres Drücken der Taste *PROG* ermöglicht die schrittweise Einstellung der Uhrzeit, zu der auf die Nachttemperatur abgesenkt werden soll. Ein weiteres Drücken der Taste *PROG* übernimmt die Einstellungen und bringt das Gerät in den

Betriebsbereitzustand. Um die Tag- und Nachttemperaturen einzustellen, ist die unterste mit einer Sonne und einem Mond beschriftete Taste zweimal innerhalb einer Sekunde zu drücken. Im ersten Schritt kann die gewünschte Raumtemperatur für den Tag mit den Pfeiltasten eingestellt werden, die gewünschte Raumtemperatur für die Nacht lässt sich im zweiten Schritt durch erneutes Drücken dieser Taste einstellen. Im letzten Schritt kann auf dieselbe Weise die Raumtemperatur angegeben werden, wenn ein Fenster geöffnet wird (diese Funktion setzt allerdings das Vorhandensein von Funk-Fensterkontakten voraus). Erneutes Drücken der *Tag-Nacht*-Taste führt wieder in den Betriebsbereitzustand. Ein zweimaliges Betätigen der Taste *FUNKTION* innerhalb einer Sekunde ermöglicht die Einstellung von Datum und Uhrzeit. Um zwischen beiden Einstellungsmöglichkeiten auszuwählen werden die Pfeiltasten benutzt und die Wahl mit der Taste *PROG* bestätigt. Diese Taste wird auch für den Wechsel zwischen Tag, Monat und Jahr sowie zwischen Stunde und Minute verwendet. Zur eigentlichen Änderung müssen die Pfeiltasten (mehrmals) gedrückt werden.

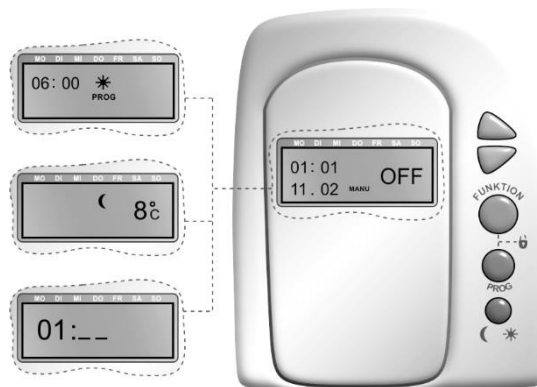


Abbildung 8.13: simuliertes Interface einer Heizungssteuerung

Die Aufgabenstellungen für die Exploration und die Steuerung der Heizungssteuerung lauten wie folgt:

Exploration: Das abgebildete Funk-Heizkörperthermostat dient zur Temperaturregelung in Einzelräumen. Es enthält ein an die Lebensgewohnheiten anpassbares Zeitprogramm, mit dem es möglich ist, einen Raum warm zu halten, wenn er genutzt wird, während zu anderen Zeiten durch Absenken der Temperatur Energie gespart wird. Finden Sie heraus, wie das Funk-Heizkörperthermostat funktioniert, bzw. eingestellt werden kann.

Steuerung: Das Heizungsthermostat soll so eingestellt werden, dass von Montags bis Freitags ab 07.00 Uhr auf eine Temperatur von 22°C geheizt und ab 21.00 Uhr auf eine Temperatur von 12°C abgesenkt wird. Am Samstag und Sonntag soll auf die Tagtemperatur von 22°C erst ab 08.00 Uhr geschaltet und auf die Nachttemperatur von 12°C ab 23.00 Uhr abgesenkt werden. Stellen Sie die geforderten Temperaturen und Zeiten am Heizungsthermostat richtig ein und stellen Sie es anschließend auf den Automatikbetrieb um.

Mobiltelefon

Mobiltelefone der früheren Generation verfügten aufgrund des Stands der Technik über wesentlich weniger Funktionen als heutige Smartphones. Die Bedienung dieser Geräte erfolgt ausschließlich über die Tasten der Hardware, da sie über keine Touchdisplays verfügen. Damit ein versehentliches Drücken von Tasten während der Aufbewahrung der Geräte in

Taschen keine Zustandsänderungen herbeiführt, sind diese Geräte mit einer Tastensperre ausgestattet, die sich mit einer gerätespezifischen Tastenkombination ein- und ausschalten lässt. Die in ihnen implementierten Funktionen beziehen sich meist auf das Telefonieren in den entsprechenden Mobilfunknetzen.





Simuliert wurde ein solches, wie oben beschriebenes Mobiltelefon (Abbildung 8.14). Um mit diesem Gerät interagieren zu können, muss zunächst die Tastensperre ausgeschaltet werden. Hierzu muss die #-Taste einmal gedrückt werden. Um eine Rufnummer aus dem Telefonbuch des Mobiltelefons zu wählen, muss die Taste  gedrückt werden. Die Probanden können zwischen sechs Telefonbucheinträgen wählen. Durch Drücken der Taste *Anrufen*, wird der vorher ausgewählte Telefonbucheintrag angerufen. Die Taste *Auflegen* beendet den Anruf wieder. Anruflisten können über die Taste *INT* aufgerufen werden. Zuerst wird die Liste der entgangenen Anrufe zum Aufruf angeboten. Vertikales Blättern zur nächsten Liste ist mit den Tasten *INT* und  möglich, da sie innerhalb eines Menüs die Funktion von Pfeiltasten nach oben bzw. unten übernehmen. Weitere aufrufbare Anruflisten sind die der angenommenen und der ausgehenden Anrufe. Um sich die Inhalte einer Liste anzeigen zu lassen, muss diese mit der rechten oberen Auswahltaste bestätigt werden. Die Listeneinträge können mit den Tasten *INT* und  durchblättert werden. Aus jedem Menü führt die linke Auswahltaste wieder zurück auf eine höhere Menüebene. Die zuletzt gewählten Telefonnummern können wiedergewählt werden, indem die linke Auswahltaste im (entsperrten) Ausgangszustand gedrückt wird. Es erscheint eine Liste mit vier Einträgen, die sich ebenfalls mit den Tasten *INT* und  durchblättern lässt. Wird die rechte Auswahltaste im (entsperrten) Ausgangszustand betätigt, können Audioeinstellungen vorgenommen werden. Für die Gesprächslautstärke kann zwischen drei Stufen und für den Akkuton zwischen vier Tönen gewählt werden.



Abbildung 8.14: simuliertes Interface eines Mobiltelefons

Die Aufgabenstellungen für die Exploration und die Steuerung des Mobiltelefons lauten wie folgt:

Exploration: Ihr Mobilfunkgerät ist in der Reparatur und Sie haben nun ein Leihgerät erhalten. Finden Sie heraus, wie dieses funktioniert.

Steuerung: Ihre Krankenkasse hat Sie vor ein paar Minuten angerufen. Sie konnten den Anruf aber aus Zeitgründen nicht entgegennehmen. Rufen Sie nun Ihren letzten Anrufer zurück.

Internetrouter

Internetrouter sind Netzwerkgeräte, die das Heimnetzwerk mit dem Internet verbinden. Erfolgt die Verbindung über einen DSL-Anschluss werden diese Geräte auch als DSL-Router bezeichnet. Die meisten der heute erhältlichen Modelle stellen für das Heimnetzwerk WLAN zur Verfügung, so dass auch die Bezeichnung WLAN-Router üblich ist. Da die Zahl der WLAN-fähigen Geräte in den Privathaushalten zunimmt, wird die Verwaltung dieser immer wichtiger. Die meisten Internetrouter bieten hierzu die Möglichkeit auf die Firmware der Geräte über einen Browser zuzugreifen. Über die Benutzeroberfläche erhält der Nutzer Informationen über vorhandene Internet- und ggf. Telefonverbindungen, die Belegungen der Geräteanschlüsse sowie aktive Geräte im Heimnetzwerk (wie Mobiltelefone, WLAN-Drucker, DECT-Telefone, Musikabspielgeräte etc.).

In Abbildung 8.15 ist ein Auszug aus der simulierten Benutzungsoberfläche der Gerätefirmware eines Internetrouters zu sehen. Die Interaktion mit der Benutzungsoberfläche der Software erfolgt dabei nicht über eine simulierte Hardware, sondern wie bei jeder Computersoftware direkt über die Maus und die Tastatur der Rechner, die die Probanden für den computerbasierten Test nutzen. Die einzelnen Hauptfunktionen sind in die Kategorien *Internet*, *Telefonie*, *Heimnetz* und *System* aufgeteilt. Für die Kategorie, die gerade gewählt ist, werden außerdem die einzelnen Unterkategorien in der Navigationsleiste auf der linken Seite angezeigt. Wenn die Simulation gestartet wird, ist die Seite *Zugangsdaten* aus der Kategorie *Internet* sichtbar. Hier lassen sich Einstellungen bezüglich des Internetanbieters, der Zugangsart sowie den Zugangsdaten vornehmen. Auf der Seite *Kindersicherung* lässt sich die Internetnutzung von verschiedenen Geräten des Heimnetzwerkes regulieren, indem zwischen drei verschiedenen Nutzungsprofilen gewählt werden kann. Die Freigabe einzelner Ports des Internetzugangs, um beispielsweise mit anderen Teilnehmern an Onlinespielen teilzunehmen, kann über die Menüseite *Freigaben* erfolgen. Die Seite *DSL-Informationen* dient lediglich der Information über die DSL-Verbindung und nicht deren Einrichtung. Alle Einstellungen, die das WLAN des Heimnetzwerkes betreffen, können auf der Seite *WLAN* vorgenommen werden. Dazu gehören die Änderungen an dem WLAN-Verschlüsselungstyp, die Einrichtung einer WLAN-Zeitschaltung, die Einrichtung eines Gastzugangs oder die Einbindung eines weiteren Repeaters, um die WLAN-Reichweite zu erhöhen. Unter der Hauptfunktion *Telefonie* finden sich Einstellungsmöglichkeiten bezüglich Anrufbeantworter, Telefonbücher, Rufsperrern, Rufumleitungen oder Weckrufe. Innerhalb der Kategorie *Heimnetz* lassen sich neue Geräte (u.a. Speichermedien, SMART-Home-Geräte) sowie Benutzer einrichten und verwalten. Systemeinstellungen wie die Einrichtung einer Nachtschaltung oder das Zurücksetzen des Internetrouters auf die Werkseinstellungen sind unter der Kategorie *System* möglich.

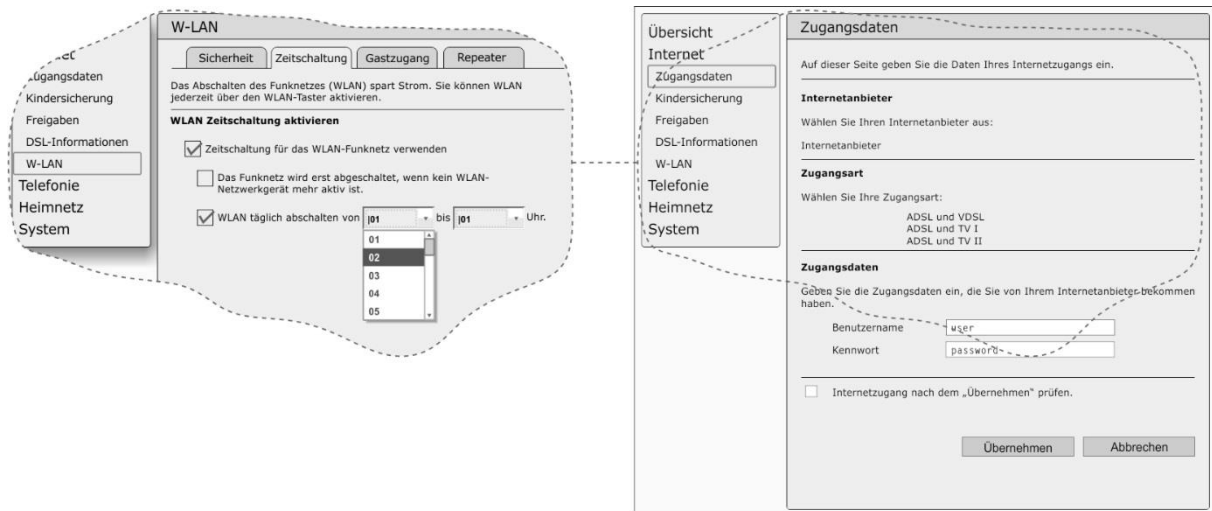


Abbildung 8.15: simulierte Interface der Software eines Internetrouters

Die Aufgabenstellungen für die Exploration und die Steuerung des Internetrouters lauten wie folgt:

Exploration: Die Benutzeroberfläche eines Internetrouters bietet vielfältige Einstellmöglichkeiten. Erkunden Sie die abgebildete Software für einen Internetrouter.

Steuerung: Um Energie zu sparen, möchten Sie das WLAN in der Nacht ausschalten. Nehmen Sie dazu die entsprechenden Einstellungen an dem Internetrouter vor.

Fahrscheinautomat

Um den öffentlichen Personennahverkehr nutzen zu können, muss vorab ein für die geplante Fahrt entsprechender Fahrschein gekauft werden. Dazu stehen an vielen Haltestellen Fahrscheinautomaten zu Verfügung, über deren Bildschirm zunächst aus einer Auswahl angebotener Fahrscheinarten der passende ausgewählt werden muss. An den meisten Automaten können mehrere Fahrschein ausgewählt und gleichzeitig gekauft werden.

Der für diese Arbeit entwickelte, simulierte Fahrscheinautomat (Abbildung 8.16) kann über das Touchdisplay des Gerätes bedient werden. In dem Ausgangszustand erhalten die Probanden eine Übersicht mit Ticketvorschlägen sowie einigen interaktiven Schaltflächen. Die Schaltfläche *Kurzstrecke* gibt eine Liste von Fahrtzielen an, die ausgehend von dem Standpunkt des Fahrscheinautomaten mit dem Tickettyp Kurzstrecke erreicht werden können. Die Schaltfläche *i* gibt eine Auflistung von Kategorien an, über die die Probanden Tarifinformationen beziehen können. Aus den unteren Menüebenen gelangt eine Person durch Drücken der Schaltfläche *Zurück* auf eine höhere Menüebene. Wird ein Ticket durch Drücken ausgewählt, gelangen die Probanden auf eine Seite, die zum Bezahlvorgang auffordert. Dort werden die angeforderten Fahrschein aufgeführt und der zu zahlende Betrag angezeigt. Ausgehend von dieser Bezahlfläche können weitere Tickets ausgewählt und der Liste zu kaufender Fahrscheinen zugefügt werden. Durch die Schaltfläche *Auswahl rückgängig*, lassen sich Fahrscheinen aus der Liste wieder entfernen. Der Kaufvorgang wird mit dem Bezahlen des Fahrscheins abgeschlossen. Hierzu können die vier rechts neben dem Fahrscheinautomaten abgebildeten *Geldscheine* genutzt werden, indem diese per

Drag and Drop auf dem unter dem Automaten-Display angebrachten Geldschlitz gezogen werden.

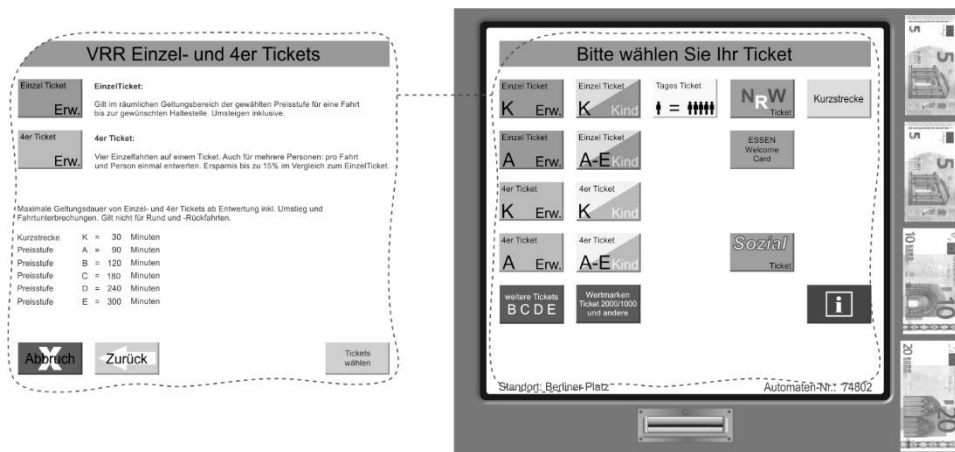


Abbildung 8.16: simuliertes Interface eines Fahrscheinautomaten

Die Aufgabenstellungen für die Exploration und die Steuerung des Fahrscheinautomaten lauten wie folgt:

Exploration: *An dem nebenstehenden Fahrscheinautomat können verschiedene Fahrscheine für den öffentlichen Nahverkehr sowie weitere Fahrscheine erworben werden. Finden Sie zunächst heraus, welche verschiedenen Fahrscheine angeboten werden und wie sie gekauft werden können.*

Steuerung: *Kaufen Sie den preisgünstigsten Fahrschein, um mit dem Bus bis zu der Haltestelle Essen Hbf fahren zu können. Bezahlen Sie mit den abgebildeten Scheinen.*

Fotoentwicklungsautomat

Die Erstellung von Abzügen digitaler Fotos ist an Selbstbedienungsautomaten möglich, die in vielen Geschäften zu finden sind. Dabei können die Nutzer zwischen verschiedenen Fotoprodukten und -formaten auswählen sowie sich die Produkte entweder selbst ausdrucken oder sie an ein Fotolabor schicken, die diese dann den Wünschen entsprechend ausbelichten. Um die digitalen Fotos an die Fotostation zu übermitteln benötigt man ein Speichermedium wie beispielweise einen USB-Stick, eine CD bzw. DVD oder eine SD-Karte.

In der Interaktion mit dem simulierten Fotoentwicklungsautomaten (Abbildung 8.17) müssen die Probanden zunächst entscheiden, ob sie Fotos selbst vor Ort ausdrucken oder diese an ein Fotolabor senden möchten. Wenn die Probanden die Optionen *Fotos sofort* über das Automaten-Display gewählt haben, werden ihnen daraufhin verschiedene Produktkategorien (*Fotos & Fotosets, Fotogrufskarten, Poster, Fotokalender*) angeboten. In der Simulation wurde ausschließlich die Option *Fotos & Fotosets* umgesetzt. Wählen die Probanden diese aus, können sie nachfolgend zwischen verschiedenen Bildformaten auswählen und im folgenden Fenster dann das Speichermedium angeben, von dem die Bilder übertragen werden sollen. In der Simulation steht ein USB-Stick zur Verfügung, der rechts neben dem Automaten abgebildet ist. Dieser muss nach Aufforderung per Drag and Drop auf den USB-Anschluss des Gerätes gezogen werden. Die auf dem Stick gespeicherten Bilder erscheinen in einer Vorschau, aus der entweder einzelne oder alle Bilder für den Druck ausgewählt werden. In dieser Vorschau ist

es nachträglich möglich einzelne Bilder in anderen Formaten, als das zuvor angegebene zu drucken. Einzelne Bilder können durch Anklicken in die Auswahl hinzugenommen und durch die Taste + die Anzahl der zu druckenden Exemplare erhöht werden. Soll die Auswahl gedruckt werden, ist die Schaltfläche *Weiter zum Druck* anzuklicken. Es erscheint eine Liste der bestellten Produkte, deren Einzelpreise sowie der Gesamtpreis. Drücken der Taste *Fotos drucken* beendet den Bestellvorgang und löst den Druck aus.



Abbildung 8.17: simuliertes Interface eines Fotoentwicklungsautomaten

Die Aufgabenstellungen für die Exploration und die Steuerung des Fotoentwicklungsautomaten lauten wie folgt:

Exploration: Mit dem abgebildeten Fotoentwickler lassen sich professionelle Abzüge von Fotos, die auf einem Speichermedium gespeichert sind, erstellen. Versuchen Sie möglichst alle Funktionen und Einstellungsmöglichkeiten zu erfassen.

Steuerung: Lassen Sie von dem Fotoentwickler von allen Bildern, die auf dem abgebildeten USB-Stick gespeichert sind, jeweils 2 Abzüge in dem Format 15x20 cm sofort ausdrucken.

Online-Banking

Für die Abwicklung von Bankgeschäften bieten Banken ihren Privatkunden neben dem Telefonbanking auch das Online-Banking an, für das entweder ein PC, ein Smartphone oder ein ähnliches elektronisches Endgerät benötigt wird. In der Regel erfolgt der Zugriff auf den Bankrechner über die Webseite der Bank, indem sich die Nutzer mit vorher beantragten Zugangsdaten dort ausweisen müssen. Per Online-Banking kann der aktuelle Bankkontostand eingesehen, zurückliegende und teils zukünftige Transaktionen geprüft, Kontoauszüge abgerufen, Überweisungen getätigt, Daueraufträge eingerichtet, Sparpläne angelegt und Depots eingerichtet werden. Die elektronische Unterschrift, die für jede der genannten Tätigkeiten erforderlich ist, wird durch die Eingabe von TAN-Nummern getätigt, die die Kunden durch die Zusendung von TAN-Listen, durch einen zusätzlichen TAN-Generator oder durch den Versand von SMS auf das Mobiltelefon erhalten.

In der Simulation des Online-Banking-Portals (Abbildung 8.18) können die Probanden durch verschiedene Seiten navigieren. Auf der zu Beginn der Simulation aufgerufenen Seite *Konten und Depots* ist das eingerichtete Girokonto mit der entsprechenden Kontonummer sowie dem aktuellen Kontosaldo zu sehen. Es stehen auf dieser Seite die Funktionen *Überweisung*, *Umsatzanzeige*, *Zinsen* sowie *Adressänderung* zur Verfügung, die über den

entsprechenden Button aufgerufen werden können. Soll eine Überweisung getätigt werden, müssen dazu in dem dann angezeigten Eingabeformular alle Felder ausgefüllt werden. Nach Anklicken des Buttons *Übernehmen* werden die Probanden aufgefordert eine bestimmte TAN-Nummer einzugeben, die sie in der unter dem Monitor abgebildeten TAN-Liste finden. Die Eingabe wird mit dem Button *Übernehmen* bestätigt und es erscheint die Seite *Umsatzanzeige*, die eine Übersicht der letzten getätigten Überweisungen gibt. Eine Liste der eingerichteten Daueraufträge findet sich auf der Seite *Daueraufträge*, auf die man nur gelangt, indem vorher die Funktion *Überweisung* gewählt wurde. Dort kann durch Drücken der Taste *Neuer Dauerauftrag* ein Dauerauftrag angelegt werden. Der Vorgang ist dabei mit dem zur Tätigung einer Überweisung identisch. Die Seite *Zinsen* informiert über die Zinserträge der letzten zwei Jahre. Um die Adresse des Kontoinhabers ändern zu können ist zunächst der Button *Adressänderung* zu drücken, es folgt die Anzeige des aktuell hinterlegten Namens, der Postanschrift sowie der E-Mail-Adresse. Der Button *Ändern* ermöglicht eine Änderung aller Felder, bis auf den Namen der kontoinhabenden Person. Unter dem Eintrag *Post-Box* können Nachrichten, die von der Bank dort hinterlegt werden, abgespeichert werden. Über die Seite *Service* werden die Funktionen Adressänderung, Kontodetails und Zinsen noch einmal zusammengefasst.

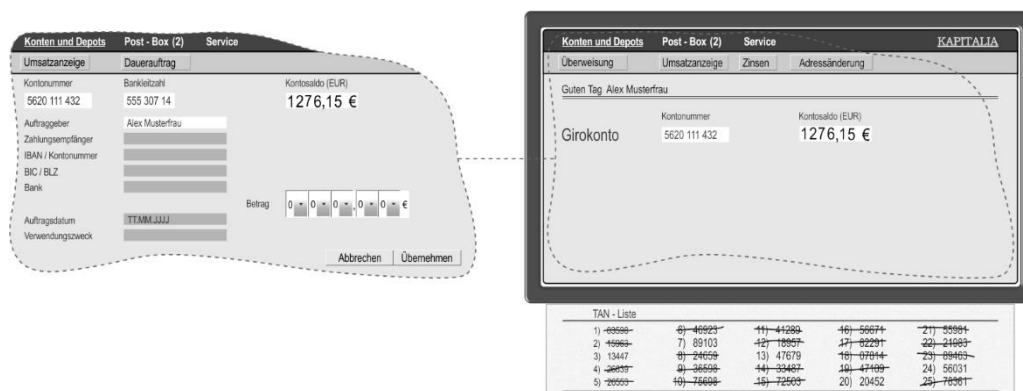


Abbildung 8.18: simuliertes Interface eines Online-Banking Portals

Die Aufgabenstellungen für die Exploration und die Steuerung des Online-Banking Portals lauten wie folgt:

Exploration: Ihre Hausbank präsentiert Ihnen hier ihr neues Online-Banking Portal. Navigieren Sie so durch das abgebildete Portal, dass Sie alle Funktionen und Informationen kennenlernen.

Steuerung: Richten Sie folgenden Dauerauftrag ein: Empfänger: Wohnungsgesellschaft
Kontonummer: 123345324; Bankleitzahl: 50030040; Bank: Sparkasse Entenhausen;
Verwendungszweck: Miete Wohnung 04; Erste Ausführung: 01.01.2015; letzte Ausführung: ---;
Betrag: 300,00 €; Turnus: monatlich.

8.5.2 Beschreibung des zusammengestellten Tests

Für die Zusammenstellung der einzelnen oben beschriebenen Items zu einem Gesamttest müssen einige Aspekte bedacht werden: die Reihenfolge der Items in dem Test, die äußere Gestaltung des Testmaterials, die vorgesehene Testzeit sowie die Testinstruktion. Sowohl im Rahmen der Klassischen als auch der Item-Response-Theorie wird gefordert, dass die Bearbeitung eines Items nicht von der Beantwortung eines anderen Items abhängig sein darf

(Moosbrugger & Kelava, 2012b, S. 68). Übertragen auf den im Rahmen dieser Arbeit entwickelten Problemlösetest bedeutet dies, dass beispielsweise die Bedienung des simulierten Radioweckers nicht davon abhängig sein darf, wie ein Proband zuvor mit dem Parkscheinautomaten interagiert hat. Auch wenn keine logischen Abhängigkeiten zwischen den Items bestehen, so müssen dennoch Positions- und Reihenfolgeeffekte berücksichtigt werden. Die Veränderung der Schwierigkeit eines Items aufgrund seiner Positionierung im Test wird als Positionseffekt bezeichnet, der vor allem Items am Testanfang (warming-up-Phase) und am Testende (Ermüdung, Zeitmangel oder nachlassende Testmotivation) betreffen (Rost, 2004, S. 69). Als Reihenfolgeeffekt wird die Beeinflussung der Schwierigkeit durch die Items, die vorher bearbeitet wurden, bezeichnet. Wurden beispielsweise bereits viele einfache Items bearbeitet, können folgende schwierige Items aufgrund der Übung leichter gelöst werden (ebd., S. 70). Solche Effekte können nicht ganz beseitigt werden und solange diese dazu führen, dass ein Item für alle Probanden leichter oder schwerer wird, ist die lokale stochastische Unabhängigkeit nicht verletzt (ebd.).

In Leistungstests werden die Items in der Regel nach aufsteigender Schwierigkeit sortiert, um eine Überforderung der Probanden bereits zu Beginn zu vermeiden. Sehr leichte Items werden an den Testanfang gestellt, um die Probanden für die weitere Bearbeitung zu motivieren (Moosbrugger & Kelava, 2012b, S. 68). Da die Schwierigkeit der einzelnen neu entwickelten Items noch nicht bekannt ist, werden die Items zunächst anhand von (aus den schwierigkeitsbestimmenden Merkmalen begründeten) Annahmen zu ihrer Schwierigkeit in eine Reihenfolge gebracht:

1. Parkscheinautomat
2. Receiver
3. Radiowecker
4. Fotodrucker
5. Kaffeemaschine
6. Multifunktionsdrucker
7. Waschmaschine
8. Küchenradio
9. Navigationsgerät
10. Festnetztelefon
11. Spülmaschine
12. Rollladensteuerung
13. Heizungssteuerung
14. Mobiltelefon
15. Internetrouter
16. Fahrscheinautomat
17. Fotoentwicklungsautomat
18. Online-Banking Portal

Das Layout des Gesamttests muss so angepasst sein, dass es die Bearbeitung des Tests erleichtert. Unterstützend wirkt dabei eine einfache und übersichtliche Gestaltung des Testmaterials (Pospeschill, 2010, S. 67). Aus diesem Grund wurde ein für alle Items gleichbleibendes Layout gewählt (Abbildung 8.19). Das Browserfenster, in dem die einzelnen

Items präsentiert werden, ist hierbei zweigeteilt. Innerhalb der rechten Seite des Fensters, die annähernd dreiviertel der Gesamtansicht einnimmt, läuft die entsprechende Simulation ab. In dem verbleibenden linken Teil der Testansicht erhalten die Probanden die jeweilige itemspezifische Instruktion. An erster Stelle stehen der Name des Gerätes, das gerade zu bedienen ist, sowie direkt darunter die Art der Bedienung, die aktuell im Fokus steht. Jedes Item muss zunächst exploriert werden (Phase Exploration), um Informationen zu generieren und darauf aufbauend Eingriffswissen zu erwerben und direkt im Anschluss nach einer konkreten Zielvorgabe gesteuert werden (Phase Steuerung). Darunter, mit ein paar Zeilen Abstand wird die jeweilige Aufgabenstellung dargeboten, aus der Informationen zu dem Gerät und dem Bedienungsziel entnommen werden können. Um von der Explorationsphase eines Gerätes zur Steuerungsphase oder von der Steuerungsphase zur Explorationsphase des nächsten technischen Gerätes wechseln zu können, dient der Weiter-Button unterhalb der Aufgabenstellung. Ein Zurückspringen zu einer vorherigen Phase oder eines vorherigen Items ist nicht vorgesehen.

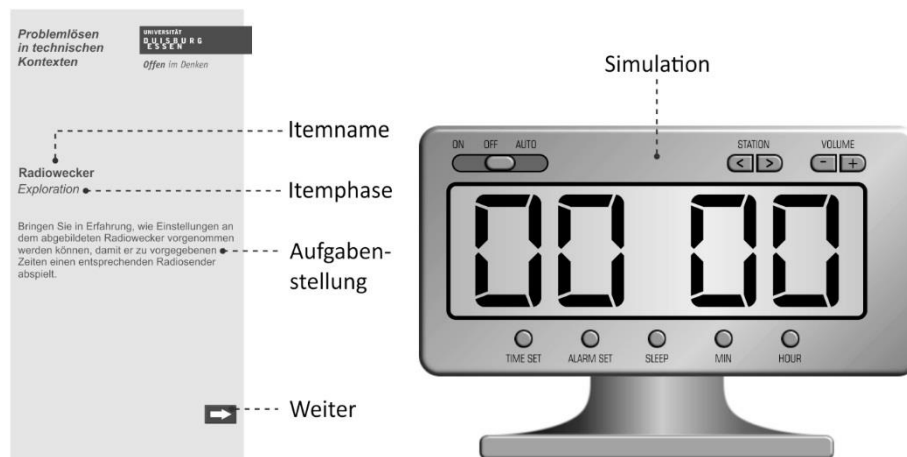


Abbildung 8.19: gleichbleibendes Layout für alle Items des computerbasierten Tests zum Umgang mit technischen Geräten am Beispiel des Items Radiowecker

Der entwickelte Problemlösetest soll zwar als Niveau- und nicht als Speedtest eingesetzt werden, dennoch muss eine Begrenzung für die komplette Testdauer festgelegt werden. Lienert & Raatz (1998, S. 164) schlagen für die Zeitbegrenzung eines Tests, der als Niveautest durchgeführt wird, die Zeit vor, in der 80 bis 90 % aller Probanden alle Aufgaben bearbeiten konnten. Da für den neuentwickelten Problemlösetest noch keine solche Zeit angegeben werden kann, wird für die Exploration eines jeden einzelnen Items eine Begrenzung von zehn Minuten und für die Steuerung eines jeden Items eine Begrenzung von fünf Minuten festgelegt. Mit einer Testlänge von 18 Items ergibt sich somit eine maximale Gesamtdauer von 270 Minuten (4,5 Stunden). Diese Testbegrenzungen können in den einzelnen definierten Tasks innerhalb des *CBA ItemBuilders* schon hinterlegt werden. Für den in dieser Arbeit entwickelten Test ist außerdem im Fall der abgelaufenen Testzeit eine Meldung vorgesehen, die über den Ablauf der Zeit informiert: *Die Zeit für die Exploration (Steuerung) ist abgelaufen. Durch Klicken des ⇨-Buttons auf der linken Seite gelangen Sie zur Steuerungsaufgabe (zum nächsten technischen Gerät).*

Zu Beginn der Testdurchführung müssen die Testteilnehmer generelle (also alle Testitems betreffende) Handlungsanweisungen erhalten. Diese Testanweisung (Instruktion) enthält

ferner Informationen über das Forschungsvorhaben, in dem die Testentwicklung eingebettet ist, den Aufbau des Tests, die Art der darin enthaltenen Items sowie eine Erläuterung des Antwortmodus. Um sicher zu gehen, dass die Testteilnehmer diese Anweisungen verstanden haben, enthält diese außerdem ein Übungsitem, das vor den eigentlichen Items bearbeitet wird.

8.5.3 Beschreibung des Scorings

Die Bewertung des Probandenverhaltens im Umgang mit den computersimulierten technischen Systemen wird, aufgrund der unterschiedlichen Interaktionsziele, getrennt nach Exploration und Steuerung vorgenommen. Ziel des *Explorierens* ist der Erwerb von Wissen über die Bedienung des jeweiligen technischen Gerätes, so dass die Interaktionen der Probanden danach ausgerichtet sein sollten, möglichst viele Informationen zu generieren. Informationsgenerierung findet durch Interaktionen statt, die zu einer Änderung des Systemzustandes führen. Je größer die Anzahl der erreichten Systemzustände, desto größer ist die Anzahl der generierten Informationen. Die absolute Anzahl der explorierten Systemzustände kann jedoch kein Maß für die Leistung der Probanden in der Explorationsphase sein, da sich die technischen Systeme in der Anzahl ihrer Systemzustände (ihrer Komplexität) unterscheiden. Um die Explorationsleistung der Probanden durch einen numerischen Wert auszudrücken, wird die Anzahl der explorierten Systemzustände an der Gesamtzahl der vorhandenen Zustände des jeweiligen Systems relativiert (Abbildung 8.20). Das sich ergebende Maß der Explorationsvollständigkeit kann demnach Werte zwischen 0 und 1 annehmen, wobei der Wert 0 für keine Exploration und der Wert 1 für vollständige Exploration steht.

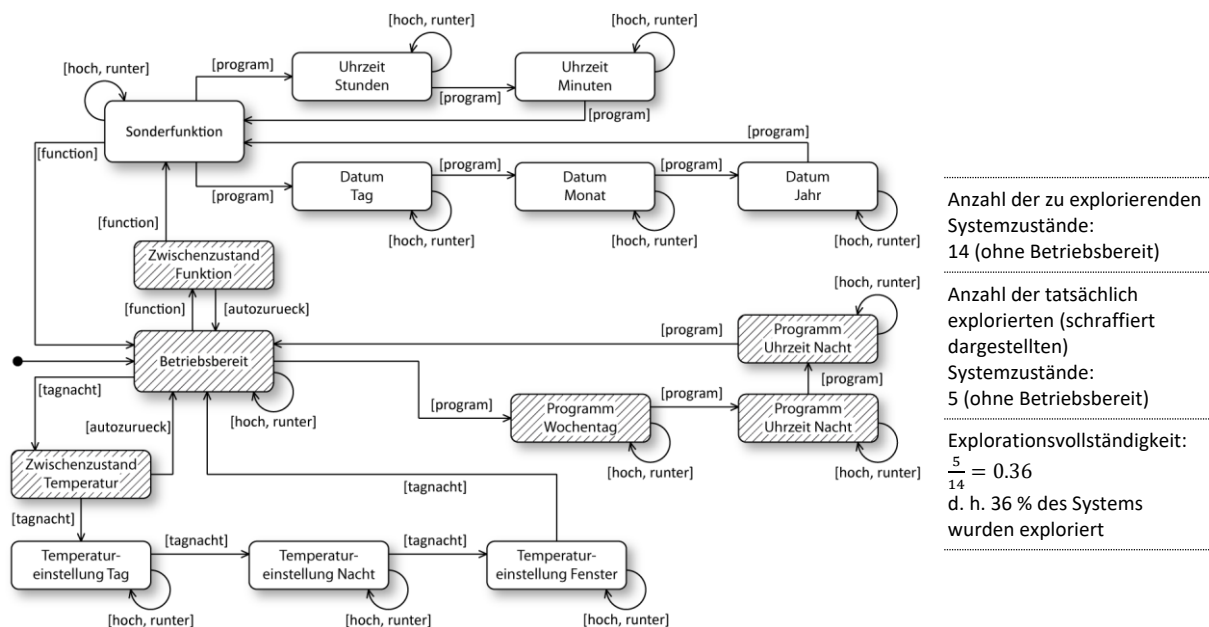


Abbildung 8.20: Beispiel für die Ermittlung der Explorationsvollständigkeit anhand der relativen Anzahl an explorierten Systemzuständen in einer Heizungssteuerung (die Zustände Start und Betriebsbereit werden nicht mitgezählt, da diese Zustände automatisch auch ohne Interaktionen erreicht werden)

Die Rohwertbestimmung für die Leistung eines Probanden/einer Probandin in der *Systemsteuerung* kann prinzipiell auf unterschiedlichen Wegen erfolgen. Da der

computerbasierte technische Problemlösetest ein, im Rahmen dieser Arbeit neu entwickeltes Testinstrument ist, kann vorab noch nicht angegeben werden, welche Methode der Rohwertbestimmung geeigneter (in Bezug auf Verteilungsannahmen, Modellkonformität etc.) ist. Im Folgenden werden deshalb zunächst die möglichen Methoden der Rohwertbestimmung für die Steuerungsphase vorgestellt; ein Vergleich dieser unterschiedlichen Indikatorwerte erfolgt später im Kapitel 11.7 und 13.

In jeder Steuerungsphase müssen die Probanden ein vorgegebenes Ziel erreichen, das sich in einigen Items in weitere Teilziele zerlegen lässt. Für das Erreichen eines jeden Teilziels erhalten die Probanden 1 Punkt, so dass sich die in Tabelle 8.3 dargestellten Maximalpunktzahlen für jedes Item ergeben. Die Leistung der Probanden kann also zum einen danach bewertet werden, wie viele der (Teil-)Ziele, bezogen auf die mögliche Zielanzahl erreicht wurden. Zum anderen kann die Probandenleistung derart bewertet werden, dass nur bei Erreichen aller (Teil-)Ziele 1 Punkt vergeben wird. Diese dichotome Auswertungsweise bewertet die Leistung der Probanden im steuernden Umgang mit den technischen Systemen weniger differenziert, als die zuvor genannte polytome Auswertungsmethode: Letztere ist jedoch nicht bei allen Items anwendbar.

Tabelle 8.3: In der Steuerungsphase maximal zu erreichende itemabhängige Punktzahl sowie die Anzahl der hierfür nötigen Bearbeitungsschritte

Item	Maximale Punktzahl	nötige Anzahl an Bearbeitungsschritten
Waschmaschine	4	11
Parkscheinautomat	1	39
Spülmaschine	1	5
Fotoentwicklungsautomat	1	11
Multifunktionsdrucker	1	4
Kaffeemaschine	2	11
Küchenradio	1	9
Navigationsgerät	3	101
Fotodrucker	4	19
Radiowecker	5	71
Festnetztelefon	1	21
Rollladensteuerung	1	10
Heizungssteuerung	17	99
Mobiltelefon	1	5
Internetrouter	9	15

Eine weitere Möglichkeit der Rohwertermittlung, bezieht neben den Resultaten der Problembearbeitung außerdem die Qualität des Problemlöseprozesses mit ein. Je weniger Interaktionen Probanden für die Problemlösung benötigen, desto besser sollte demnach ihre Leistung bewertet werden. Das in der Steuerungsphase vorgegebene Ziel eines jeden Items kann auf einem direkten Weg erreicht werden, der durch eine Mindestanzahl an Bearbeitungsschritten charakterisiert ist (Tabelle 8.3). Die Größe der Differenz zwischen dieser mindestens benötigten Bearbeitungsschrittzahl und der tatsächlich von den Probanden benötigten Anzahl an Bearbeitungsschritten kann in Kategorien eingeteilt werden, für die weitere Punkte vergeben werden. Nach diesem Bewertungsschema (Tabelle 8.4) erhalten

Probanden, die das vorgegebene Steuerungsziel nicht erreicht haben 0 Punkte. Probanden, die das Ziel erreicht haben, hierfür aber mehr als 100 % der Mindestanzahl an Bearbeitungsschritten benötigten 1 Punkt, die Probanden, die mehr als 50 % der Mindestanzahl an Bearbeitungsschritten benötigten 2 Punkte und die Probanden, die zwischen 0 und 50 % der Mindestanzahl an Bearbeitungsschritten benötigten 3 Punkte. Die tatsächliche benötigte Anzahl an Bearbeitungsschritten wird aus den während der Testbearbeitung erstellten Logfiles entnommen.

Tabelle 8.4: Bewertung der Problemlöseleistung in der Steuerungsphase mit der Berücksichtigung der für die Lösung benötigten Anzahl an Bearbeitungsschritten bei gegebener Mindestanzahl an Bearbeitungsschritten

Lösung	Punkte
Falsch	0
Richtig, mit über 100 % mehr Bearbeitungsschritte als nötig	1
Richtig, mit über 50 % mehr Bearbeitungsschritte als nötig	2
Richtig, mit weniger als 50 % mehr Bearbeitungsschritte als nötig	3

8.6 Erprobung des Testinstrumentes

Bevor das entwickelte Testinstrument eingesetzt wird, um die Validität des Kompetenzkonstrukts *problemlösender Umgang mit technischen Alltagsgeräten* zu prüfen, muss es zunächst selbst erprobt werden. Ziel der Erprobung ist die Aufdeckung von technischen Problemen bei der Testdurchführung, von Programmfehlern in den Simulationen sowie von Konstruktionsmängeln bei der Testinstruktion, bei der Formulierung der Aufgabenstellung und der Itemanordnung, um diese anschließend korrigieren zu können (vgl. Pospeschill, 2010, S. 69). Erprobt werden sollte neben der eigentlichen Testdurchführung außerdem die anschließende Auswertung bzw. die verwendete Messskala. Um Rückschlüsse aus der Erprobung ziehen zu können, sind Beobachtungen während der Testdurchführung zu protokollieren, retrospektive Befragungen der Probanden durchzuführen und die entstandenen Logfiles zu analysieren.

Nur unter möglichst realistischen Bedingungen kann die Erprobung des Testinstrumentes aufschlussreiche Hinweise liefern. Die Stichprobe hierfür sollte also entweder aus der Zielgruppe stammen oder dieser möglichst ähnlich sein (vgl. Moosbrugger & Kelava, 2012b, S. 70). Weil davon auszugehen ist, dass das Interesse an technischen Sachverhalten einen Einfluss auf die Techniknutzung ausübt, soll der Einfluss des Interesses in dieser Arbeit zunächst weitgehend konstant gehalten werden. Für die Haupterhebung werden daher Probanden gesucht, die ein Interesse an Technik bzw. an technischen Geräten haben. Zu dieser Zielgruppe gehören beispielsweise Studierende, die ein technisches Fach studieren. Sie eignen sich zudem als Stichprobe, weil sie mit verhältnismäßig geringem Aufwand für die Teilnahme an dem Test gewonnen werden können. Für die erste Erprobung des Testinstrumentes sind ebenfalls technisch interessierte Probanden zu gewinnen. Aus forschungsökonomischen Gründen wurden hierzu zunächst Personen ausgewählt, die zu diesem Zeitpunkt zur Verfügung standen. Dabei handelt es sich um zwei verschiedenen Personengruppen: Zuerst wird der Test an wissenschaftlichen Mitarbeitern und studentischen Hilfskräften aus einem technischen Fachbereich der Universität Duisburg-Essen erprobt. Daraufhin bearbeiten Teilnehmer eines Kraftwerksmeisterkurses der Kraftwerksschule Essen den Test. Nach dieser

ersten Erprobungsphase werden eventuell vorhandene Fehler beseitigt und gegebenenfalls Verbesserungen vorgenommen. Das Testinstrument wird nach der Revision einer weiteren Gelegenheitsstichprobe, bestehend aus Teilnehmern eines weiteren Kraftwerksmeisterkurses zur Bearbeitung vorgelegt. Der nächste Abschnitt geht auf den Ablauf der beiden Erprobungsphasen, die aufgedeckten Probleme und Fehler sowie den daraus gezogenen Konsequenzen ein.

8.6.1 Erste Testerprobungsphase

Die erste Testerprobung wurde von 8 Universitätsmitarbeitern (4 wissenschaftliche Mitarbeiter, 4 studentische Hilfskräfte) aus einem technischen Fach als Gruppentest durchgeführt. Für die Auslieferung des Tests standen 8 baugleiche Laptops zur Verfügung, die entsprechend den Anforderungen für das Ausführen der Simulationen konfiguriert wurden. Als Laufumgebung für den zusammengestellten Test diente die zum CBA ItemBuilder Version 5.02 passende Execution Environment (EE), die zunächst auf allen Rechnern installiert werden musste. Der in der EE integrierte JBoss® Server ist dann für die Zusammenstellung und für die Durchführung des Testes manuell zu starten. Weiterhin benötigt sowohl die Administration als auch die Durchführung des Tests einen Browser. Hierzu wurde die Version 31 des Mozilla® Firefox® installiert. Da einige der simulierten technischen Geräte bei bestimmten Interaktionen Musik oder Töne abspielen sollen, wird außerdem für den Browser der Adobe® Flash® Player benötigt, der in der Version 14 auf den Laptops installiert wurde. Damit es nicht zu einer gegenseitigen Störung der Probanden durch die abgespielten Laute kommt, wurden an den Laptops Kopfhörer angeschlossen, die von den Probanden während der Testbearbeitung zu tragen waren. Zum Zeitpunkt der ersten Testerprobung konnten die Items *Internetrouter*, *Fahrscheinautomat*, *Fotoentwicklungsautomat* und das *Online-Banking Portal* aufgrund technischer Probleme noch nicht für den Test verwendet werden. Die restlichen 14 Items wurden den Probanden in der nachstehenden Reihenfolge präsentiert:

1. Parkscheinautomat
2. Receiver
3. Radiowecker
4. Fotodrucker
5. Kaffeemaschine
6. Multifunktionsdrucker
7. Waschmaschine
8. Küchenradio
9. Navigationsgerät
10. Festnetztelefon
11. Spülmaschine
12. Rollladensteuerung
13. Heizungssteuerung
14. Mobiltelefon

Der Test wurde in der Execution Environment so eingerichtet, dass die Probanden bevor sie mit der Bearbeitung beginnen konnten, einen Benutzernamen eintragen mussten, unter dem die Execution Environment während des Tests die Logfiles der einzelnen Tasks abspeichert.

Die Probanden der ersten Testerprobung wurden gebeten, hier die Nummer des Rechners einzutragen, an dem sie den Test bearbeiteten. Nach erfolgter Eingabe wurde die erste Aufgabe in dem Browserfenster aufgerufen. Die Hälfte der Probanden konnte den Test in dem Zeitraum von 1 h 09 min abschließend bearbeiten. Innerhalb einer Dauer von 1 h 25 min schlossen alle Probanden den Test ab.

Während der Testbearbeitung wurden bei einigen technischen Systemen Reaktionen beobachtet, die nicht intendiert waren und deren Ursache in der Beschreibung der Zustandsübergänge lag. Als problematisch erwies sich bei einer Person außerdem die Benutzung der Tastatur während der Testbearbeitung. Nach dem Drücken der *Enter*-Taste schloss sich das Browserfenster, in dem der Test angezeigt wurde. In der Folge musste der Test erneut aufgerufen werden und die Person sich ein weiteres Mal mit einem neuen Benutzernamen anmelden. Die vorab bereits bearbeiteten Items wurden in dieser Bearbeitung übersprungen, indem der *Weiter*-Button entsprechend oft gedrückt wurde.

Bevor eine weitere Stichprobe den computerbasierten Test erprobte, wurden die aufgedeckten Fehler zweier simulierter technischer Geräte beseitigt. Die von den Probanden in einer retrospektiven Befragung vorgeschlagenen Verbesserungen werden erst nach der zweiten Erprobung umgesetzt. Zum einen wurde empfohlen, die Anzahl der noch zu bearbeitenden Items sichtbar zu machen und zum anderen die Itemnamen zweier technischer Geräte innerhalb der linken Aufgabenstellung zu ändern, da diese schon etwas über die Aufgabestellung in der Steuerungsphase verraten (*Rufumleitung* für das Festnetztelefon und *Rückruf* für das Mobiltelefon).

Die zweite Gelegenheitsstichprobe bestand aus 20 angehenden Kraftwerksmeistern, die zum Testzeitpunkt an einem Kurs der Kraftwerksschule in Essen teilnahmen. Alle 20 Teilnehmer sind männlich und zum Zeitpunkt der Testerprobung zwischen 24 und 45 Jahren alt ($MW = 31.75$, $SD = 6.70$). Für die Auslieferung des Tests, der analog zur ersten Erprobung als Gruppentest durchgeführt wurde, stellte die Kraftwerksschule zunächst 24 Desktop-Computer mit Monitoren, Tastaturen und Mäusen zur Verfügung. Bei den Geräten handelte es sich um seit längerer Zeit nicht genutzte Computer, deren Betriebssysteme und installierte Software seit der letzten Nutzung nicht mehr aktualisiert wurden. Die bereits während der Konfiguration der Rechner festgestellte unzureichende Verarbeitungsgeschwindigkeit führte dazu, dass ihre Eignung für die Testauslieferung in Frage gestellt wurde. Der nötige Austausch der Geräte erfolgte zum einen durch die 8 Laptops, die bereits in der ersten Erprobung eingesetzt wurden und zum anderen durch weitere 10 Laptops verschiedener Hersteller und 2 Desktop-Computer, die aus dem aktuellen Bestand der Kraftwerksschule stammten. Alle verwendeten Computer erhielten die gleiche Konfiguration wie die zur ersten Testerprobung eingesetzten Computer.

Änderungen in dem zusammengestellten Test gegenüber der ersten Erprobung ergaben sich zum einen aus der Behebung der aufgedeckten Programmierfehler in zwei simulierten technischen Geräten und der Behebung der technischen Probleme in den Items *Internetrouter*, *Fahrscheinautomat*, *Fotoentwicklungsautomat* und *Online-Banking Portal*, sodass diese in der zweiten Erprobung ebenfalls eingesetzt werden konnten.

Insgesamt wurden 18 Items in der nachstehenden Reihenfolge zu einem Test zusammengeführt:

1. Parkscheinautomat
2. Receiver
3. Radiowecker
4. Fotodrucker
5. Kaffeemaschine
6. Multifunktionsdrucker
7. Waschmaschine
8. Küchenradio
9. Navigationsgerät
10. Festnetztelefon
11. Spülmaschine
12. Rollladensteuerung
13. Heizungssteuerung
14. Mobiltelefon
15. Internetrouter
16. Fahrscheinautomat
17. Fotoentwicklungsautomat
18. Online-Banking Portal

Vor Beginn der Testbearbeitung mussten die Probanden einen Benutzernamen in ein Browserfenster eingeben, unter dem die Execution Environment die generierten Logfiles abspeichert. In der zweiten Testerprobung konnten die Testteilnehmer den einzugebenden Namen frei wählen. Erst die Analyse der Logfiles ergab, dass Umlaute in dem Dateinamen für die Logfiles unzulässig sind. Für die nachfolgende Testbearbeitung wurde maximal eine Zeit von 2 h 30 min benötigt, die meisten der Probanden (90 %) schloss die Bearbeitung der 18 Items nach 01 h 58 min ab.

Im Verlauf der Testbearbeitung deckten Testteilnehmer Programmierfehler in einigen wenigen Simulationen auf, die in dem ersten Testdurchlauf nicht festgestellt wurden. Diese Beobachtungen wurden während der Bearbeitung protokolliert, so dass das fehlerhafte Verhalten der Simulationen im Anschluss überprüft und die ursächlichen Fehler eliminiert werden konnten. Beobachtet wurde außerdem, dass alle Probanden in der Simulation des Parkscheinautomaten versuchten, die neben dem Automaten abgebildeten Geldmünzen über die Funktion Drag and Drop in den Automaten zu ziehen. Der Einwurf der Münzen wurde in der Simulation jedoch nicht über diese Funktion realisiert, worüber die Teilnehmer während des Tests daraufhin informiert wurden. In einem Fall kam es zum Schließen des Browserfensters, nachdem eine Person auf der Tastatur die *Escape*-Taste gedrückt hat. Im Anschluss an die Testdurchführung wurden die Logfiledaten von allen Rechnern gespeichert und zu Erprobungszwecken ausgewertet. Durch die Auswertung ergaben sich weitere hilfreiche Hinweise in Bezug auf Verständnisschwierigkeiten, Programmierfehler und Messfehler, die entsprechend zu ändern waren.

Zu den vorgenommenen Änderungen gehört beispielsweise die Implementierung der Drag and Drop-Funktion in der Simulation des Parkscheinautomaten, um Münzen in den

Automaten werfen zu können. In der Simulation der Spülmaschine wurden die als Symbole verwendeten Buchstaben zur Kennzeichnung der Einstellungsmöglichkeit der Klarspülermenge geändert. Anlass war die in der Aufgabenstellung der Steuerungsphase geforderte Änderung des Wasserhärtegrades, für die der Großbuchstabe *H* verwendet wird. Die Probanden änderten jedoch die Klarspülmenge, die mit dem Kleinbuchstaben *h* gekennzeichnet war. Um zukünftig eine Verwechslung zu vermeiden, wird zur Kennzeichnung der Klarspülermenge der Buchstabe *d* verwendet.

In dem linken Bereich des Browserfensters, der für die itemspezifischen Instruktionen vorgesehen ist, wird zudem eine Information über die Anzahl der noch zu bearbeitenden Items ergänzt. In diesem Bereich wurden außerdem die Namen zweier Items derart geändert, dass sie keine Hinweise mehr auf das in der Steuerung zu erreichende Handlungsziel geben. Eine weitere Modifikation betraf die Anzahl der in dem Test verwendeten Items. Weil in der Hauptstudie der Einsatz weiterer Testinstrumente vorgesehen ist, muss die Testdauer für den entwickelten Problemlösetest verringert werden, indem die Itemmenge von 18 auf 15 reduziert wird. Für die Selektion von Items und die anschließende Anpassung der Itemreihenfolge wurden die relativen Lösungshäufigkeiten sowie die Trennschärfe der einzelnen Items in den beiden Gelegenheitsstichproben analysiert und als Anhaltspunkt verwendet. Die drei Items *Receiver*, *Online-Banking Portal* und der *Fahrscheinautomat* wurden aufgrund ihrer negativen beziehungsweise fehlenden Item-Skala-Korrelationen aus der Testzusammenstellung entfernt (Tabelle 8.5). Das Item *Radiowecker*, dessen Trennschärfe ebenfalls einen negativen Wert aufwies, wurde dennoch in dem Test belassen, da die negative Trennschärfe wahrscheinlich auf den Fehler in der Programmierung zurückzuführen ist, ebenso wie die geringe Trennschärfe des Items *Spülmaschine*.

Aufgrund der oben genannten Änderungen in der Simulation der Spülmaschine ist davon auszugehen, dass die Lösungshäufigkeit für dieses Item steigen wird, was dazu führt, dass der Anteil an Items, deren Lösungshäufigkeit unter 50 % liegt, geringer wird. Um dem entgegenzuwirken, wird versucht, durch eine Änderung in der Ausprägung der schwierigkeitsbestimmenden Merkmale Items schwieriger zu gestalten. Das Item *Radiowecker* wird aufgrund dessen intransparenter gestaltet, indem die Beschriftung zweier Bedienelemente „verkratzt“ dargestellt wird, so dass diese nicht mehr zu lesen sind. Für das Item *Internetrouter* wird die Aufgabenstellung der Steuerungsphase derart verändert, dass mehr Schritte zur Aufgabenbewältigung nötig werden.

Tabelle 8.5: relative Lösungshäufigkeiten der einzelnen Items und deren Trennschärfe (Ergebnisse aus der zweiten Erprobung)

Item	Relative Lösungshäufigkeit	Korr. Item-Skala-Korrelation	
Waschmaschine	.95	.15	
Receiver *	.90	-.13	
Fotoentwicklungsautomat	.85	.51	
Multifunktionsdrucker	.80	.27	
Radiowecker	.80	-.42	Mängelbeseitigung in der Simulation, Reduzierung der Transparenz
Kaffeemaschine	.75	.13	Mängelbeseitigung in der Simulation
Parkscheinautomat	.75	.31	
Fahrscheinautomat *	.70	-.07	
Küchenradio	.65	.34	
Online-Banking Portal *	.65	.10	
Navigationsgerät	.60	.66	
Internetrouter	.50	.28	Änderung in der Aufgabenstellung für die Steuerung: Um Energie zu sparen, möchten Sie das WLAN in der Nacht (00:00-05:00 Uhr) ausschalten. Außerdem möchten Sie in der Nacht nicht durch ein eventuell klingelndes Telefon gestört werden. Nehmen Sie dazu die entsprechenden Einstellungen an dem Internetrouter vor.
Fotodrucker	.50	.48	
Festnetztelefon	.45	.16	
Rollladensteuerung	.35	.54	
Heizungssteuerung	.30	.52	
Mobiltelefon	.20	.46	
Spülmaschine	.05	.02	Änderung in der Simulation

Anmerkung: * Items werden entfernt

8.6.2 Zweite Testerprobungsphase

Um sicherzugehen, dass mit der vorgenommenen Testrevision alle bekannten Mängel beseitigt wurden und keine neuen entstanden sind, wurde der geänderte Test noch einmal einer Stichprobe von 11 Teilnehmern eines weiteren Kraftwerksmeisterkurses zur Bearbeitung vorgelegt. Bei den ausschließlich männlichen Kursteilnehmern handelt es sich um eine Gruppe von Probanden im Alter zwischen 27 und 46 Jahren ($MW = 32.18$; $SD = 5.58$). Da seit der letzten Testerprobung eine neue Version des CBA ItemBuilders erschienen war (V 6.02), wurden die Items in dieser neuen Version abgespeichert und in der zugehörigen Execution Environment ausgeliefert. Als Browser wurde der Mozilla® Firefox® in der Version 35 und der Adobe® Flash® Player in der Version 16 installiert. Der zusammengestellte Test bestand nach der Entfernung von drei Items aus den folgenden 15 Items:

1. Waschmaschine
2. Parkscheinautomat
3. Spülmaschine
4. Fotoentwicklungsautomat
5. Multifunktionsdrucker
6. Kaffeemaschine
7. Küchenradio
8. Navigationsgerät
9. Fotodrucker
10. Radiowecker
11. Festnetztelefon
12. Rollladensteuerung
13. Heizungssteuerung
14. Mobiltelefon
15. Internetrouter

Für die Bearbeitung des Tests benötigten die Probanden erwartungsgemäß mit 1 h 38 min weniger Zeit als die in der zweiten Erprobung ermittelte Testdauer von 1 h 58 min. In der Phase der Testdurchführung konnten keine Mängel bezüglich der Simulationen, Aufgabenstellungen oder ähnliches festgestellt werden. Lediglich die Analyse der Logfile-Dateien deckte einen Softwarefehler der Execution Environment für die Items des CBA ItemBuilders Version 6.02 auf. In einigen Fällen dauerte die Speicherung der aufgezeichneten Verhaltensdaten so lange, dass diese in die Dateien der folgenden bereits aufgerufenen Task geschrieben wurden, wodurch die Aufzeichnungen der Folgetask nicht gespeichert wurden. Da mit einer Behebung des Softwarefehlers bis zum Zeitpunkt der Haupterhebung nicht sicher gerechnet werden konnte, wurden nach dieser dritten Testerprobung alle Items in der Version 5.02 des CBA ItemBuilders gespeichert und mit der entsprechenden Execution Environment in der Haupterhebung ausgeliefert.

8.7 Testinstruktion

Zur Gewährleistung der Durchführungsobjektivität soll die Instruktion zur Durchführung des Tests bei allen Anwendungen nahezu identisch sein. Die Informationen, die die Probanden vor der Testbearbeitung erhalten, sind durch die Erfahrungen in den drei Probedurchläufen entstanden und lassen sich stichpunktartig folgendermaßen zusammenfassen:

- Der Test an dem Sie heute teilnehmen, dient zur Erforschung der Kompetenz im Umgang mit technischen Systemen und wurde im Fach Technologie und Didaktik der Technik an der Universität Duisburg-Essen entwickelt.
- Dabei soll der Test nicht nur erfassen, ob jemand in der Lage ist, Probleme im Umgang mit unbekanntem technischen Geräten zu lösen, sondern auch, wie jemand dabei vorgeht.
- Der Test simuliert dazu verschiedene technische Geräte und ist außerdem in der Lage, jede Interaktion mit dem technischen Gerät aufzuzeichnen.
- Der Test ist absolut anonym. Ich kenne Sie nicht und kann hinterher keine Rückschlüsse auf Ihre Person ziehen.
- Der Test dauert ca. eineinhalb bis zwei Stunden ohne Unterbrechung.
- Der Test enthält zunächst einen Fragebogen, in dem Sie Daten über Ihre Person, wie Geburtsjahr, schulischer und beruflicher Bildungsweg und die Erfahrungen im Umgang mit Problemen angeben.
- Benutzen Sie nur die Tastatur, wenn Sie einen Text in dem Fragebogen eingeben müssen.
- Danach dürfen Sie die Tastatur auf keinen Fall mehr benutzen.
- Innerhalb des Fragebogens und innerhalb des ganzen Tests kommen Sie immer über einen blauen Button mit einem hellen Pfeil nach rechts zu der nächsten Seite oder Aufgabe.
- Sie können in dem Test nie zurück. Bitte versuchen Sie das auch nicht über die Tastatur. Der Test wird sonst gestoppt und Sie müssen noch einmal von vorne beginnen.
- Nach dem Fragebogen werden Ihnen nach und nach 15 simulierte technische Geräte präsentiert.
- Es handelt sich hierbei um technische Geräte, die aus dem Alltag oftmals bekannt sind.
- Da die Erfahrung mit solchen Geräten immer eine große Rolle bei dem Umgang spielt, bekommen Sie vor jeder Simulation zunächst eine Frage dazu gestellt. Mit dem Weiter-Button geht es dann weiter.
- Jedes dieser Geräte dürfen Sie dann zunächst ausprobieren und intensiv kennenlernen. Diese Phase nennt sich immer Explorationsphase. In dem Test wird die Aufgabenstellung immer auf der linken Seite angezeigt, während die Simulation auf der rechten Seite im Bildschirm erscheint.
- Der Test enthält technische Geräte, die sich nicht immer so verhalten wie Sie das erwarten würden. Es sind absichtlich auch ein paar Schwierigkeiten eingebaut. Wenn Sie also Schwierigkeiten bei der Bedienung haben, hat es nicht damit zu tun, dass der Test nicht richtig funktioniert oder das Gerät kaputt ist oder einen Programmierfehler enthält.
- Manche Buttons reagieren zeitverzögert, manche Buttons müssen mehrmals hintereinander gedrückt werden. Manche Elemente können per Drag & Drop (durch Klicken und Ziehen) verschoben werden. Manche Buttons lassen sich nur in bestimmten Zuständen betätigen und reagieren in anderen wiederum gar nicht. All das sollen Sie in dieser Explorationsphase erkunden.
- Nach dieser Probierphase, die Sie mit dem weiter-Button auch frühzeitig beenden können, wird Ihnen in der anschließenden Steuerungsphase eine konkrete Bedienungsaufgabe gestellt. Sie müssen also zum Beispiel eine konkrete Einstellung in dem Gerät vornehmen oder eine Funktion ausführen. Diese Steuerungsaufgabe sollte so schnell wie möglich und natürlich auf dem kürzesten Weg erfolgen, denn diese Wege und Zeiten fließen in das Ergebnis Ihrer Problemlösekompetenz ein. Je schneller Sie sind und je weniger Schritte Sie für die richtige Lösung eines Problems benötigen, desto besser wird das Testergebnis.
- Wenn Sie zwischendurch trinken oder essen möchten, tun Sie dies bitte nur zwischen den einzelnen Simulationen, wenn Ihre Erfahrungen zu dem jeweiligen Gerät abgefragt werden und auf keinen Fall während der Steuerungsphase, in der Sie die Aufgabe schnellstmöglich erledigen sollen.
- Manche technische Geräte spielen Musik ab. Damit Ihre Sitznachbarn nicht durch diese gestört werden, sind an den Rechnern Kopfhörer angeschlossen, die Sie gerne durch Ihre eigenen ersetzen dürfen, aber dann auch tragen sollten, um die Reaktionen der Geräte wahrnehmen zu können.
- Nun bleibt noch Zeit für Fragen, die wir Ihnen beantworten werden.

9 Datenerhebung

Zur Prüfung der Konstruktvalidität des *problemlösenden Umgangs mit technischen Alltagsgeräten* werden neben dem entwickelten Testinstrument, dessen Beschreibung in dem vorherigen Kapitel erfolgte, weitere Instrumente benötigt, die die angrenzenden Konstrukte erfassen. Im folgenden Abschnitt werden zunächst die Tests beschrieben, die zum Zweck der Validität des im Rahmen dieser Arbeit untersuchten Konstruktes eingesetzt werden. Im Anschluss erfolgt die Beschreibung der Population, in der die in Abschnitt 7 aufgestellten Hypothesen gelten sollen, sowie die Stichprobe, an der diese Hypothesen überprüft werden.

9.1 Weitere Testinstrumente

9.1.1 CFT 20-R

Fluide Intelligenz, als die Fähigkeit sich in neuen Situationen anzupassen, ist eine grundlegende Voraussetzung für den Erwerb von Wissen und damit für das Lösen neuartiger Probleme (Kapitel 5.1.1). Damit hat fluide Intelligenz den Charakter einer generellen Fähigkeit, die sich in Tests zeigt, in denen eine Anpassung an neue Situationen gefordert ist und in denen kristalline Intelligenz keinen Vorteil bietet (Rost, 2009, S. 52). Aufgaben, die sich zur Messung der generellen Intelligenz besonders eignen, sind die von Raven entwickelten Progressiven Matrizen, die ausschließlich bildhafte Figuren enthalten.

Eine solche sprachfreie und bildungs- bzw. kulturunabhängige Intelligenztestung soll auch mit dem ursprünglich von Cattell entwickelten Grundintelligenztest möglich sein (Guthke, 1996, S. 49f.), in dem sich solche Matrizenaufgaben wiederfinden. Er strebte eine Intelligenzmessung an, die nicht von soziokulturellen, erziehungsbedingten oder ethnischen Faktoren beeinflusst ist und wählte demnach für seinen Test die Bezeichnung *Culture Fair Test (CFT)* (Holling, Preckel & Vock, 2004, S. 89). Basierend auf dem Intelligenzkonzept von Cattell existiert eine Reihe Culture Fair Tests, mit denen fluide Intelligenz zuverlässig und valide gemessen werden können (Rost, 2009, S. 53). Die Grundintelligenz-Skala 1 (CFT 1) ist für Kinder bis zu einem Alter von 10 Jahren konzipiert (Schweizer, 2006b, S. 80). Der Einsatz der Grundintelligenz-Skala 2 (CFT 20 bzw. CFT 20-R) ist für Kinder ab 8 ½ Jahren und für Erwachsene bis 60 Jahren vorgesehen (Weiß, 2006), ebenso wie die Grundintelligenz-Skala 3, die im Gegensatz zur Skala 2 einen höheren Schwierigkeitsgrad aufweist (Holling, Preckel & Vock, 2004, S. 89; Schweizer, 2006b, S. 80).

Für die Erfassung fluider Intelligenz als Einflussfaktor auf die Kompetenz zum Lösen von Problemen wird in der vorliegenden Studie die Grundintelligenz-Skala 2 (CFT 20-R) verwendet. Die revidierte Fassung (CFT 20-R) der Skala 2 unterscheidet sich von der bisherigen Version CFT 20 in der Anzahl der Items, deren Anordnung und Reihenfolge sowie den Zeiten für die einzelnen Subtests (Jacobs & Petermann, 2007, S. 110). Der CFT 20-R, der als Gruppen- oder Einzeltest durchgeführt werden kann, besteht aus zwei Teilen, die sich jeweils aus vier Subtests zusammensetzen. Ein Subtest enthält verschiedene, in zeichnerischer Form dargestellte und nach Schwierigkeiten geordnete Einzelaufgaben (Weiß, 2006). Die Antworten sind jeweils aus einer Auswahl von fünf Antwortalternativen zu wählen und auf einem Antwortbogen einzutragen (Jacobs & Petermann, 2007, S. 109):

1. Reihenfortsetzen

Eine Reihe von drei Bildern soll logisch fortgeführt werden. Als Antwort stehen fünf Bildalternativen zur Auswahl (Abbildung 9.1).

2. Klassifikationen

In einer Reihe von fünf Bildern, soll dasjenige ausgewählt werden, das eine wesentliche Eigenschaft nicht mit den anderen teilt.

3. Matrizen

In vorgegebenen Vier-Felder und Neun-Felder-Tafeln ist ein Feld durch die Auswahl aus fünf Antwortalternativen logisch zu ergänzen.

4. Topologien

In Bildern mit abstrakten Formen sind ein oder mehrere Punkte eingezeichnet. Aus den vorgegeben Antwortbildern ist dasjenige auszuwählen, in dem ebenfalls ein oder mehrere Punkte nach analogen Vorgaben zu dem vorgegeben Bild eingezeichnet werden könnten.



Abbildung 9.1: Ein Beispielitem des Untertests aus dem CFT 20-R (Grundintelligenztest Skala 2 Revision; Weiß (2006)); die Aufgabenstellung lautet die Figurenreihe logisch richtig fortzusetzen (Hornke, Amelang & Kersting, 2011, S. 21)

Für den ersten Testteil, der insgesamt 56 Items (15 Items Reihenfortsetzen, 15 Items Klassifikationen, 15 Items Matrizen und 11 Items Topologien) umfasst, ist eine exakte Testzeit von 14 Minuten (4-4-3-3) vorgesehen, die um 1 Minute je Subtest für leistungsschwache Probanden verlängert werden kann. Der zweite Testteil enthält insgesamt 45 Items (12 Items Reihenfortsetzen, 12 Items Klassifikationen, 12 Items Matrizen und 9 Items Topologien) wodurch sich eine exakte Testzeit von 12 Minuten (3-3-3-3) ergibt. Bei Berücksichtigung der allgemeinen Testeinführung und der einzelnen Subtestinstruktionen ist für den ersten Testteil eine Testdauer von 37 Minuten (ohne Testzeitverlängerung) und für den zweiten Testteil eine Testdauer von 25 Minuten einzuplanen. Um die Gesamttestdauer, die sich aus den einzelnen zu einer Testbatterie zusammengestellten Tests ergibt, möglichst kurz zu halten, wird von der Möglichkeit Gebrauch gemacht, eine Kurzform des CFT 20-R zu verwenden. Diese Kurzform besteht allein aus dem Testteil 1, womit der Nachteil einer verringerten Reliabilität verbunden ist. Da diese für den Teil 1 mit $r_{\text{Teil1}} = .92$ immer noch sehr hoch ausfällt ($r_{\text{gesamt}} = .96$), ist der Nachteil vertretbar (Weiß, 2006).

9.1.2 Selbstwirksamkeitserwartungen

Die Kompetenz, erfolgreich in unbekanntem Situationen agieren und neue Probleme lösen zu können, erfordert neben kognitiven Voraussetzungen stets auch motivationale Aspekte (Kapitel 5.1.3). Ausubel, Novak & Hanesian (1981, S. 664) stellen fest, dass die Einstellungen erfolgreicher Problemlöser gegenüber dem Nachdenken positiver ist, sie mehr Selbstvertrauen in die eigene Fähigkeit Probleme zu lösen haben und sich durch auftretende Schwierigkeiten nicht so leicht ermutigen lassen. Denn eine hohe Selbstwirksamkeit erhöht die Anstrengungsbereitschaft und die Ausdauer bei der Bewältigung neuer Aufgaben und

Probleme und führt so zu einer größeren Erfolgswahrscheinlichkeit. Personen, die über ein geringeres kognitives Potenzial verfügen, sich aber durch eine erhöhte Ausdauer und eine größere Arbeits- und Leistungsbereitschaft auszeichnen, sind in vielen Fällen erfolgreicher als Personen, die zwar über das kognitive Potenzial verfügen, jedoch wenig zielorientiert und motiviert sind (OECD, 2013b, S. 64).

Die Erwartungen an die Wirksamkeit des eigenen Handelns basieren größtenteils auf direkten eigenen Erfahrungen und sind demnach vor allem von den spezifischen Anforderungen in einer Situation abhängig (Wirth, 2004). Um den Einfluss, den die Selbstwirksamkeitserwartung auf die Leistungen im problemlösenden Umgang mit unbekanntem technischen Systemen ausübt, zu prüfen, muss die Selbstwirksamkeit aufgabenbezogen, das heißt in Bezug zum Problemlösen erfasst werden. Die in der PISA Studie 2012 (OECD, 2013b, S. 65) erstmalig eingesetzten Skalen *Perseverance* (Ausdauer) und *openness to problem solving* (Aufgeschlossenheit) erscheinen hierfür geeignet und werden deshalb für die Erfassung der Selbstwirksamkeitserwartungen in dieser Arbeit mit einigen wenigen Modifikationen übernommen.

PISA misst die *Ausdauer* der Probanden in Bezug auf das Bearbeiten von Aufgaben bzw. das Lösen von Problemen anhand ihrer Zustimmung auf die Aussage nach der sie glauben einer Person ähnlich zu sein, die schnell aufgibt, wenn sie mit einem Problem konfrontiert wird; die schwierige Probleme beiseiteschiebt; die in Aufgaben, die sie anfängt, interessiert bleibt; die solange an einer Aufgabe arbeitet, bis alles perfekt ist und die mehr tut, als man von ihr erwartet, wenn sie mit einem Problem konfrontiert wird (vgl. OECD, 2013b, S. 65). Die Zustimmung auf diese Aussagen kann zwischen: Diese Person ist „fast wie ich“, „mir sehr ähnlich“, „mir ein bisschen ähnlich“, „mir nicht ähnlich“ oder „überhaupt nicht so wie ich“ variieren. Die *Aufgeschlossenheit gegenüber dem Problemlösen* wird in PISA über die Zustimmung auf die Aussagen nach der die Probanden finden, sie seien einer Person ähnlich, die mit einer großen Menge an Informationen umgehen kann, die Dinge sehr schnell versteht, die versucht Erklärungen für Dinge zu suchen, die Fakten einfach miteinander verbinden kann und die es mag komplexe Probleme zu lösen (vgl. OECD, 2013b, S. 67). Für die in PISA verwendete Skala *Perseverance* (Ausdauer) wird eine Reliabilität von $r = .67$ (für Deutschland) und von $r = .80$ für den OECD Median berichtet. Für die in PISA verwendete Skala *Openness to problem solving* (Aufgeschlossenheit) wird eine Reliabilität von $r = .81$ (für Deutschland) und von $r = .82$ für den OECD Median berichtet (OECD, 2014a, S. 336).

Die Modifikationen der Skalen, die für diese Arbeit verwendet werden, betreffen zum einen die eigene Übersetzung der Wortlaute der Itemaussagen und der Antwortalternativen ins Deutsche sowie die Reduzierung der Antwortstufen von fünf auf vier Antwortalternativen, um einen nicht eindeutig zu interpretierenden Mittelpunkt der Merkmalsausprägung zu vermeiden (Tabelle 9.1).

Tabelle 9.1: Itemwortlaut der Skalen Ausdauer und Aufgeschlossenheit gegenüber Problemlösen (vgl. OECD, 2013b)

Item	Itemwortlaut ^a
<i>Ausdauer</i>	
SWE_1_1	Wenn ich mit einem Problem konfrontiert werde, gebe ich schnell auf.
SWE_1_2	Schwierige Probleme schiebe ich oft beiseite.
SWE_1_3	Ich bleibe interessiert bei den Aufgaben, die ich anfangen.
SWE_1_4	Ich arbeite solange an einer Aufgabe, bis alles perfekt ist.
SWE_1_5	Wenn ich mit einem Problem konfrontiert werde, tue ich mehr als von mir erwartet wird.
<i>Aufgeschlossenheit gegenüber Problemlösen</i>	
SWE_2_1	Ich kann mit einer großen Menge an Informationen umgehen.
SWE_2_2	Ich verstehe Dinge sehr schnell.
SWE_2_3	Ich versuche Erklärungen für Dinge zu suchen.
SWE_2_4	Ich kann einfach Fakten miteinander verbinden.
SWE_2_5	Ich mag es Probleme zu lösen.

Anmerkungen: ^a Einleitender Text: Wie gut treffen die folgenden Aussagen auf Sie zu?; SWE = Selbstwirksamkeitserwartung

9.1.3 Need for Cognition

Neben dem Bedürfnis nach Selbstvertrauen und Selbstwirksamkeit führt auch das Bedürfnis, das eigene Potenzial auszuschöpfen, Ziele erreichen zu können und das Bedürfnis nach Neugier zur intrinsischen Motivation in einer Situation (Kapitel 5.1.3). Je stärker das Motiv der kognitiven Neugier desto stärker das Bedürfnis nach Erkenntnisgewinn. Die Tendenz einer Person sich mit dem Denken zu beschäftigen und Spaß an dieser Tätigkeit zu haben, beschreiben Cacioppo & Petty (1982) mit dem Konstrukt Need for Cognition. Weil angenommen wird, dass Need for Cognition ein Motiv ist, das sich insbesondere in der Motivation zur Bewältigung kognitiver Anforderungen und dem Lösen von Problemen zeigt, soll sein Einfluss in dieser Studie untersucht werden.

Für die Erfassung von Need for Cognition haben Cacioppo & Petty (1982) eine Skala vorgestellt, die zunächst 45 Items beinhaltete. Das Ausmaß an Zustimmung zu den präsentierten Aussagen können die Probanden auf einer neunstufigen Likert-Skala angeben (von +4 = „very strong agreement“ bis -4 = „very strong disagreement“). Diese Langversion wurde im Zuge ihrer Untersuchung auf 34 Items reduziert (Cacioppo & Petty, 1982, S. 123). In einer weiteren Studie erfolgte für eine Kurzversion der Skala Need for Cognition eine weitere Reduzierung der Anzahl auf 18 Items (Cacioppo, Petty & Kao, 1984). Eine Übersetzung der ursprünglichen Langversion der englischen Need for Cognition Skala nahmen die Autoren Bless, Fellhauer, Bohner et al. (1991) vor und reduzierten ihn ähnlich wie Cacioppo & Petty (1982) anhand der Eliminierung schwach ladender Items auf 33 Items. Die Antworten wurden mittels einer siebenstufigen Likert-Skala (von +3 = „trifft ganz genau zu“ bis -3 = „völlig unzutreffend“) erfasst. Um eine möglichst effiziente Erfassung zu ermöglichen erstellten Bless, Wänke, Bohner et al. (1994) aus 16 der enthaltenen Items eine Kurzform der Need for Cognition Skala. Eine Validierung der deutschen Version der Need for Cognition Skala erfolgte anhand positiver Zusammenhänge zwischen Need for Cognition und

wissenschaftlichem Interesse und Leistungsmotivation sowie gefundenen negativen Zusammenhängen zwischen Need for Cognition und Misserfolgsmotivation und der überdauernden Stimmungslage (Beißert, Köhler, Rempel et al., 2014, S. 5).

Eine weitere Adaption nimmt Preckel (2014) mit dem Ziel eine deutschsprachige Need for Cognition Skala für Kinder ab 10 Jahren und Jugendliche zu entwickeln vor. Hierfür entnimmt er der Kurzversion der deutschen Need for Cognition Skala 15 Items und weitere 4 Items aus einer französischen Version der Skala für Kinder. Die Antworten auf die insgesamt 19 Items werden mit einer fünfstufigen Skala (von 1 = trifft nicht zu bis 5 = trifft zu) erfasst (Tabelle 9.2). Für die Reliabilität der Skala werden je nach Stichprobe Werte zwischen $r = .89$ und $.92$ berichtet.

Tabelle 9.2: Itemwortlaut der Skala Need for Cognition (vgl. Preckel, 2014)

Item	Itemwortlaut ^a
NFC_1	Es macht mir viel Spaß, mir Lösungen für Probleme auszudenken.
NFC_2	Ich würde lieber eine wichtige Aufgabe lösen, die schwierig ist und Nachdenken erfordert, als eine Aufgabe, die zwar wichtig ist, aber nicht viel Nachdenken erfordert.
NFC_3	Ich mag Situationen, in denen ich mit gründlichem Nachdenken etwas erreichen kann.
NFC_4	Ich mag keine Situationen, in denen ich mich auf mein Denken verlassen muss, um etwas zu erreichen.
NFC_5	Für mich ist es besonders schön, wenn ich eine wichtige Aufgabe erledigt habe, die viel Nachdenken erfordert hat.
NFC_6	Ich würde lieber etwas tun, bei dem ich wenig nachdenken muss, als etwas, bei dem ich viel nachdenken muss.
NFC_7	Ich mag es nicht, angestrengt und stundenlang nachzudenken.
NFC_8	Ich denke nur nach, wenn ich muss.
NFC_9	Nachdenken macht mir keinen Spaß.
NFC_10	Ich mag keine Situationen, in denen ich intensiv über etwas nachdenken muss.
NFC_11	Ich habe es gern, wenn mein Leben voller kniffliger Aufgaben ist, die ich lösen muss.
NFC_12	Ich mag komplizierte Probleme lieber als einfache Probleme.
NFC_13	Gibt es ein Problem, dann genügt es mir, einfach die Lösung zu kennen; mir ist es nicht wichtig, die Gründe für die Lösung des Problems zu verstehen.
NFC_14	Mir genügt es, dass etwas funktioniert und mir ist egal, wie oder warum es funktioniert.
NFC_15	Ich sage mir oft, dass man gut und lange nachdenken muss, um die beste Lösung für ein Problem zu finden.
NFC_16	Ich erledige gerne Aufgaben, bei denen man viel nachdenken muss.
NFC_17	Ich bin jemand, der sehr gerne nachdenkt.
NFC_18	Ich denke gerne über ein Problem nach, selbst wenn ich weiß, dass mein Nachdenken an dem Problem nichts ändern wird.
NFC_19	Wenn ich mir in den Kopf setze, die Lösung für ein schwieriges Problem zu finden, schaffe ich das auch oft.

Anmerkungen: ^a Einleitender Text: Wie gut treffen die folgenden Aussagen auf Sie zu?; NFC = Need for Cognition

9.1.4 ICT-Literacy und gerätespezifisches Vorwissen

Das in dieser Arbeit im Fokus stehende Konstrukt des problemlösenden Umgangs mit technischen Alltagsgeräten wurde in Abschnitt 6 als kontextuelle Problemlösekompetenz spezifiziert. Situative Faktoren innerhalb eines Kontextes beeinflussen das Handeln und Verhalten einer Person ebenso wie die oben bereits genannten kontextübergreifenden Fähigkeiten und Motive (Kapitel 5.2.2). Handeln erfordert Aufmerksamkeits- und Wahrnehmungsprozesse, die stark von den eigenen Erfahrungen und Handlungsgewohnheiten gelenkt werden. Aber auch die Verarbeitung und Speicherung neuer Informationen ist erst durch die Verknüpfung an bereits vorhandene Wissensstrukturen möglich. Aufgrund des angenommenen Einflusses des Vorwissens auf die technische Problemlösekompetenz soll das Vorwissen im Rahmen der durchgeführten Studie ebenfalls erfasst werden. Da unter Vorwissen all das problembezogene Wissen zu verstehen ist, das einer Person vor Beginn des Problemlöseprozesses zur Verfügung steht und damit auch weniger strukturierte Wissensaspekte wie Alltagserfahrungen gemeint sind, ist die Erfassung des relevanten Vorwissens nicht einfach.

Weil sowohl geräteübergreifendes als auch gerätespezifisches Wissen als Vorwissen für den problemlösenden Umgang mit technischen Alltagsgeräten relevant sein können, werden diese Kategorien in dieser Arbeit getrennt voneinander erfasst. Als geräteübergreifendes Vorwissen, das relevant für den Umgang mit den simulierten technischen Systemen angenommen wird, ist die Fähigkeit im Umgang mit Informations- und Kommunikationstechnologien (ICT-Literacy). ICT-Literacy stellt eine wichtige Fähigkeit dar, „um alltägliche informationsbezogene Anforderungen und Problemstellungen erfolgreich bewältigen zu können“ (Goldhammer, Kröhne, Keßel et al., 2014, S. 10). Vier der simulierten technischen Geräte sind dem Bereich der ICT entnommen, so dass von einem Einfluss der ICT-Literacy auszugehen ist. Um diese zu messen, können entweder Selbstbeurteilungen der Probanden oder papier- bzw. computerbasierte Tests verwendet werden (ebd).

Aus testökonomischen Gründen wird für die Erfassung der ICT-Literacy in dieser Studie das Verfahren der Selbstbeurteilung mit einem Fragebogen gewählt. Ein solcher Fragebogen wurde in der PISA-Studie 2012 (OECD, 2013a, S. 248) eingesetzt. Der *ICT Familiarity Questionnaire* umfasst dabei fünf verschiedene Aspekte: *Verfügbarkeit von ICT, Generelle Computernutzung, ICT-Nutzung außerhalb der Schule, ICT-Nutzung in der Schule und Einstellungen gegenüber von Computern als Hilfsmittel für die Schule* (ebd.). Aus diesem Fragebogen wird lediglich die Skala *ICT-Nutzung außerhalb der Schule* verwendet und der Schulbezug aus der Itemformulierung herausgenommen. Die Antworten der Probanden werden anhand einer fünfstufigen Skala erfasst (1 = niemals bis 5 = täglich) (Tabelle 9.3). Für die aus 10 Items bestehende Skala wird eine Reliabilität von $r = .75$ für die deutsche Stichprobe und $r = .78$ für den OECD Median berichtet (OECD, 2014b, S. 338).

Tabelle 9.3: Itemwortlaut der Skala ICT – Nutzung (vgl. OECD, 2013a)

Item	Itemwortlaut ^a
ICT_1	Spielen von Computerspielen (für Einzelspieler)
ICT_2	Spielen von Onlinespielen (mehrere Spieler erforderlich)
ICT_3	Nutzen von E-Mails
ICT_4	Online-Chats
ICT_5	Nutzen von sozialen Netzwerken
ICT_6	Internetsurfen aus Spaß (z. B. Ansehen von Videos bei Youtube)
ICT_7	Lesen von tagesaktuellen Nachrichten
ICT_8	Beschaffen von Informationen aus dem Internet (Locations, Events...)
ICT_9	Download von Musik, Filme, Spiele oder Software
ICT_10	Upload von eigenen erstellten Inhalten zum Teilen (Podcasts, Musik, Software etc.)

Anmerkungen: ^a Einleitender Text: Wie oft nutzen Sie den Computer für die folgenden Aktivitäten?; ICT = Informations- und Kommunikationstechnologien

Das gerätespezifische Vorwissen lässt sich nicht vollständig erfassen. Als relevant für den Umgang mit technischen Geräten ist vor allem das sogenannte Eingriffswissen, das als prozedurales Wissen in Form von Wenn-Dann Regeln repräsentiert ist. Dieses Wissen vor der Bearbeitung abzufragen ist kaum möglich und vor dem Hintergrund einer ökonomischen Testdurchführung auch nicht anzuraten. Da für den Erwerb solcher Prozeduren Übung nötig ist, die in der Regel im Alltag durch Erfahrungen im Umgang mit den Geräten erworben wird, kann diese Erfahrung durch Items erfasst werden.

Wieviel Erfahrung eine Person im Umgang mit einem spezifischen technischen Gerät (wie z. B. einem Backofen) hat, wird für jedes der in dem Problemlösetest simulierten technischen Systeme mit einem Einzelitem (Single-Item) erfasst. Der Erhebungsaufwand dieser Single-Item-Messungen ist sehr gering, wodurch nicht die Gefahr eines Motivationsverlustes seitens der Probanden durch redundante Items besteht (Weiber & Mühlhaus, 2014, S. 112). Bei Single-Items wird das zu erfassende Konstrukt selbst in die Frageformulierung aufgenommen und dessen Ausprägung mit Hilfe einer Ratingskala erhoben (ebd.). Die Ausprägung der gerätespezifischen Erfahrung wird für jedes Gerät mit einer drei- bis vierstufigen Ratingskala erfasst, deren Abstufungen zusätzlich durch konkrete Falldarstellungen verdeutlicht werden (Tabelle 9.4).

Tabelle 9.4: Itemwortlaut der Skala Erfahrungen im Umgang mit technischen Geräten

Waschmaschine
Wie vertraut sind Sie im Umgang mit Waschmaschinen?
<ul style="list-style-type: none"> • Ich weiß nicht, wie eine Waschmaschine zu betätigen ist. Jemand anderes wäscht meine Wäsche. • Ich bediene eine Waschmaschine regelmäßig und nutze hierbei generell die Grundfunktionen (max. drei verschiedene Programme). • Ich habe schon verschiedene Waschmaschinenmodelle bedient und kenne mich gut mit ihren Funktionen aus.

Parkscheinautomat

Wie häufig müssen Sie in Ihrem Alltag einen Parkscheinautomaten bedienen?

- Ich nutze die öffentlichen Verkehrsmittel und musste noch nie mit einem Parkscheinautomaten umgehen.
 - Ich parke ab und zu in parkscheinpflichtigen Zonen in meiner Region und bin in der Lage die dort vorhandenen Parkscheinautomaten zu bedienen.
 - Ich bin sehr oft mit dem Auto unterwegs und habe bereits Parktickets an unterschiedlichen Parkscheinautomaten in verschiedenen Städten gekauft.
-

Spülmaschine

Wie bewerten Sie Ihre Fähigkeiten im Umgang mit Spülmaschinen?

- Ich habe noch nie eine Spülmaschine betätigt.
 - Ich bin in der Lage die Grundprogramme einer Spülmaschine auszuwählen und diese einzustellen.
 - Ich kenne meine Spülmaschine sehr gut und weiß auch, wie ich benutzerdefinierte Einstellungen vornehmen kann.
 - Ich musste bereits viele verschiedene (ältere und neuere Modelle) von Spülmaschinen bedienen und vorhandene Einstellungen anpassen.
-

Fotoentwicklungsautomat

Wie vertraut sind Sie im Umgang mit Selbstbedienungsautomaten für die Fotoentwicklung?

- Ich habe noch nie einen Fotoentwickler bedient.
 - Ich habe bereits (mehrmals) einen Fotoentwicklungsautomaten bedient. Ich habe jedoch nicht alle Entwicklungsmöglichkeiten ausprobiert.
 - Ich habe bereits verschiedene Möglichkeiten der Fotoentwicklung (vorherige Bildbearbeitung, Posterentwicklung etc) an verschiedenen Fotoentwicklungsautomaten genutzt.
-

Multifunktionsdrucker

Welche Erfahrungen haben Sie bereits im Umgang mit sogenannten Multifunktionsgeräten für das Kopieren, Scannen und Drucken von Dokumenten gemacht?

- Bislang habe ich noch keine Erfahrung mit ihnen.
 - Ich nutze hin und wieder ein Multifunktionsgerät zum Kopieren.
 - Ich nutze ein Multifunktionsgerät, dessen Funktion ich sehr gut kenne.
 - Ich kenne verschiedene Modelle und nutze sie regelmäßig. Ich kann jede Funktion von ihnen bedienen.
-

Kaffeemaschine

Wie vertraut sind Sie im Umgang mit Kaffeefullautomaten?

- Ich trinke keinen Kaffee bzw. ich trinke zwar Kaffee, bereite ihn aber nicht selbst zu.
 - Ich trinke Kaffee, den ich mir aber mithilfe einer halbautomatischen Kaffeemaschine zubereite (Filterkaffeemaschine, Pad- oder Kapselmaschine).
 - Ich habe bereits einen Kaffeefullautomaten bedient und kenne die Grundfunktionen.
 - Ich bin mit unterschiedlichen Kaffeefullautomaten vertraut und bin auch in der Lage kleinere Störungen des Automaten zu beheben.
-

Küchenradio

Wie häufig nutzen Sie in Ihrem Alltag ein Küchenradio?

- Ich besitze kein Küchenradio und kenne auch nicht den Unterschied zu einem gewöhnlichen Radio.
 - Ich habe schon einmal ein Küchenradio genutzt, aber nur um Musik zu hören. Andere Funktionen habe ich noch nie ausprobiert.
 - Ich nutze ein Küchenradio regelmäßig und kann zusätzliche Funktionen, ohne in die Beschreibung zu schauen, bedienen.
-

Navigationsgerät

Wie schätzen Sie Ihre Fähigkeiten im Umgang mit Kraftfahrzeug-Navigationssystemen aus?

- Ich habe noch nie ein Navigationsgerät bedient (weder in einem Fahrzeug noch in einem mobilen Gerät).
 - Ich habe bereits ein Navigationsgerät in einem mobilen Gerät (z. B. Smartphone) genutzt aber noch nie in einem Fahrzeug.
 - Ich habe bereits ein Navigationsgerät in einem Fahrzeug bedient und kenne die zum Navigieren erforderlichen Grundfunktionen.
 - Der Umgang mit verschiedenen Modellen von Navigationsgeräten ist mir vertraut. Ich nutze nicht nur die Grundfunktionen, sondern auch erweiterte Einstellungen.
-

Fotodrucker

Wie bewerten Sie Ihre Fähigkeit bezüglich der Bedienung von Fotodruckern?

- Ich besitze keinen eigenen Fotodrucker. Wenn ich Bilder auf Papier erhalten möchte, lasse ich sie entwickeln.
 - Ich habe bereits mehrmals mit einem Fotodrucker Fotos gedruckt, nutze aber nicht alle Funktionen des Fotodruckers.
 - Ich weiß mit verschiedenen Fotodruckermodellen umzugehen und verfüge über genug Erfahrung, meine Fotos nach Bedarf auch anpassen zu können.
-

Radiowecker

Wie häufig benutzen Sie in Ihrem Alltag einen Radiowecker?

- Ich benutze noch einen analogen Wecker und weiß nicht, wie ein Radiowecker zu bedienen ist.
 - Ich benutze den Wecker als Zusatzfunktion in einem anderen Gerät (Mobiltelefon, Stereoanlage etc.).
 - Ich verwende einen Radiowecker, dessen Einstellungen ich nicht oft verändere.
 - Ich kenne verschiedene Radioweckertypen und bin imstande die meisten ihrer Funktionen zu nutzen.
-

Festnetztelefon

Wie kompetent sind Sie im Umgang mit Festnetztelefonen?

- Ich habe kein Festnetztelefon, sondern telefoniere ausschließlich über mein Handy.
 - Ich nutze regelmäßig ein Telefon und kenne mich mit den Grundfunktionen aus.
 - Ich kenne verschiedene Telefonmodelle mit ihren verschiedenen Einstellungsmöglichkeiten und kann sie ohne Probleme bedienen.
-

Rollladensteuerung

Wie bewerten Sie Ihre Fähigkeit in der Bedienung von Rollladensteuerungen?

- Ich habe noch nie eine Rollladensteuerung bedient.
 - Ich bin mithilfe von Rollladensteuerung imstande, die Rollläden manuell hoch und runter zu fahren.
 - Ich bin nicht nur in der Lage die Grundfunktionen meiner Rollladensteuerung zu nutzen, sondern sie auch nach meinen Bedürfnissen einzurichten.
-

 Heizungssteuerung

Wie hoch ist Ihre Fähigkeit bei der Verwendung von Heizungssteuerungen?

- Ich habe noch nie eine Heizungssteuerung bedient. Ich stelle meine Heizkörper nach Bedarf manuell ein.
 - Ich verwende eine Heizungssteuerung für meine Heizkörper und kenne ihre Grundfunktionen.
 - Ich habe bereits verschiedene Heizungssteuerungen bedient und kann sie nach meinen Bedürfnissen einrichten.
-

 Mobiltelefon

Wie schätzen Sie Ihre Fähigkeit im Umgang mit Handys ein?

- Ich besitze zwar ein Handy, verwende aber nur Grundfunktionen wie die des Telefonierens und Versendens von Nachrichten.
 - Ich kenne mein Handy sehr gut und kann es nach meinen Bedürfnissen entsprechend einrichten und nutzen.
 - Ich bin fähig, viele verschiedene (ältere und neuere) Handymodelle zu bedienen und kann Anderen bei Problemen mit diesen Geräten helfen.
-

 Internetrouter

Wie schätzen Sie Ihre Fähigkeiten im Umgang mit Internetroutern ein?

- Ich habe zu Hause kein Internet und deswegen auch keinen Internetrouter.
 - Ich nutze zwar einen Internetrouter, weiß aber nicht wie man ihn einrichtet oder dessen Einstellungen ändern kann.
 - Ich habe einen Internetrouter bereits eingerichtet und kann manche Änderungen in seinen Einstellungen vornehmen.
 - Ich kenne mich mit verschiedenen Internetroutern und ihren Verwendungsmöglichkeiten sehr gut aus. Meine Kenntnisse reichen über die eines durchschnittlichen Anwenders hinaus.
-

9.1.5 Allgemeine Problemlösefähigkeit

Ein weiteres Konstrukt, von dem ein Einfluss auf den Umgang mit technischen Geräten angenommen wird, ist die allgemeine bzw. fachübergreifende Problemlösefähigkeit. Sie beschreibt eine Fähigkeit, die zur Lösung einer großen Anzahl verschiedener Probleme eingesetzt werden kann und damit nicht problemspezifisches Wissen und Strategien, sondern allgemeinere Denk- und Problemlösestrategien. Für die Erfassung solcher allgemeinen Strategien wurden ausgehend von der Kritik an den, in den Anfängen komplexer Problemlöseforschung eingesetzten realitätsnahen Computerszenarien, von Funke (1992a) vorwissensfreie, komplexe Szenarien auf Basis formaler Modelle¹ entwickelt. Bei der Entwicklung dieser Szenarien wird zunächst die Struktur festgelegt und erst im Anschluss erfolgt die inhaltliche Einbettung. Mängel bezüglich der fehlenden Objektivität bei der Bewertung der Lösungsgüte sowie der fehlenden Operationalisierung der Problemmerkmale konnten damit zwar behoben werden, aber erst mit den von Greiff (2012) zu minimal komplexen Szenarien weiterentwickelten Problemstellungen konnte das Problem fehlender Reliabilitätsangaben der vorherigen One-Item-Messungen gelöst werden.

Nach Art der verwendeten Formalismen werden die zur Erfassung allgemeiner Problemlösefähigkeit entwickelten Problemszenarien *MicroDYN* und *MicroFIN* genannt. Ein Item wird als MicroDYN-Item bezeichnet, wenn die dahinterliegende Struktur ein lineares

¹ Formale Modelle werden im Abschnitt 4.3 beschrieben.

Gleichungssystem ist. Erfolgt die Entwicklung eines Mikroitems auf Basis finiter Automaten wird es als MicroFIN-Item bezeichnet (Neubert, Kretzschmar, Wüstenberg et al., 2015). Da auch die im Rahmen dieser Arbeit entwickelten Items zur Erfassung technischer Problemlösekompetenz finite Automaten zur Grundlage haben, soll die allgemeine Problemlösefähigkeit mit einem Test bestehend aus MicroFIN-Items erfasst werden. Hierfür wird auf bereits von der Forschergruppe um Greiff entwickelte Items zurückgegriffen.

Der zusammengestellte allgemeine Problemlösetest umfasst ein Beispielszenario sowie sechs weitere Items zur Bearbeitung (Tabelle 9.5). Jedes der präsentierten Szenarien muss dabei von den Probanden zunächst exploriert werden. Das Wissen, das die Probanden in dieser maximal vierminütigen Explorationsphase erwerben, wird im Anschluss abgefragt. Die Wissensabfrage setzt sich aus vier Verifikationsaufgaben (Kapitel 3.2) sowie aus zwei Identifikationsaufgaben zusammen (Buchner & Funke, 1993). In den Verifikationsaufgaben ist für jeden der präsentierten Zustände zu entscheiden, ob es sich um einen tatsächlich aufgetretenen Zustand handelt. In den Identifikationsaufgaben muss anhand einer präsentierten Intervention und dem sich daraus resultierenden Zustand der ursprüngliche Ausgangszustand angegeben werden. Für die Identifizierung der Ausgangszustände werden Informationsvorschläge unterhalb des Szenarios dargeboten, von denen dann die richtigen Informationen an die entsprechende Position im Szenario zu bringen sind. Im Anschluss an die Wissensabfrage erfolgt die Steuerungsphase, in der die Probanden zwei nacheinander vorgegebene Ziele in dem Szenario mit möglichst wenigen Bearbeitungsschritten erreichen sollen. Im Anschluss an die beiden Steuerungsaufgaben folgt die Explorationsphase des nächsten Szenarios.

Tabelle 9.5: Beschreibung der Szenarien des Tests zur Erfassung allgemeiner Problemlösefähigkeit

Szenario	Kurzbeschreibung
Fernbedienung (Instruktionsitem)	Die Probanden müssen herausfinden, wie die gezeigte Fernbedienung zu bedienen ist.
1 Gewächshaus	Herauszufinden ist, welche Düngungsmittel in Abhängigkeit der Jahreszeit zu welchen Erträgen bei dem Anbau von Kürbissen in einem Gewächshaus führen.
2 Aquarium	Es sollen die Einflüsse der Bestandteile im Wasser des gezeigten Aquariums auf den Bestand der in ihm lebenden Fische entdeckt werden.
3 Planomat	Durch die Veränderung von Gebäudestandorten sollen Kenntnisse über die Passung von Gebäuden zueinander gewonnen werden.
4 Konzertomat	Durch die Variation von Musikart, Bühnenart und Eintrittspreis soll der Einfluss dieser Variablen auf die Zuschaueremenge bei einem Konzert bestimmt werden.
5 Pflanzomat	Herauszufinden ist, inwieweit Kombinationen aus Wasser-, Licht- und Düngemenge auf die Blätter- und Blütenmenge eines Baumes Einfluss haben.
6 Waschomat	Durch die Verstellung farbiger Regler und dem Drücken von Tasten für den Verschmutzungsgrad von Wäsche an einer Waschmaschine, soll der Einfluss auf die Sauberkeit der gewaschenen Wäsche sowie die Umweltverträglichkeit bestimmt werden.

Die Explorationsphase wird nicht prozessorientiert, sondern ausschließlich ergebnisorientiert ausgewertet. Zwar werden die Anzahl der, in dieser Phase vorgenommenen Interaktionen sowie die Zeit, innerhalb der mit dem Szenario in dieser Phase interagiert wird, aufgezeichnet, diese Daten fließen jedoch nicht in die Rohwertermittlung ein. Stattdessen wird das Wissen, das nach der Explorationsphase über das explorierte Szenario vorhanden ist, abgefragt. Die Wissensabfrage erfolgt dabei durch zwei unterschiedliche Aufgabentypen: *state identification* und *initial state*. Der Aufgabentyp *state identification* zu jedem Szenario enthält insgesamt vier Items, bei denen die Probanden für jeden der gezeigten Systemzustände angeben sollten, ob dieser in dem Szenario auftreten kann. Die Antwort wird durch Klicken auf *Ja* oder *Nein* gegeben und dichotom ausgewertet. Aus diesen vier Items werden je zwei Items zu einem Subscore zusammengefasst¹. Nur wenn beide der zusammengefassten Items richtig gelöst sind, wird der jeweilige Subscore als *gelöst* (1 Punkt) gewertet, sobald eines der Items falsch beantwortet wird, wird der Subscore als *nicht gelöst* gewertet (0 Punkte). Aus diesen beiden Subscores wird das arithmetische Mittel gebildet. Der Gesamtscore für den Aufgabentyp *state identification* eines jeden Szenarios kann demnach die Werte 0; 0.5 und 1 annehmen. Der Aufgabentyp *initial state* enthält zwei Items, in denen die Probanden bis zu 8 Informationen (variiert szenarioabhängig) richtig zuordnen müssen. Eine Aufgabe wird nur dann als gelöst (1 Punkt) bewertet, wenn alle Informationen richtig positioniert sind. Der Gesamtscore für den Aufgabentyp *initial state* berechnet sich ebenfalls aus dem arithmetischen Mittel der beiden Items und kann die Werte 0; 0.5 und 1 annehmen. Die Bewertung der gesamten Phase der Wissensabfrage erfolgt anhand des Mittelwertes aus dem Gesamtscore für den Aufgabentyp *state identification* und dem Gesamtscore für den Aufgabentyp *initial state*.

In der sich anschließenden Steuerungsphase müssen die Probanden ein vorgegebenes Ziel in dem Szenario mit möglichst wenigen Bearbeitungsschritten erreichen. Die Steuerungsaufgaben werden zum einen danach ausgewertet, ob das vorgegebene Ziel erreicht wurde und zum anderen danach, wie viele Bearbeitungsschritte dafür benötigt wurden (0 Punkte = nicht gelöst; 1 Punkt = gelöst, aber mehr Schritte als nötig; 2 Punkte = gelöst). Für den Gesamtscore der Steuerungsphase wird der Punktwert an der maximalen Punktzahl relativiert, sodass sich Werte von 0; 0.5 und 1 ergeben.

¹ Item 1 und Item 2 werden zu einem Subscore zusammengefasst und Item 3 und Item 4.

9.2 Testdesign

Die oben beschriebenen Testinstrumente und Skalen sollen in der folgenden Studie in einer bestimmten für alle Probanden identischen Abfolge eingesetzt werden. Aufgrund der genau einzuhaltenden Testzeit bei der Durchführung des CFT 20-R ist dieser von allen Teilnehmern eines Durchführungstermins zeitgleich zu bearbeiten. Optimal für die Durchführung des CFT 20-R ist demnach die Positionierung an den Anfang der verwendeten Testbatterie direkt nach der Begrüßung der Probanden. Für die gesamte Testzeit des CFT 20-R sind 40 min vorgesehen. Im Anschluss füllen die Probanden die computerbasierten Fragebögen zu den soziodemografischen Angaben, der Need for Cognition-Skala, der Skala zur ICT-Nutzung sowie den Skalen zur Selbstwirksamkeitserwartung aus. Hiernach folgt der Einsatz des Problemlösetests im Umgang mit technischen Geräten. Vor der Explorationsphase eines technischen Systems wird das jeweilige gerätespezifische Vorwissen abgefragt, auf die Explorationsphase folgt dann die Steuerungsphase. Diese Abfolge (gerätespezifisches Vorwissen - Exploration - Steuerung) ist für alle präsentierten 15 technischen Alltagsgeräte gleich. Aufgrund der, aus den vorher durchgeführten Testerprobungen gesammelten Erfahrungen werden für die Dauer des technischen Problemlösetests etwa 90 min eingeplant. Direkt im Anschluss an den Problemlösetest bearbeiten die Probanden den allgemeinen Problemlösetest, der eine eigene Instruktion mit einem Beispiel sowie sechs folgende Problemszenarien enthält. Die Bearbeitungsdauer eines Items wird von den Autoren mit 5 min angegeben, so dass sich eine ungefähre Testzeit von 30 min ergibt. Abbildung 9.2 fasst den Ablauf der in dieser Studie eingesetzten Messinstrumente sowie die für sie benötigten Testzeiten zusammen.

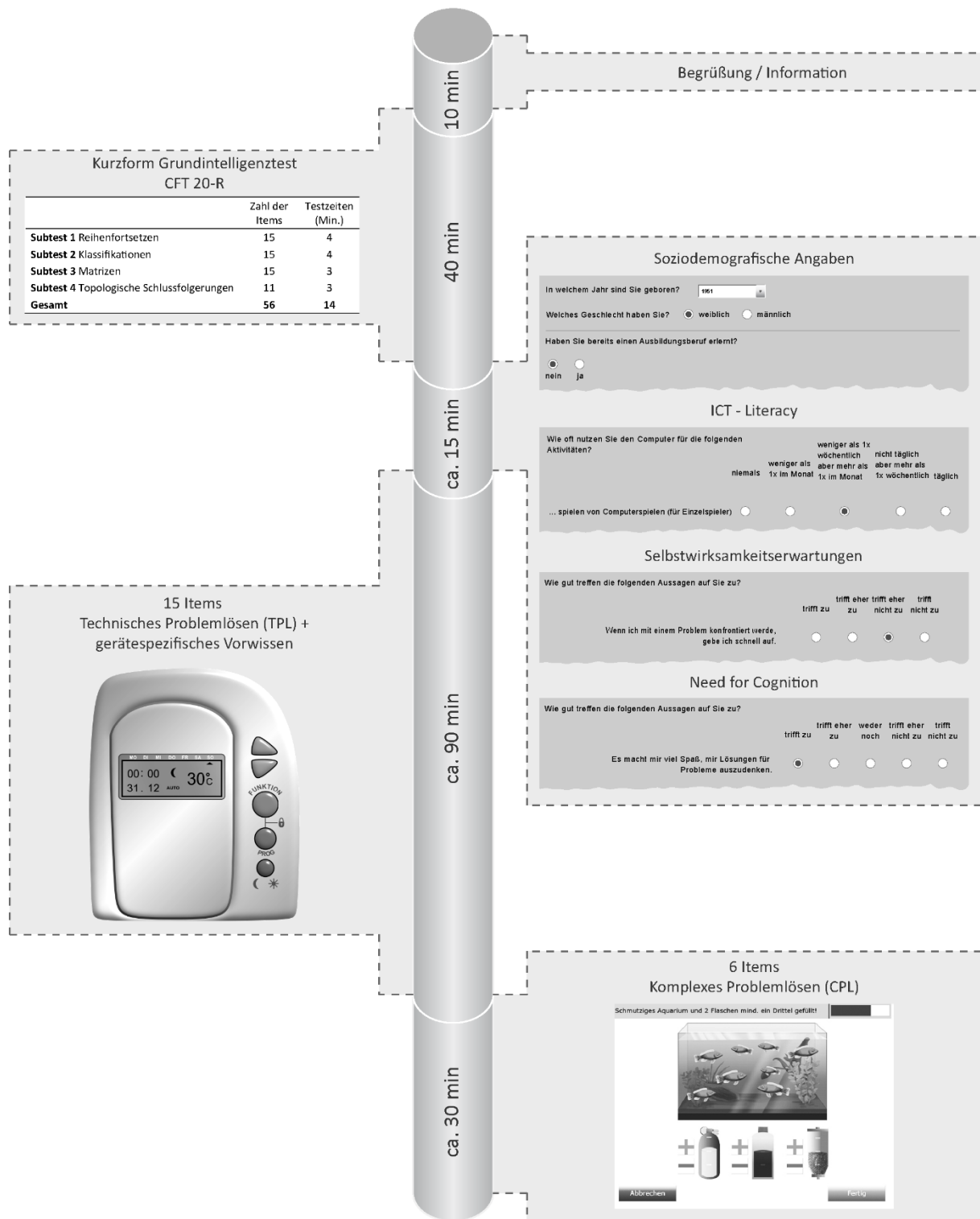


Abbildung 9.2: Abfolge und Testzeiten der eingesetzten Testinstrumente und Skalen

9.3 Stichprobe

Die Gültigkeit des in Abschnitt 6.6 formulierten Kausalmodells sowie die Gültigkeit der daraus abgeleiteten Hypothesen (Abschnitt 7) wird zunächst nur für die Population technisch interessierter Personen postuliert. Eine solche Eingrenzung muss vorgenommen werden, da der Einfluss, den das Interesse an Technik oder technischen Artefakten auf den problemlösenden Umgang mit technischen Alltagsgeräten ausübt, in dieser Studie nicht

untersucht wird. Die Stichprobe, an der die Gültigkeit der getroffenen Aussagen geprüft wird, stammt aus dieser Population.

Die Auswahl der Stichprobe erfolgte nach einer nicht-probabilistischen Methode, das heißt die Auswahl wurde nicht mit einem statistischen Zufallsverfahren getroffen, sondern basierte auf der willkürlichen Auswahl von Fällen von Studierenden, die ein technisches Fach studieren. Bei dieser Gelegenheitsstichprobe handelt es sich um eine Selbstselektionsstichprobe, da sich die Personen auf einen Teilnahmeaufruf meldeten. Hierzu wurde ein Plakat erstellt, das zum einen in gedruckter Form ausgegangen und zum anderen in digitaler Form über E-Learning-Plattformen, Soziale Netzwerke und E-Mail-Newsletter an Studierende versendet wurde.

Der erforderliche Umfang der Stichprobe ist abhängig von der Größe des Effekts, der in einem statistischen Test noch nachweisbar sein soll, bzw. davon, wie präzise Parameterschätzungen ausfallen sollen (Döring & Bortz, 2016a, S. 294). Um beispielsweise eine mittlere korrelative Effektgröße von $r = .30$ für einen Zusammenhang zwischen zwei Variablen statistisch (mit $\alpha = .05$ und $1-\beta = .80$) absichern zu können, wird eine Stichprobengröße von $n = 64$ benötigt (Berechnung mit G*Power 3.1.9.2 (Faul, Erdfelder, Buchner et al., 2009)). Für die Prüfung des in Kapitel 6.6 aufgestellten Strukturmodells sollte das Verhältnis zwischen der Stichprobengröße und der Anzahl frei zu schätzender Parameter etwa 5:1 betragen (Bentler & Chou, 1987). Schon bei einem geringen Freiheitsgrad von beispielsweise $df = 50$ wird demnach eine Stichprobengröße von $n = 250$ benötigt.

Neben den Überlegungen zur statistischen Absicherung gehen in die Bestimmung des Stichprobenumfangs auch Überlegungen bezüglich der verfügbaren Ressourcen ein. Um Probanden für die Teilnahme an der Studie zu gewinnen, wurde eine Vergütung von 30 € pro Person vorgesehen. Weil die Gesamttestzeit mit drei Stunden sehr lang ist, ist die Testleistung stark von der Motivation abhängig. Es wurde daher versucht, eine High-Stakes-Testsituation zu schaffen, indem außerdem unter den erfolgreichsten 10 Problemlösern ein Apple® iPad® Air2 verlost wurde. Für die Durchführung der Studie standen insgesamt 5.000,00 € zur Verfügung, so dass sich daraus eine maximale Stichprobengröße von 150 Personen ergab ($[5.000,00 \text{ €} - 500,00 \text{ €}] / 30,00 \text{ €/Person} = 150 \text{ Personen}$).

Für die Teilnahme an der Studie konnten insgesamt 147 Studierende eines technischen Faches gewonnen werden, von denen etwa ein Drittel weiblich ist (64,63 % männlich; 35,37 % weiblich). Das Durchschnittsalter der Probanden betrug zum Zeitpunkt der Datenerhebung $M = 26.59$ Jahre ($SD = 5.89$ Jahre). Das Fachsemester, in dem sich die Studierenden zum Erhebungszeitpunkt befanden, streute zwischen dem ersten und dem zwölften Semester ($M = 6.84$; $SD = 5.89$). Tabelle 9.6 fasst die Daten der Stichprobe zusammen.

Um den Probanden eine anonyme Studienteilnahme zusichern zu können, wurden keine Daten erhoben, aus denen Rückschlüsse auf die teilnehmende Person gezogen werden können. Damit die Ergebnisse aus dem papierbasierten CFT 20-R sowie die Ergebnisse aus den computerbasierten Skalen und Testinstrumenten fallbasiert zusammengefügt werden konnten, wird eine eindeutige Kennung benötigt. Die Probanden erhielten deshalb vor Durchführungsbeginn einen zufallsgenerierten, sechsstelligen Code aus lateinischen Kleinbuchstaben und Ziffern. Dieser Code musste von den Probanden zum einen auf dem Antwortbogen des CFT 20-R und zum anderen als Benutzername zur Anmeldung für die

computerbasierte Testbatterie eingetragen werden. Genutzt wurde der Code außerdem, um das Ergebnis der Verlosung des Apple® iPad® Air2 unter den zehn besten Problemlösern bekanntzugeben.

Tabelle 9.6: Stichprobendaten

	<i>n</i>	<i>M (Min;Max)</i>	<i>SD</i>
Alter	147	26.59 (17;58)	5.89
Fachsemester ^a	141	6.84 (1;12)	3.37
Studiengang			
Ingenieurwesen	33		
Elektrotechnik und Informationstechnik	2		
Informatik / Wirtschaftsinformatik	14		
Lehramt Technik	92		
Umwelttechnik	4		
Sonstige	2		
Geschlecht			
männlich	95		
weiblich	52		

Anmerkungen: *n* = Anzahl der Probanden; *M* = Mittelwert; *SD* = Standardabweichung; *Min* = Minimum; *Max* = Maximum

^aausgeschlossen *n* = 6, deren Angabe 12+ lautet

9.4 Zeitlicher Ablauf

Zu der Studienteilnahme wurde ab Februar 2015 aufgerufen. Interessierte Personen meldeten sich dazu per E-Mail mit Angabe ihres Studienganges und dem gewünschten Zeitraum für die Teilnahme (in der Vorlesungszeit im März oder in der vorlesungsfreien Zeit ab April bis Mai) an. Für die Durchführung der Studie standen insgesamt 30 Laptops zur Verfügung, so dass es nicht möglich war, die Daten von 150 Personen zeitgleich zu erheben. Die Datenerhebung war außerdem von der individuellen Terminplanung der interessierten Personen abhängig, so dass die Datenerhebung zu verschiedenen Zeitpunkten stattfinden musste. Hierzu erhielten die für die Studie in Frage kommenden Personen in einer nachfolgenden Antwort-E-Mail einen Link zu einer Onlineplattform, auf der sie einen der dort vorgeschlagenen Termine auswählen oder einen weiteren Termin vorschlagen konnten. Die aus dem beschriebenen Verfahren resultierenden Erhebungstermine in dem Erhebungszeitraum von März bis Mai 2015 mit der Anzahl der teilnehmenden Personen sind Abbildung 9.3 zu entnehmen.

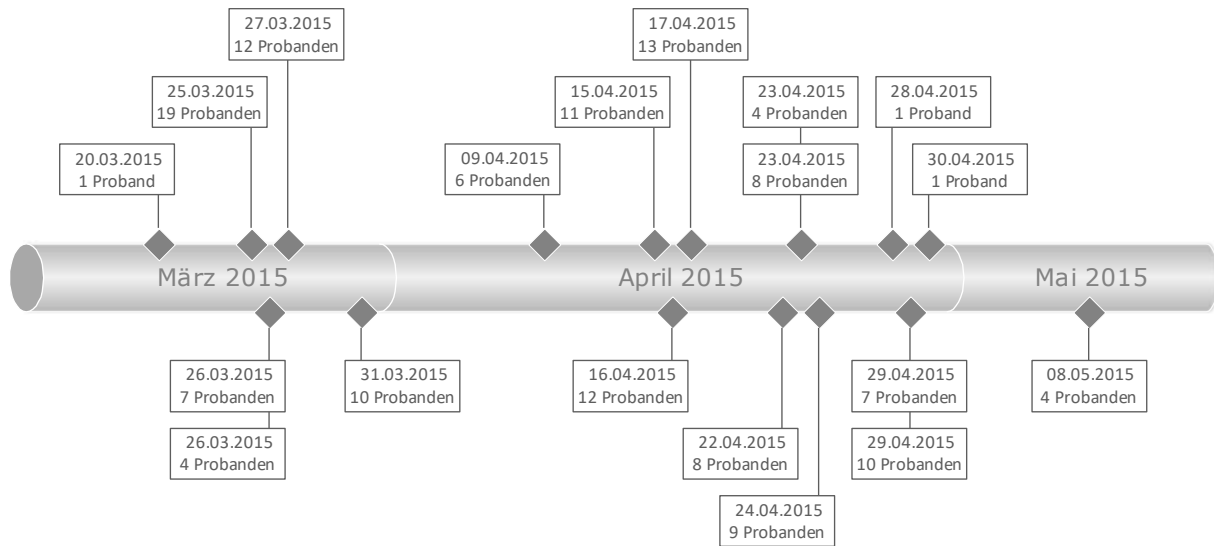


Abbildung 9.3: Erhebungszeitraum

10 Validitätsbetrachtungen

In der Literatur, die sich der Validität empirischer Untersuchungen widmet, wird diese als die Fähigkeit einer Untersuchung bezeichnet, zu einer richtigen Entscheidung über die zu untersuchende Hypothese zu kommen (vgl. Westermann, 2000). So sollen falsche Hypothesen als falsch und richtige Hypothesen als richtig erkannt werden. Es wird hier in der Regel zwischen interner und externer Validität unterschieden (Westermann, 2000; Bortz & Döring, 2006). Von interner Validität sprechen Bortz & Döring (2006), wenn die Ergebnisse einer Untersuchung kausal eindeutig interpretierbar sind; es also keine anderen Erklärungen für das Zustandekommen der Ergebnisse gibt. Die externe Validität betrifft hingegen die Generalisierbarkeit der Untersuchungsergebnisse auf weitere Personengruppen oder andere Situationen als die untersuchten (vgl. ebd., S. 502).

Zu den Einflussgrößen, die die externe Validität gefährden können, zählen die Autoren unter anderem eine mangelnde *instrumentelle Validität*, unter der sie die Güte des Testinstrumentes, das zu erfassen, was es erfassen soll, verstehen (vgl. ebd., S. 504). Die Validität eines Testinstrumentes gilt als das wichtigste Testgütekriterium (Pospeschill, 2010; Bortz & Döring, 2006; Schmidt-Atzert & Amelang, 2012). Die beiden anderen Hauptgütekriterien Objektivität und Reliabilität sind zwar notwendig, damit ein Test überhaupt valide sein kann, jedoch handelt es sich bei ihnen nicht um hinreichende Bedingungen, das heißt ein Test, der zwar über eine hohe Objektivität und Reliabilität verfügt, muss nicht zwingend valide sein (vgl. Pospeschill, 2010).

Die Frage nach der Validität muss vor allem gestellt werden, wenn, wie im vorliegenden Fall, ein latentes Konstrukt erfasst werden soll, dessen Indikatoren nicht definitorisch sind, sondern lediglich auf das Konstrukt schließen lassen (vgl. Pospeschill, 2013, S. 50ff.). Die im Rahmen der Testanwendung entstehenden Messwerte müssen dazu gemäß dem operational definierten Merkmal interpretiert werden. Dabei muss das Ziel der Ergebnisinterpretation nicht immer das gleiche sein, sondern kann sich „auf

- das Bewerten des Ergebnisses,
- das Verallgemeinern des Ergebnisses,
- das Extrapolieren des Ergebnisses auf andere Bereiche,
- das (kausale) Erklären eines Testergebnisses und
- das Fällen von weiterführenden Entscheidungen als Konsequenz aus dem Testergebnis“

(Hartig, Frey & Jude, 2012, S. 144 nach Kane 2001) beziehen. Messick (1995, S. 741) empfiehlt daher von Validität nicht als Eigenschaft eines Tests, sondern als Eigenschaft der interpretierten Testwerte zu sprechen. Auch Häcker, Leutner & Amelang (1998) sprechen von Validierung als einen Prozess, in dem beurteilt wird, „[...] wie angemessen, wie bedeutsam und wie nützlich die spezifischen Schlußfolgerungen sind, die aus Testwerten gezogen werden können“ (Häcker, Leutner & Amelang, 1998, S. 10). Die Begriffe Schlussfolgerungen und Interpretationen haben in diesem Zusammenhang eine ähnliche Bedeutung. Je nachdem, wie ein Testwert interpretiert wird, können daraus entsprechende Schlussfolgerungen abgeleitet werden. Die Konsequenz, die sich aus der Definition von Validität als

Testwertinterpretation ergibt, ist dass es nicht nur eine Validität gibt, sondern so viele wie mögliche Testwertinterpretationen bzw. Schlussfolgerungen.

Um einen Testwert sinnvoll interpretieren zu können, bedarf es eines Vergleichsmaßstabs. Der Testwert für sich allein lässt noch keine Schlussfolgerungen auf die Ausprägung des zu erfassenden Merkmals zu (Goldhammer & Hartig, 2012, S. 175). Wenn eine Person zum Beispiel in einem Ängstlichkeitstest den Wert 20 erreicht hat, lassen sich aus diesem Wert keine Aussagen darüber treffen, wie ängstlich diese Person ist. Erst ein Vergleich mit beispielsweise den Testwerten anderer Personen oder dem Mittelwert einer Bezugsgruppe gibt Auskunft über die zu untersuchende Merkmalsausprägung.

In der Kompetenzforschung wird die Interpretation von Ergebnissen, die aus der Anwendung eines Kompetenztestes entstehen, danach unterschieden, womit sich die Ergebnisse vergleichen lassen. Wird der Testwert einer Person mit denen aus einer genormten Stichprobe verglichen, spricht man von normorientierter Testwertinterpretation. Hierfür wird zu dem individuellen Testwert ein Normwert bestimmt, mit dem eine Person hinsichtlich ihrer Merkmalausprägung innerhalb einer Bezugsgruppe eingeordnet werden kann (Goldhammer & Hartig, 2012, S. 175f.). Für die Autoren Helmke & Hosenfeld (2004) reichen normorientierte Aussagen jedoch nicht aus, die Leistungen von Schülerinnen und Schülern im Rahmen von Vergleichsarbeiten (beispielsweise für die Verbesserung von Unterrichtsqualität) zu bewerten. Sie plädieren im Hinblick auf bildungspolitische Forderungen zur Erreichung von Bildungsstandards zum einen für die schnelle Entwicklung von Kompetenzmodellen und zum anderen, falls diese noch nicht existieren, für vorläufige Aussagen zu Kompetenzstufen (ebd., S. 61). Hierfür müssen die Testwerte aber kriteriumsorientiert interpretiert werden, das bedeutet, sie werden mit einem vorher festgelegten und zum Beispiel durch Bildungsstandards definierten Kriterium verglichen (Jung, 2010, S. 196). Bei der kriteriumsorientierten Testwertinterpretation interessiert, ob eine Person ein Kriterium erfüllt (ein Bildungsziel erreicht) hat und nicht wie viele Personen es ebenfalls erreicht haben, die Interpretation ist also unabhängig von der Bezugsgruppe (Jung, 2010, S. 196; Goldhammer & Hartig, 2012, S. 182). Für die Validierung als einen Prozess, in dem versucht wird zu belegen, dass die (norm- oder kriteriumsorientierten) Schlussfolgerungen, die man aus gewonnenen Testergebnissen zieht, richtig sind, gibt es verschiedene Zugänge (Häcker, Leutner & Amelang, 1998, S. 10; Pospeschill, 2010, S. 25ff.), die sich üblicherweise anhand drei Validitätsaspekte unterscheiden lassen (z. B. Fisseni, 2004, S. 63; Lienert & Raatz, 1998, S. 10f.; Bortz & Döring, 2006, S. 200f.): Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität.

Die *Inhaltsvalidität* entscheidet, ob die in einem Test verwendeten Items „repräsentativ sind für alle theoretisch denkbaren Items, die zur Erfassung des Konstrukts geeignet sind“ (Schmiemann & Lücken, 2014, S. 109). Wenn dem so ist, kann von dem Verhalten bei der Bearbeitung der vorliegenden Testaufgaben auf das Verhalten bei der Bearbeitung einer größeren Aufgabenmenge geschlossen werden (Hartig, Frey & Jude, 2012, S. 149). Belege für eine hinreichende Inhaltsvalidität setzen ein in der Konzeptualisierung ausreichend präzise definiertes Konstrukt voraus (Weiber & Mühlhaus, 2014, S. 157; Fisseni, 2004, S. 63). Die Beurteilung der Passung zwischen dem zu messenden Konstrukt und den konstruierten Testaufgaben ist Aufgabe der Inhaltsvalidierung und kann zum einen durch eine streng

theoriegeleitete Itementwicklung sichergestellt oder zum anderen von Experten beurteilt werden (Wilhelm & Kunina, 2009, S. 319).

Mit *Kriteriumsvalidität* wird der Beleg für die Richtigkeit der Schlussfolgerungen bezeichnet, die aus dem Verhalten der Testperson in dem Test Aussagen über ein Verhalten außerhalb der Testsituation tätigen (Moosbrugger & Kelava, 2012a, S. 18; Jenßen, Dunekacke & Blömeke, 2015, S. 16). Kriteriumsvalidität liegt dann vor, wenn das Testergebnis des zu erfassenden Konstruktes mit einem oder mehreren korrespondierenden Kriterien zusammenhängt (Bühner, 2011, S. 63; Bortz & Döring, 2006, S. 200). Soll ein Test beispielsweise Aussagen über die fachliche Eignung einer Person für ein Technikstudium zulassen, so sollten die Testwerte mit dem Kriterium *Schulnote im Fach Technik* korrelieren. Weiber & Mühlhaus (2014, S. 157) sehen ein Problem bei der Kriteriumsvalidierung darin, ein geeignetes Kriterium zu finden. Vorausgesetzt, es kann ein Außenkriterium benannt werden, so muss es, wenn es sich um ein hypothetisches Konstrukt handelt, selbst valide und reliabel erfasst sein, damit eine Validierung mit diesem Kriterium überhaupt sinnvoll ist (Bortz & Döring, 2006, S. 201; Lienert & Raatz, 1998, S. 221). Bortz & Döring (2006, S. 201) empfehlen aus diesem Grund, wenn es möglich ist, einen Test an mehreren Außenkriterien zu validieren.

Je nachdem wann das Außenkriterium zeitlich verfügbar ist, wird die Kriteriumsvalidität in prognostische Validität und Übereinstimmungsvalidität unterschieden (z. B. bei Cronbach & Meehl, 1955, S. 282; Moosbrugger & Kelava, 2012a, S. 18; Bortz & Döring, 2006, S. 200f.; Fisseni, 2004, S. 66). Kann die Messung des Kriteriums erst zu einem späteren Zeitpunkt erfolgen, so können die zu validierenden Testwerte dazu dienen, das Außenkriterium vorauszusagen. Lässt sich ein Zusammenhang zwischen dem Verhalten in dem Test und dem späteren Verhalten außerhalb der Testsituation durch ein Außenkriterium feststellen, ist Prognosevalidität gegeben, das heißt aus den gewonnenen Testwerten lässt sich auf ein späteres Verhalten schließen (Weiber & Mühlhaus, 2014, S. 158). Liegen zwischen dem Konstrukt und dem zum annähernd gleichen Zeitpunkt erfassten Außenkriterium bedeutsame Zusammenhänge vor, so spricht man von Übereinstimmungsvalidität (Amelang & Schmidt-Atzert, 2006, S. 154).

Lassen sich keine Kriterien finden, die selbst als ausreichend valide angesehen werden können, ist der einzige Weg einen infiniten Regress zu vermeiden, derjenige, die *Konstruktvalidität* zu untersuchen (Cronbach & Meehl, 1955, S. 282). Dieser Validitätsaspekt zielt auf die theoretische Legitimation des von einem Test erfassten Konstruktes (Moosbrugger & Kelava, 2012a, S. 16). Interpretationen von Testergebnissen gelten dann als konstruktvalide, wenn sich aus dem supponierten Konstrukt Hypothesen ableiten lassen, die durch empirische Befunde gestützt werden (Bortz & Döring, 2006, S. 201; Hartig, Frey & Jude, 2012, S. 153). Nur wenn die abgeleiteten Hypothesen nicht widerlegt werden können, darf zunächst davon ausgegangen werden, dass der Test das angenommene Konstrukt erfasst und nicht ein anderes eventuell ähnliches Konstrukt. Für die Bestimmung der Konstruktvalidität, unter der viele Autoren auch alle anderen Validitätsaspekte fassen (z. B. Messick, 1995, S. 742; Fisseni, 2004, S. 68), muss das hypothetische Konstrukt bereits ausformuliert und in ein nomologisches Netzwerk eingebettet sein, das die Beziehungen des Konstruktes mit anderen Konstrukten darstellt. Aufgabe der Konstruktvalidität ist es, das postulierte nomologische Netzwerk schrittweise zu überprüfen und Fehler darin aufzudecken (Cronbach & Meehl, 1955). Hierzu

werden aus dem formulierten theoretischen Konzept Hypothesen abgeleitet, die dann anhand der beobachteten Testwerte überprüft werden. Hartig, Frey & Jude (2012, S. 154) weisen darauf hin, dass die Aussagen zur Konstruktvalidität aufgrund dieses hypothetisch-deduktiven Vorgehens immer nur vorläufig sein können und dass die Annahme, dass ein bestimmtes Konstrukt für die Entstehung eines Testergebnisses verantwortlich ist, nur solange gerechtfertigt ist, bis sie falsifiziert wird.

Für die strenge Prüfung des nomologischen Netzwerkes, sowie die aus dem Konstrukt abgeleiteten Annahmen im Rahmen der Konstruktvalidierung eignen sich alle empirische Methoden, die hypothesenprüfend vorgehen (ebd., S. 156). Die Wahl einer geeigneten Validierungsstrategie richtet sich nach dem zu analysierenden Aspekt der Übereinstimmung zwischen Theorie und Empirie. Dabei lassen sich prinzipiell die folgenden drei Aspekte unterscheiden (vgl. Jenßen, Dunekacke & Blömeke, 2015; Bortz & Döring, 2006; Fisseni, 2004; Lienert & Raatz, 1998), wobei sich diese nicht gegenseitig ausschließen, sondern ergänzen sollten (Jenßen, Dunekacke & Blömeke, 2015):

- Prüfung der Verortung des Testkonstruktes im nomologischen Netzwerk (Zusammenhang zu anderen Konstrukten),
- Logisch-inhaltliche Analyse des Konstruktes (Konstruktstruktur),
- Analyse der Konstruktitems (Konstruktrepräsentation).

Für die Überprüfung der Zusammenhänge des Testkonstruktes zu anderen Konstrukten werden korrelative Ansätze genutzt, wobei ein Konstrukt mit einem ähnlichen Konstrukt höher korrelieren sollte (konvergente Validität) als mit einem eher fremden Konstrukt (diskriminante Validität) (Moosbrugger & Kelava, 2012a, S. 17; Lienert & Raatz, 1998, S. 227). Wird beispielsweise ein neuer computerbasierter Intelligenztest entwickelt, der in der Durchführung und Auswertung ökonomischer zu handhaben ist als konventionelle Intelligenztests, sollten die aus seinem Einsatz gewonnenen Testdaten mit denen aus einem papierbasierten Intelligenztest sehr hoch korrelieren, um von den Testergebnissen auf das dahinterliegende Konstrukt Intelligenz schließen zu dürfen. Zum Nachweis diskriminanter Validität geht es nicht darum einen fehlenden bzw. geringen Zusammenhang zu irgendwelchen offensichtlich fremden Konstrukten zu zeigen (Moosbrugger & Kelava, 2012a, S. 17), sondern solche Konstrukte auszuwählen, die möglicherweise ähnlich sind, aber mit dem entwickelten Testinstrument nicht erfasst werden sollen (Schmiemann & Lücken, 2014, S. 116). Für das Beispiel des entwickelten computerbasierten Intelligenztests ließe sich wohl kaum erwarten, dass die Testergebnisse mit den Leistungen im Hochsprung korrelieren. Denkbar, aber nicht wünschenswert wäre eine Korrelation mit dem Konstrukt ICT-Literacy, also der Fähigkeit im Umgang mit Informations- und Kommunikationstechnologien, weil die Durchführung des entwickelten Intelligenztests computerbasiert erfolgt.

Mit der logisch-inhaltlichen Analyse kann geprüft werden, ob die theoretisch angenommene Struktur des Konstruktes, das durch die Items operationalisiert wurde, auch in dem Testverhalten zum Ausdruck kommt. Mit der inneren Struktur des Konstruktes eng verbunden sind Annahmen über dessen Dimensionalität, die sich durch die Untersuchung der Zusammenhänge zwischen den einzelnen Items zeigen lässt (vgl. Wilhelm & Kunina, 2009, S. 319; Borg & Staufenbiel, 2007, S. 336). Wenn die Items lediglich ein homogenes Merkmal

erfassen, kann mithilfe von Faktorenanalysen geprüft werden, ob diese auch tatsächlich nur auf einem Faktor laden (ebd.).

Auf Ebene der Testitems lässt sich auch mit Hilfe experimenteller Ansätze Konstruktvalidität untersuchen (Bortz & Döring, 2006, S. 202; Hartig, Frey & Jude, 2012, S. 156), indem theoriegeleitet Annahmen über individuelle Unterschiede im Testverhalten getroffen und anhand der gewonnenen Daten überprüft werden. Individuelle Unterschiede im Testverhalten werden dabei vor allem durch die Unterschiede in der Schwierigkeit der Testitems verursacht (Embretson, 1983, S. 180). Können aus dem präzisierten Konstrukt Hypothesen über diese Schwierigkeitsunterschiede zwischen den einzelnen Items vorab formuliert werden, dient ihre Überprüfung der Konstruktvalidierung (Hartig & Frey, 2012, S. 44). Im Hinblick auf Kompetenzkonstrukte müssen die spezifischen Eigenschaften und Merkmale der Situation, in der sich die entsprechende Kompetenz zeigen soll, und die sich daraus ergebenden Anforderungen an die Testperson bekannt sein und systematisch beschrieben werden (Jenßen, Dunekacke & Blömeke, 2015, S. 17). Der Vorteil der Analyse von schwierigkeitsbestimmenden Itemmerkmalen für die Validität von Testwerten wenig erforschter Konstrukte, ist deren unproblematische Zugänglichkeit, im Gegensatz zu den korrelativen Ansätzen, bei denen andere Testskalen vorhanden sein müssen, um einen oder keinen Zusammenhang zeigen zu können. Die verschiedenen Strategien zur Konstruktvalidierung stehen aber nicht in Konkurrenz zueinander wie die Autoren Jenßen, Dunekacke & Blömeke (2015) zu Recht bemerken:

„Verschiedene Strategien der Konstruktvalidierung schließen sich nicht gegenseitig aus, sondern sollten sich ergänzen. Welche Strategien sich zur Unterstützung der theoriebasierten Interpretation von spezifischen Testwerten empfehlen, hängt davon ab, wozu präzise theoretische Annahmen existieren, ob also fundierte Hypothesen über eine dimensionale Struktur formuliert werden können oder über Anforderungen, mit denen die Schwierigkeit von Aufgaben erklärt werden können, bzw. über Zusammenhänge mit anderen Variablen“ (Jenßen, Dunekacke & Blömeke, 2015, S. 17).

10.1 Schwierigkeitsbestimmende Merkmale

Die vorab formulierten (theoriegeleiteten) Annahmen über das unterschiedliche Verhalten von Personen infolge der Schwierigkeitsunterschiede zwischen den Items, sollen anhand des tatsächlich beobachteten Verhaltens überprüft werden. Der Abschnitt 6.4 beschreibt hierzu Merkmale auf System- und Situationsebene, von denen ein Einfluss auf die Schwierigkeit beim explorierenden und steuernden Umgang mit technischen Geräten angenommen wird. Um den theoretisch begründeten Zusammenhang zwischen den Problemmerkmalen und der Problemschwierigkeit untersuchen zu können, muss das Theoriemodell in ein geeignetes statistisches Modell überführt werden. Denn im Gegensatz zu Theoriemodellen sind statistische Modelle in der Lage die Modellgrößen mathematisch berechnen oder schätzen zu können (vgl. Urban & Mayerl, 2011, S. 17). Die Anwendung mathematischer Rechenoperationen¹ erfordert eine weitere Komplexitätsreduzierung der bereits realitätsvereinfachenden Theoriemodelle. Im einfachsten Fall wird der Zusammenhang

¹ Auch in einem Rechenmodell, das viele Parameter berücksichtigt und daher sehr komplex ist, verbleiben noch Differenzen zwischen dem Theoriemodell und dem statistischen Modell.

zwischen Ursache und Wirkung als eine lineare Beziehung angenommen (vgl. Opp, 2010), so dass sich statistische Modelle durch lineare Gleichungssysteme darstellen und sich die Modellparameter mithilfe dieser Gleichungssysteme berechnen lassen.

10.1.1 Regressionsanalyse

Als Statistikmodell für die Prüfung des Wirkzusammenhangs der angenommenen schwierigkeitsbestimmenden Merkmale auf die empirisch zu beobachtende Schwierigkeit eignet sich die Regressionsanalyse (vgl. Hartig, 2007, S. 90; Gut-Glanzmann, 2012, S. 146). Basierend auf dem allgemeinen linearen Modell lassen sich mit regressionsanalytischen Verfahren Beziehungen zwischen Variablen modellieren. Bei der Modellformulierung wird diejenige Variable, auf die ein Einfluss ausgeübt wird, als *abhängige Variable* und diejenigen Variablen, die einen Einfluss ausüben als *unabhängige Variablen* bezeichnet (Urban & Mayerl, 2011, S. 26). Ob eine Variable abhängig oder unabhängig ist, ergibt sich allein aus dem jeweiligen Theoriemodell; die Festlegung muss bereits vor der Regressionsanalyse erfolgen (vgl. Schendera, 2014, S. 37; Opp, 2010, S. 25). Die Entwicklung der Probleme, die in dem Test zur Erfassung des *problemlösenden Umgangs mit technischen Systemen* zusammengefasst sind, erfolgte anhand der Variation der System- und Situationsmerkmale. Die sich nach der Erstellung des Testes nicht mehr verändernden Merkmale stellen somit die unabhängigen Variablen in dem statistischen Modell dar. Es wird angenommen, dass eine höhere Ausprägung eines der Merkmale (z. B. eine größere Anzahl an Bedienelementen) dazu führt, dass die Schwierigkeit in der Bedienung eines technischen Systems steigt. Die zu beobachtende Schwierigkeit stellt damit die abhängige Variable in dem statistischen Modell dar.

Die abhängige Variable *Itemschwierigkeit* lässt sich damit in eine Summe (Linearkombination) von gewichteten Ausprägungen in den unabhängigen Variablen zerlegen, bzw. durch sie hervorsagen (vgl. Moosbrugger, 2011, S. 4). Weil sich die abhängige Variable durch die unabhängigen Variablen vorhersagen lassen, werden letztere in Statistikmodell auch als *Prädiktorvariablen* bezeichnet, während für die abhängige Variable der Begriff *Kriteriumsvariable* verwendet wird (vgl. ebd., S. 5; Schendera, 2014, S. 36).

Formel (10.1) beschreibt das lineare Modell in allgemeiner Form:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon \quad (10.1)$$

y_i = empirische Schwierigkeit von Item i

β_0 = Regressionskonstante (Achsenabschnitt oder englisch: intercept)

β_j = Regressionsgewicht für Merkmal j

k = Anzahl der Itemmerkmale

x_{ij} = Ausprägung des Merkmals j in Item i

ε = Verbleibende Abweichung zwischen erwarteter und empirischer Itemschwierigkeit

Die Regressionsgewichte¹ β_j bringen den Einfluss eines Merkmals auf die Itemschwierigkeit zum Ausdruck und werden in der Regressionsanalyse so geschätzt, dass die empirischen

¹ Synonym werden die Begriffe Regressionskoeffizienten und Regressionsparameter verwendet.

Itemschwierigkeiten bestmöglich durch die Regressionsgewichte und den zugehörigen Merkmalsausprägungen vorhergesagt werden. Eine bestmögliche Vorhersage wird erreicht, wenn die verbleibende Abweichung ε zwischen der im Modell erwarteten und der empirischen Schwierigkeit minimal wird (Hartig, 2007, S. 91). Der Regressionskoeffizient β_0 beschreibt einen von den Prädiktorvariablen unabhängigen Betrag in der Ausprägung der Kriteriumsvariablen (Moosbrugger, 2011, S. 5) und wird ebenfalls durch die Regressionsanalyse geschätzt. Eine einfache Regression untersucht lediglich den Einfluss einer unabhängigen auf eine abhängige Variable, während die multiple Regression die Einflüsse mehrerer unabhängigen Variablen gleichzeitig schätzt, indem sie die jeweils anderen Einflussgrößen konstant hält (vgl. Wolf & Best, 2010, S. 607).

Schätzverfahren

Die Herausforderung jeder Regressionsanalyse besteht darin, die Regressionskoeffizienten so zu schätzen, dass die durch sie geschätzten Werte für die abhängige Variable \hat{y} möglichst gut den beobachteten Werten für die abhängige Variable y entsprechen (Wolf & Best, 2010, S. 613). Die Aufgabe der Regressionsanalyse ist eine lineare Funktion zu finden, bei der diese Abweichungen (Residuen) möglichst klein sind. Im Fall einer einfachen Regression (eine unabhängige Variable) lässt sich die Funktion als eine Gerade, im Fall von zwei unabhängigen Variablen als eine Ebene darstellen. Multiple Regressionen mit mehr als zwei unabhängigen Variablen lassen sich grafisch nicht mehr veranschaulichen (m-dimensionaler Unterraum) (vgl. Moosbrugger, 2011, S. 25). Das Standardverfahren für das Finden einer solchen linearen Funktion für (lineare) Regressionen ist die *Methode der kleinsten Quadrate* (vgl. ebd.; Urban & Mayerl, 2011, S. 14; Moosbrugger, 2011, S. 29). Diese Methode bestimmt die Regressionsparameter so, dass die Summe der quadrierten Abweichungen zwischen y und \hat{y} minimal wird (Wentura & Pospeschill, 2015, S. 23). Das Quadrieren der Residuen führt dabei zu einer stärkeren Gewichtung von größeren Abweichungen und zur Vermeidung der Kompensierung positiver und negativer Abweichungen (Backhaus, Erichson, Plinke et al., 2008, S. 63). Die Verwendung der Kleinst-Quadrate-Methode ist jedoch an Voraussetzungen geknüpft, die vor dem Einsatz des Verfahrens zu prüfen sind.

Voraussetzungen

Der Einsatz der linearen Regression auf Basis der Methode der kleinsten Quadrate setzt mindestens Intervallskalenniveau der abhängigen Variable voraus (vgl. Wolf & Best, 2010, S. 608). Weist die abhängige Variable nur Nominal- oder Ordinalskalenniveau auf, so wird stattdessen die logistische Regression verwendet. Die Itemschwierigkeit, die abhängige Variable, deren Beeinflussung durch die System- und Situationsmerkmale in der hier eingesetzten Regressionsanalyse untersucht werden soll, ist intervallskaliert. Da in dem hier verwendeten statistischen Modell der Zusammenhang zwischen den schwierigkeitsbestimmenden Merkmalen und der Itemschwierigkeit als linear angenommen wird, ist Linearität eine weitere Voraussetzung, die vorab zu prüfen ist (vgl. Schendera, 2014, S. 7). Die Art der Beziehung zwischen zwei intervallskalierten Variablen lässt sich gut graphisch mithilfe von Streudiagrammen erkennen (vgl. Lewis-Beck & Lewis-Beck, 2016, S. 8 f.), in denen die abhängige Variable auf der y-Achse und die unabhängige Variable auf der x-Achse aufgetragen wird. Die Daten werden als Punkte in das Diagramm eingetragen und sollten (wenn Linearität vorliegt) entlang einer gedachten Gerade angeordnet sein (ebd., S. 9).

Das Ausmaß der Streuung der Punkte um die gedachte Gerade zeigt die Stärke des Zusammenhangs an.

Dass die Methode der multiplen Regression die Einflüsse der unabhängigen Variablen auf die abhängige Variable schätzen kann, indem sie die jeweils anderen Prädiktoren konstant hält, setzt voraus, dass diese Prädiktorvariablen voneinander unabhängig sind (vgl. Schendera, 2014, S. 104). Zwischen den, in dem Modell der multiplen Regression definierten unabhängigen Variablen dürfen als eine weitere Bedingung also keine Multikollinearitäten vorliegen. Das bedeutet, dass eine Variable nicht eine Linearkombination aus anderen in dem Modell enthaltenen Variablen sein darf (vgl. Achen, 1982, S. 32). Zur Aufdeckung einer Multikollinearität können die Korrelationskoeffizienten zwischen den unabhängigen Variablen betrachtet werden (Hedderich & Sachs, 2016, S. 771; Backhaus, Erichson, Plinke et al., 2008, S. 89), denn eine Multikollinearität geht auch immer mit einer hohen Korrelation der Variablen einher (Schendera, 2014, S. 104). Eine hohe Korrelation ist jedoch nicht zwingend ein Indikator für Multikollinearität (Hedderich & Sachs, 2016, S. 771), so dass weitere Informationen berücksichtigt werden müssen. Eine Möglichkeit ist jede unabhängige Variable selbst als abhängige Variable in einer linearen multiplen Regression durch die anderen unabhängigen Variablen zu analysieren (Backhaus, Erichson, Plinke et al., 2008, S. 89). Für jeden Prädiktor x_j kann damit ein Bestimmtheitsmaß R_j^2 angegeben werden (S. 188). Aus diesem wird anhand der Unbestimmtheit $(1-R_j^2)$ der Varianz-Inflationsfaktor (VIF) berechnet:

$$VIF(x_j) = \frac{1}{1 - R_j^2} \quad (10.2)$$

Je größer der lineare Zusammenhang zwischen den unabhängigen Variablen, desto weiter nähert sich R_j^2 dem Wert 1 und umso größer wird der Wert für $VIF(x_j)$ (Schendera, 2014, S. 105). Für $VIF(x_j) > 10$ ist für die Autoren Hedderich & Sachs (2016, S. 771) eine Multikollinearität zu vermuten (siehe auch Lewis-Beck & Lewis-Beck, 2016, S. 79), während für $VIF(x_j) > 100$ sicher von ihr auszugehen ist (vgl. Rudolf & Müller, 2012, S. 76).

Für den Umgang mit Multikollinearität werden von den Autoren Chatterjee & Price (1995) verschiedene Ansätze vorgestellt. Korrelieren zwei Prädiktorvariablen nahezu perfekt miteinander, sollte eine der interkorrelierenden Variablen aus dem Regressionsmodell entfernt werden (Schendera, 2014, S. 129), da durch sie keine zusätzlichen Informationen gewonnen werden. Gibt es theoretische Gründe die miteinander korrelierenden Variablen in der Modellgleichung zu belassen, empfiehlt es sich die vorhandene Multikollinearität nicht zu ignorieren, sondern entsprechend zu berücksichtigen. Die Modellgleichung muss dazu sorgfältig interpretiert werden, damit die abhängige Variable nicht falsch vorhergesagt wird (vgl. Schendera, 2014, S. 131). Eine weitere Möglichkeit besteht in der Gewichtung der multikollinearen Prädiktorvariablen direkt in der Regressionsgleichung anhand des ermittelten Verhältnisses zwischen ihnen (ebd. S. 131).

Die Bestimmung von Konfidenzintervallen für die Regressionskoeffizienten setzt außerdem eine Normalverteilung der Residuen voraus (vgl. Bortz & Schuster, 2010, S. 198; Fromm, 2012, S. 88) und verbietet, dass diese in Abhängigkeit voneinander (Autokorrelation) und in Abhängigkeit zur beobachteten Variablen (Heteroskedastizität) systematisch variieren (vgl. Fromm, 2012, S. 87f.). Die Residualanalyse ist demnach eine weitere vorbereitende Maßnahme

für die eigentliche Regressionsanalyse. Moosbrugger schlägt zu diesem Zweck vor, eine graphische Begutachtung der Residuen mit Hilfe von Residualplots vorzunehmen, in denen die vorhergesagten Werte für die abhängige Variable \hat{y}_i auf der x-Achse und die Residuen auf der y-Achse abgetragen werden (Moosbrugger, 2011, S. 91).

Bei vorliegender Heteroskedastizität (Streuungsungleichheit) streuen die Residuen mit einem systematischen Muster um ihren Mittelwert, die Streuung ist also nicht überall gleich, sondern variiert in Abhängigkeit von den Werten von \hat{y}_i (vgl. Urban & Mayerl, 2011, S. 244). Heteroskedastizität ist für Urban & Mayerl ein Hinweis auf Mängel in der Formulierung der Modellgleichung. Ungleich streuende Residuen zeigen „[...]“, dass es noch einen bedeutenden Anteil von systematischer Variation in der abhängigen Y-Variablen gibt, der in der Regressionsschätzung nicht berücksichtigt werden konnte und deshalb zu systematisch variierenden Residuen führt“ (Urban & Mayerl, 2011, S. 249). In diesem Fall muss das Regressionsmodell neu spezifiziert werden (ebd.).

Autokorrelationen (also Korrelationen der Residuen untereinander) entstehen vor allem bei Längsschnittstudien, in denen Variablen über einen längeren Zeitraum hinweg beobachtet werden (Urban & Mayerl, 2011, S. 260) und in denen nicht berücksichtigte Einflussgrößen über die Zeit hinweg stabil sind (vgl. Ohr, 2010, S. 649). Da im Rahmen dieser Arbeit weder zeitliche noch kontextuelle Veränderungen modelliert werden, ist eine Autokorrelation der Residuen nahezu auszuschließen (vgl. dazu Backhaus, Erichson, Plinke et al., 2008, S. 98). Auf den Nachweis von Autokorrelationen der Residuen wird hier deshalb nicht weiter eingegangen. Für die Prüfung der Normalverteilung der Residuen empfiehlt sich die Ausgabe der Residuen-Histogramme (vgl. Riedwyl & Ambühl, 2000, S. 11) oder ein Normal-Probability-Plot (Ohr, 2010, S. 667). Ausreißer bei kleinen Stichproben üben einen besonderen Einfluss auf die Regressionsergebnisse aus und sollten ebenfalls vorab genauer betrachtet werden (Jann, 2009, S. 102). Ausreißer können nämlich eine Hebelwirkung auf die Regressionsfunktion ausüben.

Regressionstechnik

Die praktische Durchführung der eigentlichen Regressionsanalyse bezeichnet man auch als Regressionstechnik. Technisch gibt es verschiedene Methoden, unabhängige Variablen in die Regressionsgleichung einzubeziehen. Die Autoren Eid, Gollwitzer & Schmitt unterscheiden hierbei zwischen zwei generellen Strategien: die Auswahl von Variablen aufgrund theoretischer Überlegungen sowie die datengesteuerte Auswahl von Variablen (2013, S. 629). Die theoretische Auswahl ist einer datengesteuerten vorzuziehen, da nur eine inhaltlich-theoretische Modellbildung einer theoretisch-kritischen Prüfung standhalten kann (Schendera, 2014, S. 103). Es werden diejenigen unabhängigen Variablen in das Regressionsmodell aufgenommen, von denen man aus theoretischen Überlegungen einen Einfluss auf die abhängige Variable erwartet (Eid, Gollwitzer & Schmitt, 2013, S. 629). Gibt es mehrere Variablen, die eine inhaltliche Ähnlichkeit aufweisen (ohne dass zwischen ihnen eine Multikollinearität vorliegt), kann es sinnvoll sein, sie zu Blöcken zusammenzufassen und nacheinander in die Regression einzubeziehen (Riedwyl & Ambühl, 2000). Nachteil dieser a-priori-Auswahl von unabhängigen Variablen ist, dass eventuell (statistisch) nicht bedeutsame oder zu viele Prädiktoren in die Regressionsgleichung aufgenommen werden (Leonhart, 2010, S. 198). Bei der datengesteuerten Auswahl an Prädiktoren, werden schrittweise nur solche in die Regressionsgleichung genommen, die sich als relevant für die Aufklärung der Varianz in

der abhängigen Variable erweisen, wobei versucht wird, die Anzahl an unabhängigen Variablen zu minimieren. Die Relevanz ergibt sich dabei nicht anhand theoretischer Überlegungen, sondern ausschließlich daraus, ob sie einen signifikanten Beitrag zur Vorhersage der Kriteriumsvariablen leistet (Eid, Gollwitzer & Schmitt, 2013, S. 629) und wird aus diesem Grund auch kritisiert (vgl. Urban & Mayerl, 2011, S. 113).

Für Untersuchungen, in denen noch kein empirisch geprüftes Theoriemodell vorhanden ist, bietet sich eine Kombination beider Strategien an (vgl. Leonhart, 2010, S. 198). Mit den aus einem Theoriemodell abgeleiteten, potenziellen Prädiktorvariablen wird eine multiple Regressionsanalyse durchgeführt, in der iterativ Variablen entfernt werden, die keinen bedeutenden Beitrag zur Vorhersage der abhängigen Variable leisten. Dieses Verfahren wird auch für die Analyse der Schwierigkeiten im explorierenden und steuernden Umgang mit technischen Systemen gewählt. Zunächst werden die theoretisch erwarteten Einflüsse auf die Itemschwierigkeit anhand der Problemmerkmale, denen sie zugeordnet werden können, in Blöcke gefasst. Für die Explorationsphase ergeben sich die folgenden Blöcke:

Tabelle 10.1: Blockweise Aufnahme unabhängiger Variablen in die Regressionsgleichung zur Bestimmung der Itemschwierigkeit

Explorationsschwierigkeit	Steuerungsschwierigkeit
Komplexität	Explorationsvollständigkeit
Vernetztheit	Situation - Komplexität
Transparenz	Situation - Vernetztheit
Dynamik	

Interpretation

Ziel der Regressionsanalyse ist die Schätzung der Regressionsparameter, um sie anschließend inhaltlich im Sinne der Konstruktvalidierung interpretieren zu können. Die Regressionskoeffizienten geben an, um welchen Betrag sich der Wert der vorhergesagten abhängigen Variable \hat{y}_i (die Itemschwierigkeit) ändert, wenn sich die unabhängige Variable (das System- oder Situationsmerkmal) um eine Einheit ändert (Eid, Gollwitzer & Schmitt, 2013, S. 613). Um solche Vorhersagen treffen zu können, werden die Regressionsparameter jeweils auf Grundlage der Metrik des entsprechenden Merkmals und der abhängigen Variable y berechnet (Diaz-Bone, 2013, S. 197). Dadurch lassen sich die Parameter aber nicht miteinander vergleichen und anhand der Größe eines Regressionskoeffizienten darf nicht auf die Wichtigkeit der betreffenden Variable geschlossen werden (vgl. Backhaus, Erichson, Plinke et al., 2008, S. 65).

Werden die Regressionsgewichte anschließend standardisiert, lassen sich die Itemmerkmale anhand ihres Einflusses auf die Itemschwierigkeit miteinander vergleichen (vgl. Backhaus, Erichson, Plinke et al., 2008, S. 65). Je größer der Betrag eines standardisierten Regressionskoeffizienten, umso größer ist der Einfluss des Merkmals, auf den sich der Koeffizient bezieht. Standardisierte Regressionskoeffizienten > 0 zeigen bei zunehmender Merkmalsausprägung zunehmende Werte der Itemschwierigkeit¹ an, während

¹ Hohe Werte für die Itemschwierigkeit bedeuten, dass ein Item häufiger gelöst wird, also einfacher ist.

standardisierte Koeffizienten < 0 bei zunehmender Merkmalsausprägung abnehmende Werte der Itemschwierigkeit anzeigen. Regressionskoeffizienten, deren Werte nahe 0 liegen, zeigen, dass das betreffende Merkmal keinen Einfluss auf die Itemschwierigkeit hat (vgl. Schendera, 2014, S. 59).

Modellgüte

Um zu klären, wie gut das statistische Modell die beobachteten Daten überhaupt vorhersagen kann, ist seine Güte zu prüfen. Backhaus, Erichson, Plinke et al. (2008) gliedern die Überprüfung in zwei Bereiche. Im ersten Teil soll das statistische Modell der Regressionsfunktion als Ganzes geprüft werden, um zu klären wie gut das gesamte Modell die Kriteriumsvariable erklären kann. Im zweiten Teil ist die Frage zu beantworten, wie gut einzelne Prädiktorvariablen zur Vorhersage bzw. Erklärung der Kriteriumsvariablen beitragen.

Ein globales Gütemaß zur Prüfung der Regressionsfunktion als statistisches Modell stellt der Determinationskoeffizient R^2 dar und beschreibt, wie gut das Modell zu den beobachteten Daten *passt* (engl: *to fit*) und stellt damit ein Maß für den Modellfit dar. Er berechnet sich durch den Quotient aus der systematischen Varianz s_y^2 und der Gesamtvarianz s_y^2 :

$$R^2 = \frac{s_y^2}{s_y^2} \quad (10.3)$$

Die systematische Varianz s_y^2 ist die Varianz, die durch die Prädiktorvariablen erklärt werden kann. Sie ist in den seltensten Fällen identisch mit der Gesamtvarianz s_y^2 , da es immer auch eine durch die Residuen verursachte unsystematische Varianz s_E^2 gibt. Die Gesamtvarianz ist also die Summe aus systematischer und unsystematischer Varianz. Je größer der Anteil an systematischer Varianz, desto geringer der Anteil der unsystematischen Varianz und umso besser kann die Gesamtvarianz der beobachteten Daten durch die Prädiktoren x_j erklärt werden.

R^2 kann Werte zwischen 0 und 1 annehmen, wobei ein Wert gleich 0 bedeutet, dass die unabhängigen Variablen nicht geeignet sind, die Unterschiede in der abhängigen Variable zu erklären. Ist der Determinationskoeffizient gleich 1, lassen sich alle Unterschiede in der abhängigen Variable auf die Unterschiede in den unabhängigen Variablen zurückführen (Eid, Gollwitzer & Schmitt, 2013, S. 570; Chatterjee & Price, 1995, S. 64f.). Je mehr unabhängige Variablen in die Regressionsfunktion aufgenommen werden, desto größer wird der Anteil der systematischen Varianz und damit auch der Determinationskoeffizient R^2 (Schendera, 2014, S. 41), auch wenn der zusätzliche Anteil an systematischer Varianz zufallsbedingt ist (Backhaus, Erichson, Plinke et al., 2008, S. 71). Diese Verschätzung des Modellgütemaßes wird bei einer kleinen Zahl von Freiheitsgraden (in der die Stichprobengröße bzw. Fallzahl enthalten ist) noch bedeutender (ebd.). Im Fall großer Prädiktorenzahlen und/oder kleiner Zahlen an Beobachtungswerten wird für die Modellgüte deshalb außerdem ein korrigierter

Determinationskoeffizient R^2_{korrr} angegeben (Urban & Mayerl, 2011, S. 170):

$$R^2_{\text{korrr}} = R^2 - \frac{k \cdot (1 - R^2)}{n - k - 1} \quad (10.4)$$

n = Anzahl der Fälle

k = Anzahl der unabhängigen Variablen

$n - k - 1$ = Anzahl der Freiheitsgrade

Das Ergebnis für den Determinationskoeffizienten lässt sich mit dem F -Test statistisch absichern, indem die Nullhypothese, dass der Determinationskoeffizient null ist (und keine der unabhängigen Variablen einen Einfluss auf die abhängige Variable haben), überprüft wird (Bortz & Schuster, 2010, S. 348; Backhaus, Erichson, Plinke et al., 2008, S. 73; Diaz-Bone, 2013, S. 222):

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - k - 1}{k} \quad (10.5)$$

Liegt der berechnete F -Wert oberhalb eines kritischen F -Wert¹, so ist er statistisch signifikant und die Nullhypothese kann mit entsprechender Irrtumswahrscheinlichkeit zurückgewiesen (Urban & Mayerl, 2011, S. 154) und die Alternativhypothese angenommen werden. Diese behauptet, dass zumindest eine unabhängige Variable einen Einfluss in der Regressionsfunktion hat und der Determinationskoeffizient damit von Null verschieden ist (Diaz-Bone, 2013, S. 222).

Im Anschluss an die globale Prüfung der Regressionsfunktion, erfolgt die der einzelnen Regressionskoeffizienten. Wenn sich hier zeigt, dass eine der Prädiktorvariablen keinen signifikanten Beitrag zur Vorhersage leistet, wird von den Autoren Backhaus, Erichson, Plinke et al. (2008) empfohlen, diese aus der Regressionsfunktion zu entfernen. Ein geeignetes Kriterium, um den Einfluss der Prädiktoren zu analysieren ist der t -Test (ebd., S. 76; Urban & Mayerl, 2011, S. 148; Chatterjee & Price, 1995, S. 15). Mit dem t -Test wird die Nullhypothese, nach der ein einzelner Regressionskoeffizient keinen Einfluss auf das Kriterium ausübt und deshalb gleich 0 ist, überprüft (vgl. Eid, Gollwitzer & Schmitt, 2013, S. 624). Zur Berechnung der Signifikanz einzelner Regressionsparameter mithilfe des t -Testes wird deren Standardfehler benötigt (Urban & Mayerl, 2011, S. 145). Unter der Annahme, dass die Nullhypothese gilt, erhält man die t -Werte, indem die geschätzten Regressionsparameter b_j durch ihren geschätzten Standardfehler s_{b_j} geteilt werden (Bortz & Schuster, 2010, S. 350):

$$t = \frac{b_j}{s_{b_j}} \quad (10.6)$$

Diese Werte folgen bei Gültigkeit des Regressionsmodells einer t -Verteilung, mit den Freiheitsgraden $df = n - k - 1$ (ebd.; Wolf & Best, 2010, S. 620). Zur Ermittlung der Signifikanz (ob die Nullhypothese verworfen werden muss), wird der berechnete empirische t -Wert unter Berücksichtigung einer Irrtumswahrscheinlichkeit mit einem theoretischen t -Wert aus der t -Verteilung verglichen. Ist der Betrag des empirischen t -Werts größer als der theoretische

¹ Der kritische F -Wert, der von dem Signifikanzniveau und den Freiheitsgraden $df_1 = k$ und $df_2 = n - k - 1$ abhängig ist, kann standardisierten Tabellen entnommen werden.

Grenzwert, dann muss die Nullhypothese abgelehnt werden und der geschätzte Regressionsparameter ist signifikant, d. h. die zugehörige Prädiktorvariable hat einen bedeutenden Einfluss auf die Kriteriumsvariable (vgl. Urban & Mayerl, 2011, S. 149).

10.1.2 Bestimmung der Merkmalsausprägungen

Da sich die Steuerungsphase an die Explorationsphase anschließt und angenommen wird, dass das Verhalten in der Explorationsphase einen Einfluss auf die Steuerungsleistung hat, wird für jede Phase separat eine Itemanalyse durchgeführt. Das Verhalten in der Explorationsphase wird dabei bestimmt durch die Ausprägungen der als schwierig angenommenen Systemmerkmale, das Verhalten in der Steuerungsphase außerdem durch die Ausprägung der Merkmale, die durch die konkrete Problemstellung resultieren. Um im Sinne der Konstruktrepräsentation überprüfen zu können, ob die a-priori angenommenen Merkmale auch tatsächlich die Schwierigkeit im Umgang mit den technischen Systemen beeinflussen, müssen die Ausprägungen der System- und Situationsmerkmale für jedes technische Gerät klar definiert sein. Hierzu wird zunächst die Vorgehensweise bei der Ermittlung der Merkmalsausprägungen in den einzelnen Items beschrieben, bevor dann eine Übersicht über die Ausprägungen der Merkmale in den jeweiligen Items gegeben wird, die in den Regressionsanalysen verwendet werden.

Die Systemmerkmale, von denen vor allem in der Explorationsphase ein schwierigkeitserzeugender Einfluss vermutet wird, lassen sich zur besseren Operationalisierung in die Komponenten der Hardware und Software einteilen und dann den Merkmalskategorien Komplexität, Vernetztheit, Dynamik und Transparenz zuordnen. Für eine bessere Übersicht werden zunächst die Ausprägungen auf Seiten der Hardware, dann die Ausprägungen seitens der Software und zuletzt die Ausprägungen der Situationsmerkmale beschrieben.

Hardware

Die Ausprägungen der Variablen in den Kategorien Komplexität, Vernetztheit und Dynamik können eindeutig durch Zählen festgestellt werden. Die Variable *Anzahl der Bedienelemente* umfasst alle Tasten und Regler der Hardware, die betätigt werden können, um mit den technischen Systemen zu interagieren. Diese Elemente können gezählt werden. Bedienelemente, die prinzipiell die gleiche Bedienhandlung ausführen, sie aus Einfachheitsgründen aber in unterschiedliche Richtungen aufteilen, werden als ein Bedienelement gezählt. Als Beispiel sind die beiden Tasten *plus* und *minus* zur Einstellung der Lautstärke an einem Radio zu nennen, die als ein Bedienelement gezählt werden.

Die Vernetzung eines technischen Systems auf Hardwareebene wird erreicht, wenn für eine große Anzahl an Bedienhandlungen nur wenige Bedienelemente zur Verfügung stehen. Um alle Bedienhandlungen ausführen zu können, müssen Bedienelemente mit mehreren Funktionen belegt sein. Diese können dann erreicht werden, indem Bedienelemente mehrfach gedrückt oder in unterschiedlichen Systemzuständen (bspw. ausgeschalteter Systemzustand) betätigt werden. Wie vernetzt ein technisches Gerät auf Hardwareebene ist, wird berechnet durch die Anzahl der mehrfachbelegten Bedienelemente bezogen auf die Gesamtzahl der Bedienelemente und stellt ein relatives Maß der *Mehrfachbelegung von Bedienelementen* dar, das Werte zwischen 0 und 1 annehmen kann.

Die Merkmalskategorie *Transparenz* beschreibt wie sichtbar und verständlich die Hardwarekomponenten eines technischen Systems und ihre Betätigungen sind. Die Variable *Beschriftung der Bedienelemente* kennzeichnet die Verständlichkeit von Beschriftungen und Symbole für Tasten und Regler für die Nutzer. Die Ausprägung dieser Variable lässt sich jedoch nicht wie die der anderen Merkmalskategorien durch objektives Zusammenzählen ermitteln, sondern hängt mit individuellen Voraussetzungen wie die Vorerfahrungen der Probanden zusammen. Um die Merkmalsausprägung dennoch zu bestimmen, wurde im Rahmen einer Bachelorarbeit an der Universität Duisburg-Essen eine Wissensabfrage zu der Verständlichkeit der Bedienelementebeschriftungen der eingesetzten technischen Geräte durchgeführt. Der hierzu erstellte Fragebogen enthält zu jeder der Simulationen ein Screenshot, auf dem das jeweilige technische Gerät mit all seinen Bedienelementen auf Hardwareebene und ihren Beschriftungen und Symbolen zu sehen ist. Die Befragten mussten für jedes abgebildete Bedienelement angeben, welche Funktion sie bei Betätigung des jeweiligen Elementes erwarten würden. In dem Fragebogen wurden keine Antwortmöglichkeiten vorgegeben, sondern diese mussten von den Befragten frei formuliert werden. Die gegebenen Antworten wurden im Anschluss an die Befragung danach bewertet, ob sie richtig oder falsch waren. Pro befragter Person wurden die Richtigantworten aufsummiert und durch die Anzahl der vorhandenen Bedienelemente dividiert. Aus diesen individuellen Maßen für die Verständlichkeit der Beschriftung der Bedienelemente wurde der Mittelwert über alle individuellen Maße berechnet. Der so ermittelte Wert stellt das Maß für die Ausprägung der *Verständlichkeit der Bedienelementebeschriftung* eines technischen Gerätes dar.

Die zweite Variable des Merkmals *Transparenz*, die *Rückkopplung* von Bedienelementen beschreibt die Sichtbarkeit der erfolgten Bedienhandlungen. Die Rückmeldung von betätigten Reglern kann dabei hardwareseitig durch die veränderte Stellung von Reglern, bei betätigten Tasten oder Knöpfe durch akustische oder optische Signale wahrgenommen werden. Die Ausprägung der Variable *Rückkopplung* in einem technischen System wird bestimmt durch die Anzahl der Bedienelemente, die bei Betätigung eine sichtbare Rückmeldung geben, bezogen auf die Gesamtzahl der in einem System vorhandenen Bedienelemente.

Eine weitere Variable, die die *Transparenz* von technischen Systemen erhöht und damit die Schwierigkeit im Umgang mit diesen Geräten beeinflusst, stellt die *Anzahl der Bedienelementarten* dar. Je mehr unterschiedliche Bedienelemente zur Verfügung stehen, die durch ihr spezifisches Aussehen die Bedienung eines Gerätes selbsterklärender machen, desto einfacher wird die Handhabung. In den im Rahmen des technischen Problemlösekompetenztests entwickelten Simulationen von technischen Geräten sind maximal drei unterschiedliche *Arten von Bedienelementen* implementiert. Alle Geräte, die über Bedienelemente verfügen, sind mindestens mit Tasten bzw. Knöpfen ausgestattet.

Tabelle 10.2: Ausprägungen der schwierigkeitsbestimmenden Systemmerkmale der Hardware in den einzelnen technischen Geräten

Item	Komplexität	Vernetztheit	Transparenz		
	Anzahl Bedienelemente	Mehrfachbelegung Bedienelemente	Beschriftung Bedienelemente	Rückkopplung	Anzahl Bedienelementarten
(1) Waschmaschine	10	.30	.81	.70	1
(2) Parkscheinautomat	5	.00	.98	.20	2
(3) Spülmaschine	16	.31	.50	.38	1
(4) Fotoentwicklungsautomat	3	.00	.97	.33	2
(5) Multifunktionsdrucker	7	.00	.71	.43	1
(6) Kaffeemaschine	11	.00	.72	.00	3
(7) Küchenradio	15	.33	.63	.00	1
(8) Navigationsgerät	18	.06	.86	.00	1
(9) Fotodrucker	12	.08	.72	.08	2
(10) Radiowecker	10	.20	.50	.00	2
(11) Festnetztelefon	12	.17	.66	.67	1
(12) Rollladensteuerung	7	.00	.81	.00	1
(13) Heizungssteuerung	5	.80	.38	.00	1
(14) Mobiltelefon	9	.00	.99	.00	1
(15) Internetrouter	n. b.	n. b.	n. b.	n. b.	1

Anmerkungen: n. b. = nicht bewertbar

Einige Geräte haben zusätzlich Schieberegler, über die eine Interaktion mit den Systemen erfolgen kann. Bei den Systemen, für deren Bedienung außerdem andere Gegenstände benötigt werden, wird deren Verwendung durch die in Computersimulationen intuitiv zu gebrauchende Drag and Drop-Funktion realisiert. Diese *Art von Bedienelement* ist beispielweise in dem Item des Parkscheinautomaten umgesetzt worden, indem das für die Ziehung eines Parkscheins benötigte Kleingeld per Drag and Drop in dem dafür vorgesehenen Geldeinwurfschlitze gebracht werden muss.

Tabelle 10.2 zeigt für die Systemmerkmale der Hardware die jeweiligen Ausprägungen in den einzelnen simulierten technischen Geräten. In der Simulation des technischen Gerätes Internetrouter ist keine Hardware zu sehen, sondern allein das Softwareinterface, so dass die Ausprägung von Variablen, die die Bedienelemente der Hardware betreffen, nicht bewertet werden können.

Software

Analog zur Systemkomponente der Hardware lassen sich die Ausprägungen der Variablen der Merkmalskategorien Komplexität, Vernetztheit und für die Systemkomponente Software zusätzlich die Ausprägung der Variablen der Kategorie Dynamik durch Zählen bestimmen. Die *Menütiefe* gibt die maximale Anzahl der Ebenen an, durch die die Nutzer in dem Menü navigieren können, wobei die oberste Ebene eines Menüs, das sogenannte Hauptmenü, nicht mitgezählt wird. Die *Menübreite* beschreibt die maximale Anzahl der Menüpunkte auf einer Navigationsebene. Eine weitere Variable, die zu einer Komplexitätserhöhung des technischen Systems führt, ist die *Anzahl der Zustände*, die ein solches System annehmen kann. Der Zustand eines Systems wird beschrieben durch Zustandsgrößen, durch deren Änderungen es möglich ist, einen Zustand in einen anderen Zustand zu überführen. Die Anzahl aller möglichen Systemzustände lässt sich jedoch nur aus der Programmieroberfläche entnehmen, da einige Systemzustände für die Nutzer nicht sichtbar sind. Die tatsächlich für die Nutzer *sichtbaren Zustände* werden als eine separate Variable der Komplexität erfasst. Die Anzahl der Variablen, die als Zustandsgrößen das System beschreiben und mit denen Zustandsänderungen herbeigeführt werden, beschreiben ebenso wie die Zustandsanzahl den quantitativen Umfang und damit die Komplexität eines technischen Systems. Auch bei der Variablenanzahl wird unterschieden zwischen *allen vorhandenen Variablen* in einem System und den von den Nutzern sichtbaren und auch *veränderbaren Variablen*. Die Ausprägung beider Merkmale können der Programmierumgebung entnommen werden.

Die Merkmalskategorie Vernetztheit auf Ebene der Systemsoftware wird durch die Variablen *Menüvernetzung*, *vorhandene Variablenbedingungen* und *Reversibilität* operationalisiert. In einem stark vernetzten Menü sind Menüpunkte aus einer Menüebene mit anderen Menüpunkten aus einer anderen Menüebene miteinander verknüpft. Die Anzahl dieser Verknüpfungen lässt sich zählen und durch den Bezug zu der Gesamtzahl aller Menüpunkte erhält man als relatives Maß die Ausprägung der *Menüvernetzung*. Kann eine Variable nur verändert werden, wenn eine andere Variable einen bestimmten Wert aufweist, so liegt eine *Variablenbedingung* vor. Je größer die Anzahl solcher Bedingungen, desto stärker ist die Vernetzung des technischen Systems. Lassen sich vorgenommene Interaktionen mit einem technischen System, wie beispielsweise die Navigation durch ein Menü oder die Veränderung von Variablen direkt wieder rückgängig machen, so handelt es sich um eine reversible Bedienhandlung. Müssen

Nutzer erst durch alle Menüpunkte navigieren, um wieder an den gewünschten Ausgangspunkt zu gelangen, fehlt die Reversibilität und die Bedienung des entsprechenden Gerätes wird erschwert. Das Maß für die Ausprägung der *Reversibilität* wird gebildet, indem die Anzahl der reversiblen Interaktionen durch die Gesamtzahl aller Interaktionen geteilt wird.

Die Merkmalskategorie Dynamik kann einzig mit Variablen auf Seiten der Software realisiert werden. Die Ausprägung der *zeitverzögerten Reaktionen* eines technischen Systems auf Bedienhandlungen in einem Item lässt sich durch ihre Anzahl in dem entsprechenden Gerät erfassen. Die Dauer der Zeitverzögerung spielt zwar ebenfalls eine entscheidende Rolle für die Schwierigkeit im Umgang mit einem technischen Gerät, diese wurde bei der Entwicklung der Items jedoch nicht berücksichtigt, so dass keine Varianz in der Verzögerungsdauer existiert. Die innerhalb eines Systems vorhandenen *Eigendynamiken* werden ebenfalls durch Zählen ermittelt. Die Größe des Zeitraumes innerhalb der die Nutzer mit dem System interagieren können, ohne dass eine Eigendynamik wirksam wird, wurde ebenfalls nicht variiert.

Die Ausprägungen der *Menüverständlichkeit* und der *Rückkopplung* von getätigten Bedienhandlungen durch das Menü als Variablen der Transparenz werden auf die gleiche Weise bestimmt, wie die Verständlichkeit der Bedienelementebeschriftung und die Rückkopplung durch die Bedienelemente. Die Ergebnisse aus der Durchführung der beschriebenen Verfahren zur Bestimmung der Merkmalsausprägungen für die Software der technischen Systeme sind in Tabelle 10.3 und Tabelle 10.4 zusammengefasst.

Tabelle 10.3: Ausprägungen der schwierigkeitsbestimmenden Systemmerkmale der Software in den einzelnen technischen Geräten

Item	Komplexität						Vernetztheit		
	Menütiefe	Menübreite	Anzahl Zustände	Anzahl sichtbarer Zustände	Anzahl Variablen	Anzahl veränderbarer Variablen	Menü- vernetzung	vorhandene Bedingungen	Reversibilität
(1) Waschmaschine	1	3	13	4	12	7	1.50	5	.66
(2) Parkscheinautomat	0	1	10	2	15	2	.00	1	.80
(3) Spülmaschine	0	0	12	7	20	7	1.43	1	.75
(4) Fotoentwicklungsautomat	5	4	25	18	179	6	2.43	2	.43
(5) Multifunktionsdrucker	3	4	36	26	277	23	2.87	1	.57
(6) Kaffeemaschine	0	0	22	8	32	10	.00	4	.33
(7) Küchenradio	2	7	37	13	47	28	3.42	0	.09
(8) Navigationsgerät	5	5	55	49	198	28	7.66	1	.33
(9) Fotodrucker	3	3	16	11	108	58	1.08	3	1.00
(10) Radiowecker	1	2	29	6	11	6	2.00	1	.60
(11) Festnetztelefon	3	8	63	55	194	14	1.45	0	.55
(12) Rollladensteuerung	4	4	51	28	51	15	1.04	1	.35
(13) Heizungssteuerung	0	3	19	14	47	24	1.27	4	.38
(14) Mobiltelefon	4	2	28	12	2	0	2.16	0	.89
(15) Internetrouter	2	5	43	41	103	65	2.75	0	.98

Tabelle 10.4: Ausprägungen der schwierigkeitsbestimmenden Systemmerkmale der Software in den einzelnen technischen Geräten

Item	Dynamik		Transparenz	
	Zeitverzögerung	Eigendynamik	Menüverständ.	Rückkopplung
(1) Waschmaschine	1	0	.80	.10
(2) Parkscheinautomat	0	0	1.00	1.00
(3) Spülmaschine	0	0	.21	.94
(4) Fotoentwicklungsautomat	1	0	.56	1.00
(5) Multifunktionsdrucker	1	0	.20	1.00
(6) Kaffeemaschine	0	0	.21	1.00
(7) Küchenradio	0	2	.50	1.00
(8) Navigationsgerät	0	0	.58	.88
(9) Fotodrucker	2	0	.61	1.00
(10) Radiowecker	0	0	.66	.80
(11) Festnetztelefon	0	0	.06	.67
(12) Rollladensteuerung	0	0	.37	1.00
(13) Heizungssteuerung	0	2	.28	.97
(14) Mobiltelefon	0	0	.60	1.00
(15) Internetrouter	1	0	.53	1.00

Situationsmerkmale

Zusätzlich zu den Systemmerkmalen beeinflusst in der Steuerungsphase eine konkrete Aufgabenstellung den Umgang mit den technischen Systemen. Die Aufgabenstellungen unterscheiden sich in ihrer Komplexität zunächst darin, wie viele Schritte nötig sind und wie viele Variablen verändert werden müssen, um die gegebene Aufgabe zu lösen. Die in dem System vorhandene Vernetzung, die sich durch eine *Mehrfachbelegung von Bedienelementen* oder durch *Variablenbedingungen* zeigt, wirkt in der Steuerungsphase als schwierigkeitsbestimmend, wenn sie für die Problemlösung auch berücksichtigt werden muss. Das Komplexitätsmerkmal *Anzahl der Lösungsschritte* wird für eine differenziertere Analyse noch einmal nach der Gesamtzahl aller benötigten Lösungsschritte und der Zahl voneinander verschiedenen Lösungsschritte unterschieden. Die durch Zählen ermittelten Ausprägungen der Merkmale, die in der Steuerungsphase wirksam werden, finden sich in Tabelle 10.5.

Mit den in den Tabellen 10.2 - 10.5 aufgeführten Merkmalsausprägungen in den jeweiligen Items und den nach der Durchführung des Testes bestimmten empirischen Itemschwierigkeiten lassen sich mithilfe der Regressionsanalyse die Regressionsgewichte berechnen. Anhand der standardisierten Regressionsgewichte können die einzelnen als schwierigkeitsbestimmend angenommenen Merkmale miteinander verglichen werden, um Aussagen über die Stärke des Einflusses eines Merkmals auf die Itemschwierigkeit treffen zu können.

Tabelle 10.5: Ausprägungen der schwierigkeitsbestimmenden Situationsmerkmale der einzelnen technischen Geräte

Item	Komplexität			Vernetztheit	
	Anzahl aller Lösungsschritte	Anzahl unterschiedlicher Lösungsschritte	einzustellende Variablen	zu berücksichtigende mehrfachbelegte Bedienelemente	zu berücksichtigende Bedingungen
(1) Waschmaschine	11	11	4	1	0
(2) Parkscheinautomat	39	6	2	0	1
(3) Spülmaschine	5	4	1	2	0
(4) Fotoentwicklungsautomat	11	12	4	0	1
(5) Multifunktionsgerät	4	6	1	0	0
(6) Kaffeemaschine	11	13	2	0	3
(7) Küchenradio	9	8	2	0	0
(8) Navigationsgerät	101	15	3	1	1
(9) Fotodrucker	19	19	2	0	1
(10) Radiowecker	71	7	5	2	2
(11) Festnetztelefon	21	9	2	1	0
(12) Rollladensteuerung	10	9	1	0	0
(13) Heizungssteuerung	99	38	16	0	2
(14) Mobiltelefon	5	3	0	0	0
(15) Internetrouter	15	23	9	0	0

10.2 Dimensionalität technischer Problemlösekompetenz

Ein weiterer Aspekt zur Überprüfung der Konstruktvalidität betrifft die logisch-inhaltliche Analyse des Konstruktes, das heißt die Analyse der Konstruktstruktur. Die Frage nach der Struktur eines Konstruktes ist gleichbedeutend mit der Frage nach den Subdimensionen (Facetten, Faktoren), die ein Konstrukt umfasst. Im Abschnitt 6 wurde das Kompetenzkonstrukt *problemlösender Umgang mit technischen Alltagsgeräten* anhand des zu erreichenden Interaktionsziels in den explorierenden Umgang mit technischen Systemen (Systemexploration) und den steuernden Umgang mit technischen Systemen (Systemsteuerung) differenziert. Die Kompetenz in der jeweiligen Umgangsform stellt dabei eine eigenständige Subdimension der zu erfassenden Problemlösekompetenz dar, wobei *nicht* postuliert wird, dass es eine zentrale Dimension, ein „g“-Faktor gibt, der als gemeinsame Ursache für das kompetente Verhalten in den jeweiligen Umgangsformen angesehen werden kann. Vielmehr handelt es sich bei dem zu erfassenden Konstrukt um ein Sammelbegriff (vgl. Steinmetz, 2014, S. 38) für die genannten Subdimensionen *Systemexploration* und *Systemsteuerung*.

Diese Subdimensionen stellen latente Merkmale dar, für deren Erfassung im Abschnitt 8.5.1 entsprechende Indikatoren (Items) vorgestellt wurden. Die angenommenen Beziehungen zwischen den latenten Variablen und ihren jeweiligen manifesten Indikatorvariablen werden in sogenannten *Messmodellen* spezifiziert, während die Zusammenhänge zwischen den jeweiligen latenten Variablen in sogenannten *Strukturmodellen* konkretisiert werden (Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 950). Im Rahmen der Validitätsuntersuchung sind die (sich auf die Konstruktstruktur beziehenden) theoretischen Annahmen, die im Abschnitt 6.6 getroffen wurden, empirisch zu prüfen. Dabei werden sowohl die Anzahl der postulierten Dimensionen als auch die Zuordnung der Items zu den einzelnen Dimensionen analysiert. Dieser Aspekt der Konstruktvalidität wird auch als faktorielle Validität bezeichnet (Döring & Bortz, 2016a, S. 479).

Von Items, die nur einer latenten Variable zugeordnet sind und damit ein und dasselbe Merkmal messen sollen, kann erwartet werden, dass sie inhaltlich ähnlich (homogen) sind und somit empirisch relativ hoch miteinander korrelieren (ebd.). Deshalb können für eine erste Betrachtung der Dimensionalität eines Tests oder einer Skala die bivariaten Item-Interkorrelationen analysiert werden, die bei vorliegender Eindimensionalität positive, signifikant von Null abweichende Werte aufweisen sollten. Ein quantitatives Maß für die Homogenität lässt sich über den Mittelwert der Item-Interkorrelationen bilden; dieser sollte bei einem homogenen Gesamttest zwischen .20 und .40 betragen (ebd.).

Die gleichzeitige Betrachtung der Homogenität von Items innerhalb eines Messmodells (konvergente Validität) und der Heterogenität von Items zwischen Messmodellen (diskriminante Validität) wird mit *konfirmatorischen Faktorenanalysen* möglich (Urban & Mayerl, 2014, S. 136). Konfirmatorische Faktorenanalysen (CFA – Confirmatory Factor Analysis) führen die Zusammenhänge zwischen den manifesten Variablen auf wenige dahinterliegende latente Variablen (Faktoren) zurück. Dazu wird schon vor der Analyse das Modell der konfirmatorischen Faktorenanalyse mit seiner Anzahl an Faktoren und der Zuordnung der manifesten Variablen zu den Faktoren theoretisch definiert (Bühner, 2011,

S. 381). Diese bereits a priori festgelegte Zusammenhgangsstruktur wird in der CFA auf ihre Passung zu den empirischen Daten hin untersucht (Arzheimer, 2016, S. 41) und gegebenenfalls mit alternativen Modellen (sofern diese ebenfalls theoretisch abgeleitet wurden) verglichen (Bühner, 2011, S. 380).

Die Stärke des Einflusses, den eine latente Variable ξ auf eine manifeste Variable x ausübt, wird als Ladung λ bezeichnet. Ziel der CFA ist unter anderem die Schätzung dieser Faktorladungen, die analog zu Regressionskoeffizienten zu interpretieren sind (Urban & Mayerl, 2014, S. 54). Wird der Wert der latenten Variable ξ um eine Einheit erhöht, so steigt der Wert der manifesten Variable um den Wert λ (Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 951). Ein Vergleich der so zu interpretierenden unstandardisierten Faktorenladungen ist aufgrund der unterschiedlichen Metriken nicht möglich, weshalb die Faktorladungen oftmals standardisiert werden. Je größer der Wert für eine standardisierte Faktorladung ist, desto besser lässt sich die manifeste Variable durch die latente Variable erklären. Durch Quadrieren der standardisierten Faktorladung erhält man die entsprechende Varianz der manifesten Variable, die durch die latente Variable erklärt werden kann (als Kommunalität bezeichnet). Somit gilt die Höhe der standardisierten Faktorladungen als Hinweis auf die Validität des angenommenen Konstruktes (Steinmetz, 2014, S. 45). Als Daumenregel für anzustrebende standardisierte Faktorladungen nennen Urban & Mayerl (2014, S. 54) Werte $> .50$ und Wang & Wang (2012, S. 39) Werte $\geq .30$.

Neben den Faktorladungen λ werden in der CFA weitere Modellparameter geschätzt, zu denen die Varianzen der latenten Variablen $\text{Var}(\xi)$, die Messfehlervarianzen $\text{Var}(\delta)$ der einzelnen manifesten Variablen und eventuell vorhandene Kovarianzen zwischen mehreren latenten Variablen $\text{Cov}(\varphi)$ gehören. Um die unbekanntes Modellparameter, deren Anzahl abhängig von dem postulierten Modell ist, ermitteln zu können, müssen einerseits genügend empirische Informationen zur Verfügung stehen und andererseits müssen für die latenten Variablen Einheiten festgelegt werden (Kline, 2011, S. 124). Die erste Bedingung ist erfüllt, wenn die Anzahl s der Einträge in der sich aus den empirischen Daten ergebenden Kovarianzmatrix S der manifesten Variablen gleich oder größer als die Anzahl t der zu schätzenden Modellparameter ist, so dass die Anzahl der Freiheitsgrade $df = s - t \geq 0$ ist. Mit der Anzahl k an manifesten Variablen ergibt sich für die Anzahl empirischer Informationen $s = \frac{k(k+1)}{2}$ (Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 964). Ist der Freiheitsgrad $df = 0$, ist das Modell gerade identifiziert, während das Modell bei einem Freiheitsgrad $df > 0$ überidentifiziert ist (Kline, 2011, S. 125). Für die Erfüllung der zweiten Bedingung, nach der die latenten Variablen eine Einheit aufweisen müssen, um deren Varianzen schätzen zu können, gibt es zwei Möglichkeiten. Eine Möglichkeit ist, die Faktorladung einer manifesten Variable der entsprechenden latenten Variablen auf den Wert 1 zu setzen, so dass die latente Variable deren Skalierung übernimmt (Moosbrugger & Kelava, 2012b, S. 336). Eine alternative Möglichkeit ist die Varianz der latenten Variable auf den Wert 1 zu setzen und diese damit zu standardisieren. Dieser Ansatz hat für die vorliegende Arbeit den Vorteil, dass alle, der noch unbekanntes Faktorladungen geschätzt und für diese Signifikanztests durchgeführt werden können (Bühner, 2011, S. 398).

Neben dem Fixieren der Varianzen der latenten Variablen auf einen Wert können je nach messtheoretischen Hypothesen weitere Restriktionen vorgenommen werden. Alle in der CFA

spezifizierten Modelle sind Messmodelle der Klassischen Testtheorie, die einen linearen Zusammenhang zwischen der manifesten und der latenten Variable annehmen. Hier werden ähnlich wie in der Item-Response-Theorie die Messmodelle unter anderem danach unterschieden, wie viele Parameter variieren dürfen und demnach frei geschätzt werden können. Das restriktivste Messmodell ist das τ -äquivalente Modell, das gleiche Schwierigkeit und gleiche Trennschärfe für die manifesten Variablen voraussetzt (Rost, 2004, S. 110). In der CFA werden deshalb alle Faktorladungen auf den gleichen Wert und alle Messfehlervarianzen auf den gleichen Wert fixiert (Reinecke, 2014, S. 98). Das essentiell τ -äquivalente Messmodell geht nur von der Annahme gleicher Trennschärfen aus; die Schwierigkeiten der einzelnen manifesten Variablen können unterschiedlich sein (Rost, 2004, S. 108). In dem spezifizierten essentiell τ -äquivalenten Messmodell werden deshalb gleiche Werte für die Faktorladungen angenommen (Reinecke, 2014, S. 98). Das am wenigsten restriktive Messmodell ist das kongenerische Modell, bei dem sich die manifesten Variablen sowohl hinsichtlich der Schwierigkeit, als auch hinsichtlich der Trennschärfe unterscheiden dürfen (Rost, 2004, S. 109). Es werden also weder Restriktionen für die Faktorladungen noch für die Messfehlervarianzen vorgenommen (Reinecke, 2014, S. 98). Das kongenerische Messmodell ist das am häufigsten zugrunde gelegte Messmodell, obwohl es ausschließlich für manifeste Variablen mit metrischer Skala geeignet ist (Rost, 2004, S. 110).

Die Ermittlung der Modellparameter, also die Stärken und Richtungen der angenommenen Zusammenhänge in dem Messmodell, erfolgt aus den empirischen Daten (Arzheimer, 2016, S. 57). Ist das Modell genau identifiziert ($df = 0$) können die Modellparameter eindeutig berechnet werden, ist das Modell hingegen überidentifiziert ($df > 0$) und enthält Restriktionen, müssen die Parameter geschätzt werden (Reinecke, 2014, S. 100). Die zur Verfügung stehenden Schätzmethode bestimmen die Modellparameter so, dass die sich aus ihnen ergebende theoretische Kovarianzmatrix möglichst gut mit der tatsächlich beobachteten Kovarianzmatrix übereinstimmt, die Diskrepanz zwischen ihnen also möglichst gering wird (Moosbrugger & Kelava, 2012b, S. 337). In Abhängigkeit von Stichprobengröße, Verteilung der Itemantworten sowie dem Skalenniveau der Items stehen verschiedene, auch als Diskrepanzfunktionen (Reinecke, 2014, S. 100) bezeichnete Schätzmethode zur Verfügung (Bühner, 2011, S. 407).

Mit dem am häufigsten eingesetzten *Maximum-Likelihood* (ML)-Schätzverfahren werden die Parameter so geschätzt, dass die Wahrscheinlichkeit die empirische Kovarianzmatrix zu beobachten maximal wird (Arzheimer, 2016, S. 58). Da diese als Datenbasis zur Parameterschätzung dient und Kovarianzen multivariat normalverteilte Variablen voraussetzen, kann die ML-Methode auch nur unter dieser Bedingung eingesetzt werden (Bühner, 2011, S. 431). Multivariate Normalverteilung setzt voraus, dass jede einzelne Variable individuell normalverteilt ist, was nur für kontinuierliche, d. h. annähernd intervallskalierte Daten zutreffen kann (Arzheimer, 2016, S. 59). Wird die Normalverteilungsannahme verletzt, werden die Standardfehler der Modellparameter unterschätzt und es kommt zu einer Überhöhung des χ^2 -Wertes (Finney & DiStefano, 2006). Im Fall leichter bis mittelgradiger Abweichungen von der Normalverteilung kann das von Satorra & Bentler (1988) vorgestellte robuste ML-Verfahren (MLM) eingesetzt werden. Da die ML-Schätzung vor allem empfindlicher auf Abweichungen der Kurtosis als auf Abweichungen der Schiefe reagiert (Urban & Mayerl, 2014, S. 142), wird bei der robusten ML-Schätzung der Wert für die

χ^2 -Statistik in Abhängigkeit der Größe der Kurtosis reduziert und analog die Werte der Standardfehler erhöht (ebd., S. 141). Neben dieser sogenannten Satorra-Bentler-Korrektur gibt es außerdem die Yuan-Bentler-Korrektur des χ^2 -Wertes und der Standardfehler, die in dem MLR-Schätzverfahren Verwendung findet und sowohl robust gegenüber Verletzungen der Normalverteilungsannahme als auch gegenüber Verletzungen der Annahme der Unabhängigkeit der Beobachtungen ist (Yuan & Bentler, 2000). Die zu korrigierenden Abweichungen von der Normalverteilung dürfen jedoch nicht zu hoch sein, damit das verwendete robuste ML-Schätzverfahren zu hinreichend genauen und zuverlässigen Schätzwerten führt. Urban & Mayerl (2014, S. 142) geben mit Verweis auf Boomsma & Hoogland (2001, S. 146) Grenzwerte für die univariate Schiefe von bis zu ± 2.0 und für die univariate Kurtosis von bis zu ± 7.0 , mit Verweis auf Byrne (2010) sowie Nevitt & Hancock (2001) Werte von bis zu ± 0.6 für die Schiefe und von bis zu ± 1.0 für die Kurtosis wieder. Der Einsatz robuster ML-Schätzverfahren führt auch bei ordinalen Antwortformaten mit einer geringen Breite (minimal aber 5 Stufen) noch zu zuverlässigen Schätzwerten (Olsson, 1979), auch wenn die Stichprobengröße nur etwa $n = 200$ beträgt (Curran, West & Finch, 1996).

Bei Indikatoren, die weniger als 5 Ausprägungen aufweisen, das heißt auch bei dichotomen Indikatoren führen Schätzverfahren auf Basis der Maximum-Likelihood-Methode zu erheblichen Verzerrungen bei den Schätzwerten (Urban & Mayerl, 2014, S. 145). Hier empfiehlt es sich als Diskrepanzfunktion die Weighted-Least-Square (WLS)-Funktion zu verwenden, die zur Schätzung der Modellparameter Informationen über die Verteilungen aus der gewichteten asymptotischen Kovarianzmatrix nutzt (Reinecke, 2014, S. 106). Damit werden vorab keine Verteilungsannahmen gemacht, was ein Vorteil der Schätzmethode ist. Jedoch ist für die Berechnung der asymptotischen Kovarianzmatrix eine hohe Stichprobengröße erforderlich, weshalb häufig robuste WLS-Schätzverfahren (z. B. WLSMV) zum Einsatz kommen, die auch für kleinere Stichprobengrößen geeignet sind (Muthén, 1993).

Gemeinsames Ziel der genannten Schätzmethoden ist die Modellparameter für das postulierte Mess- bzw. Strukturmodell so zu bestimmen, dass die Diskrepanz zu den empirischen Daten dabei möglichst gering wird (Moosbrugger & Kelava, 2012b, S. 337). Als Indiz für die Stabilität der geschätzten Modellparameter gelten die Standardfehler (SE) eines jeden geschätzten Modellparameters. Die Standardfehler sind Schätzwerte für die Varianz der einzelnen Parameterschätzungen (Urban & Mayerl, 2014, S. 58). Je geringer der Standardfehler, desto eher handelt es sich bei der Parameterschätzung um ein statistisch signifikantes Ergebnis. Mit der Division des geschätzten Parameterwerts durch seinen Standardfehler ergibt sich der t -Wert, dessen Verteilung sich bei großen Stichproben der z -Verteilung annähert. t -Werte, die größer als $|1.96|$ sind, zeigen ein signifikantes Ergebnis an (bei $\alpha = .05$) (ebd., S. 59).

Selbst bei bestmöglicher Schätzung der Parameter verbleiben jedoch häufig Abweichungen zwischen modellimplizierter und empirischer Kovarianzmatrix (Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 957), so dass die Größe der Diskrepanz erste Hinweise auf die Modellgüte gibt. Werden die Residuen, also die einzelnen Abweichungen zwischen modellimplizierter und empirischer Kovarianzmatrix durch ihren jeweiligen Standardfehler dividiert, ergeben sich ebenfalls z -verteilte Werte, die nach Urban & Mayerl (2014, S. 89) kleiner als $|2.58|$ sein sollten.

Ein Gesamtmaß für die Modellabweichungen ergibt sich aus der Diskrepanz F_{min} , die nach der Schätzung der Modellparameter von der jeweils gewählten Diskrepanzfunktion noch verbleibt. Dazu wird aus dem Diskrepanzwert die Prüfgröße $T = (N - 1) F_{min}$ gebildet, die bei ausreichend hoher Stichprobe χ^2 -verteilt ist (Bühner, 2011, S. 407). Je größer der so ermittelte χ^2 -Wert ist, desto eher liegt eine Diskrepanz zwischen dem Theoriemodell und den empirischen Daten vor (Reinecke, 2014, S. 102). Ist der χ^2 -Wert außerdem signifikant ($p < .01$), sind diese Abweichungen nicht zufällig (Bühner, 2011, S. 407). Für eine gute Modellpassung sollte der χ^2 -Wert also möglichst klein sein und nach einer Daumenregel etwa kleiner als zweimal die Anzahl der Freiheitsgrade df ($\chi^2 = 2 df$ bzw. $\frac{\chi^2}{df} < 2$) (Moosbrugger & Kelava, 2012b, S. 337). Da der χ^2 -Wert nicht nur aus der Diskrepanz, sondern auch aus der Stichprobengröße berechnet wird, kann es sein, dass bei einer großen Stichprobe auch kleine Abweichungen signifikant werden (Arzheimer, 2016, S. 62f.), umgekehrt kann es aber auch sein, dass durch eine zu kleine Stichprobe größere Abweichungen nicht signifikant werden (Bühner, 2011, S. 420). Aufgrund der stichprobenabhängigen Empfindlichkeit des χ^2 -Wertes gegenüber Modellabweichungen empfehlen Urban & Mayerl (2014, S. 93) zwei weitere deskriptive Modellgütemaße zu berichten.

Eines der alternativen Modellgütemaße ist der *Comparative Fit Index* (CFI), der nicht die modellimplizierte und die empirische Kovarianzmatrix miteinander vergleicht, sondern das spezifizierte Theoriemodell mit einem sogenannten Unabhängigkeitsmodell (Basismodell), bei dem keine Beziehungen zwischen den Variablen angenommen werden, die entsprechenden Zusammenhänge demnach auf Null fixiert sind (Wentura & Pospeschill, 2015, S. 205). Der Vergleich der beiden Modelle erfolgt über das von 1 abgezogene Ergebnis eines Bruches, bei dem das spezifizierte Modell im Zähler steht und das Basismodell im Nenner (Kline, 2011, S. 208). Je größer der durch die Aufgabe von Freiheitsgraden resultierende Unterschied zwischen den beiden Modellen ist, desto größer wird der berechnete Wert für den CFI, der maximal 1 werden kann (Bühner, 2011, S. 427). Bei einer akzeptablen Modellpassung sollte der Wert für den CFI $\geq .90$ sein (Raykov & Marcoulides, 2006).

Ein weiteres alternatives Modellgütemaß stellt der *Root Mean Square Error of Approximation* (RMSEA) dar, der die Residuen zwischen der modellimplizierten und der empirischen Kovarianzmatrix ähnlich wie der χ^2 -Wert zu einem Maß zusammenfasst, dabei aber die Anzahl der Freiheitsgrade sowie die Stichprobengröße berücksichtigt (Urban & Mayerl, 2014, S. 96). Werte bis $.10$ gelten als akzeptabel, Werte zwischen $.05$ und $.08$ als Hinweis auf einen guten, Werte kleiner als $.05$ als Beleg für einen sehr guten Fit (Arzheimer, 2016, S. 63f.). Während der χ^2 -Wert bei großen Fallzahlen oft sehr hoch wird, führen zu geringe Fallzahlen dazu, dass der RMSEA-Wert schnell zu groß wird und damit einen schlechten Modellfit anzeigt (Yu, 2002, S. 161). Neben der Höhe des Wertes für den RMSEA sollten außerdem die Grenzen für das 90 % Konfidenzintervall angegeben werden, da sie Aussagen über die Präzision der Schätzung des RMSEA-Wertes zulassen (MacCallum, Browne & Sugawara, 1996). Urban & Mayerl (2014, S. 97) empfehlen für das 90 % Konfidenzintervall einen maximalen Wert von $.05$ für die untere Grenze und einen maximalen Wert von $.08$ für die obere Grenze. Je enger das Konfidenzintervall, desto genauer ist die Schätzung des RMSEA-Wertes. Bei kleinen Stichprobengrößen fällt die Breite des Konfidenzintervalls jedoch generell größer aus (Kline, 2011, S. 206). Ein von Muthén (1998) vorgeschlagenes Modellgütemaß

speziell für kategoriale Daten, für die die Parameter mit den robusten WLSM-Verfahren geschätzt werden, ist der WRMR-Index. Dieser fasst analog zum RMSEA-Wert die Residuen zu einem Abweichungsmaß zusammen (Reinecke, 2014, S. 119). WRMR-Werte < 1.0 zeigen für Modelle mit kategorialen Daten (aber auch mit kontinuierlichen Daten) einen akzeptablen Modellfit an (Yu, 2002, S. 155).

10.3 Strukturmodell zur Verortung der Problemlösekompetenz im Umgang mit technischen Alltagsgeräten

Nachdem die angenommene Konstruktstruktur mittels konfirmatorischer Faktorenanalyse geprüft wurde, kann die Verortung des Konstruktes in das theoretische Kausalmodell (nomologische Netzwerk) vorgenommen werden. Solch ein komplexes Wirkungsgefüge von Variablen, wie es in Abschnitt 6.6 postuliert wird, lässt sich mithilfe von Strukturgleichungsmodellen formalisieren und anschließend empirisch testen (vgl. Weiber & Mühlhaus, 2014, S. 7). Weil die konfirmatorische Faktorenanalyse bereits als ein Sonderfall eines vollständigen Strukturgleichungsmodells gilt (Steinmetz, 2014, S. 73), werden an dieser Stelle lediglich die noch nicht im vorherigen Abschnitt besprochenen methodischen Aspekte berücksichtigt.

Im Gegensatz zu konfirmatorischen Faktorenanalysen, die ihren Fokus auf das Spezifizieren und Testen von Messmodellen legen, rücken im Strukturgleichungsmodell die gerichteten oder kausalen Beziehungen zwischen den latenten Variablen in den Vordergrund (Borg & Staufenbiel, 2007, S. 247). Damit diese in dem Strukturgleichungsmodell analysiert werden können, sind sie zunächst als solche zu konkretisieren. Hierbei geht es nicht wie bei der Korrelation nur darum einen Zusammenhang anzugeben, sondern außerdem darum, eine Aussage über die Richtung des Zusammenhangs zu treffen (Wentura & Pospeschill, 2015, S. 196). Solche Ursache-Wirkungs-Beziehungen von Variablen setzen zwar Korrelationen zwischen ihnen als notwendige Bedingung voraus, Korrelationen allein stellen aber keine hinreichende Bedingung für Kausalität dar (vgl. Rudolf & Müller, 2012, S. 340). Zwei weitere Voraussetzungen, die erfüllt sein müssen, um von Kausalität sprechen zu können, sind das zeitliche Vorangehen der potenziellen Ursache vor der Wirkung sowie der Ausschluss anderer Variableneinflüsse (Arzheimer, 2016, S. 43). Diese lassen sich jedoch nur statistisch kontrollieren, wenn sie bekannt sind, gemessen werden und nicht in Abhängigkeit der kausal zu interpretierenden Variable variieren (ebd., S. 44).

Sind diese Bedingungen erfüllt, lassen sich in Strukturgleichungsmodellen Messfehlereinflüsse in den beobachteten Variablen systematisch berücksichtigen (Christ & Schlüter, 2012, S. 49). Solche zufälligen Messfehler verringern nämlich die Reliabilität von Messungen, wodurch es dann zu einer Verzerrung der geschätzten Zusammenhänge von messfehlerbehafteten Variablen kommt. Aus der von Bollen (1989) gezeigten Abhängigkeit der Korrelation zweier manifester Indikatorvariablen von ihren Reliabilitäten sowie den wahren Korrelationen zwischen ihren latenten Variablen ergibt sich die folgende Formel, nach

der Korrelationskoeffizienten in den Strukturgleichungsmodellen zu korrigieren sind (Urban & Mayerl, 2014, S. 19):

$$r_{FY1,FY2} = \frac{r_{Y1,Y2}}{\sqrt{Rel_{Y1} \cdot Rel_{Y2}}} \quad (10.7)$$

$Y1, Y2$ = Indikatorvariablen

$FY1, FY2$ = latente Variablen

r = Korrelationskoeffizient

Rel = Reliabilität

Somit erlauben Strukturgleichungsmodelle die Schätzung der *wahren* Beziehungen zwischen latenten Variablen, ohne die in den manifesten Variablen beinhalteten zufälligen Messfehler (Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 946; Eid, Gollwitzer & Schmitt, 2013, S. 927).

Die Überführung der im Abschnitt 6.6 theoretisch angenommenen Variablenbeziehungen in ein vollständiges Strukturgleichungsmodell bedarf nach Borg & Staufenbiel (2007, S. 248) der Einhaltung gewisser Regeln, mit denen eine eindeutige Kommunikation über das Modell gewährleistet werden soll. So sollten explizit alle Variablen in das Strukturmodell eingebracht werden, die in dem Theoriemodell bedacht wurden. Das heißt, nicht das Konstrukt des *problemlösenden Umgangs mit technischen Alltagsgeräten* wird als eine latente Variable modelliert, sondern die entsprechenden Subdimensionen *Systemexploration* und *Systemsteuerung* (vgl. Steinmetz, 2014, S. 36). Diese beiden latenten Variablen sind, ebenso wie die Subdimensionen *Exploration* und *Steuerung* des Konstruktes allgemeine Problemlösefähigkeit, die in dem Strukturmodell zu erklärenden Variablen und werden *endogene* Variablen genannt. Endogene Variablen werden in dem grafischen Modell auf der rechten Seite abgebildet (siehe Abbildung 10.1). Auf der linken Seite des Modell stehen die *exogenen* Variablen, also die Variablen, die einen Einfluss auf andere Variablen ausüben und selbst nicht in dem Modell erklärt werden (vgl. Wentura & Pospeschill, 2015, S. 199; Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 953; Rudolf & Müller, 2012, S. 348).

Jede, in dem Modell vorhandene latente Variable muss durch ein Messmodell mit den entsprechenden manifesten Indikatorvariablen verknüpft werden. Die beobachteten Variablen werden somit in einen von der latenten Variable erklärten und einen auf Messfehler oder andere Konstrukte zurückführbaren Anteil aufgeteilt (Bortz & Schuster, 2010, S. 447). Für den nicht durch die latente Variable erklärten Anteil wird pro Indikatorvariable eine Residualvariable (die ebenfalls als latente Variable betrachtet wird) eingeführt und in dem Modell durch einen auf sie gerichteten Pfeil dargestellt (Borg & Staufenbiel, 2007, S. 249). Dem Vorteil einer messfehlerbereinigten Zusammenhangsanalyse zwischen exogenen und endogenen Variablen steht jedoch eine hohe Anzahl unbekannter Modellparameter entgegen, deren stabile Schätzung einen großen empirischen Datensatz benötigt (Arzheimer, 2016, S. 44). Um die Anzahl zu schätzender Modellparameter zu verringern, werden diejenigen Konstrukte, die in der vorliegenden Studie durch einen Fragebogen erfasst wurden, in dem Modell nicht als latente, sondern als manifeste exogene Variablen eingebracht. Sie sind in dem grafischen Modell durch Rechtecke dargestellt.

Nach dem Einbezug aller interessierender Variablen in das Strukturgleichungsmodell sind die Zusammenhänge zwischen ihnen zu präzisieren. Für die exogenen Variablen, für die man in der Regel auf spezifische kausale Hypothesen verzichtet (Bortz & Schuster, 2010, S. 436), werden bivariate Beziehungen angenommen (in dem Modell als gebogene Doppelpfeile dargestellt). Es werden jedoch nur diejenigen exogenen Variablen als korreliert betrachtet, für die in dem Theoriemodell auch ein solcher Zusammenhang begründet wurde. So wird in der vorliegenden Arbeit kein Zusammenhang zwischen Intelligenz und den Skalen der Selbstwirksamkeit (Ausdauer und Aufgeschlossenheit gegenüber Problemlösen) und Need for Cognition und auch kein Zusammenhang zwischen Need for Cognition und dem Vorwissen (ICT-Nutzung und gerätespezifisches Vorwissen) postuliert.

Die kausalen Einflüsse (auch als Effekte bezeichnet) zwischen einzelnen Variablen werden in dem Strukturmodell mit geraden Pfeilen symbolisiert, deren Richtungen nach Ursache und Wirkung definiert werden. Grundsätzlich lassen sich drei verschiedene Arten von Effekten unterscheiden: direkte, indirekte und totale Effekte (Bortz & Schuster, 2010, S. 443). Der direkte Effekt bezeichnet einen Einfluss, den eine Variable direkt auf eine abhängige Variable hat, während indirekte Effekte solche Einflüsse kennzeichnen, die eine Variable vermittelt über eine oder mehrere andere Variablen (sogenannte Mediatorvariablen) auf eine abhängige Variable ausübt (Eid, Gollwitzer & Schmitt, 2013, S. 929). Während es zwischen zwei Variablen nur einen direkten Effekt geben kann, sind zwischen zwei Variablen mehrere indirekte Effekte möglich. Der totale Effekt ist der Gesamteffekt, den eine Variable auf eine andere haben kann. Er wird durch die Summe der direkten und indirekten Effekte gebildet (Bortz & Schuster, 2010, S. 443).

Zwischen den endogenen Variablen selbst werden keine Kovarianzen angenommen (Borg & Staufenbiel, 2007, S. 250), da diese eine Funktion der Varianzen und Einflüsse der exogenen Variablen sind (Steinmetz, 2014, S. 7). Da aber nicht davon auszugehen ist, dass alle Variablen erfasst und in dem Strukturmodell berücksichtigt wurden, die einen Einfluss auf die allgemeine Problemlösefähigkeit und die technische Problemlösekompetenz ausüben, verbleiben unerklärte Varianzanteile, die durch die Angabe von Residualvariablen zum Ausdruck kommen (Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 953). Statt Kovarianzen zwischen den endogenen Variablen werden Kovarianzen zwischen ihren Residualvariablen angenommen, die „weitere, nicht im Modell adressierte, gemeinsame Ursachen der betreffenden Variablen“ repräsentieren (Steinmetz, 2014, S. 7). Diese Kovarianzen werden für die Residualvariablen der endogenen Variablen Systemexploration und Systemsteuerung, der Exploration und der Steuerung des allgemeinen Problemlösens als auch für die konstruktübergreifende Exploration und die konstruktübergreifende Steuerung angenommen (gebogene Doppelpfeile auf die Residualvariablen).

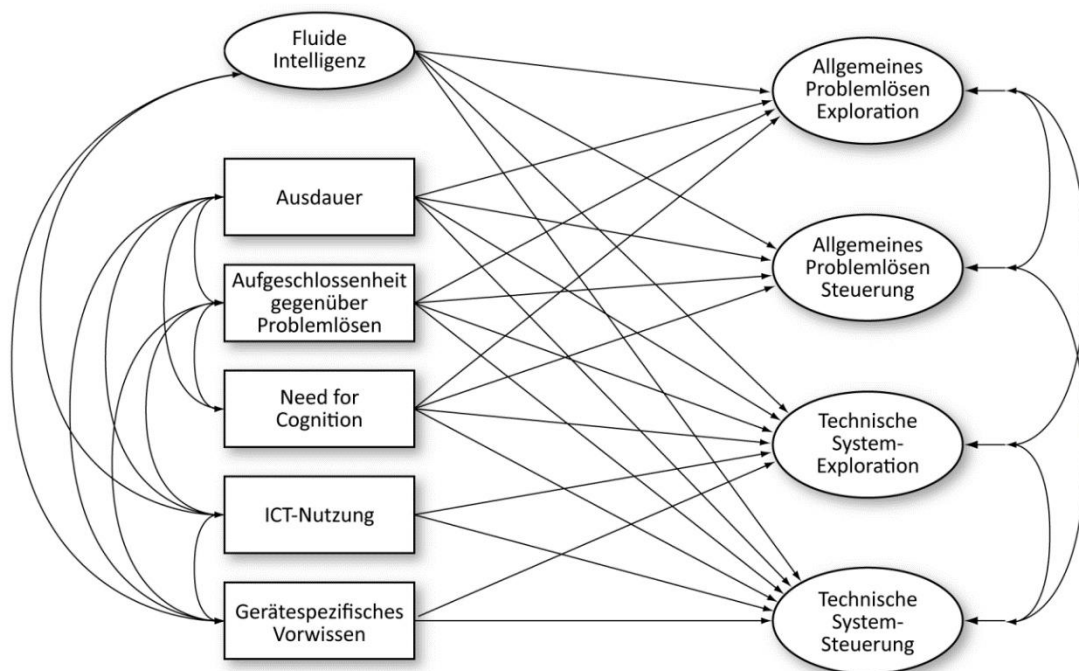


Abbildung 10.1: Zu prüfendes Strukturmodell zur Verortung technischer Problemlösekompetenz (die Messmodelle für die latenten Variablen sind hier nicht dargestellt)

Komplexe Strukturmodelle, wie das hier aufgestellte Theoriemodell (Abbildung 10.1) lassen sich mit einem Gleichungssystem aus mehreren Regressionsgleichungen berechnen, bei der die Beziehungen zwischen allen Variablen simultan berücksichtigt werden (Backhaus, Erichson, Plinke et al., 2008, S. 511). Die hieraus resultierende (vom Modell implizierte) Kovarianzmatrix kann jedoch keine Belege für die Richtigkeit des spezifizierten Modells liefern, so dass die Übereinstimmung mit den beobachteten Daten analysiert werden muss (Urban & Mayerl, 2014, S. 14). Bei einem solch strikt konfirmatorischen Vorgehen wird das gesamte Modell mit allen darin angenommenen Beziehungen entweder falsifiziert oder vorläufig beibehalten (Reinecke, 2014, S. 1). Hierfür werden die Modellparameter genauso wie in der konfirmatorischen Faktorenanalyse, die ja ein Spezialfall eines Strukturgleichungsmodells darstellt, derart geschätzt, dass die Unterschiede zwischen der modellimplizierten und der empirischen Kovarianzmatrix möglichst gering werden (Eid, Gollwitzer & Schmitt, 2013, S. 930). Das im Abschnitt 10.2 vorgestellte Schätzverfahren zur Prüfung der Dimensionalität technischer Problemlösekompetenz kann daher auch für die Analyse der Strukturgleichungsmodellierung genutzt werden und wird hier nicht mehr weiter besprochen.

Anhand der Modellgütemaße, die die Passung zwischen dem theoretischen Modell und den empirischen Daten widerspiegeln, kann beurteilt werden, ob die aufgestellten Zusammenhangshypothesen, auf denen das Gesamtmodell beruht, insgesamt beibehalten werden können (Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 967). Diese Modell-Fits entsprechen ebenfalls den bereits im vorherigen Kapitel vorgestellten Maßen. Muss das obige Modell aufgrund zu großer Abweichungen ($\chi^2/df > 2$; CFI $< .90$; RMSEA $> .08$ und WRMR > 1) abgelehnt werden, sind entsprechende Modifikationen vorzunehmen, die dann

an einer neuen Stichprobe überprüft werden müssen (Bortz & Schuster, 2010, S. 435). Solche Modifikationen sind jedoch nicht mehr Teil der vorliegenden Arbeit und werden allenfalls in die abschließende Diskussion eingebracht.

Kann aufgrund der Modellgütemaße von einer insgesamt ausreichenden Passung des Modells zu den Daten ausgegangen werden, lassen sich die einzelnen Modellparameter auch inhaltlich interpretieren (Werner, Schermelleh-Engel, Gerhard et al., 2016, S. 958). Hypothesen in Bezug auf einzelne Zusammenhänge können geprüft werden, indem der geschätzte Parameterwert durch seinen Standardfehler dividiert wird. Der sich ergebende t -Wert zeigt bei einem Wert größer als $|1.96|$ ein signifikantes Ergebnis an (Eid, Gollwitzer & Schmitt, 2013, S. 937). Von einer zu differenzierten Deutung von Pfadgewichten bei einer großen Anzahl von zu schätzender Modellparameter aus nur wenigen Daten raten Borg & Staufenbiel (2007, S. 265) ab. Sie empfehlen vielmehr die Ergebnisse in solchen Fällen in einem *groben Raster* zu betrachten, das heißt lediglich zu analysieren, welche der Pfadkoeffizienten groß sind und welche Vorzeichen sie aufweisen.

Für die Berechnung des vollständigen, in Abbildung 10.1 dargestellten Strukturgleichungsmodells, wird ausschließlich auf die Fälle zurückgegriffen, die alle Tests vollständig bearbeitet haben. Von den insgesamt 147 Probanden konnte lediglich 94 Personen der Test zum allgemeinen Problemlösen zur Bearbeitung vorgelegt werden. Einzelne als *zufällig fehlend* zu deklarierende Werte reduzieren die Anzahl weiter auf 77 Beobachtungen. Sowohl die zur Dimensionalitätsprüfung eingesetzte konfirmatorische Faktorenanalyse als auch die Analyse des Strukturgleichungsmodells werden mit der Software R (Version 3.3.1) (R core Team, 2016) und dem R-Paket lavaan (Version 0.5-20) (Rosseel, 2015) durchgeführt.

Teil IV: Forschungsergebnisse

11 Deskriptive Befunde

Das folgende Kapitel fasst die Ergebnisse der Erhebung deskriptiv zusammen. Die Daten werden dabei zunächst separat nach dem Messinstrument, mit dem sie erhoben wurden analysiert. Die Analyse beinhaltet sowohl Maße der zentralen Tendenz und der Streuung als auch Angaben zu den relativen Häufigkeiten. Fehlende Werte und der Umgang mit ihnen werden an den entsprechenden Stellen berichtet. Die Annahme, dass die Merkmale, die mit den jeweiligen Messinstrumenten erhoben wurden, in der untersuchten Stichprobe normalverteilt sind, wurde mit unterschiedlichen Tests überprüft, deren Ergebnisse ebenfalls Inhalt dieses Abschnittes sind. Die speziellen statistischen Voraussetzungen zur Überprüfung der Validität werden erst in den jeweiligen Auswertungen in Kapitel 12-15 analysiert.

11.1 Fluide Intelligenz

Alle 147 Probanden bearbeiteten die Kurzform des CFT 20-R zur Erfassung fluider Intelligenz. Die von den Probanden auf den Antwortbögen gegebenen richtigen Antworten wurden für jeden der vier Subtests *Reihenfortsetzen*, *Klassifikationen*, *Matrizen* und *Topologien* ausgezählt. Die vier Subtestwerte wurden anschließend zum Gesamtestrohwert aufsummiert. Dieser kann theoretisch einen Wert zwischen 0 und 56 annehmen. Der Mittelwert des CFT 20-R Rohwertes für die untersuchte Gelegenheitsstichprobe von Studierenden eines technischen Faches beträgt $MW = 41.39$ ($SD = 6.55$) (Tabelle 11.1). Die Rohwerte der Probanden wurden mithilfe vorhandener Normtabellen in altersabhängige IQ-Werte transformiert, deren mittlerer Wert auf 100 IQ-Punkte genormt ist. Für die untersuchte Stichprobe beträgt der Mittelwert $MW = 108.61$ IQ-Punkte, wobei die Werte im Vergleich zur Normstichprobe ($SD_{norm} = 15.00$) etwas stärker streuen ($SD = 16.95$). Das bezogen auf die Normstichprobe, höhere fluide Intelligenzniveau ist aufgrund der stark selektierten, leistungsstärkeren Stichprobe plausibel.

Tabelle 11.1: Deskriptive Statistik und Subtest-Rohwert-Korrelationen der Ergebnisse des eingesetzten CFT 20-R

	<i>n</i>	<i>MW</i>	<i>SD</i>	α	Klassifikation	Matrizen	Topologien	Rohwert
Reihenfortsetzen	147	12.33	1.93	.53	.49**	.53**	.50**	.79**
Klassifikation	147	10.60	2.39	.64		.53**	.41**	.80**
Matrizen	147	11.31	2.09	.59			.41**	.79**
Topologien	147	7.15	2.01	.62				.73**
Rohwert	147	41.39	6.55	.83				

Anmerkungen: *n* = Stichprobengröße; *MW* = Mittelwert; *SD* = Standardabweichung; α = Cronbachs Alpha; ** $p < .01$

Um zu prüfen, ob sich die erhobenen Intelligenztestdaten (IQ-Werte) normalverteilen, wurden zunächst die Koeffizienten für die Schiefe ($g_1 = -.23$) und den Exzess ($g_2 = -.32$) betrachtet. Diese deuten auf eine leichte linksschiefe und breitgipflige Verteilung hin. Ob es sich um signifikante Abweichungen von der Normalverteilung handelt, wird anhand des Critical Ratios *c.r.* (die Division der Koeffizienten durch ihren Standardfehler) geprüft. Diese als *z*-Werte zu interpretierenden Critical Ratios, sollten bei Vorliegen einer Normalverteilung $\pm 1,96$ (5 % Niveau) nicht überschreiten (vgl. Reisinger, Svecnik & Schwetz, 2012, S. 12). Für die Schiefe ergibt sich ein *c.r.*(g_1) von -1.16 und für den Exzess ein *c.r.*(g_2) von -.82, beide weichen demnach nicht signifikant von den Koeffizienten der Normalverteilung ab. Die

Prüfgröße $D = .07$, die sich aus der Anwendung des Kolmogorov-Smirnov-Tests (mit Signifikanzkorrektur nach Lilliefors) ergibt, der den maximalen Abstand zwischen der Normalverteilung und der empirischen Verteilung misst, wird tendenziell signifikant ($p = .09$; $df = 147$), was gegen die Nullhypothese des Vorliegens eines normalverteilten Merkmals spricht. Ein weiterer Test zur Prüfung der Normalverteilungsannahme, der Shapiro-Wilk-Test, der vor allem bei kleinen Stichproben ($n < 50$) eingesetzt wird (ebd. S. 18) zeigt mit $W = .989$; $p = .30$ keine signifikante Abweichung von einer Normalverteilung. Das Ergebnis des Shapiro-Wilk-Tests wird in den folgenden Überprüfungen der Normalverteilungsannahme nur dann genannt, wenn es von dem Ergebnis des Kolmogorov-Smirnov-Tests abweicht.

Tabelle 11.2: Deskriptive Statistik der, in genormte IQ-Werte transformierten Ergebnisse des eingesetzten CFT 20-R

	<i>n</i>	<i>MW</i> (<i>MIN</i> , <i>MAX</i>)	<i>SD</i>	95 % <i>CI</i>	<i>g</i> ₁	<i>SE</i>	<i>c.r.</i> (<i>g</i> ₁)	<i>g</i> ₂	<i>SE</i>	<i>c.r.</i> (<i>g</i> ₂)	K-S-Test	
											<i>D</i>	<i>p</i>
fluide Intelligenz IQ	147	108.61 (62, 145)	17	[105.84, 111.34]	-.23	.20	-1.16	-.32	.40	-.82	.07	.09

Anmerkungen: *n* = Stichprobengröße; *MW* = Mittelwert; *MIN* = Minimum; *MAX* = Maximum; *SD* = Standardabweichung; 95 % *CI* = Konfidenzintervall für den Mittelwert; *c.r.* = Critical Ratio; *g*₁ = Schiefe; *g*₂ = Exzess; *SE* = Standardfehler; K-S-Test = Kolmogorov-Smirnov-Test auf Prüfung der Normalverteilung; *D* = Prüfgröße des Kolmogorov-Smirnov-Tests mit $df = 147$; *p* = *p*-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen

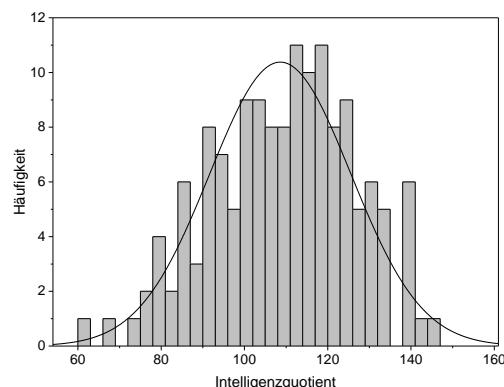


Abbildung 11.1: Häufigkeiten der erreichten IQ-Werte der $n = 147$ Probanden mit eingezeichneter Normalverteilungskurve

11.2 Need for Cognition

Die Skala Need for Cognition wurde von allen 147 Probanden am Computer bearbeitet. Da es nicht möglich war, mit fehlenden Angaben auf die präsentierten 19 Aussagen zum nächsten Testteil zu gelangen, wiesen die Daten keine fehlenden Werte auf. Die Antworten der Probanden auf die Einzelitems der Skala erfolgte durch die Wahl einer Antwort aus je fünf Antwortalternativen („trifft nicht zu“, „trifft eher nicht zu“, „weder noch“, „trifft eher zu“, „trifft zu“), die mit Punkten zwischen 0 und 4 bewertet wurden, wobei ein höherer Punktwert für eine größere Ausprägung an Need for Cognition steht. Die relativen Häufigkeiten der gewählten Antwortkategorien sind in Tabelle 11.3 zusammengefasst. Die Kategorie 0 wurde zur Beantwortung des Items 3 und des Items 19 nicht und zur Beantwortung der anderen

Items nur vereinzelt gewählt. Hingegen häufen sich die Probandenantworten in der höheren Kategorie 3.

Tabelle 11.3: relative Häufigkeiten der gewählten Kategorien aller Items der Skala Need for Cognition (n = 147)

Item	Kategorien				
	0	1	2	3	4
NFC_1	1	5	18	50	25
NFC_2	4	16	28	37	16
NFC_3	0	7	12	52	28
NFC_4*	4	14	13	50	19
NFC_5	1	5	10	42	43
NFC_6*	4	14	22	45	16
NFC_7*	4	24	24	38	10
NFC_8*	2	8	10	46	35
NFC_9*	1	3	10	35	51
NFC_10*	1	10	14	42	33
NFC_11	4	29	29	34	3
NFC_12	7	29	29	31	5
NFC_13*	4	15	12	48	21
NFC_14*	2	8	11	50	29
NFC_15	7	17	22	45	9
NFC_16	1	15	26	54	4
NFC_17	1	7	15	61	16
NFC_18	3	18	16	48	16
NFC_19	0	5	12	59	24

Anmerkungen: Einleitender Text: Wie gut treffen die folgenden Aussagen auf Sie zu?;
Codierung der Skala: 0 = trifft nicht zu; 1 = trifft eher nicht zu; 2 = weder noch; 3 = trifft eher zu; 4 = trifft zu; * = umgekehrt codierte Items; Zeilensummen, die von dem Wert 100 abweichen, sind auf Rundungsfehler zurückzuführen

Aus der Summe der einzelnen Itemwerte wurde für jeden Probanden ein Skalenwert gebildet, der als Maß für die individuelle Ausprägung an Need for Cognition interpretiert wird. Die Skalenwerte befinden sich dabei in einem theoretisch möglichen Wertebereich zwischen 0 und 76. Der kleinste tatsächlich beobachtete Skalenwert von $MIN = 26$ sowie ein Mittelwert von $MW = 51.24$ ($SD = 9.34$) deuten darauf hin, dass sich die Skalenwerte in der Stichprobe nicht über den gesamten Wertebereich $[0, 76]$ verteilen. Das Histogramm der absoluten Häufigkeiten (Abbildung 11.2) zeigt, dass sich die Skalenwerte am Ende des Wertebereiches häufen, das Merkmal Need for Cognition in der Stichprobe also eher stark ausgeprägt ist. Diese Beobachtung lässt sich ad hoc mit der Selektion der Stichprobe begründen. So kann zum einen davon ausgegangen werden, dass Studierende generell eine höhere Need for Cognition haben und zum anderen, dass bei Personen, die freiwillig an einem Test zur Erfassung technischer Problemlösekompetenz teilnehmen, das Bedürfnis kognitive Anforderungen zu bewältigen hoch ausgeprägt ist.

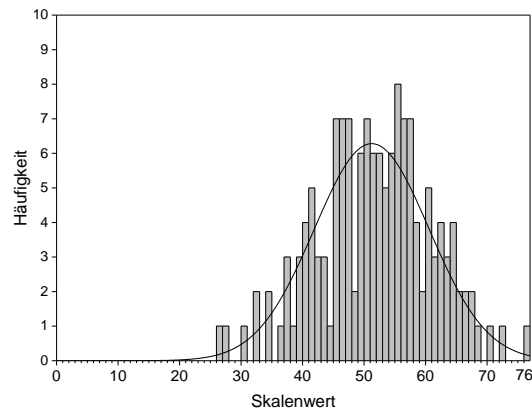


Abbildung 11.2: Häufigkeiten der Testwerte der Skala Need for Cognition der $n = 147$ Probanden mit eingezeichneter Normalverteilungskurve

Die Verteilung des Merkmals Need for Cognition in der Stichprobe wird für den *empirischen* Skalenswertebereich, der sich von $MIN = 26$ bis $MAX = 76$ erstreckt, überprüft. Der Koeffizient für die Schiefe $g_1 = -.17$ deutet auf eine leichte Linksschiefe der Merkmalsverteilung und der Koeffizient für den Exzess $g_2 = -.05$ auf eine minimale Abweichung von der Wölbung einer Normalverteilung hin (Tabelle 11.4). Die Critical Ratios beider Koeffizienten ($c.r.(g_1) = -.83$; $c.r.(g_2) = -.11$) überschreiten nicht den kritischen Wert von $|1.97|$, sodass davon ausgegangen werden kann, dass sich weder die Schiefe noch der Exzess der empirischen Verteilung signifikant von denen einer Normalverteilung unterscheiden. Das Ergebnis der zusätzlichen Prüfung der empirischen Verteilung auf Normalverteilung mit dem Kolmogorov-Smirnov-Test bringt ebenfalls keine Hinweise auf eine Verletzung der Normalverteilungsannahme hervor ($D = .05$; $p = .20$; $df = 147$).

Tabelle 11.4: Deskriptive Statistik der Skala Need for Cognition

	n	MW (MIN, MAX)	SD	95 % CI	g_1	SE	c.r.(g_1)	g_2	SE	c.r.(g_2)	K-S-Test		
											D	p	α
NFC-Gesamttest	147	51.24 (26, 76)	9.34	[49.72, 52.77]	-.17	.20	-.83	-.05	.40	-.11	.05	.20	.84

Anmerkungen: n = Stichprobengröße; MW = Mittelwert; MIN = Minimum; MAX = Maximum; SD = Standardabweichung; 95 % CI = Konfidenzintervall für den Mittelwert; c.r. = Critical Ratio; g_1 = Schiefe; g_2 = Exzess; SE = Standardfehler; K-S-Test = Kolmogorov-Smirnov-Test auf Prüfung der Normalverteilung; D = Prüfgröße des Kolmogorov-Smirnov-Tests mit $df = 147$; p = p-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; α = Cronbachs Alpha

Die mit der Methode der internen Konsistenz gemessene Reliabilität der Skala beträgt Cronbachs $\alpha = .84$ und liegt damit unter denen von Preckel (2014, S. 68) in verschiedenen Stichproben gemessenen Reliabilitäten von $\alpha = .89$ bis $.92$. Obgleich die verwendete Skala das Konstrukt Need for Cognition in der vorliegenden Stichprobe mit einer verminderten Genauigkeit misst, als in der von Preckel untersuchten Stichprobe von 745 Schülerinnen und Schülern der Klassen 5 bis 7 deutscher Schulen, ist die Messgenauigkeit mit $\alpha = .84$ ausreichend.

11.3 Selbstwirksamkeitserwartung

Die Erfassung der Selbstwirksamkeitserwartungen (SWE) gegenüber dem Problemlösen erfolgte mit den beiden Skalen *Ausdauer* und *Aufgeschlossenheit gegenüber dem Problemlösen*. Beide Skalen wurden von allen 147 Probanden computerbasiert beantwortet. Durch die

technische Realisierung eines automatischen Hinweisfensters bei fehlenden Antworten konnten fehlende Werte in den Daten dieser Skala von vornherein verhindert werden. Die Zustimmung auf die je Skala vorgelegten fünf Aussagen erfolgte durch die Wahl einer von vier Antwortkategorien („trifft nicht zu“, „trifft eher nicht zu“, „trifft eher zu“, „trifft zu“). Tabelle 11.5 fasst die Häufigkeiten der gewählten Antwortkategorien zusammen. In der Skala Aufgeschlossenheit gegenüber dem Problemlösen (SWE_2) wird die Kategorie 0 („trifft nicht zu“) für die Zustimmung auf zwei der fünf Items von keinem Probanden gewählt. Das betrifft die beiden Aussagen: *Ich versuche Erklärungen für Dinge zu suchen* sowie *Ich kann einfach Fakten miteinander verbinden*. Insgesamt wird die Kategorie 2 („trifft eher zu“) für den Grad der Zustimmung auf alle Aussagen am häufigsten gewählt.

Tabelle 11.5: relative Häufigkeiten der gewählten Kategorien aller Items der Skalen Ausdauer und Aufgeschlossenheit gegenüber dem Problemlösen (n = 147)

Item	Kategorien			
	0	1	2	3
SWE_1_1*	1	7	56	35
SWE_1_2*	3	23	53	21
SWE_1_3	3	14	56	26
SWE_1_4	4	25	52	19
SWE_1_5	5	32	52	11
SWE_2_1	1	16	64	19
SWE_2_2	1	15	71	14
SWE_2_3	0	8	45	48
SWE_2_4	0	9	70	21
SWE_2_5	1	14	56	29

Anmerkungen: Einleitender Text: Wie gut treffen die folgenden Aussagen auf Sie zu?; SWE_1: Items der Skala Ausdauer; SWE_2: Items der Skala Aufgeschlossenheit gegenüber dem Problemlösen; Codierung der Skalen: 0 = trifft nicht zu; 1 = trifft eher nicht zu; 2 = trifft eher zu; 3 = trifft zu; * = umgekehrt codierte Items; Zeilensummen, die von dem Wert 100 abweichen, sind auf Rundungsfehler zurückzuführen

Der für jeden Probanden individuelle Skalenwert für jede der beiden Skalen wird durch eine Summierung der einzelnen Itemwerte gebildet, sein Wertebereich reicht von $MIN = 0$ bis $MAX = 15$. Der Median für die vierstufige Skala *Ausdauer* beträgt $MD = 10$ und für die Skala *Aufgeschlossenheit gegenüber Problemlösen* $MD = 11$. Der empirische Wertebereich beider Skalen reicht bis zu dem maximalen Skalenwert von 15, für die Skala *Ausdauer* beginnt sie bei dem Wert 4 und für die Skala *Aufgeschlossenheit* bei dem Wert 6. Beide Skalen weisen einen positiven Wert für die Schiefe der Verteilung auf, der bei der Skala *Ausdauer* mit $g_1 = .13$ eine größere Abweichung von einer Normalverteilung aufweist, als bei der Skala *Aufgeschlossenheit* mit $g_1 = .08$. Die empirische Verteilung des Merkmals in der Stichprobe ist im Gegensatz zu einer symmetrischen Normalverteilung ein wenig rechtsschief. Für beide Verteilungen wird das Critical Ratio jedoch nicht signifikant ($c.r.(g_1) < |1.96|$). Das Maß für den Exzess der Verteilung g_2 weist für beide empirischen Merkmalsverteilungen einen negativen Wert auf, der zeigt, dass die beiden Verteilungen breitgipfliger sind als eine Normalverteilung. Das Critical Ratio als standardisiertes Maß bleibt unterhalb des kritischen Wertes für die Signifikanz. Die maximale Abweichung zwischen der jeweiligen empirischen

Verteilung und der Normalverteilung ist dennoch so groß, dass dieses Distanzmaß D aus dem Kolmogorov-Smirnov-Test für beide Verteilungen signifikant wird ($p < .05$, $df = 147$).

Tabelle 11.6: Deskriptive Statistik der Skalen *Ausdauer* und *Aufgeschlossenheit* gegenüber Problemlösen

	n	Mdn (MIN, MAX)	SD	g_1	SE	$c.r.(g_1)$	g_2	SE	$c.r.(g_2)$	K-S-Test		
										D	p	α
SWE_1	147	10 (4, 15)	2.29	.13	.20	.67	-.17	.40	-.44	.15	.00	.62
SWE_2	147	11 (6, 15)	1.94	.08	.20	.41	-.27	.40	-.67	.11	.00	.63

Anmerkungen: n = Stichprobengröße; Mdn = Median; MIN = Minimum; MAX = Maximum; SD = Standardabweichung; $c.r.$ = Critical Ratio; g_1 = Schiefe; g_2 = Exzess; SE = Standardfehler; K-S-Test = Kolmogorov-Smirnov-Test auf Prüfung der Normalverteilung; D = Prüfgröße des Kolmogorov-Smirnov-Tests mit $df = 147$; p = p -Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; α = Cronbachs Alpha

Die Reliabilität der Skala *Ausdauer* beträgt in der hier untersuchten Stichprobe $\alpha = .62$ und in der Skala *Aufgeschlossenheit* gegenüber Problemlösen $\alpha = .63$. Beide Skalen wurden in der PISA-Studie 2012 (OECD, 2014b) mit je fünf Antwortkategorien eingesetzt. Die für Deutschland berichteten Maße der internen Konsistenz werden mit $\alpha = .67$ für die Skala *Ausdauer* und $\alpha = .81$ für die Skala *Aufgeschlossenheit* angegeben. Ursachen für die Abweichungen können zum einen in den unterschiedlich untersuchten Stichproben und zum anderen in der vorgenommenen Änderung an der Antwortkategorienanzahl liegen.

11.4 ICT-Literacy

Um den Einfluss geräteübergreifenden Vorwissens auf den problemlösenden Umgang mit technischen Alltagsgeräten zu untersuchen, wurde das geräteübergreifende Vorwissen mit einem Fragebogen zur Nutzung von Informations- und Kommunikationstechnologien (ICT) erfasst. Der aus der PISA-Studie 2012 stammende Fragebogen (OECD, 2014b) umfasst 10 Items, die durch die Wahl einer Antwort aus fünf Antwortalternativen beantwortet werden („niemals“, „weniger als 1x im Monat“, „weniger als 1x wöchentlich, aber mehr als 1x im Monat“, „nicht täglich, aber mehr als 1x wöchentlich“, „täglich“). In der vorliegenden Studie wurde die Skala zur Nutzung von ICT von den Probanden am Computer bearbeitet, wodurch der bereits oben genannte Vorteil der Vermeidung von fehlenden Werten einhergeht. Die Betrachtung der in Tabelle 11.7 zusammengefassten Nennungshäufigkeiten der Antwortkategorien für die einzelnen Items der Skala liefert ein heterogenes Bild über die Nutzung von Informations- und Kommunikationstechnologien. Während manche ICT-Bereiche (z. B. ICT_2: *Spielen von Onlinespielen, bei denen mehrere Spieler erforderlich sind*) von der überwiegenden Anzahl der Probanden *niemals* genutzt werden (Kategorie 0), gibt es andere ICT-Bereiche (z. B. ICT_3: *Nutzen von E-Mails*) die von der überwiegenden Anzahl der Probanden *täglich* genutzt wird (Kategorie 4). Für den letztgenannten ICT-Bereich gibt es keine Nennungen in den Antwortkategorien 0 und 1.

Tabelle 11.7: relative Häufigkeiten der gewählten Kategorien aller Items der Skala ICT-Nutzung ($n = 147$)

Item	Kategorien				
	0	1	2	3	4
ICT_1	35	32	9	16	9
ICT_2	53	19	5	15	8
ICT_3	0	0	5	20	75
ICT_4	24	12	9	29	26
ICT_5	12	7	10	21	50
ICT_6	1	3	14	40	42
ICT_7	4	2	10	31	52
ICT_8	1	3	11	42	42
ICT_9	7	31	28	22	11
ICT_10	55	26	13	4	2

Anmerkungen: Einleitender Text: Wie oft nutzen Sie den Computer für die folgenden Aktivitäten?; Codierung der Skala: 0 = niemals; 1 = weniger als 1x im Monat; 2 = weniger als 1x wöchentlich, aber mehr als 1x im Monat; 3 = nicht täglich aber mehr als 1x wöchentlich; 4 = täglich; Zeilensummen, die von dem Wert 100 abweichen, sind auf Rundungsfehler zurückzuführen

Die Summe über alle Itemwerte einer Person bildet den individuellen Skalenwert, der innerhalb eines theoretischen Wertebereichs von $MIN = 0$ bis $MAX = 40$ liegen kann. Die in der untersuchten Stichprobe beobachteten Daten verteilen sich um einen Mittelwert von $MW = 23.54$ mit einer Standardabweichung von $SD = 5.91$ (Tabelle 11.8). Der Koeffizient für die Schiefe weist mit $g_1 = -.34$ auf eine deutliche Linksschiefe der empirischen Verteilung hin, wohingegen der Koeffizient des Exzess mit $g_2 = .05$ keine Abweichung der Wölbung von einer Normalverteilung zeigt. Während das Critical Ratio für die Abweichung der Schiefe nicht signifikant wird ($c.r.(g_1) = -1.70$), spricht das Ergebnis des Kolmogorov-Smirnov-Tests für eine signifikante Abweichung der Verteilung des ICT-Merkmals in der Stichprobe von einer Normalverteilung ($D = .08$; $p = .02$, $df = 147$). Die Reliabilität der Skala ICT-Nutzung weist mit $\alpha = .70$ keine große Differenz zu der von der OECD (2014b, S. 338) berichteten Reliabilität von $\alpha = .75$ auf.

Tabelle 11.8: Deskriptive Statistik der Skala Nutzung von Informations- und Kommunikationstechnologien

	n	MW (MIN, MAX)	SD	95 % CI	g_1	SE	c.r.(g_1)	g_2	SE	c.r.(g_2)	K-S-Test		
											D	p	α
ICT Nutzung	147	23.54 (9, 36)	5.91	[22.58, 24.51]	-.34	.20	-1.70	.05	.40	.12	.08	.02	.70

Anmerkungen: n = Stichprobengröße; MW = Mittelwert; MIN = Minimum; MAX = Maximum; SD = Standardabweichung; 95 % CI = Konfidenzintervall für den Mittelwert; c.r. = Critical Ratio; g_1 = Schiefe; g_2 = Exzess; SE = Standardfehler; K-S-Test = Kolmogorov-Smirnov-Test auf Prüfung der Normalverteilung; D = Prüfgröße des Kolmogorov-Smirnov-Tests mit $df = 147$; p = p-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; α = Cronbachs Alpha

11.5 Gerätespezifisches Vorwissen

Neben dem geräteübergreifenden Vorwissen wird der Umgang mit technischen Alltagsgeräten außerdem durch das gerätespezifische Vorwissen beeinflusst. Um dieses spezifische Vorwissen zu erfassen, wurde die bereits vorhandene Erfahrung in Bezug auf die in dem technischen Problemlösetest zu bedienenden technischen Alltagsgeräte mit je einem Item erfasst. Die Ausprägung ihrer Erfahrungen gaben die Probanden auf einer (je nach

technischem System) drei oder vierstufigen Ratingskala mit Falldarstellungen an (die genauen Itemwortlaute sind Tabelle 9.4 zu entnehmen). Die ursprünglich bei den Items 3, 5, 6, 8, 10 und 15 vier Antwortkategorien wurden aus Gründen der Redundanz von Kategorien und der geringen Nennungshäufigkeit zu je drei Kategorien zusammengefasst. Die relativen Nennungshäufigkeiten der resultierenden drei Kategorien (0 = keine bis wenig Erfahrung, 1 = durchschnittlich viel Erfahrung, 2 = viel Erfahrung) sind in Tabelle 11.9 aufgeführt. Bei 9 von 15 technischen Alltagsgeräten gab der größte Teil der Probanden an, über durchschnittlich viel Erfahrung zu verfügen. Keine oder nur wenig Erfahrung hatte der größte Teil der Probanden bei 6 der technischen Alltagsgeräte. So gaben beispielsweise 86,39 % der Probanden an, noch nie eine Heizungssteuerung bedient zu haben und ihren Heizkörper nach Bedarf manuell einzustellen (GspV_13).

Tabelle 11.9: relative Häufigkeiten der gewählten Kategorien aller Items der Skala gerätespezifisches Vorwissen (n = 147)

Item	Kategorien		
	0	1	2
GspV_1	18	65	16
GspV_2	29	60	12
GspV_3*	10	87	3
GspV_4	56	43	1
GspV_5*	8	91	1
GspV_6*	38	61	1
GspV_7	52	44	4
GspV_8**	34	61	5
GspV_9	79	21	0
GspV_10*	20	78	2
GspV_11	35	57	8
GspV_12	56	42	1
GspV_13	86	13	1
GspV_14	26	47	27
GspV_15**	57	39	3

Anmerkungen: Codierung der Skala: 0 = keine bis wenig Erfahrung; 1 = durchschnittlich viel Erfahrung; 2 = viel Erfahrung; * ursprüngliche Kategorien 1 und 2 wurden zu Kategorie 1 zusammengefasst; ** ursprüngliche Kategorien 0 und 1 wurden zu Kategorie 0 zusammengefasst; Zeilensummen, die von dem Wert 100 abweichen, sind auf Rundungsfehler zurückzuführen

Zum einen wird vermutet, dass die Bedienungserfahrungen, die eine Person mit einem technischen Gerät wie beispielsweise einer Waschmaschine bereits gesammelt hat, ihr in der Handhabung einer anderen, als der ihr bekannten Waschmaschine helfen kann. Je ähnlicher das neue zu bedienende dem bekannten System ist, desto nutzbringender kann die Vorerfahrung sein. Inwieweit das Vorwissen der Probanden über ein spezifisches technisches Gerät tatsächlich den Umgang mit einem anderen aber ähnlichem Gerät in dem computerbasierten Test beeinflusst hat, wird in Kapitel 15 besprochen.

Weil aber zum anderen außerdem davon auszugehen ist, dass bereits erworbene Erfahrungen im Umgang mit einem speziellen technischen System (z. B. Waschmaschine) relevant bei der

Bedienung anderer auch funktionsfremder technischer Geräte (z. B. Spülmaschine) sein kann (weil sich daraus ein systemübergreifendes Wissen extrahiert hat), wird aus den Antworten auf die Einzelitems ein Gesamtscore gebildet, der das Ausmaß an gerätespezifischem Vorwissen angibt. Der Wertebereich dieser Gesamtskala erstreckt sich theoretisch von 0 bis 30, wobei die Werte in der Stichprobe zwischen 1 und 21 variieren. Der mit $Mdn = 10$ unter dem theoretischen Wert von 15 liegende Median für das gerätespezifische Vorwissen zeigt die bereits in den Häufigkeitsangaben zu sehende geringe bis mittlere Vorwissensausprägung der Probanden. Die Verteilung der empirischen Werte ist im Vergleich zu einer Normalverteilung mit $g_1 = .51$ linksschief und mit $g_2 = .90$ etwas breitipflig. Das Critical Ratio beider Koeffizienten zeigt eine signifikante Abweichung von einer Normalverteilung an, was durch das signifikante Ergebnis ($D = .12$; $p < .001$ mit $df = 147$) des Kolmogorov-Smirnov-Test bestätigt wird.

Tabelle 11.10: Deskriptive Statistik der Skala gerätespezifisches Vorwissen

	n	Mdn	MIN	MAX	g ₁	SE	c.r.(g ₁)	g ₂	SE	c.r.(g ₂)	K-S-Test		
											D	p	α
gerätespezifisches Vorwissen	147	10	1	21	.51	.20	2.56	.90	.40	2.27	.12	.00	.67

Anmerkungen: n = Stichprobengröße; Mdn = Median; MIN = Minimum; MAX = Maximum; c.r. = Critical Ratio; g₁ = Schiefe; g₂ = Exzess; SE = Standardfehler; K-S-Test = Kolmogorov-Smirnov-Test auf Prüfung der Normalverteilung; D = Prüfgröße des Kolmogorov-Smirnov-Tests mit $df = 147$; p = p-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; α = Cronbachs Alpha

11.6 Allgemeine Problemlösefähigkeit

Der computerbasierte Test zur Erfassung allgemeiner Problemlösefähigkeit wurde 94 der 147 Probanden zur Bearbeitung vorgelegt. Die Bearbeitung des Tests erfolgte dabei in den drei Phasen Exploration, Wissensabfrage und Steuerung. Der problemlösende Umgang mit den simulierten Szenarien in der Explorationsphase wurde nicht direkt bewertet, sondern lediglich die Antworten auf die sich anschließenden Wissensfragen. Die Wissensabfrage erfolgte durch je vier Verifikations- und je zwei Identifikationsaufgaben¹. Jede der vier Verifikationsaufgaben wurde dichotom (0/1) ausgewertet. Jeweils zwei der vier Items werden zu einem Subscore² zusammengefasst. Werden beide der zusammengefassten Verifikationsaufgaben richtig gelöst, beträgt der Subscore 1, ansonsten 0. Aus den so ermittelten Subscores wird das arithmetische Mittel, der Itemwert für die Verifikationsaufgabe gebildet. Dieser kann demnach die diskreten Werte 0; 0.5; und 1 annehmen. In den beiden Identifikationsaufgaben müssen jeweils bis zu sieben Informationen (variiert je nach Szenario) richtig zugeordnet werden. Nur wenn alle Informationen richtig positioniert wurden, wird der Subscore mit einem Punkt bewertet, ansonsten beträgt der Subscore 0. Aus diesen beiden Subscores für die Identifikationsaufgabe wird das arithmetische Mittel gebildet, so dass der Itemwert für diesen Aufgabentyp ebenfalls die diskreten Werte 0; 0.5 und 1 annehmen kann. Die sich an die Wissensabfrage anschließende Steuerungsaufgabe wird mit 0 Punkten bewertet, wenn das vorgegebene Steuerungsziel nicht erreicht wurde, 1 Punkt, wenn das Ziel mit mehr als den nötigen Bearbeitungsschritten und 2 Punkte, wenn das Ziel auf dem direktesten Bearbeitungsweg erreicht wurde. Dieser Punktwert wird anschließend durch die maximale

¹ Eine Beschreibung der unterschiedlichen Aufgabentypen finden sich in Kapitel 9.1.5.

² Die Verifikationsaufgaben 1 und 2 werden zu einem Subscore zusammengefasst sowie die Verifikationsaufgaben 3 und 4.

Punktzahl von 2 dividiert, so dass das Ergebnis der Steuerungsphase ebenfalls die diskreten Werte von 0; 0.5 und 1 annehmen kann. Tabelle 11.11 zeigt die relativen Häufigkeiten der jeweiligen Lösungskategorien für die Verifikations-, die Identifikations- und die Steuerungsaufgaben.

Tabelle 11.11: relative Lösungshäufigkeiten der unterschiedlichen Aufgaben für die Wissensabfrage nach der Exploration und für die Steuerungsaufgabe im allgemeinen Problemlösetest

Item	Wissen nach der Exploration								
	Verifikationsaufgabe (n = 92)			Identifikationsaufgabe (n = 92)			Steuerungsaufgabe (n = 86)		
	.00	.50	1.00	.00	.50	1.00	.00	.50	1.00
Item1	8	35	58	35	27	38	22	46	33
Item2	9	42	49	24	22	54	11	46	44
Item3	8	15	77	25	13	62	12	43	45
Item4	13	40	47	14	33	53	25	11	64
Item5	14	42	44	40	41	18	21	42	37
Item6	27	53	20	28	55	16	57	30	13

Anmerkungen: n = Stichprobengröße; Zeilensummen, die von dem Wert 100 abweichen, sind auf Rundungsfehler zurückzuführen

Die unterschiedlichen Stichprobengrößen ergeben sich durch fehlende Werte, die zum einen aufgrund softwareseitig nicht gespeicherter Logfiles sowie durch vereinzelt nicht bearbeitete Aufgaben entstanden. Die relativen Lösungshäufigkeiten für die Verifikations- und Identifikationsaufgaben deuten darauf hin, dass die eingesetzten Testitems dieser Aufgabentypen für die untersuchte Stichprobe eine geringe Schwierigkeit hatten. Für die Steuerungsaufgaben zeigt sich ein anderes Bild, hier variiert die Schwierigkeit der Items stärker.

Die Bildung individueller Testwerte wird für die drei Aufgabentypen separat vorgenommen, indem die für jedes Item erreichten Punkte summiert werden. Somit ergeben sich für jede Person zwei Testwerte für das nach der Explorationsphase erworbene Wissen und ein Testwert für die Leistung in der Steuerungsphase. Der theoretische Wertebereich aller drei Testwerte erstreckt sich jeweils von 0 bis 6 Punkten. Für das nach der Explorationsphase vorhandene Systemwissen, das mit Verifikationsaufgaben erfasst wurde, beträgt der Mittelwert für die $n = 92$ Probanden $MW = 4.08$ ($SD = 1.04$). Die bereits anhand der relativen Lösungshäufigkeiten sichtbare Leichtigkeit der Items des Aufgabentyps Verifikation zeigt sich so auch in der nach rechts verschobenen Verteilung der beobachteten Daten (Tabelle 11.12). Der Koeffizient für die Schiefe der empirischen Verteilung weicht mit $g_1 = -0.48$ von einer Normalverteilung ab. Der kritische Wert für das Critical Ratio wird nur knapp verfehlt ($c.r.(g_1) = -1.92$). Das mit dem Kolmogorov-Smirnov-Test ermittelte Distanzmaß $D = .15$ weist jedoch auf eine signifikante Abweichung von einer Normalverteilung hin ($p < .05$; $df = 92$). Für das nach der Exploration vorhandene Systemwissen, das mit Identifikationsaufgaben erfasst wurde, beträgt der Mittelwert $MW = 3.38$ ($SD = 1.63$), die beobachteten Daten decken den theoretischen Wertebereich vollständig ab ($MIN = 0$; $MAX = 6$). Der hohe Wert für den Exzesskoeffizienten $g_2 = -0.99$, dessen Critical Ratio $c.r.(g_2) = -1.98$ signifikant wird sowie das Histogramm der absoluten Lösungshäufigkeiten in Abbildung 11.3 zeigen, dass sich die beobachteten Daten im Vergleich zu einer Normalverteilung sehr breitipflig verteilen. Diese Abweichung von der Normalverteilung zeigt sich auch in dem signifikant werdenden Ergebnis des Kolmogorov-Smirnov-Test ($D = .13$; $p < .05$; $df = 92$).

Tabelle 11.12: Deskriptive Statistik für die unterschiedlichen Aufgaben der Wissensabfrage nach der Exploration sowie für die Steuerungsaufgabe im allgemeinen Problemlösetest

	n	MW (MIN, MAX)	SD	95 % CI	g ₁	SE	c.r.(g ₁)	g ₂	SE	c.r.(g ₂)	K-S-Test		
											D	p	α
Verifikation	92	4.08 (1, 6)	1.04	[3.86, 4.29]	-.48	.25	-1.92	.03	.50	.14	.15	.00	.47
Identifikation	92	3.38 (0, 6)	1.63	[3.04, 3.72]	-.32	.25	-1.29	-.99	.50	-1.98	.13	.00	.79
Steuerung	86	3.54 (.50, 6)	1.31	[3.26, 3.82]	-.35	.26	-1.33	-.53	.51	-1.03	.14	.00	.65

Anmerkungen: n = Stichprobengröße; MW = Mittelwert; MIN = Minimum; MAX = Maximum; SD = Standardabweichung; 95 % CI = Konfidenzintervall für den Mittelwert; c.r. = Critical Ratio; g₁ = Schiefe; g₂ = Exzess; SE = Standardfehler; K-S-Test = Kolmogorov-Smirnov-Test auf Prüfung der Normalverteilung; D = Prüfgröße des Kolmogorov-Smirnov-Tests mit df = n; p = p-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; α = Cronbachs Alpha

Die von n = 86 Probanden erreichten Testwerte in der Steuerungsphase des allgemeinen Problemlösetests verteilen sich um einen Mittelwert von MW = 3.54 (SD = 1.31). Auch diese Verteilung weicht in ihrer Schiefe (g₁ = -.35) und ihrem Exzess (g₂ = -.53) von einer Normalverteilung ab. Zwar werden die entsprechenden Critical Ratios nicht signifikant, jedoch das Ergebnis des Kolmogorov-Smirnov-Tests (D = .14; p < .05, df = 86).

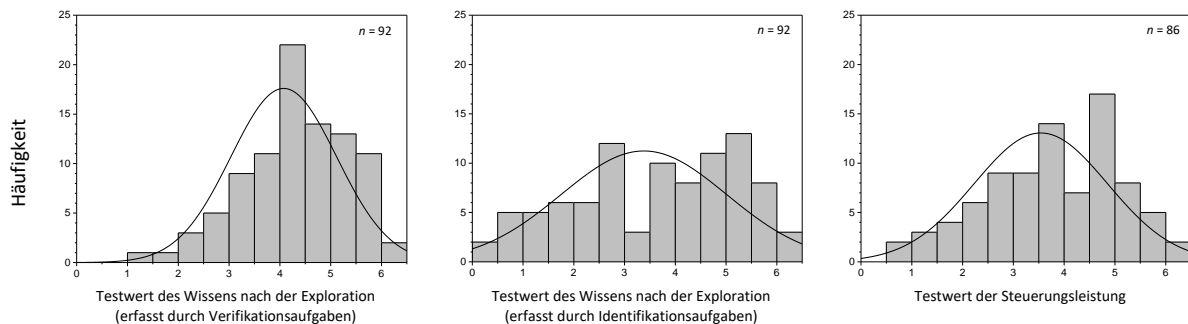


Abbildung 11.3: Häufigkeit der beobachteten Testwerte für die Aufgabentypen Verifikation, Identifikation und Steuerung des allgemeinen Problemlösetests

Die interne Konsistenz, als Maß für die Reliabilität beträgt für die Wissensabfrage mit Verifikationsaufgaben lediglich $\alpha = .47$, für die Wissensabfrage mit Identifikationsaufgaben hingegen $\alpha = .79$. Die geringe Messgenauigkeit der Verifikationsaufgaben zeigt sich auch in einem geringeren korrelativen Zusammenhang zwischen dem erfassten Wissen nach der Exploration und der Wissensanwendung in der Steuerungsphase (Tabelle 11.13). Dieser Zusammenhang beträgt zwischen dem nach der Exploration mit Verifikationsaufgaben erfassten Wissen und der Steuerung $r = .54$; während der Zusammenhang zwischen dem durch Identifikationsaufgaben erfassten Wissen und der Steuerung $r = .64$ beträgt.

Tabelle 11.13: Interkorrelation zwischen dem Wissen nach der Exploration (Verifikation und Identifikation) und der Wissensanwendung in der Steuerung im allgemeinen Problemlösetest

	Identifikation	Steuerung
Verifikation	.67**	.54**
Identifikation	-	.64**
Steuerung	-	-

Anmerkungen: manifeste Produkt-Moment-Korrelation; Korrelationen mit Steuerung sind einseitig; ** p < .01

Wird in der weiteren Betrachtung der allgemeinen Problemlösefähigkeit zwischen den latenten Variablen Wissen und Steuerung differenziert, gehen in das Messmodell der latenten Variable ausschließlich die Identifikationsitems als manifeste Variable ein.

11.7 Problemlösekompetenz im Umgang mit technischen Alltagsgeräten

Der computerbasierte Test zur Erfassung des problemlösenden Umgangs mit technischen Alltagsgeräten lag allen 147 Probanden zur Bearbeitung vor. Ihr in dem Test gezeigtes Verhalten wurde nach den unterschiedlichen zu erreichenden Interaktionszielen vorgenommen. Zur Bewertung des explorierenden Umgangs mit den technischen Systemen wurde die Vollständigkeit der Systemexploration herangezogen. Bei der Explorationsvollständigkeit handelt es sich um eine annähernd kontinuierliche Variable, die Werte zwischen 0 und 100 % (bzw. zwischen 0 und 1) annehmen kann. Um zwischen den Leistungen der Probanden bei der Systemsteuerung differenzieren zu können, stehen drei Scoringalternativen zur Verfügung. Die einzelnen Items können dichotom (falsch/richtig), polytom (Anzahl der erreichten Teilziele/Anzahl aller zu erreichenden Teilziele) oder dichotom, mit der Bewertung der Problemlösegröße (0; 1; 2; 3) ausgewertet werden¹.

Sowohl die Exploration als auch die Steuerung eines technischen Systems konnte von den Probanden vor Ablauf einer maximalen Bearbeitungszeit beendet werden, indem ein Weiter-Button betätigt wurde. Ein Datum wurde dann als fehlender Wert deklariert, wenn dieser Weiter-Button unmittelbar (innerhalb weniger Sekunden) nach dem Aufruf des entsprechenden Items gedrückt wurde, so dass davon auszugehen war, dass Probanden entweder nach Beendigung eines Items den Weiter-Button gedrückt haben um zu dem nächsten zu gelangen, eine verlängerte Ladezeit aber dazu geführt hat, dass einzelne Probanden diesen Button ein weiteres Mal gedrückt haben oder durch das ähnliche Aussehen der Explorations- und Steuerungsitems eines technischen Systems, letzteres übersehen haben. In acht der fünfzehn Items entstanden für den explorierenden Umgang jeweils zwei bis drei und in fünf Items für den steuernden Umgang je ein bis drei fehlende Werte.

¹ Eine genaue Beschreibung des Scorings findet sich im Kapitel 8.5.3.

Tabelle 11.14: Explorationsvollständigkeit für den explorierenden Umgang und relative Lösungshäufigkeiten für den steuernden Umgang mit technischen Systemen in dem computerbasierten technischen Problemlösetest

Item	Explorationsvollständigkeit					Steuerungsleistung										
						dichotome Auswertung			teilweise polytome Auswertung*				dichotome Auswertung berücksichtigt Bearbeitungsschritte**			
	<i>n</i>	<i>MW</i>	<i>SD</i>	<i>MIN</i>	<i>MAX</i>	<i>n</i>	falsch	richtig	<i>MW</i>	<i>SD</i>	<i>MIN</i>	<i>MAX</i>	0	1	2	3
TPL_1	144	.77	.19	.23	1.00	147	40	60	.88	.16	.50	1.00	40	20	9	31
TPL_2	145	.88	.20	.00	1.00	147	14	86	.86	.35	.00	1.00	14	16	1	69
TPL_3	145	.37	.13	.07	.71	147	90	10	.10	.30	.00	1.00	90	10	0	0
TPL_4	145	.64	.24	.06	1.00	147	15	85	.85	.36	.00	1.00	15	24	10	52
TPL_5	145	.50	.23	.03	1.00	146	36	64	.64	.48	.00	1.00	36	61	1	1
TPL_6	147	.77	.15	.12	1.00	147	30	70	.82	.29	.00	1.00	30	22	25	23
TPL_7	147	.45	.33	.06	1.00	144	22	78	.78	.42	.00	1.00	22	78	0	0
TPL_8	147	.32	.17	.02	.73	147	34	66	.74	.41	.00	1.00	34	48	13	5
TPL_9	147	.67	.16	.09	.91	147	57	43	.66	.38	.00	1.00	57	32	7	4
TPL_10	147	.86	.17	.22	1.00	146	65	35	.55	.41	.00	1.00	65	15	3	16
TPL_11	147	.26	.16	.02	.62	147	69	31	.31	.47	.00	1.00	69	20	6	5
TPL_12	145	.44	.23	.15	1.00	147	82	18	.18	.38	.00	1.00	82	17	1	0
TPL_13	147	.73	.24	.00	1.00	146	74	26	.62	.36	.00	1.00	74	5	8	13
TPL_14	145	.47	.20	.00	.95	146	56	44	.44	.50	.00	1.00	56	21	9	14
TPL_15	145	.18	.11	.08	.62	147	72	28	.75	.25	.11	1.00	72	23	5	0

Anmerkungen: *n* = Stichprobengröße; *MW* = Mittelwert; *SD* = Standardabweichung; *MIN* = Minimum; *MAX* = Maximum; *die folgenden Items können nur dichotom ausgewertet werden: TPL_2, TPL_3, TPL_4, TPL_5, TPL_7, TPL_11, TPL_12, TPL_14; ** 0 = falsch; 1 = richtig, mit über 100% mehr Bearbeitungsschritten als nötig, 2 = richtig, mit über 50% mehr Bearbeitungsschritten als nötig, 3 = richtig, mit weniger als 50% mehr Bearbeitungsschritten als nötig; TPL = Technisches Problemlösen

Die relativen Häufigkeiten der erreichten Itemwerte für den explorierenden Umgang sowie für den steuernden Umgang, getrennt nach den drei Auswertungsalternativen, sind in Tabelle 11.14 zusammengefasst. Die Mittelwerte der Explorationsvollständigkeit der einzelnen technischen Systeme reichen von $MW_{TPL15} = .18$ bis $MW_{TPL2} = .88$. Das Schwierigkeitsspektrum der Items für den explorierenden Umgang wurde dabei ausreichend abgedeckt. Die relativen Häufigkeiten der beiden Kategorien richtig und falsch für die dichotome Auswertung der Systemsteuerung deuten auf eine mittlere Testschwierigkeit hin. Sieben Items weisen Lösungshäufigkeiten oberhalb von 50 % und acht Items Lösungshäufigkeiten unterhalb von 50 % auf. Das, in dem Item 12 vorgegebene Steuerungsziel die Displaysperre in dem Bedienpanel einer Rollladensteuerung einzustellen, konnte von nur 17.69 % der Probanden erreicht werden, das Item 12 zählt damit für die Systemsteuerung zu den schwierigen Items. Das in dem Item 2 vorgegebene Steuerungsziel ein Ticket an einem Parkscheinautomaten zu kaufen konnte hingegen von 85.71 % der Probanden erreicht werden, womit das Item 2 für die Systemsteuerung das leichteste Item für die untersuchte Stichprobe ist.

Für sieben der fünfzehn Steuerungsisems konnten die Steuerungsziele in Teilziele differenziert werden, für deren Erreichung Teilpunkte vergeben wurden, somit ließen sich diese Items auch danach auswerten, wie groß der relative Anteil der erreichten Teilziele ist. Die Mittelwerte für die polytome Auswertung zeigen, dass die Gesamttestschwierigkeit durch diese Auswertemethode sinkt, während bei dichotomer Auswertung acht Items eine Lösungshäufigkeit von über 50 % aufweisen, sind es bei polytomer Auswertung sogar elf Items. Für die Items, für die keine polytome Auswertung möglich ist, entsprechen die angegebenen Mittelwerte bei polytomer Auswertung den relativen Häufigkeiten aus der Kategorie *richtig* bei dichotomer Auswertung.

Wird neben dem vollständigen Erreichen eines Steuerungsziels (dichotome Auswertung) außerdem die Güte der Problemlösung berücksichtigt, indem die Differenz zur der minimal nötigen Anzahl an Bearbeitungsschritten in das Scoring einfließt, ergeben sich höhere Schwierigkeiten für die Steuerung der Systeme. Die relative Häufigkeit, für die Vergabe von drei Punkten (beim Erreichen des Steuerungsziels mit bis zu 50 % mehr Interaktionsschritte als nötig), beträgt für vier der Items 0 %. Für den größten Teil der Items, wurden über 100 % mehr an Interaktionen benötigt, um das entsprechende Steuerungsziel zu erreichen. Eine detaillierte Analyse der Itemschwierigkeiten wird in Kapitel 13 vorgenommen.

Für die Bildung individueller Testwerte für die Systemexploration werden die jeweils erreichten Explorationsvollständigkeiten aufsummiert und durch die Anzahl der fünfzehn Items dividiert, so dass wieder ein Maß entsteht, dessen Wertebereich sich von 0 bis 1 erstreckt. Der Testwert für die Systemsteuerung wird durch das Aufsummieren der einzelnen Itemwerte gebildet, und nimmt für die dichotome und die polytome Auswertung Werte zwischen 0 und 15 und für die dichotome Auswertung, die die Anzahl benötigter Interaktionen berücksichtigt, Werte zwischen 0 und 45 an.

Hinweise auf die Verteilung der beobachteten Testwerte lassen sich Tabelle 11.15 entnehmen. Für die Leistung im explorierenden Umgang mit technischen Systemen weicht der Koeffizient der Schiefe mit $g_1 = .01$ nicht und der Koeffizient des Exzess mit $g_2 = -.28$ geringfügig von einer Normalverteilung ab. Obwohl das Critical Ratio beider Koeffizienten gegen eine signifikante Abweichung spricht, wird das Distanzmaß $D = .08$ des Kolmogorov-Smirnov-Tests mit

$df = 136$ signifikant. Der zu Beginn des Kapitels erwähnte Shapiro-Wilk-Test, der vor allem bei kleinen Stichproben eingesetzt wird, spricht wiederum gegen eine Verletzung der Normalverteilung ($W = .99$; $p = .16$; $df = 136$).

Tabelle 11.15: deskriptive Statistiken des explorierenden und steuernden Umgangs mit den simulierten technischen Alltagsgeräten im technischen Problemlösekompetenztest

	<i>n</i>	<i>MW</i> (<i>MIN</i> , <i>MAX</i>)	<i>SD</i>	95 % <i>CI</i>	g_1	<i>SE</i>	<i>c.r.</i> (g_1)	g_2	<i>SE</i>	<i>c.r.</i> (g_2)	K-S-Test	
											<i>D</i>	<i>p</i>
Exploration	136	.57 (.30, .81)	.11	[.55, .58]	.01	.21	.06	-.28	.41	-.68	.08	.02
Steuerung ^a	141	7.54 (.00, 14.00)	2.74	[7.08, 7.99]	-.33	.20	-1.62	.03	.41	.06	.11	.00
Steuerung ^b	141	9.26 (1.55, 14.20)	2.51	[8.84, 9.68]	-.64	.20	-3.16	.46	.41	1.14	.07	.07
Steuerung ^c	141	13.29 (.00, 27.00)	5.68	[12.34, 14.24]	.06	.20	.30	-.37	.41	-.92	.06	.20

Anmerkungen: *n* = Stichprobengröße; *MW* = Mittelwert; *MIN* = Minimum; *MAX* = Maximum; *SD* = Standardabweichung; 95 % *CI* = Konfidenzintervall; *c.r.* = Critical Ratio; g_1 = Schiefe; g_2 = Exzess; *SE* = Standardfehler; K-S-Test = Kolmogorov-Smirnov-Test auf Prüfung der Normalverteilung; *D* = Prüfgröße des Kolmogorov-Smirnov-Tests mit $df = n$; *p* = *p*-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; ^a = dichotome Auswertung; ^b = polytome Auswertung; ^c = dichotome Auswertung berücksichtigt benötigte Bearbeitungsschrittzahl

Die Verteilung der beobachteten Daten für die dichotom ausgewertete Steuerungsleistung weicht ebenfalls geringfügig von einer Normalverteilung ab. Der Koeffizient für die Schiefe beträgt $g_1 = -.33$ und der Koeffizient für den Exzess $g_2 = .08$ und auch hier wird das Ergebnis des Kolmogorov-Smirnov-Tests, aber auch das Ergebnis des Shapiro-Wilk-Tests signifikant. Die in einem Histogramm aufgetragenen Häufigkeiten der beobachteten Testwerte der dichotom ausgewerteten Steuerungsleistung (Abbildung 11.4) lassen keine gravierenden Abweichungen bezüglich der Schiefe erkennen, so dass davon auszugehen ist, dass einzelne Ausreißer die Abweichungen verursacht haben.

Der erheblich höhere Mittelwert (bei gleichbleibendem theoretischen Wertebereich) der polytom ausgewerteten Steuerungsleistung wird in Abbildung 11.4 durch eine nach rechts verschobene Verteilung deutlich. Der Koeffizient für die Schiefe $g_1 = -.64$ deutet auf eine linksschiefe Verteilung hin, dessen Critical Ratio mit $c.r.(g_1) = -3.16$ signifikant wird. Der maximale Abstand zwischen der empirischen und der Normalverteilung, der durch den Kolmogorov-Smirnov-Test gemessen wird, beträgt $D = .07$ und wird tendenziell signifikant ($p = .07$; $df = 141$). Während die Verteilung der Testwerte bei polytomer Auswertung in Richtung des höheren Wertebereichs verschoben ist, führt eine dichotome Auswertung, bei der die Anzahl der benötigten Interaktionen berücksichtigt werden, zu einer deutlichen Verschiebung in Richtung des niedrigeren Wertebereichs. Obwohl theoretisch ein maximaler Testwert von 45 Punkten denkbar ist, liegt das Maximum der tatsächlich beobachteten Testwerte bei 27 Punkten. In diesem eingeschränkten Wertebereich folgen die Daten weitestgehend einer Normalverteilung ($D = .06$; $p > .20$, $df = 141$).

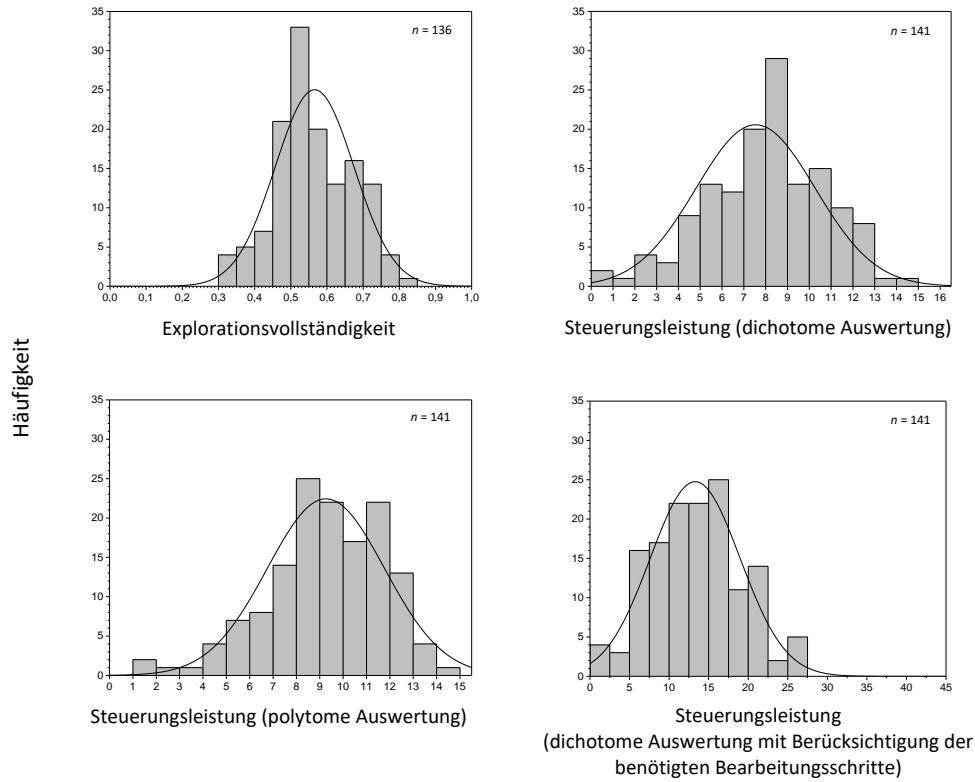


Abbildung 11.4: Häufigkeiten der beobachteten Testwerte für die Exploration und die Steuerung technischer Geräte

Eine Betrachtung der Messgenauigkeit der einzelnen Teile des konstruierten Tests folgt im Kapitel 12, das außerdem auf weitere Testgüteaspekte wie die Objektivität, die Item- und Stichprobenunabhängigkeit eingeht. Itemschwierigkeiten und ihre Ursachen werden dann im Abschnitt 13 analysiert. Hierbei werden die Testwerte, die aus der dichotomen Bewertung der Steuerungsphase, die außerdem die Anzahl der Interaktionen berücksichtigt, entstehen, noch einmal aufgegriffen, da sich bei dieser Auswertungsmethode die Gesamttestschwierigkeit erheblich verändert. Da eine polytome Steuerungsauswertung nicht für alle Items möglich ist, werden aus Gründen der Konsistenz bei den weiteren Analysen ausschließlich die dichotomen Ergebnisse der Steuerungsleistung verwendet.

11.8 Erste Zusammenhangsbetrachtungen

Das Ziel einen Test zur Erfassung der Kompetenz im problemlösenden Umgang mit technischen Systemen zu entwickeln, erfordert die Prüfung der theoretischen Legitimation eines solchen Konstruktes, sprich die Prüfung der Konstruktvalidierung. Hierzu beschäftigt sich der Abschnitt 5 mit den theoretischen Beziehungen zu angrenzenden Konstrukten, die im Abschnitt 6.6 in einem nomologischen Netzwerk zusammengefasst wurden. Für die im Kapitel 15 erfolgende Konstruktvalidierung, bei der das postulierte nomologische Netzwerk schrittweise auf Fehler überprüft wird, enthält Kapitel 7 entsprechende Hypothesen.

Voraussetzung für das Vorhandensein der in den Hypothesen postulierten kausalen Beziehungen zwischen den Variablen des nomologischen Netzwerkes ist das Bestehen von empirischen Zusammenhängen. Eine kausale Beziehung zwischen zwei Variablen lässt sich nur dann nicht verwerfen, wenn die unabhängige und die abhängige Variable miteinander korrelieren: „A bivariate association is neither a necessary nor sufficient condition for a causal relation“ (Bollen, 1989, S. 57). In diesem Abschnitt sollen zunächst die korrelativen Zusammenhänge zwischen den einzelnen manifesten Variablen betrachtet werden, die für die Prüfung des kausalen Zusammenhangs vorausgesetzt werden. Dabei werden zunächst keine mediatorischen oder moderierenden Effekte berücksichtigt, diese sind Inhalt des Abschnitts 15.

Eine erste Zusammenhangsbetrachtung erfolgt für die Variablen des gerätespezifischen Vorwissens und den explorierenden und steuernden Umgang mit technischen Alltagsgeräten. Angenommen wird, dass sich das jeweils spezifische Vorwissen positiv auf die Systemexploration auswirkt, das heißt je höher das jeweilige Vorwissen, desto vollständiger wird das entsprechende System exploriert. Die sich aus den erhobenen Daten ergebenden Zusammenhänge sind in Tabelle 11.16 wiedergegeben. Mit Ausnahme zweier technischer Systeme (Item 3 und Item 5) kann für keines der Geräte ein Einfluss des Vorwissens auf den explorierenden Umgang mit dem entsprechenden System festgestellt werden. Der mit einer Methode nach Jäger (1974) über die Z-Transformation gemittelte Korrelationskoeffizient beträgt für die Korrelation zwischen dem gerätespezifischen Vorwissen und der Explorationsvollständigkeit $\bar{r} = .01$. Zwar ist der Zusammenhang zwischen dem Vorwissen und der Exploration bei den zwei ausgenommenen Items signifikant von Null verschieden, die Höhe der Korrelation ist jedoch gering. Sie beträgt für das Item 3 (die Exploration des technischen Systems Spülmaschine) $r = .20$ und für das Item 5 (die Exploration des Multifunktionsdruckers) $r = -.27$. Während die Korrelation in dem einen Fall positiv ist, das heißt ein hohes Vorwissen hat zu einem hohen Wert an Explorationsvollständigkeit geführt (Kausalität darf aufgrund der zeitlichen Ordnung der Variablen angenommen werden), ist sie in dem anderen Fall negativ, ein hohes Vorwissen hat hier zu einer geringeren Explorationsvollständigkeit geführt.

Tabelle 11.16: bivariate Korrelationen (Spearman) zwischen dem gerätespezifischen Vorwissen und der Vollständigkeit der Systemexploration und der Steuerungsleistung für die 15 technischen Systeme

Exploration	Gerätespezifisches Vorwissen														
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
TPL_1	.08														
TPL_2		-.15													
TPL_3			.20**												
TPL_4				-.09											
TPL_5					-.27**										
TPL_6						.00									
TPL_7							.14								
TPL_8								.11							
TPL_9									-.06						
TPL_10										.10					
TPL_11											.05				
TPL_12												.12			
TPL_13													-.10		
TPL_14														-.02	
TPL_15															.10
Steuerung															
TPL_1	.11														
TPL_2		-.11													
TPL_3			.20**												
TPL_4				-.05											
TPL_5					.00										
TPL_6						.15*									
TPL_7							.07								
TPL_8								.02							
TPL_9									-.14*						
TPL_10										.13					
TPL_11											.08				
TPL_12												.12			
TPL_13													.00		
TPL_14														-.01	
TPL_15															.26**

Die Analyse des Einflusses des gerätespezifischen Vorwissens auf den steuernden Umgang mit den einzelnen technischen Alltagsgeräten führt zu ähnlichen Ergebnissen. Lediglich für vier Items (Item 3, Item 6, Item 9 und Item 15) wird der Korrelationskoeffizient zwar signifikant, seine Höhe ist jedoch mit einem maximalen Wert von $r = .26$ als gering einzustufen. Der gemittelte Korrelationskoeffizient über alle 15 Korrelationen beträgt für den Zusammenhang zwischen dem gerätespezifischen Vorwissen und dem steuernden Umgang mit technischen Systemen $\bar{r} = .05$. Ob ein, über alle technischen Systeme hinweg generalisiertes Vorwissen wiederum einen Einfluss auf das Explorieren und Steuern eines technischen

Gerätes hat, wird im Abschnitt 15 beantwortet, indem die Items des gerätespezifischen Vorwissens als Einzelindikatoren zu einer latenten Variable zusammengefasst und deren Einfluss auf die Systemexploration und die Systemsteuerung betrachtet wird.

Eine weitere Zusammenhangsbetrachtung, die an dieser Stelle vorgenommen wird, ist die der Systemexploration und der Systemsteuerung. Angenommen und in Abschnitt 14 geprüft wird, dass Systemexploration und Systemsteuerung zwei voneinander abgrenzbare Interaktionsformen mit jeweils unterschiedlichen Interaktionszielen sind. Weil für die gezielte Steuerung eines technischen Systems jedoch Wissen vorausgesetzt wird, dessen Erwerb Ziel der Systemexploration ist, erfolgt die Systemsteuerung zeitlich nach der Systemexploration, so dass zwar angenommen wird, dass es sich bei ihnen um getrennte Interaktionsformen handelt, diese aber aufeinander bezogen sind. Tabelle 11.17 enthält die nach Pearson berechneten Korrelationskoeffizienten für den Zusammenhang zwischen Systemexploration und der nach den bereits vorher beschriebenen unterschiedlich ausgewerteten Ergebnissen für die Systemsteuerung.

Tabelle 11.17: Skaleninterkorrelationen der manifesten Variablen Systemexploration und Systemsteuerung

	Steuerung ^a	Steuerung ^b	Steuerung ^c
Exploration	.55**	.57**	.66**
Steuerung ^a		.96**	.89**
Steuerung ^b			.84**
Steuerung ^c			

Anmerkungen: ** $p < .01$; zweiseitige Interkorrelationen (Pearson) zwischen den einzelnen Steuerungstestwerten basieren auf $n = 141$; einseitige Korrelationen (Pearson) zwischen Explorations- und den Steuerungstestwerten basieren auf $n = 131$;
^a = dichotome Auswertung; ^b = polytome Auswertung; ^c = dichotome Auswertung berücksichtigt benötigte Bearbeitungsschrittzahl

Zwischen den, nach unterschiedlichen Methoden gebildeten Steuerungsvariablen besteht mit bis zu $r = .96$ ein sehr hoher Zusammenhang, der für Variablen, die dasselbe messen sollen auch anzustreben ist. Die Höhe des signifikanten Zusammenhangs zwischen der Systemexploration und der dichotom ausgewerteten Steuerungsleistung beträgt $r = .55$. Der mit $r = .57$ höher ausfallende Korrelationskoeffizient zwischen der Systemexploration und der polytom ausgewerteten Steuerungsleistung beruht auf einer größeren Variationsbreite der so ermittelten Steuerungsergebnisse. Während die Differenz der Korrelationskoeffizienten für den Zusammenhang zwischen Systemexploration und Systemsteuerung zwischen der dichotomen und der polytomen Auswertung des Steuerungsergebnisses gering ist, fällt sie zwischen dichotomer und der Anzahl an benötigten Interaktionen berücksichtigender dichotomer Auswertung mit $r = .66$ etwas größer aus. Fließt in die Bewertung des Steuerungsergebnisses die Anzahl der für die Steuerung benötigten Bearbeitungsschritte mit ein, so ist eine vorherige Systemexploration also umso bedeutsamer.

Eine Zusammenschau aller Korrelationen zwischen den einzelnen manifesten Skalen- und Testwerten enthält Tabelle 11.18. Der erwartete Einfluss der Intelligenz auf die Exploration und die Steuerung der technischen Systeme als auch der fachübergreifenden Problemszenarien deutet sich in den signifikanten Korrelationen der entsprechenden Variablen an. Die Korrelationskoeffizienten für den Zusammenhang zwischen Intelligenz und allgemeiner Problemlösefähigkeit fallen erwartungskonform höher aus, als für den

Zusammenhang zwischen Intelligenz und technischer Problemlösekompetenz. Keinen Zusammenhang gibt es scheinbar zwischen Intelligenz und den beiden Variablen des Vorwissens: ICT-Nutzung und das gerätespezifische Vorwissen. Dieser Zusammenhang sollte noch einmal anhand des Strukturmodells in Kapitel 15 untersucht werden, da die dortigen Zusammenhänge zwischen den latenten Variablen messfehlerbereinigt sind und demnach höher ausfallen können als zwischen den hier betrachteten manifesten Variablen. Gleiches gilt auch für die gering ausfallende Korrelation zwischen dem gerätespezifischen Vorwissen und der Systemsteuerung und der hier nicht feststellbaren Verbindung zwischen ICT-Nutzung und dem technischen Problemlösen.

Ein signifikanter Zusammenhang kann zwischen den beiden Skalen zur Erfassung der Selbstwirksamkeit sowie zwischen den beiden und dem Konstrukt Need for Cognition ausgemacht werden. Von diesen drei Skalen hat jedoch lediglich die Aufgeschlossenheit gegenüber dem Problemlösen einen Einfluss auf das technische Problemlösen, allerdings einen negativen Einfluss auf das Explorationsverhalten.

Tabelle 11.18: Interkorrelationen zwischen allen manifesten Skalen- und Testwerten

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) fluide Intelligenz	--								
(2) Need for Cognition	.18 *	--							
<i>Vorwissen</i>									
(3) gerätespezifisches Vorwissen	.11	.27 **	--						
(4) ICT-Nutzung	.03	-.05	.17 *	--					
<i>Selbstwirksamkeit</i>									
(5) Ausdauer	-.06	.42 **	.22 **	.02	--				
(6) Aufgeschlossenheit gegenüber Problemlösen	.13	.54 **	.47 **	.14	.32 **	--			
<i>Allgemeine Problemlösefähigkeit</i>									
(7) Wissen nach der Exploration	.55 **	.06	.22 *	.16	-.15	.04	--		
(8) Steuerungsleistung	.58 **	-.04	.18	.08	-.23 *	.07	.64 **	--	
<i>Technische Problemlösekompetenz</i>									
(9) Explorationsvollständigkeit	.35 **	-.04	.01	-.03	-.08	-.20 *	.47 **	.51 **	--
(10) Steuerungsleistung	.49 **	-.03	.21 *	.12	-.09	.04	.57 **	.58 **	.55 **

Anmerkungen: * $p < .05$; ** $p < .01$

12 Güte des Messinstruments zur Erfassung des problemlösenden Umgangs mit technischen Alltagsgeräten

Bevor in den nächsten Abschnitten verschiedene Validitätsaspekte des theoretisch angenommenen Konstruktes geprüft werden, liegt der Fokus in diesem Kapitel auf der Analyse des entwickelten Testinstruments, mit dem das zu validierende Konstrukt erfasst werden soll. Da angenommen wird, dass sich der problemlösende Umgang mit technischen Alltagsgeräten in die Dimensionen *Systemexploration* und *Systemsteuerung* differenzieren lässt, werden die als latente Merkmale zu betrachtenden Dimensionen im Folgenden zunächst getrennt voneinander untersucht. Ob es sich bei den supponierten Dimensionen tatsächlich um voneinander abgrenzbare Merkmale handelt, ist dann Inhalt des Kapitels 14.

Für jedes der beiden latenten Merkmale wird geprüft, ob sich die jeweils 15 Indikatoren, mit denen sie erfasst werden, in ihrer Trennschärfe unterscheiden. Hierfür wird zunächst für beide Merkmale ein kongenerisches Messmodell spezifiziert, bei dem sowohl die Itemschwierigkeiten als auch die Faktorladungen variieren dürfen. Dieses wird dann gegen ein restringiertes essentiell τ -äquivalentes Messmodell getestet, bei dem die Faktorladungen gleichgesetzt werden, das heißt, die Items sich in ihren Diskriminationsfähigkeiten nicht unterscheiden (Eid & Schmidt, 2014, S. 313). Berichtet werden für die entsprechend besser passenden Messmodelle außerdem die Itemschwierigkeiten sowie die Reliabilitäten.

12.1 Messmodell der Systemexploration

Bei den Indikatoren, die den explorierenden Umgang mit technischen Geräten erfassen, handelt es sich um die relativen Systemexplorationsvollständigkeiten der fünfzehn simulierten Technikgeräte, deren Skala als annähernd kontinuierlich angenommen werden kann. Um die Stabilität der Faktorladungen über unterschiedliche Strukturmodelle hinweg betrachten zu können, wird zur Parameterschätzung des Messmodells der Systemexploration das gleiche Schätzverfahren eingesetzt, wie es für das Messmodell der Systemsteuerung notwendig ist. Da die Indikatoren der Systemsteuerung dichotome Variablen sind, bei denen nicht von einer Normalverteilung ausgegangen werden kann, können keine Likelihood-basierten Schätzverfahren eingesetzt werden (Reinecke, 2014, S. 109), da ansonsten die χ^2 -Statistik und die Standardfehler der einzelnen Parameterschätzungen verzerrt würden (Finney & DiStefano, 2006, S. 273).

Das prinzipiell für kategoriale und nicht normalverteilte Daten geeignete WLS (weighted least squares)-Verfahren erfordert für die Berechnung der asymptotischen Kovarianzmatrix (Gewichtungsmatrix), mit der die Parameter geschätzt werden, eine hohe Stichprobengröße (Kline, 2011, S. 178), die in dieser Arbeit mit $n = 147$ nicht zu erreichen ist. Stattdessen wird eine modifizierte Variante des WLS-Verfahrens genutzt, bei der ausschließlich auf die Elemente auf der Diagonalen der Gewichtungsmatrix zurückgegriffen wird, um die Parameter zu schätzen (Finch & French, 2015, S. 40). Die Informationsreduktion führt jedoch zu verzerrten Standardfehlern und Teststatistiken, so dass bei diesem sogenannten DWLS (diagonally weighted least squares)-Verfahren eine Korrektur dieser nötig wird (Wirth & Edwards, 2007, S. 65). Das DWLS-Verfahren bietet sich vor allem an, wenn einige oder alle der endogenen Indikatorvariablen in einem Modell kategorial und nichtnormalverteilt sind (Lei

& Wu, 2015, S. 169) sowie die Anzahl der Faktoren und Indikatoren in dem Modell groß sind (Wirth & Edwards, 2007, S. 64).

Ob sich die einzelnen Items in ihrer Eignung, zwischen Personen differenzieren zu können, deren Kompetenz im explorierenden Umgang mit technischen Geräten verschieden ausgeprägt sind, unterscheiden, wird durch den Vergleich von zwei Messmodellen geprüft. Das erste (essentiell τ -äquivalente) Modell M1 nimmt an, dass sich die Items zwar in ihrer Schwierigkeit unterscheiden, nicht aber in ihrer Trennschärfe. Die Faktorladungen, die in dieser Arbeit als Trennschärfekoeffizienten betrachtet werden¹, werden deshalb gleichgesetzt. Das zweite (kongenerische) Modell M2 hingegen nimmt an, dass sich die Items sowohl in ihrer Schwierigkeit, als auch in ihrer Trennschärfe unterscheiden, weshalb in diesem Modell ein Gleichsetzen der Faktorladungen ausbleibt. Im ersten Schritt wird die empirische Gültigkeit beider Modelle geprüft, dessen Ergebnisse Tabelle 12.1 zu entnehmen sind. Fehlende Werte werden hierbei listenweise ausgeschlossen, so dass die genannten Fitstatistiken auf $n = 136$ beruhen. Die nach Satorra-Bentler korrigierte χ^2 -Statistik wird mit einem cut-off-Kriterium von $p < .01$ für beide Modelle signifikant, so dass von Fehlspezifikationen in den Modellen ausgegangen werden muss. Die weiteren Fitwerte deuten auf eine akzeptable Passung zwischen den empirischen Daten und dem kongenerischen Messmodell hin (CFI = .97; WRMR = .69), wohingegen sie für die Passung des essentiell τ -äquivalenten Messmodell M1 einen deutlichen Missfit zeigen (CFI = .71; WRMR = 1.70).

Tabelle 12.1: Fitstatistiken für die Messmodelle der Systemexploration ($n = 136$)

	χ^2	df	p	χ^2/df	RMSEA (90 % CI)	CFI	WRMR	$\Delta\chi^2 (\Delta df)$	ΔCFI
M2: Systemexploration kongenerisches Messmodell	133	90	.002	1.47	.06 (.02, .09)	.97	.69		
M1: Systemexploration essentiell τ -äquivalentes Messmodell	466.23	104	.000	4.48	.12 (0.15, 0.18)	.71	1.70	105.37 (14)***	-.25

Anmerkungen: n = Stichprobengröße; χ^2 = Prüfgröße der Modellabweichung (Nach Satorra-Bentler korrigierte χ^2 -Werte); df = Anzahl der Freiheitsgrade; p = p-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; RMSEA = Root Mean Square Error of Approximation; 90 % CI = Konfidenzintervall; CFI = Comparative Fit Index; WRMR = Weighted Root Mean Square Residual; die Differenzen $\Delta\chi^2 (\Delta df)$ sind die korrigierten Differenzen der beiden χ^2 Werte; *** $p < .001$

Aussagen über den direkten Vergleich der beiden Modelle sind jedoch nur durch einen statistischen Test möglich, bei dem die Nullhypothese das restriktivere Modell aus Sparsamkeitsgründen beizubehalten, geprüft wird. Der Vergleich zweier χ^2 -Statistiken, die auf denselben Daten beruhen, kann mit dem χ^2 -Differenzentest erfolgen. Bei den zu vergleichenden Modellen muss es sich allerdings um genestete Modelle handeln, das heißt, dass das eine Modell nur durch Restriktionen aus dem anderen Modell hervorgehen darf. Das essentiell τ -äquivalente Messmodell M1 ist ein durch die Restriktion der Faktorladungen aus dem kongenerischen Modell M2 hervorgegangenes Modell, so dass die Anwendung des χ^2 -Differenzentests möglich ist.

Da aufgrund des verwendeten DWLS-Schätzverfahrens die χ^2 -Werte mit einem Skalierungsfaktor korrigiert wurden, dessen Höhe eine Funktion der modellimplizierten

¹ Weitere Ursachen für unterschiedliche Faktorladungen werden bei Eid, Gollwitzer & Schmitt (2013, S. 836) besprochen.

Gewichtungsmatrix, der multivariaten Kurtosis und der Anzahl der Freiheitsgrade des Modell ist (West, Finch & Curran, 1995, S. 65), müssen die für den Test benötigten χ^2 -Differenzen ebenfalls korrigiert werden, weil sie selbst nicht mehr einer χ^2 -Verteilung folgen (Satorra & Bentler, 2001). Das signifikante Ergebnis (Tabelle 12.1) des χ^2 -Differenzentests bestätigt die bereits anhand der anderen Modellfitwerte vorgenommene Vermutung, dass ein Messmodell, das unterschiedlich trennscharfe Items annimmt, die empirischen Daten besser erklären kann, als ein Modell, das von gleich hohen Trennschärfekoeffizienten ausgeht. Hierfür spricht auch die Differenz des CFI, die nach Cheung & Rensvold (2002) erst bei einem betraglich größeren Wert als $-.01$ gegen die Ablehnung der Nullhypothese spricht. Die Differenz des CFI beträgt zwischen den beiden hier getesteten Modell $\Delta\text{CFI} = -.25$, so dass das restriktivere Modell nicht beizubehalten ist.

Die geschätzten Werte der standardisierten Faktorladungen (bei auf 1 fixierter Varianz der latenten Variable) aller 15 Items sind in Tabelle 12.2 wiedergegeben. Sie betragen für die meisten der Items $\lambda \geq .30$, lediglich das Item 3 liegt mit $\lambda = .29$ unter diesem Wert. Alle geschätzten Faktorladungen sind mit $p < .001$ signifikant. Die Höhe der Varianz der einzelnen Itemwerte, die durch die latente Variable erklärt werden kann, wird durch das Maß der Kommunalität wiedergegeben, das sich durch das Quadrieren der einzelnen Faktorladungen bildet. Die Kommunalitäten der einzelnen Items können ebenfalls Tabelle 12.2 entnommen werden.

Tabelle 12.2: standardisierte Faktorladungen, Itemleichtigkeiten, Messfehlervarianzen und Itemreliabilitäten des kongenerischen Messmodells M2 der Systemexploration (n = 136)

	stand. Faktorladung	Itemleichtigkeit	Fehlervarianz	Kommunalität
Item 1	.60	.79	.64	.36
Item 2	.30	.89	.91	.09
Item 3	.29	.38	.92	.09
Item 4	.57	.66	.68	.32
Item 5	.73	.51	.46	.54
Item 6	.49	.78	.76	.24
Item 7	.40	.46	.84	.16
Item 8	.73	.33	.47	.53
Item 9	.57	.68	.68	.33
Item 10	.47	.86	.78	.22
Item 11	.76	.27	.43	.57
Item 12	.56	.45	.69	.31
Item 13	.54	.75	.71	.29
Item 14	.59	.48	.65	.35
Item 15	.44	.19	.81	.19

Anmerkungen: n = Stichprobengröße; Alle Parameter und Varianzen sind mit $p < .001$ signifikant

Die Betrachtung der standardisierten Faktorladungen und der Kommunalitäten zeigt, dass die Items 2 und 3 des Tests das latente Konstrukt nur unzureichend abbilden und damit erfassen können; die Varianzen in den Testwerten dieser Items können nur zu 9 % von der latenten Variable erklärt werden. Der Anteil der Varianzen, die nicht durch die latente Variable erklärt

werden können, werden als Messfehlervarianzen bezeichnet (vgl. Wang & Wang, 2012, S. 40); diese werden für jedes Item ebenfalls in Tabelle 12.2 wiedergegeben. Die Kommunalität eines Items als der Anteil der aufgeklärten Varianz ist ein Maß für die Reliabilität eines Items, also die Genauigkeit mit der ein Item die latente Variable messen kann. Aus den einzelnen Itemreliabilitäten des kongenerischen Messmodells lässt sich mit McDonalds Omega (McDonald, 1970) ein Maß für die Gesamtreliabilität berechnen, dieses beträgt für die Items der Systemexploration $\omega = .85$.

Weil es sich um ein, im Rahmen dieser Arbeit entwickeltes Testinstrument handelt, das in der beschriebenen (Pilotierungs-) Studie zum ersten Mal zum Einsatz kam, leistet das beschriebene sowie die noch folgenden Messmodelle einen Beitrag für die Itemselektion und die im Anschluss dieser Arbeit folgende Testüberarbeitung. So sollten ausschließlich die Items, die über eine ausreichende Diskriminationsfähigkeit beziehungsweise Itemreliabilität verfügen, behalten werden. Das Selektieren von Items hat zum Ziel die Gesamttestreliabilität zu erhöhen (Eid, Gollwitzer & Schmitt, 2013, S. 842f.). Entfernt werden sollten Items, deren Faktorladungen nicht signifikant und deren standardisierte Faktorladungen unterhalb eines Wertes von $\lambda = 30$ liegen (Wang & Wang, 2012, S. 39). Wie der grafischen Veranschaulichung des kongenerischen Messmodells für die Systemexploration in Abbildung 12.1 zu entnehmen ist, sollte demnach das Item 3 aus dem Test entfernt und dies für das Item 2, das genau auf der Grenze liegt, überlegt werden. Um den Aspekt der diskriminanten Validität im Abschnitt 15 näher zu untersuchen, ist ein reliables Testinstrument nötig, so dass an der entsprechenden Stelle die Itemselektion noch einmal aufgegriffen wird.

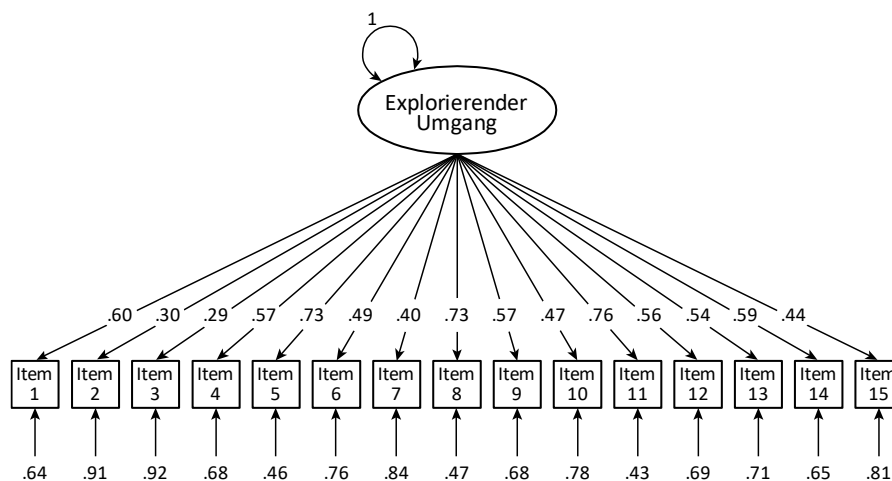


Abbildung 12.1: kongenerisches Messmodell M2 für die Systemexploration mit geschätzten Faktorladungen und Messfehlervarianzen ($n = 136$)

Die nicht standardisierten Itemleichtigkeiten entsprechen bei der vorgenommenen Restriktion (standardisierte latente Variable) im essentiell τ -äquivalenten und kongenerischen Messmodell den tatsächlich beobachteten Mittelwerten der Items (Eid, Gollwitzer & Schmitt, 2013, S. 839). Bei den, in Tabelle 12.2 dargestellten Itemleichtigkeiten handelt es sich demnach um die Mittelwerte der Explorationsvollständigkeiten. Dabei wird das Schwierigkeitsspektrum von 19 % (Item 15) bis 89 % (Item 2) ausreichend abgedeckt. Eine differenzierte Analyse der Itemschwierigkeiten und der sie verursachenden Itemmerkmale ist Inhalt des Kapitels 13.

12.2 Messmodell der Systemsteuerung

Der steuernde Umgang mit technischen Alltagsgeräten wird über Items erfasst, die von den Probanden das Erreichen eines vorgegebenen Steuerungsziels verlangen. Dabei müssen die Probanden das in der Systemexploration erworbene Eingriffswissen nutzen, womit die Position der Steuerungsisems nach den zugehörigen Explorationsitems festgelegt ist. Die dichotome Bewertung eines Items nach dem Erreichen bzw. Nichterreichen des gegebenen Steuerungsziels, erfordert ein für kategoriale Daten geeignetes Verfahren zur Parameterschätzung der unten spezifizierten Messmodelle. Verwendet wird das unter 12.1 bereits genannte DWLS-Verfahren mit nach Satorra-Bentler korrigierter χ^2 -Statistik und robusten Standardfehlern.

Zunächst wird durch einen Modellvergleich geprüft, ob sich die einzelnen Steuerungsziele der Systemsteuerung in ihrer Diskriminationsfähigkeit voneinander unterscheiden. In Analogie zu den oben analysierten Messmodellen der Systemexploration werden für diesen Zweck zwei genestete Messmodelle spezifiziert. In dem ersten (essentiell τ -äquivalente) Modell M1 werden die Faktorladungen der latenten Variable auf die einzelnen manifesten Indikatoren gleichgesetzt, womit postuliert wird, dass sich die Items hinsichtlich ihrer Trennschärfe nicht unterscheiden. In dem zweiten (kongenerischen) Messmodell M2 wird diese Restriktion nicht vorgenommen. Die Fitstatistiken der beiden Messmodelle für die Systemexploration sind in Tabelle 12.3 wiedergegeben. Fälle mit fehlenden Werten wurden listenweise aus der Berechnung ausgeschlossen. Sowohl für das essentiell τ -äquivalente als auch für das kongenerische Messmodell wird die nach Satorra-Bentler- χ^2 -Statistik signifikant und deutet damit auf eine unzureichende Passung zwischen der modellimplizierten und der empirischen Kovarianzmatrix hin. Der Wert für den RMSEA, der die Residuen zwischen den beiden Matrizen zusammenfasst, weist hingegen bei beiden Modellen nicht auf eine starke Abweichung hin. Der WRMR-Index, der analog zum RMSEA, die Residuen zu einem Abweichungsmaß zusammenfasst, dabei aber die kategoriale Natur der Items berücksichtigt, bleibt für das kongenerische Messmodell $M2 < 1.0$, für das essentiell τ -äquivalente Modell M1 hingegen deutet auch dieser Wert auf eine unzureichende Modellpassung hin. Der CFI, der das Theoriemodell mit einem Modell, in der die Variablen nicht zusammenhängen vergleicht, liegt bei beiden Modellen unterhalb des cut-off-Kriteriums von $CFI \geq .90$.

Tabelle 12.3: Fitstatistiken für die Messmodelle der Systemsteuerung ($n = 141$)

	χ^2	df	p	χ^2/df	RMSEA (90 % CI)	CFI	WRMR	$\Delta\chi^2 (\Delta df)$	ΔCFI
M2: Systemsteuerung kongenerisches Messmodell	135.17	90	.001	1.50	.06 (.04, .08)	.84	.95		
M1: Systemsteuerung essentiell τ -äquivalentes Messmodell	147.38	104	.003	1.42	.06 (.03, .07)	.85	1.08	16.72 (14) $p = .272$	-.006

Anmerkungen: n = Stichprobengröße; χ^2 = Prüfgröße der Modellabweichung (Nach Satorra-Bentler korrigierte χ^2 -Werte); df = Anzahl der Freiheitsgrade; p = p-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; RMSEA = Root Mean Square Error of Approximation; 90 % CI = Konfidenzintervall; CFI = Comparative Fit Index; WRMR = Weighted Root Mean Square Residual; die Differenzen $\Delta\chi^2 (\Delta df)$ sind die korrigierten Differenzen der beiden χ^2 Werte

Der Vergleich zwischen den beiden genesteten Messmodellen wurde mit einem χ^2 -Differenzentest durchgeführt, der die Differenz der beiden nach Satorra-Bentler

korrigierten X^2 -Werte korrigiert. Das nichtsignifikante Ergebnis des X^2 -Differenzentests ($\Delta X_{\text{korrt}}^2 = 16.72$; $\Delta df = 14$; $p = .27$) sowie die geringe Differenz zwischen den CFIs ($\Delta CFI = .006$) spricht für das Beibehalten des restriktiveren Modells, also des Messmodells, das für alle Indikatoren gleiche Diskriminationsparameter annimmt. Für die im Kapitel 14 durchzuführende Dimensionalitätsprüfung, das heißt der Prüfung, ob sich die angenommenen Subdimensionen Systemexploration und Systemsteuerung überhaupt voneinander trennen lassen, sollten sich die beiden zu vergleichenden Modelle nicht in ihren Restriktionen unterscheiden. Da sich für die Systemexploration das kongenerische Modell als besser zu den empirischen Daten passend herausgestellt hat, wird auch für die Systemsteuerung ein solches Modell verwendet, bei dem die Trennschärfekoeffizienten variieren dürfen.

Für dieses kongenerische Messmodell der Systemsteuerung sind die standardisierten Faktorladungen (bei auf den Wert 1 fixierter Varianz der latenten Variable) Tabelle 12.4 zu entnehmen. Wie durch die bessere Passung des essentiell τ -äquivalenten Messmodells zu vermuten ist, weisen die standardisierten Faktorladungen der Steuerungsisems ähnlich hohe Werte auf.

Tabelle 12.4: standardisierte Faktorladungen, Thresholds, Messfehlervarianzen und Itemreliabilität des kongenerischen Messmodells M2 der Systemsteuerung (n = 141)

	stand. Faktorladung	Thresholds	Fehlervarianz	Kommunalität
Item 1	.49	-.30	.76	.24
Item 2	.55	-1.14	.69	.31
Item 3	.27	1.25	.93	.07
Item 4	.44	-1.10	.81	.19
Item 5	.66	-.34	.56	.44
Item 6	.45	-.57	.80	.20
Item 7	.44	-.77	.81	.19
Item 8	.37	-.43	.86	.14
Item 9	.36	.17	.87	.13
Item 10	.48	.39	.77	.23
Item 11	.54	.45	.71	.29
Item 12	.70	.93	.51	.49
Item 13	.66	.61	.56	.44
Item 14	.24	.15	.94	.06
Item 15	.39	.55	.85	.15

Anmerkungen: n = Stichprobengröße; Alle Parameter und Varianzen sind mit $p < .05$ signifikant

Über die Hälfte aller standardisierten Faktorladungen liegen in dem Bereich zwischen .30 und .50. Die Faktorladungen zweier Items (Item 3 und Item 14) liegen unter dem Wert von .30. Bereits in der Systemexploration zeigte sich für das Item 3 eine geringe Faktorladung; der explorierende und steuernde Umgang mit diesem technischen Gerät (Spülmaschine) lässt sich demnach nur zu einem kleinen Teil auf die dahinterliegende latente Variable zurückführen und wird wohl durch weitere an dieser Stelle nicht erfasste Konstrukte beeinflusst. Die Faktorladung des Item 3 ist außerdem nur tendenziell signifikant ($p = .054$). In beiden Interaktionsphasen handelt es sich bei dem Item 3 um ein besonders schwieriges Item, das von

nur wenigen Personen gelöst werden konnte. Worin die Ursachen der geringen Lösungshäufigkeit dieses Items liegen, wird in den nächsten Abschnitten besprochen. Durch die zum größten Teil niedrigen bis mittelhohen Faktorladungen kann auch der geringe Wert des CFI erklärt werden, denn dieser ist stark von der Höhe der Faktorladungen beeinflusst (Bühner, 2012). Die Kommunalitäten der Items 3 und 14 (Tabelle 12.4) sind entsprechend gering und die der Messfehlervarianz entsprechend hoch. Das zugehörige Messmodell ist in Abbildung 12.2 dargestellt. Der sich aus den Itemreliabilitäten des kongenerischen Messmodells berechnete Wert für die Gesamttestreliabilität nach McDonalds Omega beträgt für die Steuerungsphase $\omega = .71$.

Statt der Itemleichtigkeiten werden in dem Fall kategorialer Daten, wie sie für die Steuerungsleistungen vorliegen, sogenannte Thresholds angegeben (Tabelle 12.4). Ausgehend von der Annahme, dass die dichotomen Steuerungsleistungen eine kontinuierliche latente Variable messen, geben die Thresholds (Schwellenwerte) die Grenzwerte für die Wertebereiche der kontinuierlichen Variablen an. Damit ist es möglich die einzelnen Indikatorwerte den jeweiligen Abschnitten der kontinuierlichen Variable zuzuordnen. Diese Schwellenwerte sind in dem vorliegenden Fall (zwei Kategorien und keine zusätzlichen Prädiktoren) als Z-Werte zu betrachten, aus denen die prozentuale Lösungshäufigkeit für jedes Item anhand der Wahrscheinlichkeitsverteilung einer Standardnormalverteilung ermittelt werden kann (Urban & Mayerl, 2014, S. 72). Hohe negative Werte stehen für leichte Items, Werte um Null für mittelschwere Items und hohe positive Werte für schwierige Items.

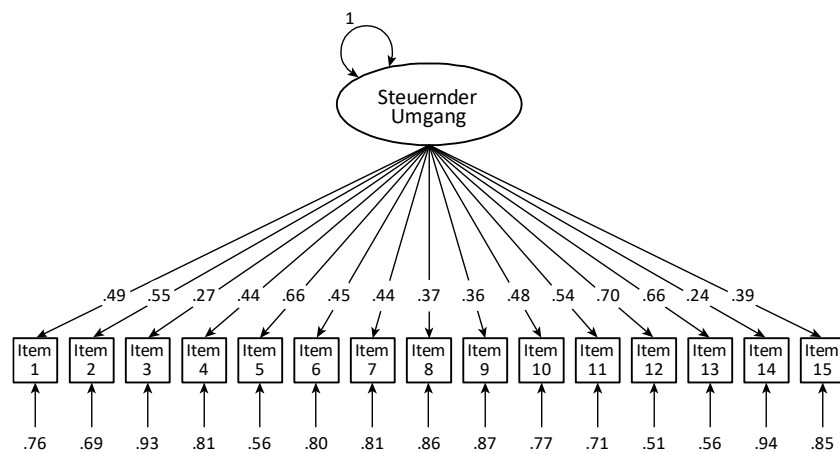


Abbildung 12.2: kongenerisches Messmodell M2 für die Systemsteuerung mit geschätzten Faktorladungen und Messfehlervarianzen ($n = 141$)

13 Schwierigkeitsbestimmende Merkmale

Beantwortung der Hypothesen 1.1 - 1.6

Die Analyse der Konstruktrepräsentation stellt einen Teil der Konstruktvalidierung dar. Hierzu werden die vorab formulierten Annahmen über den Einfluss von System- und Situationsmerkmalen auf die Unterschiede in der Schwierigkeit im Umgang mit technischen Systemen geprüft. Dabei stellt die tatsächlich beobachtete Schwierigkeit die abhängige Variable in dem gewählten Statistikmodell der multiplen linearen Regression dar. Die Varianz in der empirischen Schwierigkeit soll durch die System- und Situationsmerkmale größtmöglich vorhergesagt werden, sie stellen die Prädiktorvariablen in dem Statistikmodell dar. Die Anwendung der Regressionsanalyse bzw. die Interpretation ihrer Ergebnisse ist an bestimmte Voraussetzungen gebunden, deren Erfüllungen vorab zu untersuchen sind. Demzufolge geht das folgende Kapitel zunächst auf die deskriptiven Befunde der empirischen Itemschwierigkeiten und die Prüfung der Bedingungen ein, bevor anschließend die Ergebnisse der Regressionsanalyse berichtet werden.

13.1 Empirische Itemschwierigkeiten

Die Schwierigkeit im Umgang mit technischen Systemen ist in Abhängigkeit von der Situation, in der mit den Systemen interagiert wird, zu analysieren. Die empirischen Itemschwierigkeiten der 15 simulierten Alltagsgeräte werden deshalb getrennt nach Explorations- und Steuerungsphase ausgewertet und stellen in den anschließenden Regressionsanalysen die abhängigen Variablen dar. Die Durchführung der Regressionsanalysen erfolgt ebenfalls getrennt für die einzelnen Interaktionsformen, da angenommen wird, dass in Abhängigkeit des Interaktionsziels jeweils andere Merkmale die Itemschwierigkeit beeinflussen.

13.1.1 Explorationsphase

Die Problemstellung in der Explorationsphase fordert die ausführliche Erkundung der simulierten technischen Systeme. Die Leistung der Problemlöser in dieser Phase wird danach bewertet, wie vollständig sie die Geräte erkunden (Kapitel 8.5.3). Für jeden neuen Zustand, in den die problemlösende Person ein technisches System bringt, erhält sie einen Punkt. Die Summe der erreichten Punkte (und damit die Anzahl der voneinander unabhängigen erreichten Systemzustände) wird anschließend durch die maximal mögliche Anzahl an Zuständen, die das jeweilige System annehmen kann, dividiert. Auf diese Weise ergibt sich für jedes Item und jede Person ein Maß für die Explorationsvollständigkeit, das Werte zwischen 0 und 1 annehmen kann.

$$\text{Explorationsvollständigkeit} = \frac{\text{Anzahl besuchter Systemzustände}}{\text{Anzahl aller Systemzustände}} \quad (13.1)$$

Hat eine Person nicht mit dem System interagiert und konnte das System damit nicht in einen anderen Zustand als den Startzustand bringen, so ist der Wert für die Explorationsvollständigkeit gleich 0. Schafft es eine Person alle in dem System möglichen Zustände zu erkunden, ist der Wert für die Explorationsvollständigkeit gleich 1.

Grundsätzlich ist davon auszugehen, dass sich die Schwierigkeit eines Items in dem Verhalten der Probanden im explorierenden Umgang der technischen Geräte zeigt und dementsprechend auch in dem Maß für die Explorationsvollständigkeit. Je schwieriger ein Item ist, desto geringer wird die durchschnittliche Explorationsvollständigkeit sein, je leichter ein Item, desto vollständiger kann es exploriert werden. Die Itemschwierigkeit in der Phase der Exploration ergibt sich aus der Summe der Explorationsvollständigkeiten über alle Personen, dividiert durch die Anzahl der Personen, die das jeweilige Item bearbeitet haben:

$$\text{Itemschwierigkeit}_{i_{\text{Exploration}}} = \frac{\sum_{i=1}^N \text{Explorationsvollständigkeit}_i}{N} \quad (13.2)$$

Um im weiteren Verlauf untersuchen zu können, welche der angenommen Systemmerkmale den explorierenden Umgang mit den technischen Systemen erschweren, muss die Itemschwierigkeit über die 15 technischen Geräte ausreichend breit variieren. Abbildung 13.1 zeigt die Itemschwierigkeit in der Explorationsphase aller technischen Geräte. Dabei sind die technischen Geräte in der Reihenfolge ihrer empirischen Schwierigkeiten (Explorationsvollständigkeiten) angeordnet und nicht anhand ihrer Position in dem Test.

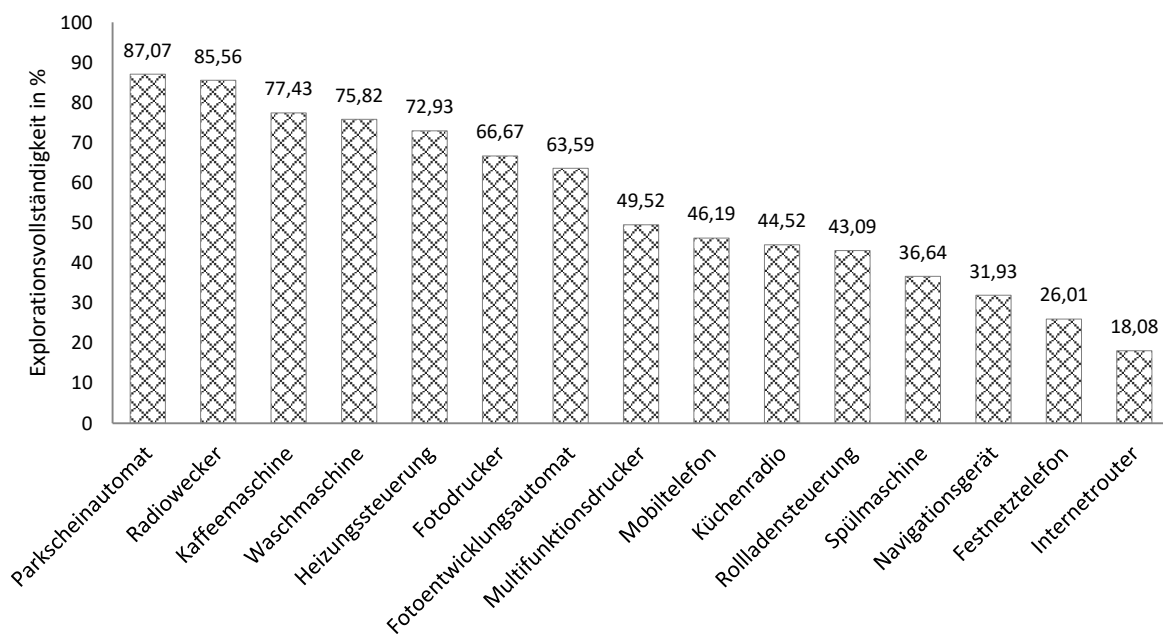


Abbildung 13.1: Itemschwierigkeiten der Exploration technischer Geräte, sortiert nach ihrer empirischen Schwierigkeit (Explorationsvollständigkeit), von links nach rechts werden die Items entsprechend schwieriger zu explorieren ($n = 136$)

Wie Abbildung 13.1 zu entnehmen, wurde keines der Items von allen Probanden vollständig exploriert. Festzustellen ist außerdem, dass kein Item so schwer ist, dass es niemand explorieren konnte. Die Variationsbreite der empirischen Schwierigkeit in der Phase der Exploration wird durch die in dem Test eingesetzten Items weitgehend abgedeckt. Große Lücken in dem Schwierigkeitsspektrum lassen sich nicht erkennen. Die durch die Explorationsschwierigkeiten aller Items operationalisierte mittlere Testschwierigkeit in dieser Phase beträgt .55.

13.1.2 Steuerungsphase

An die Exploration eines jeden technischen Systems schließt direkt die Steuerungsphase an. Während die Testpersonen in der Explorationsphase die am Computer präsentierten technischen Geräte ausprobieren und erkunden sollen, haben sie in der Steuerungsphase konkrete Bedienhandlungen auszuführen, um die in der Aufgabenstellung genannten Handlungsziele zu erreichen. Die Leistung der Probanden in dieser Phase wird in einem ersten Schritt danach bewertet, ob sie die geforderten Ziele erreichen. Wenn eine Testperson alle in einer Aufgabenstellung vorgegebenen Ziele erreicht, erhält sie für das entsprechende Item 1 Punkt, wird keines der Ziele oder lediglich eine Teilmenge der Ziele erreicht, wird kein Punkt vergeben. Es ist davon auszugehen, dass die geforderten Ziele bei leichten Items von vielen Testpersonen erreicht werden, während die Anforderungen bei schweren Items von nur wenigen Personen bewältigt werden. Die empirische Schwierigkeit eines Items ist somit kongruent mit der relativen Lösungshäufigkeit, die sich aus der Anzahl der Richtiglösungen dividiert durch die Anzahl der Probanden, die das entsprechende Item bearbeitet haben, berechnet:

$$\text{Itemschwierigkeit}_{\text{Steuerung},I} = \frac{\text{Anzahl der Probanden, die das Item gelöst haben}}{N} \times 100 \% \quad (13.3)$$

Abbildung 13.2 stellt die im Anschluss an die Testdurchführung so ermittelten Schwierigkeiten für jedes Item dar. Die Items sind entsprechend ihrer Schwierigkeit von leicht (links) nach schwer (rechts) angeordnet und befinden sich ausnahmslos innerhalb des Schwierigkeitsbereiches $> .10$ bis $< .90$. Besonders leichte oder besonders schwierige Items, die sich in den äußersten Randbereichen ($> .90$ und $< .10$) des Schwierigkeitsbereiches befinden würden, gibt es nicht. Das verbleibende Schwierigkeitsspektrum wird von den eingesetzten Items ausreichend gut abgedeckt. Einzig die empirischen Schwierigkeiten der Items *Parkscheinautomat* und *Fotoentwicklungsautomat* haben annähernd gleiche Werte (85,71 % und 85,03 %). Die mittlere Testschwierigkeit, operationalisiert durch die relative Lösungshäufigkeit aller Items, beträgt in dieser Phase .49.

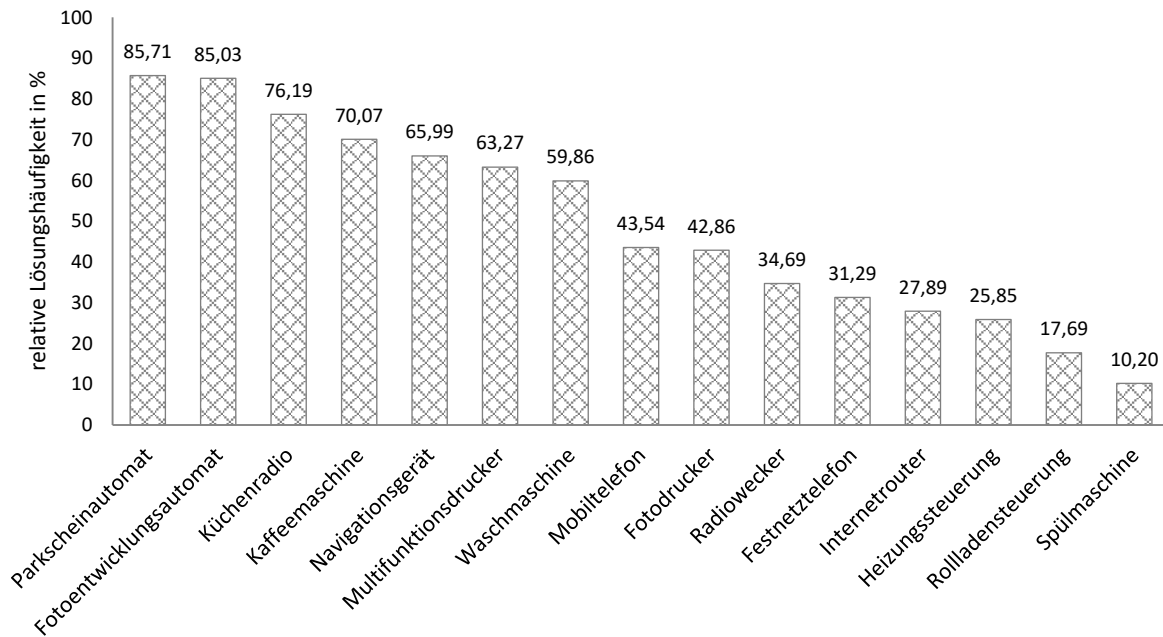


Abbildung 13.2: Itemschwierigkeiten der Steuerung technischer Geräte, berechnet aus der relativen Lösungshäufigkeit, sortiert nach ihrer empirischen Schwierigkeit, von links nach rechts werden die Items entsprechend schwieriger zu steuern ($n = 141$)

Das Testverhalten von Probanden unterscheidet sich in Abhängigkeit zur Itemschwierigkeit nicht nur in der Lösungshäufigkeit, sondern auch darin, wie viel Zeit bzw. wie viele Bearbeitungsschritte die Probanden für das Erreichen der jeweiligen Ziele benötigen. Für die Lösung eines besonders schwierigen Items (das überhaupt von nur wenigen Personen gelöst werden kann), müssen die Probanden weitaus mehr mit dem technischen System interagieren, als für die Lösung überhaupt nötig ist. Diesen Sachverhalt sollte die empirische Itemschwierigkeit berücksichtigen. Die Testleistung in der Steuerungsphase wird demzufolge in einem zweiten Schritt neben dem Erreichen des Steuerungsziels auch danach ausgewertet, wie viele Bearbeitungsschritte über die für die Problemlösung erforderlichen Schritte hinaus nötig waren (Kapitel 8.5.3). Jedes in der Problemstellung geforderte Ziel in der Steuerungsphase kann auf einem eindeutig festgelegten kürzesten Weg erreicht werden. Die Länge dieses kürzesten Weges, angegeben in der Zahl an Bearbeitungsschritten, ist abhängig von der jeweiligen Steuerungsaufgabe. Die Größe der Differenz zwischen der Mindestanzahl von Interaktionsschritten und der tatsächlich benötigten Anzahl wird nach dem folgenden Verfahren in die Bewertung der Steuerungsleistung implementiert:

Tabelle 13.1: Bewertung der Steuerungsleistung

Lösung	Punkte
Falsch	0
Richtig, aber über 100 % mehr Bearbeitungsschritte	1
Richtig, aber über 50 % mehr Bearbeitungsschritte	2
Richtig und maximal 50 % mehr Bearbeitungsschritte	3

Zur Ermittlung der empirischen Itemschwierigkeit sind für jedes Item die erreichten Punkte aller Probanden aufzusummieren und durch die Anzahl der Probanden, die das Item bearbeitet haben zu dividieren. Die aus dieser Berechnung resultierenden Werte liegen in einem Wertebereich von 0 bis 3 Punkten. Eine Division mit 3 und eine Multiplikation mit 100 % ergibt die relative Lösungshäufigkeit in %:

$$\text{Itemschwierigkeit}_{\text{Steuerung,II}} = \frac{\text{Summe der erreichten Punkte}}{N} \times \frac{100\%}{3} \quad (13.4)$$

Abbildung 13.3 zeigt die, die Interaktionszahl berücksichtigenden Itemschwierigkeiten der einzelnen technischen Geräte, sortiert nach absteigender relativer Lösungshäufigkeit (von leicht nach schwer).

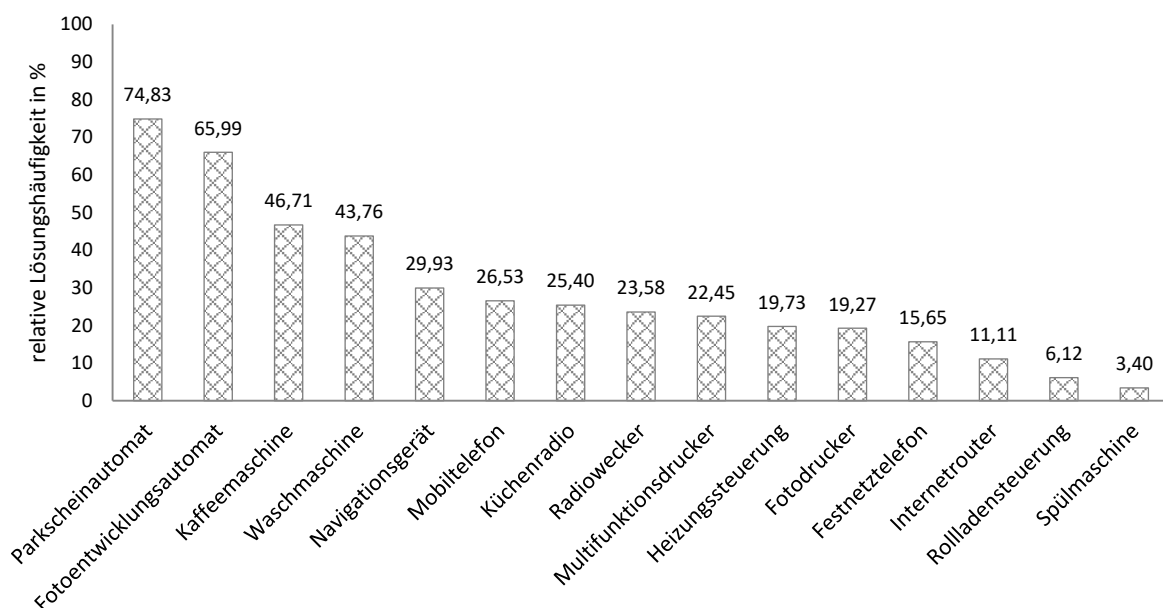


Abbildung 13.3: Itemschwierigkeiten der Steuerung technischer Geräte berücksichtigt die für die Lösung benötigten Interaktionsschritte, sortiert nach ihrer empirischen Schwierigkeit, von links nach rechts werden die Items entsprechend schwieriger zu steuern ($n = 141$)

Die Items *Rollladensteuerung* und *Spülmaschine* weisen relative Lösungshäufigkeiten unter 10 % auf und können daher als besonders schwere Items bezeichnet werden. Im Unterschied zum dichotomen Auswertungsverfahren liegen hier 13 von 15 Items unterhalb des Schwierigkeitsindex von .50, so dass dieser Bereich des Schwierigkeitsspektrums mehr als ausreichend von den Items abgedeckt ist. Oberhalb von .50 liegen hingegen lediglich die zwei Items *Parkscheinautomat* und *Fotoentwicklungsautomat*, dieser Bereich ist also nur unzureichend durch Items vertreten. Die gesamte Testschwierigkeit in der Steuerungsphase beträgt bei Berücksichtigung der für die Lösung benötigten Interaktionsschritte .29.

Der Vergleich der Itemschwierigkeiten, die aus den unterschiedlichen Auswertungsverfahren resultieren, zeigt ein Absinken des Schwierigkeitsindex für alle Items. Der Betrag, um den die relative Lösungshäufigkeit bei Berücksichtigung der Interaktionsanzahl abnimmt, ist jedoch nicht für alle Items gleich groß. Abbildung 13.4 zeigt die einzelnen technischen Systeme mit ihren nach den unterschiedlichen Verfahren berechneten Itemschwierigkeiten.

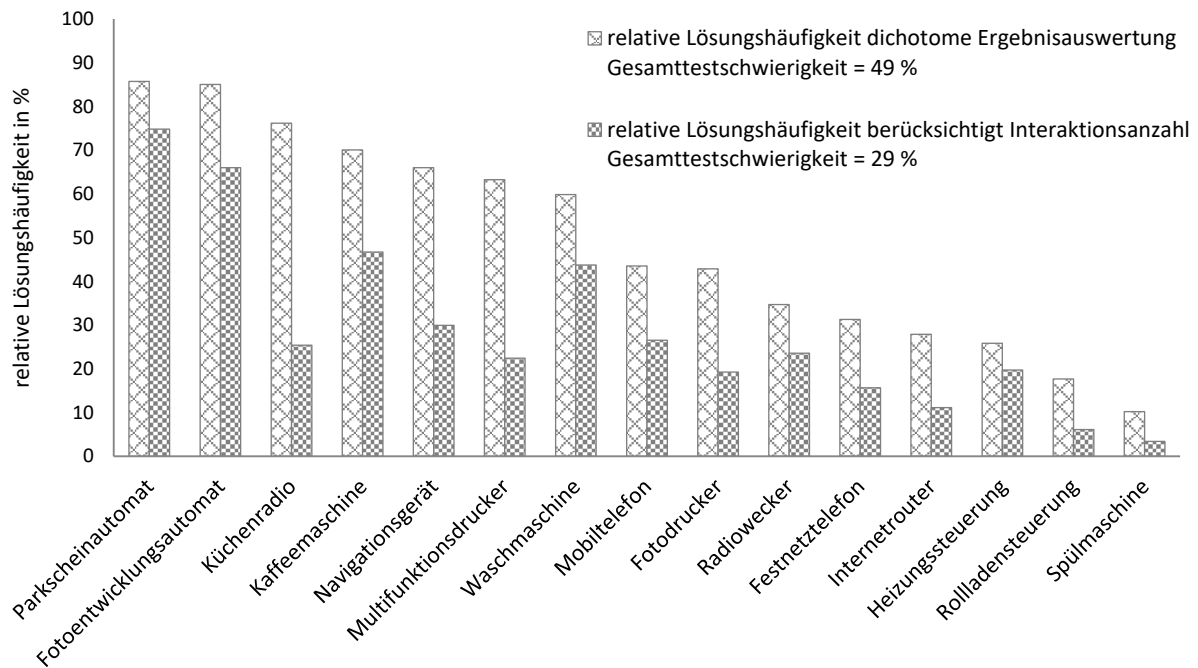


Abbildung 13.4: Vergleich der unterschiedlich berechneten empirischen Itemschwierigkeiten der Systemsteuerung ($n = 141$)

Während die Differenz zwischen den Schwierigkeiten bei dichotomer und interaktionszahlberücksichtigender Leistungsbewertung für einige Items eher gering ist (z. B. *Parkscheinautomat*, *Heizungssteuerung*), fällt diese für andere Items hingegen größer aus (z. B. *Küchenradio*, *Multifunktionsdrucker*). Letztere waren für die Testpersonen zwar zu lösen (hohe relative Lösungshäufigkeit bei dichotomer Ergebnisauswertung), sie konnten die Lösungen jedoch nur mit einem erheblichen Mehraufwand an Bearbeitungsschritten erreichen (geringe Lösungshäufigkeit bei Berücksichtigung der Interaktionsanzahl). Diesen auswertungsbedingten Schwierigkeitsunterschied bei einzelnen Items lässt sich m. E. nur mit dem Vorhandensein von Merkmalen begründen, die einen Einfluss auf die Schwierigkeit bei der Steuerung der Geräte haben. Obwohl die Items das Schwierigkeitsspektrum bei dichotomer Ergebnisauswertung der Steuerungsphase gleichmäßiger abdecken, wird für die Analyse der schwierigkeitsbestimmenden Merkmale die unter Berücksichtigung der Interaktionsschritte ausgewerteten Schwierigkeiten verwendet, mit der Absicht die genannten Differenzen in der Analyse aufklären zu können.

13.2 Prüfung der Voraussetzungen für die multiple lineare Regression

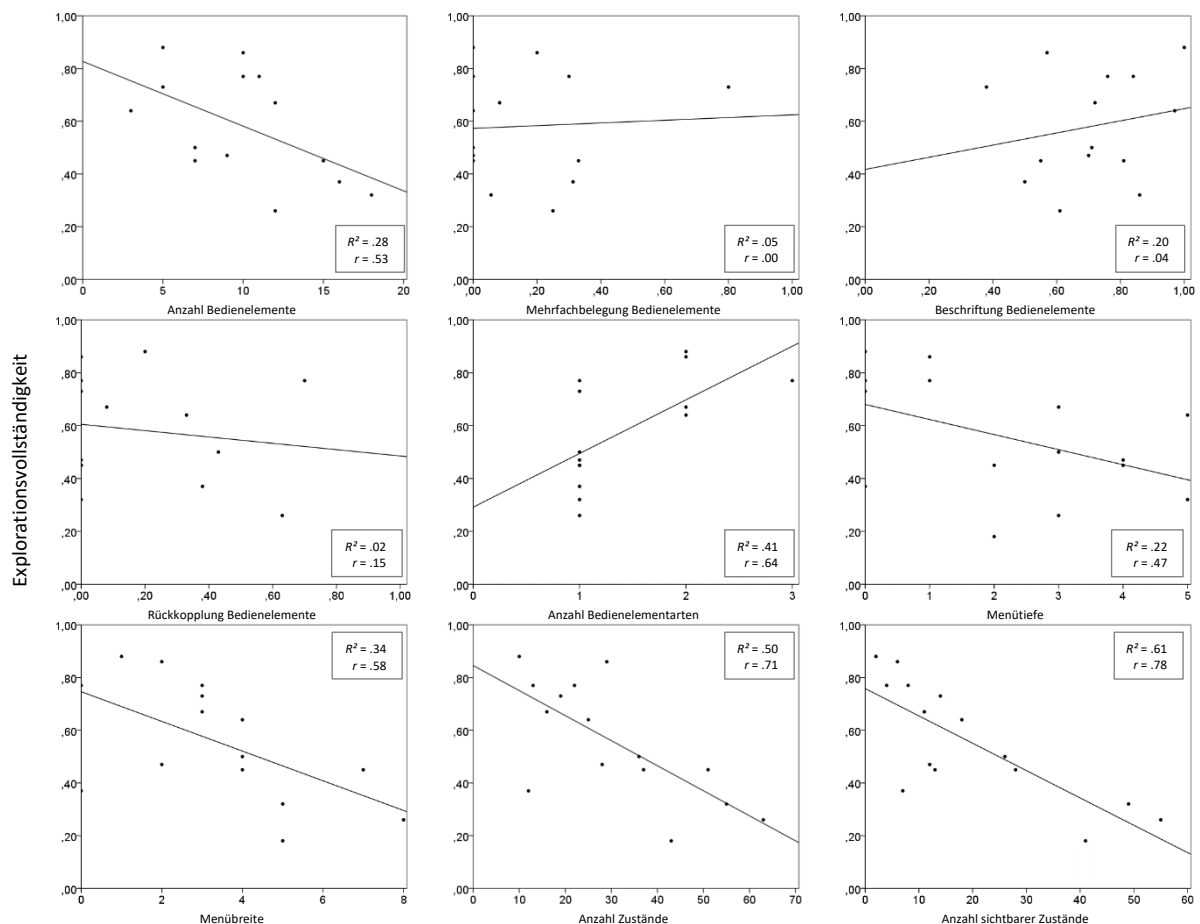
Die Verwendung des statistischen Modells der linearen Regression ist an bestimmte Voraussetzungen gebunden, deren Erfüllungen vor dem Einsatz zu prüfen sind. In dem folgenden Kapitel werden die Ergebnisse dieser Überprüfung vorgestellt. Zunächst beginnt der Abschnitt mit der Untersuchung der linearen Zusammenhänge zwischen den unabhängigen Variablen und der empirischen Itemschwierigkeit und setzt dann mit der Kontrolle, ob zwischen den unabhängigen Variablen eventuell eine lineare Abhängigkeit vorliegt, fort. Abschließend werden die Ergebnisse der Residualanalysen vorgestellt.

13.2.1 Explorationsphase

Linearität

Der Zusammenhang zwischen den als schwierig angenommenen System- und Situationsmerkmalen und der tatsächlich beobachteten Schwierigkeit wird als eine lineare Beziehung modelliert. Um zu prüfen, ob diese Annahme überhaupt begründet ist, wird der Zusammenhang zwischen jedem Itemmerkmal und der Schwierigkeit mithilfe eines Streudiagramms dargestellt, in das außerdem die lineare Regressionsgerade eingezeichnet ist. Wie gut die eingezeichnete Gerade den beobachteten Daten entspricht, lässt sich außerdem durch den Determinationskoeffizienten R^2 angeben, der in dem Fall einer einzigen unabhängigen Variable dem quadrierten Korrelationskoeffizient r entspricht.

Die Systemmerkmale, die keinen Einfluss auf die Schwierigkeit im explorierenden Umgang mit technischen Geräten haben, lassen sich leicht anhand der in Abbildung 13.5 gezeigten Streudiagramme bzw. ihrer Regressionsgeraden identifizieren und werden von weiteren Analysen ausgeschlossen. Hierzu zählen die *Mehrfachbelegung von Bedienelementen*, die *Rückkopplung* von betätigten Bedienelementen, die *Reversibilität* vorgenommener Eingaben, die *Zeitverzögerungen* in der Systemreaktion sowie die *Eigendynamiken* eines Systems. Für Merkmale, in deren Streudiagrammen der Verlauf der eingezeichneten Gerade vor allem durch Ausreißer bestimmt wird, wird der lineare Zusammenhang zur Explorationsschwierigkeit ebenfalls in Frage gestellt. Hiervon sind die Merkmale *Menüvernetzung* und *Rückkopplungen durch das Menü* betroffen.



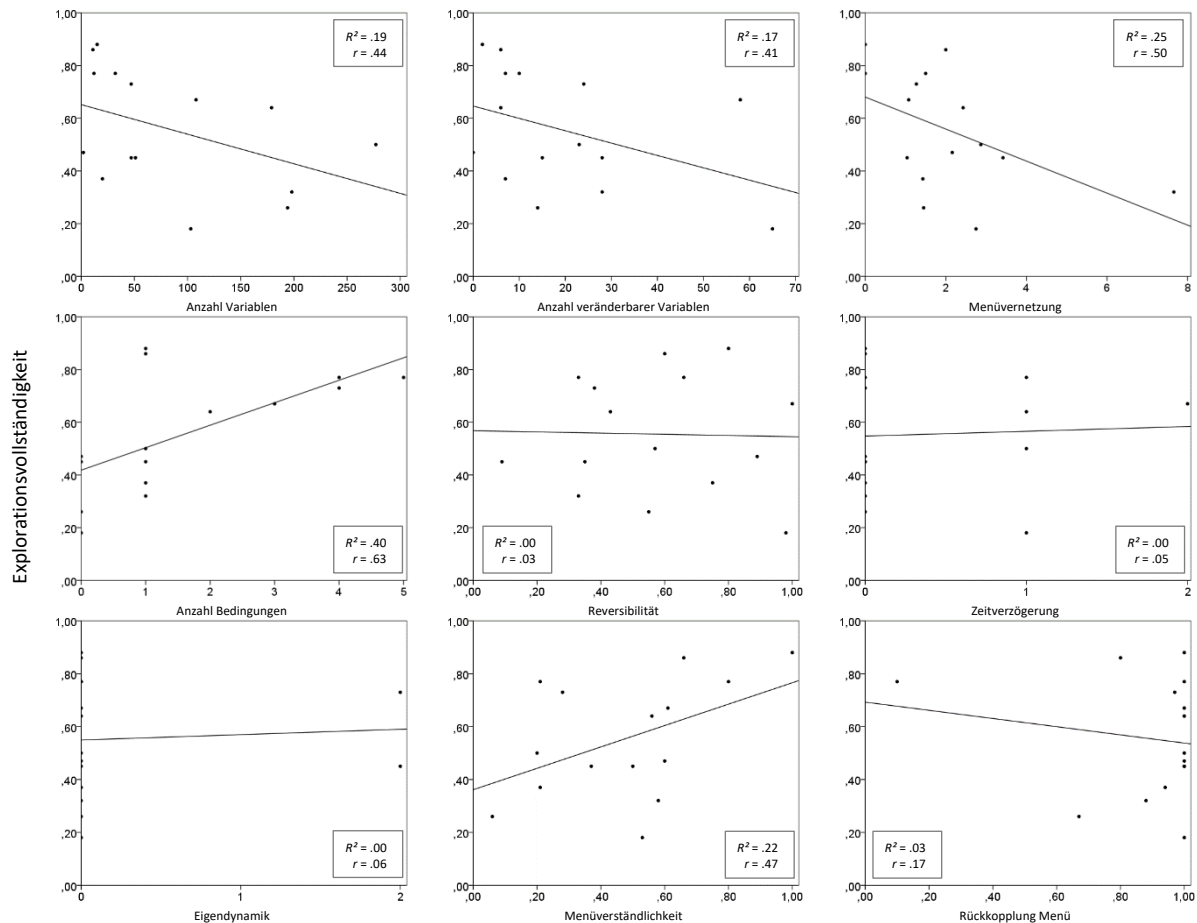


Abbildung 13.5: Streudiagramme der als schwierigkeitsbestimmend angenommenen Systemmerkmale auf die empirische Schwierigkeit in der Systemexploration ($n = 15$)

Neben den Merkmalen auf Ebene der einzelnen technischen Systeme, wird für die Bekanntheit der Geräte aus dem Alltag der Probanden ebenfalls ein Einfluss auf den explorierenden Umgang mit ihnen erwartet. Je bekannter ein technisches Gerät aus dem Alltag ist, desto anschlussfähiger ist es in Bezug zu dem Vorwissen der Probanden. Es ist davon auszugehen, dass eine geringe Vorwissenskompatibilität dazu führt, dass Geräte ausführlicher exploriert werden, da Wissensdefizite ausgeglichen werden müssen. Um die Bedeutung des als linear angenommenen Zusammenhangs in der Regressionsanalyse zu untersuchen, muss die Linearität für dieses inhaltliche Systemmerkmal ebenfalls vorab geprüft werden. Ihr Vorwissen in Bezug auf die einzelnen technischen Systeme haben die Probanden anhand von drei- bis vierstufigen Items selbst eingeschätzt. Die relativen Häufigkeiten in der Nennung dieser Vorwissenskategorie bilden ein Maß für die Vorwissenskompatibilität. Ein Item, bei dem der größte Teil der Personen angab, noch keine bzw. nur wenige Erfahrungen im Umgang mit dem genannten Gerät zu verfügen, ist im Alltag dieser Probanden nur wenig bekannt und demnach nur wenig mit dem Vorwissen der Stichprobe kompatibel. Bei keinem der in dem technischen Problemlösetest simulierten technischen Geräte gab die Mehrheit der der Probanden an, viel Erfahrung zu haben, so dass das Merkmal *Vorwissenskompatibilität* lediglich die zwei Ausprägungen *gering* und *mittel* aufweist. Der lineare Zusammenhang lässt sich daher nicht mit einem Streudiagramm zeigen, sondern mit einem Boxplot, das in Abbildung 13.6 zu sehen ist.

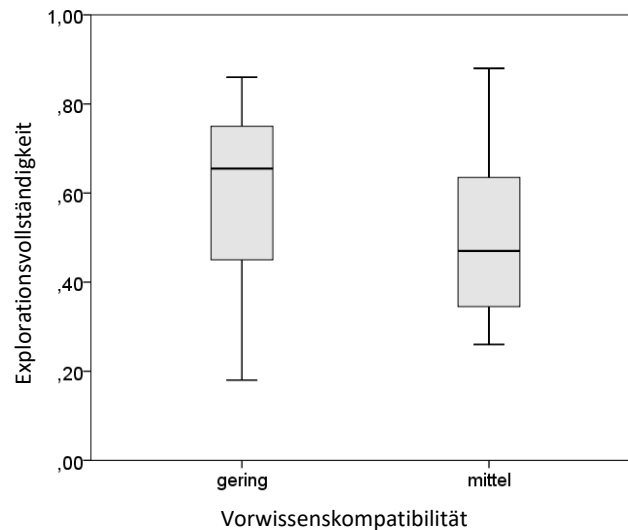


Abbildung 13.6: Boxplot der als schwierigkeitsbestimmend angenommenen Vorwissenskompatibilität auf die Explorationsvollständigkeit ($n = 15$)

Für technische Geräte, mit denen Personen im Alltag weniger Erfahrungen sammeln (geringe Vorwissenskompatibilität), liegt die Box, in deren Bereich sich 50 % der Daten befinden, etwas höher als für die Items mit mittlerer Vorwissenskompatibilität, diese Items werden von den Probanden also vollständiger exploriert. Diese Information lässt sich auch aus der Lage des Medians entnehmen, der die gesamte Datenmenge in zwei Hälften teilt, in denen jeweils 50 % der Daten liegen. Für die Annahme, dass es einen linearen Zusammenhang zwischen der Vorwissenskompatibilität und dem explorierenden Umgang mit den präsentierten technischen Systemen gibt, lassen sich keine Hinweise für deren Widerlegung finden. Die Voraussetzung für die Aufnahme des Merkmals in die Regressionsanalyse ist demnach erfüllt.

Multikollinearität

Die nach der Linearitätsanalyse verbleibenden Systemmerkmale werden im nächsten Schritt auf Multikollinearität untersucht. Variablen, die linear abhängig voneinander sind, weisen hohe Korrelationskoeffizienten auf, so dass der Blick in einem ersten Schritt auf die Korrelationsmatrix der Systemmerkmale zu richten ist (Tabelle 13.2). Bei Betrachtung der Interkorrelationen der Merkmale fällt auf, dass vor allem auf Softwareebene zwischen den Merkmalen der Komplexität ein Zusammenhang zu beobachten ist. Zwischen den Merkmalen *Anzahl der Zustände* in einem System und *Anzahl der sichtbaren Zustände* in einem System ist dieser Zusammenhang am höchsten; der Korrelationskoeffizient nach Pearson beträgt $r = .90$. Aus theoretischer Sicht überrascht dieses Ergebnis nicht, da die Anzahl der sichtbaren Zustände in einem technischen System eine Teilmenge aller vorhandenen Zustände in dem System ist. Aus Gründen der offensichtlich vorliegenden Abhängigkeit, wird von den beiden Merkmalen die gesamte *Anzahl der Zustände* nicht weiter betrachtet.

Tabelle 13.2: Interkorrelation schwierigkeitsbestimmender Systemmerkmale in der Phase der Exploration nach Entfernung der Variablen, die keinen (linearen) Einfluss auf die Explorationsvollständigkeit haben ($n = 15$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Hardware										
Komplexität										
(1) Anzahl Bedienelemente	—									
Transparenz										
(2) Anzahl Bedienelementarten	-.22	—								
Software										
Komplexität										
(3) Menütiefe	.02	-.27	—							
(4) Menübreite	.17	-.50	.53*	—						
(5) Anzahl Zustände	.28	-.39	.60*	.77**	—					
(6) Anzahl sichtbarer Zustände	.28	-.41	.57*	.73**	.90**	—				
(7) Anzahl Variablen	.01	-.18	.59*	.56*	.55*	.69**	—			
(8) Anzahl veränderbarer Variablen	.31	-.05	.14	.38	.22	.38	.31	—		
Vernetztheit										
(9) Anzahl Bedingungen	-.22	.34	-.38	-.41	-.57*	-.46	-.22	-.07	—	
Transparenz										
(10) Menüverständlichkeit	-.18	.18	-.01	-.21	-.40	-.42	-.38	-.06	.08	—

Anmerkungen: $n =$ Itemanzahl; * $p < .05$; ** $p < .01$

Weiterhin korrelieren die Merkmale *Menütiefe*, *Menübreite*, *Anzahl der Variablen* und *Anzahl der sichtbaren Zustände* in der Kategorie Software/Komplexität miteinander. Zwar weisen die Korrelationskoeffizienten nur Werte mittlerer Höhe auf ($.50 < r < .90$), die Beziehung dieser Merkmale wird dennoch genauer untersucht. Die Durchführung partieller Korrelationen, bei denen jeweils eine der genannten Merkmalsvariablen konstant gehalten wird, zeigt, dass das Merkmal *Anzahl der sichtbaren Zustände* für das Zustandekommen der Zusammenhänge zwischen den anderen Variablen verantwortlich ist (Tabelle 13.3). Dieser Sachverhalt spricht dafür, das Merkmal ebenfalls aus den weiteren Analysen auszuschließen.

Tabelle 13.3: Partielle Korrelation der schwierigkeitsbestimmenden Merkmale der Komplexität auf Softwareebene unter Konstanthaltung der Variable Anzahl sichtbarer Zustände ($n = 15$)

Kontrollvariable	Software	(1)	(2)	(3)
Anzahl sichtbarer Zustände	Komplexität			
	(1) Menütiefe	—		
	(2) Menübreite	.20	—	
	(3) Anzahl Variablen	.34	.10	—

Residualanalyse

Eine weitere Voraussetzung für die eigentliche Regressionsanalyse beziehungsweise für die Bestimmung der Konfidenzintervalle der Regressionskoeffizienten ist eine Normalverteilung sowie eine homogene Streuung der Residuen in Abhängigkeit der beobachteten Itemschwierigkeit. Die Streuung der Residuen lässt sich grafisch anhand von Streudiagrammen beurteilen. Hierfür stehen zwei alternative Streudiagramme zur

Verfügung. Das erste mögliche Diagramm gibt die Werte der Ausprägung des jeweiligen schwierigkeitsbestimmenden Merkmales und die Residuen der abhängigen Variable (die Itemschwierigkeit) wieder. Eine zweite Möglichkeit enthält die vorhergesagten (geschätzten) Werte für die abhängige Variable und die Residuen der abhängigen Variable. In beiden Diagrammen werden die standardisierten Residuen auf der Ordinatenachse abgetragen.

Die Überprüfung der Normalverteilung und Streuungsungleichheit (Heteroskedastizität) erfolgte lediglich für die nach der Überprüfung der Linearität und Multikollinearität noch verbleibenden Merkmale, die in die Regressionsanalyse hineingenommen werden. Abbildung 13.7 zeigt für die genannte erste Möglichkeit die entsprechenden Residualplots.

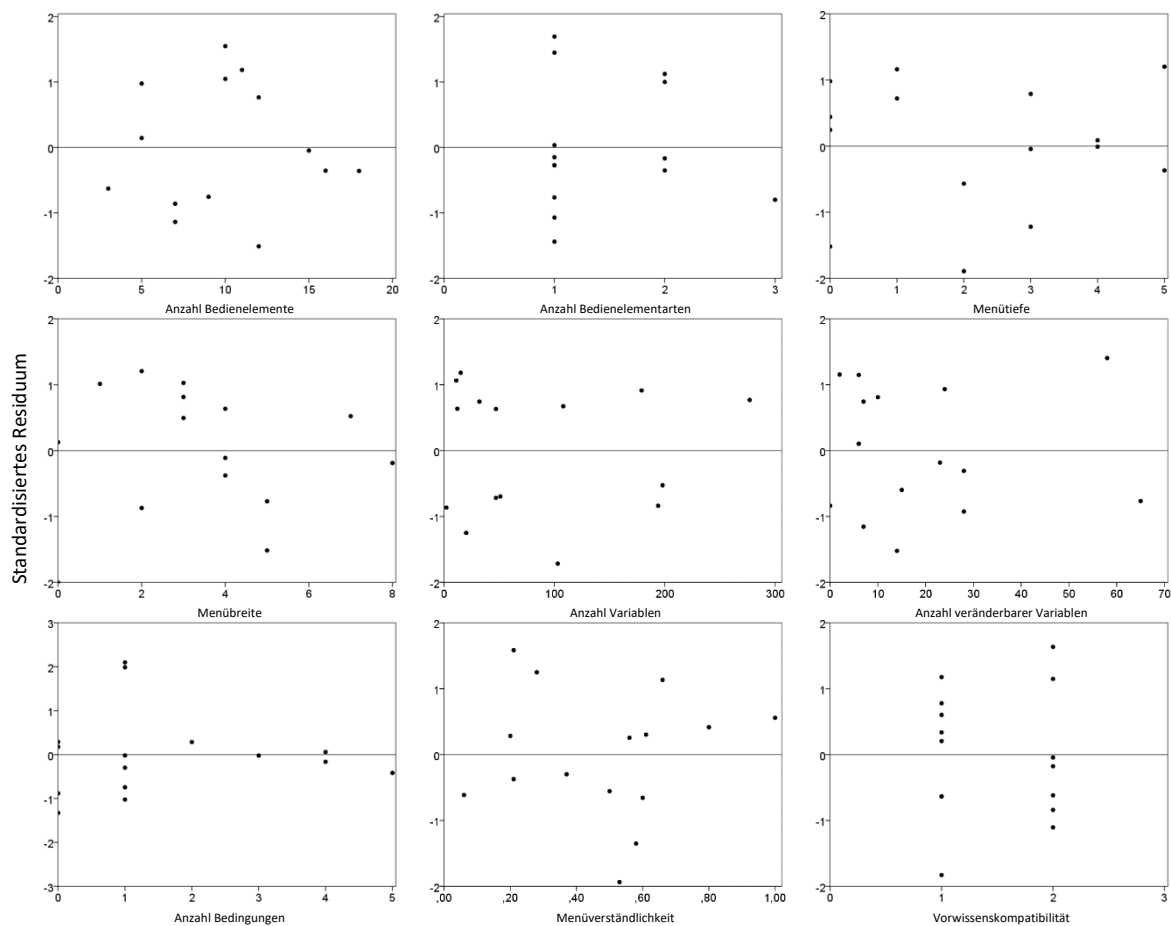


Abbildung 13.7: Streudiagramme zur Beurteilung von Normalverteilung und Heteroskedastizität der Residuen für die verbleibenden als schwierigkeitsbestimmend angenommenen Merkmale der Systemexploration ($n = 15$)

Normalverteilung der Residuen ist gegeben, wenn die Werte gleichmäßig oberhalb und unterhalb der Nulllinie verteilt sind. Für alle in Abbildung 13.7 zu sehenden Diagramme lässt sich keine systematische Verteilung um die Nulllinie ausmachen, das heißt die Werte streuen zufällig und es kann von Normalität ausgegangen werden. Eine vorliegende Heteroskedastizität ließe sich in den Streudiagrammen an einer trichterförmigen Streuung der Werte erkennen. In einem solchen Fall würde der Streubereich der standardisierten Residuen mit stärker werdender Merkmalsausprägung immer größer oder kleiner werden. Für das Systemmerkmal *Anzahl der Bedingungen* kann das Vorliegen von Heteroskedastizität nicht ausgeschlossen werden. Mit zunehmender Anzahl an vorhandenen Variablenbedingungen

auf Softwareebene der technischen Systeme streuen die Residuen immer weniger. Abgesehen von einer möglichen Heteroskedastizität für dieses Merkmal können die Ergebnisse der Regressionsanalyse für die Explorationsphase nach der Überprüfung von Linearität, Multikollinearität und der Residualanalyse als vermutlich auf einem linearen Modell basierend bewertet und interpretiert werden.

13.2.2 Steuerungsphase

In der Steuerungsphase wird der Umgang mit den technischen Systemen neben den bereits in der Explorationsphase berücksichtigten Systemmerkmalen außerdem durch konkrete Problemstellungen beeinflusst. Die in der Steuerungsphase vorgegebenen Handlungsziele unterscheiden sich darin, wie viele Bearbeitungsschritte nötig sind, um die konkrete Problemstellung zu lösen. Die *Anzahl der Lösungsschritte* stellt damit ein potenziell schwierigkeitsbestimmendes Merkmal der Problemsituation dar. Die Bearbeitungsschritte zur Steuerung eines technischen Systems unterscheiden sich in einigen Fällen kaum oder gar nicht voneinander. So benötigt eine Person, will sie die Weckzeit eines Radioweckers von 00:00 Uhr auf 07:00 Uhr ändern, insgesamt acht Schritte, von denen sieben aber identische Bedienhandlungen sind (einen Knopf immer wieder drücken, damit der Wert für die Stunde um je eine Stunde nach oben verändert wird). Die *Anzahl tatsächlich unterschiedlicher Lösungsschritte* ist demnach ebenfalls ein schwierigkeitsbestimmendes Merkmal der Problemsituation. Weitere Merkmale, die die Schwierigkeit bei der Erreichung der vorgegebenen Handlungsziele erhöhen, sind die *Anzahl der hierfür einzustellenden Variablen* sowie die dabei zu berücksichtigenden *mehrfachbelegten Bedienelemente* und die in der Software vorhandenen *Variablenbedingungen*. Die auf Systemebene als schwierigkeitserzeugend angenommenen Merkmale wurden bereits in der Explorationsphase berücksichtigt und ihr Einfluss auf die Schwierigkeit im steuernden Umgang mit den Geräten wird nicht mehr für jedes Merkmal einzeln untersucht. Vielmehr werden diese Merkmale in der Variable *Explorationsvollständigkeit* zusammengefasst und in die Regressionsgleichung aufgenommen. Berücksichtigt werden muss zudem, ob Probanden in der Explorationsphase diejenigen Zustände erkundet haben, die für die Lösung der Problemstellung in der Steuerungsphase auch tatsächlich relevant sind (*Steuerungsspezifische Exploration*). Denn es ist davon auszugehen, dass technische Geräte leichter zu bedienen sind, wenn über die konkrete Bedienhandlung bereits Informationen generiert wurden.

Linearität

Die Überprüfung der Annahmen über den Unterschied im steuernden Umgang mit technischen Geräten aufgrund unterschiedlicher Situationsmerkmale, erfolgt analog zur Explorationsphase mithilfe einer linearen Regression, so dass auch an dieser Stelle die Erfüllung der Voraussetzungen geprüft werden muss. Die Linearität wird dabei grafisch anhand von Streudiagrammen mit eingezeichneter Regressionsgerade untersucht, in denen die unabhängigen Variablen auf der x-Achse und die abhängige Variable der Steuerungsschwierigkeit auf der y-Achse abgetragen sind (Abbildung 13.8). Für das Merkmal *Anzahl der Lösungsschritte* kann kein linearer Zusammenhang zur Schwierigkeit in der Steuerungsphase festgestellt werden. Diese Variable wird deshalb von der Regressionsanalyse ausgeschlossen. Der Zusammenhang von weiteren unabhängigen Variablen (*Anzahl unterschiedlicher Lösungsschritte*, *Anzahl einzustellender Variablen* und zu berücksichtigende

Bedingungen) ist ebenfalls in Frage zu stellen, bis auf das Merkmal *zu berücksichtigende Bedingungen* werden sie aber zunächst nicht aus dem Modell genommen, um ihre Bedeutung in Anwesenheit der anderen Merkmale auf die Itemschwierigkeit zu prüfen. Das Merkmal *zu berücksichtigende Bedingungen* zeigt in dem Streudiagramm einen positiven Zusammenhang zur Steuerungsleistung. Dieser Befund widerspricht theoretischen Annahmen und da er im Rahmen dieser Arbeit nicht sachlogisch begründet werden kann, wird das Merkmal von der Regressionsanalyse ausgeschlossen.

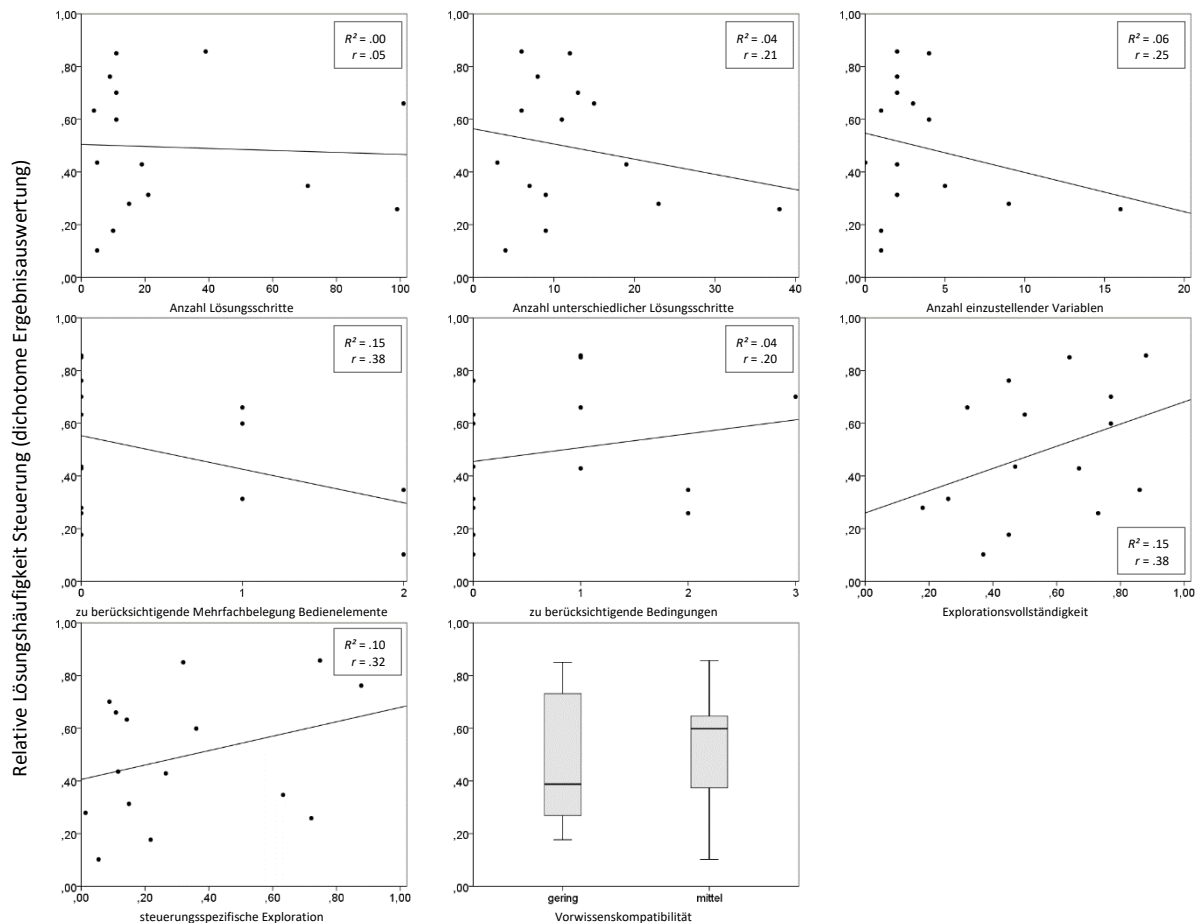


Abbildung 13.8: Streudiagramme und Boxplot der als schwierigkeitsbestimmend angenommenen Merkmale auf die (dichotom ausgewertete) empirische Schwierigkeit der Systemsteuerung ($n = 15$)

Multikollinearität

Für die nach der Linearitätsprüfung übriggebliebenen Problemmerkmale muss darüber hinaus noch gezeigt werden, dass keine Multikollinearität vorliegt. Hierzu werden die Korrelationskoeffizienten in Tabelle 13.4 betrachtet. Es zeigt sich ein hoher Zusammenhang zwischen dem Merkmal *Anzahl unterschiedlicher Lösungsschritte* und *Anzahl einzustellender Variablen* ($r = .89$). Interpretieren lässt sich diese Beziehung mit der Tatsache, dass eine größer werdende Anzahl von Variablen, die für das Erreichen des Handlungszieles verändert werden müssen, auch zu einer Erhöhung der Anzahl an Lösungsschritten führt. Die ersichtliche lineare Abhängigkeit des Merkmals *Anzahl unterschiedlicher Lösungsschritte* wird durch die Entfernung des Merkmals *Anzahl einzustellender Variablen* beseitigt. Zwischen den anderen Merkmalen der Situation lässt sich anhand der Korrelationskoeffizienten keine Multikollinearität ausmachen.

Tabelle 13.4: Interkorrelation schwierigkeitsbestimmender Merkmale der Systemsteuerung ($n = 15$)

	(1)	(2)	(3)
Situationsmerkmale			
Komplexität			
(1) Anzahl untersch. Lösungsschritte	—		
(2) Anzahl einzustellender Variablen	.89**	—	
Vernetztheit			
(3) zu berücksichtigende Mehrfachbelegung Bedienelemente	-.30	-.10	—

Anmerkungen: $n =$ Itemanzahl; ** $p < .01$

Residualanalyse

Die Überprüfungen der Normalverteilung und der Streuung der Residuen in Abhängigkeit der Merkmalsausprägungen erfolgen für das Komplexitätsmerkmal *Anzahl unterschiedlicher Lösungsschritte* sowie für die Vernetztheitsvariablen *zu berücksichtigende Mehrfachbelegung von Bedienelementen* und *zu berücksichtigende Variablenbedingungen*, außerdem für die *Explorationsvollständigkeit*, die *steuerungsspezifische Exploration* und die *Vorwissenskompatibilität* der technischen Geräte.

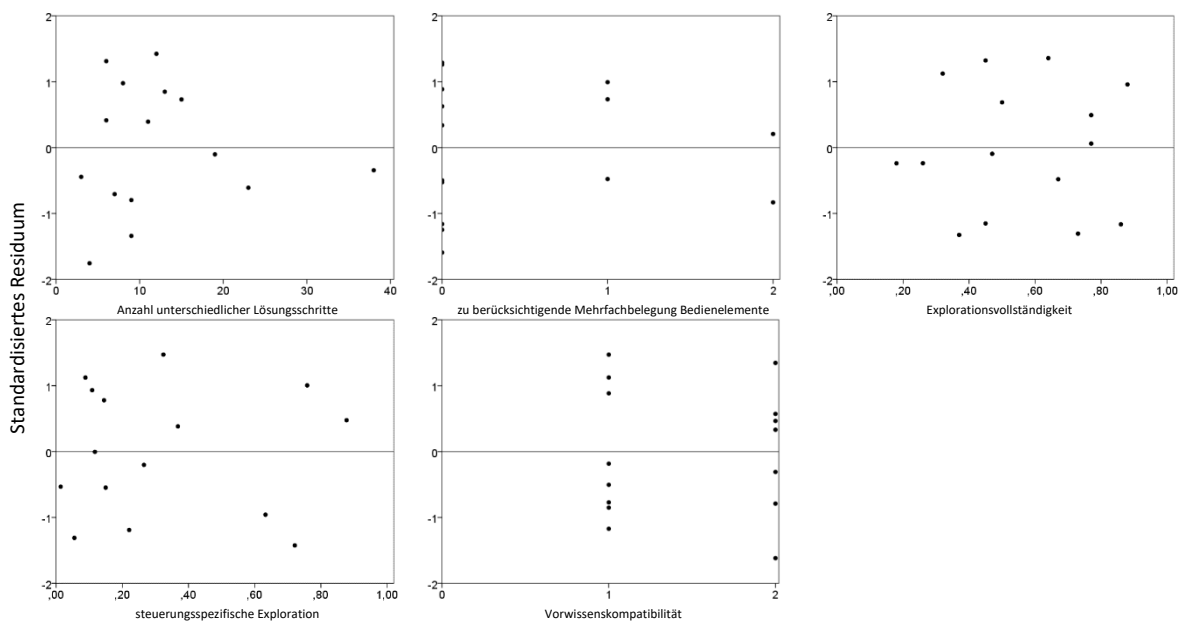


Abbildung 13.9: Residualplots zur Beurteilung von Normalverteilung und Heteroskedastizität der Residuen für die verbleibenden als schwierigkeitsbestimmend angenommenen Merkmale der Systemsteuerung (dichotome Ergebnisauswertung) ($n = 15$)

In den in Abbildung 13.9 zu sehenden Residualplots verteilen sich die Werte für die Residuen annähernd gleichmäßig ober- und unterhalb der eingezeichneten Nulllinie (Mittelwert der Residuen). Eine Verletzung der Normalverteilung kann nicht konstatiert werden. Eine systematische Streubreite der Residuen in Abhängigkeit der Ausprägung des jeweils zu untersuchenden Merkmals kann für das Merkmal *Anzahl unterschiedlicher Lösungsschritte* nicht völlig ausgeschlossen werden. Die Datenpunkte, die sich rechts von dem Wert 18 Lösungsschritte befinden, streuen weit weniger als die Datenpunkte links von diesem Wert, jedoch handelt es sich auch nur um drei Punkte, die höhere Werte für die benötigte Anzahl unterschiedlicher Lösungsschritte aufweisen.

13.3 Regressionsergebnisse und Modellgüte

13.3.1 Explorationsphase

Nach der Überprüfung der Voraussetzungen für die Durchführung einer linearen Regression verbleiben die als schwierigkeitsbestimmend angenommenen formalen Systemmerkmale: *Anzahl der Bedienelementarten, Menüverständlichkeit, Anzahl der Bedingungen, Anzahl der Bedienelemente, Menütiefe, Anzahl sichtbarer Zustände* sowie *Anzahl der veränderbaren Variablen*. Als inhaltliches Merkmal wird außerdem die *Vorwissenskompatibilität* in die Regressionsanalyse genommen. Die Aufnahme der genannten Merkmale in die Regressionsgleichung erfolgte blockweise, im ersten Schritt sind die Variablen des Merkmals *Transparenz*, im zweiten Schritt ist die Variable des Merkmals *Vernetztheit*, im dritten Schritt sind die Variablen der *Komplexität* und im letzten Schritt ist die inhaltliche Variable *Vorwissenskompatibilität* aufgenommen worden. Jede Hinzunahme eines Blockes stellt ein Regressionsmodell dar, für das entsprechende Gütemaße angegeben und die miteinander verglichen werden können. Tabelle 13.5 zeigt eine Gegenüberstellung der vier Modelle mit den darin enthaltenen Prädiktorvariablen und ihren entsprechenden Modellparametern. Der Wert für den Varianz-Inflationsfaktor (VIF) als Maß der Kollinearitätsstatistik liegt bei allen Prädiktorvariablen unter 3.30 (der Mittelwert über alle VIF-Werte ist 1.95). Es kann also davon ausgegangen werden, dass die in die Regressionsgleichung aufgenommenen Variablen voneinander linear unabhängig sind.

Beide Prädiktorvariablen für das Merkmal *Transparenz* im Modell 1 haben einen positiven signifikanten Einfluss auf die Explorationsvollständigkeit. Sowohl die *Anzahl unterschiedlicher Bedienelementarten* auf Hardwareebene ($\beta = .56$) als auch die *Verständlichkeit der Menüs* auf Softwareebene ($\beta = .45$) erleichtern demnach den explorierenden Umgang mit den im Test simulierten technischen Geräten. Selbst bei Hinzunahme der Variable *Anzahl der Variablenbedingungen* der Software als Merkmal der *Vernetztheit* verlieren die Variablen der *Transparenz* kaum an Bedeutung. Die in dem Modell 2 hinzugekommene Variable *Anzahl der Variablenbedingungen* hat einen gleichermaßen bedeutenden und signifikanten Einfluss ($\beta = .40$) auf die Itemschwierigkeit in der Explorationsphase, jedoch nicht wie anfangs vermutet einen negativen, sondern einen positiven Einfluss. Kommen im Modell 3 jedoch die Variablen des Merkmals *Komplexität* hinzu, verliert die Variable des Merkmals *Vernetztheit* deutlich an Einfluss (von $\beta = .40$ auf $\beta = .19$). Aus dem Block der Komplexitätsmerkmale haben die Variablen *Anzahl der Bedienelemente* der Hardware ($\beta = -.35$) und die *Menütiefe* der Software ($\beta = -.35$) einen größeren Einfluss auf die Explorationsvollständigkeit. Neben den Regressionskoeffizienten der *Transparenz*-variablen wird in diesem Modell aber lediglich der Regressionskoeffizient der Variable *Anzahl der Bedienelemente* signifikant ($p < .05$). Die Bedeutung der im letzten Modell 4 dazugekommenen Variable *Vorwissenskompatibilität* auf die Itemschwierigkeit in der Explorationsphase kann als eher gering eingeschätzt werden ($\beta = .12$).

Tabelle 13.5: Regressionskoeffizienten und globale Modellgütemaße für die blockweise Analyse des Einflusses von Merkmalen auf die Itemschwierigkeit der Systemexploration ($n = 15$)

	Modell 1 Transparenz			Modell 2 + Vernetztheit			Modell 3 + Komplexität			Modell 4 + Vorwissenskompatibilität		
	$B(SE)$	β	t	$B(SE)$	β	t	$B(SE)$	β	t	$B(SE)$	β	t
(Konstante)	.16(.11)		1.52n.s.	.14(.09)		1.59n.s.	.50(.10)		4.79**	.47(.11)		4.29**
Anzahl Bedienelementarten	.18(.06)	.56	2.93*	.13(.05)	.43	2.51*	.10(.04)	.31	2.87*	.09(.04)	.27	2.17n.s.
Menüverständlichkeit	.35(.15)	.45	2.37*	.34(.12)	.44	2.75*	.31(.09)	.40	3.29*	.34(.10)	.44	3.31*
Anzahl Bedingungen				.05(.02)	.40	2.41*	.02(.02)	.19	1.58n.s.	.02(.02)	.19	1.56n.s.
Anzahl Bedienelemente							-.02(.01)	-.35	-3.07*	-.02(.01)	-.32	-2.67*
Menütiefe							-.04(.02)	-.35	-2.40n.s.	-.04(.02)	-.38	-2.48n.s.
Anzahl sichtbarer Zustände							.00(.00)	-.06	-.35n.s.	.00(.00)	-.01	-.03n.s.
Anzahl veränderbarer Variablen							.00(.00)	.08	.67n.s.	.00(.00)	.03	.21n.s.
geringe Vorwissenskompatibilität										.05(.06)	.12	.85n.s.
	$R^2 = .61; F = 8.65^{**}; R_{korr}^2 = .54$			$R^2 = .75; F = 10.22^{**}; R_{korr}^2 = .68$			$R^2 = .94; F = 14.24^{**}; R_{korr}^2 = .88$			$R^2 = .95; F = 11.99^{**}; R_{korr}^2 = .87$		

Anmerkungen: n = Itemanzahl; B = unstandardisierter Regressionskoeffizient; SE = Standardfehler; β = standardisierter Regressionskoeffizient; t = t -Wert des unstandardisierten Regressionskoeffizienten; R^2 = Determinationskoeffizient; F = F -Wert; R_{korr}^2 = korrigierter Determinationskoeffizient; Signifikanzniveau: * $p < .05$; ** $p < .01$; n.s = nicht signifikant

Für die Klärung, welches der vier Modelle die beobachteten Daten besser erklären bzw. vorhersagen kann, werden die globalen Modellgütemaße der einzelnen Modelle miteinander verglichen. In dem ersten Modell beträgt der Determinationskoeffizient $R^2 = .61$, das heißt, dass mit den Variablen der Transparenz 61 % der Varianz der Explorationsschwierigkeit aufgeklärt werden kann. Je mehr Variablen in die Regressionsgleichung aufgenommen werden, desto größer wird der Wert für R^2 . Im Modell 4 beträgt die Varianzaufklärung 95 % ($R^2 = .95$). Aufgrund der sehr geringen Anzahl an Items ($n = 15$), ist an dieser Stelle außerdem das korrigierte Bestimmtheitsmaß R_{korr}^2 , das die Varianzaufklärung im Verhältnis zur Anzahl an Prädiktorvariablen und der Anzahl an Items bewertet, von Bedeutung. Der Wert von R_{korr}^2 ist im Modell 3 am höchsten ($R_{korr}^2 = .88$) und nimmt durch eine weitere Variable im Modell 4 wieder etwas ab ($R_{korr}^2 = .87$). Das Ergebnis für den F -Test, der den Determinationskoeffizient statistisch absichert, ist in allen vier Modellen signifikant, wobei der F -Wert im Modell 3 am größten wird ($F = 14.24$).

Aus dem Modell 3, das am besten zu den beobachteten Daten passt, werden die Merkmale entfernt, deren Einfluss auf die Explorationsschwierigkeit im Vergleich zu den anderen Merkmalen als gering einzuschätzen ist. Das trifft für die Variablen *Anzahl der Variablenbedingungen*, *Anzahl sichtbarer Zustände* und *Anzahl veränderbarer Variablen* zu. Es verbleiben damit nur zwei Variablen der Transparenz und zwei Variablen der Komplexität in der Regressionsgleichung, mit denen noch 90 % der Varianz in der Explorationsschwierigkeit aufgeklärt werden kann ($R^2 = .90$, $F = 20.44$, $R_{korr}^2 = .86$). Die Ausprägung der Merkmale *Anzahl unterschiedlicher Bedienelementarten* ($\beta = .37$, $p = .01$) und *Menüverständlichkeit* ($\beta = .42$, $p < .001$) führt zu einer Verringerung der Explorationsschwierigkeit (Erhöhung der Explorationsvollständigkeit), während die Ausprägung der Merkmale *Anzahl an Bedienelementen* ($\beta = -.37$, $p = .01$) und *Menütiefe* ($\beta = -.43$, $p < .001$) zu einer Erhöhung der Explorationsschwierigkeit (Verringerung der Explorationsvollständigkeit) führt (Tabelle 13.6).

Tabelle 13.6: Regressionskoeffizienten und globale Modellgütemaße der Analyse schwierigkeitsbestimmender Merkmale der Systemexploration nach Entfernung von Variablen mit nichtsignifikanten Regressionskoeffizienten ($n = 15$)

Prädiktoren	B	SE	β	t	p
(Konstante)	.54	.10		5.49	.00
Anzahl Bedienelementarten	.12	.04	.37	3.27	.01
Menüverständlichkeit	.32	.08	.42	3.88	.00
Anzahl Bedienelemente	-.02	.01	-.37	-3.38	.01
Menütiefe	-.05	.01	-.43	-3.97	.00

$R = .95$; $R^2 = .90$; $R_{korr}^2 = .86$; $F = 20.44$; $p < .001$

Anmerkungen: n = Itemanzahl; B = unstandardisierter Regressionskoeffizient; SE = Standardfehler; β = standardisierter Regressionskoeffizient; t = t -Wert des unstandardisierten Regressionskoeffizienten; R^2 = Determinationskoeffizient; F = F -Wert; R_{korr}^2 = korrigierter Determinationskoeffizient

13.3.2 Steuerungsphase

Die Vorgehensweise bei der Regressionsanalyse zur Überprüfung der Einflüsse der Merkmale, für die angenommen wird, dass sie den steuernden Umgang mit technischen Geräten erschweren, ist analog zur Regressionsanalyse für die Explorationsphase. Für die Steuerungsphase gibt es jedoch im Gegensatz zur Explorationsphase zwei mögliche Methoden zur Auswertung der empirischen Schwierigkeit. Die erste bewertet die Leistung der Probanden nur anhand des Steuerungsergebnisses (dichotome Leistungsbewertung), während die zweite Möglichkeit die Anzahl der für die Richtiglösung benötigten Bearbeitungsschritte berücksichtigt (Leistungsbewertung berücksichtigt Anzahl an Interaktionen). Mit dem Ziel den durch die verschiedenen Auswertemethoden verursachten Unterschied in der Gesamttestschwierigkeit, durch die Situations- und Systemmerkmale zu erklären, wird für jede der beiden berechneten Itemschwierigkeiten eine separate Regressionsanalyse durchgeführt.

Die nach der Prüfung der Voraussetzungen für die Durchführung einer linearen Regression verbleibenden Merkmale der Problemsituation sind die benötigte *Anzahl unterschiedlicher Lösungsschritte* sowie die *zu berücksichtigenden Mehrfachbelegungen von Bedienelementen*. Die Systemmerkmale werden in der Variable der *Explorationsvollständigkeit* zusammengefasst, die gleichermaßen wie die *steuerungsspezifische Exploration* mit in die Regressionsgleichung eingeht. Als inhaltliches Merkmal, das bereits in der Explorationsphase berücksichtigt wurde, wird die *Vorwissenskompatibilität* der technischen Systeme als eine weitere Ursache für die Varianz in der Itemschwierigkeit vermutet.

In beiden Regressionsgleichungen werden die genannten Prädiktorvariablen blockweise in die Regressionsgleichung aufgenommen, zuerst das Merkmal der Komplexität (*Anzahl unterschiedlicher Lösungsschritte*), im zweiten Schritt das Merkmal der Vernetztheit (*zu berücksichtigende Mehrfachbelegung von Bedienelementen*), im dritten Schritt die beiden Variablen zum Explorationverhalten (*Explorationsvollständigkeit* und *steuerungsspezifische Exploration*) und im letzten Schritte das Inhaltsmerkmal *Vorwissenskompatibilität*. Eine Übersicht der einzelnen Modelle mit den einzelnen Regressionskoeffizienten sowie den entsprechenden globalen Modellgütemaßen für die beiden Regressionsanalysen sind in Tabelle 13.7 (Itemschwierigkeit nach dichotomer Leistungsbewertung) und in Tabelle 13.8 (Leistungsbewertung berücksichtigt die Anzahl an Interaktionen) dargestellt.

Tabelle 13.7: Regressionskoeffizienten und globale Modellgütemaße für die blockweise Analyse des Einflusses von Merkmalen auf die Itemschwierigkeit der Systemsteuerung (dichotome Leistungsbewertung) ($n = 15$)

	Modell 1 Komplexität			Modell 2 + Vernetztheit			Modell 3 + Explorationsverhalten			Modell 4 + Vorwissenskompatibilität		
	<i>B</i> (<i>SE</i>)	β	<i>t</i>	<i>B</i> (<i>SE</i>)	β	<i>t</i>	<i>B</i> (<i>SE</i>)	β	<i>t</i>	<i>B</i> (<i>SE</i>)	β	<i>t</i>
(Konstante)	.56(.11)		5.07***	.69(.12)		5.69***	.48(.19)		2.53*	.49(.19)		2.52*
Anzahl unterschiedlicher Lösungsschritte	-.01(.01)	-.21	-.78n.s.	-.01(.01)	-.36	-1.39n.s.	-.01(.01)	-.39	-1.56n.s.	-.01(.01)	-.31	-1.08n.s.
zu berücksichtigende mehrfachbeleg. Bedienelemente				-.16(.09)	-.49	-1.90n.s.	-.16(.08)	-.48	-1.91n.s.	-.17(.09)	-.52	-1.98n.s.
Explorationsvollständigkeit							.32(.32)	.29	.98n.s.	.34(.33)	.30	1.02n.s.
steuerungsspezifische Exploration							.15(.25)	.17	.59n.s.	.18(.26)	.20	.67n.s.
geringe Vorwissenskompatibilität										-.10(.14)	-.22	-.73n.s.
	$R^2 = .04; F = .60; R_{korr}^2 = -.03$			$R^2 = .27; F = 2.17; R_{korr}^2 = .14$			$R^2 = .43; F = 1.89; R_{korr}^2 = .20$			$R^2 = .46; F = .60; R_{korr}^2 = .16$		

Anmerkungen: n = Itemanzahl; B = unstandardisierter Regressionskoeffizient; SE = Standardfehler; β = standardisierter Regressionskoeffizient; t = t -Wert des unstandardisierten Regressionskoeffizienten; R^2 = Determinationskoeffizient; F = F -Wert; R_{korr}^2 = korrigierter Determinationskoeffizient; Signifikanzniveaus: * $p < .05$; ** $p < .01$; *** $p < .001$; n.s = nicht signifikant

Tabelle 13.8: Regressionskoeffizienten und globale Modellgütemaße für die blockweise Analyse des Einflusses von Merkmalen auf die Itemschwierigkeit der Systemsteuerung (Leistungsbewertung berücksichtigt Anzahl der Interaktionen) (n = 15)

	Modell 1 Komplexität			Modell 2 + Vernetztheit			Modell 3 + Explorationsverhalten			Modell 4 + Vorwissenskompatibilität		
	B(SE)	β	t	B(SE)	β	t	B(SE)	β	t	B(SE)	β	t
(Konstante)	.33(.09)		3.48**	.40(.11)		3.65**	.09(.15)		.64n.s.	.11(.15)		.73n.s.
Anzahl unterschiedlicher Lösungsschritte	.00(.01)	-.14	-.50n.s.	-.01(.01)	-.25	-.87n.s.	-.01(.01)	-.27	-1.14n.s.	.00(.01)	-.16	-.61n.s.
zu berücksichtigende mehrfachbeleg. Bedienelemente				-.10(.08)	-.35	-1.24n.s.	-.09(.07)	-.34	-1.45n.s.	-.11(.07)	-.39	-1.64n.s.
Explorationsvollständigkeit							.56(.25)	.61	2.25*	.59(.25)	.63	2.34*
steuerungsspezifische Exploration							.00(.20)	.01	.02n.s.	.03(.20)	.05	.17n.s.
geringe Vorwissenskompatibilität										-.11(.11)	-.28	-1.04n.s.
	$R^2 = .02; F = .25; R_{korr}^2 = -.06$			$R^2 = .13; F = .91; R_{korr}^2 = -.01$			$R^2 = .50; F = 2.54; R_{korr}^2 = .31$			$R^2 = .56; F = 2.27; R_{korr}^2 = .31$		

Anmerkungen: n = Itemanzahl; B = unstandardisierter Regressionskoeffizient; SE = Standardfehler; β = standardisierter Regressionskoeffizient; t = t-Wert des unstandardisierten Regressionskoeffizienten; R^2 = Determinationskoeffizient; F = F-Wert; R_{korr}^2 = korrigierter Determinationskoeffizient; Signifikanzniveau: * $p < .05$; ** $p < .01$; *** $p < .001$; n.s = nicht signifikant

Sowohl die Ausprägungen des Komplexitätsmerkmals als auch die des Vernetztheitsmerkmals haben einen negativen Einfluss auf die Schwierigkeit der Items in der Steuerungsphase, unabhängig davon wie die empirische Schwierigkeit berechnet wurde. Je größer demnach die nötige *Anzahl der unterschiedlichen Bearbeitungsschritte* und die der *zu berücksichtigenden mehrfachbelegten Bedienelemente* ist, desto schwieriger wird der steuernde Umgang mit den technischen Geräten. In beiden Regressionsanalysen hat von diesen beiden Merkmalen das der Vernetztheit den stärkeren Einfluss auf die Itemschwierigkeit. In dem Modell 4, das alle Merkmale für die Analyse der dichotom ausgewerteten Itemschwierigkeiten enthält, beträgt der standardisierte Regressionskoeffizient für die Variable *Anzahl unterschiedlicher Lösungsschritte* $\beta = -.31$ und der für die Variable *zu berücksichtigende mehrfachbelegte Bedienelemente* $\beta = -.52$. Für beide Merkmale wird das Ergebnis des t-Tests, der diesen Einfluss gegen die Nullhypothese prüft, nicht signifikant.

Das Verhalten in der Explorationsphase hat in beiden Analysen einen positiven Einfluss auf die Itemschwierigkeit. Sowohl die Koeffizienten der *Explorationsvollständigkeit*, die zum größten Teil durch die Systemmerkmale erklärt werden kann, als auch die der *steuerungsspezifischen Exploration* haben ein positives Vorzeichen und erleichtern demnach den Umgang mit den Geräten bei einer gegebenen Steuerungsaufgabe. In der Stärke des Einflusses auf die dichotom ausgewertete Itemschwierigkeit unterscheiden sich beide Variablen, sie beträgt im letzten Modell für die *Explorationsvollständigkeit* $\beta = .30$ und für die *steuerungsspezifische Exploration* $\beta = .20$. Weist ein technisches Gerät eine nur geringe Vorwissenskompatibilität auf, so führt das zu einer durchschnittlich geringeren Lösungshäufigkeit von Problemen in der Steuerung der hier simulierten technischen Geräte ($\beta = -.22$). In der Regressionsanalyse zur Erklärung der Unterschiede in der Explorationsschwierigkeit, hatte die geringe Vorwissenskompatibilität jedoch einen positiven, wenn auch geringen Einfluss ($\beta = .12$). Auch für diese Merkmale wird keiner der *t*-Werte signifikant. Die VIF-Werte aller Prädiktorvariablen liegen unter 1.54, der Mittelwert über alle VIF-Werte ist 1.40. Diese Werte gelten für beide Regressionsanalysen, da sich diese nicht anhand der aufgenommenen Prädiktorvariablen unterscheiden, sondern ausschließlich anhand der abhängigen Variablen.

Ein über beide Regressionsanalysen greifender Vergleich der unstandardisierten Regressionskoeffizienten zeigt deutliche Unterschiede in der Größe des Einflusses einzelner Merkmale auf die entsprechenden Itemschwierigkeiten. Die Variable *Explorationsvollständigkeit* hat auf die dichotom ausgewertete Steuerungsschwierigkeit ($B = .34$) einen geringeren Einfluss als auf die Schwierigkeit, die die Anzahl der für die Lösung benötigten Interaktionen berücksichtigt ($B = .59$) hat. Für das Merkmal *steuerungsspezifische Exploration* ist es umgekehrt, diese hat auf die dichotom ermittelte Schwierigkeit ($B = .18$) der Items in der Steuerungsphase einen größeren Einfluss, als auf die Anzahl der Interaktionen einberechnende Itemschwierigkeit ($B = .03$).

In beiden Analysen kann das jeweils 4. Modell, das alle unabhängigen Variablen umfasst, im Gegensatz zu den jeweils drei anderen Modellen, den größten Anteil an der Varianz in der Itemschwierigkeit erklären. In der ersten Analyse, bei der die abhängige Variable die dichotom ermittelte Steuerungsschwierigkeit darstellt, können so 46 % der Unterschiede in den Schwierigkeiten der einzelnen technischen Geräte vorhergesagt werden ($R^2 = .46$,

$R_{\text{korr}}^2 = .16$). In der zweiten Analyse, in der die Interaktionszahl berücksichtigende Steuerungsschwierigkeit die abhängige Variable bildet, können 56 % Varianz aufgeklärt werden ($R^2 = .56$, $R_{\text{korr}}^2 = .31$). Das Ergebnis für den F -Test wird in keinem der Regressionsmodelle signifikant.

Die deutliche Differenz zwischen dem Determinationskoeffizient R^2 und dem korrigierten Determinationskoeffizient R_{korr}^2 weist auf eine für die geringe Anzahl an Items zu hohe Zahl an Prädiktoren hin, so dass aus den Regressionsgleichungen jeweils die Prädiktoren herausgenommen werden, deren standardisierter Regressionskoeffizient den betragsmäßig geringsten Wert aufweist. Das ist in der ersten Regression (dichotome Leistungsbewertung) die Variable *steuerungsspezifische Exploration* und in der zweiten Regression (Leistungsbewertung berücksichtigt Interaktionszahl) die Variable *Anzahl unterschiedlicher Lösungsschritte*.

Es folgt eine erneute Durchführung der linearen Regressionsanalysen, deren Ergebnisse in Tabelle 13.9 und Tabelle 13.10 zusammengefasst sind. Für die Analyse der Itemschwierigkeit, die sich aus der dichotomen Bewertung der Steuerungsleistung bildet, können keine besseren Modellgütemaße berichtet werden. Weder der Einfluss, der in dem Modell enthaltenen Prädiktorvariablen auf die Itemschwierigkeiten noch der Determinationskoeffizient R^2 als globales Modellgütemaß sind signifikant. In der zweiten Analyse, wird nur das Regressionsgewicht der Variable *Explorationsvollständigkeit* signifikant ($t = 1.71$; $p = .01$). Obwohl die anderen Merkmale keinen signifikanten Beitrag zur Erklärung der Schwierigkeitsunterschiede leisten können, wird R^2 für dieses Modell signifikant ($F = 4.28$; $p = .03$). Die Differenz zwischen R^2 und R_{korr}^2 ist durch das Entfernen der Variable *Anzahl unterschiedlicher Lösungsschritte* deutlich geringer geworden ($R^2 = .54$; $R_{\text{korr}}^2 = .41$).

Tabelle 13.9: Regressionskoeffizienten und globale Modellgütemaße der Analyse schwierigkeitsbestimmender Merkmale der Systemsteuerung (dichotome Leistungsbewertung) nach Entfernung von Variablen mit geringem Einfluss auf die Itemschwierigkeit ($n = 15$)

Prädiktoren	B	SE	β	t	p
(Konstante)	.47	.19		2.52	.03
Anzahl unterschiedlicher Lösungsschritte	-.01	.01	-.30	-1.08	.31
zu berücksichtigende Mehrfachbelegung Bedienelemente	-.17	.09	-.52	-2.05	.07
Explorationsvollständigkeit	.46	.27	.41	1.71	.12
geringe Vorwissenskompatibilität	-.09	.14	-.19	-.66	.53

$R = .66$; $R^2 = .44$; $R_{\text{korr}}^2 = .21$; $F = 1.93$; $p = .18$

Anmerkungen: n = Itemanzahl; B = unstandardisierter Regressionskoeffizient; SE = Standardfehler; β = standardisierter Regressionskoeffizient; t = t -Wert des unstandardisierten Regressionskoeffizienten; R^2 = Determinationskoeffizient; F = F -Wert; R_{korr}^2 = korrigierter Determinationskoeffizient

Tabelle 13.10: Regressionskoeffizienten und globale Modellgütemaße der Analyse schwierigkeitsbestimmender Merkmale der Systemsteuerung (Leistungsbewertung berücksichtigt Anzahl der Interaktionen) nach Entfernung von Variablen mit geringem Einfluss auf die Itemschwierigkeit (n = 15)

Prädiktoren	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>
(Konstante)	.07	.12		.55	.59
zu berücksichtigende Mehrfachbelegung Bedienelemente	-.10	.06	-.37	-1.69	.12
Explorationsvollständigkeit	.62	.19	.67	3.18	.01
geringe Vorwissenskompatibilität	-.14	.09	-.34	-1.54	.15

$R = .73$; $R^2 = .54$; $R_{kor}^2 = .41$; $F = 4.28$; $p = .03$

Anmerkungen: *n* = Itemanzahl; *B* = unstandardisierter Regressionskoeffizient; *SE* = Standardfehler; β = standardisierter Regressionskoeffizient; *t* = *t*-Wert des unstandardisierten Regressionskoeffizienten; R^2 = Determinationskoeffizient; *F* = *F*-Wert; R_{kor}^2 = korrigierter Determinationskoeffizient

14 Strukturanalyse technischer Problemlösekompetenz

Beantwortung der Hypothesen 2.1 - 2.3

Der in diesem Abschnitt zu untersuchende Validitätsaspekt betrifft die logisch-inhaltliche Analyse der theoretisch angenommenen Konstruktstruktur. Mit der Struktur des Konstruktes verbunden sind die Annahmen über dessen Dimensionalität, also der Anzahl latenter Variablen, die für das Zustandekommen der Itemantworten ursächlich sind. Mit der logisch-inhaltlichen Analyse soll geprüft werden, ob die theoretisch angenommene Struktur des Konstruktes auch in dem Testverhalten der Probanden zum Ausdruck kommt.

Für das Konstrukt der Problemlösekompetenz im Umgang mit technischen Alltagsgeräten wurden im Abschnitt 6.3 zwei nach dem Interaktionsziel differenzierte Dimensionen formuliert. In diesem Abschnitt wird geprüft, ob sich eine solche zweidimensionale Struktur mit den Subdimensionen *Systemexploration* und *Systemsteuerung* auch empirisch ermitteln lässt. Zu diesem Zweck werden mithilfe konfirmatorischer Faktorenanalysen drei alternative Modelle getestet und miteinander verglichen, die im Folgenden zunächst vorgestellt werden.

Für die Systemsteuerung müssen die Probanden das in der Systemexploration erworbene Eingriffswissen nutzen, womit zum einen die Position der Steuerungselements nach den zugehörigen Explorationselements festgelegt wird und zum anderen von einem Einfluss der Systemexploration auf die Systemsteuerung auszugehen ist. Die spezifizierten Modelle unterscheiden sich darin, in welcher Form dieser Zusammenhang zwischen der Systemexploration und der Systemsteuerung gegeben ist.

Das erste Modell M1 (Abbildung 14.1 a) geht davon aus, dass der Zusammenhang zwischen der Systemexploration und der Systemsteuerung so groß ist, dass beide latenten Merkmale dasselbe Merkmal erfassen und deshalb nicht zwischen diesen beiden Subdimensionen differenziert werden sollte. Als Restriktion wird die Korrelation zwischen den latenten Variablen Systemexploration und Systemsteuerung auf den Wert 1 fixiert. Weder die Faktorladungen der einzelnen Indikatoren noch die Itemschwierigkeiten werden fixiert, sondern frei geschätzt. Die Unkorreliertheit der Messfehler wird in dem Modell M1 unterstellt. Das zweite Modell M2 (Abbildung 14.1 b) nimmt zwar einen Zusammenhang zwischen Systemexploration und Systemsteuerung an, jedoch wird nicht unterstellt, dass es sich bei den beiden Subdimensionen um eine einzige latente Variable handelt. Die Restriktion der auf 1 fixierten Korrelation zwischen den beiden latenten Variablen wird in dem Modell M2 aufgehoben und ihre Höhe in der konfirmatorischen Faktorenanalyse frei geschätzt. Eine weitere Restriktion, die in dem Modell M3 (Abbildung 14.1 c) aufgehoben wird, ist die Annahme, dass die Residuen unkorreliert sind. Die Residuen, also der Anteil der Varianz eines Indikators, der nicht durch die dahinterliegende latente Variable erklärt werden kann, sollten im Idealfall bei voneinander unabhängigen Indikatoren nicht miteinander korrelieren. In dem vorliegenden Test zur Erfassung des problemlösenden Umgangs mit technischen Alltagsgeräten kann diese Unabhängigkeit teilweise jedoch nicht vermieden werden. So hängt der Erfolg bei der Steuerung eines technischen Systems in dem entwickelten Test nicht nur von der Fähigkeit Wissen anzuwenden ab, sondern auch davon, wie viel Wissen in der zeitlich vorgelagerten Systemexploration erworben wurde. Diese bei allen Items gleich große Abhängigkeit drückt sich in der Höhe der Korrelation zwischen der latenten

Systemexploration und der latenten Systemsteuerung aus. Denkbar ist jedoch auch, dass die Abhängigkeit der Systemsteuerung von der Systemexploration nicht für alle technischen Systeme gleich groß ausfällt, so dass außerdem die Residuen der jeweils abhängigen Items korreliert sind. Der, von Cole, Ciesla & Steiger (2007, S. 381) allgemein als geteilte Methodenvarianz bezeichnete Effekt wird in dem Modell M3 für die jeweils auf demselben technischen Gerät basierenden Items zugelassen.

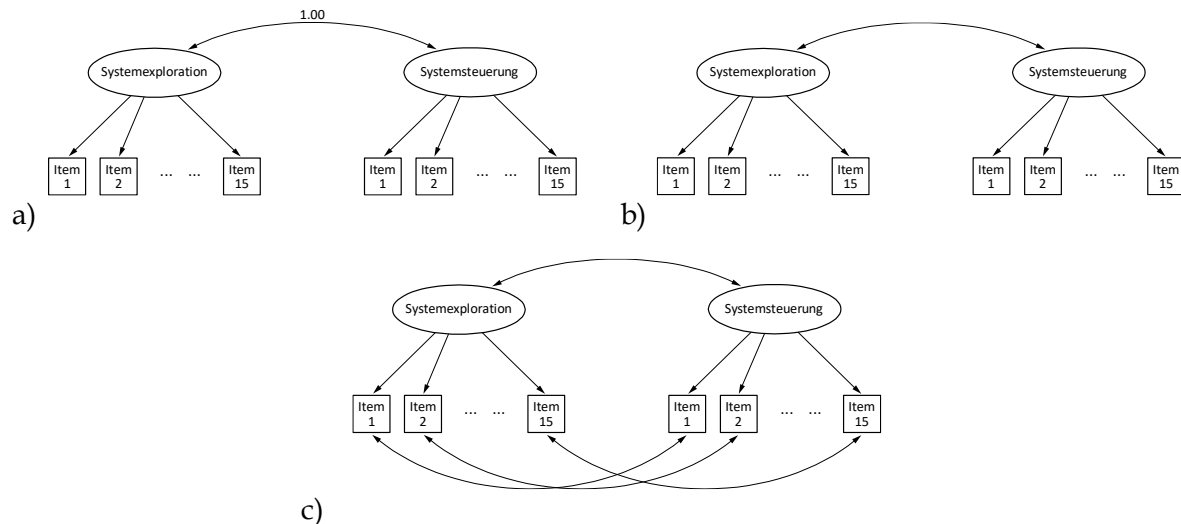


Abbildung 14.1: Modelle für die Struktur des problemlösenden Umgangs mit technischen Alltagsgeräten a) Systemexploration und Systemsteuerung erfassen dasselbe Merkmal; b) Systemexploration und Systemsteuerung sind zwei voneinander abgrenzbare Merkmale, die miteinander korrelieren; c) Systemexploration und Systemsteuerung sind zwei voneinander abgrenzbare Merkmale die miteinander korrelieren, die Residuen der Indikatoren sind teilweise korreliert

Für die Parameterschätzung der drei genannten Modelle wird das bereits in Kapitel 12 eingesetzte DWLS-Schätzverfahren, mit anschließender Korrektur der Standardfehler sowie der χ^2 -Statistik verwendet. Letztere sowie weitere Fitstatistiken für die drei unterschiedlich restringierten Modelle sind in Tabelle 14.1 zusammengefasst. Für keines der drei Modelle kann von einer fehlerfreien Modellspezifikation ausgegangen werden, da alle der angegebenen χ^2 -Werte signifikant sind.

Tabelle 14.1: Fitstatistiken des Modellvergleichs für die Dimensionalitätsprüfung des problemlösenden Umgangs mit technischen Alltagsgeräten ($n = 130$)

	χ^2	df	p	χ^2/df	RMSEA (90 % CI)	CFI	WRMR	$\Delta\chi^2$ (Δdf)	ΔCFI
M3: zweidimensionales Modell mit korrelierenden Residuen	552.18	389	.000	1.42	.06 (.04, .07)	.90	.87		
M2: zweidimensionales Modell	657.47	405	.000	1.62	.07 (.05, .08)	.87	.95	84.13 (15)***	-.039
M1: eindimensionales Modell	634.20	404	.000	1.57	.07 (.06, .08)	.85	.93	91.80 (1)***	-.013

Anmerkungen: n = Stichprobengröße; χ^2 = Prüfgröße der Modellabweichung; df = Anzahl der Freiheitsgrade; p = p -Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; RMSEA = Root Mean Square Error of Approximation; 90 % CI = Konfidenzintervall; CFI = Comparative Fit Index; WRMR = Weighted Root Mean Square Residual
Nach Satorra-Bentler korrigierte χ^2 -Werte; die Differenzen $\Delta\chi^2$ (Δdf) sind die korrigierten Differenzen der beiden χ^2 Werte; *** $p < .001$

Ein Vergleich der anderen Modellfitwerte, wie der RMSEA oder der CFI weist auf eine höhere Passung des dritten Modells M3 zu den empirischen Daten hin, als die Modelle M1 und M2. Für dieses Modell liegen die Werte des CFI = .90, des RMSEA = .06 und des WRMR = .87 in einem akzeptablen Bereich. Die Ergebnisse des χ^2 -Differenzentests, der für den direkten Vergleich jeweils zweier genesteter Modelle eingesetzt wird, sprechen jeweils gegen das restringiertere Modell. So wird die korrigierte χ^2 -Differenz zwischen den Modellen M1 und M2 mit $\Delta\chi^2 = 91.80$ sowie die korrigierte χ^2 -Differenz zwischen den Modellen M2 und M3 mit $\Delta\chi^2 = 84.13$ signifikant. Die Differenzen zwischen den CFIs sind in beiden Modellvergleichen betragsmäßig größer $\Delta\text{CFI} = -.01$ ($\Delta\text{CFI}_{M1,M2} = -.013$; $\Delta\text{CFI}_{M2,M3} = -.039$). Dieser Befund spricht für die Beibehaltung der Hypothese $H_{2.1}$ nach der Explorieren und Steuern zwei voneinander abgrenzbare Problembereiche im Umgang mit technischen Alltagsgeräten sind.

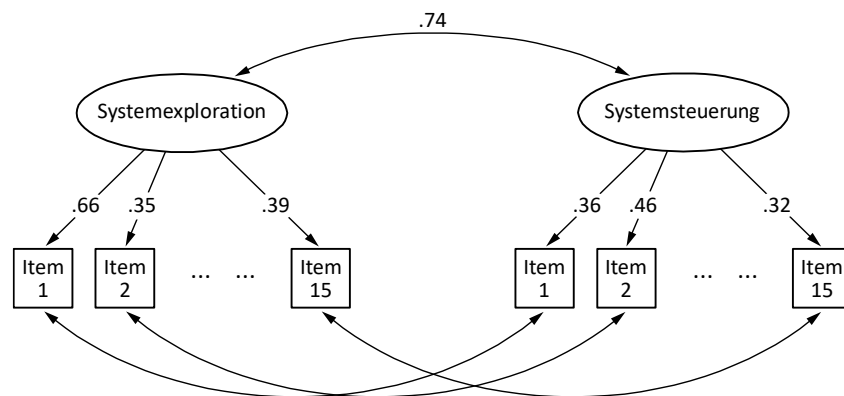


Abbildung 14.2: Ergebnis der Dimensionalitätsprüfung für das zweidimensionale Modell technischer Problemlösekompetenz mit korrelierenden Residuen ($n = 130$)

Das Ergebnis der Parameterschätzung zeigt, dass die latenten Variablen Systemexploration und Systemsteuerung mit $r = .74$ ($p < .001$) hoch korrelieren (Abbildung 14.2). Die Hypothese $H_{2.2}$, nach der eine hohe Explorationsvollständigkeit mit einer hohen Steuerungsleistung einhergeht, muss demnach ebenfalls nicht verworfen werden.

Im Vergleich zu den im Abschnitt 12 analysierten Messmodellen haben sich die Ladungen der einzelnen manifesten Indikatoren auf ihre latenten Variablen in ihrer Höhe jedoch verändert. So fallen die Faktorladungen im Messmodell der Systemexploration für das hier betrachtete zweidimensionale Modell durchweg höher aus, als für das isoliert analysierte Messmodell der Systemexploration. Die Differenz zwischen den Faktorladungen der jeweiligen Modelle beträgt jedoch maximal $\Delta\lambda = .10$. Alle Faktorladungen der Systemexploration weisen in dem zweidimensionalen Modell Werte von $\lambda \geq .30$ auf und sind signifikant (Tabelle 14.2). Im Messmodell der Systemsteuerung hingegen fallen sämtliche Faktorladungen im Gegensatz zu dem vorher isoliert betrachteten Messmodell geringer aus. Hier beträgt die Differenz sogar bis zu $\Delta\lambda = .19$. Von den 15 Faktorladungen liegen in dem zweidimensionalen Modell vier unterhalb von $\lambda \geq .30$, von denen drei nicht signifikant sind. Bei den betroffenen Items handelt es sich um das Item 3 (Steuerung der Spülmaschine), Item 7 (Steuerung des Küchenradios), Item 8 (Steuerung des Navigationsgerätes) sowie das Item 14 (Steuerung des Mobiltelefons). Diese vier technischen Systeme scheinen sich in ihren vorgegebenen Steuerungszielen derartig von den anderen zu unterscheiden, als dass neben der Wissensanwendung weitere unberücksichtigte Aspekte erforderlich sind, um das Steuerungsziel zu erreichen. Anhand der

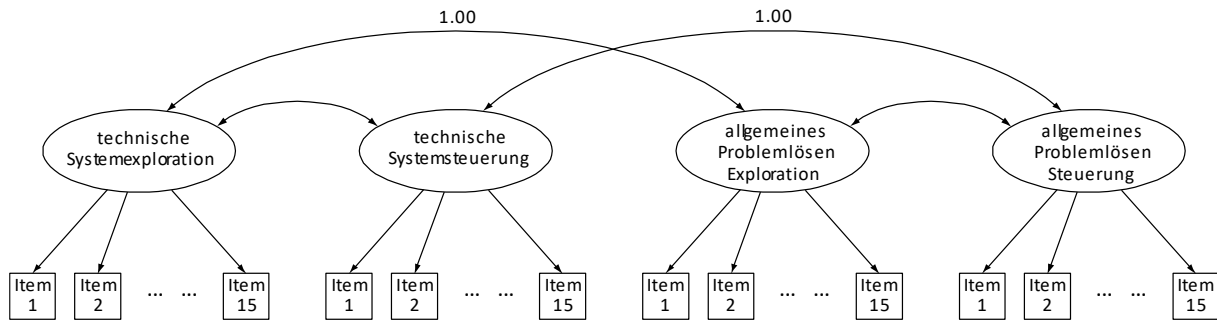
Residualkorrelationen, die ebenfalls in Tabelle 14.2 wiedergegeben werden, kann vermutet werden, dass bei zwei der technischen Geräte die Abweichungen von der Skala in dem Gerät selbst zu suchen ist (die Residuen der Exploration und der Steuerung für das Item 3 und 14 korrelieren signifikant), während bei den anderen zwei Geräten die Auffälligkeit in dem zu erreichenden Steuerungsziel liegt (die Residuen der Exploration und Steuerung für die Items 7 und 8 korrelieren nicht).

Tabelle 14.2: standardisierte Faktorladungen und Korrelationskoeffizienten der Residuen zwischen den Items der Systemexploration und der Systemsteuerung im technischen Problemlösetest (n = 130)

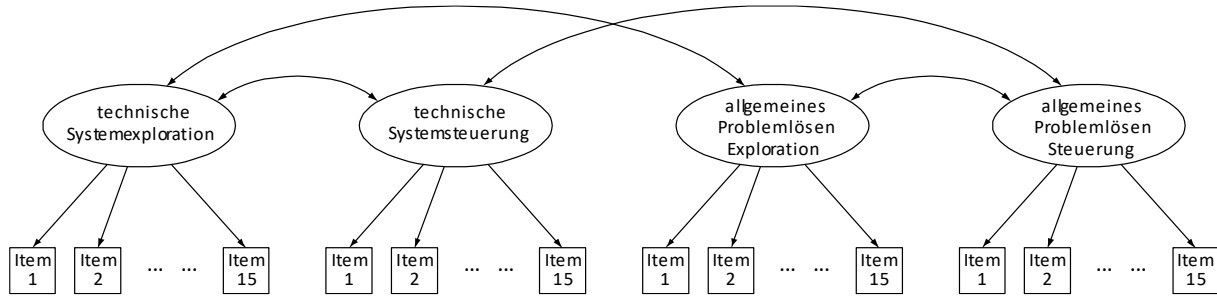
	stand. Faktorladung		Residualkorrelation
	Systemexploration	Systemsteuerung	
Item 1	.66 ***	.36 **	.34 **
Item 2	.35 ***	.46 *	.46 ***
Item 3	.30 ***	.19	.53 ***
Item 4	.50 ***	.40 **	.08
Item 5	.70 ***	.47 ***	-.05
Item 6	.45 ***	.43 ***	.28 *
Item 7	.36 **	.21	.11
Item 8	.63 ***	.29 *	-.06
Item 9	.55 ***	.31 *	.08
Item 10	.50 ***	.57 ***	.09
Item 11	.74 ***	.58 ***	.07
Item 12	.60 ***	.78 ***	.81 ***
Item 13	.55 ***	.69 ***	.00
Item 14	.59 ***	.20	.25 *
Item 15	.39 ***	.32 **	.06

Anmerkungen: n = Stichprobengröße; * p < .05; ** p < .01; *** p < .001

Zu der logisch-inhaltlichen Analyse des Konstruktes gehört auch die Prüfung des in dem Theorieteil dieser Arbeit herausgearbeiteten Kontexteinflusses auf das Problemlösen im Umgang mit technischen Alltagsgeräten. Dieser Kontexteinfluss sollte sich in der Differenzierbarkeit der Konstrukte allgemeines Problemlösen und Problemlösen im Umgang mit technischen Alltagsgeräten zeigen. Die Prüfung der diskriminanten Validität wird durch den Vergleich der folgenden zwei Modelle vorgenommen: Das Modell M4 unterscheidet Problemlösen lediglich hinsichtlich der Exploration (den Wissenserwerb) und der Steuerung (der Wissensanwendung). In welchen Kontexten diese Subdimensionen stattfinden, ist in diesem Modell unerheblich, so dass die Korrelation zwischen der latenten technischen Systemexploration und der latenten Exploration des allgemeinen Problemlösens auf 1 fixiert wird, ebenso wie die Korrelation zwischen der latenten technischen Systemsteuerung und der latenten Steuerung des allgemeinen Problemlösens (Abbildung 14.3 a). Gegenübergestellt wird ein Modell M5, das sowohl zwischen der Exploration und der Steuerung als auch zwischen technischem und allgemeinem Problemlösen unterscheidet. Hier wird angenommen, dass sich sowohl die Exploration als auch die Steuerung in technischen Kontexten von allgemein gehaltenen Problemlöseszenarien abgrenzt, die Restriktion der fixierten Korrelation zwischen den jeweiligen latenten Variablen wird hier aufgehoben, so dass die Korrelationen frei geschätzt werden können.



a)



b)

Abbildung 14.3: Modelle zur Prüfung diskriminanter Validität a) technische Systemexploration und Exploration im allgemeinen Problemlösen sind identisch, technische Systemsteuerung und Steuerung im allgemeinen Problemlösen sind identisch; b) technische Systemexploration, technische Systemsteuerung, Exploration und Steuerung im allgemeinen Problemlösen sind vier unterscheidbare latente Variablen

Beide Modelle enthalten kategoriale Variablen, so dass zu ihrer Schätzung das DWLS-Schätzverfahren eingesetzt wurde. Die korrigierten Fitstatistiken für die beiden Modelle sind in Tabelle 14.3 aufgeführt. Sowohl das Verhältnis χ^2/df als auch der RMSEA, der CFI und der WRMR liegen für beide Modelle oberhalb ihrer Cut-Off-Werte. Bis auf den signifikanten χ^2 -Wert lassen sich keine weiteren Hinweise auf eine Fehlspezifikation der Modelle finden.

Tabelle 14.3: Fitstatistiken des Modellvergleichs zur Prüfung diskriminanter Validität des problemlösenden Umgangs mit technischen Alltagsgeräten (n = 77)

	χ^2	df	p	χ^2/df	RMSEA (90 % CI)	CFI	WRMR	$\Delta\chi^2 (\Delta df)$	ΔCFI
M5: vierdimensionales Modell	786.80	659	.000	1.19	.04 (.03, .05)	.96	.82		
M4: zweidimensionales Modell	865.00	661	.000	1.30	.05 (.04, .06)	.94	.86	47.84(2)***	-.023

Anmerkungen: n = Stichprobengröße; χ^2 = Prüfgröße der Modellabweichung; df = Anzahl der Freiheitsgrade; p = p-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; RMSEA = Root Mean Square Error of Approximation; 90 % CI = Konfidenzintervall; CFI = Comparative Fit Index; WRMR = Weighted Root Mean Square Residual; Nach Satorra-Bentler korrigierte χ^2 -Werte; die Differenzen $\Delta\chi^2 (\Delta df)$ sind die korrigierten Differenzen der beiden χ^2 Werte; *** p < .001

Da das Modell 4 durch Restriktionen aus dem Modell 5 hervorgeht und es sich demnach um genestete Modelle handelt, kann für den statistischen Vergleich der χ^2 -Differenzentest eingesetzt werden. Das signifikante Ergebnis des Tests mit $\Delta\chi^2 (\Delta df) = 47.84 (2)$ spricht für eine bessere Passung zwischen der von Modell M5 induzierten und der empirischen Kovarianzmatrix. Auch die Differenz zwischen den beiden CFI-Werten ist mit $\Delta CFI = -.023$ als

ausreichend hoch einzuschätzen, um die Hypothese H_{2.3} nach der allgemeines Problemlösen und Problemlösen im Umgang mit technischen Systemen zwei voneinander abgrenzbare Problembereiche sind, zunächst beizubehalten.

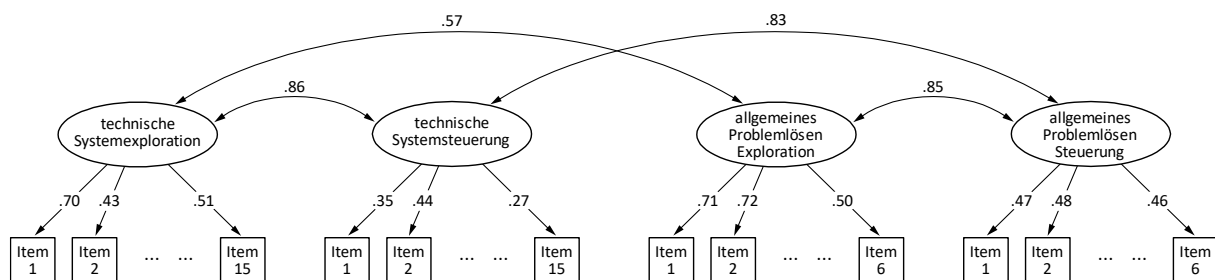


Abbildung 14.4: Ergebnis für das vierdimensionale Modell, das sowohl zwischen Exploration und Steuerung als auch zwischen technischem und allgemeinem Problemlösen differenziert. Alle eingezeichneten Korrelationen zwischen den latenten Variablen sind mit $p < .001$ signifikant ($n = 77$)

Die Höhe des Zusammenhangs zwischen den latenten Variablen Exploration und Steuerung fällt für beide Konstrukte mit $r = .86$ und $.85$ etwa gleich groß aus. Die Beziehungen zwischen den Konstrukten des technischen und des allgemeinen Problemlösen weichen jedoch auffallend voneinander ab. Während die Korrelation der Exploration beider Konstrukte mit $r = .57$ für einen mittelhohen Zusammenhang spricht, ist für die Korrelation der Steuerung beider Konstrukte mit $r = .83$ ein hoher Zusammenhang festzustellen.

15 Verortung technischer Problemlösekompetenz im nomologischen Netzwerk

Beantwortung der Hypothesen 3.1 - 3.18

Ein gewichtiger Teil dieser Arbeit beschäftigt sich mit möglichen Einflussgrößen auf den explorierenden und den steuernden Umgang mit technischen Alltagsgeräten. Während die erste Einschätzung der manifesten Variablenzusammenhänge, die im Abschnitt 11 *Deskriptive Befunde* vorgenommen wurde, lediglich geeignet ist, die Beziehungen der exogenen Variablen zu betrachten, ist für die Interpretation der kausalen Wirkungen auf die endogenen Variablen das in Abschnitt 10.3 besprochene Strukturgleichungsmodell zu analysieren. Der nachstehende Ergebnisteil befasst sich mit der Beantwortung der Hypothesen 3.1 - 3.18, die sich auf die Verortung des Konstruktes des problemlösenden Umgangs mit technischen Alltagsgeräten in das postulierte nomologische Netzwerk beziehen.

Die in diesem nomologischen Netzwerk außerdem verorteten Konstrukte sind fluide Intelligenz, die auf Problemlösen bezogene Selbstwirksamkeitserwartung, Need for Cognition, allgemeine Problemlösefähigkeit sowie ICT-Nutzung und das gerätespezifische Vorwissen. Aus Gründen des im Abschnitt 10.3 bereits angesprochenen ungünstigen Verhältnisses zwischen der Anzahl zu schätzender Modellparameter und der Anzahl beobachteter Daten wurden die Konstrukte, deren Erfassung über eine Fragebogenskala erfolgte, als manifeste Variablen in das Strukturmodell eingefügt. Skalen, die als manifeste Variablen in das Strukturmodell eingehen sind: *Ausdauer, Aufgeschlossenheit gegenüber Problemlösen, Need for Cognition, Nutzungshäufigkeit von ICT* sowie die *gerätespezifische Vorerfahrung*. Diese manifesten Variablen werden aus dem jeweiligen Skalenwert gebildet, der sich aus der Summe der einzelnen Itemwerte ergibt. Die Verteilungen der Skalenwerte in der Gesamtstichprobe mit ihren Lage- und Streuungsparametern wurden bereits im Abschnitt 11 vorgestellt.

Die kognitiven Konstrukte, für deren Erfassung keine Skalen sondern ein Test zum Einsatz kam, wurden als latente Variablen in das Strukturmodell eingebracht, deren Messmodelle vorab zunächst isoliert von anderen Konstrukten analysiert wurden. Die Ergebnisse für die latenten Variablen des explorierenden und des steuernden Umgangs mit technischen Systemen waren bereits Inhalt des Abschnittes 12 und sind nur aus Vollständigkeitsgründen noch einmal in Tabelle 15.1 aufgeführt. Als Indikatorvariablen für die latente Variable der fluiden Intelligenz dienten die Summenscores der Subtests *Reihenfortsetzen, Klassifikation, Matrizen* sowie *Topologien*. Die Ladungen der Subtests auf die latente Variable liegen zwischen den Werten .61 und .76. Der Modellfit für das Messmodell fluider Intelligenz wurde aus allen 147 Beobachtungen geschätzt und zeigt keine Fehlspezifikationen an ($X^2/df = 1.17$, CFI = 1.00, RMSEA = .03, WRMR = .19).

Das Konstrukt allgemeine Problemlösefähigkeit wurde in Anlehnung an Neubert, Kretzschmar, Wüstenberg et al. (2015) sowie Greiff, Fischer, Wüstenberg et al. (2013) anhand der zwei Subdimensionen Exploration (dort Wissenserwerb bzw. Wissen) und Steuerung (dort Wissensanwendung) differenziert. Als Indikatorvariablen für das Messmodell der Exploration dienten die sechs arithmetischen Mittel, die aus den für jedes Szenario jeweils zu

lösenden zwei Identifikationsaufgaben gebildet wurden. Für das Messmodell der Steuerung werden die sechs Steuerungssitems als Indikatoren verwendet. Die Gütemaße für die beiden Messmodelle der allgemeinen Problemlösefähigkeit sind in Tabelle 15.1 dargestellt. Die χ^2 -Statistik wird für das Messmodell der Steuerung signifikant und der χ^2 -Wert größer als zweimal die Anzahl der Freiheitsgrade ($\chi^2/df = 2.29$). Vermutlich liegen hier Modellabweichungen vor. Dafür würde auch der Wert für den RMSEA sprechen, der die Abweichungen zwischen modellimplizierter und empirischer Kovarianzmatrix zusammenfasst und dabei die Stichprobengröße sowie die Anzahl der Freiheitsgrade berücksichtigt. Dieser beträgt für das Messmodell der Steuerung im allgemeinen Problemlösen $RMSEA = .12$ und liegt damit über den für einen akzeptablen Fit geltenden Wert von $< .10$. Die Werte für den CFI und den WRMR bleiben hingegen noch in einem akzeptablen Bereich.

Tabelle 15.1: Modellgütemaße für die Messmodelle, die in dem Strukturmodell eingefügten latenten Variablen

	<i>n</i>	χ^2	<i>df</i>	<i>p</i>	χ^2/df	RMSEA (90 % CI)	CFI	WRMR
fluide Intelligenz	147	2.34	2	.31	1.17	.03 (.00, .00)	1.00	.19
technische Systemexploration	136	133.00	90	.00	1.47	.06 (.02, .09)	.97	.69
technische Systemsteuerung	141	135.17	90	.00	1.50	.06 (.04, .08)	.84	.95
allgemeines Problemlösen Exploration	92	3.91	9	.91	.43	.00 (.00, .09)	1.00	.27
allgemeines Problemlösen Steuerung	86	20.61	9	.02	2.29	.12 (.03, .21)	.91	.72

Anmerkungen: *n* = Stichprobengröße; χ^2 = Prüfgröße der Modellabweichung; *df* = Anzahl der Freiheitsgrade; *p* = *p*-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; RMSEA = Root Mean Square Error of Approximation; 90 % CI = Konfidenzintervall; CFI = Comparative Fit Index; WRMR = Weighted Root Mean Square Residual

Die zum Teil auf latenter und zum Teil auf manifester Ebene spezifizierten Variablen wurden in dem Strukturmodell so zueinander in Beziehung gesetzt, dass sie das Theoriemodell wiedergeben und die im Abschnitt 7 formulierten Hypothesen beantworten können. Demnach ist ein Zusammenhang zwischen der latenten Variable Intelligenz und dem durch die manifesten Variablen ICT-Nutzung und gerätespezifisches Vorwissen repräsentierte Vorwissen einer Person zu erwarten (Hypothese H_{3.6}). Fluide Intelligenz wirkt außerdem kausal auf den explorierenden und den steuernden Umgang mit technischen Systemen und mit allgemeinen Problemlöseszenarien (H_{3.1} - H_{3.4}), wobei Intelligenz auf die Variablen der allgemeinen Problemlösefähigkeit einen größeren Einfluss haben sollte, als auf die Variablen der technischen Problemlösekompetenz (H_{3.5}). Das über die manifesten Variablen Ausdauer und Aufgeschlossenheit gegenüber Problemlösen repräsentierte Konstrukt problemlösebezogene Selbstwirksamkeitserwartung steht im Zusammenhang mit dem Konstrukt Need for Cognition (H_{3.11}). Von beiden Eigenschaften wird angenommen, dass sie kausal auf alle abhängigen Variablen wirken (H_{3.7} - H_{3.10}, H_{3.12} - H_{3.15}). Nicht auf alle, sondern nur auf den explorierenden und steuernden Umgang mit technischen Systemen wird ein Einfluss des Vorwissens erwartet (H_{3.16} - H_{3.17}). Zudem ist davon auszugehen, dass es einen Zusammenhang zwischen dem Vorwissen und der Selbstwirksamkeitserwartung einer Person gibt (H_{3.18}).

Abbildung 15.1 zeigt das sich aus den genannten Zusammenhängen ergebene vollständige und geschätzte Strukturmodell. Die Modellgütemaße für das Gesamtmodell zeigen bis auf die signifikante χ^2 -Statistik einen guten Modellfit an ($\chi^2/df = 1.21$; CFI = .96; RMSEA = .04

[.03; .05]; WRMR = .82). Die weiteren Darstellungen beziehen sich auf die hypothesenprüfende Interpretation einzelner Zusammenhänge.

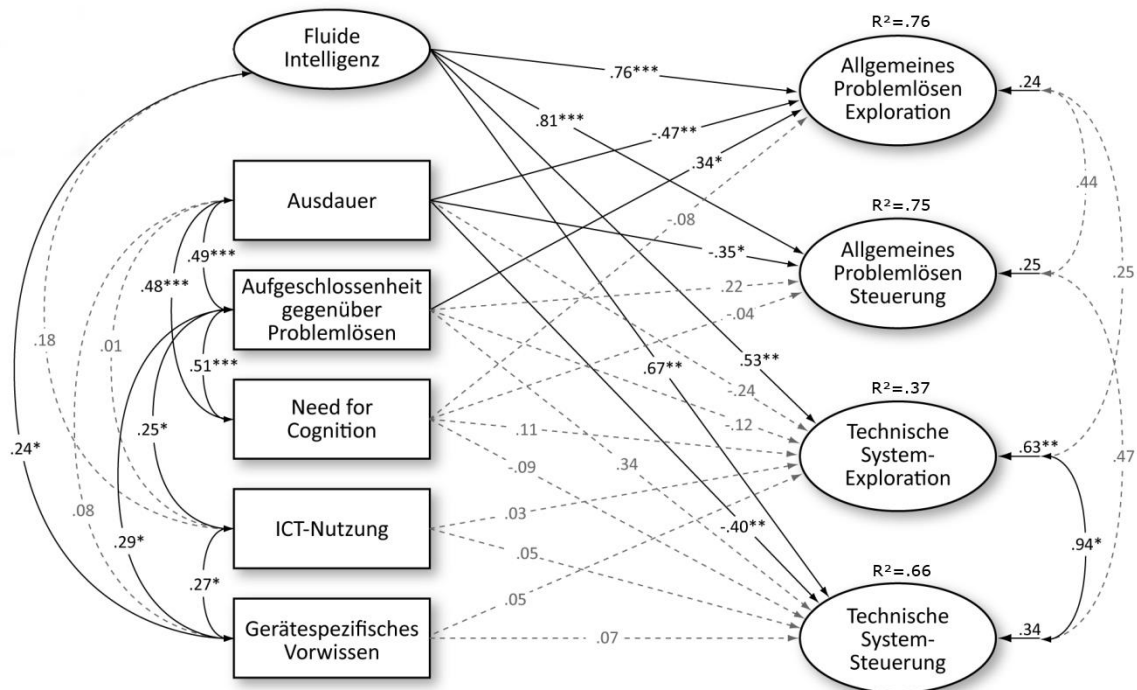


Abbildung 15.1: Strukturmodell zur Verortung technischer Problemlösekompetenz in einem Theoriemodell ($n = 77$); getrichelt dargestellte Effekte sind nicht signifikant, * $p < .05$; ** $p < .01$; *** $p < .001$; R^2 = Kommunalität; Modellfit: $\chi^2/df = 1.21$; CFI = .96; RMSEA = .04 [.03, .05]; WRMR = .82

Der stärkste Prädiktor für die Leistungen im explorierenden und steuernden Umgang mit technischen Systemen ist fluide Intelligenz. Die standardisierten Pfadkoeffizienten betragen für den Einfluss der Intelligenz auf die Systemexploration $\beta = .53$ ($p = .002$) und auf die Systemsteuerung $\beta = .67$ ($p = .003$). Diese Ergebnisse sprechen für die Beibehaltung der Hypothesen $H_{3.1}$ und $H_{3.2}$, nach denen intelligentere Personen höhere Leistungen im explorierenden und steuernden Umgang mit technischen Systemen erzielen. Die latente Korrelation zwischen fluider Intelligenz und der Exploration in allgemeinen Problemlöseszenarien wird auf $\beta = .76$ ($p < .001$) und zwischen fluider Intelligenz und der Steuerung in allgemeinen Problemlöseszenarien sogar auf $\beta = .81$ ($p < .001$) geschätzt. Die Hypothesen $H_{3.3}$ und $H_{3.4}$, nach denen intelligentere Personen auch eine höhere allgemeine Problemlösefähigkeit aufweisen, können nach diesen Befunden ebenfalls bestätigt werden. Die unterschiedlichen hohen Pfadkoeffizienten zwischen Intelligenz, den Subdimensionen technischer Problemlösekompetenz und den Subdimensionen allgemeiner Problemlösefähigkeit sprechen für einen höheren Einfluss der Intelligenz auf den explorierenden und steuernden Umgang mit allgemeinen Problemlöseszenarien als auf den Umgang mit technischen Alltagsgeräten und damit für die Beibehaltung der Hypothese $H_{3.5}$. Der vermutete Zusammenhang zwischen Intelligenz und dem Vorwissen zeigt sich nur für das gerätespezifische Vorwissen und ist mit $r = .24$ ($p = .04$) eher als gering einzuschätzen. Da das in dieser Studie erhobene gerätespezifische Vorwissen lediglich wenig strukturierte

Wissensaspekte umfasst, war ein stärkerer Zusammenhang zur Intelligenz auch nicht zu erwarten.

Die auf das Lösen von Problemen bezogene Selbstwirksamkeitserwartung wurde mit den Selbsteinschätzungsskalen *Ausdauer* bei der Bewältigung neuer Aufgaben sowie *Aufgeschlossenheit gegenüber dem Problemlösen* erfasst. Beide Skalen korrelieren zu $r = .49$ ($p < .001$) miteinander und erfassen damit unterschiedliche Facetten der problemlösebezogenen Selbstwirksamkeitserwartung. Diese Unterschiede zeigen sich auch deutlich in den Einflüssen, die sie auf die latenten Variablen des allgemeinen und technischen Problemlösens ausüben. So haben die Pfadkoeffizienten zwischen der Ausdauer bei der Aufgabenbewältigung und den endogenen Variablen ein durchweg negatives Vorzeichen, während sie zwischen der Aufgeschlossenheit gegenüber dem Problemlösen und den endogenen Variablen (mit Ausnahme technischer Systemexploration) positiv sind.

Für die Variable *Ausdauer* bedeutet das, dass Personen die sich selbst als besonders ausdauernd bei der Bewältigung neuer Anforderungen einschätzen (der genaue Wortlaut der Skala ist im Abschnitt 9.1.2 nachzulesen), tendenziell eher geringere Leistungen erbrachten. Die Pfadkoeffizienten zwischen der Variable Ausdauer und dem explorierenden und steuernden Umgang mit allgemeinen fachübergreifenden Problemszenarien betragen $\beta = -.47$ und $\beta = -.35$ und sind in beiden Fällen signifikant. Zwischen der Variable Ausdauer und der technischen Systemsteuerung besteht mit $\beta = -.40$ ein ähnlich hoher Zusammenhang, lediglich der Effekt auf die technische Systemexploration fällt mit $\beta = -.24$ gering aus und wird nicht signifikant. Die technische Systemexploration ist die einzige prozessorientiert erfasste Variable, deren Wert für die Vollständigkeit der Systemexploration aus den aufgezeichneten Logfiles extrahiert wurde. Ein vermuteter direkter Effekt der Variable Ausdauer auf die Anzahl der Interaktionen während der Systemexploration, die wiederum auf die Explorationsvollständigkeit wirkt, konnte nicht festgestellt werden ($r = -.09$; $p = .15$).

Der durch die Skala Aufgeschlossenheit gegenüber dem Problemlösen angesprochene Aspekt problemlösebezogener Selbstwirksamkeitserwartung betrifft die Fähigkeit komplexe Probleme zu lösen, mit einer großen Menge an Informationen umzugehen und Dinge schnell zu verstehen. Diese Facette der Selbstwirksamkeitserwartung scheint hingegen einen überwiegend positiven Einfluss auf die Problemlöseleistungen zu haben, wobei die Höhe der Effekte geringer ausfällt als für die Effekte der Facette Ausdauer und nur für den Effekt auf die Exploration in allgemeinen Problemlöseszenarien signifikant wird.

Um zu prüfen, ob die teilweise erwartungswidrigen Befunde auch ohne den Einbezug der anderen exogenen Variablen stabil bleiben, ist ein entsprechendes Modell analysiert worden, dessen geschätzte Modellparameter in Tabelle 15.2 enthalten sind. Zwar ändern sich die Werte der Pfadkoeffizienten geringfügig, jedoch bleiben die Vorzeichen als auch die ungefähre Größenordnung stabil. Die Beantwortung der Hypothesen $H_{3.7}$ bis $H_{3.10}$, nach denen eine hohe Selbstwirksamkeitserwartung zu einer höheren Leistung im explorierenden und steuernden Umgang mit allgemeinen Problemlöseszenarien und technischen Systemen führt, muss differenzierter erfolgen. Für den Aspekt der Ausdauer bei der Bewältigung neuer Anforderungen müssen die Hypothesen $H_{3.7}$ bis $H_{3.10}$ verworfen werden, denn es zeigt sich ein negativer Zusammenhang zwischen dieser und allen endogenen Variablen. Die Gründe für den gegensätzlich eingetretenen Zusammenhang müssen in Kapitel 16 diskutiert werden.

Für den Aspekt der Aufgeschlossenheit gegenüber dem Problemlösen kann ausschließlich die Hypothese H_{3.9} beibehalten werden. Auch wenn fast alle Effekte dieser Selbstwirksamkeitsfacette positiv sind, wird in dem Gesamtmodell (Abbildung 15.1) lediglich der Pfadkoeffizient zur Exploration der allgemeinen Problemlösefähigkeit signifikant.

Tabelle 15.2: Modellparameter des Teilmodells für den Einfluss von Selbstwirksamkeitserwartung auf allgemeines und technisches Problemlösen (Modellgütemaße: $X^2 = 874.48$; $p < .001$; $df = 727$; $X^2/df = 1.20$; $CFI = .96$; $RMSEA = .04$ [.03, .05]; $WRMR = .82$) ($n = 77$)

Hypothese	Strukturmodell	endogene latente Variable(n)	exogene Variable(n)	std. Regressionskoeffizient	SE	p	R ²
3.6 - 3.9	SWE --> technisches Problemlösen, allgemeines Problemlösen	technische Systemexploration	Ausdauer	-.21	.01	.12	.06
			Aufgeschlossenheit	-.06	.01	.66	
		technische Systemsteuerung	Ausdauer	-.44	.02	.02	.17
			Aufgeschlossenheit	.35	.02	.03	
		allgemeines Problemlösen Exploration	Ausdauer	-.43	.02	.00	.16
			Aufgeschlossenheit	.35	.02	.01	
		allgemeines Problemlösen Steuerung	Ausdauer	-.37	.02	.02	.11
			Aufgeschlossenheit	.22	.02	.14	

Anmerkungen: n = Stichprobengröße; SE = Standardfehler; p = p-Wert bei zweiseitigem Test und der Wahrscheinlichkeit von 5 %, einen Fehler 1. Art zu machen; R^2 = Kommunalität; SWE = Selbstwirksamkeitserwartung

Ein weiteres Konstrukt, von dem ein positiver Einfluss auf das Problemlösen im allgemeinen und im technischen Kontext angenommen wurde, ist der Spaß am Denken bzw. das Bedürfnis nach kognitiven Herausforderungen, das durch die Skala Need for Cognition erfasst wurde. Weil die Beantwortung der Skalenitems anhand der subjektiven Einschätzung der Probanden erfolgt, wird in Hypothese H_{3.11} ein Zusammenhang zwischen den Konstrukten Need for Cognition und Selbstwirksamkeitserwartung erwartet. Die signifikant positiven Korrelationen zwischen den manifesten exogenen Variablen Need for Cognition und der Ausdauer und der Aufgeschlossenheit gegenüber dem Problemlösen bestätigen die Hypothese H_{3.11}. So kann die Höhe des bivariaten Zusammenhangs zwischen Need for Cognition und Ausdauer mit $r = .48$ ($p < .001$) und zwischen Need for Cognition und Aufgeschlossenheit gegenüber Problemlösen mit $r = .51$ ($p < .001$) beziffert werden. Der in den Hypothesen H_{3.12} bis H_{3.15} angenommene direkte Einfluss von Need for Cognition auf die Problemlöseleistungen konnte jedoch nicht bestätigt werden. Die Höhe der Pfadkoeffizienten bleiben unter bzw. gleich dem Wert von $\beta = .11$ und werden nicht signifikant. Auch die indirekten Effekte von Need for Cognition über die Ausdauer und die Aufgeschlossenheit gegenüber dem Problemlösen auf die Problemlöseleistungen sind sehr klein (der indirekte Effekt von Need for Cognition über die Aufgeschlossenheit gegenüber dem Problemlösen auf die technische Systemexploration beispielsweise ist $\beta = .11$; $p = .13$).

Ebenfalls keinen Einfluss auf den explorierenden und den steuernden Umgang mit technischen Alltagsgeräten konnte für das Vorwissen festgestellt werden. Weder die Nutzungshäufigkeit von Informations- und Kommunikationstechnologien (ICT) noch die bereits gemachten Erfahrungen der Probanden in der Nutzung der spezifischen Geräte beeinflussen die Leistungen der Probanden in der Systemexploration und -steuerung. Die Hypothesen H_{3.16} und H_{3.17}, nach denen ein höheres Vorwissen zu einer besseren Leistung im explorierenden und steuernden Systemumgang führt, muss demnach verworfen werden. Teilweise bestätigen lässt sich hingegen die Hypothese H_{3.18}, nach der ein höheres Vorwissen

mit einer höheren Selbstwirksamkeitserwartung einhergeht. Der bivariate Zusammenhang zwischen der Skala Aufgeschlossenheit gegenüber dem Problemlösen und dem gerätespezifischen Vorwissen beträgt hier $r = .29$ ($p = .02$) und zwischen der Skala Aufgeschlossenheit gegenüber dem Problemlösen und der ICT-Nutzung $r = .25$ ($p = .04$). Keinen Zusammenhang gibt es zwischen dem Vorwissen und der Skala Ausdauer.

Insgesamt können mit den berücksichtigten exogenen Variablen 37 % Varianz der Systemexploration und 66 % Varianz der Systemsteuerung erklärt werden. Hierbei trägt die Variable fluide Intelligenz den jeweils größten Beitrag ($R^2 = .30$ bei technischer Systemexploration und $R^2 = .48$ bei technischer Systemsteuerung). Im Gegensatz zum allgemeinen Problemlösen, bei dem mindestens 75 % der Varianz aufgeklärt werden kann (bis zu 66 % durch fluide Intelligenz), verbleibt insbesondere bei der technischen Systemexploration noch ein Anteil von über 50 %, der nicht durch die berücksichtigten exogenen Variablen erklärt werden kann.

Ein hier noch einmal aufgegriffener Validitätsaspekt, der eigentlich bereits in Kapitel 14 analysiert wurde, ist der der Dimensionalität technischer Problemlösekompetenz. In dem dort ohne Berücksichtigung von Einflüssen analysierten Modell wurde eine latente Korrelation von $r = .74$ zwischen Systemexploration und Systemsteuerung berichtet. Unter Hinzunahme der latenten Variablen Exploration und Steuerung in allgemeinen Problemlöseszenarien stieg die Höhe des Zusammenhangs zwischen der technischen Systemexploration und -steuerung auf $r = .86$. Einen ähnlich hohen Zusammenhang ließ sich für die Exploration und Steuerung in allgemeinen Problemlöseszenarien feststellen ($r = .85$). Die Integration aller anderen Variablen in das Gesamtmodell (Abbildung 15.1) führt zu einer weiteren Unterschiedsreduktion zwischen der Exploration und Steuerung technischer Alltagsgeräte. Die Korrelation zwischen den verbleibenden Residuen beider Variablen beträgt hier sogar $r = .94$ ($p = .01$), was darauf hindeutet, dass der Unterschied zwischen den beiden Subdimensionen technischer Problemlösekompetenz durch die in dem Modell enthaltenen exogenen Variablen erklärt werden kann. Um die maßgeblichen Variablen identifizieren zu können, wurden über das Gesamtmodell hinaus Teilmodelle analysiert, die jeweils nur eine bzw. zwei exogene Variablen enthalten.

Ein erstes Teilmodell untersucht den ausschließlichen Einfluss fluider Intelligenz auf die endogenen Variablen (Abbildung 15.2). Hierbei zeigt sich, dass der Zusammenhang zwischen dem explorierenden und dem steuernden Systemumgang im Gegensatz zu einem Modell, das den Einfluss fluider Intelligenz unberücksichtigt lässt, auf $r = .80$ ($p = .01$) sinkt. Das bedeutet, ein Teil der Gemeinsamkeit zwischen den beiden Umgangsformen lässt sich auf fluide Intelligenz zurückführen, so dass die Höhe der Korrelation beider Residualvariablen von $r = .86$ auf $.80$ verringert wird. Der gemeinsame Anteil zwischen dem Explorieren und Steuern allgemeiner Problemlöseszenarien, der sich auf Intelligenz zurückführen lässt, ist hier weitaus größer. Die Korrelation zwischen den Subdimensionen allgemeiner Problemlösefähigkeit sinkt bei Integration fluider Intelligenz von ursprünglich $r = .85$ auf $.61$ ($p = .02$). Besonders aufschlussreich erscheint der Befund, dass ausschließlich die Variable Intelligenz für die Gemeinsamkeit zwischen dem technischen und dem allgemeinen Problemlösen ursächlich ist. Der zwischen beiden Konstrukten angenommene bivariate Zusammenhang wird bei Integration fluider Intelligenz deutlich geringer und nicht mehr signifikant. Zwischen der

Exploration in allgemeinen Problemlöseszenarien und der technischen Systemexploration beträgt der Korrelationskoeffizient nur noch $r = .28$ ($p = .20$), zwischen der Steuerung in allgemeinen Problemlöseszenarien und der technischen Systemsteuerung noch $r = .62$ ($p = .052$).

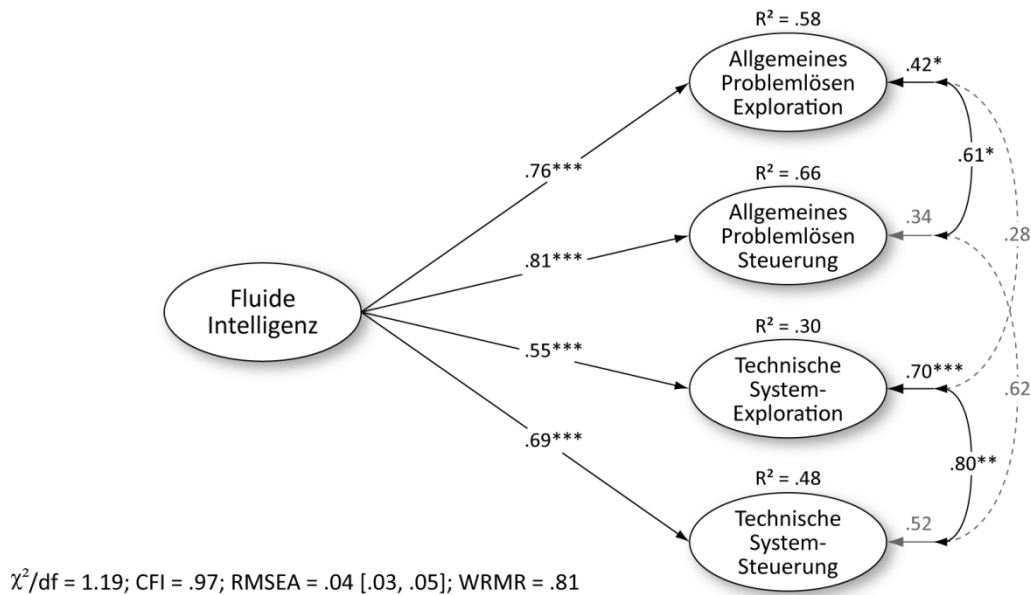
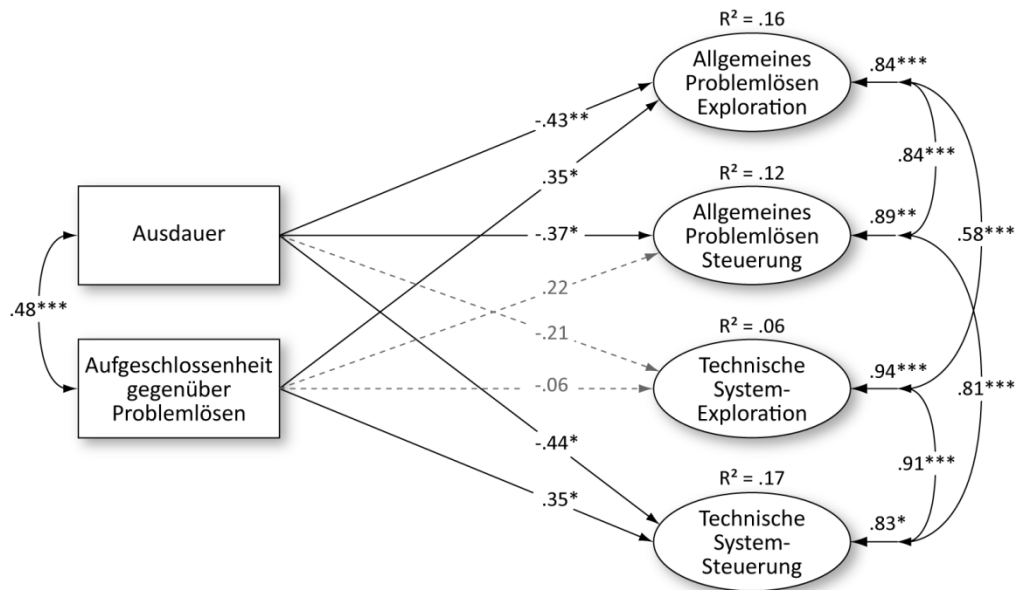


Abbildung 15.2: Einfluss fluider Intelligenz auf die Zusammenhänge zwischen Exploration und Steuerung im Umgang mit technischen Systemen und allgemeinen Problemlöseszenarien sowie auf den Zusammenhang zwischen technischem und allgemeinem Problemlösen ($n = 77$); getrichelt dargestellte Kovarianzen sind nicht signifikant; * $p < .05$; ** $p < .01$; *** $p < .001$; R^2 = Kommunalität

Ein zweites Teilmodell analysiert den Einfluss, den die Variablen der Selbstwirksamkeitserwartung, unabhängig von den anderen im Gesamtmodell enthaltenen Variablen auf den explorierenden und steuernden Umgang mit allgemeinen Problemlöseszenarien und technischen Systemen haben (Abbildung 15.3). Die Ergebnisse dieses Modells zeigen deutlich, dass die manifesten Variablen Ausdauer und Aufgeschlossenheit gegenüber dem Problemlösen einen großen Teil des Unterschiedes zwischen den Dimensionen Systemexploration und Systemsteuerung ausmachen. Bei Berücksichtigung der Selbstwirksamkeitsvariablen korrelieren die verbleibenden Residuen der beiden latenten Variablen zu $r = .91$ ($p < .001$). Statistisch begründet werden kann dieser Befund durch die unterschiedlichen Einflüsse der beiden Selbstwirksamkeitserwartungsvariablen auf die Systemexploration und die Systemsteuerung. Die nach Kontrolle der Selbstwirksamkeitserwartung verbleibende Varianz in den beiden Subdimensionen scheinen eine gemeinsame Ursache zu haben, die in dieser Studie nicht berücksichtigt wird und dafür spricht, dass beide Dimensionen eigentlich dasselbe erfassen. Der Validitätsaspekt der Dimensionalität technischer Problemlösekompetenz muss vor diesem Hintergrund noch einmal diskutiert werden (Abschnitt 16). Da die Effekte der beiden Selbstwirksamkeitserwartungsvariablen auf die beiden Dimensionen der allgemeinen Problemlösefähigkeit ähnlich hoch sind, lässt sich eine Veränderung in der Korrelation der zugehörigen Residualvariablen bei Berücksichtigung der Selbstwirksamkeitserwartung hier nicht beobachten ($r = .84$; $p < .001$).



$\chi^2/df = 1.20$; CFI = .96; RMSEA = .04 [.03, .05]; WRMR = .82

Abbildung 15.3: Einfluss von Selbstwirksamkeitserwartungen auf die Zusammenhänge zwischen Exploration und Steuerung im Umgang mit technischen Systemen und allgemeinen Problemlöseszenarien sowie auf den Zusammenhang zwischen technischem und allgemeinem Problemlösen ($n = 77$); getrichelt dargestellte Effekte sind nicht signifikant; $*p < .05$; $**p < .01$; $***p < .001$; $R^2 = \text{Kommunalität}$

Teil V: Diskussion

16 Problemlösender Umgang mit technischen Alltagsgeräten – ein eigenständiges Konstrukt?

Die in den vorangegangenen Kapiteln präsentierten Ergebnisse aus der empirischen Prüfung des Kompetenzkonstrukts werden in dem folgenden Kapitel näher diskutiert. Dazu fasst der nächste Abschnitt die Befunde anhand der in Kapitel 7 formulierten Fragestellungen zusammen, bevor sie unter Einbezug der in den Kapiteln 2 - 5 betrachteten Theorie inhaltlich interpretiert werden. Diskutiert werden hier vor allem erwartungswidrige Befunde sowie die Generalisierbarkeit der Ergebnisse. Abschließend wird die Bedeutung der Resultate für den Forschungsbereich der allgemeinbildenden Technikdidaktik sowie die Relevanz weiterer Forschungsarbeiten hervorgehoben.

16.1 Zusammenfassende Ergebnisbetrachtungen

Die Schwierigkeit im Umgang mit technischen Alltagsgeräten variiert in Abhängigkeit von dem vorliegenden technischen System und der Aufgabe, die mit dem jeweiligen Gerät erledigt werden soll. Schwierig zu bedienende Systeme können nur von wenigen Personen exploriert und gesteuert werden. Für die Exploration und die Steuerung dieser ist außerdem eine längere Bearbeitungszeit nötig, als für weniger schwierige Systeme. Diese Schwierigkeitsunterschiede galt es in dem, zur empirischen Prüfung des Kompetenzkonstrukts konstruierten Testinstrument umzusetzen. Zu diesem Zweck wurden aus der Literatur Problemmerkmale extrahiert, für die ein Einfluss auf die Schwierigkeit im Umgang mit technischen Geräten anzunehmen ist. Da die Schwierigkeitseinflüsse der angenommenen Merkmale *Komplexität*, *Vernetztheit*, *Dynamik* und *Intransparenz* auch theoretisch begründet werden können, dient die Überprüfung der vorab formulierten Einflüsse als Beitrag zur Konstruktvalidierung. Diesen Validitätsaspekt berücksichtigt die folgende Fragestellung aus Kapitel 7:

Fragestellung 1: Beeinflussen die angenommenen schwierigkeitsbestimmenden System- und Situationsmerkmale tatsächlich die Schwierigkeit im Umgang mit technischen Systemen?

Für die Beantwortung der Fragestellung wurden die schwierigkeitsbestimmenden Merkmale zunächst in System- und Situationsmerkmale unterteilt, da sie weitgehend unabhängig voneinander variiert und untersucht werden können. Die Operationalisierung der zunächst noch abstrakten Problemmerkmale erfolgte für die Systemebene getrennt anhand der Systeminterfacekomponenten Hardware und Software (Kapitel 10.1.2). Der Schwierigkeitseinfluss konnte jedoch aufgrund nicht erfüllter Voraussetzungen (linearer Zusammenhang zwischen den Variablen und der empirischen Schwierigkeit, Ausschluss von Multikollinearität, normalverteilte und homogen streuende Residuen) nicht für alle der messbaren Variablen, in die die einzelnen Merkmale übersetzt wurden, überprüft werden. Das Merkmal Komplexität eines technischen Systems wird durch die Variablen *Anzahl der Bedienelemente*, *Menütiefe*, *Anzahl sichtbarer Zustände* sowie die *Anzahl veränderbarer Variablen* repräsentiert, das Merkmal der Vernetztheit durch die Variable *Anzahl der Bedingungen* und das Merkmal Transparenz durch die Variablen *Anzahl unterschiedlicher Bedienelementarten* sowie *Menüverständlichkeit*.

Das mit diesen sieben Prädiktorvariablen gebildete Regressionsmodell konnte 95 % (korrigiert 90 %) der empirischen Schwierigkeit im explorierenden Umgang (gemessen durch die mittlere Explorationsvollständigkeit) erklären. Ein tatsächlich signifikanter Einfluss auf die Explorationsschwierigkeit ist allerdings lediglich für die Variablen der Merkmale Komplexität und Transparenz festzustellen. Verglichen an den standardisierten Regressionskoeffizienten zeigt sich für beide Merkmale ein betragsmäßig gleich hoher Einfluss auf die Explorationsschwierigkeit, der sich lediglich in dem Vorzeichen unterscheidet. Während das Merkmal Komplexität den explorierenden Umgang mit den technischen Systemen erschwert (Bestätigung von $H_{1.1}$), sorgt das Merkmal Transparenz für eine Verringerung der Explorationsschwierigkeit (Bestätigung von $H_{1.4}$). Keinen signifikanten Einfluss übte das Merkmal Vernetztheit auf die Schwierigkeit der Systemexploration aus (Ablehnung von $H_{1.2}$) und nicht zu prüfen war der Einfluss des Merkmals Dynamik ($H_{1.3}$).

Die Schwierigkeit ein technisches System zu steuern, wird neben den bereits für die Schwierigkeit der Systemexploration operationalisierten Systemmerkmalen außerdem durch das konkrete Steuerungsziel (die Situationsmerkmale) beeinflusst. Die Operationalisierung der Situationsmerkmale in messbare Variablen erfolgte ausschließlich für die Merkmale Komplexität und Vernetztheit. Als Prädiktoren in das Regressionsmodell zur Erklärung der Steuerungsschwierigkeit wurden schließlich die Variablen *Anzahl unterschiedlicher Lösungsschritte* und *zu berücksichtigende Mehrfachbelegung von Bedienelementen* neben der Explorationsvollständigkeit und der steuerungsspezifischen Exploration aufgenommen. Da sich die Schwierigkeit nicht nur in dem Erreichen des Steuerungsziels, sondern auch in den dafür benötigten Bearbeitungsschritten zeigt, wurde der Einfluss der Merkmale einmal auf die empirische Steuerungsschwierigkeit, die nur das Ergebnis berücksichtigt, und einmal auf die empirische Steuerungsschwierigkeit, die die Anzahl benötigter Interaktionen berücksichtigt, analysiert. In beiden Modellen werden die Regressionskoeffizienten der Merkmale Komplexität und Vernetztheit zwar negativ und weisen damit auf eine Erhöhung der Steuerungsschwierigkeit hin, sie werden jedoch nicht signifikant. Die Hypothesen $H_{1.5}$ und $H_{1.6}$ sind daher abzulehnen. Lediglich in dem Modell zur Vorhersage der Schwierigkeit im steuernden Umgang mit technischen Systemen, das die zur Lösung benötigte Interaktionsanzahl berücksichtigt, kann der Explorationsvollständigkeit ein signifikanter Einfluss unterstellt werden. In diesem Modell lassen sich jedoch lediglich 54 % (korrigiert 41 %) der empirischen Schwierigkeit erklären.

Sowohl der Testkonstruktion als auch der Operationalisierung der genannten schwierigkeitsbestimmenden Merkmale liegt die Vorstellung zugrunde, dass sich die Art und Weise, wie eine Person mit einem technischen System interagiert, anhand des zu erreichenden Ziels unterscheidet. Die Systemexploration dient vor allem dem Erwerb von Eingriffswissen über ein bis dahin unbekanntes Alltagsgerät, während die Steuerung eines technischen Gerätes zum Ziel hat bereits vorhandenes Wissen zur Änderung von Systemzuständen anzuwenden. Die unterschiedliche Verortung der Zielsetzungen innerhalb (Wissenserwerb) und außerhalb (Wissensanwendung) einer Person spiegelt sich in der Konstruktstruktur wider, indem die Kompetenz im Umgang mit technischen Geräten in die Subdimensionen *Systemexploration* und *Systemsteuerung* differenziert wurden. Aufgrund des postulierten Kontexteinflusses sollte sich der problemlösende Umgang mit technischen Systemen außerdem von dem allgemeiner Problemszenarien unterscheiden. Ein weiterer in dieser

Arbeit untersuchter Aspekt von Konstruktvalidität betrifft damit die theoretisch angenommene Konstruktstruktur:

Fragestellung 2: Lässt sich die logisch-inhaltliche Analyse des Konstruktes (Konstruktstruktur) empirisch bestätigen?

Ob sich eine zweidimensionale Konstruktstruktur mit den Subdimensionen Systemexploration und Systemsteuerung auch empirisch zeigen lässt, wurde mithilfe einer konfirmatorischen Faktorenanalyse überprüft. Die Spezifizierung der Messmodelle erfolgte dabei theoriegeleitet, das heißt als Indikatoren für die Systemexploration wurden die jeweiligen Explorationsaufgaben der 15 technischen Systeme und als Indikatoren für die Systemsteuerung die jeweiligen Steuerungsaufgaben festgelegt. Getestet wurden drei alternative, sich nur durch Restriktionen unterscheidende Modelle. Das Modell, das eine eindimensionale Struktur des problemlösenden Umgangs mit technischen Alltagsgeräten unterstellt, wies einen schlechteren Modellfit auf, als die eine zweidimensionale Struktur annehmenden Modelle. Ein statistischer Vergleich der Modelle unterstützt die Annahme, nach der die Systemexploration und die Systemsteuerung zwei voneinander abgrenzbare Problembereiche im Umgang mit technischen Alltagsgeräten sind (Bestätigung von H_{2.1}). Die Parameterschätzung des am wenigsten restringierten Modells zeigt außerdem, dass die latenten Variablen Systemexploration und Systemsteuerung hoch korrelieren, womit außerdem die Aussage, nach der die Explorationsleistung einen Einfluss auf die Steuerungsleistung hat zutreffend ist (Bestätigung von H_{2.2}).

Die Hinzunahme der Subdimensionen Wissen nach der Exploration und Steuerung des Konstrukts allgemeine Problemlösefähigkeit ermöglichte die Prüfung diskriminanter Validität. Ein zweidimensionales Modell, das nicht zwischen allgemeinem und technischem Problemlösen, sondern lediglich die Interaktionsformen Exploration und Steuerung unterscheidet wurde gegen ein vierdimensionales Modell, das sowohl zwischen allgemeiner und technischer Exploration als auch zwischen allgemeiner und technischer Steuerung unterscheidet, getestet. Der Test fiel zugunsten des vierdimensionalen Modells aus, so dass anzunehmen bleibt, dass sich die beiden Konstrukte trotz der zum Teil hohen Zusammenhänge empirisch voneinander trennen lassen (Bestätigung von H_{2.3}).

Eine darüber hinaus gehende Bestimmung und Eingrenzung des Konstruktes erforderte außerdem die Betrachtung von Anknüpfungspunkten und Überlappungen zu weiteren Konstrukten. Für die Verortung technischer Problemlösekompetenz in dem, die Beziehungen zwischen diesen Konstrukten darstellenden, nomologischen Netzwerk wurden diejenigen Konstrukte herangezogen, von denen auf Grundlage theoretischer und empirischer Vorarbeiten ein Einfluss anzunehmen ist (Kapitel 6.6). Neben dem bereits betrachteten Konstrukt allgemeine Problemlösefähigkeit wurden für die Konstrukte fluide Intelligenz, Selbstwirksamkeitserwartungen, Need for Cognition sowie Vorwissen bivariate und kausale Beziehungen in einem Strukturgleichungsmodell derart spezifiziert, dass sie das Theoriemodell wiedergeben. Im Rahmen der Konstruktvalidierung erfolgt die Überprüfung des postulierten nomologischen Netzwerkes:

Fragestellung 3: Lässt sich die Verortung des Konstruktes in dem nomologischen Netzwerk empirisch bestätigen?

Bereits die ersten Zusammenhangsbetrachtungen der manifesten Variablen (Kapitel 11.8) deuteten auf einen Einfluss fluider Intelligenz auf die Leistungen im explorierenden und steuernden Umgang sowohl mit technischen Systemen als auch mit allgemeinen Problemlöseszenarien hin. Die latenten Korrelationen zwischen den Konstrukten im analysierten Strukturgleichungsmodell fielen erwartungskonform noch etwas höher aus (Bestätigung von H_{3.1} – H_{3.4}). In dem Gesamtmodell und einem, nur die drei Konstrukte fluide Intelligenz, allgemeine Problemlösefähigkeit und technische Problemlösekompetenz betrachtenden Modell (Kapitel 15) zeigte sich außerdem ein stärkerer Zusammenhang zwischen Intelligenz und allgemeiner Problemlösefähigkeit als zwischen Intelligenz und technischer Problemlösekompetenz (Bestätigung von H_{3.5}). Die Berücksichtigung fluider Intelligenz führte außerdem zu einer Verringerung des Zusammenhangs zwischen allgemeinem und technischem Explorieren sowie zwischen allgemeinem und technischem Steuern, womit sich die Annahme der Separierbarkeit beider Konstrukte erneut bestätigen ließe (Bestätigung von H_{2.3}). Der erwartete Zusammenhang zwischen den Einflussvariablen Intelligenz und Vorwissen ließ sich jedoch nur für das gerätespezifische Vorwissen zeigen, keinen Einfluss übt Intelligenz hingegen auf die Nutzung von Informations- und Kommunikationstechnologien aus (partielle Bestätigung von H_{3.6}).

Die auf das Problemlösen bezogene Selbstwirksamkeitserwartung wurde durch die Variablen Ausdauer und Aufgeschlossenheit gegenüber dem Problemlösen erfasst. Die Befunde zur Prüfung des Einflusses, den diese Variablen auf die Subdimensionen des allgemeinen und technischen Problemlösens haben, fielen sehr unterschiedlich aus. Entgegen den Erwartungen zeigte sich ein durchweg negativer und größtenteils signifikanter Einfluss der Variable Ausdauer auf alle Problemlöseleistungen. Personen, die sich selbst als besonders ausdauernd bei der Bewältigung neuer Anforderungen einschätzten, erzielten tendenziell eher geringe Leistungen. Anders sehen die Befunde für die Selbstwirksamkeitsfacette Aufgeschlossenheit gegenüber dem Problemlösen aus; diese scheint einen überwiegend positiven Einfluss auf die Problemlöseleistungen zu haben, wobei die Effekte zu gering sind, um in der vorliegenden Untersuchung signifikant zu werden (Ablehnung von H_{3.7} – H_{3.10}).

Die aus methodischen Gründen angenommene Überlappung der Konstrukte Selbstwirksamkeitserwartung und Need for Cognition (Erfassung von Need for Cognition anhand einer Selbsteinschätzungsskala) ließ sich auch empirisch nachweisen. Die Korrelation beider Facetten der Selbstwirksamkeitserwartung mit dem Konstrukt Need for Cognition fiel mittelhoch aus (Bestätigung von H_{3.11}). Nicht nachweisen ließ sich jedoch der angenommene positive Effekt von Need for Cognition auf die Problemlöseleistungen (Ablehnung von H_{3.12} – H_{3.15}).

Die bereits vorhandene Erfahrung im Umgang mit den einzelnen technischen Geräten sowie die Nutzungshäufigkeit von Informations- und Kommunikationstechnologien beeinflusste wider Erwarten weder das Explorationsverhalten noch die Steuerungsleistung im Umgang mit den technischen Systemen (Ablehnung von H_{3.16} und H_{3.17}). Beide Variablen sollten das Vorwissen bzw. die Erfahrung erfassen, die Personen bereits in die Problemsituationen einbringen, ein Zusammenhang zwischen den beiden Variablen war zwar festzustellen, dieser

ist allerdings nicht sehr groß. Ähnlich hoch fiel auch der Zusammenhang zwischen den beiden Variablen des Vorwissens und der Variable Aufgeschlossenheit gegenüber dem Problemlösen aus. Für die andere Facette der Selbstwirksamkeitserwartung Ausdauer ließ sich wiederum kein Zusammenhang erkennen (partielle Bestätigung von H_{3.18}).

16.2 Theoriebezogene Ergebnisinterpretation

Um die Ergebnisse nun nachvollziehbar in die betrachtete Theorie einordnen zu können, werden die aus ihr abgeleiteten Erwartungen sowie die Ergebnisse früherer Forschungsarbeiten zum Problemlösen zunächst noch einmal aufgegriffen, bevor die Besprechung der Befunde dieser Arbeit folgt. Ursachen für erwartungswidrige Ergebnisse werden ebenso diskutiert wie mögliche Folgen und Einschränkungen der Untersuchung. Der Verlauf, in dem die Ergebnisse interpretiert werden, orientiert sich an der Reihenfolge der Fragestellungen, die im Abschnitt 7 formuliert wurden. Die in den Fragestellungen 1 & 2 angesprochene Konstruktrepräsentation und Konstruktstruktur beziehen sich vor allem auf die innere Validität des untersuchten Problemlösekonstrukts. Die entsprechenden Hypothesen wurden dabei aus der im Kapitel 2.4.2 beschriebenen kognitionspsychologischen Theorie abgeleitet. Der in der Fragestellung 3 angesprochene Validitätsaspekt bezieht sich auf die äußere Validität, also der Beziehungen des Problemlösekonstruktes zu anderen Konstrukten in dem im Kapitel 6.6 postulierten nomologischen Netzwerk. Diese Zusammenhänge wurden separat nach kontextübergreifenden und kontextuellen Einflussfaktoren im Kapitel 5 beschrieben und so auch hier diskutiert.

Interpretation der Ergebnisse vor dem Hintergrund kognitionspsychologischer Theorien

Der problemlösende Umgang mit technischen Alltagsgeräten erfordert Wissen darüber, welche Eingriffe in das technische System zu welchen Reaktionen führen und in welcher Reihenfolge diese Eingriffe vorgenommen werden müssen, um konkrete Steuerungsziele zu erreichen. Dieses Eingriffswissen lässt sich im Rahmen der ACT-R-Theorie von Anderson & Lebiere (1998a) als prozedurales Wissen beschreiben, das in Form von Produktionen repräsentiert ist (Kapitel 2.4.2.1). Zu Beginn des Wissenserwerbs enthalten diese Produktionen Leerstellen, so dass für die Handlungsausführung fehlende Aspekte aus dem deklarativen Gedächtnis repräsentiert werden müssen (ebd., Wallach & Lebiere, 1998, S. 94). Hierzu muss das in Chunks repräsentierte Faktenwissen aktiviert werden, an dem neues Wissen angebunden werden kann. Das Arbeitsgedächtnis, das den Teil des Wissens darstellt, der aktuell aktiviert ist, ist in seiner Kapazität jedoch begrenzt (Sweller, 1988, S. 265).

Aus der angenommenen Struktur des kognitiven Systems, lassen sich Erwartungen an die Schwierigkeit im problemlösenden Umgang mit technischen Systemen formulieren. *Komplexität* als quantitatives Maß wirkt deshalb schwierigkeitsbestimmend, weil die Anzahl der Produktionsschritte zum Erwerb von Wissen und dem Anwenden von Wissen erhöht ist (Daily, Lovett & Reder, 2001, S. 319). Je komplexer ein System, desto mehr Informationen müssen wahrgenommen und zunächst im deklarativen Gedächtnis an vorhandene Chunks angebunden werden. Änderungen an *vernetzten Variablen* führen zu Änderungen an anderen Stellen in dem System. Diese zusätzlichen Veränderungen müssen wiederum wahrgenommen werden, wodurch die Repräsentation größerer Bereiche des deklarativen Wissens erforderlich ist. Neben den Systemvariablen, die als zu verarbeitende Informationen die

Arbeitsgedächtnisressourcen belasten, stellen die vorhandenen Relationen zwischen ihnen weitere Anforderungen dar. Wenn die Wahrnehmung von Veränderungen durch das ausgeprägte Merkmal der *Intransparenz* behindert ist, kann keine Aktivierung erforderlichen deklarativen Wissen und damit auch keine Einbindung in den vorhandenen Leerstellen der Produktionen erfolgen. In der Folge bleiben die ausgeführten Aktionen fehlerhaft und die Steuerung des entsprechenden Systems ist erschwert. *Dynamische* Systeme belasten das Arbeitsgedächtnis des kognitiven Systems, weil ständig Zwischenergebnisse abrufbar bleiben, also Chunks länger aktiviert bleiben müssen, obwohl bereits neues Wissen repräsentiert werden muss. Die Wahrnehmung des Zusammenhangs zwischen einer getätigten Aktion und der Reaktion des Systems wird erschwert.

Neben diesen theoretischen Erwartungen sprechen auch empirische Befunde für den schwierigkeitsbestimmenden Einfluss der Merkmale Komplexität, Vernetztheit, Intransparenz und Dynamik auf den Problemlöseprozess und damit auf die Problemlöseleistungen. So konnte Hussy (1984b) knapp die Hälfte der Varianz in der Problemschwierigkeit durch die Merkmale Variablenanzahl (Komplexität), Transparenz und Vernetzung erklären, wobei sich die einzelnen Merkmale in ihrem Einfluss auf die Schwierigkeit nicht stark voneinander unterscheiden. Funke (1990), der den Grad an Eigendynamiken und Nebenwirkungen variierte, konnte zeigen, dass diese einen negativen linearen Effekt auf die Güte des erworbenen Wissens und vermittelt darüber einen indirekten Effekt auf die Güte der Systemsteuerung haben. Außerdem variierte er die Anzahl zu kontrollierender Variablen und kommt zu dem Ergebnis, dass sich die Systeme besser steuern ließen, wenn nur wenige zu kontrollierende Variablen enthalten sind. Bezüglich der Merkmale Eigendynamik und Vernetzung können auch Kluge (2004) sowie Greiff, Krkovic & Nagy (2014) einen schwierigkeitsbestimmenden Einfluss auf die Problemlöseleistungen im Umgang mit dynamischen Problemszenarien feststellen.

Alle vorgestellten Arbeiten beschränken sich jedoch auf die Untersuchung der Schwierigkeit im Umgang mit Systemen, die auf Basis linearer Gleichungssysteme entworfen wurden, wenig komplex und vor allem wenig realitätsnah sind. Ziel dieser Arbeit war es hingegen, einen computerbasierten Test zu entwickeln, der die gleichen Kompetenzen erfasst, wie sie in den entsprechenden realen Anforderungssituationen nötig sind. Mit der Verwendung finiter Automaten als formales Modell (Kapitel 4.3) zur Entwicklung der Computersimulationen gelang es, reale technische Systeme möglichst gut abzubilden. Die Berücksichtigung der schwierigkeitsbestimmenden Merkmale erforderte hier eine auf die technischen Systeme bezogene Operationalisierung (Kapitel 6.4), die nicht mit der in den genannten Arbeiten verwendeten übereinstimmt. Ein Vergleich der Befunde ist daher nur bedingt möglich.

Zeigen konnte diese Arbeit den deutlichen Einfluss des anhand der quantitativen Informationsmenge operationalisierten Merkmals der Komplexität auf das Verhalten der Systemexploration (Kapitel 13.3.1). Je mehr Funktionen in einem technischen Gerät über die Hardware und Software des Bedienerinterfaces aufgerufen und Variablen verändert werden können, desto unvollständiger gelingt eine Systemexploration. Sind diese Funktionen und Variablen außerdem für die Nutzer nicht transparent, weil beispielsweise Bedienelemente oder Menüeinträge mit unverständlichen Symbolen beschriftet sind, wird die Systemexploration weiter erschwert.

Der Grund für den nicht sichtbaren Effekt des Merkmals Dynamik auf die Schwierigkeit im Umgang mit den technischen Systemen ist in der unzureichenden Ausprägung des Merkmals in den simulierten Geräten zu finden. Die Implementierung von Zeitverzögerungen erfolgte in einigen Geräten derart, dass die Systeme auf vereinzelt Bedienhandlungen erst nach einer Zeit von ein paar Sekunden reagierten. Da diese Verzögerungszeit relativ kurz ist und sie außerdem nicht bei jeder Bedienhandlung eintritt, beeinflusst die Variable Zeitverzögerung des Merkmals Dynamik die Schwierigkeit im Umgang mit den Geräten kaum. Eigendynamiken wurden bei der Entwicklung zweier Systeme berücksichtigt, indem der längere Verbleib in einem Systemzustand dazu führt, dass sich das System auf den Ausgangszustand zurücksetzt. Da die Probanden die dort implementierte Zeitbegrenzung nicht überschritten, konnte das Merkmal nicht wirksam und ein schwierigkeitsbestimmender Einfluss nicht nachgewiesen werden.

Für die Ablehnung der Hypothese ($H_{1.2}$), nach der die Vernetzung eines technischen Gerätes dessen Exploration erschwert, können mehrere Gründe angeführt werden. Das Merkmal Vernetztheit wurde in dieser Arbeit auf Hardwareebene durch die *Mehrfachbelegung von Bedienelementen* und auf Softwareebene durch *vorhandene Bedingungen* operationalisiert (Kapitel 10.1.2). Der Einfluss, den die Mehrfachbelegung von Bedienelementen auf die Schwierigkeit im Umgang mit einem technischen Gerät ausübt, wird außerdem durch das Merkmal Intransparenz moderiert. Mehrfachbelegte Bedienelemente, deren Mehrfachbelegung aufgrund fehlender Beschriftungen nicht wahrnehmbar ist, waren schwieriger zu explorieren, als mehrfachbelegte Bedienelemente, die einen Hinweis auf ihre Mehrfachbelegung (durch eine entsprechende Beschriftung oder durch Hinweise im Menü) geben. Der negative Einfluss, der durch vorhandene Variablenbedingungen auf die Schwierigkeit im explorierenden Umgang mit den Systemen beobachtet wurde (das heißt die Items waren bei vorhandenen Variablenbedingungen einfacher zu explorieren), ließ sich erst nach der Analyse der Logfiles begründen. Wenn eine Variablenbedingung vorliegt (d. h. der Wert einer Variable kann nur geändert werden, wenn der Wert einer anderen Variable innerhalb eines bestimmten Wertebereichs liegt) so erhalten die Nutzer einen entsprechenden Hinweis, der wiederum dazu führt, dass die Probanden das System an dieser Stelle vollständiger explorieren. Durch diesen Hinweis erfolgte eine Lenkung der Personen bei der Systemexploration und führte damit zu einer Verringerung der Schwierigkeit. Nicht die Variablenbedingung selbst ist hier das schwierigkeitsbestimmende Merkmal, sondern die Transparenz. Würde ein solcher Hinweis nicht gegeben, so dürfte die Systemexploration und -steuerung wohl erheblich schwieriger werden.

Die sich auf die Explorationsvollständigkeit auswirkenden Systemmerkmale beeinflussen über die Informationen, die während der Systemexploration generiert wurden, außerdem die Schwierigkeit in der Systemsteuerung. Je unvollständiger die vorherige Systemexploration und je geringer damit die generierte Informationsmenge, desto schwieriger ist die anschließende Steuerung der technischen Systeme. Die Bedeutung dieses Einflusses wird noch größer, wenn die Steuerungsschwierigkeit nicht nur anhand der relativen Personenanzahl bewertet wird, die das vorgegebene Steuerungsziel erreichen, sondern auch anhand der für die Lösung benötigten Bearbeitungsschritte. Eine vollständigere Systemexploration erhöht somit nicht nur die Wahrscheinlichkeit, das Steuerungsziel zu erreichen, sondern führt außerdem dazu, dass sich dieses mit weniger Systemeingriffen steuern lässt.

Eine Bewertung, die neben dem erreichten Steuerungsziel außerdem die Anzahl der für die Lösung benötigten Bearbeitungsschritte berücksichtigt, kann nur zwischen Personen differenzieren, die zu einer Lösung gekommen sind bzw. das Steuerungsziel erreicht haben. Eine weitere Unterscheidung zwischen Personen, die ein Item nicht gelöst haben, kann über die so ermittelte Steuerungsleistung nicht erfolgen. Um Personen hinsichtlich ihrer Kompetenz auf einer Niveauskala einordnen zu können, sollte dieses Maß demnach nicht verwendet werden. Diese, die benötigte Interaktionszahl berücksichtigende Steuerungsleistung wurde in der vorliegenden Arbeit ausschließlich zur Analyse schwierigkeitsbestimmender Merkmale verwendet.

Für die neben der Explorationsvollständigkeit vermuteten Situationsmerkmale, die durch die Anzahl zu berücksichtigender Mehrfachbelegungen von Bedienelementen sowie die Anzahl erforderlicher Lösungsschritte operationalisiert wurden, ließ sich zwar ein Effekt auf die Steuerungsschwierigkeit erkennen, der aber vermutlich aufgrund der geringen Anzahl an Items nicht signifikant wurde. Auffällig ist unabhängig davon, dass sich fast die gesamte Varianz in der Explorationsschwierigkeit durch die als schwierigkeitsbestimmend angenommenen Merkmale erklären lassen, aber nur knapp die Hälfte der Varianz in der Steuerungsschwierigkeit. Dieser Befund spricht dafür, dass weitere eventuell subjektiv als schwierig wahrgenommene Merkmale die Anwendung von Wissen im steuernden Umgang mit den technischen Systemen beeinflussen. Für diese Annahme spricht, dass ein ähnlicher Befund bereits von den Autoren Greiff, Krkovic & Nagy (2014) berichtet wurde. Auch sie konnten nahezu die gesamte Schwierigkeit in der Wissenserwerbsphase jedoch nur einen geringen Anteil der Schwierigkeit in der Steuerungsphase durch die Merkmale Vernetztheit und Eigendynamik erklären.

Dass sich die Schwierigkeit im explorierenden und steuernden Umgang mit technischen Systemen nicht gleich gut durch die angenommenen Merkmale erklären lässt, spricht für die Annahme, dass die Systemexploration und die Systemsteuerung voneinander unterscheidbare Interaktionsformen im Umgang mit technischen Alltagsgeräten darstellen. Begründet wurde die Abgrenzbarkeit der beiden Interaktionsformen vorab mit den unterschiedlich vorliegenden Zielen, die einmal innerhalb der Person (Wissenserwerb) und einmal außerhalb der Person (Wissensanwendung) verortet sind (Kapitel 2.4.1). Ohne den Einbezug von Hintergrundvariablen ließ sich diese angenommene zweidimensionale Konstruktstruktur auch tatsächlich in den Daten zeigen. Damit stimmen die eigenen Befunde mit denen von Neubert, Kretzschmar, Wüstenberg et al. (2015) überein, die eine zweidimensionale Struktur bestehend aus Wissenserwerb und Wissensanwendung für die allgemeine Problemlösefähigkeit als zutreffender herausstellen. Obwohl die Autoren die Phase des Wissenserwerbs über die Abfrage des tatsächlich erworbenen Wissens erfassen und nicht wie in dieser Arbeit über den Prozessindikator der Explorationsvollständigkeit, berichten sie einen ähnlich hohen Zusammenhang zwischen den Dimensionen, wie er für den Zusammenhang zwischen der Systemexploration und der Systemsteuerung gezeigt wurde. Dieses Ergebnis spricht dafür, dass die Probanden während der Systemexploration nicht nur Informationen generiert, sondern diese Informationen auch im kognitiven System zu Wissen weiterverarbeitet haben. Wenn sich diese Vermutung durch weitere Forschungsarbeiten bestätigen ließe, so wäre es möglich von der Vollständigkeit der Systemexploration auf das dabei erworbene Wissen zu schließen. Der damit verbundene Vorteil ist der Verzicht auf die

sich an die Systemexploration anschließende Wissensabfrage. Diese Wissensabfrage benötigt wertvolle Testzeit, es besteht die Gefahr des Ratens bei Unsicherheit oder Nichtwissen und es kann zu Motivationsverlusten seitens der Probanden durch die nichterkannte Sinnhaftigkeit dieser Fragen kommen.

Neben den zu beobachtenden Gemeinsamkeiten, die aufgrund der ähnlichen Konstruktstruktur zwischen den beiden Problemlösekonstrukten bestehen, gibt es auch Unterschiede. Die Annahme, dass sich aufgrund des spezifischen Kontextes der Umgang mit technischen Alltagsgeräten von dem Umgang mit nahezu dekontextualisierten Problemszenarien unterscheidet, ließ sich auch durch die bessere Passung eines Modells, das zwischen den einzelnen Konstrukten differenziert, zeigen. Der Unterschied zwischen dem allgemeinen Problemlösen und dem problemlösenden Umgang mit technischen Geräten wird bei Berücksichtigung der Variablen, die einen Einfluss auf beide Konstrukte ausüben noch einmal größer, weshalb die Ergebnisse im nächsten Abschnitt unter Einbezug weiterer Konstrukte betrachtet werden.

Interpretation der Ergebnisse vor dem Hintergrund der angenommenen kontextübergreifenden und kontextuellen Einflussfaktoren

Um das im Problemlöseprozess benötigte Wissen erwerben und auch anwenden zu können, werden allgemeine kognitive Grundfähigkeiten benötigt. Von fluider Intelligenz als die Fähigkeit sich an neue Situationen anzupassen und neue Anforderungen bewältigen zu können (Cattell, 1963) bzw. allgemein als die Fähigkeit zum Lernen (Mack, 1999; Rost, 2009) wird erwartet, dass sie den Problemlöseprozess und die Problemlöseleistung in besonderem Maße beeinflusst (Kapitel 5.1.1). Das problemlösende Verhalten einer Person ist jedoch nicht ausschließlich von Intelligenz abhängig, sondern wird außerdem durch situative Faktoren beeinflusst (Kaufhold, 2006), womit der Kontext, in dem Problemlösen stattfindet, die Höhe des Zusammenhangs zwischen Intelligenz und Problemlöseleistung bestimmt. Im problemlösenden Umgang mit den simulierten technischen Geräten kann eine Person auf ihre bereits erworbenen Erfahrungen im Hinblick auf diese Geräte zurückgreifen. Die nahezu unbekanntes minimalkomplexen Problemszenarien zur Erfassung allgemeiner Problemlösefähigkeit erlauben einen solchen Rückgriff hingegen nicht, da sie in weitgehend fiktiven Kontexten eingebunden sind. Es konnte in dieser Arbeit also davon ausgegangen werden, dass Intelligenz für das Lösen allgemeiner Probleme eine höhere Vorhersagekraft hat als für das Lösen technischer Probleme, bei denen außerdem kontextuelle Faktoren eine Rolle spielen.

Diese theoretischen Annahmen werden gestützt durch empirische Arbeiten von z. B. Wittmann, Süß & Oberauer (1996) und Kersting (1999), in denen nach statistischer Kontrolle der Prädiktoren Intelligenz und Wissen keine statistisch bedeutsamen Zusammenhänge zwischen den einzelnen Leistungen aus der Bearbeitung unterschiedlicher Problemszenarien nachgewiesen werden konnten. In anderen Untersuchungen zum fachspezifischen Problemlösen bei Elektronikern und Kfz-Mechatronikern (Abele, Greiff, Gschwendtner et al., 2012) sowie bei Elektronikern der Automatisierungstechnik (Walker, Link & Nickolaus, 2015) zeigte sich, dass Intelligenz über das Fachwissen hinaus keine Vorhersage auf das Lösen der gegebenen technischen Probleme leistet.

Es wurde nicht erwartet, dass die Problemlöseleistungen im Umgang mit technischen Alltagsgeräten wie in den eben genannten Studien ausschließlich durch das bereits vorhandene Wissen bzw. die erfassten Erfahrungen bestimmt werden und Intelligenz keine bedeutende Rolle spielt. Denn im Gegensatz zu den dort vorgegebenen fachspezifischen Problemstellungen, konnten die Probanden während des Explorierens der technischen Geräte Wissensdefizite ausgleichen, ein Prozess, der wiederum Intelligenz erfordert.

Vor diesem theoretischen und empirischen Hintergrund lassen sich die gefundenen Ergebnisse zur Intelligenzeinwirkung gut einordnen. Intelligenz hat wie erwartet sowohl einen bedeutsamen Einfluss auf die Exploration und die Steuerung im Umgang mit technischen Alltagsgeräten als auch auf den Wissenserwerb nach der Exploration und die Steuerung im allgemeinen Problemlösen. Ebenfalls erwartungskonform fällt der Zusammenhang zwischen Intelligenz und technischer Problemlösekompetenz geringer aus, als zwischen Intelligenz und allgemeiner Problemlösefähigkeit. Mit dem Einflussfaktor Intelligenz kann weniger als die Hälfte der Varianz in den technischen Problemlöseleistungen erklärt werden, aber mehr als die Hälfte der Varianz in den allgemeinen Problemlöseleistungen. Auffallend ist der geringer werdende Zusammenhang zwischen dem technischen und dem allgemeinen Problemlösekonstrukt wenn der Prädiktor Intelligenz in die Betrachtungen einbezogen wird. Die Gemeinsamkeit der beiden Konstrukte ist dem Anschein nach zu einem großen Teil auf die sie beeinflussende fluide Intelligenz zurückzuführen. Das spricht für die Eigenständigkeit eines Konstruktes Problemlösekompetenz im Umgang mit technischen Alltagsgeräten.

Die Annahme, dass der spezifische Kontext, bzw. die durch ihn aktivierten Vorerfahrungen bezüglich der technischen Geräte den zu Tage getretenen Unterschied zwischen den beiden Konstrukten begründet (Kapitel 5.2), ließ sich in dieser Arbeit allerdings nicht zeigen. Als relevant für den Erwerb von Eingriffswissen wurde die Übung mit einem technischen System gesehen. Diese Übung wird im Alltag durch Erfahrungen im Umgang mit den entsprechenden Geräten gewonnen, weshalb anzunehmen war, dass die Erfahrung auch einen Einfluss auf den problemlösenden Umgang mit einem technischen System hat. Lediglich für vereinzelt technische Systeme konnte ein Einfluss der Vorerfahrung mit genau diesen Systemen auf den explorierenden und/oder steuernden Umgang festgestellt werden, nicht aber für eine über alle Geräte hinweg generalisierte Vorerfahrung.

Die Ursache für den in dieser Arbeit ausgebliebenen Zusammenhang zwischen der Nutzungserfahrung und dem Umgang mit technischen Geräten kann zum einen in der Erfassungsmethode und zum anderen in der theoretischen Annahme selbst liegen. Die Ausprägung der gerätespezifischen Erfahrung wurde für jedes technische Gerät aus dem Test mit einer drei-vierstufigen Ratingskala erfasst, deren Abstufungen zusätzlich durch konkrete Falldarstellungen verdeutlicht wurden (Kapitel 9.1.4). Eine separate Überprüfung der Qualität (d. h. der Eindeutigkeit) der Falldarstellungen konnte in dieser Arbeit nicht geleistet werden, so dass nicht auszuschließen ist, dass Probanden sich eventuell nur ungenau in eine der Kategorien einordnen ließen. Die Ausprägung der Erfahrung variierte mit einer lediglich drei Stufen enthaltenen Skala in den Ergebnissen nur unzureichend, insbesondere bei Geräten, deren Vorwissenskompatibilität als gering einzuschätzen ist. Eine ausreichend große Varianz

ist aber die Voraussetzung um einen Zusammenhang zwischen zwei Variablen feststellen zu können.

Die theoretische Annahme, dass die Vorerfahrungen förderlich für den Problemlöseprozess und die Problemlöseleistung sind, muss nicht bei jeder Person und jedem technischen Gerät zutreffen. Vorstellbar ist, dass die technischen Geräte, mit denen die Probanden bereits Erfahrungen gesammelt haben, sich hinsichtlich bestimmter Aspekte von den Geräten unterscheiden, die in dem Test simuliert wurden. Ein Transfer bereits vorhandenen Wissens und erworbener Fähigkeiten gelingt nämlich meist nur, wenn sich Ursprungs- und Zielkontext in ihren Wissensinhalten möglichst ähnlich sind (Kapitel 5.2.1). Auch wenn sich die technischen Geräte in dem hier verwendeten computerbasierten Test nicht hinsichtlich ihrer Funktionen von denen, die die Probanden aus ihrem Alltag kennen, unterscheiden, sind Abweichungen in der Art und Weise, wie diese Funktionen aufgerufen werden können, denkbar. Es kann außerdem sein, dass sich das Vorwissen der Probanden lediglich auf einen bestimmten Teil von Funktionen bezieht, in dem Test jedoch eine ganz andere Funktion aufgerufen werden musste. Ein weiteres Problem kann in der angenommenen Linearität der Beziehung zwischen der Erfahrung und dem durch sie erworbenen Wissen liegen. Ein Mehr an Erfahrung mit einem technischen System muss nicht zwangsläufig auch zu einer Ausdifferenzierung und einer Wissenserweiterung führen. Ab einem bestimmten Expertisegrad werden Wissensinhalte miteinander verknüpft und in Chunks zusammengefasst (Kapitel 2.4.2.1), wodurch es zu einer Vereinfachung der Wissensbasis kommt, bei der Wissen, das nicht mehr oft benötigt wird, auch schwerer zu aktivieren ist (Kluwe, 1997b). Um eine eventuell vorhandene Diskrepanz zwischen den technischen Systemen, mit denen Probanden bereits Erfahrungen sammeln konnten und denen, die in dem Test zu bedienen waren, erkennen zu können, wäre eine Abfrage des Grads an Übereinstimmung nach jedem Steuerungselement vorstellbar.

Probanden, die über nur wenige Erfahrungen mit den technischen Geräten verfügten, hatten in der Systemexploration die Möglichkeit diese Vorwissensdefizite auszugleichen. Wie gut diese Anforderung von einer Person bewältigt wird ist neben der kognitiven Fähigkeit außerdem von der aktualisierten Motivation abhängig, die sich aus den situativen Anreizen und den überdauernden Motiven einer Person zusammensetzt (Hesse, Spies & Lüer, 1983). Eine zurückhaltende Informationssuche lässt sich somit auch auf eine geringe Selbstwirksamkeitserwartung zurückführen. Selbstwirksamkeitserwartung, unter der die subjektive Überzeugung verstanden wird, aufgrund der eigenen Fähigkeiten zukünftige Anforderungen bewältigen und eigene Handlungen zielgerichtet steuern zu können (Bandura, 1977), ist ein wichtiges Bestimmungsstück der Motivation. Eine hohe Selbstwirksamkeitserwartung erhöht wiederum die Anstrengung und Ausdauer bei der Bewältigung neuer Aufgaben mit der Konsequenz einer größeren Erfolgswahrscheinlichkeit (Jerusalem & Mittag, 1999). Personen, die über ein geringeres kognitives Potenzial verfügen, sich aber durch eine erhöhte Ausdauer und eine größere Arbeits- und Leistungsbereitschaft auszeichnen, sind in vielen Fällen erfolgreicher als Personen, die zwar über das kognitive Potenzial verfügen, jedoch wenig zielorientiert und motiviert sind (OECD, 2013b, S. 64).

Arbeiten, in denen das Leistungsergebnis als Folge der eigenen Handlung sichtbar wird, können die Selbstwirksamkeitserwartungen der Probanden als Einflussfaktor auf die

Leistungen herausstellen. Die PISA Studie 2012 (OECD, 2013b), aus der die in dieser Arbeit verwendeten Skalen Ausdauer und Aufgeschlossenheit gegenüber Problemlösen entstammen, offenbarte einen starken positiven Zusammenhang zwischen den Selbstwirksamkeitserwartungen und den Problemlöseleistungen der Schülerinnen und Schüler, wobei dieser Effekt auf dem höheren Leistungsniveau noch deutlich größer ausfällt als auf dem niedrigen Leistungsniveau. In der vorliegenden Arbeit gelang es jedoch nicht dieses Ergebnis zu replizieren. Wider Erwarten zeigte sich ein statistisch bedeutender negativer Einfluss der selbsteingeschätzten Ausdauer bei der Bewältigung neuer Probleme auf die Problemlöseleistungen. Personen, die von sich selbst sagen, sie seien wenig ausdauernd in Bezug auf das Lösen von Problemen explorierten technische Systeme vollständiger, steuerten sie anschließend besser und erzielten höhere Leistungen im allgemeinen Problemlösetest, als Personen die vorgaben, besonders ausdauernd zu sein.

Die Verteilung der Häufigkeitsnennungen (Kapitel 11) dokumentiert eine überdurchschnittliche Ausprägung der Selbstwirksamkeitserwartung in der untersuchten Stichprobe von Studierenden. Dies kann unter anderem darin begründet sein, dass sich Studierende, aber auch Schülerinnen und Schüler tendenziell häufiger überschätzen, als unterschätzen (Pajares & Kranzler, 1995; Pajares & Miller, 1994). Während eine moderate Überschätzung der eigenen Kompetenzerwartung das Ausprobieren herausfordernder Aufgaben ermöglicht und über die Beeinflussung von Motivation und Selbstregulation den Wissenserwerb eher fördert (Pfeiffer, 2000), deutet eine starke Überschätzung auf eine unrealistische Einschätzung der eigenen Handlungs- und Kontrollmöglichkeiten hin (Maag Merki, 2004). Diese führt mit großer Wahrscheinlichkeit zu Fehlentscheidungen und damit zu fehlerbehafteten Handlungen, denn die Leistungen einer Person werden nicht nur durch ihre Selbstwirksamkeitserwartung bestimmt, sondern auch durch ihr Wissen und ihre Fähigkeiten (Bandura, 1977). Ein anderer Erklärungsgrund für den hohen Grad an Selbstwirksamkeitserwartung ist in der Selbstselektion der Stichprobe zu finden. Der Testaufruf enthielt bereits Informationen über die Testdauer und darüber, dass der problemlösende Umgang mit technischen Geräten computerbasiert erfasst werden soll. Es ist denkbar, dass sich auf diesen Aufruf vorwiegend Personen meldeten, die Spaß am Denken und an Herausforderungen haben und die von sich erwarten, dass sie in dem Test gut abschneiden.

Warum Personen, die sich selbst als besonders ausdauernd einschätzten, diese Eigenschaft nicht zur Exploration der technischen Systeme nutzten, sondern diese Phase früher beendeten als Personen mit einer geringeren Ausdauer, lässt sich nur vermuten. Durch das in der Explorationsphase noch nicht vorhandene Steuerungsziel fehlte diesen Personen möglicherweise die Herausforderung sich intensiv mit den Systemen zu beschäftigen. Denkbar ist aber auch, dass diese Personen aufgrund ihrer hohen Selbstwirksamkeitserwartung überzeugt waren, dass sie die technischen Geräte anschließend auch ohne eine ausführliche Systemexploration gut steuern können. Da jedoch die Systemexploration der Informationsgenerierung und damit dem Wissenserwerb diene, müsste die folgende Steuerungsaufgabe für die Personen, die unvollständiger explorierten, eine größere Herausforderung darstellen. Das eine tatsächlich stark ausgeprägte Ausdauer die in dieser Situation nötig würde, zu geringeren Leistungen führen soll, ist allerdings schwer nachvollziehbar.

Ein fehlender Zusammenhang zwischen der Selbstwirksamkeitserwartung als ein überdauerndes Motiv und dem Testverhalten ließe sich noch über die aktualisierte Motivation erklären. Durch die vorhandene High-Stakes-Testsituation (in der die Probanden durch die Aufsicht auf Belohnung motiviert wurden) wurden überdauernde Motive nicht wie erwartet wirksam. Ein Erklärungsansatz für den gefundenen negativen Einfluss von Ausdauer auf alle Problemlöseleistungen ist damit jedoch nicht aufgezeigt.

Ein weiteres überdauerndes Motiv, für das ein Einfluss auf die aktuelle Motivation und damit auf die Problemlöseleistungen angenommen wurde, ist Need for Cognition (Kapitel 5.1.3). Need for Cognition beschreibt die Tendenz einer Person sich mit dem Denken zu beschäftigen und Spaß an dieser Tätigkeit zu haben (Cacioppo & Petty, 1982). Das Konstrukt zeigt sich insbesondere in der Motivation zur Bewältigung kognitiver Anforderungen und dem Lösen von Problemen, wobei der Fokus auf intrinsischer Motivation liegt (Beißert, Köhler, Rempel et al., 2014). So sind Personen mit einer hohen Need for Cognition eher bereit sich kognitiv aufwändig mit einem Problem zu beschäftigen als Personen mit einer geringeren Need for Cognition (Pechtl, 2009). Den Einfluss, den Need for Cognition auf die Problemlöseleistung ausübt, bestätigen Studien sowohl für analytische Probleme (Coutinho, Wiemer-Hastings, Skowronski et al., 2005), als auch für allgemeine dynamische Probleme (Rudolph, Niepel, Wüstenberg et al., 2015). Neben dem direkten Effekt auf die Problemlöseleistung zeigen sie auch einen indirekten über die Bearbeitungszeit medierten Effekt von Need for Cognition auf die Problemlöseleistung (ebd.).

Die theoretisch begründete und in anderen Arbeiten ermittelte Bedeutung von Need for Cognition konnte in dieser Arbeit weder für die allgemeine noch für die technische Problemlöseleistung gezeigt werden. Kritisch ist zum einen die Verwendung einer für Kinder adaptierten deutschen Version der Originalskala zu sehen und zum anderen die starke Ausprägung des Merkmals in der Stichprobe und die damit verbundene geringe Varianz des Merkmals für die Zusammenhangsbetrachtungen. Weil sich Need for Cognition insbesondere in der intrinsischen Motivation einer Person zeigt, ist allerdings denkbar, dass diese aufgrund der eventuell vorgelegenen extrinsischen Motivation keine Rolle mehr spielte, d. h. auch Personen mit einer geringeren Need for Cognition haben sich in Erwartung positiver Konsequenzen in dem Test angestrengt.

Ein weiterer in dieser Hinsicht interessanter Aspekt betrifft den von Coutinho, Wiemer-Hastings, Skowronski et al. (2005) gefundenen Zusammenhang zwischen Need for Cognition, Metakognition und Feedback. Personen mit einer hohen Need for Cognition wollen Dinge verstehen und passen ihre verwendeten Strategien den Fortschritten an, die sie im Prozess des Problemlösens ständig überwachen. Dafür benötigen sie über ihre Vorgehensweise und ihre Fortschritte jedoch ein regelmäßiges Feedback. Ein solches Feedback sah der technische Problemlösetest nicht vor. Die Systemexploration wurde von den Probanden beendet, wenn sie der Ansicht waren, das System vollständig exploriert zu haben, aber spätestens nach Ablauf einer Zeit von zehn Minuten. Auch die Items der Systemsteuerung sahen kein direktes Feedback in der Art vor, dass die Probanden über eine Richtig- oder eine Falschlösung informiert wurden. Zwar ließen sich diese Informationen indirekt über die Systemreaktionen entnehmen, hierfür mussten diese aber interpretiert werden können. An dieser Stelle kann eine Überschätzung der eigenen Kompetenz wiederum negative Folgen haben. Wenn eine

Person davon überzeugt ist eine Handlung kompetent durchzuführen und sie kein Feedback darüber erhält, dass diese Handlung ein falsches Ergebnis zur Folge hatte, kann sie keine entsprechenden Anpassungen vornehmen. Personen, die von der Wirksamkeit ihrer Handlungen weniger überzeugt sind, werden diese eventuell mehrmals kontrollieren, um sicher zu sein keine Fehler gemacht zu haben. Weil es an dieser Stelle nicht möglich ist zu klären, worin letztendlich die Gründe für die erwartungswidrigen Befunde bezüglich aller motivationalen Einflussfaktoren liegen, diese Faktoren aber scheinbar den Unterschied zwischen dem explorierenden und steuernden Umgang mit technischen Geräten ausmachen, sind hierzu weitere Forschungsanstrengungen nötig, die zum einen den Einfluss des Feedbacks auf die aktuelle Motivation untersuchen sollten, als auch den Einfluss der aktuellen Motivation auf die Problemlöseleistung.

Zusammenfassende Interpretation der Ergebnisse vor dem Hintergrund des Ziels dieser Arbeit

Das in der Einleitung zum Ausdruck gebrachte Ziel in dieser Arbeit zunächst das Konstrukt einer Problemlösekompetenz im Umgang mit technischen Alltagsgeräten auszuarbeiten und auf dieser Grundlage ein Testinstrument zu entwickeln, das schließlich einer empirischen Prüfung unterzogen wird, kann als erreicht gewertet werden. Da für den zu untersuchenden Bereich noch keine ausgearbeitete Theorie vorlag, sind in Anlehnung an Schnell, Hill & Esser (2013) theoretische Ansätze aus verwandten Gegenstandsbereichen übertragen worden. Zu diesen Gegenstandsbereichen zählen zum einen die allgemeine Problemlösepsychologie und zum anderen die gewerblich-technische Berufsbildung.

Grundlegender Bestandteil bei der Ausarbeitung des Konstruktes war die Präzisierung der dabei verwendeten Begriffe, da diese in den angrenzenden Forschungsbereichen unterschiedliche Bedeutungen aufweisen oder unterschiedlich weit gefasst sind. Ein nach der Begriffspräzisierung wesentliches Element bei der Konstruktausarbeitung nahm die Formulierung theoretischer Erklärungen für das unterschiedliche Problemlöseverhalten von Personen ein, denn nach Hempel (1974) hat eine Theorie nicht nur die Aufgabe Sachverhalte zu beschreiben, sondern soll diese auch erklären und prognostizieren können. Die das Problemlöseverhalten einer Person im Umgang mit technischen Alltagsgeräten bestimmenden Faktoren wurden dabei in System-, Situations- und Personenmerkmale (Funke, 1990) differenziert. Anhand größtenteils objektiv ermittelter System- und Situationsmerkmale wurde das Konstrukt inhaltlich in seiner Struktur und seiner Repräsentativität gekennzeichnet. Dafür wurden zum einen die Merkmale Komplexität, Vernetztheit, Dynamik und Intransparenz operationalisiert, um die Schwierigkeit im Umgang mit technischen Geräten zu prognostizieren und zum anderen das Ziel der Interaktion genutzt, um die Systemexploration und die Systemsteuerung als zwei voneinander abgrenzbare Subdimensionen im Umgang mit technischen Geräten zu definieren. Personenmerkmale, von denen im Hinblick auf die Theorie und bereits vorhandener Empirie ein Einfluss auf das Problemlöseverhalten angenommen wurde, dienten der Verortung des Konstruktes in einem nomologischen Netzwerk. Die anschließende Überführung in ein theoretisches Kausalmodell ermöglichte die grafische Darstellung der angenommenen Zusammenhänge und die Ableitung von Hypothesen. Diese wurden benötigt, um das Theoriemodell, das lediglich eine vereinfachte Abbildung der Realität darstellt, im Sinne wissenschaftstheoretischer Erkenntnisgewinnung (Opp, 2014) zu überprüfen.

Mithilfe des Modells finiter Automaten sowie der Erstellung möglichst realitätsnaher Grafiken gelang unter Berücksichtigung der Variation von Systemmerkmalen die Entwicklung von computersimulierten technischen Geräten, anhand derer das unterschiedliche Verhalten der Probanden im Umgang mit solchen Geräten erfasst wurde.

Zeigen konnte diese Arbeit den deutlichen Einfluss, den einige der Merkmale auf die Schwierigkeit im explorierenden Umgang mit den technischen Geräten ausübten. Komplexität als die Menge an Informationen, die während der Systemexploration generiert werden muss, als auch eine fehlende Transparenz dieser Informationen bestimmen wie gut die Exploration eines technischen Gerätes gelingt. Sichtbar wurde dabei auch, dass das Verhalten in der Systemexploration einen Einfluss auf die Systemsteuerung ausübt, das heißt je vollständiger ein technisches Gerät exploriert wurde, desto besser gelang die anschließende Systemsteuerung. Der Einfluss der Systemexploration auf die Systemsteuerung ist dabei allerdings nicht so groß, als dass man hier von einem eindimensionalen Konstrukt ausgehen könnte. Unter Einbezug der Selbstwirksamkeitserwartung, deren Einfluss auf die Exploration und die Steuerung der technischen Alltagsgeräte unterschiedlich hoch ausfiel, zeigte sich jedoch ein überraschend großer Zusammenhang zwischen den in dieser Arbeit nicht erklärbaren Anteilen von Systemexploration und Systemsteuerung, so dass unter dieser Perspektive sogar die Annahme der Mehrdimensionalität des Konstruktes überdacht werden sollte. Dass die Gemeinsamkeiten der beiden Interaktionsformen in dieser Arbeit durch die berücksichtigten Konstrukte nicht erklärt werden konnten, spricht für die Eigenständigkeit des Konstruktes. Trotz der sie beeinflussenden Konstrukte gibt es scheinbar etwas Gemeinsames zwischen den Subdimensionen des Konstruktes, das sich auf die dahinter liegende Problemlösekompetenz im Umgang mit technischen Geräten zurückführen lässt, es aber vielleicht keinen Unterschied macht, welches Interaktionsziel erreicht werden soll.

Nicht falsifiziert werden konnte außerdem die Annahme, nach der sich die Fähigkeit einer Person allgemeine Probleme zu lösen von ihrer Fähigkeit technische Probleme zu lösen unterscheidet. Der zunächst festgestellte Zusammenhang zwischen den beiden Konstrukten wurde bei Beachtung der gemeinsamen Einflussfaktoren sichtlich geringer. Wenn fluide Intelligenz als der stärkste gemeinsame Prädiktor in die Betrachtungen einbezogen wird, verliert der Zusammenhang zwischen den beiden Konstrukten schon deutlich an Bedeutung. Im Unterschied zum technischen Problemlösen lässt sich außerdem ein großer Teil der Gemeinsamkeit zwischen dem Explorieren und dem Steuern beim allgemeinen Problemlösen durch die in dem Theoriemodell berücksichtigten Konstrukte erklären.

Insgesamt zeigen die Ergebnisse der empirischen Prüfung des ausgearbeiteten Konstruktes, dass es anscheinend zwar gelungen ist, die theoretischen Zusammenhänge innerhalb des Konstruktes zu bestätigen, aber nur zu einem Teil die angenommenen Zusammenhänge zu anderen Konstrukten. Erklärungsbedürftig bleiben hier vor allem die teils erwartungswidrigen Befunde, die sich auf den Einfluss motivationaler Aspekte im problemlösenden Umgang mit technischen Geräten beziehen. Nach der Lehre vom kritischen Rationalismus ist das Theoriemodell so umzuformulieren, dass die falsifizierten Bestandteile eliminiert und die sich bewährten Bestandteile nicht verloren gehen (Kromrey, 2009, S. 35). Die so modifizierte Theorie ist dann erneut einer kritischen Überprüfung zu unterziehen (vgl. Opp, 2014, S. 231). Jedoch können die Befunde dieser Arbeit nicht uneingeschränkt für

künftige Forschungsarbeiten genutzt werden. An den Stellen dieses Abschnittes, an denen von Abweichungen zu anderen Forschungsergebnissen berichtet wurde, wurde bereits darauf hingedeutet, dass die Rahmenbedingungen, unter denen die Studie stattfand, sich von denen anderer Untersuchungen unterscheiden. Im nächsten Abschnitt sollen diese Aspekte noch einmal aufgegriffen und daraus resultierende Einschränkungen bei der Interpretation und Generalisierbarkeit der Ergebnisse angeführt werden.

16.3 Limitationen und Ausblick

Der letzte Abschnitt dieser Arbeit erläutert die vorhandenen Einschränkungen des durchgeführten Forschungsvorhabens in Bezug auf die Interpretation der Ergebnisse und die Übertragbarkeit dieser auf andere Untersuchungskontexte. Einzelne Limitationen wurden zwar bereits in dem vorherigen Abschnitt bedacht, hier sollen sie hingegen noch einmal systematisch unter den Schwerpunktrubriken *Methode*, *Stichprobe* und *Testinstrument* dargelegt werden. Aus diesen Überlegungen werden abschließend Empfehlungen und Vorschläge für weiterführende Untersuchungen abgeleitet, die sich zum einen der Modifikation des Theoriemodells und zum anderen der Weiterentwicklung des Testinstrumentes widmen sollten. Die Fortführung der in dieser Arbeit begonnenen Validitätsuntersuchungen sind nötig, um dem Anspruch nach individueller Diagnose und Förderung von technischer Problemlösekompetenz im allgemeinbildenden Technikunterricht nachkommen zu können.

Methode

Intraindividuelle Unterschiede im Umgang mit technischen Alltagsgeräten werden in dieser Arbeit damit begründet, dass sich technische Geräte hinsichtlich der Schwierigkeit mit ihnen zu interagieren unterscheiden. Die noch weit gefassten Merkmale Komplexität, Vernetztheit, Dynamik und Intransparenz, von denen ein schwierigkeitsbestimmender Einfluss auf die Systeminteraktion angenommen wird, wurden zunächst aus der Literatur extrahiert und ihr Einfluss mithilfe kognitionspsychologischer Theorien begründet. Aus ihrer Konkretisierung für den Problemkontext technischer Alltagsgeräte ergaben sich insgesamt 18 System- und 5 Situationsmerkmale von denen ein Einfluss auf die Schwierigkeit angenommen wird. Die Konstruktion der computersimulierten technischen Systeme erfolgte jedoch nicht ausschließlich auf Basis einer gezielten Variation der Systemmerkmale, da die Simulationen möglichst realitätsnahe technische Systeme abbilden sollten. Die Berücksichtigung der System- und Situationsmerkmale bei der Itemkonstruktion reichte aus, um das Schwierigkeitsspektrum in dem Test weitgehend abzudecken. Allerdings zeigte sich auch, dass einige Merkmal nur unzureichend oder nur in Abhängigkeit anderer Merkmale variierten, so dass von den 18 als schwierigkeitsbestimmend angenommenen Systemmerkmalen lediglich von 8 Merkmalen der tatsächliche Einfluss auf die Schwierigkeit analysiert werden konnte. Um solche Abhängigkeiten zu vermeiden, sollten die Systemmerkmale bei der Konstruktion gezielter variiert werden, ohne jedoch minimal komplexe technische Systeme zu entwickeln, die im Alltag so nicht zu finden sind. Ein in dieser Arbeit nicht systematisch in der Itementwicklung und -analyse berücksichtigter Aspekt betrifft eventuelle Interaktionseffekte von einzelnen Merkmalen auf die Explorations- und Steuerungsschwierigkeit. Interessant wäre beispielsweise inwieweit das Merkmal Intransparenz als Moderatorvariable einen Einfluss auf den Zusammenhang zwischen den

Merkmale Komplexität, Vernetztheit und Dynamik und der Explorationsschwierigkeit ausübt. Wenn sich eine solche Vermutung empirisch bestätigen würde, ließen sich daraus Maßnahmen für die Förderung der Problemlösekompetenz von Schülerinnen und Schüler im allgemeinbildenden Technikunterricht ableiten.

Zu prüfen ist vorab allerdings, ob sich der gefundene Einfluss der Merkmale auf die Schwierigkeit auch im Umgang mit anderen technischen Systemen zeigt. Die Ermittlung der schwierigkeitsbestimmenden Merkmale mithilfe der multiplen linearen Regressionsanalyse wurde in dieser Arbeit an lediglich 15 Items untersucht, so dass zu fragen ist, ob ein solcher Befund auf die Gesamtheit aller technischen Alltagsgeräte generalisierbar ist. Die Anzahl an Items, die in etwa erforderlich ist, um einen Determinationskoeffizienten von $R^2 = .26$ für einen starken Effekt mit einer Teststärke von $1-\beta = .80$ als signifikant für $\alpha = .05$ nachweisen zu können, beträgt bei $k = 8$ zu analysierenden Merkmalen $n = 45$ Items (vgl. Bortz & Schuster, 2010, S. 359). Der bei der Analyse des Einflusses der 8 Systemmerkmale auf die Explorationsschwierigkeit berechnete (und korrigierte) Determinationskoeffizient ist allerdings so hoch, dass die 15 Items zur statistischen Absicherung ausreichen. Um den Effekt der 5 Situationsmerkmale auf die Schwierigkeit im steuernden Umgang mit den technischen Geräten statistisch absichern zu können, werden allerdings mindestens 19 Items benötigt. Eine größere Anzahl an Items ist jedoch auch mit einer Erhöhung der Testzeit verbunden, die wiederum einen Einfluss auf die Motivation der Probanden während der Testbearbeitung hat.

Eine weitere Beschränkung betrifft die Interpretation der Befunde zur Analyse des theoretischen Kausalmodells. Als kritisch ist vor allem der Einsatz der Skalen zur Erfassung der überdauernden Motive als Einflussfaktoren auf das Verhalten im Umgang mit technischen Systemen zu sehen. Need for Cognition wurde mit einer deutschen Adaption der Skala für Kinder erfasst. Zwar weicht die Reliabilität nur geringfügig von der berichteten Messgenauigkeit in der Zielstichprobe ab, jedoch kann nicht von einer kongruenten Verteilungsannahme ausgegangen werden. Dies betrifft gleichermaßen die Skalen Ausdauer und Aufgeschlossenheit gegenüber dem Problemlösen, die der PISA-Studie 2012 (OECD, 2013b) entnommen wurden und damit für die Zielgruppe 15-jähriger Schülerinnen und Schüler vorgesehen sind. Zu bedenken ist, dass die Güte der Übersetzung der im Original englischsprachigen Skalen Ausdauer und Aufgeschlossenheit gegenüber dem Problemlösen ins Deutsche nicht gezielt geprüft werden konnte. Problematischer als die vorgenommene Übersetzung ist die inhaltliche Ausrichtung der einzelnen Skalen zu sehen. Diese fordern zur Einschätzung der eigenen Fähigkeiten in Bezug auf das allgemeine Problemlösen auf. An welche Art von Problemen (z. B. Denksportaufgaben, mathematische Probleme etc.) die Probanden bei der Beantwortung dieser Skalen gedacht haben ist unklar. Die Selbstwirksamkeit bezogen auf das Lösen von statischen Problemen wird eventuell anders eingeschätzt als das Lösen von Problemen in dynamischen Situationen. Technische Geräte werden oftmals als eigenständige Kommunikationspartner gesehen, auf die nur bedingt Einfluss genommen werden kann und schwierig zu kontrollieren sind. In weiterführenden Forschungsarbeiten sollten diese Überlegung aufgegriffen und untersucht werden, ob die Selbstwirksamkeitserwartung bezüglich des Umgangs mit technischen Alltagsgeräten sich von der hier erfassten auf das Problemlösen bezogenen Selbstwirksamkeitserwartung unterscheidet.

Beschränkend sind die Befunde aus der Überprüfung des Theoriemodells auch vor dem Hintergrund des geringen Stichprobenumfangs zu sehen. Im Hinblick auf die anderenfalls zu hohe Parameteranzahl, die in dem Strukturgleichungsmodell anhand der geringen Datenmenge zu schätzen wäre, wurden einige der Konstrukte nicht auf latenter Ebene analysiert, sondern als manifeste Variablen in das Modell aufgenommen. Hierdurch sind Zusammenhänge allerdings messfehlerbehaftet und ihre Höhen werden eventuell unterschätzt. Dazu tragen auch die unterschiedlich schief verteilten sowie wenig Varianz aufweisenden Testwerte bei.

Stichprobe

Unter ökonomischen Gesichtspunkten konnten nicht alle Konstrukte, von denen eventuell ein Einfluss auf das Verhalten im Umgang mit technischen Geräten ausgeübt wird, erfasst werden. Das Interesse an Technik und an technischen Artefakten ist ein solcher Einflussfaktor, der nicht explizit mit einer Skala oder einem anderen Testinstrument ermittelt wurde. Berücksichtigt wurde diese Variable, indem die technische Problemlösekompetenz zunächst ausschließlich an einer Stichprobe von Studierenden eines technischen Faches untersucht wurde, da für das Studium eines technischen Faches Interesse an Technik vorausgesetzt wird. Ob das Interesse an den spezifischen technischen Sachverhalten der jeweiligen Studiengänge auch mit dem Interesse an technischen Geräten verbunden ist, konnte im Rahmen dieser Arbeit nicht geklärt werden. Mit der Auswahl der Stichprobe ist allerdings auch der Nachteil verbunden, dass sich die Befunde aus dieser Untersuchung nicht auf Personen übertragen lassen, deren Technikinteresse als gering einzuschätzen ist. Nicht generalisiert werden dürfen die Befunde außerdem auf Personen, die nicht der Gruppe von Studierenden angehören, da sich diese hinsichtlich der kognitiven Fähigkeiten, der Selbstwirksamkeitserwartungen aber auch hinsichtlich des Alters und anderen Variablen von anderen Personengruppen unterscheiden. Im Anschluss an die Beseitigung methodischer Mängel sollte die Repräsentativität der Stichprobe durch eine sukzessive Aufhebung der Selektionskriterien erhöht werden, sodass schlussendlich Aussagen über die Problemlösekompetenz von Schülerinnen und Schülern im allgemeinbildenden Technikunterricht aller Schulformen möglich sind.

Testinstrument

Mit dem in dieser Arbeit entwickelten Testinstrument ist es möglich geworden, kompetentes Handeln von Personen in dynamischen Kontexten, wie sie technische Alltagsgeräte darstellen, zu erfassen. Zwar handelt es sich bei den erstellten computersimulierten technischen Systemen um Modelle, die nur einen Ausschnitt der Realität abbilden können, aufgrund der Menge an modellierten Details, die mit denen realer Geräte übereinstimmen, lassen sich jedoch Rückschlüsse von dem Verhalten in den Simulationen auf das Verhalten in der Realität ziehen. Da die vorliegende Arbeit eine solche Annahme allerdings nicht prüfen konnte und sich auf diesbezügliche Arbeiten aus dem beruflichen Bereich stützt (Abele, Gschwendtner & Nickolaus, 2009a; Walker, Link & Nickolaus, 2015), steht eine Überprüfung dieses Validitätsaspektes für Alltagsgeräte noch aus.

Eine weitere Annahme, die im Rahmen dieser Arbeit getroffen wurde, bezieht sich auf das während der Systemexploration erworbene Wissen. Im Gegensatz zu vielen anderen Studien,

in denen das Wissen nach der Exploration der präsentierten dynamischen Systeme abgefragt wird (Kluge, 2004; Buchner, 1999; Neubert, Kretzschmar, Wüstenberg et al., 2015), erfolgt die Bewertung des Probandenverhaltens in dieser Arbeit anhand der Vollständigkeit der Systemexploration. Ob die während der Systemexploration generierten Informationen allerdings wahrgenommen und zu Wissen weiterverarbeitet wurden, ist noch zu klären. Der hohe Zusammenhang zwischen der Vollständigkeit der Systemexploration und der Leistung in der Systemsteuerung deutet zwar darauf hin, dass ein Wissenserwerb stattgefunden hat, einen Nachweis dafür kann hiermit aber nicht gegeben werden. Die Frage nach den tatsächlich wahrgenommenen Informationen ließe sich mit der Methode Eye-Tracking beantworten, bei der sich anhand von Blickbewegungen und Fixationen Aussagen über die zur Kenntnis genommenen Informationen treffen lassen (Holmqvist, Nyström & Andersson, 2011). Für den Nachweis, dass diese Informationen zu Wissen verarbeitet wurden, bleibt bislang nur die Wissensabfrage.

Begrenzend sollten die Befunde zur Validität des theoretisch entwickelten Konstruktes außerdem vor dem Hintergrund der Messgüte des konstruierten Testinstrumentes bewertet werden. Die teilweise hohen Fehlervarianzen in den Messergebnissen einzelner Items weisen darauf hin, dass außer der zu erfassenden Kompetenz im Umgang mit technischen Alltagsgeräten andere Merkmale während der Interaktion mit einem Gerät eine Rolle spielen. Problematisch dabei erscheint die unterschiedliche Höhe des Einflusses anderer Konstrukte auf den Umgang mit den einzelnen Geräten. Das Problem geringer Itemtrennschärfen entsteht vor allem bei Konstrukten die sehr breit definiert sind und durch deren Operationalisierung sehr heterogene Items entstehen. Zwar wurde das hier formulierte Kompetenzkonstrukt auf den Umgang mit technischen Alltagsgeräten eingegrenzt (bei dem der Umgang mit technischen Geräten aus beruflichen Kontexten zunächst ausgeschlossen wird), dieser Kontext ist mit den Inhaltsbereichen Selbstbedienungsautomaten, Hausgerätetechnik, Haustechnik, Multimedia und Informations- und Kommunikationstechnik allerdings noch sehr weit gefasst. Nicht nur die verschiedenen Inhaltsbereiche, sondern auch die Schwierigkeiten haben einen Einfluss auf die Trennschärfe der Items und damit auf die Gesamttestreliabilität (Bühner, 2011, S. 178). Eine Reliabilitätserhöhung mithilfe der Homogenisierung von Items zu erreichen, die dann nur einen Teil des Konstruktes und des Schwierigkeitsspektrums abdecken, erscheint nicht zielführend. Gleiches gilt für die alternative Möglichkeit die Reliabilität durch eine Vergrößerung der Itemanzahl zu erhöhen, da diese mit einer Testzeitverlängerung einhergeht. Da die Problematik der geringen Trennschärfen insbesondere Items des steuernden Umgangs mit den technischen Geräten betrifft, ist zu überlegen ob zur Erfassung dieser Interaktionsform, Steuerungsziele so formuliert werden können, dass sie eine polytome Auswertung ermöglichen, die im Gegensatz zu der dichotomen Auswertung Leistungen differenzierter erfassen kann.

Ausblick

Ogleich der genannten Limitierungen, die es bei der Ergebnisinterpretation zu berücksichtigen gilt, leistet die vorliegende Arbeit einen ersten Beitrag kompetentes Handeln mit technischen Systemen computerbasiert zu erfassen. Der Einsatz von Computersimulationen hilft dabei zu verstehen, wodurch Anforderungen in der Interaktion mit einem technischen System entstehen. Im Gegensatz zur Methode der post-hoc

Beschreibung von schwierigkeitsbestimmenden Merkmalen sind in dieser Studie *vorab* Annahmen über System- und Situationsmerkmale getroffen worden. Die Überprüfung dieser Merkmale erfolgte, indem das System selbst sowie die Umgebungsbedingungen variiert und die Wirkungen dieser Variationen analysiert wurden. Hier und in künftigen Untersuchungen haben Computersimulationen gegenüber realen technischen Systemen den Vorteil vielfältige Modifikationen zu erlauben. Bereits in dieser Arbeit konnte anhand unterschiedlich ausgeprägter Systemmerkmale ein großer Teil der Schwierigkeit im explorierenden und steuernden Umgang mit technischen Systemen erklärt werden. Die Erkenntnis, dass der größte Teil der Schwierigkeit im explorierenden Umgang mit technischen Geräten durch dessen Komplexität und mangelnde Transparenz verursacht wird, lässt sich für eine gezieltere Itemkonstruktion nutzen. Zum einen können Merkmale bei künftigen Itementwicklungen noch gezielter variiert und zum anderen mögliche Interaktionseffekte zwischen ihnen untersucht werden. Künftige Forschungsarbeiten sollten in diesem Zusammenhang außerdem die Möglichkeit unterschiedlicher Nutzungstypen im Umgang mit technischen Geräten in Betracht ziehen, die sich beispielsweise in den tatsächlich als schwierig wahrgenommenen System- und Situationsmerkmalen unterscheiden.

Die berichteten Ergebnisse haben nicht nur aus testtheoretischer Perspektive für die Validierung des Konstruktes *Problemlösender Umgang mit technischen Geräten* Relevanz, sondern sind außerdem aus fachdidaktischer Sicht bedeutend, weil sich aus ihnen Maßnahmen zur Förderung technischer Problemlösefähigkeit generieren lassen. Ein Fokus kann dabei in Abhängigkeit individueller Voraussetzungen auf die Vermittlung komplexitätsreduzierender und transparenzerhöhender Explorationsstrategien liegen. Tätigkeiten von Schülerinnen und Schüler im Technikunterricht die eigenes Handeln erlauben, erwiesen sich bereits in einer anderen Untersuchung (Walker, 2013) als besonders lernwirksam. Bevor allerdings über Fördermaßnahmen diskutiert wird, ist der kompetenzdiagnostische Prozess weiter zu optimieren. Vielversprechend erscheint hier vor allem die weitere Analyse der Prozessdaten, die während der Interaktionen mit den technischen Systemen entstehen. Für die Systemexploration konnte mit der Explorationsvollständigkeit bereits ein geeigneter Prozessindikator gefunden werden. Das in die Bewertung der Steuerungsleistung eingehende Prozessmerkmal, die zur Erreichung eines Steuerungsziels benötigten Interaktionsschritte, kann allerdings nur zwischen Personen differenzieren, die das Steuerungsziel bereits erreicht haben. Die Herausforderung bei der Prozessanalyse, deren Auflösungsgrad bis auf Interaktionsebene reichen soll, besteht bislang allerdings in dem damit verbundenen Aufwand. Eine automatische Auswertung solcher Prozessdaten dürfte in den nächsten Jahren sicherlich möglich werden, so dass die Genauigkeit der Diagnose des problemlösenden Umgangs mit technischen Geräten erhöht und damit individuell passende Fördermaßnahmen entwickelt werden können.

17 Literaturverzeichnis

- Abele, Stephan (2016). Umgang mit Komplexität: Eine bedeutende psychische Voraussetzung des domänenspezifischen Problemlösens? *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 112(1), S. 37–59.
- Abele, Stephan; Greiff, Samuel; Gschwendtner, Tobias; Wüstenberg, Sascha; Nickolaus, Reinhold; Nitzschke, Alexander, et al. (2012). Dynamische Problemlösekompetenz: Ein bedeutsamer Prädiktor von Problemlöseleistungen in technischen Anforderungskontexten? *Zeitschrift für Erziehungswissenschaft*, 15(2), S. 363–391.
- Abele, Stephan; Gschwendtner, Tobias & Nickolaus, Reinhold (2009a). Berufliche Handlungskompetenz valide erfassen - computerbasierte Simulationen technischer Systeme als innovative Diagnoseinstrumente. *Die berufsbildende Schule*, 61(9), S. 252–254.
- Abele, Stephan; Gschwendtner, Tobias & Nickolaus, Reinhold (2009b). Computerbasierte Simulationen in internationalen Vergleichsstudien: Können sie berufliche Handlungskompetenz aussagekräftig erfassen? *Berufsbildung: Zeitschrift für Praxis und Theorie in Betrieb und Schule*, 63(119), S. 24–26.
- Abele, Stephan; Walker, Felix & Nickolaus, Reinhold (2014). Zeitökonomische und reliable Diagnostik beruflicher Problemlösekompetenzen bei Auszubildenden zum Kfz-Mechatroniker. *Zeitschrift für Pädagogische Psychologie*, 28(28), S. 167–179.
- Achen, Christopher H. (1982). *Interpreting and using regression*. Beverly Hills, London: Sage Publ.
- Ackerman, Phillip L.; Beier, Margaret E. & Boyle, Mary O. (2005). Working Memory and Intelligence: The Same or Different Constructs? *Psychological Bulletin*, 131(1), S. 30–60.
- Aebli, Hans (1980). *Denken, das Ordnen des Tuns: Bd.1 Kognitive Aspekte der Handlungstheorie*. Stuttgart: Klett-Cotta.
- Amelang, Manfred & Schmidt-Atzert, Lothar (2006). *Psychologische Diagnostik und Intervention*. Berlin, Heidelberg: Springer Medizin.
- Anderson, John R. (1983). *The architecture of cognition*. Cambridge, Massachusetts: Harvard Univ. Pr.
- Anderson, John R. (2007a). *How can the human mind occur in the physical universe?* Oxford, New York: Oxford University Press.
- Anderson, John R. (2007b). *Kognitive Psychologie*. Heidelberg: Spektrum Akad. Verl.
- Anderson, John R. & Lebiere, Christian J. (1998a). Introduction. In John R. Anderson & Christian J. Lebiere (Hrsg.), *The Atomic Components of Thought* (S. 1–17). New York, London: Taylor and Francis.
- Anderson, John R. & Lebiere, Christian J. (Hrsg.) (1998b). *The Atomic Components of Thought*. New York, London: Taylor and Francis.
- Anderson, John R.; Lebiere, Christian J. & Lovett, Marsha C. (1998). Performance. In John R. Anderson & Christian J. Lebiere (Hrsg.), *The Atomic Components of Thought* (S. 57–100). New York, London: Taylor and Francis.
- Anderson, Lorin W. & Krathwohl, David R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Arbinger, Roland (2015). *Psychologie des Problemlösens*. Darmstadt: Wissenschaftliche Buchgesellschaft.

- Artelt, Cordula & Wirth, Joachim (2014). Kognition und Metakognition. In Tina Seidel & Andreas Krapp (Hrsg.), *Pädagogische Psychologie* (S. 167–192). Weinheim: Beltz.
- Arzheimer, Kai (2016). *Strukturgleichungsmodelle: Eine anwendungsorientierte Einführung*. Wiesbaden: Springer VS.
- Asseburg, Regine (2011). *Motivation zur Testbearbeitung in adaptiven und nicht-adaptiven Leistungstests*. Dissertation. Christian-Albrechts-Universität, Kiel.
- Ausubel, David P.; Novak, Joseph D. & Hanesian, Helen (1981). *Psychologie des Unterrichts: Band 2*. Weinheim: Beltz.
- Backhaus, Klaus; Erichson, Bernd; Plinke, Wulff & Weiber, Rolf (2008). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Berlin, Heidelberg: Springer.
- Badke-Schaub, Petra & Tisdale, Timothy (1995). Die Erforschung menschlichen Handelns in komplexen Situationen. In Bernd Strauß & Martin Kleinmann (Hrsg.), *Computersimulierte Szenarien in der Personalarbeit* (S. 43–56). Göttingen: Verl. für Angewandte Psychologie.
- Baier, Gundolf (2003). Erfolgsfaktor Gebrauchstauglichkeit – Anforderungen und Nutzen sowie Wege zur Hervorbringung gebrauchstauglicher Waren. *Forum Ware*, 31(1-4), S. 1–6.
- Baker, Eva L. & O’Neil, Harold F., jr. (2002). Measuring problem solving in computer environments: current and future states. *Computers in Human Behavior*, 18, S. 609–622.
- Balzert, Heide (2005). *Lehrbuch der Objektmodellierung: Analyse und Entwurf mit der UML 2*. Heidelberg: Elsevier Spektrum Akad. Verl.
- Bandura, Albert (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, 84(2), S. 191–215.
- Banse, Gerhard (1996). Technisches Handeln unter Unsicherheit - unvollständiges Wissen und Risiko. In Gerhard Banse & Käthe Friedrich (Hrsg.), *Technik zwischen Erkenntnis und Gestaltung: Philosophische Sichten auf Technikwissenschaften und technisches Handeln* (S. 105–140). Berlin: Ed. Sigma.
- Barnett, Susan M. & Ceci, Stephen J. (2002). When and Where Do We Apply What We Learn? A Taxonomy for Far Transfer. *Psychological Bulletin*, 128(4), S. 612–637.
- Barth, Carola M. & Funke, Joachim (2010). Negative affective environments improve complex solving performance. *Cognition and Emotion*, 24(7), S. 1259–1268.
- Baumann, Konrad & Lanz, Herwig (1998). *Mensch-Maschine-Schnittstellen elektronischer Geräte: Leitfaden für Design und Schaltungstechnik*. Berlin: Springer.
- Baumann, Martin; Simon, Uwe; Wonisch, Astrid & Guttenberger, Helmut (2013). Computersimulation versus Experiment: Gibt es Unterschiede im Erzeugen nachhaltigen Wissens und in der Attraktivität für die Schüler? *MNU*, 66(5), S. 305–310.
- Beckmann, Jens F. (1994). *Lernen und komplexes Problemlösen: Ein Beitrag zur Konstruktvalidierung von Lerntests*. Bonn: Holos-Verl.
- Beckmann, Jens F. & Goode, Natassia (2014). The benefit of being naive and knowing it: the unfavourable impact of perceived context familiarity on learning in complex problem solving tasks. *Instructional Science*, 42, S. 271–290.
- Beckmann, Jens F. & Guthke, Jürgen (1995). Complex Problem Solving, Intelligence, and Learning Ability. In Peter A. Frensch & Joachim Funke (Hrsg.), *Complex problem solving: The European perspective* (S. 177–200). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Beier, Margaret E. & Ackerman, Phillip L. (2005). Working Memory and Intelligence: Different Constructs. Reply to Oberauer et al. (2005) and Kane et al. (2005). *Psychological Bulletin*, 131(1), S. 72–75.

- Beierlein, Constanze; Kemper, Christoph J.; Kovaleva, Anastassiya & Rammstedt, Beatrice (2013). Kurzsкала zur Erfassung allgemeiner Selbstwirksamkeitserwartungen (ASKU). *Methoden, Daten, Analysen*, 7(2), S. 251–278.
- Beißert, Hanna; Köhler, Meike; Rempel, Marina & Beierlein, Constanze (2014). *Eine deutschsprachige Kurzsкала zur Messung des Konstrukts Need for Cognition: Die Need for Cognition Kurzsкала (NFC-K)* (GESIS-Working Papers 32).
- Bentler, Peter M. & Chou, Chin P. (1987). Practical Issues in Structural Modeling. *Sociological Methods & Research*, 16(1), S. 78–117.
- Berg, Cynthia A. & Calderone, Katerina S. (1994). The role of problem interpretations in understanding the development of everyday problem solving. In Robert J. Sternberg & Richard K. Wagner (Hrsg.), *Mind in context: Interactionist perspectives on human intelligence* (S. 105–133). Cambridge, New York: Cambridge University Press.
- Berlyne, Daniel E. (1974). *Konflikt, Erregung, Neugier: Zur Psychologie der kognitiven Motivation*. Stuttgart: Klett.
- Bernien, Maritta (1997). Anforderungen an eine qualitative und quantitative Darstellung der beruflichen Kompetenzentwicklung. In Arbeitsgemeinschaft Qualifikations-Entwicklungs-Management Berlin (Hrsg.), *Berufliche Weiterbildung in der Transformation: Fakten und Visionen* (S. 17–83). Münster, New York, München, Berlin: Waxmann.
- Berry, Dianne C. & Broadbent, Donald E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology*, 79, S. 251–272.
- Bertholet, Miriam & Spada, Hans (2004). Wissen als Voraussetzung und Hindernis für Denken, Problemlösen und Entscheiden. In Gabi Reinmann & Heinz Mandl (Hrsg.), *Psychologie des Wissensmanagements: Perspektiven, Theorien und Methoden* (S. 66–78). Göttingen: Hogrefe.
- Betsch, Tilmann; Funke, Joachim & Plessner, Henning (2011). *Denken - Urteilen, Entscheiden, Problemlösen*. Berlin, Heidelberg, New York: Springer Medizin.
- Blech, C. & Funke, Joachim (2006). Zur Reaktivität von Kausaldiagramm-Analysen beim komplexen Problemlösen. *Zeitschrift für Psychologie*, 117(4), S. 185–195.
- Bless, Herbert; Fellhauer, Roland F.; Bohner, Gerd & Schwarz, Norbert (1991). *Need for cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben* (ZUMA-Arbeitsbericht 6). Mannheim. <http://nbn-resolving.de/urn:nbn:de:0168-ss0ar-68892>. Zugegriffen: 9. Januar 2017.
- Bless, Herbert; Wänke, Michaela; Bohner, Gerd; Fellhauer, Roland F. & Schwarz, Norbert (1994). Need for Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben. *Zeitschrift für Sozialpsychologie*, 25, S. 147–154.
- Blömeke, Sigrid & Zlatkin-Troitschanskaia, Olga (2015). Kompetenzen von Studierenden: Einleitung. In Sigrid Blömeke & Olga Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (S. 7–10). Weinheim, Basel: Beltz.
- Bloom, Benjamin S. (1976). *Taxonomie von Lernzielen im kognitiven Bereich*. Weinheim: Beltz.
- Böhme, Katrin (2011). *Methodische und didaktische Überlegungen sowie empirische Befunde zur Erfassung sprachlicher Kompetenzen im Deutschen: Analyse zu den Bildungsstandards im Fach Deutsch für den Primarbereich*. Dissertation. Humboldt-Universität zu Berlin, Berlin. <http://edoc.hu-berlin.de/dissertationen/boehme-katrin-2011-12-16/PDF/boehme.pdf>. Zugegriffen: 9. Januar 2017.
- Bollen, Kenneth A. (1989). *Structural equations with latent variables*. New York: Wiley.

- Boomsma, Anne & Hoogland, Jeffrey J. (2001). The Robustness of LISREL Modeling Revisited. In Robert Cudeck, Stephen H.C. du Toit & Dag Sörbom (Hrsg.), *Structural equation models: Present and future: A Festschrift in honor of Karl Jöreskog* (S. 139–168). Lincolnwood: Scientific Software International.
- Borg, Ingwer & Staufenbiel, Thomas (2007). *Lehrbuch Theorien und Methoden der Skalierung*. Bern: Huber.
- Bortz, Jürgen & Döring, Nicola (2006). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler*. Heidelberg: Springer Medizin.
- Bortz, Jürgen & Schuster, Christof (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin, Heidelberg: Springer.
- Brandstätter, Hermann (2004). Persönliche Verhaltens- und Leistungsbedingungen. In Heinz Schuler (Hrsg.), *Lehrbuch Organisationspsychologie* (S. 257–288). Bern: Huber.
- Braun-Thürmann, Holger (2002). *Künstliche Interaktion: Wie Technik zur Teilnehmerin sozialer Wirklichkeit wird*. Wiesbaden: Westdt. Verl.
- Breuer, Klaus (2006). Kompetenzdiagnostik in der beruflichen Bildung - eine Zwischenbilanz. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 102(2), S. 194–210.
- Brézillon, Patrick (1999). Context in Problem Solving: A Survey. *The Knowledge Engineering Review*, 14(1), S. 47–80.
- Broadbent, Donald E.; FitzGerald, Peter & Broadbent, Margaret H.P. (1986). Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology*, 77(1), S. 33–50.
- Brügge, Bernd & Dutoit, Allen H. (2004). *Objektorientierte Softwaretechnik: mit UML, Entwurfsmustern und Java*. München: Pearson.
- Buchner, Axel (1999). Komplexes Problemlösen vor dem Hintergrund der Theorie finiter Automaten. *Psychologische Rundschau*, 50(4), S. 206–212.
- Buchner, Axel & Funke, Joachim (1991). *Transfer of associations in finite state automata* (Berichte aus dem Psychologischen Institut der Universität Bonn Band 17, Heft 2).
- Buchner, Axel & Funke, Joachim (1993). Finite-state Automata: Dynamic Task Environments in Problem-solving Research. *The Quarterly Journal of Experimental Psychology*, 46A(1), S. 83–118.
- Buchwald, Florian (2015). *Analytisches Problemlösen: Labor- und feldexperimentelle Untersuchung von Aspekten der kognitiven Potenzialausschöpfungshypothese*. Dissertation. Universität Duisburg-Essen, Essen.
- Buerschaper, Cornelius (2000). Strategisches Denken beim Umgang mit komplexen Problemen. Computersimulierte Szenarien im Forschungs- und Trainingskontext. In Dietmar Herz & Andreas Blätte (Hrsg.), *Simulation und Planspiel in den Sozialwissenschaften: Eine Bestandsaufnahme der internationalen Diskussion* (S. 145–180). Münster: LIT-Verl.
- Bühner, Markus (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Bühner, Markus (2012). *Lineare Strukturgleichungsmodelle*, München, 2012.
- Bühner, Markus; Kröner, Stephan & Ziegler, Matthias (2008). Working memory, visual-spatial-intelligence and their relationship to problem-solving. *Intelligence*, 36(6), S. 672–680.
- Burmeister, Kai (2009). *Komplexes Problemlösen im Kontext angewandter Eignungsdiagnostik*. Dissertation. Ernst-Moritz-Arndt-Universität, Greifswald. http://ub-ed.ub.uni-greifswald.de/opus/volltexte/2009/614/pdf/Burmeister_Kai_Komplexes_Problemloesen_im_Kontext_angewandter_Eignungsdiagnostik.pdf. Zugegriffen: 27. November 2013.

- Byrne, Barbara M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications and programming*. New York: Routledge.
- Cacioppo, John T. & Petty, Richard E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, S. 116–131.
- Cacioppo, John T.; Petty, Richard E.; Feinstein, Jeffrey A. & Jarvis, Blair W.G. (1996). Dispositional Differences in Cognitive Motivation: The Life and Times of Individuals Varying in Need for Cognition. *Psychological Bulletin*, 119(2), S. 197–253.
- Cacioppo, John T.; Petty, Richard E. & Kao, Chuan F. (1984). The Efficient Assessment of Need for Cognition. *Journal of Personality Assessment*, 48(3), S. 306–307.
- Cattell, Raymond B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), S. 1–22.
- Cattell, Raymond B. (1971). *Abilities: Their Structure, Growth, and Action*. Boston: Houghton Mifflin.
- Chatterjee, Samprit & Price, Bertram (1995). *Praxis der Regressionsanalyse*. München: Oldenbourg.
- Cheung, Gordon W. & Rensvold, Roger B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling*, 9(2), S. 233–255.
- Chi, Michelene T.H.; Glaser, Robert & Rees, Ernest (1982). Expertise in Problem Solving. In Robert J. Sternberg (Hrsg.), *Advances in the psychology of human intelligence: Volume 1* (S. 7–75). Hillsdale: Lawrence Erlbaum Associates.
- Chomsky, Noam (1970). *Aspekte der Syntax-Theorie*. Frankfurt am Main: Suhrkamp.
- Christ, Oliver & Schlüter, Elmar (2012). *Strukturgleichungsmodelle mit Mplus: Eine praktische Einführung*. München: Oldenbourg.
- Cohen, Arthur R.; Stotland, Ezra & Wolfe, Donald M. (1955). An experimental investigation of need for cognition. *Journal of abnormal and social psychology*, 51, S. 291–294.
- Cole, David A.; Ciesla, Jeffrey A. & Steiger, James H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, 12(4), S. 381–398.
- Colom, Roberto; Abad, Francisco J.; Quiroga, Ángeles M.; Shih, Pei C. & Flores-Mendoza, Carmen (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, 36(6), S. 584–606.
- Coutinho, Savia; Wiemer-Hastings, Katja; Skowronski, John J. & Britt, Anne M. (2005). Metacognition, need for cognition and use of explanations during ongoing learning and problem solving. *Learning and Individual Differences*, 15(4), S. 321–337.
- Cox, Anna L. & Young, Richard M. (2000). Device-Oriented and Task-Oriented Exploratory Learning of Interactive Devices. In Niels A. Taatgen & Jans Aasman (Hrsg.), *ICCM, University of Groningen, Groningen, Netherlands, 23-25 March 2000* (S. 70–77). Veenendaal: Universal Press.
- Cronbach, Lee J. & Meehl, Paul E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin*, 52(4), S. 281–302.
- Curran, Patrick J.; West, Stephen G. & Finch, John F. (1996). The robustness of test statistics to normality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(16-29).
- Dahm, Markus (2006). *Grundlagen der Mensch-Computer-Interaktion*. München: Pearson.

- Daily, Larry Z.; Lovett, Marsha C. & Reder, Lynne M. (2001). Modeling individual differences in working memory performance: a source activation account. *Cognitive Science*, 25(3), S. 315–353.
- Dammann, Elmar; Behrendt, Stefan; Ștefănică, Florina & Nickolaus, Reinhold (2016). Kompetenzniveaus in der ingenieurwissenschaftlichen akademischen Grundbildung - Analysen im Fach Technische Mechanik. *Zeitschrift für Erziehungswissenschaften (ZfE)*, 19(2), S. 351–374.
- Danner, Daniel; Hagemann, Dirk; Holt, Daniel V.; Hager, Marieke; Schankin, Andrea; Wüstenberg, Sascha, et al. (2011). Measuring Performance in Dynamic Decision Making: Reliability and Validity of the Tailorshop Simulation. *Journal of Individual Differences*, 32(4), S. 225–233.
- Danner, Daniel; Hagemann, Dirk; Schankin, Andrea; Hager, Marieke & Funke, Joachim (2011). Beyond IQ: A latent state-trait analysis of general intelligence, dynamic decision making, and implicit learning. *Intelligence*, 39, S. 323–334.
- Deci, Edward L. & Ryan, Richard M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), S. 223–238.
- Diaz-Bone, Rainer (2013). *Statistik für Soziologen*. Konstanz, Stuttgart: UVK-Verl.-Ges; UTB.
- Döring, Nicola & Bortz, Jürgen (Hrsg.) (2016a). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Döring, Nicola & Bortz, Jürgen (2016b). Untersuchungsdesign. In Nicola Döring & Jürgen Bortz (Hrsg.), *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler* (S. 181–220). Heidelberg: Springer.
- Dörner, Dietrich (1976). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.
- Dörner, Dietrich (1981). Über die Schwierigkeiten menschlichen Umgangs mit Komplexität. *Psychologische Rundschau*, 32, S. 163–179.
- Dörner, Dietrich (1984). Denken, Problemlösen und Intelligenz. *Psychologische Rundschau*, 35(1), S. 10–20.
- Dörner, Dietrich (1986). Diagnostik der operativen Intelligenz. *Diagnostica*, 32(4), S. 290–308.
- Dörner, Dietrich & Kreuzig, Heinz W. (1983). Problemlösefähigkeit und Intelligenz. *Psychologische Rundschau*, 34(4), S. 185–192.
- Dörner, Dietrich & van der Meer, Elke (Hrsg.) (1995). *Das Gedächtnis: Probleme - Trends - Perspektiven*. Göttingen: Hogrefe.
- Duncker, Karl (1974). *Zur Psychologie des produktiven Denkens*. Berlin, Heidelberg, New York: Springer.
- Eckert, Manfred (2000). Die Theorie technischer Systeme - Ein Ansatz zur Didaktik technischer beruflicher Fachrichtungen? In Reinhard Bader & Klaus Jenewein (Hrsg.), *Didaktik der Technik zwischen Generalisierung und Spezialisierung: Helmut Sanfleber gewidmet anlässlich der Verabschiedung von der Gerhard-Mercator-Universität Duisburg* (S. 125–140). Frankfurt am Main: GAFB.
- Eid, Michael; Gollwitzer, Mario & Schmitt, Manfred (2013). *Statistik und Forschungsmethoden: Lehrbuch. Mit Online-Materialien*. Weinheim: Beltz.
- Eid, Michael & Schmidt, Katharina (2014). *Testtheorie und Testkonstruktion*. Göttingen: Hogrefe.
- Embretson, Susan (1983). Construct Validity: Construct Representation Versus Nomothetic Span. *Psychological Bulletin*, 93(1), S. 179–197.

- Engle, Rendall W.; Tuholski, Stephen W.; Laughlin, James E. & Conway, Andrew R.A. (1999). Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *Journal of Experimental Psychology: General*, 128(3), S. 309–331.
- Erdfelder, Edgar (1994). Erzeugung und Verwendung empirischer Daten. In Theo Herrmann & Werner H. Tack (Hrsg.), *Methodologische Grundlagen der Psychologie* (S. 47–97). Göttingen, Seattle: Hogrefe.
- Ericsson, K. A. & Simon, Herbert A. (1980). Verbal reports as Data. *Psychological Review*, 87(3), S. 215–251.
- Erpenbeck, John & Rosenstiel, Lutz von (Hrsg.) (2007). *Handbuch Kompetenzmessung: Erkennen, verstehen und bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis*. Stuttgart: Schäffer-Poeschel.
- Esser, Hartmut (1996). Die Definition der Situation. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 48(1), S. 1–34.
- Faul, Franz; Erdfelder, Edgar; Buchner, Axel & Lang, Albert-Georg (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior research methods*, 41(4), S. 1149–1160.
- Finch, William H. & French, Brian F. (2015). *Latent variable modeling with R*. New York, NY: Routledge Taylor & Francis Group.
- Finney, Sara J. & DiStefano, Christine (2006). Non-Normal and Categorical Data in Structural Equation Modeling. In Gregory R. Hancock & Ralph O. Mueller (Hrsg.), *Structural equation modeling: A second course* (S. 269–314). Greenwich, Conn: IAP.
- Fisseni, Hermann-Josef (2004). *Lehrbuch der psychologischen Diagnostik: Mit Hinweisen zur Intervention*. Göttingen: Hogrefe.
- Flammer, August (2009). *Entwicklungstheorien: Psychologische Theorien der menschlichen Entwicklung*. Bern: Huber.
- Fleischer, Jens; Wirth, Joachim & Leutner, Detlev (2014). Effekte der kontextuellen Einkleidung von Testaufgaben auf die Schülerleistungen im analytischen Problemlösen und in der Mathematik. *Zeitschrift für Pädagogische Psychologie*, 28(4), S. 217–227.
- Fleischer, Jens; Wirth, Joachim; Rumann, Stefan & Leutner, Detlev (2010). Strukturen fächerübergreifender und fachlicher Problemlösekompetenz - Analyse von Aufgabenprofilen. *Zeitschrift für Pädagogik*, 56(Beiheft), S. 239–248.
- Fletcher, Stefan (2004). *Förderung der Problemlösefähigkeit zum Konstruieren mit Hilfe eines wissensbasierten Lernsystems*. Dissertation. Otto-von-Guericke-Universität Magdeburg, Magdeburg.
- Florian, Christine; Sandmann, Angela & Schmiemann, Philipp (2014). Modellierung kognitiver Anforderungen schriftlicher Abituraufgaben im Fach Biologie. *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), S. 175–189.
- Fortmüller, Richard (1997). *Wissen und Problemlösen: Eine wissenspsychologische Analyse der notwendigen Voraussetzungen für die Bewältigung von (komplexen) Problemen und Konsequenzen für den Unterricht in berufsbildenden Vollzeitschulen*. Wien: Manz-Verl.-Schulbuch.
- Frackmann, Margit & Tärre, Michael (2009). *Lernen und Problemlösen in der beruflichen Bildung: Methodenhandbuch*. Bielefeld: Bertelsmann.
- Frei, Felix; Duell, Werner & Baitsch, Christof (1984). *Arbeit und Kompetenzentwicklung: Theoretische Konzepte zur Psychologie arbeitsimmanenter Qualifizierung*. Bern: Huber.

- Frensch, Peter A. & Funke, Joachim (1995). Definitions, Traditions, and a General Framework for Understanding Complex Problem Solving. In Peter A. Frensch & Joachim Funke (Hrsg.), *Complex problem solving: The European perspective* (S. 3–25). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Freunberger, Roman & Itzlinger-Bruneforth, Ursula (2013). *Item Entwicklung: Kognitive Prozesse und Item Schwierigkeit*. Salzburg. <https://www.bifie.at/system/files/dl/Item-Entwicklung.pdf>. Zugegriffen: 9. Januar 2017.
- Frey, Andreas & Hartig, Johannes (2013). Wann sollten computerbasierte Verfahren zur Messung von Kompetenzen anstelle von papier- und bleistift-basierten Verfahren eingesetzt werden? *Zeitschrift für Erziehungswissenschaft*, 16(S1), S. 53–57.
- Friege, Gunnar & Lind, Gunter (2003). Allgemeine und fachspezifische Problemlösekompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 9, S. 63–74.
- Fromm, Sabine (2012). *Datenanalyse mit SPSS für Fortgeschrittene 2: Multivariate Verfahren für Querschnittsdaten*. Wiesbaden: VS Verl. für Sozialwissenschaften.
- Funke, Joachim (1983a). *Bereichsspezifische Effekte beim Problemlösen und der Einfluß von Vorerfahrung*. Trier.
- Funke, Joachim (1983b). Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? *Diagnostica*, 29(4), S. 283–302.
- Funke, Joachim (1984). Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen: Mit einigen Anmerkungen von D. Dörner. *Sprache und Kognition*, 3, S. 159–172.
- Funke, Joachim (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle. *Zeitschrift für Psychologie*, 193(4), S. 443–465.
- Funke, Joachim (1986). *Komplexes Problemlösen: Bestandsaufnahme und Perspektiven*. Berlin, New York: Springer.
- Funke, Joachim (1990). Systemmerkmale als Determinanten des Umgangs mit dynamischen Systemen. *Sprache und Kognition*, 9(3), S. 143–154.
- Funke, Joachim (1992a). Dealing with Dynamic Systems: Research Strategy, Diagnostic Approach and Experimental Results. *German Journal of Psychology*, 16(1), S. 24–43.
- Funke, Joachim (1992b). *Wissen über dynamische Systeme: Erwerb, Repräsentation und Anwendung*. Berlin, New York: Springer.
- Funke, Joachim (1993). Computergestützte Arbeitsproben: Begriffsklärung, Beispiele sowie Entwicklungspotentiale. *Zeitschrift für Arbeits- und Organisationspsychologie A&O*, 37(3), S. 119–129.
- Funke, Joachim (1999). Komplexes Problemlösen - Ein Blick zurück und ein Blick nach vorne. *Psychologische Rundschau*, 50(4), S. 194–197.
- Funke, Joachim (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning*, 7(1), S. 69–89.
- Funke, Joachim (2003). *Problemlösendes Denken*. Stuttgart: Verlag W. Kohlhammer.
- Funke, Joachim (2004). Neue Verfahren zur Erfassung intelligenten Umgangs mit komplexen und dynamischen Anforderungen. In Elisabeth Stern & Jürgen Guthke (Hrsg.), *Perspektiven der Intelligenzforschung* (S. 89–107). Lengerich: Pabst Science Publ.
- Funke, Joachim (Hrsg.) (2006a). *Denken und Problemlösen: Enzyklopädie der Psychologie: Themenbereich C, Theorie und Forschung: Serie 2, Kognition; Band 8*. Göttingen, Bern, Toronto, Seattle: Hogrefe.

- Funke, Joachim (2006b). Lösen komplexer Probleme. In Joachim Funke & Peter A. Frensch (Hrsg.), *Handbuch der Allgemeinen Psychologie - Kognition* (S. 439–445). Göttingen: Hogrefe.
- Funke, Joachim (2010). Complex Problem Solving: A case for complex cognition? *Cognitive Processing*, 11(2), S. 133–142.
- Funke, Joachim & Buchner, Axel (1992). Finite Automaten als Instrumente für die Analyse von wissensgeleiteten Problemlöseprozessen: Vorstellung eines neuen Untersuchungsparadigmas. *Sprache und Kognition*, 11(1), S. 27–37.
- Funke, Joachim & Gerdes, H. (1993). Manuale für Videorekorder: Auswahl von Textinhalten unter Verwendung der Theorie endlicher Automaten. *Zeitschrift für Arbeitswissenschaft*, 47(1), S. 44–49.
- Funke, Joachim & Reuschenbach, Bernd (2011). Einsatz technischer Mittel in der psychologischen Diagnostik. In Lutz F. Hornke, Manfred Amelang & Martin Kersting (Hrsg.), *Leistungs-, Intelligenz- und Verhaltensdiagnostik* (S. 595–631). Göttingen: Hogrefe.
- Funke, Joachim & Spering, Miriam (2006). Methoden der Denk- und Problemlöseforschung. In Joachim Funke (Hrsg.), *Denken und Problemlösen: Enzyklopädie der Psychologie: Themenbereich C, Theorie und Forschung: Serie 2, Kognition; Band 8* (S. 647–744). Göttingen, Bern, Toronto, Seattle: Hogrefe.
- Gebauer, Miriam M. (2013). *Determinanten der Selbstwirksamkeitsüberzeugung von Lehrenden: Schulischer Berufsalltag an Gymnasien und Hauptschulen*. Wiesbaden: Springer.
- Gelman, Rochel (1998). Domain specificity in cognitive development: universals and nonuniversals. In Michel Sabourin, Michèle Robert & Craik, Fergus I. M. (Hrsg.), *XXVI International Congress of Psychology, Montréal, 1996* (S. 557–579). Hove, New York: Psychology Press.
- Gerjets, Peter H. & Scheiter, Katharina (2003). Goal Configurations and Processing Strategies as Moderators Between Instructional Design and Cognitive Load: Evidence From Hypertext-Based Instruction. *Educational Psychologist*, 38(1), S. 33–41.
- Gilbert, John K. (2006). On the nature of 'context' in chemical education. *International Journal of Science Education*, 28(9), S. 957–976.
- Gillen, Julia & Kaufhold, Marisa (2005). Kompetenzanalysen - kritische Reflektion von Begrifflichkeiten und Messmöglichkeiten. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 101(3), S. 364–378.
- Gnahn, Dieter (2010). *Kompetenzen - Erwerb, Erfassung, Instrumente*. Bielefeld: Bertelsmann.
- Goldhammer, Frank & Hartig, Johannes (2012). Interpretation von Testresultaten und Testeichung. In Helfried Moosbrugger & Augustin Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 173–202). Berlin, Heidelberg: Springer.
- Goldhammer, Frank; Kröhne, Ulf; Keßel, Yvonne; Senkbeil, Martin & Ihme, Jan M. (2014). Diagnostik von ICT-Literacy: Multiple-Choice- vs. simulationsbasierte Aufgaben. *Diagnostica*, 60(1), S. 10–21.
- Gomez, Peter & Probst, Gilbert (1995). *Die Praxis des ganzheitlichen Problemlösens: Vernetzt denken, unternehmerisch handeln, persönlich überzeugen*. Bern: P. Haupt.
- Graube, Gabriele (2009). *Technik und Kommunikation: Ein systemischer Ansatz technischer Bildung*. Göttingen: Cuvillier.
- Gray, Wayne D. (2002). Simulated Task Environments: The Role of High-Fidelity Simulations, Scaled Worlds, Synthetic Environments, and Laboratory Tasks in Basic and Applied Cognitive Research. *Cognitive Science Quarterly*, 2(2), S. 205–227.

- Greeno, James G. (1974). Processes of Learning and Comprehension. In Lee W. Gregg (Hrsg.), *Knowledge and Cognition* (S. 17–41). Maryland: Erlbaum.
- Greif, Hajo; Mitrea, Oana & Werner, Matthias (2008). Information und technologische Handlungsfähigkeit. In Hajo Greif, Oana Mitrea & Matthias Werner (Hrsg.), *Information und Gesellschaft* (S. 49–71). s.l.: VS Verl. für Sozialwissenschaften.
- Greiff, Samuel (2012). *Individualdiagnostik der komplexen Problemlösefähigkeit*. Münster: Waxmann.
- Greiff, Samuel & Fischer, Andreas (2013a). Der Nutzen einer komplexen Problemlösekompetenz: Theoretische Überlegungen und empirische Befunde. *Zeitschrift für Pädagogische Psychologie*, 27(1), S. 27–39.
- Greiff, Samuel & Fischer, Andreas (2013b). Measuring Complex Problem Solving: An educational application of psychological theories. *Journal for Educational Research Online*, 5(1), S. 38–58.
- Greiff, Samuel; Fischer, Andreas; Wüstenberg, Sascha; Sonnleitner, Philipp; Brunner, Martin & Martin, Romain (2013). A multitrait-multimethod study of assessment instruments for complex problem solving. *Intelligence*, 41(5), S. 579–596.
- Greiff, Samuel & Funke, Joachim (2010). Systematische Erforschung komplexer Problemlösefähigkeit anhand minimal komplexer Systeme. *Zeitschrift für Pädagogik*, 56(Beiheft), S. 216–227.
- Greiff, Samuel; Holt, Daniel V. & Funke, Joachim (2013). Perspectives on Problem Solving in Educational Assessment: Analytical, Interactive, and Collaborative Problem Solving. *The Journal of Problem Solving*, 5(2), S. 71–91.
- Greiff, Samuel; Krkovic, Katarina & Nagy, Gabriel (2014). The systematic variation of task characteristics facilitates the understanding of task difficulty: A cognitive diagnostic modeling approach to complex problem solving. *Psychological Test and Assessment Modeling*, 56(1), S. 83–103.
- Greiff, Samuel; Stadler, Matthias; Sonnleitner, Philipp; Wolff, Christian & Martin, Romain (2015). Sometimes less is more: Comparing the validity of complex problem solving measures. *Intelligence*, 50, S. 100–113.
- Greiff, Samuel; Wüstenberg, Sascha; Holt, Daniel V.; Goldhammer, Frank & Funke, Joachim (2013). Computer-based assessment of Complex Problem Solving: concept, implementation, and application. *Educational Technology Research and Development*, 61(3), S. 407–421.
- Grob, Urs & Maag Merki, Katharina (2001). *Überfachliche Kompetenzen: Theoretische Grundlegung und empirische Erprobung eines Indikatorensystems*. Bern: P. Lang.
- Grote, Gudela (2005). Menschliche Kontrolle über technische Systeme - Ein irreführendes Postulat. In Katja Karrer & Klaus-Peter Timpe (Hrsg.), *Beiträge zur Mensch-Maschine-Systemtechnik aus Forschung und Praxis: Festschrift für Klaus-Peter Timpe* (S. 65–78). Düsseldorf: Symposium.
- Gruber, Hans (1999). *Erfahrung als Grundlage kompetenten Handelns*. Bern: Huber.
- Gruber, Hans; Mack, Wolfgang & Ziegler, Albert (1999). Wissen und Denken: Eine problematische Beziehung. In Hans Gruber, Wolfgang Mack & Albert Ziegler (Hrsg.), *Wissen und Denken: Beiträge aus Problemlösepsychologie und Wissenspsychologie* (S. 7–16). Wiesbaden: Deutscher Universitätsverlag.

- Gruber, Hans & Stamouli, Eleni (2015). Intelligenz und Vorwissen. In Elke Wild & Jens Möller (Hrsg.), *Pädagogische Psychologie* (S. 25–44). Berlin, Heidelberg: Springer.
- Gschwendtner, Tobias (2008). Ein Kompetenzmodell für die Kraftfahrzeugtechnische Grundbildung. In Reinhold Nickolaus & Heinrich Schanz (Hrsg.), *Didaktik der gewerblich-technischen Berufsbildung: Konzeptionelle Entwürfe und empirische Befunde* (S. 103–119). Baltmannsweiler: Schneider Hohengehren.
- Gschwendtner, Tobias; Abele, Stephan & Nickolaus, Reinhold (2009). Computersimulierte Arbeitsproben: Eine Validierungsstudie am Beispiel der Fehlerdiagnoseleistung von Kfz-Mechatronikern. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 4, S. 557–578.
- Gschwendtner, Tobias; Geißel, Bernd & Nickolaus, Reinhold (2010). Modellierung beruflicher Fachkompetenz in der gewerblich-technischen Grundbildung. *Zeitschrift für Pädagogik*, 56(Beiheft 56), S. 258–269.
- Gut-Glanzmann, Christoph (2012). *Modellierung und Messung experimenteller Kompetenz: Analyse eines large-scale Experimentiertests*. Inauguraldissertation. Universität Basel, Basel. http://edoc.unibas.ch/21284/1/Dissertation_Gut-Glanzmann.pdf. Zugegriffen: 10. Januar 2012.
- Guthke, Jürgen (1996). *Intelligenz im Test: Wege der psychologischen Intelligenzdiagnostik*. Göttingen: Vandenhoeck & Ruprecht.
- Guthke, Jürgen & Beckmann, Jens F. (2004). Intelligenz als "Lernfähigkeit" - Lerntests als Alternative zum herkömmlichen Intelligenztest. In Elsbeth Stern & Jürgen Guthke (Hrsg.), *Perspektiven der Intelligenzforschung* (S. 137–161). Lengerich: Pabst Science Publ.
- Haase, Jana (2005). *Computer aus Nutzerperspektive: Von der Nutzeranalyse zum Interface-Design*. Wiesbaden: Deutscher Universitäts-Verlag.
- Häcker, Hartmut; Leutner, Detlev & Amelang, Manfred (1998). *Standards für pädagogisches und psychologisches Testen*. Göttingen, Bern: Hogrefe; Huber.
- Hahnel, Carolin & Hacker, Miriam (2013). *Methoden der technologiebasierten Itemkonstruktion: Einführung in den CBA ItemBuilder*. Zentrum für internationale Vergleichsstudien (ZIB), Frankfurt am Main, 2013.
- Haider, Hilde (1992). Implizites Wissen und Lernen. Ein Artefakt? *Zeitschrift für experimentelle und angewandte Psychologie*, 39(1), S. 68–100.
- Haider, Hilde & Frensch, Peter A. (1997). Lernmechanismen des kognitiven Fertigkeitserwerbs. *Zeitschrift für experimentelle Psychologie*, 44(4), S. 521–560.
- Harel, David (1987). Statecharts: A Visual Formalism For Complex Systems. *Science of Computer Programming*, 8, S. 231–274.
- Hartig, Johannes (2006). Kompetenzen als Ergebnisse von Bildungsprozessen. *dipf informiert*(10), S. 2–7.
- Hartig, Johannes (2007). Skalierung und Definition von Kompetenzniveaus. In Bärbel Beck & Eckhard Klieme (Hrsg.), *Sprachliche Kompetenzen: Konzepte und Messung ; DESI-Studie (Deutsch-Englisch-Schülerleistungen-International)* (S. 83–99). Weinheim, Basel: Beltz.
- Hartig, Johannes & Frey, Andreas (2012). Konstruktvalidierung und Skalenbeschreibung in der Kompetenzdiagnostik durch die Vorhersage von Aufgabenschwierigkeiten. *Psychologische Rundschau*, 63(1), S. 43–49.
- Hartig, Johannes; Frey, Andreas & Jude, Nina (2012). Validität. In Helfried Moosbrugger & Augustin Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 143–172). Berlin, Heidelberg: Springer.

- Hartig, Johannes & Klieme, Eckhard (2006). Kompetenz und Kompetenzdiagnostik. In Karl Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127–143). Berlin, Heidelberg: Springer Medizin.
- Hayes, Nicky (1995). Kognitive Prozesse - eine Einführung. In Jochen Gerstenmaier (Hrsg.), *Einführung in die Kognitionspsychologie* (S. 11–40). München: Reinhardt.
- Hazotte, Cyril; Mayer, Hélène; Djaghoul, Younes; Latour, Thibaud; Sonnleitner, Philipp; Brunner, Martin, et al. (2011). The Genetics Lab: An Innovative Tool for Assessment of Intelligence by Mean of Complex Problem Solving. In ICIEIS 2011, Azizah A. Manaf, Shamsul Sahibuddin, Rabiah Ahmad, Salwani M. Daud & Eyas El-Qawasmeh (Hrsg.), *International Conference, ICIEIS 2011, Kuala Lumpur, Malaysia, November 14-16, 2011* (S. 296–310). Berlin: Springer.
- Heckhausen, Jutta & Heckhausen, Heinz (2006). Motivation und Handeln: Einführung und Überblick. In Jutta Heckhausen & Heinz Heckhausen (Hrsg.), *Motivation und Handeln: Mit 43 Tabellen* (S. 1–9). Heidelberg: Springer.
- Hedderich, Jürgen & Sachs, Lothar (2016). *Angewandte Statistik: Methodensammlung mit R*. Berlin, Heidelberg: Springer.
- Heinecke, Andreas M. (2012). *Mensch-Computer-Interaktion*. Berlin: Springer.
- Heinrich, Gert & Mairon, Klaus (2008). *Objektorientierte Systemanalyse*. München: Oldenbourg.
- Helmke, Andreas & Hosenfeld, Ingmar (2004). Vergleichsarbeiten - Standards - Kompetenzstufen: Begriffliche Klärung und Perspektiven. In Marold Wosnitza, Andreas Frey & Reinhold Jäger (Hrsg.), *Lernprozess, Lernumgebung und Lerndiagnostik: Wissenschaftliche Beiträge zum Lernen im 21. Jahrhundert* (S. 56–75). Landau: Verlag Empirische Pädagogik.
- Hempel, Carl G. (1974). *Grundzüge der Begriffsbildung in der empirischen Wissenschaft*. Düsseldorf: Bertelsmann Univ.-Verl.
- Herczeg, Michael (2006). *Interaktionsdesign: Gestaltung interaktiver und multimedialer Systeme*. München: Oldenbourg.
- Hesse, Friedrich W. (1982). Effekte des semantischen Kontexts auf die Bearbeitung komplexer Probleme. *Zeitschrift für experimentelle und angewandte Psychologie*, 29(1), S. 62–91.
- Hesse, Friedrich W.; Spies, Kordelia & Lüer, Gerd (1983). Einfluß motivationaler Faktoren auf das Problemlöseverhalten im Umgang mit komplexen Problemen. *Zeitschrift für experimentelle und angewandte Psychologie*, 30(3), S. 400–424.
- Hohmann, Sandra (2002). *Mensch-Maschine-Interface: Studien zu einer Theorie der Mensch-Computer-Interaktion*. Dissertation. Universität Duisburg-Essen, Duisburg-Essen.
- Holling, Heinz; Preckel, Franzis & Vock, Miriam (2004). *Intelligenzdiagnostik*. Göttingen: Hogrefe.
- Holmqvist, Kenneth; Nyström, Marcus & Andersson, Richard (2011). *Eye Tracking: A comprehensive guide to methods and measures*. Oxford: OUP Oxford.
- Hopcroft, John E.; Motwani, Rajeev & Ullman, Jeffrey D. (2011). *Einführung in Automatentheorie, formale Sprachen und Berechenbarkeit*. München: Pearson.
- Hörmann, Hans-Jürgen & Thomas, Michael (1989). Zum Zusammenhang zwischen Intelligenz und komplexem Problemlösen. *Sprache und Kognition*, 8(1), S. 23–31.
- Hörning, Karl H. (1997). Entlastung und Verunsicherung durch Alltagstechnik. In Max Kerner (Hrsg.), *Technik und Angst: Zur Zukunft der industriellen Zivilisation* (S. 217–228). Aachen: Thout.

- Hornke, Lutz F.; Amelang, Manfred & Kersting, Martin (Hrsg.) (2011). *Leistungs-, Intelligenz- und Verhaltensdiagnostik*. Göttingen: Hogrefe.
- Hubig, Christoph (2012). Bildung und Kompetenz. In Christoph Hubig & Heiner Rindermann (Hrsg.), *Bildung und Kompetenz* (S. 7–51). Göttingen: Vandenhoeck & Ruprecht.
- Hübner, R. (1989). Methoden zur Analyse und Konstruktion von Aufgaben zur kognitiven Steuerung dynamischer Systeme. *Zeitschrift für experimentelle und angewandte Psychologie*, 36, S. 221–238.
- Hussy, Walter (1984a). *Denkpsychologie - ein Lehrbuch: Geschichte, Begriffs- und Problemlöseforschung, Intelligenz (Bd. 1)*. Stuttgart: Kohlhammer.
- Hussy, Walter (1984b). *Zum Begriff der Problemschwierigkeit beim komplexen Problemlösen*. Trier: Univ. Trier, Fachber. 1, Psychologie.
- Hussy, Walter (1985). Komplexes Problemlösen - Eine Sackgasse? *Zeitschrift für experimentelle und angewandte Psychologie*, 32(1), S. 55–74.
- Hussy, Walter (1989). Intelligenz und komplexes Problemlösen. *Diagnostica*, 35(1), S. 1–16.
- Hussy, Walter (1990). Eine experimentelle Analyse des Intelligenzfaktors "Bearbeitungsgeschwindigkeit". *Diagnostica*, 36(3), S. 283–298.
- Hussy, Walter (1991). Eine experimentelle Studie zum Intelligenzkonzept "Verarbeitungskapazität". *Diagnostica*, 37(4), S. 314–333.
- Jacobs, Claus & Petermann, Franz (2007). Grundintelligenztest (CFT 20-R) von Rudolf Weiß (2006). *Diagnostica*, 53(2), S. 109–113.
- Jäger, Reinhold (1974). Methoden zur Mittelung von Korrelationen. *Psychologische Beiträge*, 16, S. 417–424.
- Jakobs, Eva-Maria (2005). Technikakzeptanz und Technikteilhabe. *Technikfolgenabschätzung - Theorie und Praxis*, 14(3), S. 68–75.
- Jann, Ben (2009). Diagnostik von Regressionsschätzungen bei kleinen Stichproben (mit einem Exkurs zu logistischer Regression). In Peter Kriwy & Christiane Gross (Hrsg.), *Klein aber fein!: Quantitative empirische Sozialforschung mit kleinen Fallzahlen* (S. 93–125). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Jenßen, Lars; Dunekacke, Simone & Blömeke, Sigrid (2015). Qualitätssicherung in der Kompetenzforschung: Empfehlungen für den Nachweis von Validität in Testentwicklung und Veröffentlichungspraxis. In Sigrid Blömeke & Olga Zlatkin-Troitschanskaia (Hrsg.), *Kompetenzen von Studierenden* (S. 11–31). Weinheim, Basel: Beltz.
- Jerusalem, Matthias & Mittag, Waldemar (1999). Selbstwirksamkeit, Bezugsnormen, Leistung und Wohlbefinden in der Schule. In Matthias Jerusalem & Reinhard Pekrun (Hrsg.), *Emotion, Motivation und Leistung* (S. 223–245). Göttingen: Hogrefe.
- Johannsen, Gunnar (1993). *Mensch-Maschine-Systeme*. Berlin, Heidelberg: Springer.
- Jude, Nina & Wirth, Joachim (2007). Neue Chancen bei der technologiebasierten Erfassung von Kompetenzen. In Johannes Hartig & Eckhard Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik: Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (S. 49–56). Bonn: Bundesministerium für Bildung und Forschung.
- Jung, Eberhard (2010). *Kompetenzerwerb: Grundlagen, Didaktik, Überprüfbarkeit*. München: Oldenbourg.
- Jurecka, Astrid & Hartig, Johannes (2007). Computer- und netzwerkbasierendes Assessment. In Johannes Hartig & Eckhard Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen*

- technologiebasierter Kompetenzdiagnostik: Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (S. 37–48). Bonn: Bundesministerium für Bildung und Forschung.
- Kaiser, A. (1998). *carte de competence: Wie lassen sich Kompetenzen feststellen?* *GdWZ*, 9(5), S. 199–201.
- Kane, Michael T.; Hambrick, David Z. & Conway, Andrew R.A. (2005). Working Memory Capacity and Fluid Intelligence Are Strongly Related Constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131(1), S. 66–71.
- Kauertz, Alexander (2008). *Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben*. Berlin: Logos.
- Kaufhold, Marisa (2006). *Kompetenz und Kompetenzerfassung: Analyse und Beurteilung von Verfahren der Kompetenzerfassung*. Wiesbaden: VS Verl. für Sozialwissenschaften.
- Kersting, Martin (1999). *Diagnostik und Personalauswahl mit computergestützten Problemlöseszenarien? Zur Kriteriumsvalidität von Problemlöseszenarien und Intelligenztests*. Göttingen: Hogrefe.
- Klieme, Eckhard; Avenarius, Hermann; Blum, Werner; Döbrich, Peter; Gruber, Hans; Prenzel, Manfred, et al. (2007). Zur Entwicklung nationaler Bildungsstandards. In Bundesministerium für Bildung und Forschung (Hrsg.), *Zur Entwicklung nationaler Bildungsstandards: Eine Expertise* (S. 7–174). Bonn, Berlin.
- Klieme, Eckhard & Hartig, Johannes (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In Manfred Prenzel, Ingrid Gogolin & Heinz-Hermann Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 11–29). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Klieme, Eckhard; Maag-Merki, Katharina & Hartig, Johannes (2007). Kompetenzbegriff und Bedeutung von Kompetenzen im Bildungswesen. In Johannes Hartig & Eckhard Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik: Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (S. 5–15). Bonn: Bundesministerium für Bildung und Forschung.
- Kline, Rex B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Klix, Friedhart (1971). *Information und Verhalten*. Bern, Stuttgart, Wien: H. Huber.
- Kluge, Annette (2004). *Wissenserwerb für das Steuern komplexer Systeme*. Lengerich, Berlin: Pabst Science Publ.
- Kluwe, Rainer H. (1988). Methoden der Psychologie zur Gewinnung von Daten über menschliches Wissen. In Heinz Mandl, Hans Spada, Hans Aebli & Friedhart Klix (Hrsg.), *Wissenspsychologie* (S. 359–385). München: Psychologie-Verl.-Union.
- Kluwe, Rainer H. (1991). Zum Problem der Wissensvoraussetzungen für Prozeß- und Systemsteuerung. In Friedhart Klix & Elke van der Meer (Hrsg.), *Experimentalpsychologie heute: 100 Jahre Zeitschrift für Psychologie* (S. 311–324). Heidelberg: Barth.
- Kluwe, Rainer H. (1997a). Acquisition of knowledge in the control of a simulated technical system. *Le Travail humain*, 60(1), S. 61–85.
- Kluwe, Rainer H. (1997b). Informationsverarbeitung, Wissen und mentale Modelle beim Umgang mit komplexen Systemen. In Karlheinz Sonntag & Niclas Schaper (Hrsg.), *Störungsmanagement und Diagnosekompetenz: Leistungskritisches Denken und Handeln in komplexen technischen Systemen* (S. 13–37). Zürich: Vdf, Hochschulverl. an der ETH Zürich.

- Kluwe, Rainer H. & Haider, Hilde (1990). Modelle zur internen Repräsentation komplexer technischer Systeme. *Sprache und Kognition*, 9(4), S. 173–192.
- Kluwe, Rainer H. & Haider, Hilde (1995). Erwerb kognitiver Fertigkeiten durch Übung. In Dietrich Dörner & Elke van der Meer (Hrsg.), *Das Gedächtnis: Probleme - Trends - Perspektiven* (S. 253–291). Göttingen: Hogrefe.
- Kluwe, Rainer H. & Reimann, Hans (1983). *Problemlösen bei vernetzten, komplexen Problemen: Effekte des Verbalisierens auf die Problemlöseleistung*. Hamburg: Bericht aus dem Forschungsbereich Pädagogik der Hochschule der Bundeswehr.
- Kluwe, Rainer H.; Schilde, Andree; Fischer, Christian & Oelierer, Norbert (1991). Problemlöseleistungen beim Umgang mit komplexen Systemen und Intelligenz. *Diagnostica*, 37(4), S. 291–313.
- Knoblich, Günther & Rhenius, Detlef (1995). Zur Reaktivität Lauten Denkens beim komplexen Problemlösen. *Zeitschrift für experimentelle Psychologie*, 42(3), S. 419–454.
- Kölbach, Eva (2011). *Kontexteinflüsse beim Lernen mit Lösungsbeispielen*. Berlin: Logos.
- Kotkamp, Uwe (1999). *Elementares und komplexes Problemlösen: Über Invarianzeigenschaften von Denkprozessen*. Lengerich: Pabst.
- Krause, Ulrike-Marie & Stark, Robin (2006). Vorwissen aktivieren. In Heinz Mandl & Helmut F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 38–49). Göttingen: Hogrefe.
- Krause, Werner (1982a). Problemlösen - Stand und Perspektiven Teil I. *Zeitschrift für Psychologie*, 190(1), S. 17–36.
- Krause, Werner (1982b). Problemlösen - Stand und Perspektiven Teil II. *Zeitschrift für Psychologie*, 190(2), S. 141–169.
- Kreuzpointner, Ludwig (2010). *Bedingungen für die Äquivalenz von Papier-Bleistift-Version und Computerversion bei Leistungstests*. Dissertation. Universität Regensburg, Regensburg.
- Kromrey, Helmut (2009). *Empirische Sozialforschung: Modelle und Methoden der standardisierten Datenerhebung und Datenauswertung*. Stuttgart: Lucius & Lucius.
- Kröner, Stephan (2001). *Intelligenzdiagnostik per Computersimulation*. Dissertation; Pädag. Hochsch., Erfurt. Münster: Waxmann.
- Kuhl, Julius (1983). Emotion, Kognition und Motivation: II. Die funktionale Bedeutung der Emotionen für das problemlösende Denken und für das konkrete Handeln. *Sprache und Kognition*, 2(4), S. 228–253.
- Kyllonen, Patrick C. & Christal, Raymond E. (1990). Reasoning Ability Is (Little More Than) Working-Memory Capacity?! *Intelligence*, 14(4), S. 389–433.
- Lang, Martin (2004). *Berufspädagogische Perspektiven netzbasierter Lernumgebungen in der betrieblichen Aus- und Weiterbildung*. Dissertation, Universität Dortmund. Bochum: Projektverlag.
- Langens, Thomas A. (2009). Leistung. In Veronika Brandstätter & Jürgen H. Otto (Hrsg.), *Handbuch der Allgemeinen Psychologie - Motivation und Emotion* (S. 217–224). Göttingen: Hogrefe.
- Lei, Pui-Wa & Wu, Qiong (2015). Estimation in Structural Equation Modeling. In Rick H. Hoyle (Hrsg.), *Handbook of structural equation modeling* (S. 164–180). New York: Guilford Press.
- Lengning, Anke (2009). Neugier und Exploration. In Veronika Brandstätter & Jürgen H. Otto (Hrsg.), *Handbuch der Allgemeinen Psychologie - Motivation und Emotion* (S. 252–257). Göttingen: Hogrefe.

- Leonhart, Rainer (2010). *Datenanalyse mit SPSS*. Göttingen: Hogrefe.
- Leutner, Detlev (2002). The fuzzy relationship of intelligence and problem solving in computer simulations. *Computers in Human Behavior*, 18, S. 685–697.
- Leutner, Detlev; Fleischer, Jens; Wirth, Joachim; Greiff, Samuel & Funke, Joachim (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsvergleichsstudien: Untersuchungen zur Dimensionalität. *Psychologische Rundschau*, 63(1), S. 34–42.
- Leutner, Detlev; Funke, Joachim; Klieme, Eckhard & Wirth, Joachim (2005). Problemlösefähigkeit als fächerübergreifende Kompetenz. In Eckhard Klieme, Detlev Leutner & Joachim Wirth (Hrsg.), *Problemlösekompetenz von Schülerinnen und Schülern: Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studien* (S. 11–19). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Leutner, Detlev; Klieme, Eckhard; Meyer, K. & Wirth, Joachim (2004). Problemlösen. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (S. 147–175). Münster: Waxmann.
- Lewis-Beck, Colin & Lewis-Beck, Michael S. (2016). *Applied regression: An introduction*. Thousand Oaks: Sage Publ.
- Lienert, Gustav A. & Raatz, Ulrich (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Lunze, Jan (2009). *Ereignisdiskrete Systeme: Modellierung und Analyse dynamischer Systeme mit Automaten, Markovketten und Petrinetzen*. München, Wien: Oldenbourg.
- Maag Merki, Katharina (2004). Lernkompetenzen als Bildungsstandards - eine Diskussion der Umsetzungsmöglichkeiten. *Zeitschrift für Erziehungswissenschaft*, 7(4), S. 537–550.
- MacCallum, Robert C.; Browne, Michael W. & Sugawara, Hazuki M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods*, 1(2), S. 130–149.
- Mack, Wolfgang (1999). Intelligenz und Wissen. In Hans Gruber, Wolfgang Mack & Albert Ziegler (Hrsg.), *Wissen und Denken: Beiträge aus Problemlösepsychologie und Wissenspsychologie* (S. 119–150). Wiesbaden: Deutscher Universitätsverlag.
- Mammes, Ingelore (2001). *Förderung des Interesses an Technik: Eine Untersuchung zum Einfluss technischen Sachunterrichts auf die Verringerung von Geschlechterdifferenzen im technischen Interesse*. Frankfurt am Main: Lang.
- Mandl, Heinz; Friedrich, Helmut F. & Hron, Aemilian (1988). Theoretische Ansätze zum Wissenserwerb. In Heinz Mandl, Hans Spada, Hans Aebli & Friedhart Klix (Hrsg.), *Wissenspsychologie* (S. 123–160). München: Psychologie-Verl.-Union.
- Maslow, Abraham H. (1970). *Motivation and personality*. New York, Evanston, London: Harper & Row, Publishers.
- Mayer, Richard E. (1979). *Denken und Problemlösen: Eine Einführung in menschliches Denken und Lernen*. Berlin, New York: Springer.
- McDonald, Roderick P. (1970). The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *The British Journal of Mathematical and Statistical Psychology*, 23(1), S. 1–21.
- Meili, Richard & Steingrüber, Hans-Joachim (1978). *Lehrbuch der psychologischen Diagnostik*. Bern: Huber.

- Merz, Ferdinand (1969). Der Einfluß des Verbalisierens auf die Leistung bei Intelligenzaufgaben. *Zeitschrift für experimentelle und angewandte Psychologie*, 16, S. 114–137.
- Messick, Samuel (1995). Validity of Psychological Assessment: Validation of Inferences From Persons' Responses and Performances as Scientific Inquiry Into Score Meaning. *American Psychologist*, 50(9), S. 741–749.
- Mielke, Rosemarie (1984). *Lernen und Erwartung: Zur Selbst-Wirksamkeits-Theorie von Albert Bandura*. Stuttgart: Huber.
- Möller, Carsten (2003). *Dreamteam Mensch-Computer: Neue Wege zur nutzergerechten Gestaltung des Mensch-Maschine-Dialogs*. Dissertation. Universität Duisburg-Essen, Essen.
- Molnár, Gyöngyvér; Greiff, Samuel & Csapó, Benő (2013). Inductive reasoning, domain specific and complex problem solving: Relations and development. *Thinking Skills and Creativity*, 9, S. 35–45.
- Moosbrugger, Helfried (2011). *Lineare Modelle: Regressions- und Varianzanalysen*. Bern: Verlag Hans Huber.
- Moosbrugger, Helfried & Kelava, Augustin (2012a). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In Helfried Moosbrugger & Augustin Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 7–26). Berlin, Heidelberg: Springer.
- Moosbrugger, Helfried & Kelava, Augustin (Hrsg.) (2012b). *Testtheorie und Fragebogenkonstruktion*. Berlin, Heidelberg: Springer.
- Müller, Horst (1993). *Komplexes Problemlösen: Reliabilität und Wissen*. Dissertation, Universität Bonn. Bonn: Holos-Verl.
- Muthén, Bengt O. (1993). Goodness of fit with categorical and other nonnormal variables. In Kenneth A. Bollen & J. S. Long (Hrsg.), *Testing structural equation models* (S. 205–234). Newbury Park: Sage Publ.
- Muthén, Bengt O. (1998). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In Linda M. Collins & Aline Sayer (Hrsg.), *New methods for the analysis of change* (S. 291–322). Washington: American Psychological Association.
- Myers, David G. (2014). Intelligenz. In David G. Myers (Hrsg.), *Psychologie* (S. 399–436). Berlin: Springer.
- Nakhosteen, C. B. (2009). *Technisches Erfahrungswissen in industriellen Produktionsprozessen*. Aachen, Dortmund: Shaker.
- Naumann, Johannes; Goldhammer, Frank; Rölke, Heiko & Stelter, Annette (2014). Erfolgreiches Problemlösen in technologiebasierten Umgebungen: Wechselwirkungen zwischen Interaktionsschritten und Aufgabenanforderungen. *Zeitschrift für Pädagogische Psychologie*, 28(4), S. 193–203.
- Neber, Heinz (1987). Angewandte Problemlösepsychologie. In Heinz Neber (Hrsg.), *Angewandte Problemlösepsychologie* (S. 1–117). Münster: Aschendorff.
- Neubert, Jonas C.; Kretschmar, André; Wüstenberg, Sascha & Greiff, Samuel (2015). Extending the Assessment of Complex Problem Solving to Finite State Automata: Embracing Heterogeneity. *European Journal of Psychological Assessment*, 31(3), S. 181–194.
- Nevitt, Jonathan & Hancock, Gregory R. (2001). Performance of Bootstrapping Approaches to Model Test Statistics and Parameter Standard Error Estimation in Structural Equation Modeling. *Structural Equation Modeling*, 8(3), S. 353–377.

- Newell, Allen & Simon, Herbert A. (1972). *Human problem solving*. Englewood Cliffs: Prentice-Hall.
- Nickolaus, Reinhold; Abele, Stephan; Gschwendtner, Tobias; Nitzschke, Alexander & Greiff, Samuel (2012). Fachspezifische Problemlösefähigkeit in gewerblich-technischen Ausbildungsberufen - Modellierung, erreichte Niveaus und relevante Einflussfaktoren. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 108(2), S. 243–272.
- Noack, Jörg (Hrsg.) (2001). *Techniken der objektorientierten Softwareentwicklung: Mit 74 Tabellen*. Berlin: Springer.
- Norman, Donald A. & Rumelhart, David E. (1978). Gedächtnis und Wissen. In Donald A. Norman, David E. Rumelhart, A. A. Abrahamson & Hans Aebli (Hrsg.), *Strukturen des Wissens: Wege der Kognitionsforschung* (S. 21–47). Stuttgart: Klett-Cotta.
- Oberauer, Klaus; Schulze, Ralf; Wilhelm, Oliver & Süß, Heinz-Martin (2005). Working memory and intelligence - their correlation and their relation: comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131(1), S. 61-5; author reply 72-5.
- Obermayer, Richard W. (1964). Simulation, Models, and Games: Sources of Measurement. *Human Factors*, 6(6), S. 607–619.
- OECD (2003). *The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills*: OECD Publishing.
- OECD (2013a). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*: OECD Publishing.
- OECD (2013b). *PISA 2012 Results: Ready to Learn: Students' engagement, drive and self-beliefs (Volume III)*: OECD Publishing.
- OECD (2013c). *Technical Report of the Survey of Adult Skills (PIAAC)*. http://www.oecd.org/site/piaac/_Technical%20Report_17OCT13.pdf. Zugegriffen: 28. Dezember 2013.
- OECD (2014a). *PISA 2012 Results: Creative Problem Solving (Volume V)*: OECD Publishing.
- OECD (2014b). *PISA 2012 Technical Report*.
- OECD PISA Deutschland (2003). *Erfassung fächerübergreifender Problemlösekompetenzen in PISA*. www.mpib-berlin.mpg.de/Pisa/Problemloesen-2.pdf. Zugegriffen: 27. August 2012.
- Ohr, Dieter (2010). Lineare Regression: Modellannahmen und Regressionsdiagnostik. In Christof Wolf & Henning Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 639–675). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Olson, Kenneth; Camp, Cameron & Fuller, Dana (1984). Curiosity and need for cognition. *Psychological Reports*, 54, S. 71–74.
- Olsson, Ulf (1979). Maximum Likelihood Estimation of the polychoric Correlation Coefficient. *Psychometrika*, 44(4), S. 443–460.
- Opp, Karl-Dieter (2010). Kausalität als Gegenstand der Sozialwissenschaften und der multivariaten Statistik. In Christof Wolf & Henning Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 9–38). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Opp, Karl-Dieter (2014). *Methodologie der Sozialwissenschaften: Einführung in Probleme ihrer Theorienbildung und praktischen Anwendung*. Wiesbaden: Imprint: Springer VS.
- Opwis, Klaus (1988). Produktionssysteme. In Heinz Mandl, Hans Spada, Hans Aebli & Friedhart Klix (Hrsg.), *Wissenspsychologie* (S. 74–98). München: Psychologie-Verl.-Union.

- Opwis, Klaus; Beller, Sieghard; Spada, Hans & Lüer, Gerd (2006). Problemlösen, Denken, Entscheiden. In Hans Spada (Hrsg.), *Lehrbuch Allgemeine Psychologie* (S. 197–275). Bern: H. Huber.
- Pahl, Gerhard; Beitz, Wolfgang; Feldhusen, Jörg & Grote, Karl-Heinrich (2007). *Konstruktionslehre: Grundlagen erfolgreicher Produktentwicklung. Methoden und Anwendung*. Berlin: Springer.
- Pajares, Frank & Kranzler, John (1995). Self-Efficacy Beliefs and General Mental Ability in Mathematical Problem-Solving. *Contemporary Educational Psychology*, 20(4), S. 426–443.
- Pajares, Frank & Miller, David M. (1994). Role of Self-Efficacy and Self-Concept Beliefs in Mathematical Problem Solving: A Path Analysis. *Journal of Educational Psychology*, 86(2), S. 193–203.
- Pechtl, Hans (2009). *Anmerkungen zur Operationalisierung und Messung des Konstrukts "need for cognition"* (Wirtschaftswissenschaftliche Diskussionspapiere 05/2009). Greifswald. <https://www.econstor.eu/bitstream/10419/41073/1/614470943.pdf>. Zugegriffen: 9. Januar 2017.
- Pfeiffer, Iris (2000). *Berufliche Umorientierung: Ressourcen und Risikofaktoren: Eine empirische Studie zum Umschulungserfolg im Kontext der beruflichen Rehabilitation*. Dissertation. Freie Universität Berlin, Berlin.
- Pfenning, Uwe & Renn, Ortwin (2010). *Nachwuchsbarometer Technikwissenschaften: Wissenschaftlicher Abschlussbericht der Universität Stuttgart*. Berlin.
- Pospeschill, Markus (2010). *Testtheorie, Testkonstruktion, Testevaluation*. München: Reinhardt.
- Pospeschill, Markus (2013). *Empirische Methoden in der Psychologie*. München, Stuttgart: Reinhardt; UTB.
- Preckel, Franzis (2014). Assessing Need for Cognition in Early Adolescence. *European Journal of Psychological Assessment*, 30(1), S. 65–72.
- Preim, Bernhard (1999). *Entwicklung interaktiver Systeme: Grundlagen, Fallbeispiele und innovative Anwendungsfelder*. Berlin: Springer.
- Preim, Bernhard & Dachselt, Raimund (2010). *Interaktive Systeme*. Heidelberg: Springer.
- Prenzel, Manfred; Häußler, Peter; Rost, Jürgen & Senkbeil, Martin (2002). Der PISA-Naturwissenschaftstest: Lassen sich die Aufgabenschwierigkeiten vorhersagen? *Unterrichtswissenschaft*, 30, S. 120–135.
- Preußler, Walburga (1998). Strukturwissen als Voraussetzung für die Steuerung komplexer dynamischer Systeme. *Zeitschrift für experimentelle Psychologie*, 45(3), S. 218–240.
- Putz-Osterloh, Wiebke (1981). *Problemlöseprozesse und Intelligenztestleistung*. Bern: Huber.
- Putz-Osterloh, Wiebke (1988). Wissen und Problemlösen. In Heinz Mandl, Hans Spada, Hans Aebli & Friedhart Klix (Hrsg.), *Wissenspsychologie* (S. 247–263). München: Psychologie-Verl.-Union.
- Putz-Osterloh, Wiebke; Bott, Birgit & Houben, I. (1988). Beeinflusst Wissen über ein realitätsnahes System dessen Steuerung? *Sprache und Kognition*, 10, S. 240–251.
- Putz-Osterloh, Wiebke & Lüer, Gerd (1981). Über die Vorhersagbarkeit komplexer Problemlöseleistungen durch Ergebnisse in einem Intelligenztest. *Zeitschrift für experimentelle und angewandte Psychologie*, 28(2), S. 309–334.
- R core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna Austria.
- Raaheim, Kjell (1988). Intelligence and Task Novelty. *Advances in the psychology of human intelligence*, 4, S. 73–97.

- Rammert, Werner (2006). *Technik, Handeln und Sozialstruktur: Eine Einführung in die Soziologie der Technik: Technical University Technology Studies Working Papers* (Technical University Technology Studies Working Papers TUTS-WP-3-2006).
- Rammert, Werner & Schulz-Schaeffer, Ingo (2002). *Können Maschinen handeln? Soziologische Beiträge zum Verhältnis von Mensch und Technik*. Frankfurt am Main: Campus.
- Raykov, Tenko & Marcoulides, George A. (2006). *A first course in structural equation modeling*. Mahwah: Lawrence Erlbaum Associates.
- Reinecke, Jost (2014). *Strukturgleichungsmodelle in den Sozialwissenschaften*. München: Oldenbourg.
- Reisinger, Monika; Svecnik, Erich & Schwetz, Herbert (2012). *Fehlende Werte und keine Normalverteilung? Tipps und Abhilfen für das quantitativ orientierte Forschen*. Wien: Facultas.
- Renkl, Alexander (1996). Vorwissen und Schulleistungen. In Jens Möller & Olaf Köller (Hrsg.), *Emotionen, Kognitionen und Schulleistung* (S. 175–190). Weinheim: Beltz.
- Renkl, Alexander (2015). Wissenserwerb. In Elke Wild & Jens Möller (Hrsg.), *Pädagogische Psychologie* (S. 3–24). Berlin, Heidelberg: Springer.
- Renkl, Alexander; Gruber, Hans; Weber, Sandra; Lerche, Thomas & Schweizer, Karl (2003). *Cognitive Load beim Lernen aus Lösungsbeispielen: (Forschungsbericht Nr. 4)*.
- Rheinberg, Falko (2006). *Motivation*. Stuttgart: Kohlhammer.
- Rheinberg, Falko; Vollmeyer, Regina & Lehnik, Anke (2000). Selbstkonzept der Begabung, Erfolgserwartungen und Lernleistung. In Friedrich Försterling, Joachim Stiensmeier-Pelster & Lily-Maria Silny (Hrsg.), *Kognitive und emotionale Aspekte der Motivation* (S. 77–97). Göttingen: Hogrefe.
- Richter, Tobias; Naumann, Johannes & Noller, Stephan (1999). Computer Literacy und computerbezogene Einstellungen: Zur Vergleichbarkeit von Online- und Paper-Pencil-Erhebungen. In Ulf-Dietrich Reips, Bernad Batinic, Wolfgang Bandilla, Michael Bosnjak, Lorenz Gräf, Klaus Moser, et al. (Hrsg.). Zürich: Online Press.
- Riedwyl, Hans & Ambühl, Mathias (2000). *Statistische Auswertungen mit Regressionsprogrammen: Lineare Regression und Verwandtes, multivariate Statistik, Planung und Auswertung von Versuchen*. München: Oldenbourg.
- Rindermann, Heiner (2012). Bildung und Kompetenz als bürgerliches Phänomen. In Christoph Hubig & Heiner Rindermann (Hrsg.), *Bildung und Kompetenz* (S. 52–122). Göttingen: Vandenhoeck & Ruprecht.
- Rölke, Heiko (2012). The ItemBuilder: A Graphical Authoring System for Complex Item Development. In Theo Bastiaens & Gary Marks (Hrsg.), *E-Learn 2012 World Conference on E-Learning in Corporate, Gouvernmet, Healthcare, and Higher Education, Montréal, Québec, 9-12 Oktober 2012* (S. 344–353). Chesapeake: Association for the Advancement of Computing in Education (AACE).
- Rollett, Wolfram (2008). *Strategieinsatz, erzeugte Information und Informationsnutzung bei der Exploration und Steuerung komplexer dynamischer Systeme*. Berlin: LIT-Verl.
- Ropohl, Günter (1990). Technisches Problemlösen und soziales Umfeld. In Friedrich Rapp (Hrsg.), *Technik und Philosophie* (S. 109–167). Düsseldorf: VDI Verlag.
- Ropohl, Günter (1999). *Allgemeine Technologie: Eine Systemtheorie der Technik*. München: C. Hanser.
- Rosseel, Yves (2015). *lavaan: An R Package for Structural Equation Modeling*.
- Rost, Detlef H. (2009). *Intelligenz*. Weinheim, Basel: Beltz.

- Rost, Jürgen (2004). *Lehrbuch Testtheorie - Testkonstruktion*. Bern: Huber.
- Rothermund, Klaus & Eder, Andreas (2011). *Allgemeine Psychologie: Motivation und Emotion*. Wiesbaden: VS Verl. für Sozialwissenschaften.
- Rudolf, Matthias & Müller, Johannes (2012). *Multivariate Verfahren: Eine praxisorientierte Einführung mit Anwendungsbeispielen in SPSS*. Göttingen: Hogrefe.
- Rudolph, Julia; Niepel, Christoph; Wüstenberg, Sascha & Greiff, Samuel (2015). *Prädiktoren komplexen Problemlösens besser verstehen: Bearbeitungszeit als partieller Mediator für den Effekt von Reasoning und Need for Cognition auf komplexes Problemlösen*. 3. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Bochum, 2015.
- Rupp, André A.; Leucht, Michael & Hartung, Ralph (2006). "Die Kompetenzbrille aufsetzen": Verfahren zur multiplen Klassifikation von Lernenden für Kompetenzdiagnostik in Unterricht und Testung. *Unterrichtswissenschaft*, 34(3), S. 195–219.
- Rüppel, Hermann; Hinnersmann, Herwig & Wiegand, Jan (1987). Problemlösen - Allgemein oder Spezifisch? In Heinz Neber (Hrsg.), *Angewandte Problemlösepsychologie* (S. 173–192). Münster: Aschendorff.
- Rupprecht, Werner (2014). *Einführung in die Theorie der kognitiven Kommunikation: Wie Sprache, Information, Energie, Internet, Gehirn und Geist zusammenhängen*. Wiesbaden: Springer Vieweg.
- Saam, Nicole J. (1997). Begriffsdefinitionen: Modell - Simulation - Theorie. In Astrid Lux-Endrich & Anne Wachsmann (Hrsg.), *Konstruierte Wirklichkeiten: Modellbildung und Simulation in den Wissenschaften*. Dokumentation der Ferienakademie für Studentinnen und junge Wissenschaftlerinnen in Tutzing vom 10. bis 12. Oktober 1997 (S. 13–17).
- Sachs, Burkhard (2005). Technikunterricht: Bedingungen und Perspektiven. http://www.eduhi.at/dl/Technikbegriff_Sachs_-_tu_100.pdf. Zugegriffen: 19. September 2012.
- Sackmann, Reinhold; Weymann, Ansgar & Hüttner, Bernd (1994). *Die Technisierung des Alltags: Generationen und technische Innovationen*. Frankfurt, New York: Campus-Verlag.
- Sandmann, Angela (2014). Lauter Denken - die Analyse von Denk-, Lern- und Problemlöseprozessen. In Dirk Krüger, Ilka Parchmann & Horst Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 179–188). Berlin, Heidelberg: Imprint: Springer Spektrum.
- Sarodnick, Florian & Brau, Henning (2011). *Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung*. Bern: Huber.
- Satorra, Albert & Bentler, Peter M. (1988). Scaling Corrections for Chi-square Statistics in Covariance Structure Analysis. In American Statistical Association (Hrsg.), *Alexandria* (S. 308–313).
- Satorra, Albert & Bentler, Peter M. (2001). A Scaled Difference Chi-Square Test Statistic for Moment Structure Analysis. *Psychometrika*, 66(4), S. 507–514.
- Schabram, Katharina (2007). *Lernaufgaben im Unterricht: Instruktionspsychologische Analysen am Beispiel der Physik*. Dissertation. Universität Duisburg-Essen, Duisburg-Essen.
- Schachtner, Christina (1997). Die Technik und das Soziale: Begründung einer subjektivitätsorientierten Technikforschung. In Christina Schachtner (Hrsg.), *Technik und Subjektivität: Das Wechselverhältnis zwischen Mensch und Computer aus interdisziplinärer Sicht* (S. 7–25). Frankfurt am Main: Suhrkamp.

- Schaper, Niclas (2009). Online-Tests aus diagnostisch-methodischer Sicht. In Heinke Steiner (Hrsg.), *Online-Assessment: Grundlagen und Anwendung von Online-Tests in der Unternehmenspraxis* (S. 17–36). Heidelberg: Springer.
- Schaub, Harald (1993). *Modellierung der Handlungsorganisation*. Bern: Huber.
- Schaub, Harald (2001). *Persönlichkeit und Problemlösen: Persönlichkeitsfaktoren als Parameter eines informationsverarbeitenden Systems*. Weinheim, Basel: Beltz.
- Schaub, Harald & Reimann, Ralph (1999). Zur Rolle des Wissens beim komplexen Problemlösen. In Hans Gruber, Wolfgang Mack & Albert Ziegler (Hrsg.), *Wissen und Denken: Beiträge aus Problemlösepsychologie und Wissenspsychologie* (S. 169–191). Wiesbaden: Deutscher Universitätsverlag.
- Schendera, Christian F.G. (2014). *Regressionsanalyse mit SPSS*. München: De Gruyter.
- Scherer, Ronny (2014). Komplexes Problemlösen im Fach Chemie: Ein domänenspezifischer Zugang. *Zeitschrift für Pädagogische Psychologie*, 28(4), S. 181–192.
- Schiefele, Ulrich & Schaffner, Ellen (2015). Motivation. In Elke Wild & Jens Möller (Hrsg.), *Pädagogische Psychologie* (S. 153–175). Berlin, Heidelberg: Springer.
- Schley, Thomas; Eigenmann, Rebecca & Siegfried, Christin (2015). Was macht ein Problem zu einem Problem? - Modellierung der Schwierigkeit von Problemlöseszenarien für den Ausbildungsberuf Indutrieauffrau/-mann. *Wirtschaft & Erziehung*, 67(4), S. 140–146.
- Schmeck, Annett (2010). *Visualisieren naturwissenschaftlicher Sachverhalte: Der Einsatz von vorgegebenen und selbst generierten Visualisierungen als Textverstehenshilfen beim Lernen aus naturwissenschaftlichen Sachtexten*. Dissertation. Universität Duisburg-Essen, Essen.
- Schmidt-Atzert, Lothar & Amelang, Manfred (2012). *Psychologische Diagnostik*. Berlin, Heidelberg: Springer.
- Schmiemann, Philipp & Lücken, Markus (2014). Validität - Misst mein Test, was er soll? In Dirk Krüger, Ilka Parchmann & Horst Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 107–118). Berlin, Heidelberg: Imprint: Springer Spektrum.
- Schmitz, Andreas & Yanenko, Olga (2014). Web Server Logs und Logfiles. In Nina Baur & Jörg Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 847–854). Wiesbaden: Springer VS.
- Schnell, Rainer; Hill, Paul B. & Esser, Elke (2013). *Methoden der empirischen Sozialforschung*. München: Oldenbourg.
- Schnotz, Wolfgang (1994). *Aufbau von Wissensstrukturen: Untersuchungen zur Kohärenzbildung beim Wissenserwerb mit Texten*. Habilitation, Univ. Tübingen. Weinheim: Beltz.
- Schooler, Jonathan W.; Ohlsson, Stellan & Brooks, Kevin (1993). Thoughts Beyond Words: When Language Overshadows Insight. *Journal of Experimental Psychology: General*, 122(2), S. 166–183.
- Schoppek, Wolfgang (1997). Wissen bei der Steuerung dynamischer Systeme - ein prozeßorientierter Forschungsansatz. *Zeitschrift für Psychologie*, 205, S. 269–295.
- Schoppek, Wolfgang & Putz-Osterloh, Wiebke (2003). Individuelle Unterschiede und die Bearbeitung komplexer Probleme. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24(3), S. 163–173.
- Schott, Franz & Ghanbari, Shahram Azizi (2009). Modellierung, Vermittlung und Diagnostik der Kompetenz kompetenzorientiert zu unterrichten - wissenschaftliche Herausforderung und ein praktischer Lösungsversuch. In Niclas Schaper, Annegret H. Hilligus & Peter

- Reinhold (Hrsg.), *Kompetenzmodellierung und -messung in der Lehrerbildung* (S. 10–27). Landau: Empirische Pädagogik.
- Schreiber, Nico; Theyßen, Heike & Schecker, Horst (2014). Diagnostik experimenteller Kompetenz: Kann man Realexperimente durch Simulationen ersetzen? *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), S. 161–173.
- Schumann, Stephan & Eberle, Franz (2011). Bedeutung und Verwendung schwierigkeitsbestimmender Aufgabenmerkmale für die Erfassung ökonomischer und beruflicher Kompetenzen. In Uwe Faßhauer, Bärbel Fürstenau & Eveline Wuttke (Hrsg.), *Grundlagenforschung zum Dualen System und Kompetenzentwicklung in der Lehrerbildung* (S. 77–89). Leverkusen: Budrich, Barbara.
- Schweizer, Fabian; Wüstenberg, Sascha & Greiff, Samuel (2013). Validity of the MicroDYN approach: Complex problem solving predicts school grades beyond working memory capacity. *Learning and Individual Differences*, 24, S. 42–52.
- Schweizer, Karl (2006a). Intelligenz. In Karl Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 2–15). Berlin, Heidelberg: Springer Medizin.
- Schweizer, Karl (2006b). Intelligenzdiagnostik. In Karl Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 70–83). Berlin, Heidelberg: Springer Medizin.
- Seeber, Susan & Nickolaus, Reinhold (2010). Kompetenz, Kompetenzmodelle und Kompetenzentwicklung in der beruflichen Bildung. In Reinhold Nickolaus, Günter Pätzold, Holger Reinisch & Tade Tramm (Hrsg.), *Handbuch Berufs- und Wirtschaftspädagogik* (S. 247–257). Bad Heilbrunn: Klinkhardt.
- Seel, Norbert M. (2000). *Psychologie des Lernens: Lehrbuch für Pädagogen und Psychologen*. München: Reinhardt.
- Seel, Norbert M. (2003). *Psychologie des Lernens: Lehrbuch für Pädagogen und Psychologen*. München: Reinhardt.
- Seidel, Tina & Krapp, Andreas (Hrsg.) (2014). *Pädagogische Psychologie*. Weinheim: Beltz.
- Seidel, Tina; Prenzel, Manfred & Krapp, Andreas (2014). Grundlagen der Pädagogischen Psychologie. In Tina Seidel & Andreas Krapp (Hrsg.), *Pädagogische Psychologie* (S. 21–36). Weinheim: Beltz.
- Seidl, Martina; Brandsteidl, Marion; Huemer, Christian & Kappel, Gerti (2012). *UML@Classroom: Eine Einführung in die objektorientierte Modellierung*. Heidelberg: dpunkt-Verl.
- Sembill, Detlef (1992). *Problemlösefähigkeit, Handlungskompetenz und emotionale Befindlichkeit: Zielgrößen Forschenden Lernens*. Göttingen: Hogrefe.
- Shannon, Claude E. & Weaver, Warren (1963). *The mathematical theory of communication*. Urbana: Univ. of Illinois Press.
- Sonnleitner, Philipp; Keller, Ulrich; Martin, Romain & Brunner, Martin (2013). Students' complex problem-solving abilities: Their structure and relations to reasoning ability and educational success. *Intelligence*, 41(5), S. 289–305.
- Sonntag, Karlheinz & Schaper, Niclas (1997). Aufgaben- und Wissensanalysen bei komplexen Diagnoseaufgaben zur Ermittlung des Lernbedarfs bei komplexen Diagnoseaufgaben. In Karlheinz Sonntag & Niclas Schaper (Hrsg.), *Störungsmanagement und Diagnosekompetenz: Leistungskritisches Denken und Handeln in komplexen technischen Systemen* (S. 95–118). Zürich: Vdf, Hochschulverl. an der ETH Zürich.

- Spada, Hans & Mandl, Heinz (1988). Wissenspsychologie: Einführung. In Heinz Mandl, Hans Spada, Hans Aebli & Friedhart Klix (Hrsg.), *Wissenspsychologie* (S. 1–16). München: Psychologie-Verl.-Union.
- Stachowiak, Herbert (1973). *Allgemeine Modelltheorie*. Wien: Springer.
- Stary, Christian (1996). *Interaktive Systeme: Software-Entwicklung und Software-Ergonomie*. Braunschweig, Wiesbaden: Vieweg.
- Stäudel, Thea (1987). *Problemlösen, Emotionen und Kompetenz: Die Überprüfung eines integrativen Konstrukts*. Regensburg: S. Roderer.
- Steinmetz, Holger (2014). *Lineare Strukturgleichungsmodelle: Eine Einführung mit R*. München: Hampp.
- Stemmann, Jennifer & Lang, Martin (2014). Theoretische Konzeption einer allgemeinen technischen Problemlösefähigkeit und Möglichkeiten ihrer Diagnose. *Journal of Technical Education*, 2(1), S. 80–101.
- Stemmann, Jennifer & Lang, Martin (2016). Wodurch wird die Interaktion mit technischen Alltagsgeräten zum Problem? Identifikation und Analyse schwierigkeitsbestimmender Merkmale im Umgang mit technischen Systemen. In Jürgen Menthe, Dietmar Höttecke, Thomas Zabka, Marcus Hammann & Martin Rothgangel (Hrsg.), *Befähigung zu gesellschaftlicher Teilhabe: Beiträge der fachdidaktischen Forschung* (S. 247–260). Münster: Waxmann.
- Stern, William (1920). *Die Intelligenz der Kinder und Jugendlichen: und die Methoden ihrer Untersuchung*. Leipzig: Barth.
- Strauß, Bernd & Kleinmann, Martin (1995). Die formale Beschreibung computersimulierter Szenarien. In Bernd Strauß & Martin Kleinmann (Hrsg.), *Computersimulierte Szenarien in der Personalarbeit* (S. 105–125). Göttingen: Verl. für Angewandte Psychologie.
- Strohschneider, Stefan (1990). *Wissenserwerb und Handlungsregulation*. Wiesbaden: Dt. Univ.-Verl.
- Strohschneider, Stefan (1991a). Kein System von Systemen! Kommentar zu dem Aufsatz Systemmerkmale als Determinanten des Umgangs mit dynamischen Systemen von Joachim Funke. *Sprache und Kognition*, 10(2), S. 109–113.
- Strohschneider, Stefan (1991b). Problemlösen und Intelligenz: Über die Effekte der Konkretisierung komplexer Probleme. *Diagnostica*, 37(4), S. 353–371.
- Strohschneider, Stefan & Schaub, Harald (1991). Können Manager wirklich so gut managen? Über die Effekte unterschiedlichen heuristischen Wissens beim Umgang mit komplexen Problemen. In Friedhart Klix & Elke van der Meer (Hrsg.), *Experimentalpsychologie heute: 100 Jahre Zeitschrift für Psychologie* (S. 325–340). Heidelberg: Barth.
- Sundre, Donna L. & Kitsantas, Anastasia (2004). An exploration of the psychology of the examinee: Can examinee self-regulation and test-taking motivation predict consequential and non-consequential test performance? *Contemporary Educational Psychology*, 29(1), S. 6–26.
- Süß, Heinz-Martin (1996). *Intelligenz, Wissen und Problemlösen: Kognitive Voraussetzungen für erfolgreiches Handeln bei computersimulierten Problemen*. Göttingen: Hogrefe.
- Süß, Heinz-Martin (1999). Intelligenz und komplexes Problemlösen. *Psychologische Rundschau*, 50(4), S. 220–228.

- Süß, Heinz-Martin & Beauducel, André (2011). Intelligenztests und ihre Bezüge zu Intelligenztheorien. In Lutz F. Hornke, Manfred Amelang & Martin Kersting (Hrsg.), *Leistungs-, Intelligenz- und Verhaltensdiagnostik* (S. 97–234). Göttingen: Hogrefe.
- Süß, Heinz-Martin; Kersting, Martin & Oberauer, Klaus (1991). Intelligenz und Wissen als Prädiktoren für Leistungen bei computersimulierten komplexen Problemen. *Diagnostica*, 37(4), S. 334–352.
- Süß, Heinz-Martin; Kersting, Martin & Oberauer, Klaus (1993). Zur Vorhersage von Steuerungsleistungen an computersimulierten Systemen durch Wissen und Intelligenz. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 14(3), S. 189–203.
- Sweller, John (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, S. 257–285.
- Sweller, John; van Merriënboer, Jeroen J.G. & Paas, Fred G.W.C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), S. 251–296.
- Taatgen, Niels A.; Gray, Wayne & Schunn, Christian (2002). A model of individual differences in skill acquisition in the Kanfer-Ackerman air traffic control task. *Cognitive Systems Research*, 3(1), S. 103–112.
- Tack, Werner H. (1995). Repräsentation menschlichen Wissens. In Dietrich Dörner & Elke van der Meer (Hrsg.), *Das Gedächtnis: Probleme - Trends - Perspektiven* (S. 53–74). Göttingen: Hogrefe.
- Tauschek, Rüdiger (2006). *Problemlösekompetenz in komplexen technischen Systemen: Möglichkeiten ihrer Entwicklung und Förderung im Unterricht der Berufsschule mit Hilfe computergestützter Modellbildung und Simulation. Theoretische und empirische Analyse in der gewerblich-technischen Berufsbildung*. Dissertation. Technische Universität Dresden, Dresden.
- Tully, Claus J. (2003). *Mensch - Maschine - Megabyte: Technik in der Alltagskultur. Eine sozialwissenschaftliche Hinführung*. Opladen: Leske + Budrich.
- Urban, Dieter & Mayerl, Jochen (2011). *Regressionsanalyse: Theorie, Technik und Anwendung*. Wiesbaden: VS Verl. für Sozialwissenschaften.
- Urban, Dieter & Mayerl, Jochen (2014). *Strukturgleichungsmodellierung: Ein Ratgeber für die Praxis*. Wiesbaden: Imprint: Springer VS.
- Vahling, Lothar (2000). Funktionale Transparenz: Voraussetzung für die Lösung des Problems Störungssuche. In Reinhard Bader & Klaus Jenewein (Hrsg.), *Didaktik der Technik zwischen Generalisierung und Spezialisierung: Helmut Sanfleber gewidmet anlässlich der Verabschiedung von der Gerhard-Mercator-Universität Duisburg* (S. 333–350). Frankfurt am Main: GAFB.
- Van den Heuvel-Panhuizen, Marja (2005). The role of contexts in assessment problems in mathematics. *For the Learning of Mathematics*, 25(2), S. 2–23.
- Van Der Linden, Dimitri; Sonnentag, Sabine; Frese, Michael & van Dyck, Cathy (2001). Exploration strategies, performance, and error consequences when learning a complex computer task. *Behaviour and Information Technology*, 20(3), S. 189–198.
- van Vorst, Helena; Dorschu, Alexandra; Fechner, Sabine; Kauertz, Alexander; Krabbe, Heiko & Sumfleth, Elke (2015). Charakterisierung und Strukturierung von Kontexten im naturwissenschaftlichen Unterricht – Vorschlag einer theoretischen Modellierung. *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), S. 29–39.

- Vollmeyer, Regina (2009). Motivationspsychologie des Lernens. In Veronika Brandstätter & Jürgen H. Otto (Hrsg.), *Handbuch der Allgemeinen Psychologie - Motivation und Emotion* (S. 335–346). Göttingen: Hogrefe.
- Vollmeyer, Regina & Burns, Bruce D. (1999). Problemlösen und Hypothesentesten. In Hans Gruber, Wolfgang Mack & Albert Ziegler (Hrsg.), *Wissen und Denken: Beiträge aus Problemlösepsychologie und Wissenspsychologie* (S. 101–118). Wiesbaden: Deutscher Universitätsverlag.
- Vollmeyer, Regina & Rheinberg, F. (1998). Motivationale Einflüsse auf Erwerb und Anwendung von Wissen in einem computersimulierten System. *Zeitschrift für Pädagogische Psychologie*, 12(1), S. 11–23.
- von der Weth, Rüdiger (1994). Wie entstehen individuelle Vorgehensstile beim Konstruieren? Die Rolle der heuristischen Kompetenz. In Gerhard Pahl (Hrsg.), *Psychologische und pädagogische Fragen beim methodischen Konstruieren: Ergebnisse des Ladenburger Diskurses von Mai 1992 bis Oktober 1993* (S. 68–88). Köln: Verl. TÜV Rheinland.
- Vonken, Matthias (2005). *Handlung und Kompetenz: Theoretische Perspektiven für die Erwachsenen- und Berufspädagogik*. Wiesbaden: VS Verl. für Sozialwissenschaften.
- Wagener, Dietrich (2001). *Psychologische Diagnostik mit komplexen Szenarios: Taxonomie, Entwicklung, Evaluation*. Lengerich: Pabst.
- Wagener, Willi & Haupt, Wolfgang (2000). Technik als Fach der gymnasialen Oberstufe. In Reinhard Bader & Klaus Jenewein (Hrsg.), *Didaktik der Technik zwischen Generalisierung und Spezialisierung: Helmut Sanfleber gewidmet anlässlich der Verabschiedung von der Gerhard-Mercator-Universität Duisburg* (S. 53–74). Frankfurt am Main: GAFB.
- Walker, Felix (2013). *Der Einfluss von Handlungsmöglichkeiten auf den Wissenserwerb bei der Durchführung technischer Experimente*. Dissertation. Universität Duisburg-Essen, Essen.
- Walker, Felix; Link, Nico & Nickolaus, Reinhold (2015). Berufsfachliche Kompetenzstrukturen bei Elektronikern für Automatisierungstechnik am Ende der Berufsausbildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 111(2), S. 222–241.
- Wallach, Dieter (1997). *Kognitionswissenschaftliche Analyse komplexer Problemlöseprozesse*. Dissertation, Saarbrücken.
- Wallach, Dieter (1998). *Komplexe Regelungsprozesse: Eine kognitionswissenschaftliche Analyse*. Wiesbaden: Deutscher Universitätsverlag.
- Wallach, Dieter & Lebiere, Christian J. (1998). Modellierung von Wissenserwerbsprozessen bei der Systemregelung. In Uwe Kotkamp & Werner Krause (Hrsg.), *Intelligente Informationsverarbeitung* (S. 93–100). Wiesbaden: Deutscher Universitätsverlag.
- Wang, Jichuan & Wang, Xiaoqian (2012). *Structural equation modeling: Applications using Mplus*. Chichester, West Sussex: Wiley.
- Weiber, Rolf & Mühlhaus, Daniel (2014). *Strukturgleichungsmodellierung: Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS*. Berlin: Springer Gabler.
- Weinert, Franz E. (2001a). Concept of Competence: A Conceptual Clarification. In Dominique S. Rychen & Laura H. Salganik (Hrsg.), *Defining and Selecting Key Competencies* (S. 45–65). Seattle: Hogrefe.
- Weinert, Franz E. (Hrsg.) (2001b). *Leistungsmessungen in Schulen*. Weinheim: Beltz.

- Weinert, Franz E. (2001c). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In Franz E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Weinheim: Beltz.
- Weiß, Reinhold (1999). Erfassung und Bewertung von Kompetenzen - empirische und konzeptionelle Probleme. In Arbeitsgemeinschaft Qualifikations-Entwicklungs-Management Berlin (Hrsg.), *Kompetenzentwicklung '99: Aspekte einer neuen Lernkultur; Argumente, Erfahrungen, Konsequenzen* (S. 433–493). Münster, München: Waxmann.
- Weiß, Rudolf H. (2006). CFT 20-R: Grundintelligenztest Skala 2. Revision. Göttingen: Hogrefe.
- Wentura, Dirk & Pospeschill, Markus (2015). *Multivariate Datenanalyse: Eine kompakte Einführung*. Wiesbaden: Springer.
- Werner, Christina S.; Schermelleh-Engel, Karin; Gerhard, Carla & Gäde, Jana C. (2016). Strukturgleichungsmodelle. In Nicola Döring & Jürgen Bortz (Hrsg.), *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler* (S. 945–973). Heidelberg: Springer.
- West, Stephen G.; Finch, John F. & Curran, Patrick J. (1995). Structural Equation Models with Nonnormal Variables: Problems and Remedies. In Rick H. Hoyle (Hrsg.), *Structural equation modeling: Concepts, issues, and applications* (S. 56–75). Thousand Oaks: Sage Publ.
- Westermann, Rainer (2000). *Wissenschaftstheorie und Experimentalmethodik: Ein Lehrbuch zur psychologischen Methodenlehre*. Göttingen: Hogrefe.
- Wilhelm, Oliver & Kunina, Olga (2009). Pädagogisch-psychologische Diagnostik. In Elke Wild & Jens Möller (Hrsg.), *Pädagogische Psychologie* (S. 307–332). Berlin, Heidelberg: Springer.
- Wilhelm, Oliver & Nickolaus, Reinhold (2013). Was grenzt das Kompetenzkonzept von etablierten Kategorien wie Fähigkeit, Fertigkeit oder Intelligenz ab? *Zeitschrift für Erziehungswissenschaft*, 16(1), S. 23–26.
- Wirth, Joachim (2004). *Selbstregulation von Lernprozessen*. Dissertation, Humboldt-Univ. Münster: Waxmann.
- Wirth, R. J. & Edwards, Michael C. (2007). Item factor analysis: current approaches and future directions. *Psychological Methods*, 12(1), S. 58–79.
- Wise, Steven L. (2009). Strategies for Managing the Problem of Unmotivated Examinees in Low-Stakes Testing Programs. *The Journal of General Education*, 58(3), S. 152–166.
- Wittmann, Werner W.; Süß, Heinz-Martin & Oberauer, Klaus (1996). *Determinanten komplexen Problemlösens*. http://www.psychologie.uni-mannheim.de/psycho2_alt/publi/ps/ber09.pdf. Zugegriffen: 14. November 2012.
- Wolf, Christof & Best, Henning (2010). Lineare Regressionsanalyse. In Christof Wolf & Henning Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 607–638). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Wollersheim, Heinz-Werner (1993). *Kompetenzerziehung: Befähigung zur Bewältigung*. Habilitation, Univ., Bonn. Frankfurt am Main, Berlin: Lang.
- Wüstenberg, Sascha; Greiff, Samuel & Funke, Joachim (2012). Complex problem solving – More than reasoning? *Intelligence*, 40(1), S. 1–14.
- Yu, Ching-Yun (2002). *Evaluating Cutoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes*. Dissertation. University of California, Los Angeles.
- Yuan, Ke-Hai & Bentler, Peter M. (2000). Three Likelihood-Based Methods For Mean and Covariance Structure Analysis With Nonnormal Missing Data. *Sociological Methodology*, 30(1), S. 165–200.

- Zeitler, Sigrid; Köller, Olaf & Tesch, Bernd (2010). Bildungsstandards und ihre Implikationen für Qualitätssicherung und Qualitätsentwicklung. In Axel Gehrman, Uwe Hericks & Manfred Lüders (Hrsg.), *Bildungsstandards und Kompetenzmodelle: Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (S. 23–36). Bad Heilbrunn: Klinkhardt.
- Zühlke, Detlef (2005). Useware-Engineering - brauchen wir eine neue Technikdisziplin? In Katja Karrer & Klaus-Peter Timpe (Hrsg.), *Beiträge zur Mensch-Maschine-Systemtechnik aus Forschung und Praxis: Festschrift für Klaus-Peter Timpe* (S. 29–34). Düsseldorf: Symposium.