

**Geschlechtsunterschiede
in Motivation und Strategiesystematik beim
selbstregulierten Erlernen eines Computerprogramms**

**Dissertationsschrift
zur Erlangung des akademischen Grades
eines Doktors der Philosophie (Dr. phil.)**

vorgelegt dem Fachbereich
Psychologie und Sportwissenschaften der
Johann Wolfgang Goethe-Universität Frankfurt am Main

von

Dipl.-Psych., Dipl.-Pol. Constanze Beierlein

Frankfurt am Main, Juni 2011

1. Gutachterin: Prof. Dr. Regina Vollmeyer
2. Gutachterin: Prof. Dr. Margarete Imhof

Meinen Eltern
Siegfried und Inge

Viele Menschen haben mich auf dem Weg meiner Promotion begleitet, bei denen ich mich herzlich bedanken möchte.

Meine Betreuerin, Prof. Regina Vollmeyer, hat mich sowohl während als auch nach meiner Tätigkeit an der Goethe-Universität kontinuierlich und in konstruktiver Weise fachlich unterstützt. Ihre wertvollen Rückmeldungen haben mir sehr geholfen, mich mit der Thematik der Arbeit kritisch auseinanderzusetzen. Besonders bedanken möchte ich mich für ihre Unterstützung bei der Beantragung eines DAAD-Doktorandenstipendiums und für die Möglichkeit eines Forschungsaufenthalts in den USA.

Bedanken möchte ich mich auch bei Prof. Margarete Imhof. Ihre Ideen und Anregungen zum Thema Geschlechtsunterschiede habe ich sehr geschätzt; sie haben die Arbeit an der Dissertation für mich noch interessanter gemacht. Über die fachliche Seite hinaus fand Margarete stets aufmunternde Worte, die insbesondere in der Endphase der Dissertation sehr motivierend für mich waren.

Ein besonderer Dank gebührt auch Prof. Siegfried Preiser. Seit meiner Diplomarbeit hatte ich die Chance, in mehreren Forschungsprojekten zu Themen der Politischen Psychologie mit ihm zusammenzuarbeiten. Seine Fähigkeit, psychologische Forschung und gesellschaftliche Verantwortung miteinander zu verbinden, hat mich immer sehr fasziniert und in besonderem Maße für meine eigene Forschung inspiriert.

Bei Prof. Eveline Wuttke und Prof. Uwe Hassler vom Fachbereich Wirtschaftswissenschaften sowie bei Dr. Siegbert Reiß vom Institut für Psychologie der Goethe-Universität möchte ich mich sehr herzlich bedanken, dass ich in ihren Veranstaltungen für meine Studien werben durfte.

Die ersten SPSS-Studien, auf denen die Fragestellung der vorliegenden Arbeit aufbaut, habe ich gemeinsam mit meiner Kollegin Dr. Anita Püttmann und meinem Kollegen Dr. Gerhard Bachmann geplant und durchgeführt. Für den sehr anregenden Austausch und die nette Zusammenarbeit möchte ich mich herzlich bedanken.

Bei meiner Kollegin Dipl.-Psych. Christina Werner habe ich immer ein offenes Ohr für meine Fragen zu statistischen Problemen gefunden. Die Gespräche mit ihr (auch über die Statistik hinaus) waren sehr aufschlussreich für mich. Dafür möchte ich mich bedanken.

Die Gespräche mit meinen Kolleginnen Dipl.-Psych. Stefanie Adam-Schwebe, Dr. Charlotte Dignath van Ewijk, Dipl.-Psych. Lisa Paus und Dr. Natalie Fischer haben mich während der gesamten Zeit meiner Promotion sehr bereichert. Die Erkenntnis, dass manche Schwierigkeiten doch nicht so einzigartig und unüberwindbar sind wie geglaubt, war eine große Erleichterung.

Besonders bedanken möchte ich mich auch bei meinen Freundinnen und Kolleginnen Dipl.-Psych. Edith Grasmück und Dipl.-Soz. Britta Kalscheuer. Sowohl die fachlichen als auch die freundschaftlichen Gespräche mit ihnen bedeuten mir viel. Sie waren für mich in der Promotionsphase sowohl lehrreich als auch wohltuend. Edith und Britta hatten für mich jederzeit ein offenes Ohr.

Eine besondere Unterstützung während der gesamten Zeit meiner Promotion habe ich von meiner besten Freundin Kathrin Sahr erfahren. Kathrin hat mich durch sämtliche Höhen und Tiefen meiner Promotionszeit eng begleitet. Durch ihre einfühlsame und humorvolle Art hat sie mir in unvergleichlicher Weise geholfen, auch in schwierigen Zeiten optimistisch zu bleiben. Dafür gebührt ihr mein ganz besonderer Dank!

Der Austausch mit Prof. Eldad Davidov zu Werten nach Shalom H. Schwartz hat mir neben meiner Promotion besonders viel Freude bereitet. Die stets nette, interessante und ermunternde Zusammenarbeit hat mein Interesse an interdisziplinärer Forschung noch mehr angefeuert. Dafür möchte ich mich sehr herzlich bedanken.

I would also like to thank Dr. Nilanjana Dasgupta for giving me the opportunity to work with her during my research visit at the University of Massachusetts at Amherst. It was an extraordinary honor and a great pleasure at the same time to take part in her Social Psychology course and to be able to discuss my doctoral thesis with her. I would further like to thank my friend and colleague Kumar Yogeeswaran with whom I enjoyed discussing a great variety of scientific, political, and cultural topics. Our conversations really inspired my critical thinking.

Als studentische Hilfskräfte haben mich Ina Ulrich, Jana Cerna, Lenka Treskonova-Matuska, Ursula Brehm und Robert Secon insbesondere bei der Durchführung der Studien zu meiner Dissertation unterstützt. Für ihre engagierte und gewissenhafte Hilfe möchte ich mich sehr bedanken. Die Zusammenarbeit hat mir viel Freude gemacht!

Meinen Eltern Inge und Siegfried Beierlein danke ich besonders für die langjährige, liebevolle Unterstützung während meiner gesamten Ausbildungszeit. Ihr Zuspruch und ihr Vertrauen bedeuten mir sehr viel. In meiner Schwester Susanna hatte ich während der Promotionszeit immer eine sehr gute Freundin und eine enge (Seelen-)Verwandte in allen Lebenslagen.

Last but not least möchte ich mich bei meinem Freund Raoul Pra für seine unbeschreibliche Geduld, seinen einzigartigen Humor und seine herzliche Fürsorge bedanken – all dies hat mir während der Promotionszeit besonders viel Kraft gegeben!

Inhaltsverzeichnis

	Seite
Zusammenfassung	15
1. Einleitung	18
1.1 Zielsetzung der Arbeit	20
1.2 Aufbau der Arbeit.....	21
2. Selbstreguliertes Lernen und Problemlösen	22
2.1 Modelle selbstregulierten Lernens	24
2.1.1 Sechs-Komponenten-Modell von Boekaerts	25
2.1.2 Drei-Schichten-Modell von Boekaerts	27
2.1.3 Prozessmodell von Zimmerman.....	29
2.2 Problemlösen	32
2.2.1 Modelle des Problemlösens	33
2.2.2 Klassifikationen von Problemen	36
2.3 Selbstreguliertes Erlernen eines Computerprogramms	37
2.4 Zusammenfassung	39
3. Kognitive Aspekte des Problemlöseprozesses	40
3.1 Problemlöse- und Explorationsstrategien	40
3.1.1 Systematische Exploration	41
3.1.2 Versuch-und-Irrtum	42
3.1.3 Rigide Exploration	43
3.1.4 Einkapselung in der Informationssuche	44
3.2 Methoden zur Erfassung der Strategiesystematik	45
3.3 Empirische Befunde und Zusammenfassung	50
4. Motivationale Aspekte des Problemlöseprozesses	55
4.1 Motivation als Ergebnis der Interaktion von Person und Umwelt	56
4.2 Flow-Erleben als Tätigkeitsanreiz	59
4.2.1 Komponenten des Flow-Erlebens	61
4.2.2 Vorbedingungen des Flow-Erlebens	62
4.2.3 Empirische Befunde und Zusammenfassung	65
4.3 Kognitiv-motivationales Prozessmodell von Vollmeyer und Rheinberg	66
4.3.1 Aspekte der aktuellen Motivation	68
4.3.2 Mediatoren	71
4.3.3 Zusammenfassung	73
5. Geschlechtsunterschiede im Umgang mit dem Computer	75

	Seite
5.1	Theoretische Erklärungsansätze für Geschlechtsunterschiede 76
5.2	Geschlechtsunterschiede in empirischen Studien 79
5.2.1	Computernutzung und Computerbildung 80
5.2.2	Explorationsstrategien am Computer 82
5.2.3	Computerbezogene Motivation und Affekte 84
5.2.4	Computerbezogene Geschlechtsstereotype und Performanz 87
5.2.5	Zusammenfassung 89
6.	Forschungsfragen und Hypothesen 91
6.1	Handlungsnaher Erfassung von Explorationsstrategien (Forschungsfrage 1) ... 92
6.2	Erklärung von Geschlechtsunterschieden mittels des Kognitiv-motivationalen Prozessmodells (Forschungsfrage 2) 95
6.3	Entwicklung von Motivation und Strategiesystematik im Prozessverlauf (Forschungsfrage 3) 98
6.4	Identifizierbarkeit von Personenclustern (Forschungsfrage 4) 101
6.5	Zusammenfassung 101
7.	Entwicklung eines Beobachtungssystems zur Erfassung der Explorationsstrategien 103 (Studie 1)
7.1	Methode 106
7.1.1	Stichprobe 106
7.1.2	Vorgehen 106
7.1.3	Materialien 109
7.1.3.1	Fragebogen zu Statistikkenntnissen 109
7.1.3.2	Fragebogen zur Computerbildung (INCOBI) 111
7.1.3.3	Flow-Kurz-Skala (FKS) 113
7.1.3.4	Problemlöseaufgabe und Performanzmaß 114
7.1.3.5	Beobachtungssystem zur Erfassung der Explorationsstrategien 117
7.1.4	Umgang mit fehlenden Werten 124
7.1.5	Wahl und Diskussion der statistischen Analyseverfahren 124
7.2	Erprobung und erste Validierung des Beobachtungssystems 127
7.2.1	Vorbereitende Analysen zur empirischen Untersuchung 127
7.2.2	Hypothese 1: Zuverlässigkeit des neu entwickelten Kategoriensystems 128
7.2.3	Hypothese 2: Nutzung von Explorationsstrategien 131
7.2.4	Hypothesen 3 und 4: Explorationsstrategien, Flow-Erleben und Leistung 133
7.3	Zusammenfassung und Diskussion 135

	Seite
8. Geschlechtsunterschiede auf der Basis des kognitiv-motivationalen Prozessmodells (Studie 2)	138
8.1 Methode	140
8.1.1 Stichprobe	140
8.1.2 Vorgehen	141
8.1.3 Materialien	142
8.1.3.1 Fragebogen zu Statistikkenntnissen und Computernutzung	142
8.1.3.2 Fragebogen zur aktuellen Motivation (FAM)	143
8.1.3.3 Flow-Kurz-Skala (FKS)	146
8.1.3.4 Problemlöseaufgabe, Performanzmaß und Erfassung der Strategiesystematik	147
8.1.4 Wahl und Diskussion der statistischen Analyseverfahren	150
8.1.5 Umgang mit fehlenden Werten	151
8.2 Ergebnisse	152
8.2.1 Vorbereitende Analysen	152
8.2.2 Hypothese 5: Geschlechtsunterschiede in Computernutzung und Statistikwissen	153
8.2.3 Hypothese 6: Geschlechtsunterschiede in der Eingangsmotivation .	155
8.2.4 Hypothese 7a: Geschlechtsunterschiede in den Explorationsstrategien	157
8.2.5 Hypothese 7b: Geschlechtsunterschiede im Flow-Erleben	159
8.2.6 Hypothese 8: Geschlechtsunterschiede in der Leistung	160
8.2.7 Hypothese 9: Flow-Erleben und Strategiesystematik als Mediatoren	162
8.2.8 Geschlechtsunterschiede im Verlauf der aktuellen Motivation	170
8.3 Zusammenfassung und Diskussion	174
9. Entwicklung von Motivation und Strategiesystematik im Prozessverlauf (Studie 3)	178
9.1 Methode	180
9.1.1 Stichprobe	180
9.1.2 Vorgehen	181
9.1.3 Materialien	182
9.1.3.1 Fragebogen zu Statistikkenntnissen	182
9.1.3.2 Fragebogen zur Computerbildung (INCOBI)	183
9.1.3.3 Fragebogen zur aktuellen Motivation (FAM)	183
9.1.3.4 Flow-Kurz-Skala (FKS)	185
9.1.3.5 Problemlöseaufgabe, Performanzmaß und Erfassung der Strategiesystematik	185

	Seite
9.1.4 Wahl und Diskussion der statistischen Analyseverfahren	187
9.1.5 Umgang mit fehlenden Werten	188
9.2 Ergebnisse	189
9.2.1 Vorbereitende Analysen zur empirischen Untersuchung	189
9.2.1.1 Replizierbarkeit von Geschlechtsunterschieden	191
9.2.1.2 Häufigkeiten der Explorationsstrategien	203
9.2.1.3 Korrelationen von Strategiesystematik, Flow und Leistung	207
9.2.2 Hypothese 10: Wirkungsbeziehungen der Variablen im Zeitverlauf .	207
9.2.3 Hypothese 11: Identifizierbarkeit von Personenclustern	214
9.3 Zusammenfassung und Diskussion	220
10. Gesamtdiskussion	225
10.1 Ergebnisdiskussion und Forschungsausblick	226
10.1.1 Zur handlungsnahen Erfassung von Explorationsstrategien (For- schungsfrage1)	226
10.1.2 Zur Erklärung von Geschlechtsunterschieden mittels des Kognitiv- motivationalen Prozessmodells (Forschungsfrage 2)	232
10.1.3 Zur Entwicklung von Motivation und Strategiesystematik im Pro- zessverlauf (Forschungsfrage 3)	237
10.1.4 Zur Identifizierbarkeit von Personenclustern (Forschungsfrage 4) ...	242
10.2 Studienübergreifende Überlegungen	244
10.3 Praktische Implikationen	247
11. Literaturverzeichnis	252
A Anhang	
A 1. Untersuchungsmaterialien der Studien 1 bis 3	A - 5
A 1.1 Begrüßung der Teilnehmer und Aufgabeninstruktion	A - 5
A 1.2 Einführung in das Statistikprogramm SPSS	A - 7
A 1.3 Fragebogen zur aktuellen Motivation (FAM)	A - 8
A 1.4 Flow-Kurz-Skala (FKS)	A - 9
A 1.5 Kurzversion des Fragebogens zur aktuellen Motivation (FAM-K)	A - 10
A 1.6 Beobachtungssystem zur Erfassung der Explorationsstrategien	A - 11
A 2. Anhang zu Studie 1	A - 15
A 2.1 Untersuchungsmaterialien	A - 15
A 2.1.1 Fragebogen zu Computervorwissen und Statistikkenntnissen	A - 15
A 2.1.2 Aufgabenblatt	A - 21

	Seite
A 2.2 Ergebnisse statistischer Analysen	A - 24
A 2.2.1 Item- und Skalenanalysen	A - 24
A 2.2.2 Voraussetzungsprüfungen	A - 30
A 2.2.3 Ergebnisse ergänzender statistischer Analysen zu den Hypothesen.	A - 31
A 3. Anhang zu Studie 2	A - 32
A 3.1 Untersuchungsmaterialien	A - 32
A 3.1.1 Fragebogen zu Computervorwissen und Statistikkenntnissen	A - 32
A 3.1.2 Aufgabenblatt	A - 35
A 3.2 Ergebnisse statistischer Analysen	A - 37
A 3.2.1 Item- und Skalenanalysen	A - 37
A 3.2.2 Voraussetzungsprüfungen und weitere statistische Analyseergeb- nisse.....	A - 47
A 3.2.3 Ergebnisse ergänzender statistischer Analysen zu den Hypothesen.	A - 48
A 4. Anhang zu Studie 3	A - 49
A 4.1 Untersuchungsmaterialien	A - 49
A 4.1.1 Fragebogen zu Computervorwissen und Statistikkenntnissen	A - 49
A 4.1.2 Aufgabenblatt	A - 54
A 4.2 Ergebnisse statistischer Analysen	A - 57
A 4.2.1 Item- und Skalenanalysen	A - 57
A 4.2.2 Voraussetzungsprüfungen	A - 69
A 4.2.3 Ergebnisse ergänzender statistischer Analysen zu den Hypothesen.	A - 70

Tabellenverzeichnis

		Seite
Tabelle 1	Operationalisierungsvarianten zur Erfassung der Strategiesystematik in ausgewählten Studien zum Erlernen eines Computerprogramms	45
Tabelle 2	Übersicht über spezifische Anforderungen der im Arbeitsblatt enthaltenen Statistikaufgaben	115
Tabelle 3	Übersicht über die Kategoriebezeichnungen und zentrale Indikatoren des Beobachtungssystems zur Erfassung von Explorationsstrategien am Computer (in Anlehnung an van der Linden, Sonnentag, Frese & van Dyck, 2001)	118
Tabelle 4	Übereinstimmungsmatrix der Kategorisierung durch zwei Beobachter in der Trainingsphase	129
Tabelle 5	Übereinstimmungsmatrix der Kategorisierungen durch zwei Beobachter (B1, B2) nach Präzision der Operationalisierung	130
Tabelle 6	Grad der Beobachterübereinstimmung (Cohens κ) zwischen zwei Beobachtern für die fünf Kategorien in der Trainings- bzw. der Anwendungsphase der Erprobung des Kategoriensystems	131
Tabelle 7	Produkt-Moment-Korrelationen der Variablen in der Stichprobe	133
Tabelle 8	Übersicht über die im Arbeitsblatt enthaltenen Statistikaufgaben („Leistung_Arbeitsblatt“)	147
Tabelle 9	Ergebnisse der Multivariaten Varianzanalyse (MANOVA) zur Häufigkeit spezifischer Computernutzung und Statistikkenntnisse getrennt nach Frauen und Männern	155
Tabelle 10	Ergebnisse der MANOVA für den Faktor Geschlecht und die vier Abhängigen Variablen (Erfolgswahrscheinlichkeit, Herausforderung, Interesse, Misserfolgsbefürchtung)	156
Tabelle 11	Beobachtete relative Häufigkeit einzelner Explorationsstrategien bzw. der Kategorie „Unterbrechung“ in der Gesamtstichprobe sowie getrennt für Frauen und Männer	157
Tabelle 12	Mittelwerte, Standardabweichungen separat für Frauen und Männer sowie Effektstärken für die Variable „Flow-Erleben zu verschiedenen Messzeitpunkten (t1 bis t3)“	160
Tabelle 13	Mittelwerte, Standardabweichungen und Effektstärken für verschiedene Leistungsindikatoren für die Gesamtstichprobe bzw. getrennt für Frauen und Männer	161
Tabelle 14	Deskriptive Statistiken und Interkorrelationen der Variablen im Pfadmodell sowie Diskrepanzen als Abweichung der empirischen von der durch das Modell geschätzten Kovarianzmatrix	165
Tabelle 15	Inferenzstatistische und deskriptive Maße zur Bewertung der globalen Modellgüte für das vorgeschlagene Modell	166

		Seite
Tabelle 16	Effektzerlegung (direkte Effekte, indirekte Effekte, Gesamteffekte) für das überidentifizierte Pfadmodell zur Vorhersage der Leistung über die Variablen Geschlecht, Flow und Nutzung Systematischer Explorationsstrategie	168
Tabelle 17	Mittelwerte, Standardabweichungen und Effektstärken für die Netto-Motivation (8 Items) vor und während der Aufgabenbearbeitung für die Gesamtstichprobe bzw. getrennt für Frauen und Männer	171
Tabelle 18	Produkt-Moment-Korrelationen der Netto-Motivation mit der Leistung in der Gesamtstichprobe	171
Tabelle 19	Ergebnisse der Varianzanalyse mit Messwiederholung für die Faktoren Geschlecht und Messzeitpunkt sowie die Abhängige Variable Netto-Motivation zu den Messzeitpunkten t0 und t1	172
Tabelle 20	Übersicht über spezifische Anforderungen der im Arbeitsblatt enthaltenen Statistikaufgaben	186
Tabelle 21	Mittelwerte und Standardabweichung in den zentralen Untersuchungsvariablen auf der Basis normativer, d.h. nicht χ -standardisierter, Messwerte getrennt nach Stichprobe und Geschlecht. Kursiv gedruckte Items bzw. Skalen weisen auf leicht unterschiedliche Operationalisierungen in den beiden Stichproben hin	192
Tabelle 22	Ergebnisse der multivariaten Varianzanalyse (MANOVA) des Faktors Geschlecht für verschiedene Indikatoren der Computererfahrung und für das Statistikvorwissen in beiden Stichproben	195
Tabelle 23	Ergebnisse der multivariaten Varianzanalyse (MANOVA) für die Nutzung der Explorationsstrategien mit F-Tests für den Faktor Geschlecht sowie den Interaktionsterm aus Geschlecht und Stichprobe	199
Tabelle 24	Ergebnisse der multivariaten Varianzanalyse (MANOVA) für die aktuelle Motivation sowie das Flow-Erleben (jeweils zum letzten Messzeitpunkt: Flow_end, Mot_end) sowie die Faktoren Geschlecht und Stichprobe inklusive der partiellen Effektstärke	201
Tabelle 25	Ergebnisse der fünf Varianzanalysen mit Messwiederholung für den Faktor Messzeitpunkt sowie jeweils für die relative Nutzungshäufigkeit der einzelnen Explorationsstrategien sowie der Kategorie Unterbrechung zu den drei Messzeitpunkten	204
Tabelle 26	Mittelwerte und Standardabweichungen, univariate Kurtosis und Schiefe, Interkorrelationen der Modellvariablen in Hypothese 10, Kovarianzen sowie standardisierte Residuen	209
Tabelle 27	Inferenzstatistische und deskriptive Maße zur Bewertung der globalen Modellgüte für das vorgeschlagene Modell	210
Tabelle 28	Effektzerlegung (direkte Effekte, indirekte Effekte) für das überidentifizierte Pfadmodell zur Vorhersage der Leistung über die Variablen Geschlecht, Flow und Nutzung Systematischer Explorationsstrategie	211
Tabelle 29	Ergebnisse der hierarchisch-agglomerativen Clusteranalyse zur Gruppierung auf der Grundlage der Explorationsstrategienutzung	215

	Seite	
Tabelle 30	Ergebnisse der Clusteranalyse für die durchschnittliche Nutzung unterschiedlicher Explorationsstrategien (in % der Häufigkeit aller Strategien)	215
Tabelle 31	Zusammenfassende Interpretation der Clusterlösung über Charakteristiken der Nutzung verschiedener Explorationsstrategien (+++ = häufige Nutzung, ++ durchschnittliche Nutzung, + seltene Nutzung)	216
Tabelle 32	Mittelwerte und Standardabweichungen in den drei Gruppen der Clusteranalyse für die Variablen Flow-Erleben und aktuelle (Netto-) Motivation sowie für die Variable Leistung	218

Abbildungsverzeichnis

	Seite	
Abb. 1	Sechs-Komponenten-Modell von Boekaerts (1997, S. 164)	27
Abb. 2	Drei-Schichten-Modell der Selbstregulation von Boekaerts (1999; Übersetzung in Anlehnung an Straka, 2006, S. 394)	28
Abb. 3	Selbstregulation als Interaktion von personeninternen, umgebungs- und verhaltensbezogenen Prozessen (Zimmerman, 2000b, S. 222; Abbildung entnommen aus Straka, 2006, S. 395)	30
Abb. 4	Phasen und Teilprozesse der Selbstregulation nach Zimmerman (2000; Übersetzung in Anlehnung an Straka, 2006, S. 396)	31
Abb. 5	Phasen der Handlungsregulation beim komplexen Problemlösen (nach Dörner, 1989/2006, S. 67; Dörner & Schaub, 1994, S. 435)	35
Abb. 6	Grundmodell der klassischen Motivationspsychologie (entnommen aus Vollmeyer, 2005, S. 11)	57
Abb. 7	Flow-Kanal-Modell von Csikszentmihalyi (1975/2000) in Anlehnung an Rheinberg (2006, S. 348)	63
Abb. 8	Kognitiv-motivationales Prozessmodell nach Vollmeyer und Rheinberg (1998, 2006; vereinfachte Darstellung mit ausgewählten Aspekten)	66
Abb. 9	Theoretisches Pfadmodell zur Vorhersage der Lernleistung auf der Basis des kognitiv-motivationalen Prozessmodells (Vollmeyer & Rheinberg, 1998, 2000, 2006)	98
Abb. 10	Theoretisches Pfadmodell zum Zusammenhang von Geschlecht und Leistung auf der Basis des kognitiv-motivationalen Prozessmodells mit der Eingangsmotivation sowie den Prozessvariablen Flow-Erleben und Systematischer Explorationsstrategie als Mediatoren	100
Abb. 11	Benutzeroberfläche des Programms SPSS 12.0 mit dem in der Studie verwendeten Beispieldatensatz	107
Abb. 12	Versuchsplan von Studie 1 im Zeitverlauf mit Angabe der erfassten Konstrukte	108
Abb. 13	Durchschnittliche relative Nutzungshäufigkeiten der einzelnen Explorationsstrategien (Systematische Strategie, Versuch-und-Irrtum, Rigide Exploration, Informationssuche) bzw. der Kategorie Unterbrechung in der Stichprobe	132
Abb. 14	Versuchsplan von Studie 2 im Zeitverlauf mit Angabe der erfassten Konstrukte	141
Abb. 15	Modifiziertes theoretisches Pfadmodell zur Vorhersage der Leistung mit den Variablen Geschlecht, Systematische Explorationsstrategie (alle Messzeitpunkte) und Flow-Erleben (Flow_End)	163
Abb. 16	Empirisches, rekursives Modell mit standardisierten Pfadkoeffizienten für den Zusammenhang der Variablen Geschlecht, systematischer Explorationsstrategie und Flow-Erleben zur Vorhersage der Leistung	167
Abb. 17	Höhe der Netto-Motivation zu Messzeitpunkt t0 (vor Beginn) und t1 (10	174

		Seite
	Min nach Bearbeitungsbeginn) getrennt für Frauen und Männer	
Abb. 18	Versuchsplan von Studie 3 im Zeitverlauf mit Angabe der erfassten Konstrukte	181
Abb. 19	Mittelwerte und Standardabweichungen für den Indikator „Leistung_Arbeitsblatt“ getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3)	195
Abb. 20	Mittelwerte und Standardabweichungen für verschiedene Indikatoren der Computererfahrung und für das Statistikvorwissen getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3)	197
Abb. 21	Mittelwerte und Standardabweichungen für die vier Subskalen der Eingangsmotivation getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3)	198
Abb. 22	Mittlere relative Häufigkeiten und Standardabweichungen für die vier Explorationsstrategien getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3)	200
Abb. 23	Mittelwerte und Standardabweichungen in der Netto-Motivation sowie dem Flow-Erleben (jeweils zum letzten Messzeitpunkt) getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3)	202
Abb. 24	Mittlere relative Häufigkeit (%) der vier Explorationsstrategien sowie der Kategorie Unterbrechung zu den drei Messzeitpunkten in der Gesamtstichprobe	205
Abb. 25	Modifiziertes theoretisches Pfadmodell zur Vorhersage der Leistung mit den Variablen Geschlecht sowie den Prozessvariablen Systematische Explorationsstrategie und Flow-Erleben über mehrere Messzeitpunkte	208
Abb. 26	Empirisches, rekursives Modell mit standardisierten Pfadkoeffizienten mit den Variablen Geschlecht, Systematischer Explorationsstrategie und Flow-Erleben zur Vorhersage der Leistung	211

Zusammenfassung

Seit den 1980er Jahren fällt dem Computer sowohl in Beruf und Bildung als auch in der Freizeit eine wachsende Bedeutung zu (z. B. Pace, 2004; Rheinberg & Trapp, 2006). Die gesellschaftlich-technische Entwicklung macht es erforderlich, dass sich Personen Computerkenntnisse selbstständig und ohne explizite Anleitung aneignen (Richter, Naumann & Horz, 2010). Frühere Studien liefern Hinweise darauf, dass sich Frauen und Männer einer neuen Computersoftware auf verschiedene Weise nähern und dass sie den Umgang mit dem Computer unterschiedlich erleben (z. B. Baloğlu & Çevik, 2008; Schaumburg, 2004). Aus diesen Unterschieden im Erleben und Verhalten können Lern- und Leistungsdifferenzen zuungunsten insbesondere der weiblichen Computernutzer resultieren (Schaumburg, 2004; Sieverding & Koch, 2009). Trotz der vielfältigen Erkenntnisse, die in den letzten Jahren zu Geschlechtsunterschieden am Computer zusammengetragen wurden, zeigen sich weiterhin bedeutsame Forschungslücken. Insbesondere eine prozessnahe Untersuchung von Lern- und Leistungsunterschieden zwischen den Geschlechtern lag bisher nicht vor (vgl. van der Linden, Sonnentag, Frese & van Dyck, 2001; Vollmeyer & Rheinberg, 2006).

Die „Digital Divide“ (Cooper, 2006) zwischen den Geschlechtern kann auf unterschiedliche Ursachen zurückgeführt werden. Modelle des selbstregulierten Lernens betonen die Rolle motivationaler, kognitiver und metakognitiver Faktoren im Lernprozess (z. B. Boekaerts, 1997, 1999; Zimmerman, 2000a, 2000b). Zur Erklärung potenzieller Lern- und Leistungsdifferenzen wurde in der vorliegenden Arbeit das kognitiv-motivationale Prozessmodell von Vollmeyer und Rheinberg (1998, 2000, 2006) zu Grunde gelegt. Dieses Modell erlaubt es, Leistungsunterschiede über die aktuelle Motivation, das Flow-Erleben und die Strategiesystematik vor bzw. während des Lernprozesses zu erklären. Vor diesem Hintergrund kulminierte das Forschungsthema der vorliegenden Arbeit in der Frage, ob Geschlechtsdifferenzen in der Performanz am Computer prozessnah auf Unterschiede in Motivation und Explorationsstrategien zurückgeführt werden können.

Es wurden vier Forschungsfragen untersucht: *Forschungsfrage 1* richtete sich auf die reliable und valide Erfassung der Explorationsstrategien beim selbstregulierten Erlernen eines Computerprogramms. *Forschungsfrage 2* betraf die Frage, inwiefern Geschlechtsunterschiede in der Leistung auf Variablen des kognitiv-motivationalen Prozessmodells (z. B. Strategiesystematik, Flow-Erleben) zurückgeführt werden können. Der Fokus von *Forschungsfrage 3* lag auf der wechselseitigen Beziehung der motivationalen und kognitiven Variablen während des Lernprozesses sowie auf deren Beziehung zur Leistung. Im Rahmen von *Forschungsfrage 4* wurde erforscht, ob sich Per-

sonen danach klassifizieren lassen, welche Explorationsstrategien sie nutzen und ob sich Männer und Frauen im Muster der Strategienutzung unterscheiden.

Insgesamt wurden drei empirische Studien durchgeführt, an denen Studierende der Wirtschaftswissenschaften sowie der Psychologie teilnahmen (Studie 1: $N = 18$; Studie 2: $N = 33$; Studie 3: $N = 92$). Die Datenerhebungen erfolgten in den Jahren 2005, 2007 und 2008. Die Studienteilnehmer hatten die Aufgabe, das ihnen unbekanntes Statistikprogramm SPSS selbstreguliert zu explorieren (vgl. Green & Gilhooly, 1990). Den Lernenden wurden bestimmte Aufgaben (Probleme) vorgegeben, die mit Hilfe des Statistikprogramms zu lösen waren. Das Vorgehen entsprach weitgehend einem „active/exploratory training“ (Keith, Richter & Naumann, 2010; Mayer, 2004). Es wurde ein längsschnittliches Design gewählt. Während des Prozesses wurden mittels Fragebogen zu mehreren Messzeitpunkten die aktuelle Motivation und das Flow-Erleben erfasst. Die Strategiesystematik wurde über die Verhaltensspuren der Teilnehmer rekonstruiert.

In *Forschungsfrage 1* stand die Erfassung der Explorationsstrategien, die Frauen und Männern beim Erkunden von SPSS nutzen, im Vordergrund. In der Arbeit ist es gelungen, ein reliables und valides Messinstrument zur Erfassung der Explorationsstrategien zu entwickeln. In Übereinstimmung mit früheren Studien (Kang & Yoon, 2008; van der Linden et al., 2001) erlaubte das neu entwickelte Kategoriensystem drei Explorationsstrategien trennscharf zu erfassen: die systematische, die rigide Explorationsstrategie sowie die Explorationsstrategie Informationssuche. Für die Explorationsstrategie Versuch-und-Irrtum wurde auch nach Modifikation des Messinstruments keine zufriedenstellende Beobachterübereinstimmung erreicht. Eine erste Validierung des Messinstruments zeigte, dass die systematische Explorationsstrategie positiv mit der Leistung und dem Flow-Erleben korrelierte. Die rigide Exploration und die Informationssuche standen dagegen in negativem Zusammenhang mit diesen Außenkriterien. Die Explorationsstrategie Versuch-und-Irrtum wies keine bzw. leicht negative Beziehungen zu den genannten Außenkriterien auf.

Forschungsfrage 2 richtete sich auf die Erklärung von Geschlechtsunterschieden mittels des kognitiv-motivationalen Prozessmodells. Hinsichtlich der Leistung und der Prozessvariablen (Systematische Exploration, aktuelle Motivation) ergaben sich inkonsistente Befunde zwischen den beiden Studien: In Studie 2 zeigten Männer erwartungsgemäß höhere Leistungen, eine höhere aktuelle Motivation und eine günstigere Strategiesystematik. Demgegenüber wurden diese Differenzen in Studie 3 nicht statistisch signifikant. Statt der starken Effekte in Studie 2 zeigten sich lediglich schwache Effekte (Cohen, 1992). Entgegen der Erwartungen ergaben sich in Studie 2 und 3 keine Geschlechtsunterschiede in der Eingangsmotivation. Somit trug die Eingangsmotivation nicht zur Erklärung von Leistungsunterschieden zwischen Männern und Frauen bei. Unterschiede im Flow-Erleben wurden ebenfalls nicht signifikant. In der Computerbildung, welche zur

Kontrolle mit erfasst wurde, zeigten Frauen jedoch in beiden Studien eine geringere Computererfahrung.

Forschungsfrage 3 richtete sich auf die Beziehung der Prozessvariablen untereinander, auf deren Entwicklung über die Zeit sowie auf deren Erklärungsbeitrag zur Leistung. Die Explorationsstrategien unterlagen im Explorationsprozess unterschiedlichen Entwicklungstendenzen. Der Anteil der systematischen Explorationsstrategie verringerte sich über die Zeit. Der Anteil der unsystematischen Explorationsstrategien blieb weitgehend stabil, während die Häufigkeit der Unterbrechungen zunahm, in denen keine Bildschirmaktivität der Teilnehmer beobachtet werden konnte. Bei leistungsschwachen Teilnehmern konnten jedoch im Prozessverlauf zunehmend unsystematische Explorationsstrategien beobachtet werden. In einer Pfadanalyse zeigte sich, dass die systematische Explorationsstrategie zu den meisten Messzeitpunkten einen positiven Effekt auf das Flow-Erleben und die Leistung hatte. Dabei leistete das Flow-Erleben wie erwartet einen inkrementellen Beitrag zur Erklärung der Leistung.

Im Rahmen von *Forschungsfrage 4* wurde exploratorisch erforscht, ob sich Personen danach klassifizieren lassen, welche Explorationsstrategien sie nutzen. In einer Clusteranalyse konnten drei Personengruppen voneinander abgegrenzt werden. Die erste Gruppe („Systematische Explorierer“) war gekennzeichnet durch eine häufige Nutzung der systematischen Explorationsstrategie. Die zweite Gruppe („Unsystematische Explorierer“) nutzte überwiegend die rigide Explorationsstrategie und die Explorationsstrategie Versuch-und-Irrtum. Bei der dritten Gruppe stand die Informationssuche im Vordergrund. Wurde neben der systematischen und der rigiden Exploration vorwiegend auf die Informationssuche zurückgegriffen, wirkte sich diese Kombination ungünstig auf die Leistung aus, nicht jedoch auf das Flow-Erleben. Die günstigste Motivation und die höchste Leistung erzielten erwartungsgemäß die „systematischen Explorierer“. Im Unterschied dazu waren die „unsystematischen Explorierer“ statistisch signifikant geringer motiviert und waren bei der Aufgabenlösung weniger erfolgreich. Die als unsystematisch eingestuften Explorationsstrategien (Informationssuche, rigide Exploration, Versuch-und-Irrtum) hatten damit unterschiedliche Effekte auf die Motivation und auf die Leistung.

Die Befunde der Arbeit wurden kritisch diskutiert und theoretisch eingeordnet. Darüber hinaus wurden inhaltliche Grenzen und methodische Schwierigkeiten der Arbeit aufgezeigt. In einem Forschungsausblick wurden Empfehlungen für zukünftige Forschungsstudien gegeben. Dabei wurde insbesondere auf Interventionsmöglichkeiten eingegangen, die das selbstregulierte Erlernen eines neuen Computerprogramms fördern können.

*Die Frau [hat] das Recht ...,
an den Kulturerrungenschaften unserer Zeit
vollen Anteil zu nehmen,
sie für die Erleichterung und Verbesserung
ihrer Lage auszunutzen,
und alle ihre geistigen und körperlichen Fähigkeiten
zu entwickeln und zu ihrem Besten anzuwenden
so gut wie der Mann.*

August Bebel (1929/1994)

1. Einleitung

Seit den 1980er Jahren fällt dem Computer sowohl in Beruf und Bildung als auch in der Freizeit eine wachsende Bedeutung zu (Ghani & Dhespande, 1994; Pace, 2004; Rheinberg & Tramp, 2006). Fortschritte in der Informationstechnologie und Automatisierungsprozesse haben seitdem zu einer enormen Verbreitung von Computern und Computeranwendungen geführt (Richter, Naumann & Horz, 2010; van der Linden, Sonnentag, Frese & van Dyck, 2001). In vielen Situationen sind Menschen heute auf die Nutzung von Computertechnologien angewiesen (z. B. beim Kauf einer Fahrkarte, bei finanziellen Transaktionen, bei der Literaturrecherche in Bibliotheken).

Die rasante Entwicklung neuer Computersoftware und die steigende Komplexität dieser Anwendungen setzen eine intensive Auseinandersetzung der Computernutzer mit dem Computer voraus. Dies gilt insbesondere für das Meistern von Lern- und Leistungsaufgaben im Studium, wie Naumann, Richter und Groeben (2001) betonen:

Kompetenzen im Umgang mit dem Computer sind inzwischen nicht nur für eine Bewältigung von naturwissenschaftlich-technischen Studiengängen erforderlich. Auch in den Sozial- und Geisteswissenschaften ist der Computer zu einem Instrument avanciert, ohne dessen Beherrschung die einschlägigen Ausbildungsgänge nur noch unter erschwerten Bedingungen erfolgreich absolviert werden können. (S. 219)

Dabei sind viele Studierende darauf angewiesen, sich neue Computeranwendungen selbst anzueignen. Lernende können sich dabei entweder an detaillierten Vorgaben orientieren und genauen Instruktionen folgen. Alternativ können sie das neue Programm erlernen, in dem sie es ohne oder mit geringen Vorgaben selbstgesteuert erkunden. Mayer (2004) spricht in diesem Zusammenhang von „exploratorischem“ Lernen. Das selbstinitiierte und selbstregulierte Erlernen eines Computerprogramms dient einerseits dazu, eigene Ressourcen effizient zu nutzen (z. B. in Form einer bedarfsgerechten Anwendung spezifischer Software) und somit Kosten zu reduzieren (z. B. Zeitersparnis, Gebühren für Fortbildungen entfallen). Anderer-

seits stellt das eigenständige Erlernen hohe Anforderungen an den Lernenden¹ – insbesondere hinsichtlich dessen Fähigkeiten, selbstreguliert zu lernen. Hasselhorn und Labuhn (2008) heben die Notwendigkeit selbstregulierten Lernens in der heutigen Gesellschaft hervor: „Um Neues zu lernen und in der sich in hohem Tempo weiterentwickelnden modernen Gesellschaft zurechtzukommen, wird es immer wichtiger, Verantwortung für den eigenen Lernfortschritt zu übernehmen, Defizite und Lücken zu erkennen und die Anstrengung aufzubringen, diese zu füllen“ (S. 32). Dies gilt in erheblichem Maße für den Bereich der Computeranwendungen. Carroll und Mack (1983) zufolge ist das Erlernen eines Computerprogramms ein aktiver Prozess, der vom Lernenden selbst gesteuert wird: „Learners do not ‚absorb‘ knowledge: they create, explore, and integrate knowledge“ (S. 43).

Seit mehreren Jahren ist das wissenschaftliche Interesse daran gestiegen, Faktoren zu identifizieren, die das selbstgesteuerte Erlernen neuer Computertechnologien beeinflussen (z. B. Carroll, Mack, Lewis, Grischkowsky, & Robertson, 1985; Keith, Richter & Naumann, 2010; Lepper & Malone, 1987; Wood, Kakebeeke, Debowski & Frese, 2000). Hinweise auf Bedingungen, welche bestimmte Personengruppen bezüglich der Computernutzung benachteiligen, sind von hoher pädagogischer Relevanz (Dickhäuser & Stiensmeier-Pelster, 2002a, 2002b). Wissenschaftler sprechen in diesem Zusammenhang von einer *Digital Divide* (Cooper, 2006; Imhof, Vollmeyer & Beierlein, 2007; Schaumburg, 2004): Dieses Schlagwort beruht auf der Annahme, dass soziodemografische Gruppen unterschiedliche Zugänge und Nutzungsbedingungen von Computern und Internet haben. Neben sozioökonomischen Faktoren (z. B. Bildungshintergrund, Einkommen) und dem Lebensalter trägt insbesondere das biologische Geschlecht einer Person zur Erklärung der Computerbildung und –nutzung bei (*Gender Gap*; Imhof et al., 2007; vgl. Cooper & Weaver, 2003; Jackson, Ervin, Gardner & Schmitt, 2001).

Wissenschaftliche Studien zeigen, dass sich Frauen im Vergleich zu Männern seltener für Karrierewege in den sogenannten MINT- bzw. STEM-Fächern² entscheiden (Kessels, 2008; Messersmith, Garrett, Davis-Kean, Malanchuk & Eccles, 2008; vgl. Stout, Dasgupta, Hunsinger & McManus, 2011). Dass sich deutlich mehr Männer als Frauen für Berufe im Bereich der Computertechnologien interessieren, wird unter anderem darauf zurückgeführt, dass Computertätigkeiten in der Gesellschaft maskulin konnotiert sind. In diesem Zusammenhang wird auch der Mangel an weiblichen Rollenvorbildern als Ursache genannt (Young, 2000; Zeldin & Pajares, 2000). Hirsch (2002) weist darauf hin, dass die gesellschaftliche Arbeitsteilung

¹ In der vorliegenden Arbeit wird stets die männliche Form verwendet, um das Lesen zu erleichtern. Die weibliche Form ist jedoch immer mit gemeint.

² MINT ist ein Akronym, das im deutschen Sprachraum verwendet wird und die Gruppe folgender Fachdisziplinen bezeichnet: Mathematik, Informatik, Naturwissenschaft, Technik. STEM kann als angelsächsisches Pendant betrachtet werden, hinter dessen Initialen folgende Fachdisziplinen stehen: *Science, Technology, Engineering, Mathematics*.

lange Zeit dazu führte, dass Frauen aus technologischen Bereichen ausgeschlossen waren. Im Vergleich zu Männern scheinen Frauen demzufolge nicht im gleichen Maße von der technologischen Entwicklung zu profitieren (Cooper, 2006; s. auch Vollmeyer & Imhof, 2007).

Potenzielle Geschlechtsdifferenzen im Umgang mit dem Computer bilden den Forschungsschwerpunkt der vorliegenden Arbeit. Die *Digital Divide* bezieht sich heute weniger auf ungleiche Möglichkeiten des Zugriffs auf den Computer und unterschiedliche Ausstattungen. Vielmehr stehen die Art und Weise der Computernutzung sowie Möglichkeiten zum Erwerb von Computerkenntnissen im Mittelpunkt der Forschung (Mercier, Barron & O'Connor, 2006). Die Arbeit befasst sich mit der Frage, ob und wenn ja, worin sich Frauen und Männer beim Erlernen eines Computerprogramms unterscheiden. Theorien selbstregulierten Lernens spezifizieren kognitive, motivationale und metakognitive Voraussetzungen oder Komponenten des Lernprozesses. Diese können zu Lern- und Leistungsdifferenzen beitragen. Im Kontext des selbstregulierten Lernens wird in dieser Arbeit das kognitiv-motivationale Prozessmodell von Vollmeyer und Rheinberg (1998, 1999, 2000, 2006) herangezogen, um mögliche Lern- und Leistungsdifferenzen zwischen den Geschlechtern zu erklären.

1.1 Zielsetzung der Arbeit

Geschlechtsunterschiede können sich im Rahmen des kognitiv-motivationalen Prozessmodells in motivationalen (z.B. Flow-Erleben) sowie kognitiven Variablen (z.B. in der Strategiesystematik) manifestieren. Der Strategiesystematik beim Explorieren eines unbekanntes Computerprogramms wird eine besondere Bedeutung beigemessen. Um potenzielle Geschlechtsunterschiede identifizierbar und messbar machen zu können, müssen Überlegungen zur Operationalisierung des Konstrukts angestellt werden. Vor diesem Hintergrund werden in dieser Arbeit mehrere Forschungsziele verfolgt. Das *erste Ziel* stellt die Entwicklung eines Messinstruments dar, mit dem Geschlechtsunterschiede in der Computernutzung abgebildet werden können. Das *zweite Ziel* besteht darin, den Erklärungsgehalt des kognitiv-motivationalen Prozessmodells für Leistungsunterschiede beim Erlernen eines unbekanntes Computerprogramms im Studium zu überprüfen. Es stellt sich die Frage, ob Geschlechtsunterschiede in der Leistung auf Differenzen in Motivation und Strategiesystematik vor bzw. während des Lernprozesses zurückführbar sind. Im Hinblick auf Förderungsmöglichkeiten hat die Arbeit des Weiteren zum Ziel, Personengruppen zu identifizieren, welche ein ungünstiges Muster im Umgang mit dem Computer zeigen und dadurch Nachteile im Lernprozess erleiden. Die Erkenntnisse der Arbeit sollen dazu dienen, mögliche Defizite aufzuzeigen und Ansatzpunkte für Förderungsmöglichkeiten zu liefern.

1.2 Aufbau der Arbeit

Im Anschluss an diese Einleitung führt *Kapitel 2* der Arbeit in die Theorie selbstregulierten Lernens ein. Hier werden drei häufig in der Literatur zitierte und diskutierte Modelle vorgestellt. Darüber hinaus wird die Relevanz selbstregulierter Lern- und Problemlöseprozesse für die spezifische Aufgabe des Erlernens eines Computerprogramms aufgezeigt. Eine wichtige Teilkompetenz selbstregulierten Lernens stellt die Anwendung zielführender, kognitiver Lern- bzw. Problemlösestrategien dar. *Kapitel 3* fasst die theoretischen Annahmen zu dieser kognitiven Komponente selbstregulierten Lernens sowie empirische Befunde zu dieser Teilkompetenz zusammen. In *Kapitel 4* werden motivationale Komponenten des Lern- bzw. Problemlöseprozesses vorgestellt und anhand des kognitiv-motivationalen Prozessmodells theoretisch eingeordnet. *Kapitel 5* gibt einen Überblick über Erklärungsansätze und empirische Befunde zu Geschlechtsunterschieden beim Nutzen und Erlernen von Computerprogrammen. Das darauffolgende *Kapitel 6* dient der Ableitung der Forschungsfragen und Hypothesen, welche in der vorliegenden Arbeit untersucht werden. Der empirische Teil der Arbeit in den *Kapiteln 7, 8* und *9* umfasst drei Studien, in deren Rahmen die oben skizzierten Forschungsanliegen der Arbeit untersucht werden. In *Kapitel 10* werden die Befunde der vorliegenden Arbeit zusammengefasst und einer kritischen Diskussion unterzogen. Darüber hinaus werden Grenzen der vorliegenden Untersuchungen aufgezeigt und ein Forschungsausblick gegeben.

2. Selbstreguliertes Lernen und Problemlösen

Das Explorieren und Erlernen eines neuen Computerprogramms kann auf unterschiedliche Weise vollzogen werden (vgl. Carroll & Mack, 1983; Debowksi, Wood & Bandura, 2001; Trudel & Payne, 1995; Keith et al., 2010). In der Literatur werden insbesondere zwei Ansätze diskutiert. Diese lassen sich danach unterscheiden, wie stark der Lernende zu eigenständiger, aktiver Exploration des neuen Programms angeregt und ermuntert wird (vgl. Mayer, 2004). Beim „guided learning“ werden die Lernenden durch umfangreiche Informationen angeleitet, wie eine Lernaufgabe gelöst oder bewältigt werden kann (Keith et al., 2010). Demgegenüber werden die Lernenden beim „exploratory learning“ als aktive Gestalter betrachtet, sie steuern ihren eigenen Lernprozess selbst und erhalten dabei vergleichsweise wenige Vorgaben (vgl. Bruner, 1961; Kirschner, Sweller & Clark, 2006).

Seit Ende der 1980er Jahre wird die Fähigkeit, selbständig Wissen erwerben und anwenden zu können, in zunehmendem Maße als wichtiges Erziehungsziel betrachtet (Boekaerts, 1996, 1997; Zimmerman, 2002; Zimmerman & Martinez-Pons, 1988; Schunk & Zimmerman, 1994). Zahlreiche Forschungsarbeiten widmen sich seitdem der Frage, wie Lernende ihren Lernprozess selbst steuern bzw. regulieren (z.B. Boekaerts & Corno, 2005; Pintrich & De Groot, 1990; Rheinberg, Vollmeyer & Rollett, 2000; Winne & Hadwin, 2008). Die Bedeutung des selbstregulierten Lernens sieht Zimmerman (2008, S. 166) insbesondere darin, dass es dem Lernenden erlaubt, Herr über den eigenen Lernprozess zu werden. Es wird angenommen, dass sich insbesondere das exploratorische Lernen günstig auf die Entwicklung von Selbstregulationskompetenzen auswirkt (vgl. Keith et al., 2010; Bell & Kozlowski, 2008). Um den Erfordernissen eines lebenslangen Lernens nachkommen zu können, erscheinen diese Kompetenzen als unabdingbar (Baumert et al., 2000; Zimmerman, 2002).

Selbstregulationsprozesse und -modelle werden in unterschiedlichen Teildisziplinen der Psychologie beforscht (Boekaerts, Maes & Karoly, 2005). Die Selbstregulationsmodelle werden genutzt, um zielgerichtetes Verhalten beschreiben und erklären zu können (Carver, 2004). Dies gilt auch für den Bereich des Computerlernens. Karoly (1993) definiert Selbstregulation folgendermaßen:

Self-regulation refers to those processes, internal and/or transactional, that enable an individual to guide his/her goal-directed activities over time and across changing circumstances. Regulation implies modulation of thought, affect, behavior, or attention via deliberate or automated use of specific mechanisms and supportive metaskills. (S. 25)

Selbstregulation wird häufig als zyklischer Prozess betrachtet, der in mehrere Phasen unterteilt werden kann (z.B. Zimmerman, 2000a, 2000b; Winne & Hadwin, 2008). Ein Basismodell der Selbstregulation stellt das zyklische Modell von Carver und Scheier (1981, 2000; siehe auch Carver, 2004) dar. In diesem Modell wird die Selbstregulation als ein kybernetischer Regelkreis beschrieben, welcher aus verschiedenen Elementen besteht, die über Rückkopplungsschleifen miteinander verbunden sind (vgl. Miller, Galanter & Pribram, 1960). Die Elemente des Regelkreises umfassen eine Eingangsfunktion (bezogen auf Wahrnehmungsinhalte), einen Vergleichswert (Ziele, Standard), eine Vergleichseinheit und eine Ausgabefunktion (Verhalten, Emotionen, Kognitionen...). Die Vergleichseinheit vergleicht die Eingangsfunktion mit dem Vergleichswert. Stellt eine Person eine Diskrepanz zwischen der Eingangsfunktion und dem Vergleichswert fest, verändert sie die Outputfunktion (z.B. ihr Verhalten). Carver (2004) sowie Carver und Scheier (1981, 2000) unterscheiden bei den Vergleichswerten Annäherungs- und Vermeidungsziele. Bei ersteren wird angestrebt, die Diskrepanz zwischen Eingangsfunktion und Vergleichswert zu reduzieren („negative Rückkopplungsschleife“). Bei Vermeidungszielen dagegen strebt die Person eine Vergrößerung der Diskrepanz an. Dazu wird die Outputfunktion entsprechend verändert („positive Rückkopplungsschleife“). Störeinflüsse können die Wirkung der Outputfunktion auf die Umwelt verstärken oder vermindern. Die Selbstregulation des Verhaltens besteht in dem Modell letztlich darin, den angestrebten Standard zu erreichen bzw. zu bewahren (Carver & Scheier, 2000, S. 43).

In der Pädagogischen Psychologie dienen Selbstregulationsmodelle dazu, Lernprozesse zu beschreiben (Spörer & Brunstein, 2006, S. 147). Das selbstregulierte Lernen wird definiert als „self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals“ (Zimmerman, 2000a, S. 14). Es ist dabei sowohl als ein willentlicher als auch ein automatisierter Prozess zu verstehen (Boekaerts, 1997, S. 163). Winne (1995) betrachtet das selbstregulierte Lernen als einen konstruktivistischen Prozess, dessen wesentliches Merkmal die Selbststeuerung eigener Lernaktivitäten durch den Lernenden ist. Lernen umfasst dabei kognitive, motivationale und metakognitive Komponenten (Spörer & Brunstein, 2006).

Die Fähigkeit, selbst reguliert zu lernen, wird von einigen Autoren als Handlungskompetenz aufgefasst (Baumert et al., 2000; Winne & Hadwin, 2008). Diese kann nach Wirth und Leutner (2008, S. 103) beschrieben werden als “a learner’s competence to autonomously plan, execute, and evaluate learning processes, which involves continuous decisions on cognitive, motivational, and behavioral aspects of the cyclic process of learning“. Die Selbstregulation ermöglicht es den Lernenden, Fähigkeiten, Wissen und Einstellungen zu erwerben, die sie auf

andere Lernkontexte übertragen und dort zielführend nutzen können (Boekaerts, 1999, S. 446).

2.1 Modelle selbstregulierten Lernens

In der Literatur werden verschiedene Modelle selbstregulierten Lernens beschrieben (z.B. Boekaerts, 1997, 1999; Schmitz, 2001; Winne & Hadwin, 2008; Zimmerman, 2000a, 2000b). Diese entstammen unterschiedlichen theoretischen Richtungen und heben unterschiedliche Aspekte hervor (z.B. motivationale, kognitive; Baumert et al., 2000). Boekaerts und Corno (2005) betonen jedoch, dass die Modelle in Bezug auf ihre grundlegenden Annahmen mehr oder weniger große Übereinstimmungen aufweisen. Winne und Perry (2000) haben eine Systematik entwickelt, nach der Modelle selbstregulierten Lernens klassifiziert werden können. Die Autoren differenzieren dabei zwischen Komponenten- und Prozessmodellen (vgl. auch Wirth & Leutner, 2008).

Komponentenmodelle. In Komponentenmodellen wird das selbstregulierte Lernen als *Fähigkeit* konzeptualisiert: „An aptitude describes a relatively enduring attribute of a person that predicts future behavior“ (Winne & Perry, 2000, S. 534). Diese Fähigkeit bezieht sich hierbei auf eine Kombination mehrerer Kompetenzen eines Lernenden. Hierzu zählen motivationale, kognitive sowie metakognitive Regulationskompetenzen (Boekaerts, 1999; Spörer & Brunstein, 2006). Diese Kompetenzen werden als vergleichsweise stabil, d.h. unabhängig von Situations- und Aufgabenmerkmalen betrachtet (Spörer & Brunstein, 2006). Es wird angenommen, dass sie sich günstig auf den Lernprozess sowie auf das Lernergebnis auswirken (Wirth & Leutner, 2008; Zimmerman, 2008).

Prozessmodelle. In Prozessmodellen wird das selbstregulierte Lernen als Serie von *Ereignissen* konzeptualisiert (Perels, Gürtler & Schmitz, 2005; Winne & Perry, 2000; Wirth & Leutner, 2008). Als ein Ereignis wird eine Veränderung eines Lernzustands in einen anderen angesehen. Ein Beispiel hierfür ist, wenn ein Lernender die richtige Lösung zu einem Problem entdeckt und diese Entdeckung selbst registriert (Winne & Perry, 2000, S. 535). Jedes einzelne Ereignis im Lernprozess kann wiederum in verschiedene Phasen oder Stufen unterteilt werden (Zimmerman, 2000a, 2000b, 2002, 2008). Auf der Basis von Prozessmodellen werden Lernen den Selbstregulationskompetenzen zugeschrieben, „sofern sie in der Lage sind, Lernstrategien unterschiedlicher Art (z.B. Aufgabenstrategien, Motivierungsstrategien, Reflexionsprozesse) bei der Bearbeitung einer Aufgabe selbstständig zu nutzen und so miteinander zu koordinieren, dass ein möglichst optimales Lernergebnis erreicht wird“ (Spörer & Brunstein, 2006, S. 147). Das selbstregulierte Lernen wird dabei als zyklischer Prozess betrachtet (Wirth & Leutner, 2008).

Um diesen unterschiedlichen theoretischen Ansätzen im Bereich des selbstregulierten Lernens Rechnung zu tragen, werden in dieser Arbeit drei Selbstregulationsmodelle exemplarisch vorgestellt. Ein in der Forschung häufig zitiertes Komponentenmodell ist das „Sechs-Komponenten-Modell“ von Boekaerts (1997; Kapitel 2.1.1), das sie zum „Drei-Schichten-Modell“ (Boekaerts, 1999; Kapitel 2.1.2) weiterentwickelte. Das sozial-kognitive Modell von Zimmerman (2000a, 2000b; Kapitel 2.1.3) zählt dagegen zu den Prozessmodellen der Selbstregulation. Die Modelle von Boekaerts (1997, 1999) wurden explizit für den Bereich selbstregulierten Lernens entwickelt. Das Modell von Zimmerman (2000a, 2000b) besitzt hingegen eine höhere Allgemeingültigkeit und wird unter anderem herangezogen, um selbstregulierte Lernprozesse zu beschreiben.

2.1.1 Sechs-Komponenten-Modell von Boekaerts

In ihrem Sechs-Komponenten-Modell fasst Boekaerts (1996, 1997) das selbstregulierte Lernen als einen komplexen und interaktiven Prozess auf. Der Prozess beinhaltet sowohl kognitive als auch motivationale Komponenten der Selbstregulation. Die motivationalen und die kognitiven Komponenten werden von Boekaerts als gleichermaßen wichtig erachtet (Baumert et al., 2000). Jede der sechs Komponenten des Modells bezieht sich dabei auf ein spezifisches Vorwissen des Lernenden im Bereich der kognitiven oder der motivationalen Selbstregulation (Wirth & Leutner, 2008, S. 102). Die unterschiedlichen Arten von Vorwissen sind wesentliche und wichtige Bestandteile selbstregulierten Lernens, die dem Lernenden in einer bestimmten Lernsituation potenziell zugänglich sind (Boekaerts, 1997, S. 163).

Boekaerts (1997) beschreibt die Komponenten der kognitiven und der motivationalen Selbstregulation auf der Ebene der Ziele, des Strategiegebrauchs sowie der Ebene des bereichsspezifischen Wissens (siehe Abbildung 1). Dieser Abschnitt widmet sich zunächst den *kognitiven* Komponenten der Selbstregulation. Die erste Komponente auf der Ebene des bereichsspezifischen Wissens umfasst das konzeptuelle (z.B. Fakten, Definitionen) und das prozedurale Wissen (z.B. bereichsspezifische Handlungsregeln). Die zweite Komponente kognitiver Selbstregulation bezieht sich auf allgemeine kognitive Strategien der Informationsverarbeitung: „The term ‚cognitive strategy‘ is used to refer to cognitive processes and behavior that students use during *actual* learning experiences to complete an assignment or to accomplish a goal implied by the academic task“ (Boekaerts, 1996, S. 105f.). Zu den kognitiven Strategien zählen Oberflächenstrategien (Wiederholung etc.), Tiefenstrategien (Elaboration, Organisation etc.), selektive Aufmerksamkeit, Dekodierung, Strukturierung sowie die Prozeduralisierung der Strategienutzung (z.B. Artelt, 2000). Die Anwendung der kognitiven Strategien erschließt neues, bereichsspezifisches Wissen. Gleichzeitig beeinflusst das bereichsspezifische Wissen

auch die Wahl der kognitiven Strategien. Die dritte und letzte Komponente kognitiver Selbstregulation stellen metakognitive Strategien dar (Baumert et al., 2000; Boekaerts, 1997). Metakognition ist ein Konzept, das sich auf „eine Reihe von Phänomenen, Aktivitäten und Erfahrungen [bezieht], die mit dem *Wissen* und der *Kontrolle* über eigene kognitive Funktionen (z.B. Wahrnehmung, Lernen, Gedächtnis, Verstehen, Denken) zu tun haben“ (Hasselhorn & Labuhn, 2008, S. 28). Metakognitive Strategien beziehen sich in erster Linie auf die Zielsetzung und –verfolgung, die Planung, Überwachung („Monitoring“), Bewertung und Steuerung eigener Lernprozesse. Sie beeinflussen die Nutzung kognitiver Strategien insofern, indem den Anforderungen der Lernsituation angemessene Strategien genutzt werden.

Die *motivationalen* Komponenten der Selbstregulation beziehen sich auf die Initiierung und Aufrechterhaltung von Lernhandlungen, Ergebnisbewertungen und Überzeugungen eigener Wirksamkeit und Fähigkeit (Spörer & Brunstein, 2006, S. 148). Boekaerts (1996, 1997) unterscheidet in ihrem Sechs-Komponenten-Modell zwischen motivationalen Überzeugungen (Komponente 4), Strategien (Komponente 5) und Regulationsstrategien (Komponente 6). Motivationale Überzeugungen beinhalten zum Beispiel selbstbezogene Kognitionen (z. B. Fähigkeitsselbstkonzepte) und motivationale Präferenzen (Zielorientierung, Interesse etc.; Baumert et al., 2000). Die Komponente „Motivationale Strategien“ bezieht sich auf die Regulation von Anstrengung, Aufmerksamkeit und Ausdauer sowie auf die Reduktion von Stress und negativen Emotionen (Baumert et al., 2000; Boekaerts, 1997). Die letzte motivationale Komponente umfasst die motivationalen Regulationsstrategien und wird von Boekaerts (1997) auf der Ebene der Ziele verortet. Sie beziehen sich auf mentale Repräsentationen der Handlungsintention und auf die willentliche Aufrechterhaltung der Lernhandlungen auch in der Gegenwart von Hindernissen und konkurrierenden Handlungstendenzen (Boekaerts, 1996).

Boekaerts (1996, 1997) zufolge weisen die einzelnen kognitiven und motivationalen Komponenten des selbstregulierten Lernens zahlreiche wechselseitige Abhängigkeiten auf. Beispielsweise beeinflussen die metakognitiven Strategien und die motivationalen Komponenten die Wahl der jeweiligen Lernstrategien. Das Wissen über den Inhaltsbereich des Lernprozesses liegt dem Wissen über die kognitiven Strategien zugrunde (vgl. Baumert et al., 2000). Dem Sechs-Komponenten-Modell von Boekaerts (1997) zufolge resultiert das selbstregulierte Lernen letztlich aus dem „Zusammenspiel der vom Lernenden selbst initiierten und gesteuerten Prozesse“ (Spörer & Brunstein, 2006, S. 148). Boekaerts (1996) betonte in einer frühen Publikation zum Sechs-Komponenten-Modell, dass dieses lediglich einen heuristischen Wert habe. Baumert et al. (2000, S. 5) machen überdies darauf aufmerksam, dass die Unterteilung der motivationalen und der kognitiven Selbstregulation in drei Ebenen (Ziele, Strategiegebrauch und Wissen) theoretische Abstriche erforderlich mache.

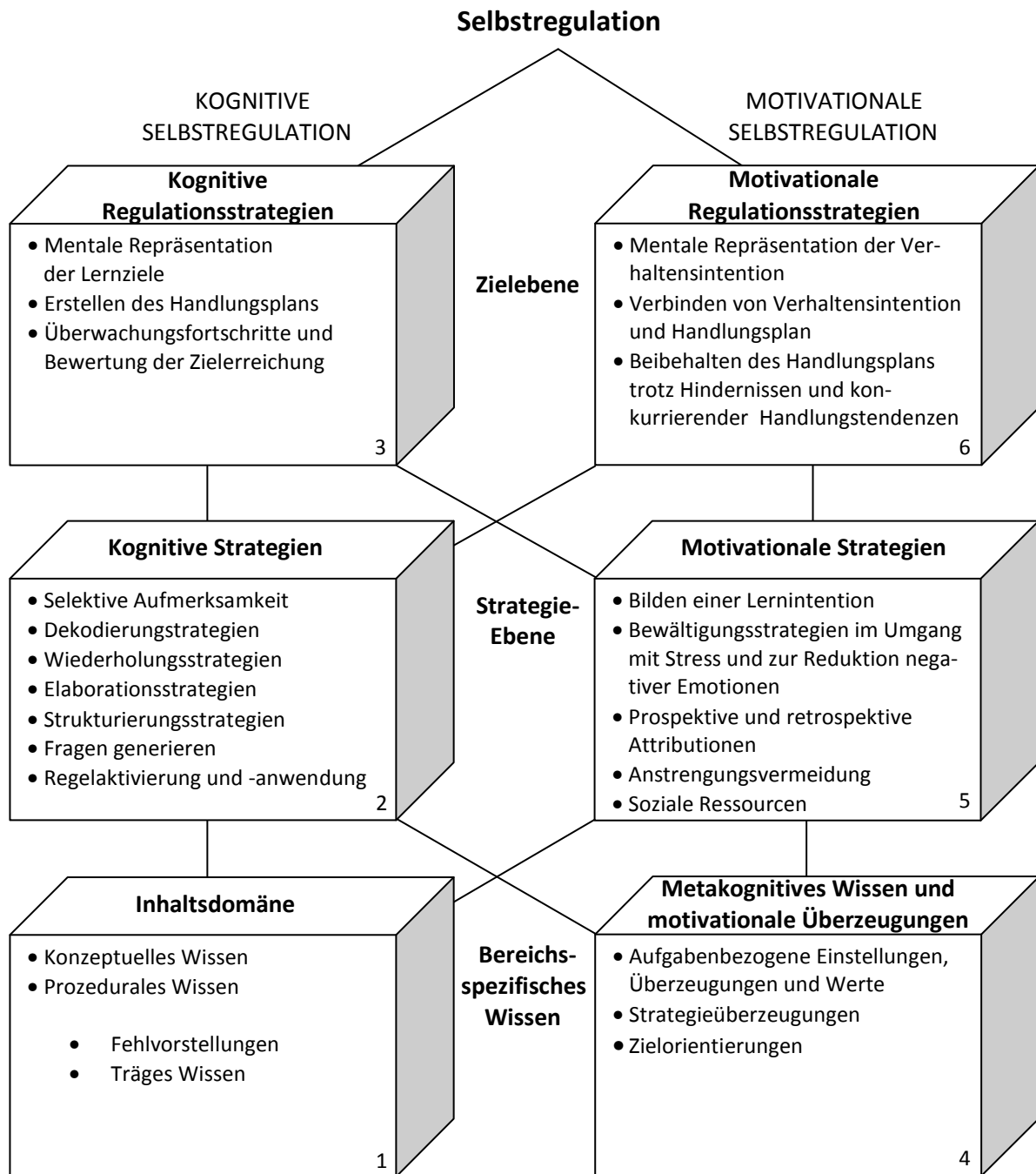


Abb. 1. Sechs-Komponenten-Modell von Boekaerts (1997, S. 164).

2.1.2 Drei-Schichten-Modell von Boekaerts

Mit dem „Drei-Schichten-Modell“ hat Boekaerts (1999) ein vereinfachtes Modell des selbstregulierten Lernens vorgelegt. Es baut auf den Postulaten des Sechs-Komponenten-Modells auf, verzichtet aber auf eine differenzierte Unterteilung der Komponenten (Baumert et al., 2000, S. 5). Das „Drei-Schichten-Modell“ geht auf Forschungsbefunde zu Lernstilen, Metakognitionen, Regulationsstilen sowie auf Theorien des Selbst zurück (Straka, 2006, S. 394). Boekaerts (1999) beschreibt in diesem Modell drei Regulationssysteme: die Regulation

des Selbst, der Lernprozesse sowie des Modus der Informationsverarbeitung (siehe Abbildung 2). Die drei Regulationssysteme lassen sich als Schichten oder Ebenen auffassen. Sie beziehen sich auf die motivationalen, metakognitiven und kognitiven Komponenten des selbstregulierten Lernprozesses.

Die erste, innerste Schicht des Selbstregulationsmodells betrifft die Regulation des Informationsverarbeitungsmodus (siehe Abb. 2). Die Regulation bezieht sich auf das Wissen über aufgabenspezifische, kognitive Strategien sowie über deren Anwendungsbedingungen (Baumert et al., 2000, S. 4; Spörer & Brunstein, 2006, S. 148). Lernende unterscheiden sich darin, wie sie sich einer Aufgabe nähern und welche kognitiven Strategien sie typischerweise einsetzen, um eine Aufgabe zu lösen (Boekaerts, 1999). Sie können dabei theoretisch zwischen kognitiven Strategien wie den oben erwähnten Tiefen- und Oberflächenstrategien wählen. Das bereichsspezifische Wissen bildet hier eine wichtige Grundlage für die Nutzung der kognitiven Strategien.

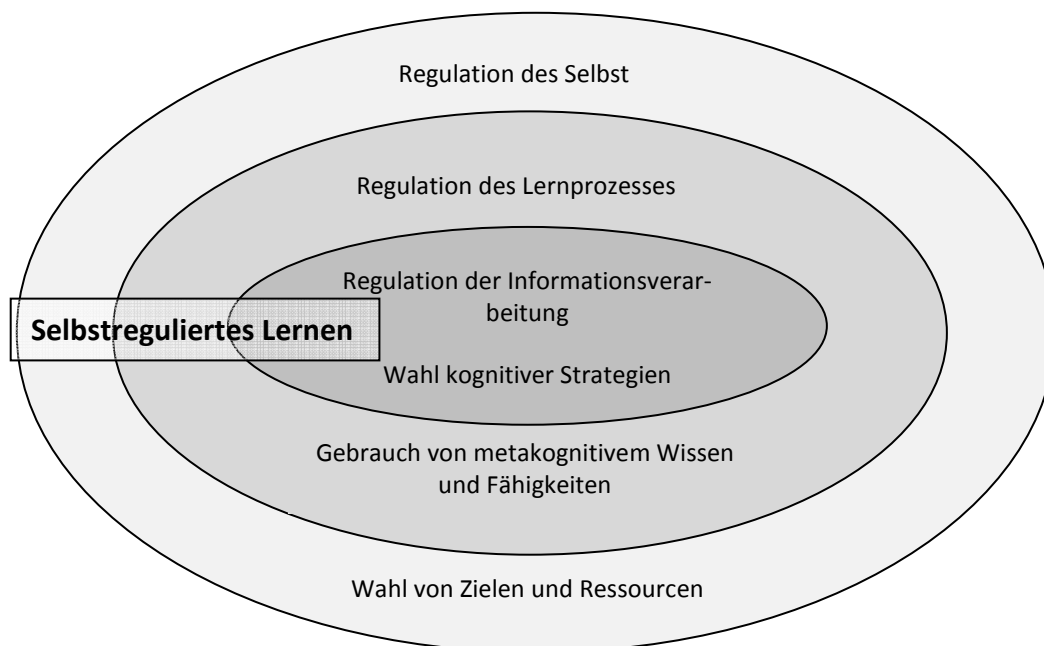


Abb. 2. Drei-Schichten-Modell der Selbstregulation von Boekaerts (1999; Übersetzung in Anlehnung an Straka, 2006, S. 394).

Das nächste Regulationssystem betrifft die Steuerung des Lernprozesses durch den Lernenden selbst (Boekaerts, 1999). Hierzu nutzt er metakognitives Wissen und Können (Baumert et al., 2000). Hierzu zählen das „Planen, Durchführen, Beobachten, Beurteilen und Korrigieren“ im Rahmen des Lernprozesses (Straka, 2006, S. 394). Ziel ist es, die eigenen Lernaktivitäten so zu organisieren, dass bereichsspezifisches Wissen erworben und Fähigkeiten erlernt werden können. Boekaerts (1999) betont die Kontextabhängigkeit der Regulation der Lernaktivitäten: Ein Lernender kann in einem bestimmten Bereich in der Lage sein, diesen

erfolgreich zu regulieren, in einem anderen dagegen nicht. Darüber hinaus kann es erforderlich sein, dass die Lernenden bei der Anwendung ihres metakognitiven Wissens durch andere Personen unterstützt werden.

Der Fokus des letzten Regulationssystems richtet sich auf die Regulation des Selbst (Straka, 2006, S. 394). Ein wichtiges Kennzeichen selbstregulierten Lernens ist nach Boekaerts (1999) das Setzen von und die Verpflichtung gegenüber selbstgesetzten Zielen durch den Lernenden. In selbst gesetzten Zielen wird deutlich, wie sich der Lernende selbst sieht. Die Verfolgung der Ziele setzt bestimmte Fähigkeiten auf Seiten des Lernenden voraus: Diese umfassen die Fähigkeit „(...) to define ongoing and upcoming activities in the light of their own wishes, needs, and expectancies, and their ability to protect their own goals from conflicting alternatives“ (Boekaerts, 1999, S. 451). Um die selbst gesetzten Ziele verfolgen zu können, muss der Lernende ausreichend Ressourcen mobilisieren.

Boekaerts (1996, 1997) hebt in ihren Modellen die Bedeutung spezifischer Fähigkeiten bzw. unterschiedlicher Arten von Vorwissen hervor, deren Zusammenspiel es dem Lernenden ermöglicht, erfolgreich selbstreguliert zu lernen. Demgegenüber steht in Zimmermans Prozessmodell die Abfolge von Phasen im Vordergrund, in denen die Anwendung des metakognitiven, kognitiven und motivationalen Vorwissens während einer konkreten Lernepisode beschrieben wird.

2.1.3 Prozessmodell von Zimmerman

Zimmerman entwickelte sein Prozessmodell der Selbstregulation auf der Basis sozial-kognitiver Theorien (vgl. Bandura, 1986). Aus sozial-kognitiver Perspektive stellt die Selbstregulation eine Interaktion aus personeninternen, verhaltens- und umgebungsbezogenen Prozessen dar (Straka, 2006, S. 395). Während des Lernprozesses können sich Veränderungen sowohl innerhalb einer Person, in deren Verhalten und in deren Lernumgebung vollziehen. Über spezifische Rückkoppelungsschleifen können diese Veränderungen (z.B. Abweichungen von einem Standard) registriert werden.

Das selbstregulierte Lernen besteht nach Zimmerman (2000b, S. 222) darin, dass der Lernende den Lernprozess sowie seine (motivationalen) Überzeugungen kontinuierlich an die Veränderungen in der eigenen Person, im Verhalten oder in der Lernumgebung anpasst. Zimmerman (2000a, 2000b, 2002) unterscheidet drei Prozesse der Selbstregulation: eine verhaltensbezogene, eine umgebungsbezogene sowie eine innere Selbstregulation (siehe Abbildung 3).

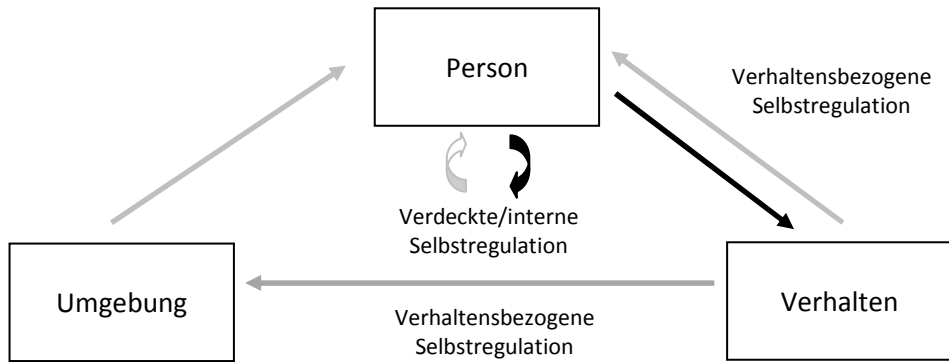


Abb. 3. Selbstregulation als Interaktion von personeninternen, umgebungs- und verhaltensbezogenen Prozessen. Die hellen Pfeile stellen Rückkoppelungsschleifen dar. Der dunkle Pfeil bezieht sich auf den Strategiegebrauch (Zimmerman, 2000b, S. 222; Abbildung entnommen aus Straka, 2006, S. 395).

Die *verhaltensbezogene Selbstregulation* bezieht sich auf die Selbstbeobachtung im Lernprozess sowie auf die Steuerung des Handlungsprozesses. Letzere zeigt sich im Einsatz unterschiedlicher Lernstrategien. Die *umgebungsbezogene Selbstregulation* betrifft demgegenüber die Beobachtung der Umgebungsbedingungen, so dass das eigene Handeln entsprechend angepasst werden kann. Auch kann der Lernende auf seine Umwelt einwirken, z.B. um sich günstige Lernbedingungen zu schaffen. Die *innere Selbstregulation* beinhaltet die Überwachung und die Anpassung der kognitiven und affektiven Zustände des Lernenden. Zimmerman (2000b, S. 223; vgl. Bandura, 1986) zufolge muss der Lernende alle drei Selbstregulationsprozesse aufeinander abstimmen, um seine persönlichen Lernziele erreichen zu können.

Das selbstregulierte Lernen in einer konkreten Lernsituation beschreibt Zimmerman (2000a, 2000b, 2002, 2008) als einen zyklischen Prozess, ähnlich jenen in kybernetischen Modellen (vgl. Miller et al., 1960; Carver, 2004; Carver & Scheier, 2000). Diesen Prozess unterteilt er in drei Phasen: Vorbereitungsphase („Forethought phase“), Handlungsphase („Performance phase“) und Selbstreflexionsphase („Self-Reflection Phase“). Abbildung 4 illustriert das Prozessmodell von Zimmerman. In jeder Phase dieses Modells finden Selbstregulationsprozesse statt. In Übereinstimmung mit Boekaerts (1997, 1999) beinhalten diese Selbstregulationsprozesse bei Zimmerman sowohl motivationale als auch kognitive und metakognitive Aspekte der Selbstregulation.

Die Selbstregulationsprozesse, die der eigentlichen Lernhandlung vorausgehen, sind in der *Vorbereitungsphase* zusammengefasst. Hierzu zählt zunächst die Aufgabenanalyse, welche in die Zielsetzung und die strategische Planung mündet. Darüber hinaus spielen selbstbezogene, motivationale Überzeugungen wie die Selbstwirksamkeitserwartung (*Self-Efficacy*; Bandura, 1997) eine entscheidende Rolle. Bandura (1986) definiert *Self-Efficacy* als „people's judgments of their capabilities to organize and execute courses of action required to attain designated

types of performances“ (S. 391). Die Aufgabenanalyse sowie die Entwicklung motivationaler Überzeugungen wie der Self-Efficacy gehen der eigentlichen Lernhandlung voraus und beeinflussen diese.

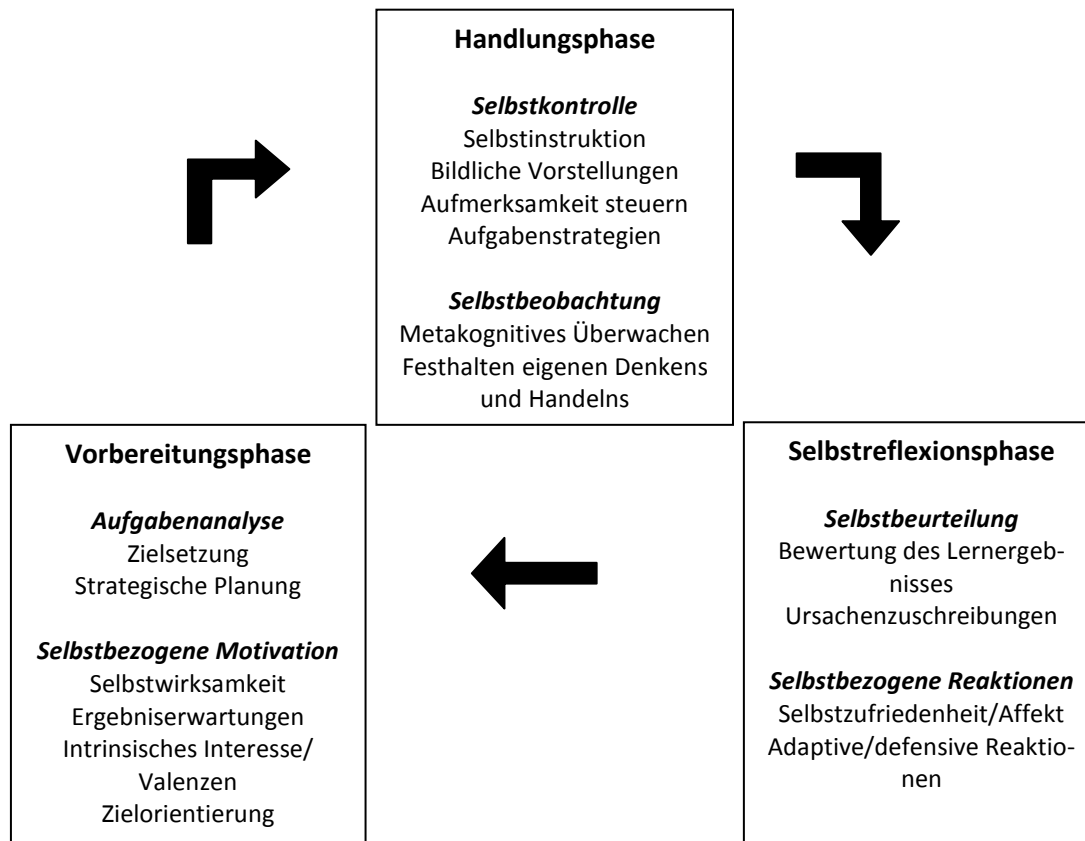


Abb. 4. Phasen und Teilprozesse der Selbstregulation nach Zimmerman (2000a; Übersetzung in Anlehnung an Straka, 2006, S. 396).

Die *Handlungsphase* betrifft das Geschehen während der Lernaktivitäten. In dieser Phase stehen kognitive und metakognitive Prozesse im Vordergrund. Zimmerman unterteilt diese Phase in Prozesse der Selbstkontrolle und Selbstbeobachtung. Die Selbstkontrolle manifestiert sich darin, dass der Lernende seine Aufmerksamkeit fokussiert. Er identifiziert Aufgabenstrategien, die er bei der Lösung der Aufgabe als zielführend betrachtet (Zimmerman & Kitsantas, 2005). Selbstgespräche dienen dazu, dass der Lernende seinen Lernprozess steuert und Kontrolle über die eigenen Lernhandlungen erhält. Handlungen können systematisch variiert werden, um diese den Erfordernissen der Lernaufgabe anzupassen. Der Lernende beobachtet den eigenen Lernvorgang und registriert Fortschritte und Probleme. Diese Selbstbeobachtung schafft die Voraussetzung für die kontinuierliche Anpassung des eigenen Lernverhaltens an die Umgebungsbedingungen bzw. an die Anforderungen der Lernaufgabe.

An die eigentlichen Lernhandlungen schließt sich die *Selbstreflexionsphase* an. Die Selbstregulationsprozesse in dieser Phase beeinflussen die Selbstregulation in zukünftigen Lernphasen (Zimmerman, 2000a, 2000b). In der Selbstreflexionsphase beurteilt der Lernende seine eigenen Lernleistungen anhand eines Standards oder Ziels. Der Lernende stellt darüber hinaus Überlegungen zu den Ursachen des Lernergebnisses an (Kausalattributionen; vgl. Weiner, 1986). Die Selbstreflexionsphase ist zudem durch selbstbezogene Reaktionen gekennzeichnet. Diese beinhalten affektive Reaktionen des Lernenden auf das eigene Lernergebnis; der Lernende verspürt Zufriedenheit oder Unzufriedenheit. Es wird angenommen, dass sich eine hohe Zufriedenheit mit der eigenen Lernleistung positiv auf die Selbstwirksamkeit und das Interesse vor Beginn einer neuen Lernepisode auswirkt (Zimmerman, 2008). Der selbstreguliert Lernende ist bestrebt, negative Emotionen in Folge von Misserfolg zu vermeiden. Defensive selbstbezogene Reaktionen wie das Prokrastinieren verhindern, dass ungünstige motivationale Orientierungen zukünftige Lernprozesse negativ beeinflussen.

Da der Lernprozess in Zimmermans Modell als Kreislauf beschrieben wird, beeinflussen die Erfahrungen einer Lernepisode das zukünftige Selbstregulationsverhalten: „This view of self-regulation is cyclical in that self-reflections from prior efforts to learn affect subsequent forethought processes (e.g., self-dissatisfaction) will lead to lower levels of self-efficacy and diminished effort during subsequent learning” (Zimmerman, 2002, S. 68).

Zimmerman und Kollegen erforschten den empirischen Zusammenhang zwischen selbstreguliertem Lernen und Lernerfolg. Insgesamt nutzten erfolgreich Lernende in stärkerem Maße Selbstregulationsprozesse (z.B. den Einsatz von Lernstrategien und die Selbstbeobachtung), um Einfluss auf den Lernverlauf und das Lernergebnis zu nehmen. Es zeigten sich höhere Lernzuwächse bei den selbstreguliert Lernenden im Vergleich zu den nicht selbstreguliert Lernenden (Schunk & Zimmerman, 1994; Zimmerman & Martinez-Pons, 1988). Der sozial-kognitiven Lerntheorie folgend, nahmen die erfolgreich Lernenden auch stärker Hilfe von Lehrkräften, Tutoren und Mitschülern in Anspruch (Zimmerman & Kitsantas, 2005; Schunk & Zimmerman, 1994). Die Ergebnisse der hier dargestellten, ausgewählten Studien belegen, dass das selbstregulierte Lernen positiv mit der Lernleistung bzw. dem Wissenserwerb assoziiert ist. Die Förderung selbstregulierten Lernens wurde folglich auch zum Gegenstand von Interventionsstudien (z.B. Cleary & Zimmerman, 2004; Dignath & Büttner, 2008; Landmann, Pöhl, & Schmitz, 2005; Perels, Gürtler & Schmitz, 2005).

2.2 Problemlösen

Lernen und Problemlösen sind in der psychologischen Forschung eng miteinander verbunden (Vollmeyer, Burns & Holyoak, 1996). Gold (2003, S. 108) zufolge kann „Lernen

durch schlussfolgerndes Denken und zielgerichtetes Handeln“ als Problemlösen aufgefasst werden. Funke und Zumbach (2006) erklären die besondere Bedeutung des Problemlösens für Lernprozesse folgendermaßen: „Der Lernende trifft ständig auf Hindernisse und Barrieren. Um diese zu überwinden, bedarf es des Nachdenkens und Problemlösens. Das Lösen von Problemen ist dabei als Handlung zu verstehen, die sich wiederum auf das Denken auswirkt“ (S. 206).

In Lehr-Lern-Kontexten werden Problemlöseaufgaben gezielt eingesetzt, um Lernprozesse positiv zu beeinflussen. Insbesondere in konstruktivistischen Lehr-Lern-Theorien stellt das Lernen durch Problemlösen eine wichtige methodische Komponente dar (Schmidt-Weigand, Hänze & Wodzinski, 2009). Die Methode ist wesentlicher Bestandteil des exploratorischen oder entdeckenden Lernens (*Discovery learning*; Bruner, 1961; Greif, 1990). Lernende werden dabei mit realistischen Problemen konfrontiert und bei der Lösung des Problems begleitet. Das selbstgesteuerte Erkunden eines Computerprogramms stellt eine Form des exploratorischen Lernens dar. Das entdeckende Lernen soll dabei helfen, Wissen über den Sachinhalt aufzubauen sowie allgemeine und spezifische Problemlösestrategien zu erwerben (Gold, 2008; Greif, 1996).

Die Beziehung zwischen Problemlösen und (selbstregulierten) Lernprozessen wird in den folgenden Abschnitten näher spezifiziert. Hierzu werden in Kapitel 2.2.1 ausgewählte theoretische Modelle des Problemlösens vorgestellt. In Übereinstimmung mit Modellen des selbstregulierten Lernens (Boekaerts, 1997, 1999; Zimmerman, 2000) berücksichtigen einzelne Problemlösemodelle sowohl motivationale, kognitive und metakognitive Aspekte. Dies trifft insbesondere für das Handlungsregulationsmodell von Dörner (2006) zu. Die Bedeutung von Problemlöseprozessen für das Erlernen eines Computerprogramms wird in Kapitel 2.2.2 aufgezeigt.

2.2.1 Modelle des Problemlösens

Ausgehend von Informationsverarbeitungstheorien lassen sich Probleme³ nach Newell und Simon (1972, S. 72) folgendermaßen umschreiben: „A person is confronted with a problem when he wants something and does not know immediately what series of actions he can perform to get it“. Formal können Probleme als Diskrepanzen zwischen einem gegebenen Problemzustand (*Ausgangszustand*) und einem erwünschten Zustand (*Zielzustand*) definiert werden (Mayer, 1992; Neber, 1987). Das eigentliche Problem besteht hierbei in der Überführung

³ Vollmeyer und Funke (1999) merken an, dass im US-amerikanischen Kontext im Gegensatz zum deutschen (Dörner, 1987) nicht zwischen „Problemen“ und „Aufgaben“ differenziert wird. In dieser Arbeit verweisen die Begriffe „Problem“ und „Aufgabe“ ebenfalls auf das gleiche Konstrukt.

(*Transformation*) des Ausgangszustands in den Zielzustand (Funke, 2006a; Newell & Simon, 1972). Die Anwendung von *Operatoren* ist ein weiteres wesentliches Element von Problemlöseprozessen. Ein Operator „bezeichnet eine Handlung, die den vorliegenden Problemzustand in einen anderen Problemzustand transformiert“ (Anderson, 2007, S. 291). Häufig sind mehrere Schritte nötig, um den Zielzustand zu erreichen. Hierzu kann der Problemlöseprozess in mehrere Teilziele zerlegt werden, für die unterschiedliche Operatoren benötigt werden. Folglich kann das Problemlösen auch als zielgerichtete Sequenz kognitiver Operationen beschrieben werden (Anderson, 2007, S. 257).

Newell und Simon (1972, S. 55f.) sehen das Problemlösen als das Absuchen eines Problemraums durch den Problemlöser an. Der Problemraum ist die interne Repräsentation der Aufgabenumgebung in einer gegebenen Situation. Er besteht „aus einem Anfangszustand (dem Problem), aus einer Menge von möglichen Zwischenzuständen, die durch die Anwendung bestimmter Operationen, Strategien oder Heuristiken erreicht werden können, und einem Zielzustand, der die Lösung des Problems markiert“ (Öllinger & Knoblich, 2006, S. 431).

Das Absuchen des Problemraums kann nach Funke (1991, zit. nach van der Linden et al., 2003, S. 483) als Explorationsprozess bezeichnet werden. Explorieren wird Oerter zufolge (1987, zitiert nach Greif, 1990, S. 151) folgendermaßen definiert: „Das Umfeld bzw. Objekt wird systematisch oder planlos durchsucht oder erforscht.“ Dabei wird die Art des Umgangs mit dem Objekt in geringerem oder stärkerem Ausmaß variiert. Ziel des Explorationsprozesses ist es, Einsicht in die Aufgabe zu erhalten und Möglichkeiten zur Lösung des Problems zu entdecken.

Aus der „europäischen“ Perspektive von Frensch und Funke (1995, S. 18; vgl. Funke & Frensch, 2007) beziehen sich Problemlöseprozesse auf Situationen, in denen Hindernisse überwunden werden müssen, um vom Anfangszustand zum Zielzustand zu gelangen. Hindernisse bestehen z. B. im mangelnden Wissen über geeignete Operatoren beziehungsweise korrekte Handlungssequenzen, die zur Überführung des Anfangszustands in den Zielzustand benötigt werden (Mayer, 1992). Darüber hinaus ist dem Problemlösenden häufig der Anfangszustand eines Problems nicht in vollem Maße bekannt. Um diese Hindernisse zu überwinden, sind mehrstufige Verhaltensaktivitäten und/oder kognitive Aktivitäten erforderlich.

Problemlösen und selbstreguliertes Lernen werden als eine Form aktiven und zweckgerichteten Verhaltens angesehen (Neber, 1987; Newell & Simon, 1972). Nach Poissant, Poellhuber und Falardeau (1994) strebt auch der Lernende an, einen Zielzustand zu erreichen: Wissen zu erwerben oder Fertigkeiten zu erlernen. Die Diskrepanz zum Anfangszustand besteht darin, dass der Lernende Kenntnisse/Fertigkeiten erwirbt, die er zu Beginn des Lernprozesses noch nicht hatte. Hierzu muss er das Lernmaterial systematisch explorieren. Problemlö-

sprozesse können das Erlernen *expliziten, deklarativen Wissens* sowie *impliziten, prozeduralen Wissens* beinhalten (vgl. Anderson, 2007). Im expliziten Lernmodus wird verbalisierbares Wissen über ein Problem erworben. Dieses Wissen liegt in Wenn-Dann-Aussagen vor. Allgemein ist implizites Wissen weniger bewusst zugänglich. Es schlägt sich jedoch häufig in (beobachtbaren) Handlungen nieder.

Selbstreguliertes Problemlösen. Ähnlich wie der Prozess selbstregulierten Lernens (Schmitz, 2001; Zimmerman, 2000) umfasst auch der Problemlöseprozess mehrere Phasen: „Komplexe Probleme setzen eine hierarchische Organisation des problemlösenden Handelns voraus. Solche hierarchischen Strukturen zur Organisation von Einzeloperationen wurden von Miller et al. (1960) als ‚Plan‘ beschrieben“ (Neber, 1987, S. 51). Modelle des Problemlösens beinhalten in der Regel vier bis fünf voneinander abgrenzbare Stufen (Dörner & Schaub, 1994; Funke, 2006b; Poissant et al., 1994; vgl. auch Anderson, 2007). Exemplarisch wird im Folgenden das von Dörner (2006; Dörner & Schaub, 1994) vorgestellte Modell der Handlungsregulation in komplexen Problemlöseprozessen näher beschrieben. In diesem Modell stehen die Steuerungs- und Kontrollmöglichkeiten des Individuums im Vordergrund. Das Modell ist in Abbildung 5 grafisch wiedergegeben.

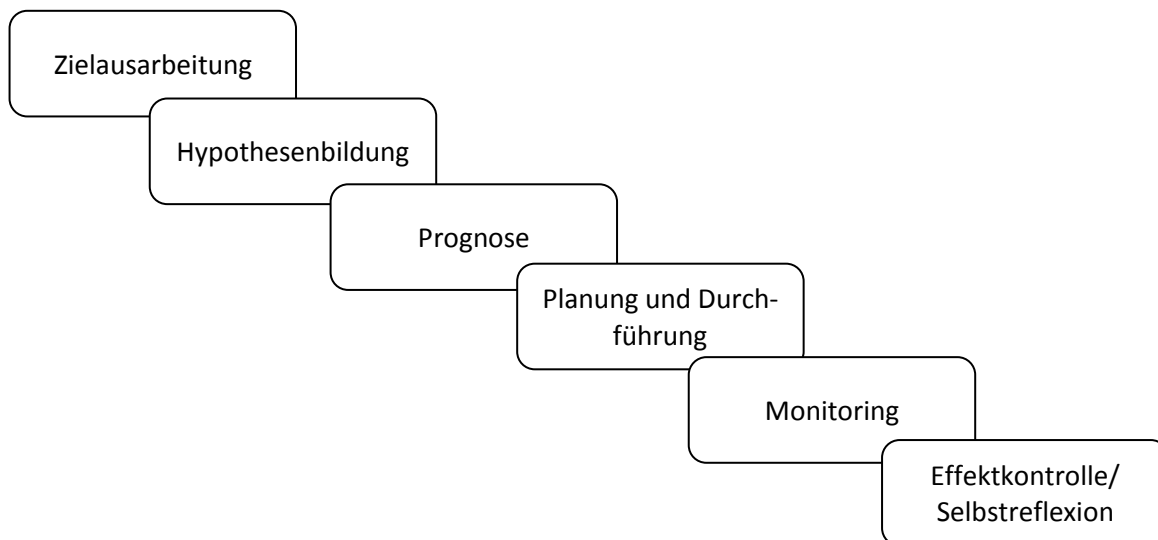


Abb. 5. Phasen der Handlungsregulation beim komplexen Problemlösen (nach Dörner, 2006, S. 67; Dörner & Schaub, 1994, S. 435).

In der ersten Handlungsphase (Zielausarbeitung) verschafft sich eine Person zunächst Klarheit über das Ziel ihrer Handlungen. Bei komplexen Problemen wird das Ziel gegebenenfalls in Teilziele unterteilt. Die nachfolgenden Handlungen werden an diesem Ziel bzw. diesen Teilzielen ausgerichtet. In der darauffolgenden Phase (*Hypothesenbildung*) holt der Problemlösende Informationen über die Struktur des Problems bzw. des Systems ein, in dem das Prob-

lem verortet ist. Hat der Problemlösende genug Informationen gesammelt, integriert er die relevanten Informationen und leitet konkrete Hypothesen zur Funktionsweise des Systems ab. Wie oben beschrieben, kann sich die Aufgabenstellung im Falle komplexer Probleme während des Prozesses verändern.

Der Problemlösende stellt dann im nächsten Schritt *Prognosen* zu diesen Veränderungen auf, um diese bei seiner Handlungsplanung entsprechend berücksichtigen zu können. Die *Handlungsplanung* erfolgt im nächsten Schritt. In dieser Phase muss eine Sequenz zielgerichteter Handlungsschritte gefunden werden. Der Problemlösende muss sich entscheiden, welche Handlungsschritte er auswählt und durchführt. Nach der Planung muss er die Handlungssequenz, für die er sich entschieden hat, umsetzen. Während der Umsetzung der Handlungen muss der Problemlösende die Umsetzung seiner Handlungen überwachen und die Effekte kontrollieren (*Monitoring*).

2.2.2 Klassifikationen von Problemen

In der Problemlöseforschung wird grundlegend zwischen „gut strukturierten“ (einfachen) und „schlecht strukturierten“ (schwierigen) Problemen differenziert (Funke & Zumbach, 2006; Mayer, 1992). Die Charakterisierung als „einfach“ bezieht sich demnach nicht in erster Linie auf den Schwierigkeitsgrad der Aufgabe. Für diese Probleme sind die Aufgabenstellung, die verfügbaren Problemlöseschritte und sowie die Lösungswege meist eindeutig formulierbar (Öllinger & Knoblich, 2006). Gleichzeitig bleibt bei einfachen Problemen die ursprüngliche Problemstellung während des Problemlöseprozesses unverändert. Demgegenüber wird ein Problem als komplex eingestuft, wenn sich die ursprüngliche Problemstellung während der Problemlösung verändert (Vollmeyer & Funke, 1999). Die Hindernisse, die bei der Transformation des Ausgangszustands in den Zielzustand überwunden werden müssen, sind dabei komplex, dynamisch oder intransparent (Frensch & Funke, 1995).

Greeno (1978; Greeno & Simon, 1988) beschreibt vier verschiedene Problemtypen, die sich in Bezug auf die notwendigen Denkopoperationen zur Transformation eines Ausgangs- in einen Zielzustand voneinander abgrenzen lassen (Anderson, 2007; Mayer, 1972, S. 6; Neber, 1987, S. 7):

1. Probleme der *Transformation*: Sowohl der Ausgangs- als auch der Zielzustand erscheinen dem Problemlösenden als klar definiert. Das Problem besteht darin, eine Handlungsabfolge zu finden, die es erlaubt, den Zielzustand zu erreichen.
2. Probleme der *Induktion einer Struktur*: Diese erfordern Prozesse des induktiven Schließens von einzelnen Beispielen oder Ereignissen auf allgemeine Regeln und Muster.

3. Probleme der *Umordnung von Elementen*: Eine Person muss eine Menge von Elementen so umordnen, dass sie der Idee des Zielzustands entsprechen.
4. Probleme der *Bewertung deduktiver Argumente*: Diese betreffen die korrekte, logische Ableitung eines Konsequens bei einem gegebenen Antecedens.

2.3 Selbstreguliertes Erlernen eines Computerprogramms

Im Sinne der oben dargestellten Problemlösetheorien kann das selbstregulierte Erlernen eines Computerprogramms als eine Problemlöseaufgabe aufgefasst werden (Kang & Yoon, 2008; Poissant et al., 1994; van der Linden et al., 2001; van der Linden, Frese & Sonntag, 2003). Im Folgenden soll dies an einem Beispiel konkretisiert werden; dem Erlernen einer Statistiksoftware. Beim Erlernen dieser Software verändert sich die Aufgabenstellung im Laufe des Problemlöseprozesses nicht, so dass nicht von einem komplexen Problem gesprochen werden kann (Funke, 2006a, 2006b; Funke & Zumbach, 2006). Das Aufgabenziel ist in der Regel klar: Möchte der Computernutzer eine Mittelwertberechnung durchführen, ist der erwünschte *Zielzustand* die Ausgabe eines korrekten Ergebnisses dieser Berechnung. Ein *Hindernis*, diesen Zielzustand zu erreichen, stellt z. B. das Auffinden der korrekten Programmoptionen dar.

Ein Novize muss zunächst die Funktionsweise des unbekanntes Statistikprogramms, die notwendigen Handlungsoptionen und -schritte explorieren, bevor er eine Statistikaufgabe mit Hilfe des Computers erfolgreich lösen kann (Newell & Simon, 1972). In dieser Phase ist der Lernende mit kognitiven Konflikten und Informationsdefiziten konfrontiert, die er bewältigen muss (Greif, 1996). Hat der Lernende keine Informationen über dieses spezielle Statistikprogramm und auch nicht über Teilziele und die korrekte Handlungssequenz zur Lösung des Problems, findet beim Erlernen eines neuen Computerprogramms häufig ein Wissenstransfer statt: Der Computernutzer erkennt ähnliche Funktionen und Strukturen wieder, die er bereits beim Erlernen einer früheren Software entdeckt und genutzt hat. Insbesondere ähnliche Visualisierungen auf der Benutzeroberfläche des Programms (z.B. Menüfunktionen) liefern ihm Hinweise darauf, wie das Programm zu bedienen ist (Dutke, 1994). Im Rahmen des Explorationsprozesses baut der Computernutzer Wissensstrukturen auf, indem er Handlungsschritte mit bekannten und unbekanntes Funktionen des Programms gezielt wiederholt und variiert (Greif, 1996).

Bezugnehmend auf Dörner und Schaub (1994, S. 435 ff.) kann das Erlernen einer Computeranwendung als Handlungsregulationsprozess beschrieben werden (siehe Kapitel 2.2.1). Die Reihenfolge der Handlungsregulationsschritte kann an folgendem Beispiel verdeutlicht werden: Ein Novize hat die Aufgabe innerhalb eines Statistikprogramms Mittelwerte für

eine ausgewählte Variable separat für zwei verschiedene Personengruppen zu berechnen. Er setzt sich das *Ziel*, eine Mittelwertanalyse durchzuführen und sich das Ergebnis im Ausgabefenster des Statistikprogramms anzeigen zu lassen. *Teilziele* sind das Auffinden der korrekten Menüoption, die Auswahl von Variablen, die Ausführung des Befehls, das Lokalisieren sowie die Interpretation der Ergebnisanzeige. Der Computernutzer muss *Hypothesen* dazu bilden, welche Objekte das Programm beinhaltet (z. B. Variablen, Daten, Fenster) und welche Aktionen durchführbar sind (z. B. Bedienen von Schaltflächen zum Ausführen von Befehlen). Darüber hinaus muss er Annahmen darüber bilden, wo Objekte und Kommandos für die Aktionen lokalisiert sind (z. B. im *Drop-Down*-Menü der Menüleiste im Datenfenster).

Diese Wissensstrukturen müssen im Explorationsprozess erworben werden, indem ein „Aufbau interner Repräsentationen mit multiplen Relationen“ (Neber, 1993, S. 206) erfolgt. Zu diesem Zweck müssen zunächst *Informationen* eingeholt werden. Im nächsten Schritt entwickelt der Computernutzer *Prognosen* darüber, zu welchem Ergebnis konkrete Handlungsweisen im Programm führen (z. B. nimmt er an, dass ein Ausführen des Befehls „Mittelwerte berechnen“ zur Anzeige der Mittelwerte im Ausgabefenster führt). Der Computernutzer entwickelt einen *Plan* zur Ausführung mindestens einer konkreten Handlungsabfolge, indem er sich zunächst vornimmt den Mauszeiger auf der Menüleiste zu positionieren und die Option „Analysieren“ anzuklicken. Danach wählt er „Deskriptive Statistiken“ aus bis er das Teilziel „Auffinden der korrekten Befehlsfunktion“ erfolgreich bewältigt hat. Wenn er sich für diesen Handlungsweg entschieden hat und diesen ausführt, muss der Computernutzer erstens seine Kompetenzen dazu nutzen, den Ausführungsprozess zu überwachen („*Monitoring*“; vgl. Hasselhorn & Labuhn, 2008; Wirth & Leutner, 2008). Zweitens muss er das Ergebnis seiner Handlungen bewerten (z. B. indem er feststellt, ob die Wahl der Befehlsfunktion dazu geführt hat, dass er die erwünschten Mittelwerte berechnen konnte). Diese *Effektkontrolle* kann dazu führen, dass sich der Computernutzer entscheidet, die Handlungsabfolge noch einmal zu verändern und einen anderen Lösungsweg zu testen (*Revision der Handlungsstrategie*). Fehlermeldungen des Computers können dabei wichtige Hinweise liefern, um das eigene Verhalten anzupassen. Dabei muss der Lernende auch eine emotionale Selbstregulation betreiben: Er muss dafür sorgen, dass er trotz negativer Emotionen wie Ärger und Ängstlichkeit weiterhin konzentriert an der Lösung des Problems arbeitet (Kay & Loverock, 2008).

Greeno (1978) zufolge entspricht das Erlernen eines Computerprogramms in erster Linie dem *Problemtyp der Transformation*. Das Aufgabenziel, das heißt der zu erzeugende Zielzustand, ist klar vorgegeben: Der Lernende will sich das Ergebnis einer Mittelwertberechnung vom Computerprogramm anzeigen lassen. Hierzu muss er die zielführende Folge an Befehlen bzw. Handlungsschritten finden. Will der Lernende nach erfolgreicher Berechnung der Mit-

telwerte weitere Maße zentraler Tendenz berechnen, kann er auf dem bisher erworbenen Wissen über die korrekte Handlungssequenz aufbauen. Da sich die erste und die zweite Aufgabe in ihrem Ausgangs- und Zielzustand ähneln, kann der Lernende die Lösungsstrategie des ersten Problems auf die Lösung des zweiten Problems übertragen. Er bildet *Analogien* zwischen der Lösung der ersten und der zweiten Aufgabe. Folglich kann das Erlernen eines Computerprogramms auch als *Problem der Induktion einer Struktur* aufgefasst werden (Neber, 1987, S. 7).

2.4 Zusammenfassung

In Kapitel 2 wurden theoretische und empirische Befunde zum selbstregulierten Lernen und Problemlösen zusammengefasst und erläutert. Aufgrund der positiven Effekte selbstregulierten Lernens auf den Lernprozess wird in der Pädagogischen Psychologie dessen Einsatz und Förderung als wünschenswert angesehen. Trotz der Unterschiede der in der Literatur vorgeschlagenen Modelle selbstregulierten Lernens, weisen diese bedeutsame Gemeinsamkeiten auf (Meece & Painter, 2008): Selbstreguliertes Lernen beinhaltet kognitive, motivationale sowie metakognitive Komponenten. Deren Wechselbeziehungen tragen zum Gelingen des Lernprozesses bei. Der selbstregulierte Lernprozess lässt sich in Phasen einteilen. Die erste Phase bezieht sich dabei auf Kompetenzen des Lernenden und auf die Vorbereitung der Lernaktivitäten. In Phase zwei (der Handlungsphase) werden in den Modellen die Strategienutzung sowie Monitoringprozesse hervorgehoben. Am Ende der Lernphase reflektiert der Lernende sein Verhalten und passt das Verhalten gegebenenfalls (neuen) Lernbedingungen an.

In der vorliegenden Arbeit steht das selbstregulierte Erlernen eines Computerprogramms im Vordergrund: Im Rahmen seiner Auseinandersetzung mit dem Computerprogramm erwirbt der Lernende selbstständig Wissen und Kompetenzen, die ihn letztlich zur zielführenden Nutzung des Computerprogramms befähigen. Die Ausführungen zum selbstregulierten Lernen und Problemlösen verdeutlichen die engen Beziehungen zwischen Lern- und Problemlöseprozessen.

Die Bearbeitung einer Aufgabe mit Hilfe eines unbekanntes Computerprogramms stellt ein Problem dar. Lernende unterscheiden sich darin, wie sie sich diesem Problem nähern und auf welche Weise sie nach Lösungswegen suchen. Interindividuelle Unterschiede manifestieren sich zum Beispiel in den Problemlösestrategien der Lernenden. Dieser Aspekt des selbstregulierten Lernens bzw. Problemlösens wird im folgenden Kapitel 3 beleuchtet.

3. Kognitive Aspekte des Problemlöseprozesses

Prozesse des selbstregulierten Lernens sowie des Problemlösens umfassen sowohl kognitive als auch nicht-kognitive Aspekte (siehe Kapitel 2.1; Boekaerts, 1997, 1999; Dörner & Schaub, 1994; Rheinberg et al., 2000). Zu den kognitiven Aspekten zählen z.B. die Arbeitsgedächtniskapazität sowie das domänenspezifische Wissen (Frensch & Funke, 1995). Lern- bzw. Problemlösestrategien nehmen unter den kognitiven Aspekten eine besondere Rolle ein (Anderson, 2007; Bielaczyc, Pirolli & Brown, 1995; Poissant et al., 1994; Vollmeyer et al., 1996). Sie werden auch als „kognitive Strategien“ bezeichnet (Boekaerts, 1996, S. 105-106). Zahlreiche Studien zum Erlernen eines Computerprogramms konnten zeigen, dass sich erfolgreiche und weniger erfolgreiche Problemlösende in erster Linie in der Qualität der von ihnen eingesetzten Strategien unterscheiden (Carroll & Mack, 1983; Dörner, 1980; van der Linden et al., 2001, 2003; Vollmeyer et al., 1996). Die Qualität der eingesetzten Lern- und Problemlösestrategien wird auch als „Strategiesystematik“ bezeichnet (Vollmeyer & Rheinberg, 1998, 2000).

Die Strategiesystematik beim Erlernen eines Computerprogramms bildet einen Forschungsschwerpunkt der vorliegenden Arbeit. Kapitel 3 bietet zunächst eine Übersicht über theoretische Annahmen zur Strategiesystematik. Darüber hinaus werden Befunde empirischer Studien zusammengefasst, in denen die Beziehung zwischen Strategiesystematik und (Lern-) Leistung erforscht wurde. Kapitel 4 ist den motivationalen Aspekten des Problemlöseprozesses gewidmet. Die motivationalen Aspekte (z.B. das Interesse am Lerngegenstand, die Misserfolgsbefürchtung) werden den nicht-kognitiven Aspekten von Lern- und Problemlöseprozessen zugerechnet (Frensch & Funke, 1995).

3.1 Problemlöse- bzw. Explorationsstrategien

Problemlösestrategien werden nach Mayer (1992, S. 395) als „plans to solve problems“ beschrieben. In Anlehnung an das Konzept der Lernstrategien⁴ umfassen Problemlösestrategien dabei die Koordination mehrerer zweckgerichteter Verhaltensweisen („Taktiken“) und Kognitionen im Hinblick auf das Erreichen eines Ziels (Jamieson-Neal & Winne, 2003; Lind & Sandmann, 2003). Mayer (1972, S. 185) bezeichnet Problemlösestrategien auch als „search strategies“. In den Arbeiten von van der Linden und Kollegen (van der Linden et al., 2001, 2003) werden die Begriffe „Problemlösestrategien“ und „Explorationsstrategien“ synonym verwendet. Die vorliegende Arbeit folgt diesem Ansatz. Explorationsstrategien gelten allge-

⁴ Als Synonyme für Strategien werden auch *Techniken*, *Methoden* oder *Taktiken* benutzt (Wild, 2000, S. 7).

mein als domainspezifisch und werden in der aktionalen Phase des selbstregulierten Lernprozesses verortet (Perels et al., 2005; Schmitz, 2001).

Das erfolgreiche Erlernen eines Computerprogramms setzt voraus, dass der Lernende auf ein Repertoire aus unterschiedlichen Explorationsstrategien zurückgreifen, eine adäquate Strategie auswählen und diese erfolgreich anwenden kann (Schmidt-Weigand et al., 2009; Spörer & Brunstein, 2006; Wirth & Leutner, 2008). Die Nutzung verschiedener Explorationsstrategien manifestiert sich in einer unterschiedlichen Herangehensweise an neue Computertechnologien (Carroll & Mack, 1983): Carroll und Carrithers (1984, zit. nach Ghani & Deshpande, 1994, S. 381) beschreiben einerseits den Typ der „reckless explorers who immediately begin to play with the system, frequently with only superficial reading of the manual“ und andererseits die „plodders who will not try anything until assured of the results“. Van der Linden et al. (2001, 2003; siehe auch Dörner, 1980; Dörner & Schaub, 1994; Schaub, 2006; Kang & Yoon, 2008) zufolge lassen sich Explorationsstrategien beim Erlernen eines Computerprogramms als *optimal* oder *suboptimal* klassifizieren. Eine optimale Strategie stellt die systematische Exploration dar. Zu den suboptimalen Explorationsstrategien zählen die Autoren die folgenden: Versuch-und-Irrtum, die rigide Exploration und die Einkapselung in der Informationssuche (vgl. Schaub, 2006). Die wesentlichen Merkmale dieser Explorationsstrategien werden im folgenden Abschnitt (3.1.1) zusammenfassend dargestellt.

3.1.1 Systematische Exploration

Die systematische Explorationsstrategie wird nach van der Linden et al. (2001, S. 190) folgendermaßen definiert: „Systematic exploration is an optimal pattern of behaviour that leads to the highest probability of successful task completion“. Personen, die systematisch explorieren, haben ein starkes Bedürfnis, den Lösungsprozess zu kontrollieren (vgl. Rotter, 1966): Der Problemlösende erkundet das Programm zielgerichtet, er generiert Hypothesen und testet diese in systematischer Weise. Wesentliche Kennzeichen einer systematischen Exploration sind dabei die Planung und Überwachung eigener Handlungen (Buchner, 1995). Dazu trägt die angemessene Beachtung der Programmrückmeldungen bei (Trudel & Payne, 1995). Der systematisch Explorierende kontrolliert darüber den Ausgang des Handlungsprozesses und reflektiert die eingesetzten Methoden. Handlungen werden den Anforderungen entsprechend flexibel angepasst, indem sich der Problemlösende verschiedene Handlungsoptionen offen hält (Funke & Zumbach, 2006). Ineffektive Handlungen werden verworfen und durch neue Handlungen ersetzt (siehe z. B. Miller et al., 1960).

Die systematische Exploration geht mit einer hohen Selbstregulationskompetenz des Problemlösenden einher: Das systematische Vorgehen ist gekennzeichnet durch eine ange-

messene Koordination von hierarchisch aufeinander aufgebauten Problemlöseschritten (vgl. Dörner, 2006; Dörner & Schaub, 1994; Poissant et al., 1994; Sternberg, 1996). Problemlöser, die systematisch explorieren, investieren ein adäquates Ausmaß an Zeit in das Sammeln von Informationen und in das Eingrenzen des Problems (Funke & Zumbach, 2006). Hohe motivationale und emotionale Kontrollfähigkeiten fördern die zielgerichtete Problemlösung insofern, als dass emotionale Ablenkungen von der Zielerreichung (z. B. Frustrationen infolge eines Misserfolges) sowie motivationale Störungen während des Problemlöseprozesses (z. B. ungünstige Attributionen von Fehlschlägen oder eine hohe Misserfolgsbefürchtung) erfolgreich reguliert werden können (vgl. Boekaerts, 1997, 1999; Funke & Zumbach, 2006; Zimmerman, 2000). Damit kann der Großteil kognitiver Ressourcen weiterhin für die Zielerreichung aufgebracht werden.

3.1.2 Versuch-und-Irrtum

Das Explorieren nach Versuch-und-Irrtum ist dadurch gekennzeichnet, dass der Problemlösende „viele verschiedene Handlungsweisen ausprobiert, bis eine davon zufällig funktioniert“ (Mayer, 1992, S. 21). Green und Gilhooly (1990, S. 105; vgl. Carroll & Mack, 1983) sprechen hierbei von unsystematischem „trial and error“, da das Vorgehen nicht durch Hypothesen geleitet ist, die systematisch getestet und bei mangelnder empirischer Evidenz verworfen werden. Stattdessen geht der Problemlösende nicht in einer logischen Weise vor, sondern ist ungeplant und aktionistisch (Kang & Yoon, 2008; Schaub, 2006). Eine Person probiert alles aus, was in der gegebenen Situation möglich ist (Dörner, 2006, S. 242). In diesem „scrambled control mode“ ist „die Wahl des nächsten Handlungsschrittes komplett unvorhersehbar und zufällig“ (Hollnagel, 1993, S. 168). Es wird in jeder Situation das erstbeste Ziel angestrebt (Schaub, 2006). Dörner (1980, S. 93; Dörner, 2006, S. 45) bezeichnet dieses Verhalten als „thematisches Vagabundieren“: Der Problemlösende wechselt schnell und häufig zwischen verschiedenen Themen ohne den Ausgang der Handlung zu Ende zu denken (Schaub, 2006). Dabei wendet sich der Problemlöser völlig neuen Themen zu oder wiederholt bereits ausgeführte Handlungsweisen. Eine Selbstreflexion, eine Ausübung von Kontrolle und eine zielführende Adaptation des eigenen Verhaltens sind nicht erkennbar (Hollnagel, 1993).

Dörner (1980, S. 93f.) interpretiert diese Explorationsstrategie als eine Art „Fluchtverhalten“, mit der Problemlösende ihren wahrgenommenen Kontrollverlust zu kompensieren hoffen und auftretenden Schwierigkeiten schnell ausweichen. Die Ursachen für dieses Verhalten sind vielfältig (van der Linden, 2001, 2003): Das Vorgehen nach *Versuch-und-Irrtum* ist häufig eine Folge mangelnden Vorwissens (Kang & Yoon, 2008). Letzteres erschwert Planungsmöglichkeiten und fördert ein ungünstiges Gefühl des Kontrollverlusts sowie niedriger Selbst-

regulationsfähigkeiten. Darüber hinaus ist die Explorationsstrategie Versuch-und-Irrtum mit negativen motivationalen und emotionalen Zuständen während des Problemlöseprozesses assoziiert (van der Linden et al., 2001).

3.1.3 Rigide Exploration

Die rigide Explorationsstrategie stellt auf den ersten Blick das Gegenteil zum Vorgehen nach Versuch-und-Irrtum dar (siehe Abschnitt 3.1.2; Dörner, 1980, S. 94): Nach van der Linden et al. (2001, S. 190) ist diese suboptimale Explorationsstrategie durch repetitive Handlungssequenzen gekennzeichnet. Eine Person, die den Problemraum rigide exploriert, kann sich nicht von einem bereits beschrittenen Lösungsweg oder der aktuellen Repräsentation des Problems lösen (Schaub, 2006). Obwohl sich die Hinweise häufen, dass der Lösungsweg nicht zielführend ist, hält der Problemlösende an ihm fest (van der Linden et al., 2003). Die rigide Exploration wird deshalb auch als „dogmatische Verschanzung“ („Entrenchment“) bezeichnet (Dörner, 1980, S. 102; Schaub, 2008, S. 449): Der Problemlösende fixiert sich auf einige wenige Handlungssequenzen, die in der Vergangenheit zu einer erfolgreichen Problemlösung beigetragen haben (Sternberg, 1996, S. 364).

Mit der Nutzung der rigiden Explorationsstrategie geht einher, dass Personen immer seltener Entscheidungen treffen, weil sie fürchten, dass diese falsch sein könnten. Auf diese Weise stagniert und begrenzt sich ihr Verhalten immer mehr (Dörner, 1980). Rigide Explorierende neigen dazu, sich bei der Wahl einer Handlungsoption durch saliente Hinweisreize und Gewohnheiten leiten zu lassen (Hollnagel, 1993, S. 169). Personen, die rigide explorieren, sind überdies weniger empfänglich für Rückmeldungen bzw. Fehlermeldungen des Programms (Green & Gilhooly, 1990; Trudel & Payne, 1995, 1996). Sie sind demgegenüber weitgehend immun (Schaub, 2006, S. 460).

Bei näherer Analyse wird deutlich, dass die rigide Explorationsstrategie ebenso wie das Vorgehen nach Versuch-und-Irrtum dazu dient, Unsicherheiten bezüglich der eigenen Lösungskompetenzen zu vermeiden (Dörner, 1980). Begünstigt wird das Auftreten dieser suboptimalen Explorationsstrategie dann, wenn der Problemraum zu groß ist, um ihn vollständig explorieren zu können (van der Linden et al., 2001). Der Problemlösende nutzt nur einen kleinen Ausschnitt der Programmoptionen. Auf diese Weise vermeidet er die Konfrontation mit eigenen Handlungsfehlern und dem Gefühl der Inkompetenz (Dörner & Schaub, 1994).

Schaub (2006) führt die rigide Exploration auf eine Diskrepanz zwischen dem inneren, mentalen Modell, das sich ein Problemlösender von dem System gemacht hat, und den realen Gegebenheiten des Systems, zurück. Wenn der Problemlösende das System erkundet, hält er zu lange an einer ursprünglichen Hypothese zur Beschaffenheit des Systems fest. Er sucht

nach Hinweisen, die sein inneres, mentales Modell der Beschaffenheit des Systems bestätigen, auch wenn diese Annahmen nicht zutreffend sind („Confirmation bias“; vgl. van der Linden et al., 2001; Nisbett & Ross, 1980). Die Planung neuer Handlungswege und die Kontrolle der Effekte eigener Handlungen sind bei der rigiden Exploration eingeschränkt (Schaub, 2006); dies führt zur Wiederholung ineffektiver Handlungsweisen.

3.1.4 Einkapselung in der Informationssuche

Die aktive Suche nach relevanten Informationen stellt eine notwendige Phase eines Problemlöseprozesses dar (Dörner & Schaub, 1994, S. 434ff.; Funke, 2006b, S. 381; Sternberg, 1996, S. 365f.). In dieser Phase sammelt der Problemlösende Informationen über das Problem, die Problemzustände und mögliche Operatoren zur Problemlösung (Dörner & Schaub, 1994). Die Informationssuche produziert einerseits mehr deklaratives Wissen, kann jedoch auch zu mehr Fragen und Unsicherheiten auf Seiten des Problemlösenden führen (van der Linden et al., 2001, S. 190). Dies kann wiederum in steigender Unsicherheit resultieren. Die Motivation, weitere Informationen über Programmfunktionen oder -optionen einzuholen, steigt (Kang & Yoon, 2008). Letztlich kann dies dazu führen, dass eine Person das eigentliche Aufgabenziel aus den Augen verliert: „In the end the gathering of information becomes the main goal“ (van der Linden et al., 2001, S. 190). Das Teilziel der Informationssuche verselbstständigt sich und der Problemlösende investiert all seine Energie in die Bearbeitung dieses Teilziels (Schaub, 2006). Eine Folge hiervon ist, dass Entscheidungen und Handlungen immer weiter aufgeschoben werden. Dies wirkt sich wiederum negativ auf die Problemlöseleistung aus.

Die Einkapselung in der Informationssuche tritt insbesondere dann auf, wenn der Problemlösende mit häufigen Misserfolge, Pannen und Enttäuschungen konfrontiert ist (Schaub, 2006, S. 449). Um das Gefühl eigener Erfolgswahrscheinlichkeit wieder zu erhöhen, wendet sich der Problemlösende leichter zu bewältigenden Aufgaben wie der Informationssuche zu (vgl. Dörner, 1980). Jedoch ist auch die Informationssuche selbst fehleranfällig. Dörner und Schaub (1994, S. 439) sprechen in diesem Zusammenhang von der „Kanalisation der Informationssammlung“ („Channelling errors“): Der Problemlösende sucht und verarbeitet ausschließlich Informationen, die zu seiner vorgefertigten mentalen Repräsentation des Problems passen.

Zur Erfassung der Strategiesystematik werden verschiedene methodische Ansätze herangezogen (vgl. Spörer & Brunstein, 2006; Winne & Perry, 2000). Das folgende Kapitel (3.2) führt zunächst in unterschiedliche Erfassungsmethoden ein, die häufig in Studien zum selbstregulierten Lernen eingesetzt werden. Die Erfassungsmethoden werden im Anschluss daran

mit den bereits beschriebenen Modellen selbstregulierten Lernens und Problemlösens in Verbindung gesetzt (siehe hierzu Kapitel 2). Stärken und Schwächen der ausgewählten Erfassungsmethoden werden schließlich vergleichend dargestellt und im Hinblick auf die Fragestellung dieser Arbeit diskutiert.

3.2 Methoden zur Erfassung der Strategiesystematik

Die Strategiesystematik im Rahmen selbstregulierter Lernprozesse kann auf vielfältige Weise operationalisiert werden. Einen Überblick über unterschiedliche methodische Ansätze zur Erfassung der Strategiesystematik bieten Boekaerts und Corno (2005), Spörer und Brunstein (2006), Winne und Perry (2000) sowie Zimmerman (2008). Zu diesen Ansätzen zählen Fragebogen, Interviews, Lautes Denken („think aloud“), Verhaltensspuren und –beobachtungen. Tabelle 1 liefert eine Übersicht über verschiedene Erfassungsmethoden in ausgewählten Studien des selbstgesteuerten, exploratorischen Erlernens eines Computerprogramms.⁵

Tabelle 1

Operationalisierungsvarianten zur Erfassung der Strategiesystematik in ausgewählten Studien zum Erlernen eines Computerprogramms

Studie	Erfassungsmethode	Umsetzung
Kang & Yoon (2008)	<i>Verhaltensbeobachtung:</i> <i>Verbale Protokolle:</i>	Videoaufzeichnung der Interaktion Interview mit „stimulated recall“
van der Linden et al. (2003)	<i>Verhaltensbeobachtung:</i> <i>Verbale Protokolle:</i>	Videoaufzeichnung der Interaktion Lautes Denken
Debowski et al. (2001)	Verhaltensspuren:	Anzahl und Qualität der Suchbegriffe
Trudel & Payne (1995)	<i>Verhaltensspuren:</i> <i>Verbale Protokolle:</i>	Anzahl der Mausklicks Lautes Denken
Green & Gilhooly (1990)	<i>Verhaltensspuren:</i> <i>Verbale Protokolle:</i>	Anzahl der Mausklicks, Eingabeprotokoll- Lautes Denken
Carroll et al. (1985)	<i>Verhaltensbeobachtung:</i> <i>Verbale Protokolle:</i>	Zeitmessungen für Aufgabenabschnitte Lautes Denken
Vollmeyer & Rheinberg (1999)	<i>Verhaltensspuren:</i> <i>Verbale Protokolle:</i>	Eingabeprotokolle Lautes Denken

Auf welche Weise die Strategiesystematik erfasst wird, ist abhängig vom zugrundeliegenden Konzept selbstregulierten Lernens (Winne & Perry, 2000; Wirth & Leutner, 2008): Wenn das selbstregulierte Lernen als *Fähigkeit* betrachtet wird (z.B. Boekaerts, 1997, 1999), kann die Strategiesystematik als Komponente selbstregulierten Lernens vergleichsweise unab-

⁵ Die aufgeführten Verfahren stellen lediglich eine Auswahl von Erfassungsmethoden dar, die in Studien zum selbstregulierten Lernen häufig eingesetzt wurden (Spörer & Brunstein, 2006; Zimmerman, 2008). Eine weitere Möglichkeit der Erfassung der Strategiesystematik am Computer bieten z.B. „Error Detection Tasks“ (Winne & Perry, 2000; siehe auch Beckwith et al., 2006).

hängig von Situationsmerkmalen erfasst werden (Spörer & Brunstein, 2006). Klassische Erfassungsmethoden im Rahmen dieses Konzepts sind z.B. Selbstberichtsverfahren wie Fragebogen oder Interviews („Offline-Erfassung“; Wirth & Leutner, 2008, S. 104f.). Wird das selbstregulierte Lernen als *Ereignis* konzeptualisiert (z.B. in Prozess- oder Phasenmodellen), so kann die Strategiesystematik über Zeit und Situationen hinweg variieren (Schmitz, 2001; Zimmerman, 2008). Dementsprechend wird sie in Bezug zu einem konkreten Lernereignis bzw. handlungsnah abgefragt („Online-Erfassung“; Wirth & Leutner, 2008, S. 106f.).

Fragebogen. Obwohl in keiner der in Tabelle 1 genannten Studien Fragebogen eingesetzt wurden, sind sie in der Forschungspraxis zum selbstregulierten Lernen die am häufigsten eingesetzten Verfahren (Jamieson-Noel & Winne, 2003; Muis, Winne & Jammieson-Neal, 2007; Spörer & Brunstein, 2006). Fragebogen erfassen die vom Lernenden selbst berichtete Nutzung von mehr oder weniger systematischen kognitiven (und metakognitiven) Strategien. Ziel der Befragung ist es, über die Selbstberichte des Lernenden Hinweise auf dessen kognitive Prozesse zu erhalten. Diese liegen der Strategiesystematik zugrunde und können durch den Forscher nicht direkt beobachtet werden (Winne & Perry, 2000). Dem Befragten werden in der Regel verschiedene Aussagen zur Nutzung unterschiedlicher Strategien vorgegeben. Die Lernenden sollen einstufen, wie stark die Aussage auf sie zutrifft (Wirth & Leutner, 2008).

Interviews. Interviews zählen zu den Verbalisierungstechniken (Winne & Perry, 2000): Die Befragten geben im Gespräch darüber Auskunft, wie sie in einer konkreten Lernsituation denken und handeln (Spörer & Brunstein, 2006). Interviews können in vielfältiger Art und Weise durchgeführt werden. Häufig werden strukturierte Interviews eingesetzt (Winne & Perry, 2000). Bei dieser Variante wird allen Befragten eine vorgegebene Reihe von Fragen zu einem eingegrenzten Themenbereich gestellt. Form und Inhalt der Fragen sind dabei für jeden Befragten gleich (Amelang & Zielinski, 2002).

Beim Interview schildert der Befragte auf der Basis seiner Erinnerungen, welche Lernstrategien er in einer spezifischen Lernsituation angewendet hat (retrospektive Erfassung). Die Gedächtnisleistung des Befragten kann dabei durch Methoden wie „stimulated recall“ (Lyle, 2003) unterstützt werden. Kurze Sequenzen der aufgezeichneten Lernsituation werden dem Befragten während des Interviews als Erinnerungshilfe vorgespielt. Der Befragte wird dann aufgefordert, anzugeben, was er in der konkreten Situation getan und gedacht hat. Der Befragte kann auch zu seinem Verhalten in einer zukünftigen, hypothetischen Lernsituation interviewt werden (prospektive Erfassung; Spörer & Brunstein, 2006).

Lautes Denken. Die Selbstauskunft über die eigenen Lernhandlungen wird als Methode des „Lauten Denkens“ bezeichnet („think aloud“; Ericson & Simon, 1980; Winne & Perry, 2000). Der Lernende wird während des Lernprozesses aufgefordert, Denkinhalte konti-

nuerlich zu verbalisieren und die eigene Strategienutzung zu kommentieren. Hierzu kann die Lernaktivität an markanten Abschnitten des Lernprozesses (z.B. nach der Beendigung eines Arbeitsschrittes) kurz unterbrochen werden (Spörer & Brunstein, 2006). Die aufgezeichneten Äußerungen des Lernenden werden im Anschluss kodiert. Steigt der Schwierigkeitsgrad der Aufgabe, neigen Lernende dazu, das Verbalisieren zu unterbrechen (Vollmeyer & Rheinberg, 1999). Um Datenausfall zu vermeiden, raten Ericson und Simon (1993) darum, die Befragten intensiv in die Technik des lauten Denkens einzuarbeiten und sie im Lernprozess an die Verbalisierungen zu erinnern.

Verhaltensspuren. Eine Alternative zu Selbstberichten (Fragebogen, Interview, Lautes Denken) stellt die Erfassung der Strategiesystematik über Verhaltensspuren bei einem konkreten Lernereignis dar (Zimmerman, 2008). Nach Winne und Perry (2000, S. 551) bezeichnen Verhaltensspuren („traces“) beobachtbare Indikatoren, die Rückschlüsse auf die kognitiven Prozesse der Lernenden zulassen. Die Verhaltensindikatoren werden theoretisch abgeleitet. Verhaltensspuren werden als direkter Ausdruck der Überzeugungen eines Lernenden in der Lernsituation betrachtet (Jamieson-Neal & Winne, 2003). Zu den Verhaltensspuren zählen zum Beispiel Tonspuren, Ereignis- und Mausclickprotokolle („log files“), Aufzeichnungen der Bildschirmoberfläche wie Navigationsspuren (Wirth & Leutner, 2008; Zimmerman, 2008).

Verhaltensbeobachtung. Wissenschaftliche, d. h. systematische Beobachtungen sind dadurch gekennzeichnet, dass sie absichtlich und geplant durchgeführt werden und auf ein bestimmtes Forschungsziel ausgerichtet sind (Bortz & Döring, 2003). Beobachtungsmethoden bieten gegenüber Selbstberichten die Möglichkeit, ein Verhalten detailliert, in seinem zeitlichen Verlauf und im jeweiligen Kontext zu erfassen (Boekaerts & Corno, 2005; Greve & Wentura, 1997). Auf diese Weise sind Handlungssequenzen nachvollziehbar (Prümper, 1991, S. 47). Häufig werden Videoaufzeichnungen (in Echtzeit) angefertigt. Bei diesen verhaltensorientierten Erhebungsverfahren findet keine direkte Interaktion zwischen Untersuchungsleiter und Untersuchungsteilnehmer statt⁶: „Der Untersucher sucht sich vielmehr Hinweise aus vorhandenem Material, die Aufschlüsse über bestimmte Verhaltensweisen der Untersuchten geben (Spurensuche)“ (Rost, 2007, S. 50).

Vergleichende Diskussion der Erfassungsmethoden. Validität und Reaktivität der einzelnen Erfassungsmethoden werden in der Forschung kontrovers diskutiert (Artelt, 2000; Ericson & Simon, 1980, 1993; Jamieson-Noel & Winne, 2003; Nisbett & Wilson, 1977; Rost, 2007; Schooler, Ohlsson & Brooks, 1993).

⁶ Eine Ausnahme stellt die teilnehmende Beobachtung dar (vgl. Diekmann, 2007; Schnell, Hill & Esser, 2008).

Fragebogen zeichnen sich insgesamt durch eine hohe Ökonomie und Reliabilität aus (Spörer & Brunstein, 2006). Jedoch werden Konstrukt- und Kriteriumsvalidität von Fragebogenverfahren kritisch beurteilt: Forschungsergebnisse zeigen, dass die mittels Fragebogen erfassten Lernstrategien nur in einem mittleren Zusammenhang mit der tatsächlichen Nutzung dieser Strategien stehen (Artelt, 2000). Des Weiteren konnten die mittels Fragebogen erfassten Lernstrategien nur einen geringen Varianzanteil in einer konkreten Leistungssituation aufklären (Jamieson-Neal & Winne, 2003; Schiefele, Streblow, Ermgassen & Moschner, 2003). Winne und Perry (2000) führen dies darauf zurück, dass Fragebogen die Strategiesystematik in der Regel als Personenmerkmal und nicht als situationsabhängiges Handlungsmerkmal erfassen. Darüber hinaus sind viele Items zur Strategiesystematik unzureichend aufgabenspezifisch bzw. wenig handlungsnah (Spörer & Brunstein, 2006). Um die Items beantworten zu können, muss der Lernende zudem sein eigenes Lernverhalten reflektieren können und über ein ausreichendes Strategiewissen verfügen (Artelt, 2000).

Im Gegensatz zu Fragebogen wird den Befragten im *Interview* häufig eine größere Freiheit bezüglich ihres eigenen Antwortverhaltens zugestanden. Spörer und Brunstein (2006; vgl. auch Zimmerman, 2008) zitieren mehrere Studien, in denen mittels Interviews Unterschiede zwischen Novizen und Experten in der Strategiesystematik aufgezeigt werden konnten. Dies wird als Beleg für die Validität der Interviews gewertet. Im Gegensatz zur Methode des Lauten Denkens werden beim Interview mit „*Stimulated Recall*“ Unterbrechungen und damit potenzielle Störungen der kognitiven Prozesse während des Lernprozesses vermieden. Die videounterstützte Befragung und die Kodierung im Anschluss an die Lernsituation stellt jedoch ein aufwändiges Verfahren dar (Lyle, 2003). Es eignet sich daher nur für kurze Phasen der Aufgabenbearbeitung (Spörer & Brunstein, 2006).

Die *Methode des Lauten Denkens* wurde besonders häufig eingesetzt, um die Strategiesystematik bei der Exploration eines Computerprogramms nachvollziehen zu können (vgl. Tabelle 1). Spörer und Brunstein (2006, S. 154) sehen den Vorteil der Methode des Lauten Denkens darin, dass sie einen unmittelbaren „Einblick in gedankliche Abläufe vermittelt, die mutmaßlich handlungssteuernd wirken“. Carroll et al. (1985) nahmen an, dass Lernende, die „laut denken“, beim Explorieren eines Computerprogramms vorsichtiger seien als in Situationen ohne diese Intervention. Das Kommentieren koste darüber hinaus Zeit, was den Lernenden benachteilige und sich negativ auf seine Lernleistung auswirken könne. Die empirischen Befunde zu den Auswirkungen des Lauten Denkens sind jedoch uneinheitlich. Studienteilnehmer, die zwischen einzelnen Problemlöseabschnitten unterbrochen wurden, um über ihre Lernstrategien Auskunft zu geben, erreichten niedrigere Leistungen als Studienteilnehmer, die dies nicht zur Auflage hatten (Schooler et al., 1993). Der abträgliche Effekt der Verbalisierung

trat jedoch nicht bei allen Problemlöseaufgaben gleichermaßen auf. In einer Studie von Trudel und Payne (1996) führten Verbalisierungen während des Lernprozesses vielmehr zu besseren Lernleistungen. Ähnliche Befunde berichten van Oostendorp und de Mul (1999), wonach das Laute Denken die Planung und Überwachung eigener Explorationshandlungen unter bestimmten Bedingungen positiv beeinflusste. In einer anderen Studie von Leow und Morgan-Short (2004) hatte die Methode weder günstige noch ungünstige Auswirkungen auf die Leseleistung der Studienteilnehmer. Vollmeyer und Rheinberg (1999) geben zu bedenken, dass das Laute Denken nicht nur für die Lernleistung, sondern auch für die Lernmotivation abträglich sein könne.

Hinsichtlich ihrer Nachteile weisen Fragebogen, Interviews, die teilnehmende Beobachtung und das Laute Denken einige Gemeinsamkeiten auf. Diese Verfahren werden allgemein den reaktiven Messverfahren zugeordnet (Rost, 2007). Reaktive Messverfahren haben den Nachteil, dass das Verhalten und Erleben, welches erfasst werden soll, durch den Einsatz dieser Erfassungsmethoden verändert werden kann. Erwartungshaltungen des Untersuchungsleiters, Reaktionen auf das Test- und Fragebogenmaterial selbst sowie bewusste und unbewusste Verfälschungstendenzen (z.B. Einfluss „impliziter Kausaltheorien“; vgl. Spörer & Brunstein, 2006, S. 152) können die Gültigkeit der Untersuchungsergebnisse beeinflussen (Bungard & Lück, 1974; Diekmann, 2007). Darüber hinaus darf das zeitliche Intervall zwischen Lernhandlungen und Selbstauskunft (Interview, *Stimulated recall*, Fragebogen) nicht zu groß sein (Lyle, 2003). Eine weitere Schwäche von Fragebogen, Interview und Lautem Denken ist, dass diese Verfahren oft hohe Anforderungen an den Befragten stellen. So müssen Befragte über Prozesse Auskunft geben, die automatisiert und damit schlecht verbalisierbar sind (Artelt, 2000; Nisbett & Wilson, 1977; Schooler et al., 1993).

Als Alternative zu (reaktiven) Erfassungsmethoden bieten sich nonreaktive Erfassungsmethoden (*Verhaltensspuren, nicht-teilnehmende Verhaltensbeobachtung*) an (Bungard & Lück, 1974; Hadwin, Nesbit, Code, Jamieson-Noel & Winne, 2007; Jamieson-Neal & Winne, 2003). Im Gegensatz zu den (reaktiven) Selbstberichtsverfahren erlaubt eine verhaltensnahe Erfassung über Verhaltensspuren (z.B. Aufzeichnungen der Bildschirmoberfläche, Mausclicksprotokolle) die tatsächliche Nutzung spezifischer Strategien nachzuvollziehen (Artelt, 1999; Hadwin et al., 2007; Zimmerman, 2008). Bei der Erfassung über Verhaltensspuren und bei anderen nonreaktiven Messverfahren wird der Anspruch erhoben, dass das zu beobachtende Verhalten durch die Aufzeichnung nicht verändert wird (Bortz & Döring, 2003; Bungard & Lück, 1974; Diekmann, 2007). Empirische Studien ergaben, dass die über Verhaltensspuren erfassten Strategien in Beziehung zum selbstregulierten Lernverhalten und zur Leistung standen (Artelt, 1999; Hadwin et al., 2007; Jamieson-Neal & Winne, 2003; Trudel & Payne, 1995; van

der Linden et al., 2001, 2003). Damit unterstützen diese Befunde die Validität dieser Operationalisierungsvariante.

3.3 Empirische Befunde und Zusammenfassung

Die Ergebnisse zahlreicher Forschungsarbeiten belegen, dass Unterschiede in der Strategiesystematik, z.B. durch die Nutzung suboptimaler und optimaler Strategien, zu interindividuellen Differenzen in der Lern- und Problemlöseleistung führen (z.B. Dörner, 1980; Dörner & Schaub, 1994; Kang & Yoon, 2008; Vollmeyer & Rheinberg, 1999, 2000; Vollmeyer et al., 1996). Interindividuelle Differenzen in der Nutzung von Explorationsstrategien lassen sich auf verschiedene Ursachen zurückführen. Hierzu zählen Unterschiede in den Lernbedingungen und in individuellen Persönlichkeitsmerkmalen der Lernenden. Im folgenden Abschnitt werden Resultate zentraler Studien vorgestellt, in welchen die Strategiesystematik sowie weitere kognitive Aspekte beim Erlernen eines Computerprogramms näher untersucht wurden. In den Studien wurden Probleme der realen Welt (*real world problems*) ausgewählt, um eine hohe ökologische Validität der Untersuchungsbedingungen zu gewährleisten (Carroll & Mack, 1983; Keith et al., 2010; Vollmeyer & Rheinberg, 1999, 2000, 2006): Die Teilnehmer hatten z. B. die Aufgabe, einen Brief mittels eines ihnen unbekanntem Textverarbeitungsprogramms zu verfassen oder eine Statistikaufgabe mittels einer neuen Statistiksoftware zu lösen. Dabei wurden Vorkenntnisse der Teilnehmer (Computererfahrung, Wissen) entweder in den Studien miterhoben oder deren Einfluss wurde über das Untersuchungsdesign konstant gehalten (z.B. Debowski et al., 2001; Green & Gilhooly, 1990; Trudel & Payne, 1995).

Bereits Anfang der 1980er Jahre untersuchten Carroll et al. (1985), wie sich das aktive, exploratorische Erlernen eines unbekanntem Textverarbeitungsprogramms auf Strategiewahl und Lernleistung auswirkt. In der Studie erstellten die Teilnehmer einen Brief mittels eines neuen Textverarbeitungsprogramms, den sie auch editierten. Bei einem Teil der Teilnehmer wurde ein aktives, exploratorisches Lernen angeregt, in dem Kärtchen mit Erklärungen und Hilfestellungen verfügbar waren, auf die diese Teilnehmer nach Bedarf zurückgreifen konnten. Die zweite Gruppe löste die Aufgabe mit Hilfe des Selbststudienmaterials des kommerziellen Anbieters.

Die Auswertung der think aloud Protokolle und der Videoaufzeichnungen der Trainingssitzungen ergaben, dass das aktive, exploratorische Lernen anhand konkreter Problemstellungen allgemein einen günstigen Effekt auf den Lernerfolg hatte: Die Teilnehmer benötigten unter diesen experimentellen Bedingungen weniger Zeit beim Erlernen wesentlicher Programmfunktionen und bei der Anwendung des erworbenen Wissens. Sie wurden außerdem schneller auf Programmfehler aufmerksam und experimentierten gezielt mit dem Programm.

Die Teilnehmer, die mit dem Handbuch zur Einführung in das Programm lernten, verbrachten dagegen viel Zeit mit dem Sammeln von Informationen. Die übermäßige Informationssuche wurde bereits als unsystematische Explorationsstrategie klassifiziert (vgl. Dörner, 1980; Kang & Yoon, 2008; van der Linden et al., 2001, 2003); sie resultierte in der Studie erwartungsgemäß in niedrigeren Lernleistungen.

Keith et al. (2010) kamen in ihren Studien zu ähnlichen Ergebnissen: Sie verglichen den Effekt zweier Trainingsmethoden beim Erlernen und Anwenden eines Textverarbeitungs- bzw. eines Präsentationsprogramms. Insbesondere wenn neuartige Probleme gelöst werden sollten („adaptive transfer“), führte das aktive, selbstregulierte Explorieren zu besseren Leistungen. Das stärker external regulierte Training, bei dem die richtigen Schrittfolgen der Problemlösung direkt vermittelt wurden, war dem aktiven, selbstregulierten Explorieren unterlegen. Diesen Effekt erklären die Autoren damit, dass beim aktiven Explorieren Selbstregulationskompetenzen gefordert sind und geübt werden. Dies begünstigt die zukünftige Problemlösung. Das aktive, selbstregulierte Explorieren reduzierte den Einfluss kognitiver Fähigkeiten sowie der Motivation auf die Transferleistung, so dass das aktive Explorieren auch für Personen mit niedrigerem Vorwissen entgegen kritischer Einschätzungen anderer Autoren (Kirschner et al., 2006) geeignet erscheint.

Dass die angeleitete Exploration, wie sie z.B. von Carroll et al. (1985) verwendet wurde, auch gegenüber der unstrukturierten Exploration überlegen ist, konnten Debowski et al. (2001; siehe auch Wood, Kakebeeke, Debowski & Frese, 2000) aufzeigen. Sie untersuchten den Einfluss von *guided exploration* und *enactive exploration* auf die Strategiesystematik und die Leistung bei der elektronischen Suche in Datenbanken. Im Gegensatz zur *guided exploration* erkunden Personen bei der *enactive exploration* das Computerprogramm ohne Vorgaben und Anleitung (Greif, 1990). Das Erkunden wird allein durch die Interessen des Computernutzers und aufkommende Probleme gesteuert. Effektive Strategien können durch Versuch-und-Irrtum, das systematische Vorgehen, die Nutzung von Hilfefunktionen und durch das Lernen aus Fehlern entdeckt werden (Keith & Frese, 2008). Ob diese unstrukturierte Explorationsmethode jedoch zu einer Verbesserung der Strategiesystematik führt, ist wesentlich von den Rückmeldungen des Programms an den Nutzer abhängig. Sind diese nicht gegeben bzw. können diese nicht korrekt interpretiert werden, wirkt sich dies negativ auf die Strategiesystematik aus (Bandura, 1997): Die Computernutzer verbleiben bei einer suboptimalen Strategiesystematik, experimentieren seltener mit verschiedenen Strategien und wiederholen stattdessen häufiger ineffektive Handlungen. Die Konsequenz ist eine unzureichende Leistung bei der elektronischen Suche, d.h. eine geringere Anzahl relevanter Suchergebnisse. Die Untersuchung von

Debowski et al. (2001) liefert empirische Belege für die erwarteten Zusammenhänge zwischen Explorationsmethode, Strategiesystematik und Suchleistung in der Datenbank.

Auch Trudel und Payne (1995, S. 308) machen darauf aufmerksam, dass bestimmte Explorationsbedingungen den Einsatz angemessener Explorationsstrategien hemmen können: „Exploratory learning can lead users into acute dead-ends, and even chronic difficulties“. Als Beispiel führen die Autoren an, dass Lernende ein Teilziel beim Erkunden eines Computerprogramms erreichen, aber später nicht mehr aus dem Gedächtnis abrufen können, mittels welcher Handlungsschritte sie es erreicht haben. Die Autoren führen dieses Unvermögen auf den simultanen Erwerb deklarativen und prozeduralen Wissens zurück (vgl. Anderson, 2007; Sternberg, 1996). Ein erfolgreiches Lernen durch Exploration setze deshalb voraus, dass der Lernende seine Ziele und die angewandten Strategien bewusst reflektiert.

Trudel und Payne (1995; siehe auch O'Hara & Payne, 1998) nahmen an, dass die Reflexion der eigenen Handlungen erhöht werden könne, wenn (unüberlegte) Handlungen im Computersystem mit Kosten verbunden seien. Sie überprüften, ob eine Begrenzung der Anzahl der von einem Computernutzer zu tätigenen Mausklicks während der Interaktion mit einer Digitaluhr die Reflexionsprozesse des Lernenden fördern kann. Die Ergebnisse ihrer Studie unterstützen die Annahme, dass die Explorationsbedingungen einen Einfluss auf die Strategiesystematik haben. Problemlösende, denen nur eine begrenzte Anzahl von Mausklicks für die Exploration zur Verfügung stand, gingen effizienter vor und lösten mehr Aufgaben. Sie entdeckten früher relevante Funktionen der Menüführung. Hinsichtlich ihres erworbenen deklarativen Wissens über das Interface und ihres prozeduralen Wissens waren diese Personen Teilnehmern, die ohne Mausklickbegrenzung explorierten, überlegen. Personen, die ohne Anleitung und Mausklick-Begrenzung explorierten, wechselten häufiger und schneller zwischen verschiedenen Funktionen der Digitaluhr. Dies wurde von den Autoren als „Vagabundieren“ interpretiert (vgl. Dörner, 1980). Nach van der Linden et al. (2001, 2003) ist dieses Verhalten Ausdruck der Versuch-und-Irrtum-Strategie. Auch wiederholten diese Teilnehmer häufiger ineffektive Handlungen. Dieses Verhalten wurde bereits oben als rigide Exploration interpretiert (vgl. van der Linden et al., 2001, 2003).

Interindividuelle Unterschiede in der Nutzung verschiedener Lernstrategien, wurden von Green und Gilhooly (1990) an Novizen beim exploratorischen Erlernen eines Statistikprogramms (MINITAB) erforscht. Den Lernenden wurden Beispielaufgaben vorgegeben, die sie in mehreren Sitzungen lösen sollten. Hierzu zählten die Dateneingabe, die Berechnung Deskriptiver Statistiken, die Durchführung von *t*-Tests, Varianzanalysen und Signifikanztests. Green und Gilhooly (1990) konnten anhand von *think aloud*- und Verhaltensprotokollen zeigen, dass sich Lernende darin unterscheiden, in welchem Ausmaß sie zwei Arten von Lern-

strategien nutzen (vgl. Carroll & Mack, 1983): „Learning by thinking“ und „Learning by doing“. *Learning by thinking* bezieht sich auf Prozesse der Reflexion über die Interaktion mit dem Computer. Unter *Learning by doing* werden Handlungen des Ausprobierens und der aktiven Interaktion mit dem Computer subsumiert. Signifikante Differenzen zwischen erfolgreichen und weniger erfolgreichen Lernenden zeigten sich hauptsächlich in der Nutzung des *learning by doing*. Erfolgreiche Lernende explorierten strukturiert, indem sie gezielt mit Bedingungen experimentierten. Dagegen explorierten weniger erfolgreiche Lernende stärker nach Versuch-und-Irrtum und hielten sich unnötig lange mit Details der Aufgabe auf. Diese Personen brachten häufiger ihre Ratlosigkeit über die nächsten erforderlichen Schritte zum Ausdruck. Sie wiederholten öfter bereits genutzte Prozeduren und schenkten Fehlermeldungen des Programms eine geringere Beachtung. Dieses Verhalten ähnelt der von van der Linden et al. (2001, 2003) sowie Kang und Yoon (2008) beschriebenen rigiden Explorationsstrategie.

Kang und Yoon (2008) untersuchten den Einfluss des Alters und des Vorwissens auf die Nutzung unterschiedlicher Explorationsstrategien. Die Teilnehmer ihrer Studie hatten die Aufgabe, zwei ihnen unbekannte, elektronische Musikgeräte zu bedienen. Die Studie ergab, dass ältere Personen tendenziell weniger systematisch explorierten: Im Vergleich zu Jüngeren wiesen sie statistisch signifikant höhere Anteile rigider Exploration auf. Sie benötigten eine größere Anzahl an Interaktionen mit dem Gerät und begingen mehr Bedienungsfehler. Dieser Alterseffekt zeigte sich auch unter Kontrolle des Vorwissens. Das Vorwissen hatte zudem einen negativen Einfluss auf die Häufigkeit der Versuch-und-Irrtum-Strategie und der übermäßigen Informationssuche. Insbesondere ältere Teilnehmer mit wenig Vorwissen berichteten eine höhere kognitive Belastung während der Interaktion mit dem elektronischen Gerät (vgl. Tuovinen & Sweller, 1999).

Ähnliche Befunde liefert eine experimentelle Studie von van der Linden et al. (2003) zur Beziehung zwischen einer hohen kognitiven Beanspruchung und der Nutzung unsystematischer Strategien. Die Autoren beobachteten den Effekt mentaler Ermüdung infolge kognitiver Beanspruchung auf das Explorationsverhalten beim Erlernen eines Tabellenkalkulationsprogramms. Erwartungsgemäß explorierten Personen mit mentaler Ermüdung das Computerprogramm weniger häufig systematisch und stärker rigide. Das Ergebnis führen die Autoren auf mangelnde Reflexionsprozesse zurück (vgl. Trudel & Payne, 1995). Entgegen der Erwartungen beeinflusste die mentale Ermüdung nicht die Häufigkeit des Versuch-und-Irrtum-Verhaltens. Auch hinsichtlich der Leistung bei der Aufgabenbearbeitung mit dem Tabellenkalkulationsprogramm zeigten sich keine statistisch signifikanten Unterschiede zur Kontrollgruppe. Van der Linden et al. (2003) begründen dies mit der kurzen Bearbeitungszeit von 15 Minuten, in der Strategieunterschiede nicht leistungswirksam werden konnten. In ihrer Korre-

lationsstudie konnten die Autoren jedoch die positive Beziehung systematischer Exploration und Lernleistung empirisch belegen. Unsystematische Explorationsstrategien (rigide Exploration, Einkapselung in der Informationssuche) gingen hypothesenkonform mit einer niedrigeren Leistung einher. Explorieren nach Versuch-und-Irrtum führte dagegen nicht zu einer Leistungsminderung, hatte jedoch insgesamt eine höhere Fehlerrate zur Folge.

Zusammenfassung. Die Befunde der vorgestellten Studien verdeutlichen, dass das selbstregulierte Explorieren eines Computerprogramms (mit Hilfestellungen) zu einer verbesserten Strategiesystematik führen kann. Die Vorgabe von Aufgaben mit zunehmender Schwierigkeit führt insbesondere bei Novizen zu einer erhöhten Strategiequalität. Jedoch zeigte sich, dass der Effekt von weiteren Bedingungen abhängig ist (z.B. den Fehlerrückmeldungen des Programms). In den Studien konnten die in Kapitel 3.1. theoretisch beschriebenen Strategien empirisch voneinander abgegrenzt werden (siehe Kapitel 3.1.). Die höhere Qualität der eingesetzten Problemlöse- und Explorationsstrategien geht dabei mit einer höheren Lernleistung einher. Insbesondere die systematische Exploration einer unbekanntem Computersoftware führt zu einem Wissenszuwachs, wohingegen unsystematische Explorationsstrategien (Versuch-und-Irrtum, Rigide Exploration, übermäßige Informationssuche) erwartungsgemäß mit einer geringeren Lernleistung assoziiert sind.

4. Motivationale Aspekte des Problemlöseprozesses

Neben kognitiven Aspekten wie der Strategiesystematik wird auch der Motivation eine besondere Bedeutung für das selbstregulierte Lernen (SRL) zugeschrieben: „Without motivation, except for reflexive behavior like the eye blink reflex, there is no behavior, including SRL“ (Winne & Hadwin, 2008, S. 297). Auch Zimmerman und Schunk (2008) betonen die Relevanz einer hohen Motivation, um die Aufmerksamkeit des Lernenden auf den Lernprozess zu richten. Dies sei insbesondere bei repetitiven Übungsaufgaben von Bedeutung. Ein Mindestmaß an Lernmotivation ist demzufolge für das zielgerichtete Lernen und Problemlösen unabdingbar (Engeser, Rheinberg, Vollmeyer & Bischoff, 2005; Wolters, 2003). In der pädagogisch-psychologischen Forschung dienen motivationspsychologische Variablen dazu, das Lern- und Leistungsverhalten von Personen zu erklären und vorherzusagen. Motivation ist dabei ein hypothetisches Konstrukt, das nicht direkt beobachtet werden kann, sondern etwas gedanklich Konstruiertes beschreibt (Vollmeyer, 2005). Darüber hinaus umfasst der Motivationsbegriff viele verschiedene Teilprozesse und Phänomene. Allgemein wird Motivation nach Rheinberg (2004, S. 15) als die aktivierende „Ausrichtung des momentanen Lebensvollzugs auf einen positiv bewerteten Zielzustand“ beschrieben. Der in Rheinbergs Definition angesprochene Zielzustand bezieht sich dabei auf erwünschte Handlungsfolgen („Folgenanreize“), auf tätigkeitszentrierte Anreize sowie auf gegenstandsbezogene Anreize (Schiefele & Streblow, 2005). Lernmotivation kann nach Schiefele (2008) definiert werden als „der in einer konkreten Situation aktuell auftretende Zustand ..., über einen gegebenen Lerngegenstand Wissen erwerben zu wollen“ (S. 38).

Motivationspsychologische Theorien unterscheiden sich unter anderem darin, welchen zentralen Einfluss sie Personenmerkmalen bzw. Anregungsbedingungen der Situation zusprechen. Aus der Sicht der Instinkt- und Triebtheorien ist das Verhalten stärker durch genetisch begründete Verhaltensvorlieben bestimmt (vgl. Freud, 1915/1952; Lorenz, 1963). Andere führen das Verhalten auf die Ausprägung unterschiedlicher Motive zurück (z.B. Leistungs-, Macht- und Anschlussmotiv; Heckhausen, 1963; McClelland, Atkinson, Clark & Lowell, 1953). Motive sind als „zeitlich stabile Personenmerkmale konzipiert und stellen eine Neigung dar, bestimmte Themen oder Gegenstände positiv oder negativ zu bewerten“ (Vollmeyer, 2005, S. 10). Schiefele (2008) bezeichnet diese Art der Motivation auch als „habituelle Motivation“.

Demgegenüber trennten Vertreter der klassischen Motivationspsychologie (z.B. Atkinson, 1957; Heckhausen, 1963; Lewin, 1926) zwischen „Motiven als überdauerndem Personenmerkmal und der je aktuellen Motivation“ (Rheinberg, 2004, S. 70). Die aktuelle Motivati-

on ist eine „situationsspezifische, aktuelle Verhaltensbereitschaft“ (Schiefele, 2008, S. 38) und resultiert wiederum aus der Interaktion von Personen- und Situationsfaktoren (siehe auch Rheinberg, Vollmeyer & Burns, 2001). Der interaktionistische Ansatz ist heute in der Motivationspsychologie nach wie vor vorherrschend (Rheinberg, 2004). Dies gilt auch für die Erklärung leistungsmotivierten Verhaltens. Als leistungsmotiviert gilt ein Verhalten, „wenn es auf die Selbstbewertung eigener Tüchtigkeit zielt - und zwar in Auseinandersetzung mit einem Gütemaßstab, den es zu erreichen oder zu übertreffen gilt“ (Rheinberg, 2004, S. 60).

Um leistungsmotiviertes Verhalten erklären zu können, werden in der vorliegenden Arbeit unterschiedliche Erklärungsvariablen (kognitive, motivationale) herangezogen. Ziel ist es, insbesondere den Anteil motivationaler Variablen beim Erlernen eines Computerprogramms und deren Einfluss auf Leistungsunterschiede genauer zu untersuchen. Stabile Personenmerkmale werden in der vorliegenden Arbeit nur in begrenztem Maße berücksichtigt (vgl. Atkinson, 1957; McClelland et al., 1953). Stattdessen liegt der Forschungsschwerpunkt der Arbeit auf der situativen Ausprägung der Motivation: Die aktuelle Motivation, welche kurz vor oder während des Lernprozesses gemessen wird und situativ variieren kann, steht im Vordergrund der Analyse. Anhand der Theorie von Lewin (1926) wird in Kapitel 4.1 zunächst der interaktionistische Ansatz der Motivationspsychologie zur Entstehung aktueller Motivation erläutert. Vollmeyer und Rheinberg (2006) zufolge wird der Einfluss der aktuellen Motivation vor Lernbeginn auf die Leistung über unterschiedliche Variablen vermittelt. Zu diesen Variablen zählt auch das von Csikszentmihalyi (1975, 2000) eingeführte Konstrukt des „Flow“-Erlebens. Dieses wird in Abschnitt 4.2. beschrieben. Über das Flow-Erleben hinaus werden im kognitiv-motivationalen Prozessmodell von Vollmeyer und Rheinberg (1998, 2000, 2006) weitere Variablen spezifiziert, welche die Leistung beeinflussen. In diesem Ansatz, der in Kapitel 4.3 dargestellt ist, wird der aktuellen Motivation während der Aufgabenbearbeitung und deren handlungsnahen Messung eine besondere Bedeutung zugeschrieben.

4.1 Motivation als Ergebnis der Interaktion von Person und Umwelt

Ob in einer konkreten Situation Lernmotivation entsteht und sich im Lernverhalten selbst bemerkbar macht, ist sowohl von den Merkmalen des Lernenden selbst als auch von den situativen Gegebenheiten abhängig. Die Wechselwirkung dieser beiden Komponenten bildet die Kernannahme der interaktionistischen Motivationstheorien. Als Wegbereiter dieser interaktionistischen Sichtweise gilt Kurt Lewin (1926, 1936).

In seiner Feldtheorie nahm Lewin an, dass Personen sich in einem Kräftefeld befinden, in dem situative Kräfte auf die Personen einwirken: „Diese Kräfte entstammen sowohl der ‚äußeren‘ Situation (der Umgebung) als auch der ‚inneren‘ Situation (der Person)“ (Beck-

mann & Heckhausen, 2006, S. 107). Unterschiedliche Kräftefelder stehen dabei für verschiedene Handlungsziele und Bedürfnisse einer Person. „Jedes Bedürfnis bildet nun innerhalb der Person ein *gespanntes System*, das nach Entspannung verlangt“, schreibt Rheinberg (2004, S. 44). Die Spannung entsteht durch äußere oder innere Gegebenheiten, die ein bestimmtes Bedürfnis aktivieren. Dem System ist immanent, dass ein Spannungsausgleich angestrebt wird. Aus der Spannung resultiert demzufolge ein gerichteter Druck. Dieser Druck sorgt dafür, dass die Person Handlungen ergreift, die zur Befriedigung des Bedürfnisses führen.

Die Bedürfnisse werden jedoch nicht von selbst handlungswirksam, sondern sie bedürfen bestimmter Anreize der Umwelt, die einen Aufforderungscharakter („Valenz“) besitzen (Lewin, 1926, S. 350). Dabei ist die subjektive Wahrnehmung der Situation durch die einzelne Person entscheidend (Beckmann & Heckhausen, 2006, S. 107). Besteht ein inneres Bedürfnis der Person und damit eine Spannung, wird die Person leichter auf solche Umweltanreize aufmerksam, die Handlungsmöglichkeiten bieten, um das Bedürfnis zu befriedigen. Die Umweltanreize treten in der Wahrnehmung besonders hervor, haben eine hohe positive Valenz und lösen damit ein entsprechendes Aufsuchungsverhalten aus (Beckmann & Heckhausen, 2006; Rheinberg, 2004). Eine negative Valenz der Umweltanreize dagegen begünstigt Vermeidungsverhalten. Wäre von vornherein keine Empfänglichkeit für diese Umweltanreize in der handelnden Person angelegt, würde die Spannung im System nicht erhöht werden. Ein Handeln würde genauso ausbleiben wie dies bei Abwesenheit des Umweltanreizes der Fall wäre. Die Motivationsstärke ist nach Lewin damit eine Funktion der Valenz (Beckmann & Heckhausen, 2006).

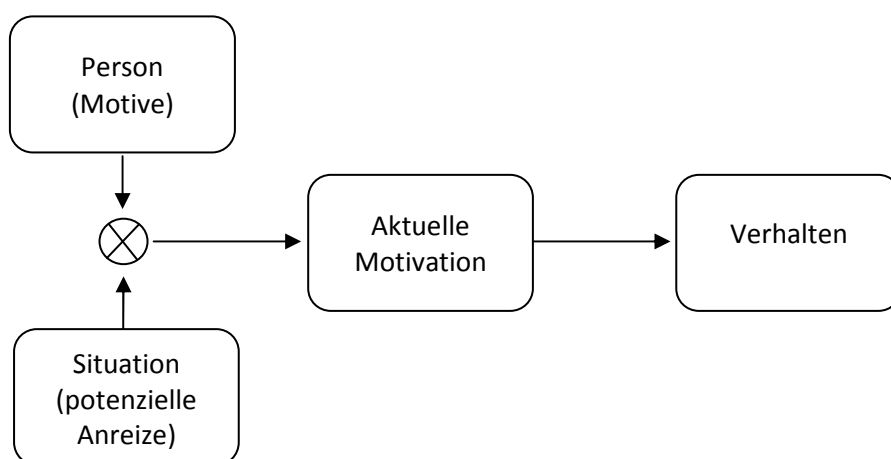


Abb. 6. Grundmodell der klassischen Motivationspsychologie (entnommen aus Vollmeyer, 2005, S. 11).

Aufgrund von Lewins Feldtheorie kann geschlossen werden, dass „weder Triebe, Gewohnheiten, Bedürfnisspannungen allein, noch situative Reize, Zwänge, Verlockungen“ zu

Verhaltenskonsequenzen führen (Rheinberg, 2004, S. 47). Lewin betont, dass Verhalten nur erklärt werden kann, wenn sowohl Personen- als auch Umweltfaktoren in der Verhaltensanalyse herangezogen werden. Lewin bringt dies mit seiner Verhaltensgleichung $V = f(P, U)$ zum Ausdruck (Lewin, 1936; vgl. auch Rheinberg, 2004, S. 47), wonach das Verhalten eine Funktion von Personen- und Umweltfaktoren ist. Das Verhalten wird von den gegenwärtigen Kräften bestimmt, die in einem Moment auf eine Person einwirken: „Nur was gegenwärtig wirkt, ist verhaltenswirksam“ (Beckmann & Heckhausen, 2006, S. 108). Die Interaktion von Personenmerkmalen und Situationsanreizen führt zur aktuellen Motivation, die wiederum das Verhalten beeinflusst. Abbildung 6 stellt das auf Lewin zurückgehende Grundmodell der klassischen Motivationspsychologie dar (vgl. Rheinberg, 2004).

Die bedeutendste Implikation des Grundmodells der klassischen Motivationspsychologie für die Erforschung von Lern- und Leistungsmotivation ist, dass die aktuelle Motivation in der Situation durch die Erfassung motivationaler Personenmerkmale („traits“) alleine nur ungenügend vorhergesagt werden kann (Vollmeyer & Rheinberg, 2003): Um verhaltenswirksam zu werden, bedürfen diese motivationalen Personenmerkmale der Anregung über eine passende Situation bzw. eine Lern- oder Problemlöseaufgabe (Rheinberg, 2004; siehe auch McClelland et al., 1953).

Die Motivation als aktivierende Zielausrichtung im Hier und Jetzt ergibt sich nach Vollmeyer und Rheinberg (2003) aus der „situativen Anregung (...) motivationale[r] *traits*“ (S. 283). Somit liefere die in einer konkreten Situation erfasste Motivation einen größeren prädiktiven Beitrag als zeitlich und situativ überdauerende Persönlichkeitsmerkmale. Boekaerts (1986) wies in diesem Zusammenhang darauf hin, dass sich auch die Erfassung motivationaler *traits* von der Messung motivationaler *states* bei Leistungs- und Lernverhalten unterscheidet. In Untersuchungen ging Boekaerts (1986, 2002) unter anderem der Frage nach, ob aktuelle Verhaltenstendenzen in Lernsituationen stärker von *traits* oder *states* vorhergesagt werden können. Sie unterschied hierbei zwischen allgemeinen motivationalen Orientierungen („general motivational orientation“; auf der Ebene von *traits*) oder aufgaben- bzw. domainspezifischen motivationalen Orientierungen („task-specific motivational orientations“ bzw. „situation specific action tendency“; Boekaerts, 1986, S. 231).

Als motivationale *traits* werden in der pädagogisch-psychologischen Forschung vor allem Motive wie das Leistungsmotiv (Brunstein & Heckhausen, 2006; McClelland et al., 1953), motivationale Zielorientierungen (Dweck & Leggett, 1988), das individuelle Interesse (Krapp, 1992) sowie Fähigkeitsselbstkonzepte (Shavelson, Hubner & Stanton, 1976) untersucht. Diese Personenfaktoren stellen hoch generalisierte Merkmale dar, die nicht oder nur in geringem Maße zwischen Situationen und Zeitpunkten variieren.

Das *State*-Maß der Motivation bezeichnen Rheinberg et al. (2001; Vollmeyer & Rheinberg, 2003) als „aktuelle Motivation“. Die aktuelle Lernmotivation resultiert aus der Passung der Anregungsgehalte einer konkreten Lern- und Leistungssituation und der Motivstruktur des Lernenden. Die aktuelle Motivation ist vergleichbar mit dem von Boekaerts (1986, S. 231) beschriebenen Konstrukt der aufgabenspezifischen motivationalen Orientierung. Rheinberg et al. (2001, S. 58) betonen in diesem Zusammenhang, dass erst die „aktuelle Motivation – und nicht die überdauernden Motive – ... einen direkten Einfluss auf das Verhalten, z.B. Lernen und Leistung“ ausübt.

Bis in die 1980er Jahre hinein wurde fast ausschließlich die Aufgabenschwierigkeit als bedeutendster Einflussfaktor der Situation auf die aktuelle Motivation beforscht; Aspekte wie der Wert der Aufgabenbearbeitung oder das inhaltliche Interesse am Aufgabenthema wurden ausgeblendet (Boekaerts, 1986, S. 229). Die Forschung zu motivationsrelevanten Aufgabenaspekten hat sich seitdem weiter entwickelt und es wurden in den letzten Jahren weitere Aufgabenmerkmale und deren Auswirkungen auf die aktuelle Motivation untersucht (z.B. Rheinberg et al., 2001). Darunter findet sich die aktuelle Motivation, deren Beitrag zur Leistung bei Lern- und Problemlöseaufgaben in Abschnitt 3.3 näher erläutert wird.

Csikszentmihalyi (1975, 2000) sowie Rheinberg und Heckhausen (1980) weisen darauf hin, dass nicht nur Merkmale der Person und der Situation motivationsrelevant sind. Auch die Ausführung der Tätigkeit selbst kann motivierend sein. Den Autoren zufolge wurden Tätigkeitsanreize in den klassischen Motivationstheorien nicht ausreichend berücksichtigt. Das Flow-Konstrukt von Csikszentmihalyi (1990, 2002) beschreibt dagegen einen Tätigkeitsanreiz, der auch für Lernhandlungen bedeutsam ist (Csikszentmihalyi & Schiefele, 1993; Engeser & Vollmeyer, 2005). Das Konstrukt wird im folgenden Abschnitt erläutert (Kapitel 4.2.). Vollmeyer und Rheinberg (1998, 2000, 2006) integrierten das Flow-Konstrukt als Mediator in ihr kognitiv-motivationales Prozessmodell. Dieses Modell wird im Anschluss an das Flow-Konstrukt vorgestellt (siehe Kapitel 4.3.).

4.2 Flow-Erleben als Tätigkeitsanreiz

Klassische Modelle der Leistungsmotivationsforschung (z.B. Atkinson, 1957; Heckhausen & Rheinberg, 1980) waren zweckrational angelegt (Rheinberg, 2004; Schiefele & Streb-low, 2005)⁷: Der Anlass zum Handeln wurde in den Folgen des Handelns verortet. Beispielsweise werden in Atkinsons Modell (1975) „ausschließlich antizipierte Emotionen der Selbst-

⁷ “Because modern psychology is concerned mainly with behavior and performance, rather than the reality of inner states of experience, psychologists largely ignore the distinction between extrinsic and intrinsic motivation” (Csikszentmihalyi, 1975, S. 5).

bewertung nach Erfolg und Misserfolg“ wie Lob oder Stolz als Anreize zu leistungsmotivierten Verhaltens berücksichtigt (Heckhausen & Rheinberg, 1980, S. 15). In seinem „erweiterten kognitiven Motivationsmodell“ postuliert Heckhausen (1977), dass Menschen insbesondere dann zum Handeln motiviert werden, wenn ihrer Erwartung nach das Ergebnis eigenen Handelns sehr wahrscheinlich mit positiven Folgen verknüpft ist. Nach McClelland et al. (1953) wird leistungsmotiviertes Handeln durch die Vorwegnahme starker positiver Affekte bei erfolgreicher Auseinandersetzung mit einem Gütemaßstab hervorgerufen (Langens, Schmalt & Sokolowski, 2005).

Demgegenüber stellte Rheinberg fest, dass sich nicht alle Formen des Leistungshandelns unter Bezug auf Folgenanreize erklären lassen. In seinem Band „Zweck und Tätigkeit“ betont der Autor (1989; vgl. auch Rheinberg, 2004, 2006) unter Bezugnahme auf Csikszentmihalyi (1996, 1975, 1990, 2000, 2002), dass neben den erwarteten Handlungsfolgen auch der Vollzug der Tätigkeit selbst Handlungsanreize bietet. Zum Beispiel zeigen Personen auch dann leistungsmotiviertes Verhalten, wenn sie erwarten, dass „die Handlung selbst von positiven Erlebenszuständen begleitet wird“ (Schiefele, 2008, S. 41). Die Tätigkeit wird ausgeführt, weil der Handlungsvollzug selbst Freude bereitet (Engeser & Vollmeyer, 2005). Diese Form der Motivation wird von mehreren Autoren als „intrinsisch“ bezeichnet (u. a. Deci & Ryan, 1985, 2000; Rheinberg, 2004, 2006; Sansone & Harackiewicz, 2000; Schiefele & Streblow, 2005): Nach Schiefele und Streblow (2005) sind besonders Explorations- und Spielhandlungen intrinsisch motivierend (vgl. auch Wood et al., 2000).

In mehreren Studien erforschten Csikszentmihalyi (1975, 2000; Csikszentmihalyi & LeFevre, 1989) sowie Rheinberg (1993, zitiert nach Engeser & Vollmeyer, 2005; Rheinberg, 1989; Rheinberg & Manig, 2003; Rheinberg & Tramp, 2006), welche Merkmale der Tätigkeit deren Ausführung besonders attraktiv machen. Mittels Interviews und standardisierter Fragebogen untersuchten sie die spezifische Anreizstruktur von Freizeitaktivitäten: Diese „autotelischen“ Aktivitäten sind nicht auf die Erfüllung eines bestimmten Zwecks ausgerichtet. Sie sind intrinsisch motiviert und werden auch unter Inkaufnahme negativer Handlungsfolgen ausgeführt.⁸ Csikszentmihalyi (1975, 2000) stellte bei seinen Studien fest, dass Personen bei der Ausführung dieser intrinsisch motivierten Handlungen ein charakteristisches Erleben von *Flow*-Zuständen berichten (Csikszentmihalyi & Schiefele, 1993; Engeser & Vollmeyer, 2005; siehe auch Rheinberg, 2004, 2006; Schiefele, 2008). Nach Rheinberg (2006) wird das *Flow*-Erleben beschrieben als ein „selbstreflexionsfreies Aufgehen in einer glatt laufenden Tätigkeit, bei der man trotz voller Kapazitätsauslastung das Gefühl hat, den Geschehensablauf noch gut

⁸ „People do get immersed in games so deeply as to forget hunger and other problems. What power does play have that men relinquish basic needs for its sake?“ (Csikszentmihalyi, 1975, S. ix, Preface).

unter Kontrolle zu haben“ (S. 345). Nach Csikszentmihalyi (1975, 2000; Csikszentmihalyi & LeFevre, 1989) zufolge ist Flow mit positiven Gefühlen verbunden (intrinsisch belohnend) und wird vom ihm mit Glückserleben gleichgesetzt.

4.2.1 Komponenten des Flow-Erlebens

Flow stellt einen einheitlichen Bewusstseinszustand dar, in welchem das Handeln als Fließen von einem Augenblick zum nächsten erlebt wird (Csikszentmihalyi & Schiefele, 1993). Flow setzt sich aus mehreren Bestandteilen zusammen. Zu den charakteristischen Merkmalen zählen die folgenden (vgl. Csikszentmihalyi, 1975, S. 49ff., 1996, S. 110ff.; siehe auch Aellig, 2004, S. 36; Csikszentmihalyi & Schiefele, 1993, S. 209f.; Engeser & Vollmeyer, 2005, S. 63; Rheinberg, 2006, S. 346):

1. Gefühl der Kontrolle des Handlungsablaufs
2. Verschmelzen von Handlung und Bewusstsein
3. Fokussierung der Aufmerksamkeit
4. Beeinträchtigung des Zeiterlebens
5. Selbstvergessenheit (Verschmelzen von Selbst und Tätigkeit)
6. Autotelische Tätigkeit

Diese Komponenten des Flow-Erlebens werden im Folgenden näher beschrieben.

Gefühl der Kontrolle des Handlungsverlaufs. Die Person hat im Flow-Zustand das Gefühl, den Handlungsablauf stets unter Kontrolle zu haben. Sie fühlt sich leistungsfähig. Die Besorgnis über Misserfolge tritt in den Hintergrund: Die Person ist so stark auf die Bewältigung der Anforderungen konzentriert, so dass Gedanken an Misserfolge nicht ins Bewusstsein gelangen. Der Handlungsablauf selbst wird als glatt und fließend erlebt und ist wie aus einer inneren Logik heraus gesteuert.

Verschmelzen von Handlung und Bewusstsein. Im Flow-Zustand ist die Person vornehmlich der Handlung und nicht ihrer Selbst als handelndem Subjekt bewusst. Die Grenzen zwischen Handlung und Bewusstsein sind aufgehoben, beide werden als Einheit erlebt. Die Aufmerksamkeit wird von der eigenen Person abgezogen. Man „vergisst“ sich selbst und die Umwelt im Flow-Zustand. Handlungen werden wie automatisch ausgeführt. Dieser Bewusstseinszustand wird jedoch immer wieder unterbrochen durch Phasen der Selbstreflexion, da er nicht über längere Zeit aufrechterhalten werden kann.

Zentrierung der Konzentration. Die Konzentration ist anstrengungsfrei und wie von selbst auf die Ausführung der Tätigkeit fokussiert. Es besteht dabei keine Notwendigkeit

zur willentlichen Regulation der Aufmerksamkeitsprozesse. Mögliche Ablenkungen werden automatisch ausgeblendet; es werden lediglich Informationen verarbeitet, die im Zusammenhang mit der Tätigkeit stehen. Die Konzentration richtet sich auf das gegenwärtige Handlungsgeschehen. Vergangene und zukünftige Ereignisse finden nur geringe Beachtung.

Beeinträchtigung des Zeiterlebens. Im Flow-Zustand verändert sich das Zeiterleben: Stunden vergehen darin wie Minuten. Auch das gegenteilige Gefühl kann eintreten, so dass kurze Zeiträume als ausgedehnt erlebt werden. Die Beziehung zwischen physikalischem Zeitvergehen und subjektiv empfundenem ist verändert. Die starke Konzentration auf die Tätigkeit strukturiert das Zeitempfinden neu.

Selbstvergessenheit. Es verschmelzen Selbst und Tätigkeit im Flow-Zustand. Gedanken an die eigene Person geraten in den Hintergrund. Flow ist gekennzeichnet durch eine verringerte (Selbst-)Reflektivität. Selbstbezogene Kognitionen wie z.B. Zweifel an den eigenen Fähigkeiten oder positive Selbstbewertungen wie Stolz, werden ausgeblendet. Jedoch bedeutet Selbstvergessenheit keinen Verlust der Wahrnehmung für innere Vorgänge: Einige psychische und physische Prozesse werden sogar intensiver erlebt. Die Selbstvergessenheit bedeutet vielmehr, dass das Selbst als Steuerungsinstanz im Tätigkeitsvollzug nicht mehr wahrgenommen wird.

Autotelische Tätigkeit. Im Flow-Zustand wird die Ausführung der Tätigkeit selbstbelohnend. Kompetenzerleben und Autonomie im Tätigkeitsvollzug sind intrinsisch motivierend und tragen zum Gefühl optimalen Funktionierens bei. Es wird kein weiterer, äußerer Anreiz benötigt, um mit der Tätigkeit fortzufahren. In der Praxis bedarf es jedoch äußerer Anreize, um die Flow-Aktivität zu beginnen. Mit der Zeit wird die Flow-Tätigkeit jedoch dann als intrinsisch erlebt und selbstinitiativ fortgeführt. Die Handlung selbst ist dabei nicht anstrengungslos, sondern kann mit erhöhten psychischen und physischen Belastungen verbunden sein.

4.2.2 Vorbedingungen des Flow-Erlebens

Prinzipiell kann Flow-Erleben bei jeder Aktivität auftreten. Einige Aktivitäten sind jedoch aufgrund ihrer Eigenschaften eher förderlich für Flow als andere. Diese Aktivitäten bieten die Möglichkeiten zum Erlernen neuer Fähigkeiten, zum Setzen von Zielen, zum Erhalten von Leistungsrückmeldungen sowie zur Handlungskontrolle (Csikszentmihalyi, 2002). Nach Engeser und Vollmeyer (2005; siehe auch Rheinberg, 2006) ist das Auftreten von Flow an zwei Vorbedingungen geknüpft:

- 1) Das Gefühl optimaler Beanspruchung,
- 2) Eindeutige Handlungsanforderungen und unverzügliche Rückmeldungen.

Die erste Vorbedingung, das Gefühl optimaler Beanspruchung, ergibt sich aus der „erlebte[n] Passung zwischen Anforderung und Fähigkeit“ (Rheinberg, 2006, S. 347). Csikszentmihalyi (2000, S. 75) geht davon aus, dass Flow dann eintritt, wenn die Handlungsanforderungen die eigenen Handlungsfähigkeiten nicht übersteigen. Übersteigen die Handlungsanforderungen die subjektiven Handlungsfähigkeiten dagegen deutlich, empfindet man zunächst Sorge gefolgt von Angst. Ist die Diskrepanz umgekehrt, d.h. die individuellen Fähigkeiten werden als beträchtlich höher eingeschätzt als die Anforderungen der Tätigkeit, dann erlebt die Person Langeweile. Flow ist demzufolge als ein Gleichgewichtszustand „beyond boredom and anxiety“ zu verstehen (Csikszentmihalyi, 1975).

Die zweite, zentrale Vorbedingung für das Auftreten von Flow betrifft die klare Struktur des Handlungsablaufs und den Erhalt unverzüglicher Rückmeldungen (Csikszentmihalyi, 1996, S. 114f.; 2002, S. 54f.). Nur wenn beides über den gesamten Handlungsverlauf hinweg gegeben ist, kann die Person völlig in der Tätigkeit aufgehen (Engeser & Vollmeyer, 2005; Schiefele & Streblov, 2005).

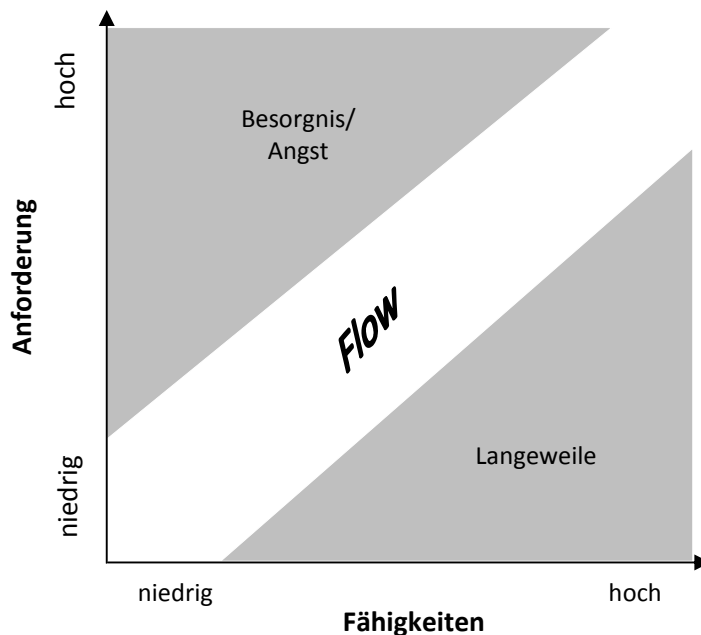


Abb. 7. Flow-Kanal-Modell von Csikszentmihalyi (1975, 2000) in Anlehnung an Rheinberg (2006, S. 348).

Abbildung 7 illustriert das Flow-Kanal-Modell von Csikszentmihalyi (1975, 2000). Der Kanal in der Mitte stellt den Bereich dar, in dem Anforderung und Fähigkeit im Gleichgewicht sind. Zum Flow-Kanal-Modell wurde kritisch angemerkt, dass das Flow-Konzept darin auf eine einzige Komponente bzw. Vorbedingung – die Passung von Anforderung und Fähigkeit –

reduziert werde (vgl. Rheinberg, 2006). Darüber hinaus sei das Erleben von Flow nicht gleichzusetzen mit der Passung von Anforderungen und Fähigkeiten; beispielsweise empfänden erfolgsmotivierte Personen diese Passung eher als förderlich als Misserfolgsmotivierte (vgl. Atkinson, 1957; Heckhausen, 1963).

Im Flow-Kanal-Modell postuliert Csikszentmihalyi (1975, 2000; 2002), dass Flow-Erleben auch bei niedrigen Anforderungen und Fähigkeiten möglich ist (vgl. Flow-Kanal-Modell, Abbildung 7). In einer Modellrevision nimmt er demgegenüber an, dass Flow nur bei überdurchschnittlich ausgeprägten Fähigkeiten und Anforderungen auftrete (vgl. Massimini & Carli, 1988). Demzufolge könnten nur Experten, jedoch nicht Anfänger Flow erleben (Engeser & Vollmeyer, 2005; Rheinberg, 2006). Rheinberg (2006) betont dagegen, dass dieser Expertise-Effekt nicht verallgemeinert werden kann, vielmehr treffe dieser nur bei „komplexere[n] Aktivitäten [zu], bei denen mehrere Basiskomponenten zunächst automatisiert werden müssen, bevor sich ein halbwegs kohärenter Handlungsablauf ergeben kann“ (S. 349). Hierzu zählen komplexe Interaktionen mit dem Computer. Anfänger kämen hier häufiger ins Stocken, was einen glatten Handlungsverlauf verhindert (Engeser et al., 2005). Bei weniger komplexen Anforderungen, wie z.B. bei Computerspielen, könnten dagegen auch Anfänger Flow empfinden.

Ob eine Aktivität geeignet ist, Flow auszulösen, hängt in erster Linie von den individuellen, subjektiven Merkmalen der handelnden Person ab. Zu diesen Merkmalen zählt insbesondere die wahrgenommene eigene Handlungskompetenz. Schiefele und Streblow (2005) bemerken, dass sich langfristig „Erfahrungen des Flow-Erlebens nur dann wiederholen, wenn die Anforderungen an die steigende Kompetenz der Person angepasst werden, was wiederum eine Steigerung ihrer Kompetenz zur Folge hat. Dies impliziert, dass Flow eine wichtige Quelle für die Kompetenzentwicklung darstellt“ (S. 51). Als leistungsmotivationaler Tätigkeitsanreiz ist Flow assoziiert mit der „Freude am eigenen optimal-effizienten Funktionieren“ (Rheinberg, 2006, S. 345) auf dem Weg zu einem Leistungsziel. Krapp (2005) geht ebenfalls davon aus, dass Personen ein Bedürfnis nach „optimalem Funktionieren“ in der Auseinandersetzung mit der Umwelt haben. Daraus ergebe sich das „Streben nach tätigkeitszentriertem Wohlbefinden“ (Krapp, 2005, S. 27).

Csikszentmihalyi (1975, 2000) unterscheidet zwischen dem voll ausgeprägten Zustand des Flow-Erlebens (*deep flow*) und einer abgeschwächten Variante von Flow, dem *microflow*. *Microflow* beinhalte nicht alle Komponenten des Flow-Erlebens, er trete auf bei einfachen, unstrukturierten Aktivitäten des Alltags. Diese Tätigkeiten spendeten keine besondere Freude, seien jedoch für das optimale Funktionieren im Alltag wichtig (Csikszentmihalyi, 2000, S. 81).

4.2.3 Empirische Befunde und Zusammenfassung

In der pädagogisch-psychologischen Forschung wird im Allgemeinen angenommen, dass Flow-Erleben eine günstige Wirkung auf den Lernprozess ausübt (Rheinberg, 2006). In mehreren Studien konnte diese Annahme empirisch untermauert werden (z. B. Csikszentmihalyi & Schiefele, 1993; Engeser et al., 2005; Rheinberg & Vollmeyer, 2003). Vollmeyer und Rheinberg (1998) kommen im Rahmen einer experimentellen Studie mit einem komplexen, dynamischen System zu dem Schluss, dass insbesondere die Flow-Komponente „anstrengungsfreie Konzentration“ die Leistung positiv beeinflusst. Zu einem ähnlichen Ergebnis gelangt Schiefele (1996): Hoch interessierte Lernende hatten beim Textlesen weniger Schwierigkeiten, sich zu konzentrieren; dies ist für Flow-Erfahrungen typisch. Dieser Befund wird durch eine Studie von Ghani und Deshpande (1994) unterstützt. Sie befragten Arbeitnehmer zu typischen Interaktionen mit dem Computer am Arbeitsplatz. Flow übte hier eine positive Beziehung auf die Konzentration beim Arbeiten aus und auf das Explorieren eines Computerprogramms: Personen mit hohem Interesse und dem Gefühl der Herausforderung waren eher bereit, ein Computerprogramm zu explorieren, auch auf die Gefahr hin, in einer Sackgasse zu enden.

Engeser et al. (2005) konnten für das universitäre Fremdsprachen- und Statistiklernen zeigen, dass das Flow-Erleben die Lernleistung auch dann vorhersagt, wenn Kompetenzunterschiede vor Lernbeginn statistisch kontrolliert werden. Das bedeutet, dass Flow über die *Expertise* einer Person hinaus einen positiven Einfluss auf die Lernleistung entfaltet. Gleichzeitig führen Rheinberg, Manig, Kliegel, Engeser und Vollmeyer (2007, S. 106) mehrere Studien an, die Hinweise darauf geben, dass Flow-Erleben eher bei Personen mit höherem Kompetenzniveau auftritt. Die Höhe des Flow-Erlebens variiert dabei mit dem Schwierigkeitsgrad der Aufgabe, wie Rheinberg und Vollmeyer (2003) im Rahmen einer experimentellen Studie mit dem Computerspiel („Roboguard“) berichten. Beide Variablen standen in einem kurvilinearen Zusammenhang: Wie aufgrund des Flow-Kanal-Modells angenommen (Csikszentmihalyi, 1975, 2000), stiegen die Flow-Werte zunächst mit den wachsenden Anforderungen. War der Punkt optimaler Beanspruchung jedoch erreicht, sanken die Flow-Werte wieder mit zunehmendem Schwierigkeitsgrad der Aufgabe.⁹

Auf das Phänomen „Flow“ war Csikszentmihalyi (1975, 2000) gestoßen, als er die Handlungsanreize von Personen untersuchte, die eine Tätigkeit ohne erkennbare Belohnung engagiert ausführten (Rheinberg, 2006; Schiefele & Streblow, 2005). Im Gegensatz zu Csikszentmihalyis Vermutung stehen Forschungsergebnisse, die von Flow-Erleben nicht nur bei

⁹ Zum Zusammenhang von Flow und der Passung von Fähigkeit und Anforderungen siehe auch Engeser und Rheinberg (2008).

intrinsischer Motivation, sondern auch bei extrinsischer berichten (Rheinberg & Vollmeyer, 2003). So standen in der Studie von Rheinberg und Vollmeyer (2003) sowohl (extrinsische) Leistungs- als auch (intrinsische) Lernziele in positiver Beziehung mit der Flow-Komponente „Absorbiertheit“ (s. auch Dweck & Leggett, 1988).

Möglichkeiten zur Förderung des Flow-Erlebens bieten sich insbesondere über die Implementation eindeutiger Handlungsanforderungen und informativer Rückmeldungen. Dies verdeutlichen die Befunde zweier empirischer Studien von Csikszentmihalyi und Schiefele (1993) zum Lernen in schulischen Settings. Die Bedeutung klarer Handlungsanforderungen konnten Rheinberg et al. (2007) auch für die Arbeitswelt zeigen: Bei Tätigkeiten mit einer klaren Zielausrichtung zeigten sich höhere Flow-Werte als bei schwacher bzw. fehlender Zielausrichtung.

4.3 Kognitiv-motivationales Prozessmodell von Vollmeyer und Rheinberg

Studien zum selbstregulierten Lernen (z. B. Boekaerts, 2002; Zimmerman, 2000; Winne & Hadwin, 2008) haben dazu beigetragen, dass dem Lernprozess selbst eine größere Aufmerksamkeit innerhalb der pädagogisch-psychologischen Forschung gewidmet wurde (Rheinberg, Vollmeyer & Rollett, 2002). Auf der Basis der theoretischen Ansätze des Selbstregulierten Lernens (z.B. Boekaerts, 1996) und der von Atkinson (1957) und Lewin (1946) entwickelten Vollmeyer und Rheinberg ihr kognitiv-motivationales Prozessmodell (Vollmeyer & Rheinberg, 1998, 1999, 2000, 2004, 2006). Ein wesentliches Ziel des Modells ist es, den spezifischen Einfluss motivationaler Prozessvariablen auf die Performanz in Lern- und Problemlöseprozessen zu beschreiben und zu erklären (Schiefele & Rheinberg, 1997). Die Erfassung motivationaler Aspekte in der konkreten Lern- und Leistungssituation selbst bietet die Möglichkeit, eine größere Varianz in der Leistung aufzuklären und den Einfluss der Motivation auf die Leistung in Lern- und Problemlöseprozessen handlungsnäher zu bestimmen (Boekaerts, 2002; Schiefele & Rheinberg, 1997).

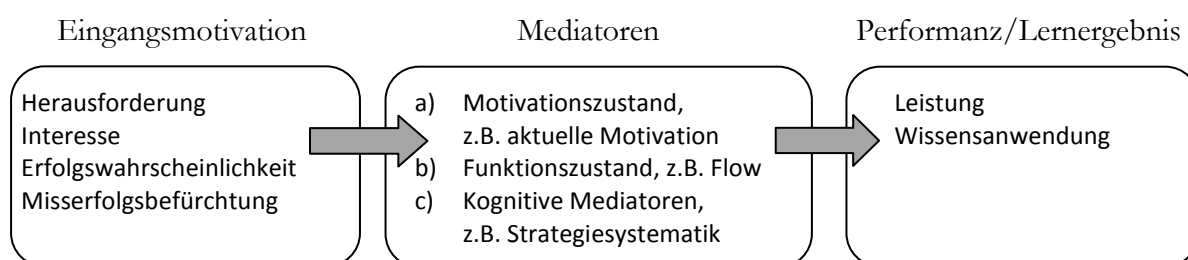


Abb. 8. Kognitiv-motivationales Prozessmodell nach Vollmeyer und Rheinberg (1998, 2006; vereinfachte Darstellung mit ausgewählten Aspekten).

Abbildung 8 illustriert die einzelnen Komponenten des kognitiv-motivationalen Prozessmodells. Die Grafik stellt ein reduziertes Modell dar, das ausschließlich jene Komponenten enthält, die in der vorliegenden empirischen Arbeit berücksichtigt wurden. Ausgangspunkt der Modellentwicklung war die Annahme, dass aktuelle Motivation durch eine Interaktion von Situations- und Personenmerkmalen entsteht (vgl. Kapitel 4.1; interaktionistische Motivations- theorie). Die Autoren postulieren, dass die aktuelle Lernmotivation durch die Konfrontation mit einer konkreten Lern- oder Problemlöseaufgabe angeregt werden muss, um verhaltens- wirksam werden zu können (Vollmeyer & Rheinberg, 2003; s. auch Boekaerts, 2002). Zu den (situativen) Einflüssen der Aufgabe zählen beispielsweise die perzipierte Schwierigkeit der Aufgabe (vgl. Atkinson, 1957; Kukla, 1974) und die Erwartung, eine Rückmeldung zur eigenen Leistungsfähigkeit zu erhalten (vgl. Vollmeyer & Rheinberg, 2005).

Die durch das Aufgabenmaterial angeregte, aktuelle Motivation vor Beginn des Lern- prozesses bezeichnen Vollmeyer und Rheinberg (2000, 2002, 2006) als *Eingangsmotivation*. Diese wird im Modell als unabhängige Variable konzeptualisiert, mit deren Hilfe die Ausprägungen in anderen Variablen vorhergesagt werden. Die Eingangsmotivation setzt sich dabei aus qualitativ unterschiedlichen Dimensionen wie dem Interesse und der Herausforderung zusammen. Da eine hohe (Eingangs-)Motivation allein keinen Einfluss auf die Performanz in Lern- und Problemlöseprozessen ausüben kann, nehmen Vollmeyer und Rheinberg (1998) an, dass motivationale Faktoren mit der konkreten „Ausführung innerer und/oder äußerer (Lern-) Aktivitäten“ (S. 12) in Verbindung stehen. Das Modell impliziert folglich, dass Unterschiede in der Höhe der Eingangsmotivation zu interindividuellen Differenzen in *Mediatorvariablen* während des Lernprozesses führen.

Zu den Mediatoren zählen im Modell in erster Linie motivationale und kognitive Variablen sowie der Funktionszustand.¹⁰ Der Zusammenhang zwischen Eingangsmotivation und Mediatoren kann an einem Beispiel verdeutlicht werden: Es wird angenommen, dass hoch interessierte Lernende bei der Lösung einer Problemaufgabe häufiger systematische Strategien einsetzen als Personen, die wenig an den Inhalten der Aufgabe interessiert sind. Für die Mediatoren wird wiederum ein direkter Zusammenhang mit der *Performanz* bzw. dem Lernergebnis postuliert. Theoretisch stellen damit die im Prozess erhobenen Variablen die besten Prädiktoren der Performanz dar.

Zusammenfassend lässt sich die Rolle der Motivation im kognitiv-motivationalen Prozessmodell folgendermaßen beschreiben: Die Eingangsmotivation wirkt sich nicht direkt auf

¹⁰ Darüber hinaus wurde untersucht, ob der Einfluss der Eingangsmotivation z.B. über metakognitive Prozesse (Bachmann, 2009; Vollmeyer & Rheinberg, 1999) oder über Persistenz (Vollmeyer & Rheinberg, 2000) vermittelt wird.

die Performanz aus. Vielmehr wird ihr Einfluss über kognitive, metakognitive und motivationale *Mediatoren* vermittelt (siehe Abb. 8). In Abschnitt 4.3.1 werden zunächst die Dimensionen der aktuellen Motivation im Rahmen der Eingangsmotivation sowie während des Prozesses näher beschrieben. Im Anschluss werden in Kapitel 4.3.2 die Mediatoren des Modells erläutert.

4.3.1 Aspekte der aktuellen Motivation

In bisherigen Studien wurden zahlreiche motivationale Konzepte identifiziert, die zu einem erfolgreichen Lern- und Problemlöseprozess beitragen (z. B. Deci & Ryan, 1985; Dweck & Leggett, 1988; Heckhausen & Rheinberg, 1980; Locke & Latham, 1990; Pajares, 1997; Schiefele & Rheinberg, 1997). Nach Boekaerts (2002, S. 80) können Merkmale der Aufgabe und der Situation (z. B. Herausforderung, Freude, Autonomie, Kompetenzerleben oder soziale Eingebundenheit) Gefühle von Langeweile und Angst hervorrufen.

Die Anzahl relevanter motivationaler Konzepte kann faktorenanalytisch auf die folgenden vier Dimensionen reduziert werden (Rheinberg et al., 2001): Erfolgswahrscheinlichkeit, Misserfolgsbefürchtung, Herausforderung und Interesse. Die Auswahl der berücksichtigten Dimensionen erfolgte dabei nach drei Kriterien (siehe Vollmeyer & Rheinberg, 2003, S. 284): Erstens sollte die erwartete Salienz der motivationalen Konzepte in Lern- und Leistungssituationen hoch sein. Zweitens sollten sich die ausgewählten motivationalen Konzepte zuverlässig gruppieren lassen. Drittens sollten sich diese Konzepte theoretischen Bedeutungssystemen zuordnen lassen.

Erfolgswahrscheinlichkeit. Dieses Konzept wurde bereits von Atkinson (1957) und Lewin et al. (1926) in die Motivationspsychologie eingeführt. Atkinson (1957) führt die aktuelle Leistungsmotivation bei der Bearbeitung einer Aufgabe auf die Verknüpfung dreier Variablen zurück: Stärke/Grundrichtung des Leistungsmotivs, Erwartungen und Anreize (siehe auch Beckmann & Heckhausen, 2006; Rheinberg, 2004). Personen unterscheiden sich in ihrer Persönlichkeit darin, ob sie bei einer Aufgabe eher anstreben, Erfolg zu erzielen („Erfolgsmotivierte“) oder Misserfolge zu vermeiden („Misserfolgsmotivierte“). Für die Erklärung und Vorhersage der aktuellen Motivation der Erfolgs- und Misserfolgsmotivierten sind insbesondere Erwartungen (Erfolgswahrscheinlichkeit bzw. Misserfolgsbefürchtung) und Anreize in einer konkreten Situation relevant. Personen schätzen die Wahrscheinlichkeit ihres Erfolgs bei einer Aufgabe umso höher ein, je leichter die Aufgabe eingestuft wird. Der umgekehrte Fall gilt für die Misserfolgsbefürchtung: Misserfolge sind umso unwahrscheinlicher, je leichter die Aufgabe ist. Beide Wahrscheinlichkeiten stellen in Atkinsons Modell darum „komplementäre Größen dar“ (Beckmann & Heckhausen, 2006, S. 131). Erfolgsmotivierte bewerten den positiven An-

reiz des Erfolgs höher als den negativen des Misserfolgs. Bei einer mittleren Erfolgswahrscheinlichkeit ist die aktuelle Motivation bei den Erfolgsmotivierten besonders hoch (Atkinson, 1957).

In Übereinstimmung mit Atkinson (1957) gehen Vollmeyer und Rheinberg (2003) davon aus, dass eine Person ihre eigenen Kompetenzen mit den Aufgabenanforderungen abgleicht, sobald sie mit diesen konfrontiert wurde: „Aus diesen, meist wohl automatisiert ablaufenden, Vergleichsprozessen resultiert der Grad an Zuversicht, mit der man die Aufgabe in Angriff nimmt“ (S. 284). Die Erfolgswahrscheinlichkeit bedingt auch im kognitiv-motivationalen Prozessmodell die positive Anreizkomponente der Leistung (*Erfolgsmotivation*; Atkinson, 1957; Heckhausen, 1963; McClelland et al., 1953). Im nomologischen Netzwerk weist das Konzept der Erfolgswahrscheinlichkeit enge Beziehungen mit weiteren Konstrukten wie Self-Efficacy (Bandura, 1997) und Kontrollüberzeugungen (Rotter, 1989; s. auch Krampen, 2000) auf. Schunk und Pajares (2004, S. 85) definieren Self-Efficacy (Bandura, 1982) als „one’s perceived capabilities to learn or perform behaviors at designated levels“. Selbstwirksamkeitserwartungen haben in Lern- und Problemlöseprozessen positive Auswirkungen auf die Anstrengung, die Persistenz, die Zielsetzung und die Leistung (Bandura, 1997; Caprara et al., 2008). Im Bereich mathematischen Problemlösens korrelieren Selbstwirksamkeitserwartungen höher mit Mathematikleistungen als z.B. Mathematikängstlichkeit oder mathematisches Vorwissen (Pajares & Graham, 1999). Überdies fanden die Autoren heraus, dass die Selbstwirksamkeit einen substanziellen eigenständigen Beitrag zur Vorhersage der Mathematikleistungen liefert.

Misserfolgsbefürchtung. Im Gegensatz zur Erfolgswahrscheinlichkeit beinhaltet das Konzept der Misserfolgsbefürchtung die negative Anreizkomponente der Leistung (*Misserfolgsmotivation*; Atkinson, 1957; Heckhausen, 1963). Personen, die eine Lern- oder Problemlöseaufgabe bearbeiten, können besorgt sein, bei der Aufgabe zu versagen bzw. eine schlechte Leistung zu zeigen. In ihren experimentellen Studien wurden McClelland et al. (1953) bereits auf diese zweite, leistungsbezogene Tendenz aufmerksam. Als Aspekt der aktuellen Motivation beschreibt diese Dimension den situativ gegebenen affektiven Zustand, der mit einem Misserfolg in der konkreten Situation assoziiert wäre (vgl. Vollmeyer & Rheinberg, 2003). Personen mit hoher Misserfolgsbefürchtung unterscheiden sich in ihren Kognitionen und Affekten sowie im Verhalten von Personen mit hoher Erfolgserwartung: Misserfolge führen sie auf mangelnde Begabung zurück, wohingegen Erfolg auf Zufall und Glück attribuiert werden (Brunstein & Heckhausen, 2006; Weiner, 2010). Sie versuchen, die negativen Selbstbewertungsaffekte (Scham, Entmutigung) infolge eines Misserfolgs zu vermeiden, indem sie extrem schwierige oder extrem leichte Aufgaben wählen, eine geringere Persistenz zeigen und Aufga-

ben schneller abbrechen (vgl. Misserfolgsmotivation; Brunstein & Heckhausen, 2006; Rheinberg, 2004).

Heckhausen (1963, S. 68) zufolge lässt sich die Erfolgstendenz („Hoffnung auf Erfolg“) mit der Misserfolgstendenz („Furcht vor Misserfolg“) in Bezug setzen: Die Differenz der beiden Motivationskomponenten bezeichnet Heckhausen als *Netto-Hoffnung*. Die Differenzbildung liegt auch für eine Verrechnung der beiden Aspekte der aktuellen Motivation (Erfolgswahrscheinlichkeit/Misserfolgsbefürchtung) nahe. In der Literatur wird die Verrechnung beider Motivationskomponenten jedoch kritisch bewertet: Eine hohe Misserfolgsbefürchtung stellt nicht unbedingt das Gegenteil einer als hoch eingeschätzten Erfolgswahrscheinlichkeit dar (Vollmeyer & Rheinberg, 2003; vgl. auch Atkinson, 1957). Darüber hinaus wirkt sich eine hohe Erfolgswahrscheinlichkeit aus der Sicht anderer Autoren (Beckmann & Heckhausen, 2006) nicht zwangsläufig positiv auf die Lernleistung aus. Gleichzeitig ist eine hohe Misserfolgsbefürchtung nicht immer von einer niedrigen Lernleistung begleitet: So sei denkbar, dass die Misserfolgsmotivation einen positiven Effekt auf die Aufgabenbearbeitung hat, in dem die Anstrengung gesteigert wird, um den Misserfolg zu vermeiden. Wie Kukla (1974) zeigen konnte, erreichen Personen mit einer hohen Erfolgsmotivation eine bessere Leistung, wenn die Aufgabe subjektiv als schwer bewertet wird. Demgegenüber erzielen Personen mit einer geringeren Erfolgsmotivation (bzw. einer erhöhten Misserfolgsmotivation) bei Aufgaben, die von ihnen als leicht bewertet werden, ein höheres Performanzniveau als hoch erfolgsmotivierte Personen. Den Befund führt Kukla auf die unterschiedliche Einschätzung eigener Kompetenz („Erfolgswahrscheinlichkeit“ nach Atkinson, 1957) sowie das Maß der Anstrengung bei der Aufgabenbearbeitung zurück.

Herausforderung. Dieser Aspekt der aktuellen Motivation bezieht sich darauf, ob die Person die Aufgabe als leistungsthematisch auffasst (Rheinberg et al., 2001). Er wurde bereits in früheren Ansätzen der Leistungsmotivationspsychologie aufgegriffen (Heckhausen, 1963; Rheinberg & Heckhausen, 1980). Deci und Ryan (1985, S. 32f.) zufolge sorgen die grundlegenden, menschlichen Bedürfnisse nach Kompetenzerleben und Selbstbestimmung dafür, dass Personen fortwährend nach Herausforderungen suchen und diese bewältigen möchten (vgl. auch Csikszentmihalyi, 1975). Die Auseinandersetzung mit einer Aufgabe wird als eine Herausforderung an die eigene Tüchtigkeit erlebt (Vollmeyer & Rheinberg, 2003). Dieser Aspekt der aktuellen Motivation steht damit in enger Beziehung zum (situativ angeregten) Leistungsmotiv (Brunstein & Heckhausen, 2006; McClelland et al., 1953). Vollmeyer und Rheinberg (2006) weisen darauf hin, dass Personen, die in einer Aufgabe gut abschneiden wollen, diese Aufgabe auch als persönlich wichtig empfinden (vgl. Eccles, 2005; Eccles & Wigfield, 1995). Wird die Aufgabe als herausfordernd empfunden, investiert der Lernende ein größeres

Maß an Anstrengung, um eventuelle Hindernisse zu überwinden (Skinner & Brewer, 2002). Darüber hinaus haben „herausgeforderte“ Lernende ein starkes Gefühl persönlicher Handlungskontrolle bei der Aufgabenlösung (Smith & Ellsworth, 1987).

Interesse. Nach Prenzel, Krapp und H. Schiefele (1986, S. 166) wird Interesse definiert als eine spezifische Beziehung zwischen Person und Gegenstand. Bei hohem Interesse wird der Gegenstand kognitiv als persönlich bedeutsam eingestuft und ist mit positiven Emotionen assoziiert (Schiefele, 2008). Rheinberg (2006) beschreibt Interesse als Motivation im Vollzug der Tätigkeit, wobei der Anreiz vor allem darin liegt, sich mit dem Gegenstand des Handelns auseinanderzusetzen. Um ein Beispiel zu geben: Ein Schüler liest Computerbücher, weil er dabei die Möglichkeit hat, sich mit dem Gegenstand (dem Computer) auseinanderzusetzen. Der Anreiz, zu lesen (Tätigkeit), ist dabei nachrangig. Nach Krapp (2005; siehe auch Renninger, Hidi & Krapp, 1992; Ainley, Hidi & Berndoff, 2002) wird Interesse danach differenziert, ob es in einer bestimmten Situation auftritt („aktuelles“ bzw. „situationales“ Interesse) oder es zeit- und situationsübergreifend ist („individuelles Interesse“). Das situationale Interesse wird durch äußere Umstände hervorgerufen (Schiefele, 2008).

Das Interesse an einem Lerngegenstand stellt nach Krapp (2005; Schiefele & Streblow, 2005; Prenzel et al., 1986; Schiefele, Krapp & Schreyer, 1993) eine der wichtigsten Bedingungen für das Auftreten intrinsischer Lernmotivation dar (z.B. Flow-Erleben; siehe Kapitel 4.2). Interesse wirkt sich positiv auf den Affekt (Hidi, Renninger & Krapp, 2004) sowie auf die Aufmerksamkeit im Lernprozess (Hidi & Renninger, 2006) aus. Auch setzen beispielsweise interessierte Studierende qualitativ höhere Lernstrategien ein (z.B. Elaborationsstrategien). Sie beziehen ihr Vorwissen stärker ein, generieren Beispiele und ziehen Schlussfolgerungen (Renninger et al., 1992; Wild, 2000). Ein erhöhtes Interesse ist zudem positiv mit dem Wissenszuwachs korreliert (Hidi & Harackiewicz, 2000). Für schulische Lernsettings konnten Schiefele et al. (1993) in ihrer Metaanalyse einen positiven Zusammenhang zwischen fachspezifischem Interesse und Lernleistung zeigen. Die Studie ergab dabei stärkere Beziehungen zwischen Interesse und Performanz für Jungen als für Mädchen.

Der Einfluss des Interesses und der anderen drei Komponenten der Eingangsmotivation wird über verschiedene Mediatoren vermittelt, die im Folgenden näher beschrieben werden.

4.3.2 Mediatoren

Vollmeyer und Rheinberg (1998, 2000, 2006; siehe auch Schiefele & Rheinberg, 1997) führen Effekte der Motivation auf den Lernerfolg auf drei Gruppen von Mediatorvariablen im Lernprozess zurück: 1) Dauer (*time on task*) und Häufigkeit der Lernaktivität, 2) die Qualität

der Lernaktivitäten, 3) den Funktionszustand sowie den Motivationszustand während der Lernaktivitäten. Im kognitiv-motivationalen Prozessmodell werden in erster Linie die Qualität des Lernprozesses (d.h. Systematik der Lern- und Problemlösestrategien), der Funktions- und der Motivationszustand des Lernenden als Mediatoren zwischen Eingangsmotivation und Lernleistung berücksichtigt. Diese ausgewählten Mediatoren werden im Folgenden näher beschrieben.

Motivationszustand. Nach Vollmeyer und Rheinberg (2003) ist die aktuelle Motivation von den Gegebenheiten in einer konkreten Lernsituation abhängig. Das bedeutet, dass die Höhe der aktuellen Motivation über den Lernprozess hinweg variieren kann. Die aktuelle Motivation bezieht sich dabei auf Bewertungen, Affekte und Interpretationen während des Lernprozesses (Vollmeyer & Rheinberg, 2006). In empirischen Studien wurde die aktuelle Motivation über die positive Valenz, die Konzentrationsfähigkeit und die Erfolgswahrscheinlichkeit/Self-Efficacy operationalisiert (Vollmeyer & Rheinberg, 1998, 1999, 2000). Theoretisch kann die aktuelle Motivation bei der Bearbeitung einer Aufgabe jedoch weitere motivationale Faktoren umfassen (vgl. Rheinberg et al., 2001). Vollmeyer und Rheinberg (2005) reduzierten die Anzahl der zu messenden Konstrukte mit der Begründung, dass der Bearbeitungsprozess dadurch weniger stark beeinträchtigt werde. Erwartungsgemäß zeigte sich in empirischen Studien, dass der Motivationszustand positiv von der Eingangsmotivation beeinflusst wird (Vollmeyer & Rheinberg, 1999, 2006). Darüber hinaus berichten die Autoren positive Beziehungen zwischen dem Motivationszustand und der Performanz.

Funktionszustand. Ähnlich wie der Motivationszustand umfasst auch der Funktionszustand in Lern- und Problemlöseprozessen verschiedene Konstrukte. Hierzu zählen zum Beispiel der Grad an positiver und negativer Aktivierung, das Maß der Erschöpfung volitionaler Ressourcen, sowie die automatische und anstrengungsfreie Konzentration (Engeser et al., 2005; Schiefele & Rheinberg, 1997). Als Indikator für den Funktionszustand wurde in Studien zum kognitiv-motivationalen Prozessmodell insbesondere das Flow-Erleben herangezogen (Engeser et al., 2005; Vollmeyer & Rheinberg, 1998, 2006). Im Zustand von Flow nimmt der Lernende den Handlungsverlauf als glatt, und wie von selbst reguliert, d.h. wie automatisiert wahr (Rheinberg et al., 2003). Flow wirkt sich insbesondere über diese Komponente der anstrengungsfreien Konzentration („Glatter Verlauf“; vgl. Rheinberg et al., 2003) und über die Absorbiertheit durch die Lernaktivität positiv auf die Lernleistung aus (Engeser et al., 2005; Vollmeyer & Rheinberg, 1998). Im Sinne des kognitiv-motivationalen Prozessmodells wird Flow durch Aspekte der Eingangsmotivation beeinflusst: Im Rahmen des universitären Fremdsprachen- bzw. Statistiklernens erwiesen sich das Interesse, die Herausforderung und

die Erfolgswahrscheinlichkeit vor Beginn des Lernprozesses als bedeutsame Prädiktoren des Flow-Erlebens (Engeser et al., 2005; Vollmeyer & Rheinberg, 1998).

Strategiesystematik. Neben dem motivationalen Zustand und dem Funktionszustand dient die Strategiesystematik im Modell als weitere vermittelnde Variable. Die Strategiesystematik zeigt an, wie systematisch der Lernende das Lernmaterial exploriert (Vollmeyer & Rheinberg, 2005). Befunde zeigen, dass kognitive Lern- und Problemlösestrategien einen direkten Effekt auf den Wissenserwerb bzw. die Wissensanwendung ausüben (Schiefele & Rheinberg, 1997; Wild, 2000). Gemäß den Annahmen des kognitiv-motivationalen Prozessmodells wird der Einsatz systematischer Strategien durch die Eingangsmotivation beeinflusst. Studienergebnisse von Vollmeyer und Rheinberg (1998) unterstützen diese Annahme empirisch: Beim komplexen Problemlösen mit dem Programm *Biology Lab* wirkte sich die Eingangsmotivation (z.B. Erfolgswahrscheinlichkeit) positiv auf die Strategiesystematik aus. In einer weiteren Studie zum *Biology Lab* (Rheinberg et al., 2002) stand die Strategiesystematik demgegenüber in keinem Zusammenhang zur Eingangsmotivation. Doch fanden die Autoren höhere Werte der aktuellen Motivation bei Personen, die häufiger systematische Strategien einsetzten. Im Allgemeinen erweisen sich systematische Strategien bei der Steuerung komplexer Systeme (Funke, 2006a, 2006b) als am erfolgreichsten: Der Einsatz systematischer Strategien beeinflusste den Wissenserwerb positiv (Vollmeyer et al., 1996; s. auch Vollmeyer & Funke, 1999). Die Nutzung einer systematischen Strategie spricht damit für eine hohe Qualität der ausgeführten Lern- und Problemlöseaktivitäten (Rheinberg et al., 2002; Vollmeyer & Rheinberg, 1998).

4.3.3 Zusammenfassung

Im kognitiv-motivationalen Prozessmodell von Vollmeyer und Rheinberg (1998, 2000, 2006) werden Leistungsunterschiede beim Lernen und Problemlösen prozessnah auf Unterschiede in der Eingangsmotivation sowie in motivationalen und kognitiven Variablen zurückgeführt. Das Modell baut dabei auf den Annahmen einer Wechselwirkung von Person und Umwelt bei der Genese von aktueller Motivation auf (vgl. Lewin, 1926; Vollmeyer, 2005): Im Gegensatz zu *Trait*-Konzeptionen der Motivation (vgl. McClelland et al., 1953) ist die Motivation bei Vollmeyer und Rheinberg als *State* konzipiert. Danach kann die durch das Aufgabenmaterial angeregte aktuelle Motivation zwischen Aufgaben, Situationen und Zeitpunkten variieren.

Die Autoren des kognitiv-motivationalen Prozessmodells identifizierten vier verschiedene Aspekte aktueller Motivation vor Beginn und während des Lernprozesses bzw. Problemlöseprozesses: Erfolgswahrscheinlichkeit, Misserfolgsbefürchtung, Herausforderung, Interesse.

Der Einfluss der vier Aspekte der Eingangsmotivation wird im Modell über die Strategiesystematik, das Flow-Erleben sowie den motivationalen Zustand vermittelt. Damit trägt das Modell der Annahme Rechnung, dass proximale Prädiktoren der Leistung eine bessere Vorhersage erlauben als zeitlich überdauernde, distale Persönlichkeitseigenschaften. Ergebnisse empirischer Studien anhand unterschiedlicher Lern- und Problemlöseaufgaben belegen, dass die Eingangsmotivation modellkonform die Mediatoren beeinflusst und sich zum Beispiel die Strategiesystematik und das Flow-Erleben wiederum auf die Leistung auswirken.

5. Geschlechtsunterschiede im Umgang mit dem Computer

Zahlreiche Forschungsarbeiten in der Psychologie thematisieren Geschlechtsunterschiede und -ähnlichkeiten¹¹ im Lernen und in der Leistung (Bussey & Bandura, 1999, 2004; Caplan, Crawford, Hyde & Richardson, 1997; Eagly, Beall & Sternberg, 2004; Eccles, 1994, 2005; Eccles, Adler & Meece, 1984; Halpern, 2000, 2004; Hyde, 2005, 2007; Hyde & Durik, 2005; Hyde & Linn, 2006; Meece & Painter, 2008). Moschner (2010) weist darauf hin, dass in früheren Studien von einer systematischen Benachteiligung von Mädchen und Frauen in Bildungsbereichen ausgegangen wurde. Dies gilt auch für den Bereich der Computerbildung und der Computernutzung (Blossfeld et al., 2009; PISA-Konsortium, 2007). Neuere Studien legen jedoch die Schlussfolgerung nahe, dass sich diese Lern- und Leistungsunterschiede zwischen den Geschlechtern in den letzten Jahren reduziert haben oder gar nicht mehr nachweisbar sind (Hyde, 2005; Meece & Painter, 2008). Meece und Painter (2008) weisen jedoch darauf hin, dass insbesondere in der Motivation und in den Einstellungen von Mädchen und Jungen beziehungsweise Frauen und Männern bedeutsame Unterschiede existieren: „Despite the findings that suggest that gender differences in academic abilities are for the most part insignificant, evidence suggests that gender differences in attitudes, expectancies, motivation, self-beliefs, goals, and self-regulation do exist in specific academic domains“ (S. 339).

Ob Geschlechtsunterschiede persistieren, scheint dabei zwischen den Lebensbereichen zu variieren: „Trotz der Angleichung zwischen den Geschlechtern in Bezug auf das Bildungsniveau (...) lassen sich jedoch in anderen Lebensbereichen weiterhin nachhaltige Ungleichheiten bzw. Unterschiede zwischen den Geschlechtern erkennen. Sie manifestieren sich auf dem Arbeitsmarkt im Rahmen geschlechtsspezifischer Segregation, d. h. in einer unterschiedlichen Berufs- und Ausbildungswahl (...)“ (Blossfeld et al., 2009, S. 24; siehe auch Hyde & Durik, 2005). Insbesondere in den mit dem Computer assoziierten Bereichen Mathematik, Informatik, Naturwissenschaft und Technik (MINT) werden weiterhin Geschlechtsunterschiede in Performanz und Motivation zuungunsten von Mädchen und Frauen berichtet (z.B. Hannover & Bettge, 1993; Kessels, Rau & Hannover, 2006; Stout et al., 2011). Mädchen erreichen in dieser Fächerkombination im Durchschnitt eine geringere Leistung als Jungen. Darüber hinaus drückt sich ein geringeres Interesse der Mädchen an den MINT-Fächern beispielsweise darin aus, dass sie diese weniger häufig als Leistungskurse wählen. Allgemein zeigen sich in einigen Bildungsbereichen mittlere Geschlechtsunterschiede, wohingegen in anderen die Effektivität schwach ist (Hyde & Durik, 2005; OECD, 2010).

¹¹ Die Verwendung des Begriffs „Geschlecht“ bezieht sich hier auf das biologische Geschlecht einer Person und nicht auf das psychologische Geschlecht. Unter „Geschlechtsdifferenzen“ werden in der vorliegenden Arbeit Unterschiede zwischen Männern und Frauen bzw. Mädchen und Jungen verstanden.

In diesem Kapitel soll der reichhaltige Forschungsstand zur „Digital Divide“ (Cooper, 2006) zwischen den Geschlechtern dargestellt werden. Im ersten Abschnitt (5.1.) werden zunächst theoretische Erklärungsansätze für potenzielle Geschlechtsunterschiede bzw. –ähnlichkeiten im Umgang mit dem Computer erläutert. Im darauffolgenden Abschnitt werden Ergebnisse ausgewählter empirischer Studien berichtet (Kapitel 5.2). Im Vordergrund steht die Frage, in welchen Bereichen Geschlechtsunterschiede empirisch beobachtet werden konnten. Hannover und Kessels (2008, S. 116) zufolge unterscheiden sich Frauen und Männer darin, was sie lernen, wie sie lernen und welches Bild sie dabei von sich als Lernenden haben (Hannover & Kessels, 2008, S. 116). Das bedeutet, dass Geschlechtsdifferenzen nicht ausschließlich Unterschiede in der Lernleistung betreffen (Meece & Painter, 2008). So können sich Geschlechtsdifferenzen beim Erlernen eines Computerprogramms in unterschiedlichen computerbezogenen Einstellungen, Erwartungen, Selbstüberzeugungen, Motivationen, Explorationsstrategien sowie in der Computerbildung und der Computernutzung manifestieren (vgl. Beckwith et al., 2006; Dickhäuser & Stiensmeier-Pelster, 2002a, 2002b, 2003; Hirsch, 2002; Meece & Painter, 2008; Richter, Naumann & Horz, 2001; Smith, Morgan & White, 2005). Im Folgenden (Abschnitte 5.2.1 bis 5.2.3) werden die empirischen Ergebnisse, nach den einzelnen Merkmalsbereichen untergliedert, vorgestellt.

5.1 Theoretische Erklärungsansätze für Geschlechtsunterschiede

Um Geschlechtsunterschiede und –ähnlichkeiten im Lern- und Leistungsbereich zu erklären, können verschiedene theoretische Ansätze herangezogen werden. Hyde und Durik (2005) führen drei Rahmenmodelle bzw. –ansätze an:

- 1) Die sozial-kognitive Theorie der Geschlechtsentwicklung und –differenzierung (Bussey & Bandura, 1999, 2004).
- 2) Das Modell der leistungsbezogenen Aufgabenwahlen (Eccles, 2005; Jacobs & Eccles, 2000; siehe auch Dickhäuser & Stiensmeier-Pelster, 2000)
- 3) Ansatz der Bedrohung durch Stereotype (*Stereotype threat*; Aronson & Steele, 2005; Spencer, Steele & Quinn, 1999)

Die Motivation und die Kompetenzerwartungen der Lernenden spielen in den drei genannten Theorieansätzen eine herausgehobene Rolle. Ausgangspunkt aller drei Ansätze ist die Annahme, dass geschlechtsspezifische Muster von Motivation und Kompetenzen im Laufe der psychosozialen Entwicklung eines Menschen entstehen und somit nicht biologisch determiniert sind (Hyde & Durik, 2005, S. 376; siehe auch Bandura, 1986; Rozell & Gardner, 2000). In den ersten beiden genannten Modellen werden Geschlechtsunterschiede in erster Linie auf Personenmerkmale (z.B. Selbstkonzepte, Interesse, Selbstwirksamkeit) zurückgeführt, die im

Laufe der Sozialisation ausgebildet werden. Der dritte Ansatz hebt dagegen die situationsspezifische Variabilität dieser Personenmerkmale hervor.

Sozial-kognitive Theorie. Bussey und Bandura (1999, 2004) nutzen Banduras (1986) sozial-kognitive Theorie, um geschlechtsspezifische Unterschiede und Ähnlichkeiten in der Entwicklung und im Lernen zu erklären. Dieser Theorie zufolge führen Sozialisationsprozesse zur kognitiven Konstruktion von Geschlechtskonzeptionen (z.B. Geschlechtsrollen) und zur Entwicklung geschlechtsspezifischer Verhaltensweisen. Die Autoren berücksichtigen in diesem Zusammenhang kognitive, soziale, affektive und motivationale Prozesse. Im Laufe ihrer Entwicklung erwerben Mädchen und Jungen Einstellungen, Interessen, Kompetenzen Kompetenzerwartungen, die in der Gesellschaft für ihr Geschlecht als angemessen angesehen werden. Mehrere Sozialisationsagenten tragen zu diesem geschlechtsbezogenen Sozialisationsprozess bei (z.B. Eltern, Gleichaltrige; Bandura & Bussey, 1999, S. 676). Die Person selbst nimmt dabei eine aktive Rolle ein: Über Selbstregulationsprozesse (z.B. die Selbstreflexion) kann die Person das eigene Denken und Handeln so ausrichten, dass beides mit den gesellschaftlichen Erwartungen an ihr Geschlecht übereinstimmt.

Im Entwicklungsmodell der triadischen, reziproken Verursachung von Bussey und Bandura (1999, 2004; Hyde & Durik, 2005) stehen Personenfaktoren, Umgebungsfaktoren sowie das Verhalten einer Person in einem wechselseitigen Zusammenhang. Zu den Personenfaktoren zählt auch die Selbstwirksamkeitserwartung (Bandura, 1986; vgl. auch Hyde & Durik, 2005). Geschlechtsspezifische Selbstwirksamkeitserwartungen werden über kognitive Informationsverarbeitungsprozesse und psychosoziale Mechanismen erworben und gefördert. Bussey und Bandura (1999, S. 685f.) nennen unter anderem die folgenden Mechanismen:

- 1) *Lernen am Modell:* Die Person beobachtet andere (z.B. Eltern, Gleichaltrige), die durch eigene Anstrengung bei der Ausführung einer Tätigkeit oder Lernhandlung erfolgreich sind. Mädchen und Jungen orientieren sich dabei eher am (erfolgreichen) Verhalten gleichgeschlechtlicher Rollenvorbilder und ahmen dieses nach.
- 2) *Lernen durch eigene Handlungserfahrung* („Mastery experience“): Eigene Erfolgserfahrungen bzw. das Erleben positiver Konsequenzen eigenen Handelns verstärken optimistische Erwartungen an die eigene Handlungsfähigkeit. Ein Beispiel: Jungen werden für die Ausführung maskulin konnotierter Tätigkeiten verstärkt und erleben sich dadurch als kompetent.
- 3) *Lernen durch Überzeugung und durch direkte Vermittlung:* Geschlechtsspezifische Denk- und Verhaltensweisen werden über verbale Instruktionen einer anderen Person vermittelt. Selbstwirksamkeitserwartungen können durch Überzeugen und Überreden erhöht

werden. Ein Beispiel: Die Aufgabenverteilung in der Gruppenarbeit wird von der Lehrkraft mit unterschiedlichen Fähigkeiten von Mädchen und Jungen begründet.

Den Selbstwirksamkeitserwartungen einer Person werden ein zentraler Einfluss auf deren (geschlechtsspezifische) Verhalten zugeschrieben (Bandura, 1986). Mädchen und Frauen sollten folglich eine niedrige Selbstwirksamkeitserwartung in maskulin konnotierten Tätigkeiten aufweisen (Meece & Painter, 2008, S. 352). Die Selbstwirksamkeitserwartung steht wiederum mit einem niedrigeren Interesse an und einer seltenen Ausführung dieser Tätigkeit in Beziehung. Durch seine Assoziation mit Mathematik und Elektronik ist der Umgang mit dem Computer in der Gesellschaft eher männlich konnotiert (Bussey & Bandura, 1999, S. 693; Hannover & Kessels, 2008). Dies sollte der Theorie zufolge zu einer niedrigeren computerbezogenen Selbstwirksamkeitserwartung der Mädchen führen (vgl. Compeau & Higgins, 1995).

Modell der leistungsbezogenen Aufgabenwahlen. Das erweiterte Erwartungs-mal-Wert-Modell von Eccles (2005; Jacobs & Eccles, 2000) ist ein allgemeines Modell, mit dem auch geschlechtsspezifische Unterschiede im Lern- und Leistungsverhalten am Computer erklärt werden können (Dickhäuser & Stiensmeier-Pelster, 2002a; vgl. auch Eccles, 1994). Geschlechtsdifferenzen werden darin vornehmlich auf unterschiedliche Erfolgserwartungen und Wertüberzeugungen zurückgeführt. Diese beiden kognitiven Variablen sind im Modell die proximalen Determinanten des (Computer-)Verhaltens (Dickhäuser & Stiensmeier-Pelster, 2000, S. 62). Das Modell impliziert, dass Personen dann das Erlernen eines neuen Computerprogramms in Angriff nehmen, wenn die Person erwartet, bei dieser Aufgabe erfolgreich zu sein und zugleich der Aufgabe einen subjektiven Wert beizumessen (Hyde & Durik, 2005, S. 376).

Neben der subjektiven Erwartung und Wertüberzeugungen werden weitere, distale Variablen wie Fähigkeitsselbstkonzepte, Attributionen und Lehrererwartungen im Modell berücksichtigt (Eccles, 2005, S. 106; Dickhäuser & Stiensmeier-Pelster, 2000, S. 60). Diese distalen Variablen beeinflussen die Erfolgserwartungen und subjektiven Wertüberzeugungen einer Person. Die Kompetenzerwartungen und Attributionen selbst werden im Modell sowohl auf bisherige Lern- und Leistungserfahrungen als auch auf kulturelle und soziale Faktoren zurückgeführt (Hyde & Durik, 2005, S. 376). Hierzu zählen kulturelle Geschlechtsrollen und Geschlechtsstereotype. Erstere schreiben den Geschlechtern bestimmte Qualitäten als angemessen zu. Geschlechtsstereotype beziehen sich auf die Konnotation bestimmter Aktivitäten als männlich oder weiblich (Hannover & Kessels, 2008). Dickhäuser und Stiensmeier-Pelster (2000, S. 62) merken an, dass das leistungsbezogene Verhalten in Eccles Modell „eher eine Folge der Interpretation der Realität durch den Schüler/die Schülerin [allgemein: den Lernen-

den] ist als eine Folge der Realität selbst“. Ähnlich wie im Modell von Bussey und Bandura (1999) wird auch hier die aktive Rolle des Lernenden im eigenen Sozialisationsprozess betont.

Bedrohung durch Stereotype (*Stereotype threat*). Neben den beiden bereits beschriebenen Ansätzen wird auch das Konzept des *Stereotype threat* (Steele & Aronson, 1995; siehe auch Aronson & Steele, 2005) zur Erklärung von gruppenbezogenen Leistungseinbußen und –defiziten genutzt. Bei diesem Konzept wird angenommen, dass Leistungsdifferenzen zwischen Männern und Frauen nicht generell auf überdauernde Kompetenzunterschiede zurückführbar sind (Hyde & Durik, 2005, S. 382). Vielmehr machen Aronson und Steele (2005) situationale Kontextfaktoren für eine geringere Leistung weiblicher Personen verantwortlich.

Ausgangspunkt für *Stereotype threat* sind kompetenzbezogene Stereotype: In der Gesellschaft wird unterstellt, dass Personen aufgrund ihrer sozialen Gruppenzugehörigkeit über spezifische Kompetenzen verfügen oder nicht verfügen. *Stereotype threat* tritt insbesondere bei negativen Stereotypen auf, d.h. wenn Personen aufgrund ihrer Gruppenzugehörigkeit Kompetenzen abgesprochen werden. Ein Beispiel hierfür ist das Stereotyp, dass die Mathematikkompetenzen von Mädchen und Frauen geringer ausgeprägt sind (Spencer, Steele & Quinn, 1999). Wird dann in einer Leistungssituation die Gruppenzugehörigkeit der Person (z.B. das Geschlecht) salient gemacht, empfindet dies die Person als Bedrohung: Sie hat die Befürchtung, das negative Stereotyp mit ihrer eigenen Leistung zu bestätigen. Es wird angenommen, dass die Bedrohung durch das Stereotyp insbesondere bei hoher Aufgabenschwierigkeit auftritt und sich ungünstig auf die Leistung auswirkt (Steele & Aronson, 1995). In ihrem Übersichtsbeitrag diskutieren Aronson und Steele (2005, S. 444ff.) mehrere Mediatoren, über welche der Effekt des Stereotyps auf die Leistung vermittelt wird. Empirische Belege werden unter anderem für folgende Mediatorvariablen berichtet: Zustandsangst, Erfolgserwartungen und Selbstregulationskapazitäten. Aronson und Steele (2005, S. 446) schlussfolgern, dass sich *Stereotype threat* über die Erhöhung des *Arousal*s durch Angst bzw. über die Erschöpfung kognitiver Ressourcen oder die Förderung Stereotyp-konformen Verhaltens negativ auf die Leistung auswirkt. Die empirischen Befunde sind hier jedoch nicht konsistent.

5.2 Geschlechtsunterschiede in empirischen Studien

Seit Anfang der 1980er Jahre widmeten sich Wissenschaftler in zahlreichen Forschungsarbeiten der Beziehung zwischen Geschlecht und Computer (z.B. Cooper & Weaver, 2003; Dickhäuser & Stiensmeier-Pelster, 2000, 2002a, 2002b, 2003; Nelson & Cooper, 1997; Shotick & Stephens, 2006; Whitley, 1997). In den folgenden Abschnitten werden Befunde zur Beziehung zwischen Geschlecht und Computer dargestellt, die für die Fragestellung der Arbeit von Bedeutung sind. Geschlechtsspezifische Unterschiede bei der Computernutzung können

sich im Erleben und Verhalten manifestieren (vgl. Rozell & Gardner, 2000). Vor diesem Hintergrund wurden die Abschnitte wie folgt gegliedert: Computernutzung und Computerbildung (Kapitel 5.2.1), Explorationsstrategien am Computer (Kapitel 5.2.2), Computerbezogene Motivation und Affekte (Kapitel 5.2.3), Computerbezogene Geschlechtsstereotype und Performance (Kapitel 5.2.4).

5.2.1 Computernutzung und Computerbildung

Computernutzung. Geschlechtsunterschiede in der Computernutzung beziehen sich auf den Umfang der Computernutzung und auf die Art und Weise des Umgangs mit dem Computer. Auch wenn sich männliche und weibliche Jugendliche seit einigen Jahren nicht mehr wesentlich in den Zugangsmöglichkeiten zu neuen Medien unterscheiden (vgl. PISA-Konsortiums, 2007; Imhof et al., 2007), zeigten sich in Studien weiterhin Geschlechtsdifferenzen in der Art der Computernutzung bzw. der Computerbildung (z.B. Blossfeld et al., 2009; Imhof et al., 2007; Koch, Müller & Sieverding, 2008; Richter, Naumann & Horz, 2001; Young, 2000).

Die Metaanalyse von Whitley (1997) ergab schwache Geschlechtseffekte hinsichtlich der Computernutzung ($d = .31$) und –vorerfahrung ($d = .21$): Männliche Teilnehmer haben danach eine höhere Computervorerfahrung und nutzen den Computer häufiger. Befunde aus der PISA-Studie aus dem Jahr 2006 verdeutlichen, dass Jungen eine längere Erfahrung mit dem Computer aufweisen als Mädchen (PISA-Konsortium, 2007, S. 291; vgl. auch Richter, Naumann & Horz, 2001). Auch in der Schule nutzen Mädchen den Computer etwas seltener als Jungen (z.B. indem sie seltener Kurse aus dem Bereich Computertechnologie wählen; Cooper, 2006). Allerdings gleichen sich die Computernutzungshäufigkeiten von Mädchen und Jungen insbesondere in der Schule allmählich an (PISA-Konsortium, 2007).

Auch in der Freizeit zeigen sich unterschiedliche Nutzungsgewohnheiten: „Nach ihren Computernutzungsgewohnheiten gefragt, gehörten in der PISA-Studie aus dem Jahre 2003 61 % der Jungen, aber nur 48 % der Mädchen zum Typus des Enthusiasten, d.h. zu den Jugendlichen, die sehr oft und sehr kompetent in ihrer Freizeit den Computer nutzen“ (Hannover & Kessels, 2008, S. 118; vgl. Meelissen & Drent, 2008). Jungen setzen den Computer für eine größere Bandbreite an Tätigkeiten ein (Blossfeld et al., 2009, S. 104). Einen frühen, außerschulischen Zugang zum Computer finden Jungen insbesondere über Computerspiele (Cooper, 2006). Schaumburg (2004) bemerkt, dass Mädchen dagegen eher solche Computeranwendungen nutzen, die keine hohe computertechnische Kompetenz erfordern (z.B. Textverarbeitungsprogramme). Jungen dagegen nutzen den Computer auch für komplexere Anwendungen

(z.B. Erstellen von Grafiken in Bildverarbeitungsprogrammen, Datentransfer, Internetrecherche, Programmieren; vgl. Cooper & Weaver, 2003).

In Studierendenstichproben zeigen sich ähnliche Tendenzen. Männliche und weibliche Studierende unterscheiden sich nicht wesentlich im Computerzugang (Imhof et al., 2007). Darüber hinaus nutzen sie in vergleichbarem Umfang den Computer im Rahmen ihres Studiums. Geschlechtsdifferenzen sind jedoch weiterhin in der Art und Häufigkeit der Nutzung in der Freizeit bemerkbar (Imhof et al., 2007; Jackson et al., 2001; Meelissen & Drent, 2008; Middendorf, 2002). Weibliche Studierende nutzen den Computer häufiger zur asynchronen, interpersonellen Kommunikation (z.B. zum Versenden von E-Mails), wohingegen ihre männlichen Kommilitonen häufiger chatten, Online-Spiele nutzen und im Internet nach Informationen suchen (Imhof et al., 2007; Jackson et al., 2001). Middendorf (2002) schlussfolgert, dass weibliche Studierende den Computer eher aufgabenorientiert nutzen, wohingegen ihre männlichen Kommilitonen den Computer gerne erkunden und für vielfältige Zwecke auch außerhalb des Studiums einsetzen.

Computerbildung. Die Computerbildung wird als die „Gesamtheit prozeduraler und deklarativer Wissensbestände verstanden, die einem Individuum einen kompetenten Umgang mit dem Computer und damit eine individuell wie sozial erfolgreiche Teilhabe an Informationstechnologien ermöglichen“ (Richter et al., 2010, S. 24). Die Computerbildung lässt sich unter anderem in deklaratives und prozedurales Computerwissen unterteilen. Zum deklarativen Wissen zählt insbesondere computertechnologisches Wissen aus dem Bereich Hardware und Betriebssystem. Das prozedurale Computerwissen bezieht sich auf Kenntnisse der Bedienung des Computers sowie einzelner Computeranwendungen. Allgemein ist die Computerbildung positiv mit der Performanz am Computer korreliert (z.B. Rozell & Gardner, 2000; Imhof et al., 2007).

Eine von Ellis et al. (2008) vorgelegte umfangreiche Übersicht empirischer Studien verdeutlicht, dass Männer im Vergleich zu Frauen über eine höhere Computerbildung verfügen. Geschlechtsunterschiede lassen sich bereits im Kindesalter nachweisen: Jungen übertreffen ihre Mitschülerinnen hinsichtlich des Computerwissens (Blossfeld et al., 2009, S. 104). Die Studienergebnisse von Schaumburg (2004) an Schülern der 9. Klasse bestätigen die Annahme, dass Jungen den Mädchen insbesondere im deklarativen Computerwissen überlegen sind. Mädchen besaßen in der Studie geringere Hardwarekenntnisse als Jungen. Hinsichtlich der Kenntnisse von Standard-Bürosoftware wie Textverarbeitungsprogrammen unterschied sich ihr deklaratives Wissen jedoch nicht statistisch signifikant von dem der Jungen. Richter, Naumann & Horz (2001) berichten für eine Studierendenstichprobe mittlere bis hohe Geschlechtseffekte hinsichtlich der Computerbildung. Männliche Studierende waren den weibli-

chen sowohl im deklarativen als auch im prozeduralen Computerwissen überlegen. Darüber hinaus waren Männer mit einer größeren Anzahl an Computeranwendungen vertraut und gaben ein höheres Sicherheitsgefühl im Umgang mit dem Computer an. Auch unter Kontrolle der bisherigen Computernutzung zeigten sich statistisch signifikante Geschlechtsdifferenzen mit mittleren Effektstärken. Auch in neueren Studien erzielten Männer in praktischen Computerwissenstests eine höhere Punktzahl als Frauen (Koch et al., 2008).

Nach Richter, Naumann & Horz (2010) stellen computerbezogene Einstellungen einen Teilaspekt der Computerbildung dar. Computerbezogene Einstellungen sind als wertende Überzeugungen konzeptualisiert und können die Computernutzung günstig oder ungünstig beeinflussen (Naumann, Richter & Groeben, 2001; Richter, Naumann & Groeben, 2001). Positive Einstellungen zum Computer standen bei Schülerinnen und Schülern zum Beispiel mit einer häufigen Computernutzung außerhalb der Schule in Zusammenhang (Meelissen & Drent, 2008). Negative Einstellungen dagegen können in einer geringeren computerbezogenen Selbstwirksamkeit und in Vermeidungsverhalten gegenüber dem Computer resultieren (Rozell & Gardner, 2000).

In mehreren Studien wird berichtet, dass positive Einstellungen zum Computer bei Mädchen und Frauen in einem geringeren Maße ausgeprägt sind (z.B. Cooper, 2006; Jackson et al., 2001; Meelissen & Drent, 2008; Schaumburg, 2004; Whitley, 1997; mit Ausnahme z.B. von Rozell & Gardner, 2000). Shotick und Stephens (2006) machen jedoch auf Studien aufmerksam, in denen keine Geschlechtseffekte gefunden wurden. Insbesondere den technologischen Aspekten des Computers stehen weibliche Personen kritisch gegenüber (Ellis, 2008, S. 449). Die Einstellung, dass der Computer eine unbeeinflussbare Maschine ist, wird von weiblichen Studierenden häufiger vertreten als von männlichen (Richter, Naumann & Horz, 2001). Darüber hinaus zeigten sich weibliche Studierende weniger überzeugt von der Nützlichkeit eigener Computerfähigkeiten (Jackson et al., 2001). Empirische Studien deuten darauf hin, dass die wahrgenommene Ermunterung durch Eltern zur Computernutzung sowie hohe Computerkompetenzen weiblicher Rollenvorbilder einen positiven Einfluss auf die Einstellungen der Mädchen ausüben und so zu einer Minderung der bisherigen Unterschiede beitragen können (Meelissen & Drent, 2008).

5.2.2 Explorationsstrategien am Computer

Geschlechtsunterschiede können sich auch im konkreten Umgang mit dem Computer manifestieren. Es ist anzunehmen, dass sich Frauen und Männer in der Art und Weise unterscheiden, in der sie sich dem Computer nähern. Dies ist wiederum von verschiedenen Sichtweisen auf den Computer abhängig. Hierzu bemerkt Schorb (1990, S. 13, zitiert nach Hirsch,

2002, S. 38): „Im Umgang von Frauen und Mädchen mit dem Computer tritt der Werkzeugcharakter der Maschine deutlich in den Vordergrund: was Frauen brauchen, eignen sie sich schnell und kompetent an – aber mehr nicht“. Cooper und Stone (1996, zitiert nach Rozell & Gardner, 2000, S. 204) berichten, dass männliche Personen demgegenüber einen „persönlichen, vertrauten“ Umgang mit dem Computer pflegen anstatt den Computer auf seine Werkzeugfunktion zu reduzieren.

Von besonderem Interesse in dieser Arbeit sind die Strategien, die Frauen und Männer zum Erlernen eines unbekanntes Computerprogramms nutzen. Hirsch zufolge (2002) neigen Jungen stärker als Mädchen dazu, nach einer „trial-and-error“ Methode zu lernen (vgl. Dörner, 1980, 2006; Dörner & Schaub, 1994). Das „männliche Lernen“ sei ein erfahrungsbasiertes Lernen, das sich ohne anfängliche Regelkenntnis oder Regelanwendung vollziehe. Beobachtungen von Lehrkräften unterstreichen, dass der Umgang von Jungen mit dem Computer selbstbewusster und experimentierfreudiger sei (Schaumburg, 2004). Im Gegensatz zu den Jungen zeigten sich Mädchen weniger experimentierfreudig, eher vorsichtig und stärker pragmatisch im Umgang mit Geräten (z.B. dem Computer; Jones et al., 2000; Schaumburg, 2004): Mädchen orientierten sich stärker an den Anweisungen der Lehrkräfte und suchten häufiger Hilfe anstelle eines selbstgesteuerten, spielerischen Umgangs mit den Geräten. Eine stärker lehrerzentrierte Lernumgebung trug in der Studie von Meelissen und Drent (2008) zu positiveren Einstellungen der Mädchen gegenüber dem Computer bei.

Das eigenständige, spielerische Erkunden von Geräten wie dem Computer wird unter anderem von Beckwith et al. (2006) als *Tinkering* bezeichnet. Es ist mit positiven Auswirkungen auf Lern- und Problemlöseprozesse assoziiert (vgl. Greif, 1990, 1996): *Tinkering* „is often generative, allows for risk-taking, is autonomous, often involves the hands, is pleasurable, is open-ended, and intrinsically motivated“ (Jones et al., 2000, S. 761). Durch diesen spielerischen Umgang lernen die Computernutzer mehr Funktionen des Computerprogramms kennen. *Tinkering* kann in verschiedenen Formen auftreten. Ist das *Tinkering* eher planlos und die Handlungsabfolge zufällig, ähnelt es einer „Versuch-und-Irrtum“-Explorationsstrategie (van der Linden et al., 2001; vgl. Kapitel 3.1.2). Demgegenüber ist das zielgerichtete *Tinkering* mit der systematischen Explorationsstrategie vergleichbar. Jones et al. (2000) fanden heraus, dass Jungen im Vergleich zu Mädchen beide Formen des *Tinkering* häufiger zum Erkunden naturwissenschaftlicher Geräte und Werkzeuge einsetzten.

Diese Divergenzen im Umgang mit dem Computer wurden auch im Erwachsenenalter beobachtet: Männer zeigen häufiger ein nicht-zweckgerichtetes *Tinkering* als Frauen. Dieser Effekt wurde allerdings nur unter der Bedingung beobachtet, in der das *Tinkering* mit niedrigen Kosten bei der Programmnutzung verbunden war (z.B. keine Restriktionen bei der Anzahl

erlaubter Mausclicks; vgl. Trudel & Payne, 1995). Entgegen der Erwartungen hatte das nicht zweckgerichtete *Tinkering* der Männer einen negativen Effekt auf die Anzahl behobener Fehler (Beckwith et al., 2006). Das *Tinkering* der Frauen war dagegen positiv mit diesem Leistungsmaß verbunden. Beckwith et al. (2006) schlussfolgern, dass sich das wiederholte Versuch-und-Irrtum-*Tinkering* der Männer kontraproduktiv auswirke, wohingegen das vorsichtige, zweckgerichtete *Tinkering* der Frauen zu einer besseren Leistung und einem größeren Wissenszuwachs führe.

5.2.3 Computerbezogene Motivation und Affekte

Computerbezogene Selbstwirksamkeit. Die computerbezogene Selbstwirksamkeit wird als bedeutende Determinante für die Nutzungshäufigkeit und die Performanz am Computer gewertet (vgl. Bandura, 1997; Cassidy & Eachus, 2002). Compeau und Higgins (1995, S. 192) definieren *Computer Self-Efficacy* folgendermaßen: „Computer self-efficacy (...) refers to a judgment of one’s capability to use a computer“. Allgemein geht die computerbezogene Selbstwirksamkeit mit effizienteren Explorations- und Navigationsstrategien, einer höheren Persistenz bei Hindernissen, einer höheren intrinsischen Motivation sowie einer niedrigen Computerängstlichkeit einher (Compeau & Higgins, 1995; Deng, Doll & Truong, 2004; Imhof et al., 2007).

In mehreren Studien werden höhere Selbstwirksamkeitserwartungen bei Männern im Vergleich zu Frauen berichtet (vgl. Pajares, 2002). Diese Diskrepanz in den Selbstüberzeugungen zeigte sich auch bei einem gleich hohen Kompetenzniveau der Geschlechter. Whitleys (1997) Metaanalyse zufolge stufen Männer ihre computerbezogene Selbstwirksamkeit höher ein als Frauen ($d = .41$). Geschlechtseffekte zeigen sich insbesondere in der Sekundarstufe und im Erwachsenenalter (vgl. Cassidy & Eachus, 2002; Jackson et al., 2001; Meelissen & Drent, 2008; Schaumburg, 2004). Auch unter Kontrolle von Computernutzungshäufigkeit und Computererfahrung ergeben sich schwache bis mittlere Geschlechtseffekte zuungunsten der Mädchen in den Bereichen computerbezogene Selbstwirksamkeit und computerbezogenes Interesse (PISA-Konsortium, 2007; vgl. Blossfeld et al., 2009; Koch et al., 2008).

Shotick und Stephens (2006) untersuchten Kompetenzerwartungen von US-Studierenden bezüglich spezifischer Computeranwendungen. Männer stufen ihre Kompetenzen insbesondere in den Bereichen höher ein, die stärker technisch oder mathematisch ausgerichtet waren. Zu diesen Anwendungen zählten unter anderem Tabellenkalkulations- und Statistikprogramme. Ähnliche Ergebnisse erbrachte eine Studie an schottischen und rumänischen Studierenden (Durndell, Haag & Laithwaite, 2000): Männliche Studierende gaben insbesondere für weiterführende und komplexere Anwendungen eine höhere Selbstwirksamkeitserwar-

tung an (z.B. Gründe finden, weshalb ein Programm auf einem PC nicht funktioniert). Die Effektstärke von Geschlechtsunterschieden scheint folglich zwischen Computeranwendungen zu variieren.

Erfolgserwartungen und subjektiver Wert. Auf der Basis einer adaptierten Version des Modells der leistungsbezogenen Aufgabenwahl (Eccles, 2005; Jacobs & Eccles, 2000) untersuchten Dickhäuser und Stiensmeier-Pelster (2002a, 2003), ob sich Männer und Frauen im subjektiven Wert und der Erfolgserwartung der Computernutzung unterscheiden. In einer Studierendenstichprobe zeigte sich, dass Frauen ihre Erfolgserwartung beim Arbeiten mit dem Computer als niedriger einstufen als Männer (Dickhäuser & Stiensmeier-Pelster, 2002a). Anders als bei der Erfolgserwartung fanden die Autoren hinsichtlich des subjektiven Werts der Computernutzung keine Hinweise auf Geschlechtsunterschiede.

In einer Schülerstichprobe gingen Dickhäuser und Stiensmeier-Pelster (2003) der Frage nach, weshalb männliche Schüler im Alter zwischen 10 und 16 Jahren häufiger Computerkurse wählten als ihre gleichaltrigen Mitschülerinnen. In dieser Stichprobe trugen sowohl die Erfolgserwartung als auch der subjektive Wert der Computernutzung (z.B. die Nützlichkeit der zu erwerbenden Computerkenntnisse; vgl. Eccles & Wigfield, 1995) zur Erklärung bei Mädchen stuften sowohl ihre Erfolgserwartung als auch den subjektiven Wert des Computerkurses geringer ein. Entsprechend niedriger lag in Studien auch die bereichsspezifische Identifikation von Mädchen und Frauen mit der Computertechnologie (Smith et al., 2005). Dickhäuser und Stiensmeier-Pelster (2003) konnten zeigen, dass Jungen bei ihren Eltern einen höheren subjektiven Wert bezüglich des Umgangs mit dem Computer berichteten als dies bei Mädchen der Fall war. Die Erfolgserwartung der Eltern wurde von Jungen und Mädchen dagegen ähnlich wahrgenommen. Erfolgserwartungen und Werte der Eltern sollten nach Eccles Modell die Ausprägungen bei Töchtern und Söhnen beeinflussen.

Computerängstlichkeit. Geschlechtsunterschiede in der Computernutzung lassen sich auch auf affektive Aspekte wie die Computerängstlichkeit („Computer Anxiety“) zurückzuführen (Todman & Day, 2006; vgl. auch Brosnan, 1998). Die Computerängstlichkeit wird einerseits als eine *State*-Furcht vor der Interaktion mit dem Computer oder beim Denken an den Computer beschrieben (Chua, Chen & Wong, 1999). In anderen Studien wurde sie als zeitlich überdauerndes Persönlichkeitsmerkmal (*Trait*) konzeptualisiert, das durch computerbezogene Angstgefühle und Besorgtheitskognitionen gekennzeichnet ist (Richter et al., 2010). Die Computerängstlichkeit war in Studien negativ mit der Computervorerfahrung, der Computer Self-Efficacy sowie der Leistung am Computer assoziiert (Brosnan, 1998; Compeau & Higgins, 1995).

Bei Frauen ist die Computerängstlichkeit tendenziell höher ausgeprägt als bei Männern: „Across a broad spectrum of activities, girls and women report that computers are not the source of fun and amusement but rather of anxiety“ (Cooper, 2006, S. 321; vgl. auch Jackson et al., 2001). In der Metaanalyse von Whitley (1997) zeigt sich, dass Frauen weniger starke positive Affekte bezüglich des Computers berichten als Männer ($d = .26$). Demgegenüber gaben weibliche Psychologie-Studierende im Vergleich zu ihren männlichen Kommilitonen ein größeres Unbehagen an, wenn sie statistische Aufgaben zuhause mit dem Computer zu lösen sollten (Cooper & Weaver, 2003). Eine Studie an Oberstufenschülern und –schülerinnen ergab, dass sich die höhere Computerängstlichkeit bei Mädchen auf zwei Aspekte bezieht: Erstens fürchteten sich die Mädchen nicht in der Lage zu sein, neue Computeranwendungen zu erlernen. Zweitens hatten sie im Vergleich zu den Jungen mehr Angst davor, den Computer zu beschädigen (Baloglu & Çevik, 2008). Todman und Day (2006, S. 858) nehmen allerdings an, dass die Computerängstlichkeit nicht primär durch das biologische Geschlecht beeinflusst ist, sondern durch das psychologische: „In a culture in which the computer is viewed as a masculine domain, high scores on femininity will be associated with high computer anxiety, regardless of biological gender“. Ergebnisse einer ersten empirischen Studie der Autoren auf der Basis partieller Korrelationsanalysen unterstützen diese Annahme (vgl. auch Brosnan, 1998).

Attributionsstil. Ursachenzuschreibungen beeinflussen „Erfolgserwartungen, Emotionen und Ausdauer“ (Dickhäuser & Stiensmeier-Pelster, 2002b, S. 44) in Lern- und Leistungssituationen. Eine Attribution von Misserfolg auf stabile und globale Ursachen wirkt sich beispielsweise ungünstig auf die Motivation der Lernenden aus. Mehrere empirische Studien liefern Hinweise darauf, dass weibliche und männliche Personen ihre Erfolge bzw. Misserfolge im Umgang mit dem Computer auf unterschiedliche Ursachen zurückführen (mit Ausnahme z.B. von Rozell & Gardner, 2000). Weibliche Personen scheinen dabei ein Attributionsmuster zu zeigen, welches mit erlernter Hilflosigkeit assoziiert ist (Cooper, 2006; Dweck, 1975): „Dieses Muster ist dadurch gekennzeichnet, dass Erfolg eher auf externale, variable Ursachen wie Glück zurückgeführt wird, Misserfolg dagegen eher auf internale, stabile Ursachen wie etwas mangelnde Begabung“ (Dickhäuser & Stiensmeier-Pelster, 2000, S. 54).

In der Studie von Nelson und Cooper (1997) führten Jungen Misserfolge am Computer zum Beispiel auf mangelnde Anstrengung oder eine fehlerhafte Diskette zurück. Ein erfolgreiches Arbeiten mit dem Computer schrieben Jungen dagegen ihren Fähigkeiten zu (Cooper, 2006). Mädchen dagegen erklärten sich ihren Erfolg am Computer mit Hilfe variabler Faktoren (z.B. „einfaches Programm“). Der Attributionsstil der Mädchen ging dabei mit einem geringeren Enthusiasmus bei der Computernutzung einher. Dickhäuser und Stiens-

meier-Pelster (2002a) gelangen bei Studierenden zu ähnlichen Attributionsmustern. Insbesondere die männlichen Studierenden zeigten einen günstigen Attributionsstil (z.B. stabil, global, internal, unkontrollierbar für Erfolg). Dieser wirkte sich wiederum positiv auf die Höhe des computerbezogenen Selbstkonzepts aus. In einer weiteren Studie zeigten weibliche Studierende allerdings ein eher günstiges Attributionsmuster für Misserfolge (instabil, internal, kontrollierbar), das nicht mit ihrer geringeren Computernutzung und niedrigen Erfolgserwartung korrespondierte (Dickhäuser & Stiensmeier-Pelster, 2002b). Dies führen die Autoren auf die spezifischen Gründe zurück, die Männer und Frauen als Ursache für den Misserfolg am Computer wahrnehmen: Männer attribuierten eher (unkontrollierbar) auf eine defekte Computerhardware („defekte Diskette“), wohingegen Frauen stärker (kontrollierbar) auf mangelnde Kenntnisse attribuierten. Die Autoren vermuten, dass die vermeintlich günstige Attribution der Frauen auf mangelnde Kenntnisse mit einem erhöhten Schamgefühl und letztlich Vermeidungsverhalten bezüglich des Computers einhergehe.

5.2.4 Computerbezogene Geschlechtsstereotype und Performanz

Der Computer ist in europäischen und nordamerikanischen Gesellschaften mehrheitlich mit dem männlichen Geschlecht assoziiert (Cooper & Weaver, 2003; Smith et al., 2005; Todman & Day, 2006). Die gesellschaftlichen Erwartungen hinsichtlich der geschlechtsspezifischen Angemessenheit von Aktivitäten, beeinflussen auch die Performanz am Computer (Cooper, 2006). Computerbezogene Geschlechtsstereotype und entsprechende Geschlechtsunterschiede im Interesse an Computer können bereits früh in der Entwicklung beobachtet werden (Mercier, Barron & O'Connor, 2006). Dabei zeigte sich bei Jungen eine stärkere Geschlechtsstereotypisierung des Computers als bei Mädchen (Meelissen & Drent, 2008; Young, 2000). Die Metaanalyse von Lytton und Romney (1991) ergab, dass Mädchen und Jungen von ihren Eltern anderen Lernangebote offeriert werden. Danach ermutigen Eltern ihre Kinder insbesondere zu geschlechtstypischen Aktivitäten und Vorlieben. Als Beispiele nennen Hannover und Kessels (2008, S. 116) das Kaufen von Technikspielzeug für Jungen und das Angebot zu Ballettunterricht für Mädchen.

In der Metaanalyse von Whitley (1997) zeigte sich, dass Schüler in der Sekundarstufe den Computer als eher für Jungen angemessen einstufen ($d = .54$). Mercier et al. (2006) forderten Schüler der 6. und 8. Klasse auf, typische Computerexperten zu malen. Entsprechend des Geschlechtsstereotyps zeichneten sowohl Schülerinnen als auch Schüler mehrheitlich männliche Personen. Ob sich die Schüler und Schülerinnen selbst als Computerexperten einstufen, war jedoch nicht in erster Linie vom Geschlecht des jeweiligen Kindes abhängig, sondern

wurde viel mehr von anderen Variablen bestimmt (z.B. Computervorerfahrung, Vertrauen in eigene Fähigkeiten etc.).

Cooper (2006) hebt hervor, dass das kompetenzbezogene Geschlechtsstereotyp bezüglich des Computers die Computerleistung der Frauen verringern könne. In einer Studie zu *Stereotype threat* zeigten Schülerinnen eine geringere Leistung beim Erstellen einer grafischen Abbildung mit einem Präsentationsprogramm, wenn sie bezüglich ihres Geschlechts geprimt wurden (Cooper, 2006). Koch et al. (2008) erforschten anhand einer konkreten Computeraufgabe, ob Frauen unter *Stereotype threat* Bedingungen ihren Misserfolg auf andere Ursachen zurückführten als Frauen in einer Kontrollbedingung ohne *Stereotype threat*. In der *Stereotype threat* Bedingung wurden weibliche und männliche Teilnehmer auf das Stereotyp aufmerksam gemacht, dass Frauen bei der Aufgabe bisher schlechtere Leistungen gezeigt hatten als Männer. Entsprechend der Erwartungen attribuierten Frauen in der *Stereotype threat* Bedingung ihren Misserfolg stärker internal (mangelnde Fähigkeiten) als Männer. In der Kontrollbedingung zeigten sich dagegen keine Geschlechtsunterschiede in der Attribution. In einer experimentellen Studie von Sieverding und Koch (2009) beobachteten Studienteilnehmer eine männliche oder weibliche Zielperson, die eine Aufgabe mit einem Textverarbeitungsprogramm lösen sollte. Die Ergebnisse zeigen, dass die Studienteilnehmer die Computerkompetenz der Zielperson sowie die Ursachenzuschreibung unabhängig von deren Geschlecht einstufen. In der Selbstbeurteilung dagegen schätzen sich weibliche Studienteilnehmer bei der Aufgabe als signifikant weniger kompetent ein als männliche Studienteilnehmer. Gleichzeitig stuften sowohl Männer als auch Frauen ihre Kompetenzen höher ein, wenn sie eine weibliche Zielperson statt einer männlichen Zielperson beobachtet hatten. Die Autorinnen schlussfolgerten daraufhin, dass das Geschlechtsstereotyp zu Computerkompetenzen weiterhin existiere, jedoch stärker auf die Beurteilung eigener Fähigkeiten bezogen sei statt auf die Beurteilung anderer Personen.

Auch in Studien, in denen kein *Stereotype threat* induziert wurde, zeigten sich Geschlechtsunterschiede in der Performanz. Allerdings ist die Befundlage zur Beziehung zwischen Geschlecht und Computerperformanz inkonsistent. Beispielsweise gaben Roy, Taylor und Chi (2004) Schülern die Aufgabe, online und offline nach Informationen zu suchen. Alle Schüler konnten ihr Wissen steigern; die Schüler, die im Web anstelle in der Schulbibliothek recherchierten, erweiterten ihr Wissen dabei am stärksten. Allerdings zeigte sich, dass die Jungen im Vergleich zu den Mädchen online deutlich mehr zielführende Informationen zur Lösung der Aufgabe fanden. Eine Studie an Studierenden ergab ebenfalls, dass männliche Studierenden den weiblichen Studierenden beim Erstellen einer Präsentation überlegen waren (Imhof et al., 2007). Demgegenüber fanden Shapka und Ferrari (2003) bei einer Formatierungsaufgabe am Computer keine Geschlechtseffekte.

Geschlechtseffekte bei der Leistung am Computer können dabei durch die Konzeption der eingesetzten Software begünstigt werden. Littleton, Light, Joiner, Messer und Barnes (1998) konnten zeigen, dass sich Mädchen und Jungen in unterschiedlicher Weise von Computersoftware angesprochen fühlen. In ihrer Studie hatte dies Auswirkungen auf die Leistung: War die Gestaltung der Software eher geschlechtsneutral, beispielsweise wenn sich Mädchen und Jungen gleichermaßen mit den Spielfiguren identifizieren konnten, dann zeigten Mädchen und Jungen ähnliche Leistungen am Computer (vgl. Cooper & Weaver, 2003).

5.2.5 Zusammenfassung

Die Hypothese der *Digital Divide* besagt, dass Frauen bezüglich des Computerzugangs sowie der Computernutzung benachteiligt sind. Diese Annahme wurde in mehreren Studien und in Bezug auf unterschiedliche Korrelate der Computernutzung empirisch überprüft. Die aufgeführten Studien liefern Hinweise darauf, dass sich Männer und Frauen im Umgang mit dem Computer weiterhin in charakteristischer Weise unterscheiden. Obwohl sich die Zugangsmöglichkeiten und die Computerausstattung zwischen den Geschlechtern seit einiger Zeit angleichen, zeigen sich weiterhin Differenzen in der Art der Computernutzung. Insbesondere im Sekundarbereich scheinen sich die Disparitäten zwischen den Geschlechtern hinsichtlich der schulischen Computernutzung in der Schule und der Computernutzung in der Freizeit zu verstärken. Jungen nutzen den Computer häufiger in der Freizeit und sind mit einer größeren Bandbreite an Computeranwendungen vertraut. Dieser Befund wurde für das Studierendentalter repliziert.

Des Weiteren persistieren Geschlechtsdifferenzen in motivationalen, affektiven und kognitiven Merkmalen. Demzufolge weisen Frauen im Vergleich zu Männern im Durchschnitt eine geringere Computerselbstwirksamkeitserwartung und ein geringeres computerbezogenes Interesse auf. Frauen geben im Allgemeinen eine höhere Computerängstlichkeit an; diese bezieht sich insbesondere auf das Erlernen neuer Anwendungen. Darüber hinaus lässt sich bei ihnen ein ungünstigerer Attributionsstil bei Erfolg und Misserfolg am Computer nachweisen.

Weibliche Personen setzen den Computer eher zweckgerichtet als Werkzeug ein, um ein konkretes Problem zu lösen. Sie orientieren sich dabei stark an Anleitungen und äußern ein stärkeres Bedürfnis nach Hilfestellung. Im Gegensatz dazu erkunden Männer den Computer spielerisch und selbstgesteuert. Dieses „männliche“ Explorationsverhalten begünstigt das Entdecken von Funktionen, die nicht unmittelbar mit der Aufgabenbearbeitung assoziiert sind. Allerdings wird für Männer auch in stärkerem Maße ein Explorationsverhalten berichtet, das mit der „Versuch-und-Irrtum“-Explorationsstrategie vergleichbar ist. Diese stand in Studien mit einer geringeren Leistung in Bezug.

Bereits früh zeigen sich in den Studien Geschlechtsunterschiede bezüglich der Geschlechtsstereotypisierung des Computers: Jungen und teilweise auch Mädchen assoziieren den Computer stärker mit dem männlichen Geschlecht. Obwohl bei Frauen das Geschlechtsstereotyp weniger stark ausgeprägt ist bzw. Frauen es sogar ablehnen, zeigt sich unter *Stereotyp threat* Bedingungen, dass das Stereotyp die Computerleistungen von Mädchen und Frauen negativ beeinflusst. Allgemein sind höhere Computerleistungen positiv mit internalen Attributionen bei Erfolg und positiven Affekten hinsichtlich der eigenen Leistung am Computer assoziiert (Rozell & Gardner, 2000, S. 215). Personen mit einer höheren Computerbildung und einer hohen Ausprägung der Computerselbstwirksamkeit erzielen am Computer ebenfalls höhere Leistungsniveaus. Zusammenfassend sprechen die zitierten Befunde dafür, dass Mädchen und Frauen hinsichtlich relevanter computerbezogener Variablen weiterhin benachteiligt sind.

6. Forschungsfragen und Hypothesen

Mehrere Studien liefern Anhaltspunkte für die Annahme, dass sich Männer und Frauen in ihrer Performanz am Computer unterscheiden (s. Kapitel 5). Gleichzeitig werden in der Literatur Geschlechtsdifferenzen in Motivation und Affekten beim Umgang mit dem Computer berichtet. Die bisherigen Forschungsergebnisse legen die Schlussfolgerung nahe, dass sich Männer und Frauen einer unbekanntem Computeranwendung in unterschiedlicher Weise nähern.

Trotz der vielfältigen Erkenntnisse, welche in den letzten Jahren zu Geschlechtsunterschieden am Computer zusammengetragen wurden, zeigen sich weiterhin bedeutsame Forschungslücken: Eine systematische Untersuchung der geschlechtsspezifischen Nutzung systematischer und unsystematischer Explorationsstrategien fehlt bisher. Von herausgehobener pädagogisch-psychologischer und gesellschaftlicher Relevanz erscheinen insbesondere Geschlechtsdifferenzen in der Anwendung einer beruflich genutzten Software (z.B. eines Statistikprogramms in den empirischen Sozialwissenschaften). Unterschiedliche Herangehensweisen könnten den Lernprozess vor allem bei weiblichen Personen in Schule, Studium und Beruf ungünstig beeinflussen. Eine weitere Grenze des aktuellen Forschungsstands betrifft die Rolle der Motivation im Lernprozess: In den bisherigen Studien wurde die Motivation mehrheitlich als *Trait* erfasst und vor oder nach dem Lern- und Arbeitsprozess am Computer gemessen. Veränderungen der Motivation im Vollzug der konkreten Lernhandlungen können unter diesen Bedingungen nicht explizit nachvollzogen werden (vgl. Bandura, 1997; Vollmeyer & Rheinberg, 1998).

Vor diesem Hintergrund kulminiert das Forschungsthema der vorliegenden Arbeit in der Frage, ob Geschlechtsdifferenzen in der Performanz am Computer prozessnah auf Unterschiede in Motivation, Affekten und Explorationsstrategien zurückführbar sind. Zur Erklärung potenzieller Leistungsdifferenzen wird das kognitiv-motivationale Prozessmodell von Vollmeyer und Rheinberg (1998, 2000, 2006) herangezogen. Dieses Modell erlaubt es, Leistungsunterschiede über die aktuelle Motivation, das Flow-Erleben und die Strategiesystematik während des Lernprozesses zu erklären (siehe Abschnitt 4.3).

Das Forschungsthema der vorliegenden Arbeit wird in vier Forschungsfragen untergliedert. Die *erste Forschungsfrage* bezieht sich auf die Erfassung unterschiedlicher Explorationsstrategien am Computer: Lassen sich mithilfe eines Beobachtungssystems unterschiedliche Explorationsstrategien valide und reliabel voneinander abgrenzen? Aufbauend auf den Arbeiten von van der Linden et al. (2001, 2003) soll ein Beobachtungssystem entwickelt werden, mit

dessen Hilfe die Explorationsstrategien von Männern und Frauen am Computer operationalisiert werden können. Die *zweite Forschungsfrage* richtet sich auf die Beschreibung und Erklärung potenzieller Geschlechtsunterschiede in der Computerleistung mit Hilfe des kognitiv-motivationalen Prozessmodells: Inwiefern können mögliche Geschlechtsunterschiede in der Performanz beim Erlernen eines Computerprogramms auf motivationale und kognitive Variablen vor und während des Lernprozesses zurückgeführt werden? Die *dritte Forschungsfrage* ist der Entwicklung und Wechselwirkung von Motivation und Strategiesystematik im Verlauf des Lernprozesses gewidmet: Beeinflussen sich die motivationalen und kognitiven Prozessvariablen während des Lernens gegenseitig? Welchen Anteil haben sie bei der Erklärung von Leistungsdifferenzen? Die *vierte Forschungsfrage* bezieht sich auf die Klassifikation der von Computernutzern genutzten optimalen (systematischen) und suboptimalen (unsystematischen) Explorationsstrategien: Können anhand von charakteristischen Mustern der genutzten Explorationsstrategien Gruppen von Computernutzern voneinander abgegrenzt werden?

Im folgenden Abschnitt 6.1 werden die Forschungsfragen im Einzelnen genauer erläutert. Auf der Basis der oben dargelegten theoretischen Annahmen und Befunde wurden Hypothesen formuliert, um die jeweilige Forschungsfrage empirisch überprüfen zu können. Die Hypothesen sind in den entsprechenden Kapiteln zu den einzelnen Forschungsfragen aufgeführt.

6.1 Handlungsnahe Erfassung von Explorationsstrategien (Forschungsfrage 1)

Wie lassen sich unterschiedliche Herangehensweisen von Männern und Frauen beim Erlernen eines spezifischen Computerprogramms operationalisieren? Vor dem Hintergrund der in Kapitel 3 dargestellten Überlegungen und Befunde wird erwartet, dass sich Individuen darin unterscheiden, auf welche Art und Weise sie ein Computerprogramm explorieren. Sowohl im kognitiv-motivationalen Prozessmodell als auch in Studien zum Problemlösen am Computer (z.B. Trudel & Payne, 1995; van der Linden et al., 2001, 2003) wird der Strategiesystematik eine zentrale Rolle eingeräumt.

Frühere Studien liefern bereits Hinweise auf Möglichkeiten zur prozessnahen Erfassung von Explorationsstrategien am Computer: Ein Beispiel stellt die Studie von van der Linden et al. (2001) dar, in der Explorationsstrategien beim Erlernen und Anwenden eines Tabellenkalkulationsprogramms erforscht wurden. Von den Autoren wurde hierzu theoriegeleitet ein Beobachtungssystem entwickelt, über das sowohl das Auftreten optimaler als auch suboptimaler Explorationsverhaltensweisen (z.B. Versuch-und-Irrtum, Rigide Exploration) messbar gemacht wurde (vgl. auch Dörner, 1980).

Da Explorationsstrategien domänenspezifisch sind, ist es erforderlich, das Beobachtungssystem für neue Problemlöseaufgaben (d.h. weitere Computeranwendungen) entsprechend zu adaptieren (vgl. Boekaerts, 1996, 1997; Perels et al., 2005). Nicht nur Programmfunktionen können sich zwischen Programmen(-versionen) unterscheiden, sondern auch das Benutzerinterface, die Navigationsmöglichkeiten und die Rückmeldungen des Programms an den Nutzer. Darüber hinaus werden Computeranwendungen aufgabenspezifisch eingesetzt, so dass zum Beispiel Briefe mit einem Textverarbeitungsprogramm erstellt werden und Präsentationsprogramme zum Erstellen von Präsentationsfolien dienen. Folglich verbinden die Computernutzer mit verschiedenen Programmfamilien unterschiedliche Aufgaben, die mit Hilfe des jeweiligen Programms bewältigt werden. Daraus ergibt sich, dass aufgabenspezifische Indikatoren für die einzelnen Explorationsstrategien abgeleitet werden müssen. Aus diesem Grund wird in der vorliegenden Arbeit ein Beobachtungssystem zur Erfassung domainspezifischer Explorationsstrategien entwickelt, das auf früheren Studien aufbaut. Das Beobachtungssystem beinhaltet Kategorien für die folgenden Explorationsstrategien: Systematische Exploration, Rigide Exploration, Versuch-und-Irrtums-Exploration, Informationssuche.

Eine zentrale Anforderung an ein neu entwickeltes Beobachtungsinstrument ist, dass es wichtige psychometrische Gütekriterien erfüllt (Wirtz & Caspar, 2002). Hierzu zählen zum Beispiel Kriterien der Messgenauigkeit, der Exklusivität sowie der Exhaustivität. Die Messgenauigkeit bezieht sich dabei auf die exakte Definition der Kategorie, damit eine genaue Zuordnung der interessierenden Verhaltensmerkmale zu den Kategorien möglich ist. Das Exklusivitätskriterium besagt, dass sich die Kategorien gegenseitig ausschließen müssen. Ein Kategoriensystem ist dann exhaustiv, wenn das interessierende Verhalten erschöpfend beschrieben werden kann. Diese Anforderungen sind bereits bei der Konstruktion des Instruments zu berücksichtigen. Ein Maß, um die Messgenauigkeit des Beobachtungssystems zu bestimmen, stellt die Beobachterübereinstimmung dar (Greve & Wentura, 1997).

Hypothese 1: Das neu entwickelte Beobachtungssystem weist eine zufriedenstellende Messgenauigkeit (*Beobachterübereinstimmung*) für die vier zu erfassenden Explorationsstrategien auf.

Vor dem Hintergrund der Ergebnisse bisheriger Studien (Kang & Yoon, 2008; Trudel & Payne, 1995, 1996; van der Linden et al., 2001, 2003) wird erwartet, dass sich die absoluten Beobachtungshäufigkeiten der vier Explorationsstrategien unterscheiden. Die Befunde legen nahe, dass die systematische Explorationsstrategie am häufigsten beobachtet wird. Die rigide Explorationsstrategie sowie die Exploration nach Versuch-und-Irrtum nahmen in Studien den

nächstgrößten Anteil an allen Beobachtungseinheiten ein (vgl. Kang & Yoon, 2008; van der Linden et al., 2001, 2003). Die (übermäßige) Informationssuche wurde demgegenüber vergleichsweise selten beobachtet.

Einschränkend muss bemerkt werden, dass die Befunde der oben zitierten Studien mehrheitlich auf Studierendenstichproben basieren. Entsprechend bezieht sich die angenommene Häufigkeitsverteilung der Explorationsstrategien auf diese Population. Für die Mehrheit der Studierenden stellt der Computer ein häufig genutztes Lern- und Arbeitsmittel dar (Middendorf, 2002, S. 5). Es ist zudem denkbar, dass Studierende im Vergleich zur Allgemeinbevölkerung ein breiteres Spektrum an studienrelevanten Computeranwendungen nutzen. Eine größere Computererfahrung ging in bisherigen Studien mit einer niedrigeren Prävalenz unsystematischer Explorationsstrategien einher. Es ist deshalb zu erwarten, dass bei Studierenden, die im Durchschnitt eine größere Computererfahrung als die Allgemeinbevölkerung besitzen, insgesamt eine niedrigere Häufigkeit der unsystematischen Exploration beobachtet wird. Hypothese 2 fasst die Annahmen zur Häufigkeitsverteilung noch einmal zusammen.

Hypothese 2: Es wird erwartet, dass die systematische Explorationsstrategie den größten Anteil an Explorationshandlungen einnimmt, gefolgt von der rigiden Explorationsstrategie und der Exploration nach Versuch-und-Irrtum. Die Informationssuche weist den kleinsten Häufigkeitsanteil aller erfassten Explorationsstrategien auf.

Die Entwicklung des Kategoriensystems umfasst auch dessen Validierung (Bühner, 2006; Lienert & Raatz, 1998). Anhand ausgewählter Kriteriumsvariablen soll geprüft werden, inwieweit das Instrument Beziehungen zu weiteren (psychologischen) Variablen während des Lernprozesses aufweist. Wie die in Kapitel 2, 3 und 4 berichteten empirischen Befunde verdeutlichen, steht die Strategiesystematik während des Lernens bzw. Problemlösens in positiver Beziehung zu motivationalen Variablen (z. B. Flow-Erleben, aktuelle Motivation). Es wird vermutet, dass insbesondere die systematische Exploration mit dem Gefühl von Kontrolle assoziiert ist, welches durch eine größere Computererfahrung begünstigt wird. Das Erleben von Kontrolle und Kompetenz steht wiederum mit dem Flow-Erleben in Beziehung. Die systematische Strategie kovariert darüber hinaus positiv mit der (Lern-)Leistung.

Im Gegensatz zu systematischen standen unsystematische Explorationsstrategien in früheren Studien mit einer ungünstigeren Motivation und einer geringeren Lernleistung in Beziehung. Eine rigide Explorationsstrategie, die durch häufige Wiederholungen nicht effektiver Handlungen und eine mangelnde Reflexion von Programmrückmeldungen gekennzeichnet

ist, sollte entsprechend zu einer niedrigeren Performanz beim Arbeiten mit einem neuen Computerprogramm führen. Die Häufigkeit, mit der einzelne Explorationsstrategien genutzt werden, wird in Beziehung zum Flow-Erleben und zur Performanz gesetzt. Die erwarteten Zusammenhänge im Rahmen einer ersten Validierung sind in Hypothese 3 und 4 formuliert.

Hypothese 3: Das Ausmaß systematischer Exploration geht mit einem höheren Flow-Erleben einher. Demgegenüber ist die häufige Nutzung unsystematischer Strategien (Versuch-und-Irrtum, rigide Exploration, übermäßige Informationssuche) von einem geringen Flow-Erleben begleitet.

Hypothese 4: Die Nutzung systematischer Strategien steht mit dem Performanzniveau bei der Aufgabenbearbeitung mit dem Computerprogramm in einer positiven Beziehung, wohingegen eine hohe Prävalenz rigider Exploration mit einem niedrigen Performanzniveau einhergeht.

6.2 Erklärung von Geschlechtsunterschieden mittels des kognitiv-motivationalen Prozessmodells (Forschungsfrage 2)

Zeigen sich Geschlechtsunterschiede bei der Exploration und Anwendung eines Computerprogramms? Und wenn ja, in welchen psychologischen Variablen sind Unterschiede bemerkbar? Wie bereits in Kapitel 5.2 ausgeführt, weisen empirische Studien mehrheitlich darauf hin, dass zwischen Männern und Frauen im Erleben und Verhalten am Computer Unterschiede bestehen. Diese manifestieren sich unter anderem darin, Mädchen und Frauen im Vergleich zu Jungen und Männern bei der Anwendung neuer Computeranwendungen weniger erfolgreich sind. Eine niedrigere computerbezogene Leistung von Frauen ging in Studien mit einer niedrigeren Motivation und ungünstigeren Affekten einher. Darüber hinaus liefern die Studien Belege für Geschlechtsunterschiede in der Strategiesystematik.

Das kognitiv-motivationale Prozessmodell erhebt den Anspruch, Leistungsunterschiede beim Lernen und Problemlösen erklären zu können (siehe Kapitel 4.3). In diesem Modell werden Leistungsdifferenzen auf Unterschiede in der Eingangsmotivation (Interesse, Erfolgswahrscheinlichkeit, Herausforderung, Misserfolgsbefürchtung) zurückgeführt (Rheinberg et al., 2001). Der Effekt der Eingangsmotivation auf die Leistung wird wiederum über kognitive und motivationale Variablen wie das Flow-Erleben, die Strategiesystematik und den motivationalen Zustand vermittelt. Im Prozessmodell wird die Motivation als Zustandsvariable berücksichtigt. Auf diese Weise können auch situative und zeitliche Veränderungen der Motivation innerhalb des Lern- bzw. Problemlöseprozesses und deren Beitrag zur Erklärung von interindividuellen Unterschieden in der Lernleistung berücksichtigt werden. Geschlechtsdiffe-

renzen in Motivation und Kognition beim Arbeiten mit dem Computer können auf der Grundlage des Modells prozessnah erfasst werden.

Die Zusammenfassung empirischer Studien in Kapitel 5.2 ergibt zunächst, dass Frauen im Vergleich zu Männern hinsichtlich der Bandbreite der von ihnen genutzten Computerprogramme eine geringere *Vorerfahrung* haben. In Bezug auf die Motivation wird angenommen, dass Frauen bedingt durch Lernerfahrungen und Geschlechtsrollenerwartungen im Durchschnitt ein geringeres *Interesse* an computerbezogenen Aktivitäten aufweisen. Es wird erwartet, dass Frauen insbesondere dann ein niedriges Lerninteresse angeben, wenn sie am Computer mit Aufgaben aus einem mathematiknahen Bereich (zum Beispiel mit Statistikaufgaben) konfrontiert sind. Studien zur *Computer Self-Efficacy* bzw. zu Kompetenzeinschätzungen am Computer zeigen des Weiteren, dass Frauen weniger von ihren Fähigkeiten am Computer überzeugt sind und stärkere Selbstzweifel äußern. Dies sollte sich in einer niedrigen Einstufung der *Erfolgswahrscheinlichkeit* vor Anwendung eines neuen Computerprogramms niederschlagen. Im Gegenzug zeigen Studienergebnisse eine höhere Computerängstlichkeit weiblicher Personen. Da eine höhere Computerängstlichkeit insbesondere beim Erlernen neuer Computeranwendungen berichtet wird, sollten Frauen beim Erlernen eines unbekanntes Programms eine erhöhte situative *Misserfolgsbefürchtung* haben. Es wird darüber hinaus angenommen, dass weibliche Personen eine geringere *Herausforderung* verspüren, sich am Computer ihre eigene Tüchtigkeit zu beweisen. In der Motivationspsychologie ist die Herausforderung auch mit dem persönlichen Wert einer Aufgabe assoziiert. Aufgrund der gesellschaftlich vermittelten Geschlechtsrollenerwartungen wird vermutet, dass Mädchen und Frauen dem Arbeiten am Computer einen geringeren persönlichen Wert zuschreiben. Die Annahmen zu Geschlechtsunterschieden in Vorerfahrung und Eingangsmotivation beim Erlernen eines neuen Computerprogramms sind in Hypothese 5 und 6 zusammengefasst:

Hypothese 5: Frauen haben eine geringere Computervorerfahrung als Männer, das heißt Frauen nutzen weniger häufig Computerprogramme wie Statistiksoftware, Tabellenkalkulationsprogramme oder Datenbanken.

Hypothese 6: Männer und Frauen unterscheiden sich in der Eingangsmotivation dahingehend, dass Frauen eine ungünstigere Motivation vor Beginn der Aufgabenbearbeitung aufweisen: Sie fühlen sich weniger herausgefordert, haben eine geringere Erfolgswahrscheinlichkeit, sind weniger interessiert und haben gleichzeitig eine höhere Misserfolgsbefürchtung.

Bisherigen empirischen Studien zufolge ist die Nutzung unterschiedlicher Explorationsstrategien beim Erlernen einer neuen Computersoftware vom Alter und Vorwissen der Befragten (Kang & Yoon, 2008), von den Explorationsbedingungen (Debowski et al., 2000; Wood et al., 2001) sowie von der Reflexion eigener Handlungen beim Explorieren (Green & Gilhooly, 1990; Trudel & Payne, 1995) abhängig. Dass auch das Geschlecht einen Einfluss auf die Strategiesystematik am Computer und damit auf die Leistung ausüben kann, wurde jedoch nur in wenigen Studien berücksichtigt und erforscht (z. B. Beckwith et al., 2006; Schaumburg, 2004). Die Lernleistung am Computer scheint dabei insbesondere durch das Ausmaß der systematischen Explorationsstrategie vorhergesagt zu werden.

Die zitierten Studien liefern außerdem Hinweise darauf, dass Frauen das Lernen nach Anweisung präferieren. Männer hingegen entdecken verschiedene Programmfunktionen selbst und probieren diese aus. Die „weibliche“ Herangehensweise führt dazu, dass Frauen weniger Funktionen am Computer kennen lernen, die sie für die Aufgabenbearbeitung gegebenenfalls benötigen. Zugleich wird angenommen, dass Männer eine höhere Computerbildung aufweisen als Frauen. Da die Computerbildung wiederum positiv mit der systematischen Explorationsstrategie assoziiert ist, wird erwartet, dass Männer stärker systematisch explorieren als Frauen. Da die systematische Strategie mit einem Gefühl der Kontrolle einhergeht, wird überdies vermutet, dass Männer Frauen auch hinsichtlich der Höhe des Flow-Erlebens übertreffen.

Hypothese 7: Männer nutzen häufiger eine systematische Explorationsstrategie beim Erlernen neuer Programme als Frauen (Hypothese 7a). Männer geben ein stärkeres Flow-Erleben an als Frauen (Hypothese 7b).

Dem kognitiv-motivationalen Prozessmodell zufolge sind beide Mediatorvariablen (Flow und Strategiesystematik) positiv mit der Performanz am Computer korreliert. Die höhere Strategiesystematik sowie das höhere Flow-Erleben der Männer sollten folglich in einer höheren Performanz resultieren (vgl. Hypothese 8).

Hypothese 8: Männer erreichen bei der Exploration von SPSS eine höhere Leistung (Punktzahl bei der Aufgabenbearbeitung) als Frauen.

Der im Prozessmodell postulierte Mediatoreffekt von Flow-Erleben und Strategiesystematik ist in Hypothese 9 formuliert.

Hypothese 9: Die Eingangsmotivation wirkt sich positiv auf das Flow-Erleben sowie auf die Häufigkeit systematischer Exploration aus. Beide Variablen haben wiederum

einen positiven Effekt auf die Performanz (*Mediationshypothese*).

Die in den Hypothesen 8 und 9 postulierten Annahmen sind in Abbildung 9 in Form eines hypothetischen Pfadmodells illustriert. Um die Komplexität der Darstellung zu reduzieren, wurden die vier Aspekte der Eingangsmotivation (Interesse, Herausforderung, Erfolgswahrscheinlichkeit, Misserfolgsbefürchtung) zu einer Variablen zusammengefasst.

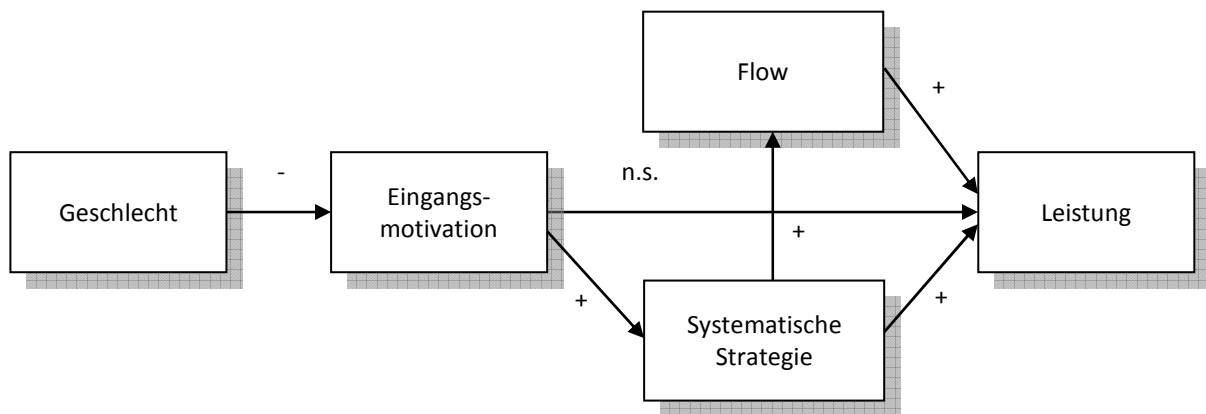


Abb. 9. Theoretisches Pfadmodell zur Vorhersage der Lernleistung auf der Basis des kognitiv-motivationalen Prozessmodells (Vollmeyer & Rheinberg, 1998, 2000, 2006). Das Modell illustriert die erwartete Rolle von Eingangsmotivation, Strategiesystematik und Flow-Erleben als Mediatoren zwischen Geschlecht und Leistung (Kodierung Geschlecht: Männer = -1, Frauen = 1). Pluszeichen spiegeln positive erwartete Pfadkoeffizienten und Minuszeichen negative erwartete Pfadkoeffizienten wider (n.s. = nicht statistisch signifikant).

6.3 Entwicklung von Motivation und Strategiesystematik im Prozessverlauf (Forschungsfrage 3)

In welcher Beziehung stehen die Prozessvariablen (Mediatoren) zueinander und zur Performanz? Forschungsfrage 3 thematisiert die individuelle Entwicklung von Motivation und Strategiesystematik im Verlauf des Lernprozesses. Im kognitiv-motivationalen Prozessmodell werden keine expliziten Annahmen zu möglichen Wechselwirkungen der Mediatoren untereinander formuliert (vgl. Bachmann, 2009). Folglich richtet sich die dritte Forschungsfrage auf die Beziehung zwischen den motivationalen und den kognitiven Mediatoren im Prozessverlauf. Ein Anliegen der Arbeit ist es zu erforschen, ob und auf welche Weise sich die Mediatoren wechselseitig beeinflussen. Des Weiteren wird untersucht, wie sich die Prozessvariablen zu unterschiedlichen Messzeitpunkten auf die abhängige Variable (Performanz) auswirken. Im Rahmen dieser Forschungsfrage soll auch geklärt werden, ob das Flow-Erleben einen direkten oder indirekten Effekt über die Strategiesystematik auf die Leistung ausübt. Die Annahmen zur Beziehung der Mediator- beziehungsweise der Prozessvariablen wurden in ein komplexes Pfadmodell integriert, das im Folgenden näher beschrieben wird (siehe Hypothese 10). Das

Modell basiert auf den in Forschungsfrage 2 postulierten Zusammenhängen und wurde um zusätzliche Annahmen zur Interaktion der Prozessvariablen im Prozessverlauf erweitert. Als Prozessvariablen wurden erneut das Flow-Erleben sowie die systematische Explorationsstrategie in das Modell aufgenommen (siehe Abschnitt 6.2).

Im theoretischen Pfadmodell (siehe Abbildung 9) werden folgende Variablenzusammenhänge postuliert: Das Geschlecht beeinflusst die Leistung indirekt über die Eingangsmotivation sowie über das Ausmaß systematischer Exploration (vgl. Abschnitt 6.2). Männer zeigen vor Beginn des Lernprozesses eine günstigere Eingangsmotivation als Frauen. Eine höhere Eingangsmotivation geht mit einer günstigeren Strategiesystematik im Zuge der Lernhandlungen einher: Folglich explorieren Männer zu allen Messzeitpunkten in höherem Ausmaß systematisch als Frauen. Die systematische Exploration beeinflusst wiederum das Flow-Erleben: Personen, die häufiger die systematische Explorationsstrategie anwenden, haben positivere Erfahrungen im Explorationsprozess. Sie fühlen sich stärker von der Aufgabe absorbiert und nehmen den Verlauf ihrer Tätigkeit als glatt wahr (Rheinberg, 2006; vgl. Schiefele & Rheinberg, 1997). Im Pfadmodell wird ein Effekt der Strategiesystematik auf das Flow-Erleben postuliert: Im Gegensatz zur Strategiesystematik tritt Flow erst im (weiteren) Verlauf der Tätigkeitsausführung.

Die Nutzung systematischer Explorationsstrategien sollte das Auffinden und die korrekte Anwendung aufgabenrelevanter Programmfunktionen begünstigen. Dies erhöht die Wahrscheinlichkeit, dass Aufgaben am Computer richtig gelöst werden. Folglich hat die systematische Explorationsstrategie zu allen Messzeitpunkten einen positiven Effekt auf die Leistung. Das Flow-Erleben zum zweiten Messzeitpunkt (t_2) trägt wiederum positiv zur Nutzung der systematischen Strategie bei: Es wird angenommen, dass sich das Gefühl, den Ablauf unter Kontrolle zu haben und bei jedem Schritt zu wissen, was zu tun ist, positiv auf die Anwendung systematischer Exploration auswirkt (vgl. Csikszentmihalyi & Schiefele, 1993; Schiefele & Rheinberg, 1997).

Zusätzlich wird erwartet, dass das Flow-Erleben kurz vor Bearbeitungsende in besonders enger Beziehung mit der Leistung steht. Diese Erwartung führte zur Aufnahme eines direkten Pfads vom Flow-Erleben zur Leistung. Das Modell impliziert, dass eine höhere Leistung mit einem höheren Flow-Erleben einhergeht. Im Allgemeinen wird der Einfluss des Flow-Erlebens auf die Leistung jedoch indirekt über die Nutzung systematischer Strategien vermittelt. Es wird erwartet, dass die Strategiesystematik im letzten Bearbeitungsabschnitt ebenfalls einen positiven direkten und indirekten Effekt auf die Leistung hat. Der indirekte Effekt ergibt sich dadurch, dass der Einfluss der Strategiesystematik partiell über das Flow-Erleben vermittelt wird.

Die beiden Mediatorvariablen Flow-Erleben und Strategiesystematik sollten sich darüber hinaus über die Messzeitpunkte hinweg als stabil erweisen: Neben den Pfaden, welche die Interaktion der beiden Variablen darstellen, wurden direkte Pfade zwischen den jeweiligen Variablen zu den beiden Messzeitpunkten aufgenommen. Personen, die zu Beginn des Prozesses systematisch explorieren, sollten dies auch über den Prozess hinweg in ähnlichem Maße tun. Im Modell wurde zusätzlich ein direkter Pfad von Geschlecht zur Leistung eingefügt, da der Effekt des Geschlechts über weitere Variablen, die nicht im Modell berücksichtigt wurden, vermittelt sein kann (z. B. Emotionen, Vorwissen). Dieser sollte jedoch nicht signifikant werden. Hypothese 10 fasst die erwarteten Zusammenhänge im Pfadmodell zusammen. Abbildung 10 stellt die theoretische Modellstruktur darüber hinaus grafisch dar.

Hypothese 10: Es wird erwartet, dass der Einfluss des Geschlechts auf die Leistung partiell über die Eingangsmotivation, die Häufigkeit der systematischen Explorationsstrategie sowie die Höhe des Flow-Erlebens zu mehreren Messzeitpunkten vermittelt wird. Die Häufigkeit der systematischen Explorationsstrategie wirkt sich günstig auf das Flow-Erleben aus. Der Effekt des Flow-Erlebens auf die Leistung wird während des Prozesses über die Strategiesystematik vermittelt. Am Ende des Lernprozesses trägt das Flow-Erleben neben der systematischen Explorationsstrategie mit einem inkrementellen Beitrag zur Erklärung der Leistung bei.

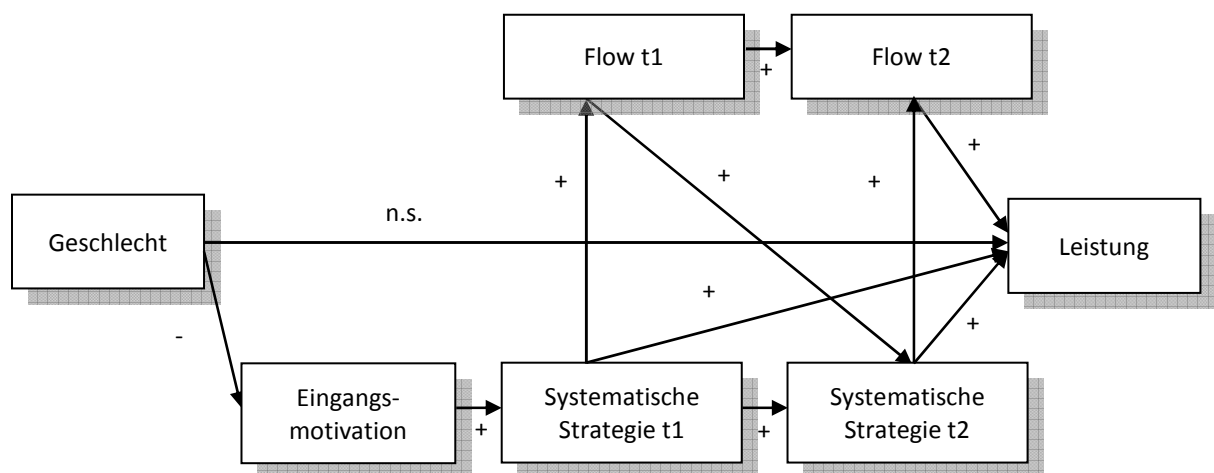


Abb. 10. Theoretisches Pfadmodell zum Zusammenhang von Geschlecht und Leistung auf der Basis des kognitiv-motivationalen Prozessmodells mit der Eingangsmotivation sowie den Prozessvariablen Flow-Erleben und Systematischer Explorationsstrategie als Mediatoren (zu Hypothese 10; Kodierung Geschlecht: Männer = -1, Frauen = 1; t1 = erster Messzeitpunkt; t2 = zweiter Messzeitpunkt).

6.4 Identifizierbarkeit von Personenclustern (Forschungsfrage 4)

Der Schwerpunkt der Forschungsfragen 1 bis 3 liegt auf den Beziehungen der Untersuchungsvariablen untereinander. Die Zusammenhänge der Variablen werden dabei auf Aggregatebene und nicht auf Ebene des Individuums analysiert (Schmitz, 2000). Beispielsweise interessiert bei diesem Vorgehen, inwiefern systematische und unsystematische Explorationsstrategien mit der Leistung zusammenhängen. Im Gegensatz zu diesem Variablen-orientierten Ansatz liegt der Fokus bei Individuen-orientierten Ansätzen auf den interindividuellen Unterschieden zwischen Individuen (vgl. von Eye & Bogat, 2006). Hierbei geht es um die Frage, was bestimmte Personen miteinander gemeinsam haben.

Bezogen auf das Thema der vorliegenden Arbeit stellt sich die Frage, ob sich Personen danach gruppieren lassen, welche Explorationsstrategien sie (vorwiegend) nutzen? Forschungsfrage 4 verfolgt einen Individuen-zentrierten Ansatz und ist der Identifizierbarkeit von Personenclustern gewidmet. Um die Komplexität zu reduzieren, wird in den Pfadmodellen in Forschungsfrage 6.2 und 6.3 lediglich das Ausmaß der systematischen Explorationsstrategie berücksichtigt. Diese hatte sich in Studien als angemessener Prädiktor der Lernleistung erwiesen (siehe Abschnitte 3.1.1, 3.3 und 4.3.2). Es ist jedoch möglich, dass sich Computernutzer im Hinblick auf ihren Einsatz unsystematischer Strategien (Versuch-und-Irrtum, Rigide Exploration, übermäßige Informationssuche) unterscheiden und bestimmte Muster erkennbar sind (siehe Hypothese 11). Denkbar ist beispielsweise, dass Männer und Frauen unterschiedliche Kombinationen von unsystematischen Explorationsstrategien nutzen. Die Validierung der Gruppenzugehörigkeit sollte nach von Eye und Bogat (2006, S. 394) an Außenkriterien erfolgen. Dazu zählt zum Beispiel der Nachweis von Gruppenunterschieden in Motivation und Leistung. Eine Differenzierung von Personenclustern könnte dabei insbesondere für die individuelle Planung von Interventionsmaßnahmen wichtig sein.

Hypothese 11: Personen lassen sich auf der Grundlage der Nutzungshäufigkeit einzelner Explorationsstrategien in Cluster klassifizieren.

6.5 Zusammenfassung

Um die vier oben genannten Forschungsfragen überprüfen zu können, wurden im Rahmen der vorliegenden Arbeit drei empirische Studien durchgeführt. Die Studien basieren auf drei separaten Stichproben und Datensätzen. *Studie 1* umfasst die Entwicklung des Kategoriensystems zur Erfassung der Explorationsstrategien am Computer. *Studie 2* ist der Erklä-

rung von Geschlechtseffekten auf der Basis des kognitiv-motivationalen Prozessmodells gewidmet. *Studie 3* umfasst die empirischen Untersuchungen zu Forschungsfragen 3 und 4. Im Rahmen der Studie wird die Interaktion der Mediatorvariablen des kognitiv-motivationalen Prozessmodells erforscht. Darüber hinaus dient Studie 3 dazu, die Identifizierbarkeit von Personenclustern zu überprüfen.

7. Entwicklung eines Beobachtungssystems zur Erfassung der Explorationsstrategien (Studie 1)

Ziel von Studie 1 war es, ein Beobachtungssystem zu entwickeln, um verschiedene Explorationsstrategien am Computer erfassen zu können (siehe Forschungsfrage 1, Abschnitt 6.1). Mit Hilfe dieses Beobachtungssystems sollten die von Novizen gezeigten Handlungen beim Erkunden einer neuen Computersoftware operationalisiert und messbar gemacht werden. Das neu entwickelte Beobachtungssystem baute auf der Studie von van der Linden et al. (2001) auf. In ihrer Publikation zur Studie beschreiben die Autoren vier Explorationsstrategien. Für jede der vier Explorationsstrategien sollten in der vorliegenden Arbeit aufgabenspezifische Indikatoren identifiziert werden. Studie 1 dient des Weiteren dazu, die Güte des neuen Beobachtungssystems im Rahmen einer ersten Validierung zu überprüfen.

Der in Studie 1 gewählte methodische Ansatz gleicht dem Vorgehen von van der Linden et al. (2001; vgl. auch Carroll & Mack, 1983): Die Studienteilnehmer hatten die Aufgabe, eine neue Software selbstreguliert zu explorieren. Den Teilnehmern wurde dabei kein umfangreiches, vorstrukturiertes Informationsmaterial zur Software zur Verfügung gestellt. Sie wurden nicht detailliert in den Umgang mit dem Programm eingeführt oder angeleitet. Stattdessen erhielten sie eine minimale Einführung in die Grundfunktionen des Programms (vgl. „guided training“, Keith et al., 2010). Das Vorgehen entsprach damit demjenigen des „active/exploratory training“ (Mayer, 2004). Den Lernenden wurden bestimmte Aufgaben (Probleme) vorgegeben, die mit Hilfe des unbekanntes Computerprogramms zu lösen waren. Die Problemlöseaufgabe bestand für die Teilnehmer zunächst darin, die Funktionen der Software zu erkunden. Das im Explorationsprozess erworbene Anwendungswissen über das Statistikprogramm sollten die Teilnehmer zur Lösung spezifischer Aufgaben nutzen. Die Methode des problemorientierten, aktiven Explorierens hat sich in mehreren Studien hinsichtlich ihrer motivationalen und kognitiven Effekte auf das Lernen von Novizen als geeignet erwiesen (siehe Kapitel 3.3; vgl. Carroll et al., 1985; Debowski et al., 2001; Keith et al., 2010).

In Anlehnung an das Vorgehen von Green und Gilhooly (1990) wurde für die Studie eine Statistiksoftware (SPSS 12.0; SPSS Inc., 2003) gewählt. Statistikprogramme stellen insbesondere in den Sozialwissenschaften eine häufig eingesetzte Computeranwendung dar, welche von Studierenden erlernt wird. SPSS bietet die Möglichkeit, Daten sowohl zu beschreiben (Deskriptive Statistiken) als auch inferenzstatistische (Signifikanz-)Testungen vorzunehmen. Das Programm SPSS besitzt eine grafische Benutzeroberfläche. Die Benutzeroberfläche ähnelt derjenigen gängiger Textverarbeitungs- sowie Tabellenkalkulationsprogramme. Aufgrund

dieser Ähnlichkeiten im Interface wird erwartet, dass die Benutzeroberfläche für Personen mit häufiger Anwendung dieser Programme intuitiv erschließbar ist.

Befehle können in SPSS ausgeführt werden, in dem Menüs und Dialogfelder mittels Mauszeiger und Tastatur bedient werden. Alternativ zur Eingabe über Dialogfelder kann das Programm über eine Befehlssyntax gesteuert werden. Beim Aufrufen von SPSS wird automatisch das Datenfenster des Statistikprogramms sowie ein leeres Ausgabefenster geöffnet. Das Datenfenster beinhaltet zwei Blätter: das Datenblatt, das den aktuellen Datensatz zeigt, und die Variablenansicht, in der die einzelnen Variablen des Datensatzes sowie deren Attribute (z. B. Skalenniveau, Wertelabels) angelegt sind. Im Ausgabefenster werden hintereinander die Ergebnisse der SPSS-Berechnungen dargestellt.¹²

Green und Gilhooly (1990) nutzten zur Erfassung der Strategiesystematik die Methode des lauten Denkens sowie Protokolle, in denen die Interaktionen des Computernutzers mit dem Computerprogramm aufgezeichnet wurden. Trotz der vielfältigen Vorteile der Technik des lauten Denkens werden von anderen Autoren auch Nachteile der Auswertung verbaler Protokolle genannt (siehe Abschnitt 3.2). Verbalisierungen im Prozess können eine höhere Reflexionstätigkeit auslösen und das erfasste Verhalten beeinflussen. Für die vorliegende Studie wurden bei Einsatz von *think-aloud*-Techniken (verzerrende) Effekte auf die Motivation sowie auf die Strategiesystematik erwartet: Im Gegensatz zu früheren Studien (z.B. Kang & Yoon, 2008; Trudel & Payne, 1995, 1996; van der Linden et al., 2001, 2003) wurde deshalb auf Techniken des lauten Denkens während des Explorationsprozesses verzichtet. Van der Linden et al. (2001, 2003) haben in ihren Studien die von Green und Gilhooly (1990) eingesetzten Methoden um Videoaufzeichnungen der Bildschirmoberfläche erweitert. Mit Hilfe dieser Videoaufzeichnungen können Verhaltensspuren des Computernutzers genutzt werden, um die Strategiesystematik zu messen. Die Aufzeichnung aller Bildschirmaktivitäten während des Lernprozesses erlaubt zudem, detaillierte Analysen der Strategiesystematik in einzelnen Zeitabschnitten der Exploration vornehmen zu können. Um die Strategiesystematik zu erfassen, wurde in der vorliegenden Arbeit lediglich das Computernutzerverhalten auf Basis von Videoaufnahmen mittels des neu entwickelten Kategoriensystems kodiert.

Als Maß für die Güte des neu entwickelten Beobachtungssystems wurde in der Studie die Beobachterübereinstimmung ermittelt. Der Grad der Übereinstimmung von Beobachterurteilen ist ein Maß für die Zuverlässigkeit des Beobachtungssystems. Einen wichtigen Hinweis auf die Zuverlässigkeit des Instruments bietet auch die Häufigkeit, mit der einzelne Strategien beobachtet werden. Zudem wurde das Kategoriensystem in Studie 1 einer ersten Vali-

¹² Nähere Informationen zu SPSS sowie zu aktuelleren Programmversionen sind unter folgendem Link abrufbar: <http://www.spss.com/de/software/statistics/> Zugriff am 27.7.2008.

dierung unterzogen. Anhand ausgewählter Kriteriumsvariablen wurde geprüft, in wie weit das Instrument das Verhalten misst, das es messen soll. Wie die in Kapitel 3.3 berichteten empirischen Befunde verdeutlichen, steht die Strategiesystematik in positiver Beziehung sowohl zu motivationalen Variablen während des Lern- oder Problemlöseprozesses als auch zur Leistung. Die Häufigkeit der Nutzung einer spezifischen Explorationsstrategie wurde folglich in Beziehung gesetzt zur Ausprägung im Flow-Erleben und zur Leistung.

Zusammenfassend wurden in Studie 1 die folgenden Fragen untersucht (s. auch Abschnitt 6.1):

1. Genügt das neu entwickelte Beobachtungssystem der Anforderung, verschiedene Explorationsstrategien zuverlässig voneinander abzugrenzen?
2. Wie häufig werden die einzelnen Explorationsstrategien in der Stichprobe beobachtet?
3. Können die theoretischen erwarteten Beziehungen der einzelnen Explorationsstrategien mit ausgewählten Validitätskriterien empirisch belegt werden?

Zur besseren Nachvollziehbarkeit der in Studie 1 zu testenden Annahmen sind die bereits in Abschnitt 6.1 dargestellten Hypothesen an dieser Stelle erneut aufgelistet:

Hypothese 1: Das neu entwickelte Beobachtungssystem weist eine zufriedenstellende Messgenauigkeit (*Beobachterübereinstimmung*) für die vier zu erfassenden Explorationsstrategien auf.

Hypothese 2: Es wird erwartet, dass die systematische Explorationsstrategie den größten Anteil an Explorationshandlungen einnimmt, gefolgt von der rigiden Explorationsstrategie und der Exploration nach Versuch-und-Irrtum. Die Informationssuche weist den kleinsten Häufigkeitsanteil aller erfassten Explorationsstrategien auf.

Hypothese 3: Das Ausmaß systematischer Exploration geht mit einem höheren Flow-Erleben einher. Demgegenüber ist die häufige Nutzung unsystematischer Strategien (Versuch-und-Irrtum, rigide Exploration, übermäßige Informationssuche) von einem geringen Flow-Erleben begleitet.

Hypothese 4: Die häufige Nutzung systematischer Strategien steht mit einem höheren Performanzniveau bei der Aufgabenbearbeitung mit dem Computerprogramm in Beziehung, wohingegen eine hohe Prävalenz rigider Exploration mit einem niedrigen Performanzniveau einhergeht.

7.1 Methode

7.1.1 Stichprobe

Im Rahmen einer Statistikveranstaltung im Studienfach Psychologie sowie einer Einführungsveranstaltung in die Wirtschaftspädagogik an der Goethe-Universität Frankfurt wurden im Sommersemester 2007 insgesamt 18 Studierende für die Studie geworben. Die Teilnahme an der Studie war freiwillig. Es handelte sich damit um eine Gelegenheitsstichprobe (Huber, 2005, S. 118; vgl. Rost, 2007). Es wurden zur Studie ausschließlich Personen zugelassen, die in Bezug auf das Statistikprogramm SPSS 12.0 noch keine Vorerfahrungen hatten (*Novizenstatus*; vgl. Naumann et al., 2001).

In den Veranstaltungen konnten ausschließlich Frauen für die Studie gewonnen werden, davon 16 Studierende der Psychologie und zwei Studierende der Wirtschaftspädagogik. Dies ist maßgeblich der Geschlechterzusammensetzung der beiden Studiengänge geschuldet, die durch eine überwiegende Mehrheit weiblicher Studierender gekennzeichnet ist. Der Altersdurchschnitt der Teilnehmer lag bei $M = 21.4$ Jahren ($SD = 1.23$). Die durchschnittliche Fachsemesteranzahl betrug $M = 1.94$ ($SD = 2.1$); 88.2 Prozent der Teilnehmer waren im ersten oder im zweiten Fachsemester.

Um die Studierenden für die Teilnahme an der Studie zu gewinnen, wurden die Studierenden auf die Möglichkeit hingewiesen, ein studien- und berufsrelevantes Computerprogramm kennen zu lernen. Darüber hinaus hatten sie Gelegenheit an einer dreistündigen Computerschulung des in der Studie eingesetzten Programms SPSS teilzunehmen und dafür eine Bescheinigung zu erhalten. Den Psychologie-Studierenden wurden außerdem Versuchspersonenminuten gemäß den Studienanforderungen für die Teilnahme gutgeschrieben.

7.1.2 Vorgehen

Die Studie wurde in den Räumen des Instituts für Psychologie der Goethe-Universität Frankfurt durchgeführt. Es konnten bis zu drei Personen gleichzeitig an der Studie teilnehmen. Nach der Begrüßung der Teilnehmer in den Laborräumen des Instituts (siehe Anhang A, Abschnitt A 1.1), beantworteten die Teilnehmer zunächst einen Fragebogen zu ihren Statistikkenntnissen und zu ihrer Computererfahrung (siehe Anhang A 2.1.1). Das Vorgehen diente dem Zweck, den Vorwissensstand der Teilnehmer in Statistik sowie die Computererfahrung der Teilnehmer kontrollieren zu können. Im Anschluss an die Beantwortung der Fragebögen erhielten die Teilnehmer eine circa einminütige Einführung in die Benutzeroberfläche der Statistiksoftware SPSS 12.0 (siehe Abbildung 11).

7. Studie 1: Entwicklung eines Beobachtungssystems

nr	name	geschl	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	beur	ag
1	Volker R	0	11	11	10	11	11	11	10	12	11	10	7	2915.00
2	Sigrid K	1	9	9	3	9	4	4	3	4	3	2	2	2125.00
3	Elmar M	0	6	5	11	6	11	5	11	6	3	11	10	2775.00
4	Jochen H	0	7	9	10	11	11	11	10	12	11	11	9	3065.00
5	Otto R	0	5	4	6	5	4	11	5	11	11	6	3	1825.00
6	Elke M	1	9	9	3	9	4	3	2	3	4	4	2	1195.00
7	Sarah K	1	3	3	4	4	3	8	4	9	9	5	3	2625.00
8	Peter T	0	5	6	4	5	6	11	6	11	11	6	2	5625.00
9	Gudrun M	1	3	4	9	3	9	4	9	3	2	11	10	2025.00
10	Siglinde P	1	9	9	3	9	4	3	3	4	3	3	2	2625.00
11	Werner W	0	5	6	11	6	11	5	11	5	6	11	10	2325.00
12	Achim Z	0	11	11	10	11	11	10	10	12	11	11	9	2785.00
13	Dieter K	0	7	6	5	4	6	11	6	11	11	6	2	3295.00
14	Boris P	0	4	5	10	4	10	5	10	5	4	10	9	1675.00
15	Silke W	1	10	10	9	10	10	9	10	9	10	10	10	2525.00
16	Clara T	1	6	5	3	4	4	10	4	10	10	5	2	1725.00
17	Manfred K	0	10	10	5	10	4	5	4	4	4	5	4	3865.00
18	Richard M	0	4	5	10	4	10	4	10	4	4	10	9	1675.00
19	Patrick M	0	11	11	10	11	11	7	10	9	11	10	7	2915.00
20	Tajana H	1	9	9	3	9	4	3	3	4	3	2	2	2125.00
21	Thorsten P	0	6	5	11	6	11	4	11	6	3	11	10	2775.00
22	Oliver G	0	7	9	10	11	11	9	10	11	11	11	9	3065.00
23	Fabian H	0	5	4	6	5	4	8	5	12	11	6	3	1825.00
24	Lisa E	1	9	9	3	9	4	11	2	9	4	4	2	1195.00
25	Melanie C	1	3	3	4	4	3	6	4	9	9	5	3	2625.00
26	Peter P	0	5	6	4	5	6	6	6	11	11	6	2	5625.00
27	Judith O	1	3	4	9	3	9	7	9	5	2	11	10	2025.00
28	Christina Y	1	9	9	3	9	4	4	3	4	3	3	2	2625.00
29	Heinrich T	0	5	6	11	6	11	9	11	7	6	11	10	2325.00
30	Siegfried Z	0	11	11	10	11	11	4	10	11	11	11	9	2785.00
31	Robert S	0	7	6	5	4	6	8	6	11	11	6	2	3295.00
32	Adrian E	0	4	5	10	4	10	10	10	5	4	10	9	1675.00
33	Susann B	1	10	10	9	10	10	9	10	9	10	10	10	2525.00
34	Elena J	1	6	5	3	4	4	9	4	10	10	5	2	1725.00
35	Dorian G	0	10	10	5	10	4	8	4	6	4	5	4	3865.00
36	Gustav I	0	4	5	10	4	10	11	10	7	4	10	9	1675.00

Abb. 11. Benutzeroberfläche des Programms SPSS 12.0 mit dem in der Studie verwendeten Beispieldatensatz.

Die Teilnehmer arbeiteten an einem PC, an dem beide Fenster des Programms (Datenfenster und Ausgabefenster) bereits geöffnet waren. Das Ausgabefenster war leer, das Datenfenster beinhaltete den eigens für die Untersuchung zusammengestellten Datensatz. In diesem Datensatz waren mehrere Variablen (z. B. „Zahlengedächtnis“, „Wort- und Redegewandtheit“, „Geschlecht“) und Werte von fiktiven Befragten enthalten. Der genaue Wortlaut der Einführung in das Programm SPSS ist im Anhang (Abschnitt A 1.2) nachzulesen. Neben der Tastatur lagen Notizblätter und Stifte für die Teilnehmer bereit. Sie wurden darauf hingewiesen, dass sie diese Materialien ebenfalls nutzen können.

Nach der Einführung in SPSS wurde den Teilnehmern das Arbeitsblatt mit fünf Textaufgaben (siehe Anhang, Abschnitt A 2.1.2) vorgelegt. Dieses Vorgehen hatte zum Ziel, dass die Teilnehmer vor Beginn Verständnisfragen zu den Aufgaben stellen konnten. Aufgabe der Teilnehmer war es, möglichst viele Aufgaben des Arbeitsblattes innerhalb von 25 Minuten mit Hilfe von SPSS richtig zu lösen. Die Teilnehmer wurden darüber informiert, dass ihr Verhalten während des Arbeitens mit SPSS von zwei Computerprogrammen aufgezeichnet wird.

Das Ablaufen beider Programme war von den Teilnehmern während der SPSS-Sitzung nicht zu bemerken, um eine größtmögliche Unauffälligkeit der Beobachtung zu gewährleisten und das interessierende Verhalten nicht durch den Beobachtungsvorgang zu beeinflussen (Bortz & Döring, 2003; Greve & Wentura, 1997).

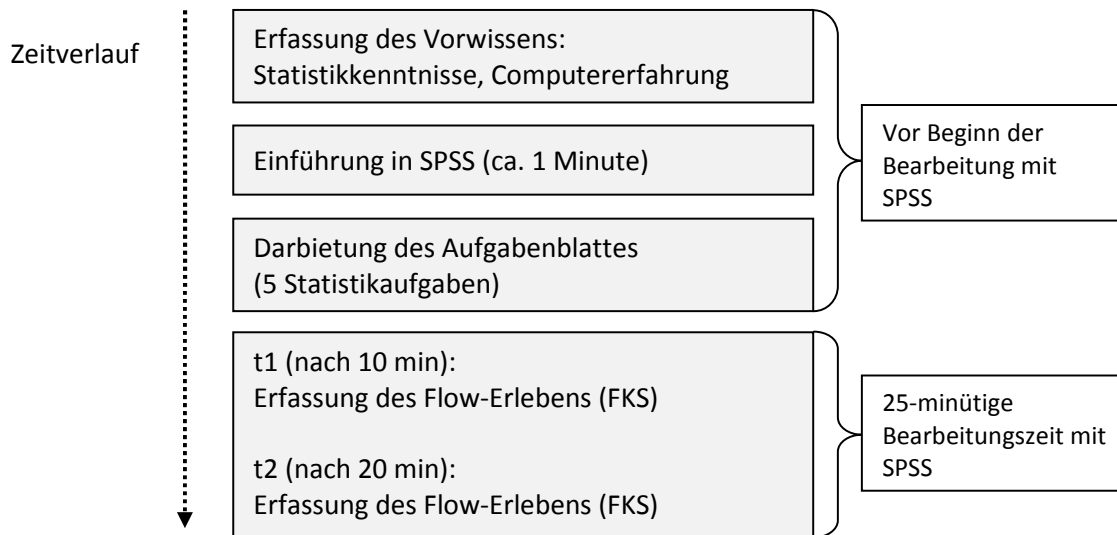


Abb. 12. Versuchsplan von Studie 1 im Zeitverlauf mit Angabe der erfassten Konstrukte.

Der Ablauf der Studie ist in Abbildung 12 noch einmal grafisch dargestellt. Im Anschluss an die Instruktion füllten die Teilnehmer den Fragebogen zu ihren Statistikkenntnissen sowie zu ihrer Computererfahrung aus (siehe Kapitel 7.1.3.1 sowie 7.1.3.2). Danach hatten die Teilnehmer 25 Minuten Zeit, die Aufgaben mit SPSS zu lösen. Während der Bearbeitungszeit mit SPSS wurden die Teilnehmer nach 10 Minuten (Messzeitpunkt t1) und nach 20 Minuten (Messzeitpunkt t2) kurz unterbrochen, um das Ausmaß des Flow-Erlebens mit Hilfe der Flow-Kurz-Skala (FKS; Rheinberg, Vollmeyer & Engeser, 2003; siehe Kapitel 7.1.3.3) zu messen. Dazu bekamen sie einen entsprechenden Fragebogen (Papier-und-Bleistift) von der Versuchsleiterin vorgelegt, die dazu kurz den Raum betrat. Nach Beantwortung des Fragebogens konnten die Teilnehmer den Bearbeitungsprozess fortsetzen.

Vor Beginn jeder Aufgabe wurden die Teilnehmer aufgefordert, die Uhrzeit auf dem Arbeitsblatt zu notieren. Diese Information sollte die Auswertung der Videoaufzeichnungen vereinfachen, indem der Arbeitsprozess in Abschnitte unterteilt werden konnte. Die Aufgaben sollten der Reihenfolge nach bearbeitet werden. Dieses Vorgehen diente dem Zweck, die Aufgabenbearbeitung zwischen den Teilnehmern zu standardisieren. Darüber hinaus sollte sichergestellt werden, dass die Anforderungen an die Teilnehmer im Laufe des Explorationsprozesses zunehmen. Mit der Bearbeitung der nächsten Aufgabe sollte erst begonnen werden, wenn

die Teilnehmer für die vorangehende Aufgabe eine Lösung gefunden und notiert hatten. Nach 25 Minuten wurde die Untersuchung abgebrochen. Insgesamt dauerte die Durchführung ca. 60 Minuten. Am Ende der Untersuchung wurden die Teilnehmer über die Ziele der Studie aufgeklärt und auf die Einladung zur SPSS-Schulung aufmerksam gemacht.

Wenn zwei Teilnehmer gleichzeitig an der Studie teilnahmen, arbeiteten diese in getrennten Räumen. In dem Fall, dass drei Personen gleichzeitig an der Studie teilnahmen, wurden die drei Personen auf zwei Räume aufgeteilt. Die beiden Personen in einem Raum wurden so platziert, dass sie gegenseitig keinen Einblick auf die Bildschirmoberfläche nehmen konnten. Die Teilnehmer wurden darüber hinaus angehalten, während des Versuchs nicht miteinander zu kommunizieren. Die Versuchsleiterin hielt sich während der Aufgabenbearbeitung in einem Zwischenraum auf, der durch eine Tür abgetrennt war. Dadurch konnte sichergestellt werden, dass sich die Teilnehmer während des Arbeitsprozesses nicht gegenseitig beeinflussten.

7.1.3 Materialien

Im Folgenden werden die in der Studie verwendeten Materialien und Skalen vorgestellt. Zunächst werden Konstruktion und Aufbau der einzelnen Skalen beschrieben. Jedes Messinstrument wurde einer statistischen Item- und Skalenanalyse unterzogen, in welcher der Schwierigkeitsindex (Bühner, 2006, S. 83ff.; Kelava & Moosbrugger, 2007, S. 74ff.; Lienert & Raatz, 1998, S. 57f.) und die Trennschärfe (Bühner, 2006, S. 95ff.; Mummendey, 1995, S. 73f.) der Items sowie die Reliabilität der Skala (Schermelleh-Engel & Werner, 2007, S. 114 f.) überprüft wurden. Vor Beginn der item- und skalenanalytischen Überprüfung der Skalen sowie der statistischen Datenanalyse, wurde die Vollständigkeit des Datensatzes überprüft.

7.1.3.1 Fragebogen zu Statistikkenntnissen

Das Statistikvorwissen der Teilnehmer wurde mit Hilfe eines subjektiven und eines objektiven Maßes erfasst: Die Teilnehmer machten im ersten Teil des Fragebogens Angaben zu ihren Kenntnissen und Vorerfahrungen und beantworteten in zweiten Teil zusätzlich einen Vorwissenstests mit statistischen Fragen (siehe Anhang, Abschnitt A 2.1.1).

Im ersten Teil gaben die Teilnehmer Anzahl und Art der von ihnen besuchten Statistikveranstaltungen an der Universität an. Hierzu zählten die im Curriculum obligatorischen Statistikvorlesungen I und II, sowie die zugehörigen Tutorien. Zusätzlich kreuzten die Teilnehmer an, ob sie bereits in der Schule Statistikunterricht gehabt hatten. Den Teilnehmern wurde eine 5-stufige Likertskala mit den Endpolen „sehr gut“ (1) und „mangelhaft“ (5) vorgelegt. Darauf stuften sie ihre theoretischen Statistikkenntnisse ein. Darüber hinaus kreuzten die

Teilnehmer an, auf welche Inhalte sie ihre theoretischen Statistikkenntnisse beziehen konnten: auf Probleme der Deskriptiven Statistik, der inferenzstatistischen Überprüfung von Hypothesen und auf die grafische Darstellung der statistischen Auswertungen.

Im zweiten Teil beantworteten die Teilnehmer einen ad hoc entwickelten Vorwissenstest, der aus neun Mehrfach-Wahl-Aufgaben („Multiple Choice“; Jonkisz & Moosbrugger, 2007) bestand. Jede Mehrfach-Wahl-Aufgabe beinhaltete vier Antwortmöglichkeiten, von denen mindestens eine Antwort korrekt war. Die restlichen Antworten fungierten als Distraktoren. Die Teilnehmer waren instruiert, alle richtigen Antworten anzukreuzen. Richtig angekreuzte bzw. richtig freigelassene Antwortkästchen wurden mit 1 kodiert. Pro Aufgabe konnten bis zu vier Punkte erreicht werden. Der maximal zu erreichende Punktwert im Vorwissenstest betrug 36 Punkte. Ein Beispielitem ist das folgende: „Empfindlich gegenüber Ausreißern unter den Beobachtungen ist... a) der Modus (Modalwert), b) das Arithmetische Mittel, c) der Median, d) die Streuung“. Die korrekte Lösung ist hier Antwort d.

Die Überprüfung der Reliabilität ergab lediglich eine niedrige interne Konsistenz von Cronbachs $\alpha = .42$ (vgl. Bühner, 2006, S. 140). Den Trennschärfeanalysen zufolge trugen vor allem die Items 4 ($r_{it} = .11$), 6 ($r_{it} = .04$) und 8 ($r_{it} = -.19$) zur niedrigen Reliabilität der Skala bei: Die Items stehen wenig bzw. in einem negativen Zusammenhang mit dem aggregierten Wert in dieser Skala. Die Items erwiesen sich dadurch als ungeeignet für die Skala (vgl. Bühner, 2006, S. 95). Der Ausschluss der drei Items aus der Skala hatte eine deutliche Erhöhung der Reliabilität der Skala zur Folge ($\alpha = .62$), die weiterhin als niedrig, jedoch als akzeptabel angesehen wurde (vgl. Peterson, 1994).

Die Trennschärfekoeffizienten der Skala lagen nach Reduktion auf sechs Items im unteren bis mittleren Bereich zwischen $r_{it} = .21$ und $.57$. Werte unter $r_{it} = .40$ werden dabei als weniger gut eingestuft (vgl. Kelava & Moosbrugger, 2007, S. 82). Nur zwei Items überschritten hinsichtlich ihrer Trennschärfe diese Grenze (Item 3 und Item 7). Zudem zeigte sich, dass Item 2 kritische Werte hinsichtlich der Schiefe (-2.71) und des Exzesses (5.98) erreichte. Das Weglassen des Items hätte eine Reduktion der internen Konsistenz auf $\alpha = .58$ bedeutet. Aus diesem Grund wurde von einer weiteren Kürzung der Skala abgesehen. Item 2 wies zudem eine (relativ zu den anderen Items gesehene) akzeptable Trennschärfe auf. Die Itemschwierigkeiten der sechs verbleibenden Items lagen zwischen $P_i = 59\%$ und 95% . Sie spiegeln damit eine mittelschwere bis leichte Beantwortbarkeit der Items wider (vgl. Mummendey, 1995, S. 73). Insgesamt ist der Statistikvorwissenstest bezüglich seiner messtheoretischen Eigenschaften als wenig zufriedenstellend zu bewerten.

7.1.3.2 Fragebogen zur Computerbildung (INCOBI)

Um die Computerkenntnisse und -vorerfahrung der Teilnehmer zu kontrollieren, wurden Items aus dem Inventar zur Computerbildung (INCOBI) von Richter, Naumann und Groeben (1999) entnommen. Das INCOBI stellt ein umfangreiches Instrument dar, mit dem unterschiedliche Aspekte der Computerbildung (*Computer Literacy*) sowie computerbezogene Einstellungen mit vier beziehungsweise acht Skalen erfasst werden. Das Inventar ist nach Angabe der Autoren besonders für Studierendenpopulationen geeignet und erwies sich als valides Messinstrument zur Vorhersage der Computernutzung (Naumann et al., 2001). Zur Beantwortung der Items hatten die Teilnehmer unbegrenzt viel Zeit. Die Beantwortungsdauer der ausgewählten Items der beiden INCOBI-Skalen lag bei ca. 10-15 min.

Aus dem INCOBI wurden zwei Skalen ausgewählt: Die Skala „Vertrautheit mit Computeranwendungen“ (VECA) sowie die Skala „Praktisches Computerwissen“ (PRACOWI). Die VECA erfasst auf 5-stufigen Ratingskalen die Selbsteinschätzung der Befragten, wie vertraut diese im Vergleich zu anderen Studierenden mit 12 verschiedenen Computeranwendungen sind (Naumann et al., 2001). Um die Ökonomie zu erhöhen und den Fragebogen besser an die Fragestellung der Untersuchung adaptieren zu können, wurde der Fragebogen leicht gekürzt: Aus den im Fragebogen VECA aufgeführten Computerprogrammen wurden die Selbsteinschätzungen des Umgangs mit Computern allgemein sowie des Umgangs mit Terminplanungsprogrammen eliminiert. Das in Terminplanungsprogrammen praktizierte Handlungswissen wurde für die Transferleistung auf ein neu zu erlernendes Statistikprogramm als weniger bedeutsam eingestuft. Darüber hinaus wurde die Instruktion sprachlich geringfügig vereinfacht. Die in der Studie eingesetzten Items der Skalen VECA und PRACOWI sind im Anhang aufgeführt (siehe Abschnitt A 2.1.1, Teil II und III).

Im VECA beurteilen die Teilnehmer wie vertraut sie im Vergleich zu anderen Studierenden mit der genannten Computeranwendung (z. B. Textverarbeitungsprogrammen oder Datenbanken) sind. Ihr Urteil stufen Sie auf der folgenden Skala ein: „weit unterdurchschnittlich“ (Kodierung: 0), „unterdurchschnittlich“ (1), „durchschnittlich“ (2), „überdurchschnittlich“ (3), „weit überdurchschnittlich“ (4). Aus den Selbsteinschätzungen zu den vorgegebenen Computeranwendungen wird ein Summenwert von maximal 50 Punkten gebildet. Höhere Werte spiegeln dabei eine größere Vertrautheit mit Computeranwendungen wider.

Mit Hilfe der Skala PRACOWI wird das Nutzungswissen für gängige Anwendungen abgefragt, z.B. für Textverarbeitungsprogramme, und es werden Handlungsroutinen für Programmfamilien wie die Windows-Anwendungen erfasst (Richter, Naumann & Groeben, 2001, S. 4). Die PRACOWI-Skala umfasst im Original 12 Multiple-Choice-Items, bei denen jeweils

eine Antwort als ideale, beste Antwort gilt und drei weitere Antwortoptionen als Distraktoren dienen. In den einzelnen Items werden Aufgaben beziehungsweise Probleme beschrieben, welche für den alltäglichen Umgang mit dem Computer als typisch gelten (Naumann et al., 2001). Dieses Handlungswissen wurde von der Autorin für das Erlernen von Computerprogrammen, z. B. im Umgang mit auftretenden Fehlern oder Problemen sowie für die Transferleistung, als relevant eingestuft. Der Teilnehmer hat in den PRACOWI-Items die Aufgabe, die richtige Handlungsoption beim Auftreten des geschilderten Computerproblems anzukreuzen. Die Skala berücksichtigt die Erkenntnis, dass prozedurales Wissen schwer erfassbar ist: Es werden Prozeduren abgefragt, bei denen das am Anfang des Lernprozesses erworbene deklarative Wissen noch leicht zugänglich ist. Ein Beispielitem für die Skala PRACOWI ist das folgende: „Ihr Computer ist abgestürzt und Sie wollen ihn möglichst ‚schonend‘ neu starten. Was tun Sie? a) Ich drücke den ‚reset‘-Knopf, b) Ich drücke die Tastenkombination ‚Strg‘ + ‚Alt‘ + ‚Entf‘, c) Ich drücke die Tastenkombination ‚Ende‘ + ‚Enter‘, d) Ich schalte den Computer aus und wieder ein, e) weiß nicht.“ Die korrekte Lösung ist Antwort b).

Um die Ökonomie der in der Studie eingesetzten Materialien zu erhöhen, wurden auch aus der Skala PRACOWI einzelne Items nach ihrem Iteminhalt ausgewählt. Auswahlkriterium war, wie relevant das in dem Item abgefragte Handlungswissen für die erfolgreiche Bearbeitung der SPSS-Aufgaben eingestuft wurde. Der Skala PRACOWI wurden insgesamt sechs Items entnommen. Auf ein „Eisbrecher-Item“ wurde dabei verzichtet (Richter, Naumann & Groeben, 2001). Für jede richtige Antwort wurde ein Punkt vergeben. Neben dem Ankreuzen einer der vier Antwortmöglichkeiten, hatten die Teilnehmer die Option, die „weiß nicht“-Kategorie anzukreuzen. Der aggregierte Wert der Skala bildete sich aus der Summe der richtig gelösten Items. Falsche Antworten sowie „weiß nicht“-Antworten wurden mit 0 kodiert.

Die von Naumann et al. (2001, S. 226) bzw. Richter, Naumann und Groeben (2001, S. 8) berichteten internen Konsistenzen für die Skala PRACOWI lassen auf gute Reliabilitäten schließen (Cronbachs $\alpha = .90$ bzw. $.88$). Das gleiche gilt für die Skala VECA (Cronbachs $\alpha = .86$ bzw. $.91$). In der vorliegenden Studie resultierte die Selektion von sechs Items der PRACOWI-Skala in einer unzufrieden stellenden internen Konsistenz von Cronbachs $\alpha = .52$. Diese weicht deutlich von der von Richter, Naumann und Groeben (2001) bzw. Naumann et al. (2001) berichteten Reliabilität des Messinstrumentes ab. Die Reduktion der Itemanzahl von 12 auf 6 Items kann hierbei als mögliche Erklärung herangezogen werden. Da das Item 1 (Im Original Item 4) eine besonders niedrige Trennschärfe auswies ($r_{it} = .02$) wurde das Item entfernt. Zwei der fünf verbleibenden Items erreichten höhere Trennschärfe-Koeffizienten als $.40$ (vgl. Tabelle A – 3 im Anhang, Abschnitt 2.2.1). Die Eliminierung des Items führte zu einer leichten Erhöhung der Reliabilität (Cronbachs $\alpha = .57$).

Die 10 Items der Skala *VECA* erzielten dagegen eine zufriedenstellende interne Konsistenz von Cronbachs $\alpha = .77$ (vgl. Bühner, 2006; Schermelleh-Engel & Werner, 2007). Die Trennschärfen der Mehrheit der *VECA*-Items variierten in einem Bereich zwischen $r_{it} = .43$ und $.83$. Kelava und Moosbrugger (2007, S. 84) bewerten Trennschärfeindizes zwischen $.40$ und $.70$ als gut. Der Schwierigkeitsindex aller Items der Skala lag zwischen $P_i = 10\%$ und 64% . Einen besonders niedrigen Schwierigkeitsindex erreichten Item 5 („Statistikprogramme“, 10%) sowie Item 3 („Programmiersprachen“, $P_i = 13\%$). Bei Item 5 war dies aufgrund der Stichprobenselektion (Novizenstatus in Bezug auf SPSS, Studienanfänger) erwartet worden. Item 3 wies zusätzlich zur niedrigen Trennschärfe eine erhebliche Schiefe von 2.05 und einen bedeutsamen Exzess von 5.20 auf (vgl. Bühner, 2006). Insgesamt fällt die messtheoretische Beurteilung der Skala *VECA* deutlich günstiger aus als hinsichtlich der Skala *PRACOWI* (siehe Tabelle A – 4, Anhang, Abschnitt 2.2.1).

7.1.3.3 Flow-Kurz-Skala (FKS)

Um die Motivation, konkret das Flow-Erleben, während des Problemlöseprozesses erfassen zu können, wurde die Flow-Kurz-Skala (FKS) von Rheinberg et al. (2003) eingesetzt. Die FKS enthält 10 Items mit 7-stufigem Antwortmodus mit den Endpolen „trifft nicht zu“ (1) bis „trifft zu“ (7). Mit Hilfe der FKS kann das Flow-Erleben bei beliebigen Tätigkeiten erfasst werden (Rheinberg, 2006, S. 349). Durch mehrmalige Vorgabe während des Lernprozesses ermöglicht es die Skala durch ihre Kürze, den Erlebenszustand zeitnah zu erfassen und dabei die Tätigkeitsunterbrechung möglichst zu begrenzen. Die Skala erwies sich mit Werten um Cronbachs $\alpha = .90$ als zuverlässig (vgl. Rheinberg et al., 2003, S. 266).

Die 10 Items der Skala bildeten in Faktorenanalysen eine zweifaktorielle Struktur (Rheinberg et al., 2003, S. 268 f.): Sechs Items luden auf einem Faktor, der von den Autoren als „Glatter Verlauf“ interpretiert wurde (Beispielitem: „Ich weiß bei jedem Schritt, was ich zu tun habe“). Die Items beziehen sich auf den automatischen, selbstregulierten Ablauf von Handlungen. Die restlichen vier Items bildeten den Faktor „Absorbiertheit“ (Beispielitem: „Ich bin ganz vertieft in das, was ich gerade mache“). Dieser Faktor umfasst Aussagen, welche den Grad der Absorption mit der Tätigkeit messen. Für beide Subskalen des FKS kann ein eigenständiger Summenwert berechnet werden. Durch die starke Varianzaufklärung des ersten Faktors ist nach den Autoren auch eine Generalfaktorlösung gerechtfertigt, so dass es gleichsam zulässig ist, einen Summenwert über alle 10 Items des Fragebogens zu bilden.

Die Überprüfung der FKS ergab eine zufriedenstellende interne Konsistenz von Cronbachs $\alpha = .77$ (Messzeitpunkt t_1 , $N = 18$) bzw. $\alpha = .83$ (Messzeitpunkt t_2 , $N = 14$). Die Reliabilität der Skala wurde unter Bezug auf Schermelleh-Engel und Werner (2007, S. 129f.)

als zufriedenstellend eingestuft. Die Trennschärfekoeffizienten lagen zwischen $r_{it} = -.07$ und $.70$ (t1) bzw. $r_{it} = .09$ und $.67$ (t2). Item 3 („Ich merke gar nicht, wie die Zeit vergeht“) wies zu t1 eine negative Trennschärfe auf ($r_{it} = -.12$). Zum zweiten Messzeitpunkt erreichte das Item demgegenüber eine Trennschärfe von $r_{it} = .44$, so dass das Item auch aufgrund der angemessenen Reliabilität beibehalten wurde. Umgekehrt verhielt sich der Fall bezüglich Item 1 („Ich fühle mich optimal beansprucht“): Zu t1 lag die Trennschärfe des Items bei $r_{it} = .63$; zu t2 erreichte das Item lediglich $r_{it} = .09$. Alle weiteren Trennschärfen lagen zu Messzeitpunkt t2 in dem von Kelava und Moosbrugger (2007) genannten Gütebereich. Zu Messzeitpunkt t1 unterschritten einzelne Items die Untergrenze von $r_{it} = .40$. Die Schwierigkeitsindices der Items variierten zum ersten Messzeitpunkt zwischen $P_i = 41\%$ und 89% . Zum zweiten Messzeitpunkt betrug der Schwierigkeitsindex eines Items mindestens $P_i = 43\%$ und höchstens $P_i = 85\%$. Die Bandbreite der Itemschwierigkeiten lag damit zu beiden Messzeitpunkten im empfohlenen Bereich (vgl. Mummendey, 1995, S. 73). Die statistischen Kennwerte der Flow-Kurzskala sind im Anhang (Tabellen A – 5 und A – 6, Abschnitt A 2.2.1 des Anhangs) aufgeführt.

7.1.3.4 Problemlöseaufgabe und Performanzmaß

Problemlöseaufgabe. Die Teilnehmer bekamen ein Arbeitsblatt vorgelegt, das fünf Statistikaufgaben beinhaltete. Alle Aufgaben sollten mit SPSS 12.0 gelöst werden. Die Anforderungen der Aufgaben werden im Folgenden separat beschrieben. Die *erste* Aufgabe bestand darin, einen Mittelwert und eine Standardabweichung für eine bestimmte Variable zu berechnen. Die *zweite* Aufgabe beinhaltete das Erstellen einer Häufigkeitstabelle. Die *dritte* Aufgabe erforderte die Durchführung einer bivariaten Korrelationsanalyse, die Interpretation des Ergebnisses sowie das Erstellen eines passenden Streudiagramms. Die *vierte* Aufgabe verlangte eine möglichst exakte Reproduktion eines gruppierten Balkendiagramms, welches auf dem Arbeitsblatt dargestellt war. In der *fünften* Aufgabe sollten die Teilnehmer Deskriptive Statistiken für eine Variable berechnen lassen (z.B. einen Mittelwert, Varianz, Spannweite der Verteilung).

Die Ergebnisse zu den Aufgaben wurden von den Teilnehmern auf dem Arbeitsblatt notiert beziehungsweise im SPSS-Ausgabefenster gespeichert (Bsp. Aufgabe 4: Gruppiertes Balkendiagramm im Ausgabefenster anzeigen lassen). Angaben zu den einzelnen Aufgaben sowie zur Punkteverteilung sind in Tabelle 2 als Übersicht dargestellt. Die konkrete Aufgabeninstruktion auf dem Arbeitsblatt ist im Anhang (Abschnitt A 2.1.2) nachzulesen.

Die Textaufgaben waren auf das theoretische Statistikwissen der Teilnehmer abgestimmt und wurden entsprechend der Vorkenntnisse der jeweiligen Stichprobe angepasst. Als Referenz diente hierbei der Stand der Statistikvorlesung bzw. der erfolgreiche Besuch der

Pflichtveranstaltungen im Bereich Statistik. Die Reihenfolge der Items wurde so gewählt, dass der Schwierigkeitsgrad kontinuierlich von Aufgabe zu Aufgabe anstieg. Voruntersuchungen hatten bereits Hinweise darauf geliefert, wie schwierig die Aufgaben zu bearbeiten waren.

Tabelle 2

Übersicht über spezifische Anforderungen der im Arbeitsblatt enthaltenen Statistikaufgaben (Aufg.; N = 18)

Aufg.	Anforderung auf dem Arbeitsblatt ¹	Max.	Prozent
1	Angabe von Mittelwert und Standardabweichung für die Gesamtstichprobe, sowie getrennt für Frauen und Männer ($n = 18$)	2	78
2	Erstellen einer Häufigkeitstabelle ($n = 14$)	4	72
	1 Punkt: korrekte Tabelle		
	1 Punkt: korrekte Tabelle, korrekte Variable		
	2 Punkte: Angabe von Häufigkeiten aus einer Häufigkeitstabelle		
3	Berechnung einer Korrelation ($n = 12$)	6	67
	1 Punkt: Korrelationskoeffizient		
	1 Punkt: Signifikanzniveau		
	2 Punkte: Interpretation des Ergebnisses		
	Erstellen eines Streudiagramms		
	0 Punkte: keine Grafik		
	1 Punkt: zu den Daten passende Grafik		
	2 Punkte: korrektes Streudiagramm		
4	Erstellen eines gruppierten Balkendiagramms ($n = 12$)	3	56
	1 Punkt: zu den Daten passende Grafik		
	2 Punkte: Balkendiagramm		
	3 Punkte: Gruppiertes Balkendiagramm		
5	Berechnung von Deskriptiven Statistiken ($n = 8$) (Mittelwert, Maximum, Standardabweichung, Varianz, Range)	5	33

Anmerkung. Max. = maximal zu erreichende Punktzahl pro Aufgabe auf dem Arbeitsblatt; % = relativer Anteil der Personen, welche die jeweilige Aufgabe vollständig bearbeitet haben. ¹Der genaue Wortlaut der Aufgabenstellung ist im Anhang, Abschnitt A 2.1.2 nachzulesen.

Performanzindikator. Als Leistungsindikator diente die Summe der Punkte, die bei der Lösung der Statistikaufgaben auf dem Arbeitsblatt erreicht wurde. In dieses Maß ging sowohl die korrekte Produktion des Ergebnisses mit Hilfe von SPSS ein als auch die korrekte Interpretation der im Ausgabefenster angezeigten Ergebnisse. Punkte wurden für das Anzeigen eines Ergebnisses im SPSS-Ausgabefenster (z. B. eines Balkendiagramms), für die korrekte Angabe aufgabenrelevanter Kennwerte auf dem Arbeitsblatt (z. B. des Produkt-Moment-Korrelationskoeffizienten r) sowie für die richtige Interpretation der statistischen Kennwerte (z. B. der Interpretation eines positiven signifikante Korrelationskoeffizienten als Hinweis

darauf, dass beide Variablen in einem positiven Zusammenhang stehen) vergeben. Insgesamt konnten maximal 20 Punkte erreicht werden.

Tabelle A – 7 im Anhang (Abschnitt A 2.2.1) gibt Mittelwert, Standardabweichung, Trennschärfe, Schwierigkeit sowie Schiefe und Exzess der Verteilungen in den einzelnen Aufgaben wieder. Aufgrund der Aufgabenstellung konnten nicht alle Aufgaben von allen Teilnehmern bearbeitet und gelöst werden. Allein 4 Teilnehmer (ca. 22%) haben bis zum Ende der SPSS-Sitzung ohne Erfolg an Aufgabe 1 gearbeitet. Nur 6 von 18 Teilnehmern (ca. 33%) haben alle Aufgaben vollständig bearbeitet (siehe Tabelle 2). Für die einzelnen Aufgaben wurde aus diesem Grund ein Schwierigkeitsindex mit Inangriffnahmekorrektur berechnet. Als Grundlage der Berechnungen wurde der korrigierte Itemmittelwert verwendet. Der Itemmittelwert wurde nur für diejenigen Teilnehmer bestimmt, welche die Aufgabe bearbeitet hatten (vgl. Bühner, 2006, S. 86; Kelava & Moosbrugger, 2007, S. 77). Die Aufgabenschwierigkeiten variierten zwischen 50 und 88%. Inhaltlich bedeutet dies, dass im Durchschnitt die Hälfte bis ca. 90% der möglichen Punkte von den Teilnehmern, die die Aufgabe bearbeitet haben, erreicht wurden. Aufgabe 1 wies mit $P = 50\%$ den niedrigsten Schwierigkeitsindex auf. Dies ist darauf zurückzuführen, dass sowohl leistungsstarke als auch leistungsschwache Personen die Aufgabe bearbeiteten. Aufgabe 5 mit einem Schwierigkeitsindex von $P = 88\%$ wurde dagegen nur von Personen in Angriff genommen, welche die Aufgabe davor bearbeitet hatten.

Bei der Berechnung der weiteren Deskriptiven Statistiken, der Reliabilität sowie der Trennschärfe wurde auf die Inangriffnahme-Korrektur verzichtet. Um einen Datenverlust bei der Berechnung dieser Kennwerte zu vermeiden, erhielten alle Teilnehmer, welche eine spezifische Aufgabe nicht bearbeitet hatten, Null Punkte für diese Aufgabe. Schiefe und Exzess lagen innerhalb der von West, Finch und Curran (1995) postulierten Grenzen von kleiner 2 für die Schiefe und kleiner sieben für den Exzess (siehe Tabelle A - 7). Der Leistungsindikator erreichte eine interne Konsistenz von Cronbachs $\alpha = .82$, die als zufriedenstellend eingestuft wird (vgl. Bühner, 2006, S. 140). Die Trennschärfen der Aufgaben variierten zwischen $r_{it} = .54$ und $.86$. Diese liegen im mittleren bis hohen Bereich und überschreiten die von Kelava und Moosbrugger (2007, S. 84) vorgeschlagene untere Grenze von $.40$. Die Aufgaben weisen damit zufriedenstellende Beziehungen zum Summenwert des Leistungsindikators auf.

Bei der Reliabilitäts- und Trennschärfebestimmung sollte Folgendes beachtet werden: Ein Vorteil des gewählten Vorgehens liegt darin, dass der gesamte Stichprobenumfang für die Reliabilitätsanalyse und die Trennschärfebestimmung zur Verfügung stand. Nachteilig ist jedoch die beschränkte Aussagekraft dieser Kennwerte. Für die Berechnung wurden Personen, die die Aufgabe nicht bearbeitet hatten, Null Punkte zugewiesen. Dadurch konnte nicht ermit-

telt werden, welche Korrelation zwischen Aufgabe und Gesamtpunktwert bestanden hätte, wenn sämtliche Aufgaben von allen Teilnehmern tatsächlich bearbeitet worden wären.

7.1.3.5 Beobachtungssystem zur Erfassung der Explorationsstrategien

Kriterien wissenschaftlicher Beobachtung. Der Beobachtungsgegenstand (vgl. Greve & Wentura, 1997, S. 82f.) der vorliegenden Untersuchung stellte das Explorationsverhalten von Personen am Computer dar. Um das interessierende Verhalten erfassen zu können, wurde eine systematische Beobachtung mit Hilfe eines Kategoriensystems durchgeführt. Dabei wurden spezifische Aspekte des Wahrnehmungsfeldes von Beobachtern genauer betrachtet (Greve & Wentura, 1997, S. 12). Konkret wurden aus dem Verhaltensstrom definierte Ereignisse extrahiert und dokumentiert (Rost, 2007, S. 74). Die Beobachtungen wurden mit Hilfe von Zeichen („Codes“) abgebildet und erhielten so Bedeutung (Greve & Wentura, 1997, S. 13). Die Beobachtung musste den Kriterien der Replizierbarkeit und Objektivität genügen, d. h. die Beobachtung musste auch von anderen Beobachtern wiederholt durchführbar sein bzw. diese müssen zu gleichen Kodierungen gelangen können (Bakeman, 2000, S. 139). Ziel der Kategorienbildung war die Aufgliederung des Verhaltensstromes in eindeutig abgrenzbare, trennscharfe Einheiten.

Einflüsse von Beobachtungsfehlern. Bei Beobachtungen können Fehler auftreten (vgl. Greve & Wentura, 1997, S. 44 ff.). Diese können auf den Beobachter selbst (z. B. in Form von Wahrnehmungs- oder Interpretationsfehlern) auf die Beobachtung als solche (z. B. Fehler durch Veränderung des Beobachtungsgegenstandes durch den Einfluss der Beobachtung) und auf Probleme des Beobachtungssystems (z. B. in Form zu enger oder zu weiter Kategorien, mangelnder Eindeutigkeit) zurückgeführt werden. Das Kategoriensystem wurde auf der Basis theoretischer und empirischer Befunde zum Problemlösen allgemein und speziell zu Explorationsstrategien beim Erlernen von Computerprogrammen durch Novizen entwickelt (vgl. auch Bakeman, 2000, S. 141). Dabei wurden die Hinweise von Greve und Wentura (1997, S. 74f.) zur Vorbeugung von Beobachtungsfehlern beachtet. Nach Ableitung der Kategorien aus der Theorie wurden diese handlungsnah definiert (Operationalisierung) und Verhaltensindikatoren für die jeweilige Explorationsstrategie formuliert (Greve & Wentura, 1997, S. 98f.).

Datengrundlage. Videoaufzeichnungen sämtlicher Bildschirmaktivitäten der Teilnehmer bildeten die Datengrundlage der Verhaltensbeobachtung. Zu den beobachteten Aktivitäten zählten das Verschieben des Mauszeigers und das Öffnen oder Anklicken von Fenstern (Daten-, Hilfe-, Ausgabe- oder Syntaxfenster), der Menüleiste sowie das Bedienen von Schaltflächen im aktuell aktiven Fenster. Als Indikatoren dienten Verhaltensweisen, die als charakteristisch für die jeweilige Explorationsstrategie eingestuft wurden. Die Aufzeichnung

des Bildschirmaktivitäten erfolgte über das Screenshotprogramm ScreenVirtuoso 2.50 von Fox Magic Software (2006).

Tabelle 3

Übersicht über die Kategoriebezeichnungen und zentrale Indikatoren des Beobachtungssystems zur Erfassung von Explorationsstrategien am Computer (in Anlehnung an van der Linden et al., 2001)

Kategoriebezeichnung	Indikatoren (Beispiele)
Systematische Exploration (SE)	<ul style="list-style-type: none"> • Auswahl einer Menüoption nach deren Bezeichnung • Aufrufen einer Funktion, Auswahl von Variablen für das Analyseverfahren (Verschieben in die Box), „ok“-Knopf drücken, Ergebniskontrolle im Ausgabefenster • Bei einem Fehler oder bei Auftreten eines nicht erwünschten Ergebnisses werden die Einstellungen im Programm variiert, der Befehl erneut durchgeführt und das Ergebnis kontrolliert
Versuch und Irrtum (TE)	<ul style="list-style-type: none"> • Auswahl der Menüoptionen ohne erkennbaren Bezug zur Aufgabenstellung bzw. nicht in einer logischen Abfolge • Auswahl von Variablen ohne inhaltlichen Bezug • Schneller Wechsel der Navigation zwischen Funktionen
Rigide Exploration (RE)	<ul style="list-style-type: none"> • Ineffektive Handlungen, die sich mehrmals als nicht zielführend erwiesen haben, werden mehr als zweimal wiederholt, z. B. das Markieren von Spalten zur Auswahl von Variablen • Trotz Fehlermeldung wird der gleiche Lösungsweg erneut beschritten
Informationssuche (IS)	<ul style="list-style-type: none"> • Aufrufen der Hilfefunktion oder des Lernprogramms • Suchen und Lesen von Informationen
Unterbrechung (U) ¹	<ul style="list-style-type: none"> • Mindestens 10-sekündiges Pausieren des Mauszeigers in einem Fenster zum Beispiel bei folgenden Ereignissen: <ul style="list-style-type: none"> ○ Ausfüllen eines Fragebogens ○ Niederschrift einer Aufgabenlösung auf dem Arbeitsblatt ○ Lesen einer Aufgabenstellung ○ Auftreten eines Computerproblems • Ausnahme: Das Pausieren im Hilfefenster wurde unter „Informationssuche“ kodiert

Anmerkungen. ¹ Die Kategorie „Unterbrechung“ spiegelt keine Explorationsstrategie wider; jedoch diente sie zur Kodierung von Zeiten, in denen keine Bildschirmaktivität zu beobachten war.

Die vier Explorationsstrategien, die mittels des Kategoriensystems erfasst wurden, werden im Folgenden näher beschrieben (vgl. auch Tabelle 3). Die *systematische Exploration* eines Computerprogramms umfasst solche Verhaltensweisen, welche mit einer hohen Wahrscheinlichkeit zu richtigen Ergebnisausgaben und damit zu einer erfolgreichen Aufgabenbearbeitung in SPSS führen. Beim Erkunden der Computersoftware ist erkennbar, dass der Programmnutzer Hypothesen über die Funktionsweise des Programms gebildet hat und Annah-

men über Lösungswege zu einer Aufgabe entwickelt. Die Hypothesen werden von dem Programmnutzer in einer logischen Abfolge überprüft und die entsprechenden Ergebnisse evaluiert (van der Linden et al., 2001). Es ist erkennbar, dass der Programmnutzer Annahmen darüber bildet, wo aufgabenrelevante Funktionen in dem Computerprogramm lokalisiert sind und wie diese zur Aufgabenlösung nutzbar sind. Das systematische Explorieren manifestiert sich für den Beobachter konkret darin, dass der Computernutzer Funktionen in strukturierter Weise ausprobiert, zum Beispiel verändert er die Einstellungen einer Programmfunktion und führt diesen Befehl erneut aus (van der Linden et al., 2003). Im Anschluss daran vergleicht der Computernutzer das Ergebnis mit dem Aufgabenziel, indem er die im Ausgabefenster angezeigten Ergebnisse aufruft und bewertet.

Im Gegensatz zum planvollen, systematischen Vorgehen ist das Problemlösen nach der Strategie *Versuch-und-Irrtum* dadurch gekennzeichnet, dass dem Erkunden des Computerprogramms keine handlungsleitenden Hypothesen zugrunde liegen (Green & Gilhooly, 1990). Eine ausreichende Reflexion von programm-basiertem Feedback wird bei Vorgehen nach Versuch-und-Irrtum nicht beobachtet (van der Linden et al., 2003). Bei diesem Vorgehen wird ein exzessiver Ressourceneinsatz in Kauf genommen: Ein Indikator hierfür ist das Ausmaß der Interaktionen mit dem Programm. Dieses wurde unter anderem von Trudel und Payne (1995, 1996) über die Anzahl der Mausklicks operationalisiert. Kennzeichen systematischen Vorgehens, wie zum Beispiel das Überprüfen verschiedener Optionen, um den optimalen Problemlöseweg zu entdecken, sind selten. Vorhersagen von Beobachtern über die nächste gewählte Funktion (d. h. den Lösungsweg) eines Programmnutzers sind bei diesem Vorgehen in der Regel nicht möglich. Stattdessen beurteilen Beobachter das Verhalten des Programmnutzers als unsystematisch. Der Explorierende scheint bei der Aufgabenbearbeitung häufig zwischen Teilzielen zu wechseln. Dazu ruft der Programmnutzer in schnellem Wechsel voneinander unabhängige Funktionen auf (Trudel & Payne, 1995, 1996; van der Linden et al., 2003, S. 484 f.). Das Vorgehen ist mit dem von Dörner (1980, S. 93; Schaub, 2006, S. 464) beschriebenen Phänomen des thematischen Vagabundierens vergleichbar.

Im Falle *rigider Exploration* (auch *Dogmatic Entrenchment* bzw. *repetitives Verhalten*; vgl. Dörner & Schaub, 1994; Green & Gilhooly, 1990) wiederholen Personen häufig Handlungssequenzen, obwohl sich diese bereits mehrfach als nicht zielführend erwiesen haben (Trudel & Payne, 1995; van der Linden et al., 2001, S. 190f.). Bei der Exploration eines Computerprogramms zeigt sich dies z. B. dadurch, dass hintereinander die gleiche Schaltfläche angeklickt wird, obwohl diese nicht aktiviert ist und sich infolgedessen keine Veränderung im Programm manifestieren kann. Es wird zu lange an falschen Hypothesen über die Funktionsweise des Programms oder über die Systemstruktur festgehalten; neue Lösungswege innerhalb des Prog-

ramms werden nicht ausprobiert. Rückmeldungen des Programms werden nicht ausreichend verarbeitet und evaluiert (Green & Gilhooly, 1990). Im Gegensatz zur Strategie nach Versuch-und-Irrtum ist die Bandbreite der aufgerufenen Funktionen bei der rigiden Exploration sehr viel enger. Beide Strategien führen jedoch zu einer niedrigeren Performanz beim Erlernen des Computerprogramms (Green & Gilhooly, 1990).

Eine weitere Explorationsstrategie ist die *Suche nach Informationen*. Dies ist der Fall, wenn Computernutzer im Programm das Hilfemenü oder das Lernprogramm aufrufen. Stößt der Computernutzer beim Anwenden eines Computerprogramms auf Hürden, besteht in vielen Programmen die Option, mittels eines Index Informationen zu bestimmten Funktionen des Programms einzuholen. Diese Suche nach Informationen kann zielführend sein, wenn nach aufgabenrelevanten Stichwörtern gesucht wird und Informationen nach ihrer Nützlichkeit und Relevanz bewertet und selektiert werden (vgl. Dörner & Schaub, 1994). Die Informationssuche im Programm kann jedoch neue Fragen aufwerfen, z. B. wenn die dargebotenen Informationen nicht hinsichtlich der genannten Kriterien bewertet werden können (vgl. van der Linden et al., 2001, S. 190). Die Informationen, welche vom Hilfeprogramm ausgegeben werden, werden gelesen, unabhängig davon, ob diese für den Erwerb oder die Anwendung von Wissen über die Software relevant sind oder nicht. Die Suche nach Informationen wird zum alleinigen Ziel im Programm. Zielführende Handlungsweisen, beispielsweise die im Hilfeprogramm dargestellten Lösungsschritte zu implementieren, werden aufgeschoben. Erkennbar wird die „Einkapselung“ in der Informationssuche (vgl. Schaub, 2006; vgl. auch Dörner & Schaub, 1994, S. 439) durch das unverhältnismäßige Aufwenden von Zeit zur Informationsrecherche.

Um auch Phasen der Inaktivität der Computernutzer mittels des Kategoriensystems erfassen zu können, wurde eine weitere Kategorie aufgenommen, welche keine Explorationsstrategie widerspiegelt. Phasen der Inaktivität waren dadurch gekennzeichnet, dass in einem Beobachtungsabschnitt weder der Mauszeiger bewegt wurde noch Eingaben getätigt oder Befehle per Tastatur oder Syntax ausgeführt wurden. Dieser Art der „Aktivitätsunterbrechung“ konnten mehrere Ursachen zugrunde liegen: Die Abwendung vom Programm während der Beantwortung eines Fragebogens oder während des Notierens von Ergebnissen auf dem Arbeitsblatt. Des Weiteren könnten Computerprobleme¹³ Aktivitäten im Programm verhindern oder verlangsamen. Derartige Pausen der Explorationsaktivitäten wurden mit der Kategoriebezeichnung *Unterbrechung* belegt. Zur besseren Abgrenzung dieser Unterbrechungs-

¹³ Als ein Computerproblem wurde der Fall definiert, wenn die Ausführung eines von den Teilnehmern aufgerufenen Befehls besonders viel Zeit in Anspruch nahm oder sich aufgrund der Überlastung ein Absturz des Betriebssystems ergab. In solchen Fällen wurde der Computer mit allen relevanten Anwendungen von der Testleiterin heruntergefahren und erneut gestartet.

phasen wurden die Teilnehmer aufgefordert, bei Eintragung von Ergebnissen auf dem Arbeitsblatt sowie bei der Beantwortung von Fragebogen den Mauszeiger an einer bestimmten Stelle des Bildschirms zu platzieren. Durch diese Vorgabe sollte nachvollziehbar werden, ob Unterbrechungen durch Aufgabenanforderungen zustande gekommen waren.

Darüber hinaus wurde eine *Restkategorie* eingeführt, um diejenigen Beobachtungsausschnitte klassifizieren zu können, für die eine Zuordnung in die Kategorien nicht eindeutig vorgenommen werden konnte. In Tabelle 3 sind die Kategorien des neu entwickelten Beobachtungssystems mit Beispielen der zentralen Verhaltensindikatoren dargestellt. Eine ausführlichere Sammlung der Indikatoren befindet sich im Anhang (Abschnitt A 1.6)

Beobachtungseinheiten. Aufgrund früherer empirischer Befunde sowie praktischer Überlegungen wurde das Beobachtungsgeschehen in bedeutungsvolle, formale Einheiten gegliedert (Greve & Wentura, 1997, S. 82 f.); die Kodierungseinheiten wurden auf Zeitintervalle festgelegt. Bei dieser Vorgabe klassifiziert der Beobachter das Verhalten in jedem Zeitintervall. Die Entscheidung zugunsten dieser Vorgehensweise stützte sich auf die folgende Begründung von van der Linden et al. (2003):

In complex tasks, such as our exploratory computer task, often there are no easy observable begin or endpoints; therefore, we decided to use fixed time intervals of 20 sec as coding units. First, coders decided whether a 20-s interval contained exploration behavior. If that was the case the coders assessed whether exploration behaviour fell into one of . . . three exploration categories. (S. 487)

Abweichend von van der Linden et al. (2003) wurde das Beobachtungsintervall auf 30 Sekunden ausgedehnt. Eine Einteilung in Sequenzen von 30 wurde aufgrund früherer empirischer Befunde als ausreichend angenommen (vgl. Todtenhaupt, 2007). Die Kategorisierung der Verhaltensweisen eines Teilnehmers wurde in der ersten Studie in der Regel nach 25 min beendet. Der gesamte Beobachtungszeitraum in Studie 1 umfasste demzufolge 50 Kodierungseinheiten. Eine Ausnahme bestand jedoch dann, wenn während der Bearbeitung mit SPSS Computerprobleme auftraten. In diesem Fall wurden dem Teilnehmer zusätzliche Bearbeitungsminuten zur Kompensation des Zeitausfalls gewährt. Die Verhaltensausschnitte wurden auf dem Video über die Grenze von 25 Minuten hinaus kategorisiert.

Pro Beobachtungseinheit wurden die Explorationshandlungen einer Kategorie zugeordnet. Dabei wurde diejenige Explorationsstrategie gewählt, welche den größten zeitlichen Anteil der Beobachtungseinheit umfasste. Um ein Beispiel zu geben: Ein Teilnehmer explorierte in einem Beobachtungsintervall 20 Sekunden in systematischer Weise; 10 Sekunden nutzte er die Strategie Versuch-und-Irrtum. Der Anteil mit systematischer Exploration nimmt

in diesem Zeitintervall einen größeren Anteil ein, so dass sein Verhalten als systematisch klassifiziert wurde.

In dem Fall, dass zwei Kategorien den gleichen zeitlichen Anteil an der Beobachtungseinheit einnahmen, wurde folgendermaßen verfahren:

1. Wenn die Systematische Exploration gemeinsam mit einer anderen Kategorie auftrat, wurde eine Kodierung zugunsten der Systematischen Strategie vorgenommen.
2. Wenn Versuch-und-Irrtum gemeinsam mit Rigider Exploration zu beobachten war, wurde die Kategorie Versuch-und-Irrtum vergeben.
3. Wenn Versuch-und-Irrtum beziehungsweise Rigide Exploration zu gleichen Anteilen mit der Kategorie „Unterbrechung“ beobachtet wurden, dann wurde den Kategorien Rigide Exploration beziehungsweise Versuch-und-Irrtum der Vorzug gegeben.

Dieses Vorgehen diente erstens dem Ziel, dem Auftreten systematischer Verhaltensweisen eine besondere Bedeutung zuzuweisen. Zweitens sollten aktive Verhaltensweisen bei der Erkundung des Programms gegenüber Verhaltensweisen mit einer geringen Interaktionsrate mit dem Programm bevorzugt werden. Zu ersteren zählen das systematische Testen von Hypothesen sowie das unsystematische Ausprobieren nach Versuch-und-Irrtum. Zu letzteren gehören die Rigide Exploration sowie die Unterbrechungen. Zusammenfassend wurde diejenige Kategorie, welche ein systematisches, aktiveres Interaktionsverhalten mit dem Programm widerspiegelt, der zweiten Gruppe an Verhaltensweisen vorgezogen.

Beobachtertraining. Wie oben beschrieben, können bei Beobachtungen verschiedene Beobachtungsfehler zu einer Abweichung zwischen dem tatsächlichen Beobachtungsgeschehen und dem Beobachtungsprotokoll führen (Greve & Wentura, 1997, S. 56). Um die Auftretenshäufigkeit dieser Beobachtungsfehler auf Seiten der beobachtenden Person zu kontrollieren und die Anwendungsobjektivität des Beobachtungssystems zu gewährleisten, wurde ein Beobachtertraining durchgeführt (Bortz & Döring, 2003, S. 273 f.; Greve & Wentura, 1997, S. 76 ff.; Rost, 2007, S. 76). Nach Bortz und Döring (2003, S. 274) ist das Ziel des von Beobachtertrainings zunächst die Einführung in das Konzept der Untersuchung und den theoretischen Hintergrund, wobei die Bekanntgabe expliziter Hypothesen vermieden werden sollte. Darüber hinaus dient die Schulung von Beobachtern dazu, Hinweise zur Art und Weise der Datenaufzeichnung sowie zu den Daten selbst zu geben. Das primäre Ziel ist die Einweisung von Beobachtern in das Kategoriensystem (Kategorien und Indikatoren) selbst. Der Nutzen des Beobachtertrainings geht jedoch Bortz und Döring (2003) zufolge über eine reine Einweisung in das Kategoriensystem hinaus:

Schon das Beobachtertraining kann als eine grobe und vorläufige Form der Validierung des Beobachtungsplans angesehen werden. Wenn bereits in der Trainingsphase mit den Beobachtern keine Einigung über die Bedeutung von Indikatoren und Kategorien zu erzielen ist, so dürfte die Eindeutigkeit der Kodierungen sehr zu wünschen übrig lassen. (S. 274)

In einer etwa dreistündigen Beobachterschulung (*Trainingsphase*) wurde eine studentische Hilfskraft, welche mit dem Versuchsablauf und der -durchführung vertraut war, als Beobachter in das Kategoriensystem eingeführt. Der Beobachter wurde zunächst in das Konzept der Untersuchung eingeweiht. Anschließend wurde dem Beobachter die Handanweisung vorgelegt, in welcher die einzelnen Kategorien definiert und in der Beispiele für Indikatoren aufgelistet waren. Nachdem der Beobachter diese gelesen hatte, wurden zur Veranschaulichung der Kategorien zwei Videos ausgewählt. Die Ergebnisse der Kodierung zu diesen Videos gingen nicht in die spätere Berechnung der Beobachterübereinstimmung ein. Die Videos wurden von der Hilfskraft gemeinsam mit der Autorin dieser Arbeit kodiert, um die Umsetzung der Kodierungsregeln zu üben. Nach jedem Abschnitt wurde das Video gestoppt und abwechselnd wurde von dem einen oder dem anderen Beobachter zuerst mitgeteilt, für welche Kategorie sich der jeweilige Beobachter entschieden hatte. Bei fehlender Übereinstimmung wurde diese umfassend diskutiert und das Kategoriensystem gegebenenfalls einer Überarbeitung unterzogen. Ziel war die Präzisierung der Kategorien (zum Beispiel indem bestimmte Verhaltensweisen als Indikatoren in das Kategoriensystem neu aufgenommen wurden).

Um Versuchsleiterartefakte bei der Kodierung zu vermeiden (vgl. Bortz & Döring, 2003; Bungard & Lück, 1974), wurde den Beobachtern vor der Beobachtung kein Hinweis zur Performanz des jeweiligen Teilnehmers zur Verfügung gestellt. Die Aufzeichnung der Aufgabenlösungen in SPSS während der Bearbeitung mit SPSS brachte es jedoch mit sich, dass den Beobachtern während des Kodierungsprozesses Informationen über das Lern- und Leistungsniveau eines Teilnehmers zugänglich waren. Es ist nicht auszuschließen, dass diese Informationen die Kategorisierung beeinflusst haben.

In einem zweiten Schritt der *Trainingsphase* wurde das Kategoriensystem von beiden Beobachtern getrennt anhand von fünf Videos angewendet. In der *Anwendungsphase* wurden 13 weitere Videos von beiden Beobachtern kodiert; dies entsprach insgesamt 18 zweifach kodierten Videos. Die Auswertungsergebnisse wurden verglichen, in dem der Grad der Beobachterübereinstimmung berechnet wurde (vgl. hierzu allgemein Wirtz & Caspar, 2002). Die Ergebnisse zur Beobachterübereinstimmung sind in Kapitel 7.2.2 dargestellt.

7.1.4 Umgang mit fehlenden Werten

Bei Vorliegen fehlender Angaben, die dem Kriterium (vollständig) zufällig fehlender Werte entsprachen (*Missing at random* bzw. *Missing completely at random*; siehe u. a. Graham, Cumsile & Elek-Fisk, 2002, S. 89; Lüdtke, Robitzsch, Trautwein & Köller, 2007, S. 104), wurde eine Korrektur vorgenommen. Angesichts der geringen Anzahl fehlender Werte wurde die von Sjitsma und van der Ark (2003, S. 514) empfohlene *Two-Way-Imputation* zur Missingskorrektur durchgeführt. Bei der *Two-Way-Imputation* wird der fehlende Wert durch den Skalenmittelwert der Person ersetzt, wobei gleichzeitig der Zeitreihenmittelwert des Items und der Gesamtscore der Stichprobe in der Skala berücksichtigt wird. Im Unterschied zur einfachen Imputation (*Ipsative Mean Imputation*; vgl. Schafer & Graham, 2002, S. 158) können bei der *Two-Way-Imputation* Differenzen in den Itemschwierigkeiten sowie dem Niveau der Skalenwerte in der Stichprobe Rechnung getragen werden. Beide Aspekte können bei der *Ipsative Mean Imputation* zu einer systematischen Verzerrung des Skalenwertes einer Person führen, insbesondere wenn Werte bei einem Item mit hoher Itemschwierigkeit fehlen.

7.1.5 Wahl und Diskussion der statistischen Analyseverfahren

Beobachterübereinstimmung. Alle Verhaltensdaten der Untersuchungsteilnehmer wurden mit Hilfe des eigens entwickelten Kategoriensystems von zwei Beobachtern kodiert (vgl. Kapitel 7.1.3.5). Die Urteile der beiden Beobachter lassen sich tabellarisch in einer Übereinstimmungsmatrix darstellen: Hierzu werden die Urteilshäufigkeiten in ein Vierfelderschema übertragen. In diesem sind in den Spalten die einzelnen Häufigkeiten der von Beobachter 1 vergebenen Kategorien aufgeführt. In den Zeilen finden sich die entsprechenden Häufigkeiten der von Beobachter zwei vergebenen Kategorien (vgl. Bortz, Linert & Boehnke, 2008, S. 450-451).

Um die Beobachterübereinstimmung zu quantifizieren, können verschiedene zufalls-korrigierte Übereinstimmungsmaße berechnet werden (vgl. Cohen, 1960; Greve & Wentura, 1997, S. 96 ff.; Hsu & Field, 2003; Wirtz & Caspar, 2002, S. 45 ff.). Hierzu zählt Cohens κ . Dieser Koeffizient gibt die Beobachterkonkordanz bei nominalskalierten Daten wieder. Bei der Berechnung von Cohens κ wurde auf eine differenzielle Gewichtung der Nichtübereinstimmungen verzichtet (vgl. Brosius, 2008, S. 427 ff.; Wirtz & Caspar, 2002, S. 78).

Cohens κ kann theoretisch Werte zwischen 1.0 und -1.0 annehmen. Ein Koeffizient nahe 1.0 spiegelt dabei eine hohe Übereinstimmung der Beobachterurteile wider. Für den empirischen κ – Koeffizienten kann ein Test auf asymptotische Signifikanz durchgeführt werden (vgl. Bortz et al., 2008, S. 452 f.; Wirtz & Caspar, 2002, S. 75). Nach Landis und Koch (1977)

gilt ein κ zwischen .61 und .80 als beachtliche (*substantial*) Übereinstimmung und ein κ zwischen .81 und 1.00 wird als (fast) perfekte Übereinstimmung eingestuft. Bortz und Döring (2003, S. 277) sprechen bei κ -Werten von über .70 von guter Übereinstimmung. Nach Wirtz und Caspar (2002, S. 59; vgl. auch Fleiss & Cohen, 1973) gilt ein Cohens Kappa von .60 bis .74 als „gut“ und über .75 als „sehr gut“. Dabei muss beachtet werden, dass die Bewertung der Höhe des Koeffizienten abhängig ist von den zu beobachtenden Merkmalen und Beobachtungsobjekten.

Kritikpunkte an Cohens κ -Koeffizient beziehen sich nach Klauer (1996, S. 102 f.) auf die Interpretierbarkeit des Kappa-Koeffizienten im Falle überzufälliger Übereinstimmung, die Willkürlichkeit der Zufallskorrektur, paradoxe Verhaltensweisen des Koeffizienten sowie auf die Problematik unterschiedlicher Basisraten bei den Beurteilern (vgl. zu diesem Punkt auch Wirtz & Caspar, 2002, S. 60 f.). Die Cohens κ inhärenten Problematiken können durch das Ausweichen auf andere Koeffizienten der Beurteilerübereinstimmung jedoch nur teilweise kompensiert werden.¹⁴ Hsu und Field (2003) empfehlen vor diesem Hintergrund die Berechnung von Cohens Kappa gegenüber anderen Koeffizienten der Beobachterübereinstimmung. Dieser Empfehlung wird in der vorliegenden Arbeit gefolgt.

Gütekriterien der Multi-Item-Skalen. Die Höhe der *Reliabilität* der eingesetzten Messinstrumente wurde über Cronbachs Alpha bestimmt (Cortina, 1993), wenn das Messinstrument mindestens drei Items umfasste. Dieser Koeffizient gibt in der Regel die untere Grenze der Reliabilität an und gibt die mittleren Itembeziehungen wider unter Berücksichtigung der Itemanzahl (Rost, 2007, S. 156; vgl. auch Bühner, 2006, S. 131). Die *Trennschärfe* der Items wurde mit der *part-whole*-Korrektur vorgenommen (vgl. Bühner, 2006, S. 95). Die *Itemschwierigkeit* der intervallskalierten Persönlichkeitsskalen (z.B. Ratings auf der Flow-Kurz-Skala) wurde nach der Formel (4.8) von Moosbrugger und Kelava (2007, S. 80; vgl. auch Amelang & Zielinski, 2002, S. 121) berechnet. Bei Niveautests (z.B. bei Multiple-Choice-Fragen des Statistiktests) wurde zur Schwierigkeitsberechnung Formel (4.6) verwendet. Die Inangriffnahme wurde berücksichtigt, indem bei der Berechnung nur Daten von Personen eingingen, welche die Aufgabe bearbeitet hatten (Amelang & Zielinski, 2002, S. 124). Zufallseinflüssen bei Multiple-Choice-Tests mit mehreren richtigen Antworten wurde Rechnung getragen, in dem falsch angekreuzte Antworten negativ auf die bei der Aufgabe erreichte Punktzahl angerechnet wurden.

¹⁴ Klauer (1996) hat auf den Annahmen von Cohen beruhend einen weiteren, anspruchsvolleren Ansatz zur Bestimmung der Beobachterübereinstimmung vorgeschlagen, in dem der Übereinstimmungsprozess und der Fehlerprozess modelliert werden.

Statistische Voraussetzungsprüfungen. Die Anwendung statistischer Analyseverfahren macht eine Prüfung erforderlich, ob die empirischen Daten die Voraussetzungen erfüllen. Dieses Vorgehen wird vor allem bei Studien mit einem kleineren Stichprobenumfang von $N \leq 30$ empfohlen (Bortz et al., 2008, S. 81). Zu diesem Zweck müssen nach Bortz et al. (2008, S. 79 ff.) das Messniveau der erhobenen Daten, die Erfüllung der mathematisch-statistischen Voraussetzungen der entsprechenden Verfahren sowie die Robustheit dieser Verfahren gegenüber Voraussetzungsverletzungen berücksichtigt werden. Eine Mindestvoraussetzung für parametrische Analyseverfahren (z. B. eines Mittelwertvergleichs über t -Test oder einer Produkt-Moment-Korrelation) ist das Vorliegen intervallskalierter Daten. Kritisch ist diese Voraussetzung nach Bortz et al. (2008, S. 80) bei subjektiven Einschätzungen zu bewerten. In der vorliegenden Studie stellen die Items des VECA, des FAM sowie der FKS subjektive Einschätzungen psychologischen Erlebens dar. Es wird angenommen, dass durch den gewählten Antwortmodus der Skalen eine äquidistante Stufenfolge vorliegt, so dass die Voraussetzung des Intervallskalenniveaus für diese Selbsteinschätzungsskalen erfüllt ist.

Zentrale mathematisch-statistische Voraussetzungen für parametrische Analyseverfahren (z. B. Pearson-Produkt-Moment-Korrelationen, t -Tests) sind die Normalverteilung der Kennwerte in der Population sowie die Varianzhomogenitätsvoraussetzung bei Stichprobenvergleichen (Bortz, 1993; Kline, 2005). Die Normalverteilungsannahme wurde für die einzelnen Variablen über den Kolmogoroff-Smirnov-Anpassungstest überprüft (vgl. Bortz, 1993, S. 319). Dieser untersucht, ob der maximale Abstand zwischen der empirischen Verteilungsfunktion und der Modellverteilung (hier: der Normalverteilung) innerhalb der kritischen Grenzen für eine bestimmte Stichprobengröße liegt.

Zusammenhangsanalysen. Um lineare Zusammenhänge zwischen Variablen zu testen, wurden Pearson-Produkt-Moment-Korrelationen durchgeführt (Bortz, 2005, S. 113 ff.). Der Produkt-Moment-Korrelationskoeffizient inklusive seines Signifikanztests hat sich auch im Falle von Voraussetzungsverletzungen (z. B. Abweichungen von der Normalverteilung und der Homoskedastizität; vgl. Bortz, 2005, S. 213 ff.) als relativ robust erwiesen (Havlicek & Peterson, 1977). Um das Signifikantwerden sämtlicher bivariater Korrelationen und damit eine Alpha-Fehler-Inflation zu vermeiden, wurden die Zusammenhänge unter den Variablen nicht wahllos, sondern aufgrund theoretischer Überlegungen überprüft (Stelzl, 2005).

Software für die Datenanalyse. Zur Auswertung der Daten wurde das Statistikprogramm SPSS 17.0 (SPSS Inc, 1993-2007) verwendet.

7.2 Erprobung und erste Validierung des Beobachtungssystems

7.2.1 Vorbereitende Analysen zur empirischen Untersuchung

Zur Vorbereitung des empirischen Kernteils dieser Studie (siehe Abschnitte 7.2.2 ff.) wurden einige statistische Analysen vorgeschaltet. Diese beziehen sich auf Häufigkeit von und Umgang mit Missing Data. Im Folgenden wird berichtet, ob die Verteilungen in den zentralen Untersuchungsvariablen die allgemeinen Voraussetzungen für die angewendeten statistischen Analyseverfahren (z. B. Normalverteilungsannahme) erfüllen. Die Erfüllung spezifischer Voraussetzungen (z. B. Varianzhomogenität bei Mittelwertvergleichen über t -Tests) wird im jeweiligen Abschnitt geprüft, in dem die Ergebnisse des entsprechenden Tests berichtet werden. Um die Stichprobe zu charakterisieren, wird ein Überblick über die Deskriptiven Statistiken der zentralen Untersuchungsvariablen gegeben.

Fehlende Werte und Missingskorrektur. Im Datensatz der Studie 1 wurden jeweils zwei fehlende Angaben in der Skala Praktisches Computerwissen (PRACOWI) sowie der Skala Vertrautheit mit Computeranwendungen (VECA) mit der Two-Way-Imputation ersetzt (siehe Abschnitt 7.2.4). Der Anteil der ersetzten Werte entsprach damit 2% (PRACOWI) beziehungsweise 1% (VECA) aller Werte der jeweiligen Skala.

Vier Teilnehmer schlossen die Bearbeitung des Aufgabenblattes bereits vor Ende der 25-minütigen Bearbeitungszeit ab. Dementsprechend konnten bei ihnen keine Werte in der Flow-Kurz-Skala zum letzten Messzeitpunkt erhoben werden. Die fehlenden Werte wurden aufgrund der Systematik, die zum Datenausfall geführt hat, nicht ersetzt. Zum zweiten Messzeitpunkt reduzierte sich demzufolge die Datengrundlage für die statistischen Auswertungen um die Angaben der vier Personen. Nach Little (1995, S. 1114) ist davon auszugehen, dass der systematische Ausfall bei bestimmten, längsschnittlichen Untersuchungsdesigns als „planned drop-out“ und damit als „vorwiegend zufällig“ eingestuft werden kann. Der Datenausfall in der Variablen Flow-Erleben zum letzten Messzeitpunkt ist folglich von der Ausprägung in dieser Variablen unabhängig. Dies ist eine wichtige Voraussetzung für die weitere Analyse der Daten (vgl. z.B. Kline, 2005; Lüdtke et al., 2007).

Um eine ausreichende Datenbasis für die Analyse mit den interessierenden Variablen zu gewährleisten, wurde für das Flow-Erleben eine neue Variable („flow_end“) geschaffen (vgl. Kapitel 7.1.3.3). Diese enthielt die Werte aller Teilnehmer zum jeweils letzten Messzeitpunkt vor der individuellen Beendigung der Aufgaben. Für dieses Vorgehen sprach außerdem, dass auf diese Weise auch das Flow-Erleben bei Personen mit einem höheren Leistungsniveau in die Analysen einbezogen werden konnte.

Statistische Voraussetzungsprüfungen. Tabelle A - 8 im Anhang (Abschnitt A 2.2.2) gibt eine Übersicht über Mittelwerte, Standardabweichung, Schiefe und Exzess in den Verteilungen der Variablen. Darüber hinaus ist das Ergebnis des Kolmogorov-Smirnov-Tests auf Abweichung von der Normalverteilung aufgeführt. Für keine der getesteten Skalen ergaben sich statistisch signifikante Abweichungen der Summenwerte von der Normalverteilung. Lediglich die Verteilung in der Explorationsstrategie „Informationssuche“ näherte sich dem Grenzwert ($Z = 1.37, p = .05$), unterschritt diesen jedoch nicht.

Deskriptive Statistiken. Die Ausprägungen der einzelnen Variablen in der Stichprobe (vgl. Tabelle A - 8 im Anhang, Abschnitt A 2.2.2) lassen sich wie folgt zusammenfassen: Der Skalenmittelwert des VECA zeigt, dass sich die Teilnehmer mit den abgefragten Computeranwendungen im Vergleich zu ihren Kommilitonen leicht unterdurchschnittlich vertraut fühlten: Auf einer Skala von 0 bis 4 erreichten sie einen Mittelwert von $M = 1.44$ ($SD = 0.54$). Hierbei muss jedoch beachtet werden, dass sowohl die Vertrautheit mit Standardanwendungen im Studium (Textverarbeitungsprogrammen) erfasst wurde als auch mit seltenen Anwendungen wie Programmiersprachen und Statistikprogrammen. Im Durchschnitt konnten die Teilnehmer ca. 40% der Aufgaben der Skala zur praktischen Computerbildung (PRACOWI) richtig lösen ($M = 0.38$). Das mit dieser Skala erfasste Handlungswissen variierte relativ stark zwischen den Teilnehmern ($SD = 0.28$). Das Ausmaß des Flow-Erlebens zum letzten Bearbeitungszeitpunkt (Flow_end) während der Bearbeitung mit SPSS lag mit einem Mittelwert von $M = 3.06$ ($SD = 0.80$) leicht unterhalb des theoretischen Durchschnitts ($M = 4$) der 7-stufigen Skala.

7.2.2 Hypothese 1: Zuverlässigkeit des neu entwickelten Kategoriensystems

Wie in Abschnitt 7.1.3.5 beschrieben, wurde das Beobachtungssystem in zwei Phasen (Trainings- und Anwendungsphase) von zwei unabhängigen Beobachtern zur Kategorisierung der Explorationsstrategien eingesetzt. Anhand der fünf ausgewählten Videos wurde erstmals in der Trainingsphase (vgl. Bortz & Döring, 2003, S. 276-277) die Übereinstimmung der Beobachtungen für nominalskalierte Daten geprüft: Cohens κ betrug für die 198 vorgenommenen Kodierungen = .68 ($p < .01$). Dieser Grad an Übereinstimmung kann nach Wirtz und Caspar (2002, S. 59) als gut bezeichnet werden. Im Gegensatz hierzu wird nach der von Bortz und Döring (2003, S. 277) vorgelegten Richtlinie bei einem solchen κ -Wert jedoch noch nicht von einer guten Übereinstimmung gesprochen.

Die Übereinstimmung der Beobachterurteile sowie deren Divergenzen in der Trainingsphase sind in Tabelle 4 wiedergegeben. Wie die Tabelle zeigt, wurde die Restkategorie von beiden Beobachtern nur selten vergeben (5 von 198 Kodierungen). Nur eine geringe An-

zahl der Beobachtungen wurde der Kategorie Versuch-und-Irrtum zugeordnet. Hier zeigten sich zudem größere Abweichungen zwischen den beiden Beobachtern. Die Zuweisungen zur Kategorie Informationssuche erfolgten dagegen in nahezu vollständiger Konkordanz.

Tabelle 4

Übereinstimmungsmatrix der Kategorisierung durch zwei Beobachter in der Trainingsphase (B1 bzw. B2; Anzahl der Beobachtungen: $N = 198$ aus fünf Videoaufzeichnungen)

		B 1						Gesamt (B 1)
		1	2	3	4	5	6	
B 2	Kategorie							
	1 Systematisch	58	5	12	0	1	0	76
	2 Versuch und Irrtum	3	8	5	0	0	0	16
	3 Rigide Exploration	3	7	36	1	0	0	47
	4 Informationssuche	0	0	0	24	0	0	24
	5 Unterbrechung	4	0	1	0	22	3	30
	6 Restkategorie	0	1	2	0	0	2	5
Gesamt (B 2)		68	21	56	25	23	5	198

Um zu prüfen, ob die Gütekriterien auch für die einzelnen Kategorien zufriedenstellend sind (Bortz et al., 2008, S. 463 f.), wurde die Übereinstimmung der Beobachter in spezifischen Kategorien berechnet. Nach der Empfehlung von Wirtz und Caspar (2002, S. 77 f.) wurde zunächst bestimmt, wie häufig die spezifische Kategorie (z. B. systematische Exploration) beobachtet wurde. Diese Kategorie bildete nun die „Zielkategorie“, für welche die Beobachterübereinstimmung bestimmt werden sollte. In einer „Restkategorie“ wurden dann die Häufigkeiten der restlichen Kategorien aufsummiert. Beide Summen wurden in eine 2x2-Übereinstimmungsmatrix mit den beiden Kategorien Zielkategorie und Restkategorie übertragen. Für die Kategorien Systematische Exploration ($\kappa = .70$), Informationssuche ($\kappa = .98$) und Unterbrechung ($\kappa = .80$) ergaben sich zufriedenstellende bis ausgezeichnete Übereinstimmungen. Für die Kategorien Rigide Exploration ($\kappa = .58$) sowie Versuch-und-Irrtum ($\kappa = .37$) erwiesen sich die Beobachterurteile jedoch als unzureichend konkordant. Alle Kappa-Koeffizienten wurden auf dem Niveau von $p < .01$ signifikant. Aufgrund der Ergebnisse wurde für die Trainingsphase geschlussfolgert, dass die Trennschärfen zwischen den Kategorien Versuch-und-Irrtum, Systematische Exploration sowie Rigide Exploration noch nicht zufriedenstellend sind. Die 30sec-Beobachtungseinheit stellte sich als angemessen heraus.

Das Beobachtungssystem wurde im Anschluss an die ersten Videoanalysen einer Revision unterzogen. Dabei wurde angestrebt, die Eindeutigkeit der Zuordnung und damit auch die Reliabilität zu erhöhen. Im Hinblick auf die Zuverlässigkeit der Beobachtungskategorien

gelten nach Gellert (1955, S. 184, zit. nach Greve & Wentura, 1997, S. 75) folgende Richtlinien: „Je weniger Kategorien, je präziser definiert, je weniger zur Klassifikation an Schlussfolgerung und Interpretation notwendig, desto größer wird die Reliabilität der Daten sein“. Daraufhin wurde die Restkategorie entfernt und die operationale Definition der verbleibenden Kategorien präzisiert, um inhaltliche Überlappungen weiter einzugrenzen.

In der *Anwendungsphase* des Kategoriensystems wurden 13 weitere Videos von den beiden Beobachtern getrennt analysiert und die Ergebnisse erneut miteinander verglichen. In der folgenden Übereinstimmungsmatrix (siehe Tabelle 5) sind die Ergebnisse der Kodierungen in der Anwendungsphase zusammengefasst. Die Beobachterübereinstimmung erreichte hier mit $\kappa = .72$ (Asymptotische Signifikanz: $p = .001$) sowohl nach Greve und Wentura (1997) sowie nach Bortz und Döring (2003) eine gute Konkordanz zwischen den Beobachtern über alle Kategorien hinweg. Die Bewertung der Konkordanz in den einzelnen Kategorien ergab für die Systematische Exploration ($\kappa = .67$), die Rigide Exploration ($\kappa = .66$), die Informationssuche ($\kappa = .96$) sowie die Kategorie Unterbrechung ($\kappa = .87$) zufrieden stellende bis ausgezeichnete Übereinstimmungen. Alle Kappa-Koeffizienten für die einzelnen Kategorien waren auf dem Niveau von $p < .01$ signifikant.

Tabelle 5

Übereinstimmungsmatrix der Kategorisierungen durch zwei Beobachter (B1, B2) nach Präzision der Operationalisierung (Anzahl der Beobachtungen: N = 607 aus 13 Videoaufzeichnungen)

		B 1					
B 2		1	2	3	4	5	Gesamt (B 1)
	Explorationsstrategie						
	1 Systematisch	208	9	28	0	15	260
	2 Versuch und Irrtum	19	17	5	0	1	42
	3 Rigide Exploration	19	12	104	0	5	140
	4 Informationssuche	0	0	1	24	0	25
	5 Unterbrechung	2	2	2	1	133	140
	Gesamt (B 2)	248	40	140	25	154	607

Im Vergleich zur Trainingsphase konnte die Güte des Instrumentes in den Kategorien Rigide Exploration und Unterbrechung leicht verbessert werden (siehe Tabelle 6). Die Messgenauigkeit in der Kategorie Versuch-und-Irrtum konnte demgegenüber auch nach Revision der operationalen Definition der Kategorie nicht erhöht werden ($\kappa = .37$). Die Reproduzierbarkeit der Beobachtungen dieser Explorationsstrategie blieb unzureichend. Dies steht in Widerspruch zu den bisherigen empirischen Ergebnissen: Die in früheren empirischen Studien

verwendete Operationalisierung der Kategorie hatte sich als ausreichend reliabel erwiesen (vgl. van der Linden et al., 2001, 2003). Darüber hinaus zeigte sich in diesen Studien, dass die Auftretenswahrscheinlichkeit von Versuch-und-Irrtum beim Erkunden eines Computerprogramms genauso hoch war wie bei der rigiden Exploration. Die bisherigen empirischen Ergebnisse unterstützen die Relevanz dieser Kategorie, so dass sie weiterhin bei der Erfassung von Explorationsstrategien berücksichtigt wurde. Mögliche Erklärungen für diese inkonsistenten Ergebnisse zwischen der eigenen im Vergleich zu früheren Untersuchungen werden in der Diskussion zu dieser Studie erläutert (siehe Abschnitt 7.4).

Tabelle 6

Grad der Beobachterübereinstimmung (Cobens κ) zwischen zwei Beobachtern für die fünf Kategorien in der Trainings- bzw. Anwendungsphase der Erprobung des Kategoriensystems

Kategorie	Trainingsphase (N = 198)	Anwendungsphase (N = 607)
Systematische Exploration	.70	.67
Versuch-und-Irrtum	.37	.37
Rigide Exploration	.58	.66
Informationssuche	.98	.96
Unterbrechung	.80	.87

Anmerkungen. N gibt die Anzahl der Kodierungen in der jeweiligen Phase der Erprobung wieder.

Die fünf in der Trainingsphase kodierten Videos wurden nach Revision der Operationalisierungen erneut kodiert. Daraufhin wurden die 13 plus fünf Datensätze aus der Anwendungs- und Trainingsphase für die Analysen zu einem Datensatz zusammengefasst. Um die relative Häufigkeit jeder Explorationsstrategie zu erfassen, wurde der Anteil einer einzelnen Explorationsstrategie an der Anzahl aller Beobachtungseinheiten relativiert.

7.2.3 Hypothese 2: Nutzung von Explorationsstrategien

Gemäß Hypothese 2 wurde erwartet, dass bei studentischen Novizen das systematische Explorationsverhalten den größten Anteil beim Erkunden eines neuen Computerprogramms einnehmen, gefolgt von rigidem Verhalten und dem Vorgehen nach Versuch-und-Irrtum. Die Explorationsstrategie Informationssuche sollte in Relation zu allen anderen erfassten Strategien die niedrigste Prävalenzrate aufweisen. Abbildung 13 gibt den Anteil der einzelnen Strategien grafisch wieder:

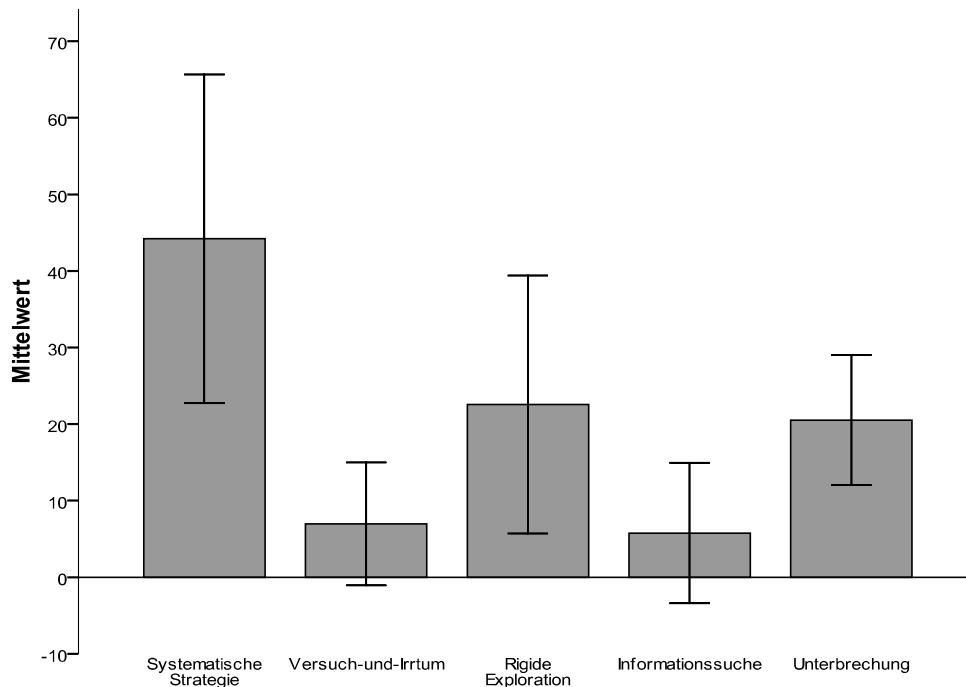


Abb. 13. Durchschnittliche relative Nutzungshäufigkeiten der einzelnen Explorationsstrategien (Systematische Strategie, Versuch-und-Irrtum, Rigide Exploration, Informationssuche) bzw. der Kategorie Unterbrechung in der Stichprobe ($N = 18$). Die Balken geben die Standardabweichungen in der jeweiligen Variablen wieder.

Die Teilnehmer verbrachten im Durchschnitt ungefähr 44% ($SD = 21.48$) der Bearbeitungszeit in SPSS mit dem systematischen Erkunden des Programms. Die unsystematische, rigide Exploration wies mit einer relativen Häufigkeit von circa 23% ($SD = 16.85$) den zweitgrößten Zeitanteil auf. Das Vorgehen nach Versuch-und-Irrtum wurde mit einer durchschnittlichen relativen Häufigkeit von ca. 7% ($SD = 8.02$) nur selten beobachtet. Die gleiche Tendenz zeigte sich für die Informationssuche mit ungefähr 6% ($SD = 9.17$). Der Anteil der Zeit, in der kein Explorationsverhalten der Teilnehmer auf der Bildschirmoberfläche erkennbar war („Unterbrechung“), belief sich im Durchschnitt auf ca. 21% ($SD = 8.52$) der Beobachtungseinheiten.

Wie die Fehlerbalken illustrieren, variierte der individuelle Anteil an systematischer bzw. rigider Exploration stark zwischen den Teilnehmern. Hypothese 2 konnte im Hinblick auf den relativen Anteil systematischer (d. h. optimaler) und rigider Explorationsweisen untermuert werden. Der Anteil der Kategorien Versuch-und-Irrtum sowie Informationssuche unterschied sich dagegen kaum. Nur ein geringer Anteil der Beobachtungseinheiten wurde mit diesen beiden Kategorien klassifiziert.

7.2.4 Hypothesen 3 und 4: Explorationsstrategien, Flow-Erleben und Leistung

Um erste Hinweise auf die Validität des Kategoriensystems zu erhalten, wurde in Hypothese 3 und 4 die Nutzung der Explorationsstrategien mit motivationalen sowie Leistungsvariablen in Verbindung gesetzt. Hypothese 3 impliziert, dass die Häufigkeit der Nutzung der optimalen, systematischen Explorationsstrategien von einem positiven Erlebenszustand begleitet wird. Demgegenüber sollte eine häufige Nutzung suboptimaler Strategien (Versuch-und-Irrtum, rigides Verhalten, übermäßige Informationssuche) negativ mit dem Flow-Erleben assoziiert sein. In Tabelle 7 sind die Interkorrelationen der Variablen aufgeführt.

Tabelle 7

Produkt-Moment-Korrelationen der Variablen in der Stichprobe ($N = 18$)

Variable	Interkorrelationen							
	2	3	4	5	6	7	8	9
1 Statistikvorwissenstest (6 Items)	-.06	.33	.60**	-.48*	-.59**	-.28	.62**	.62**
2 VECA		.46	-.17	.10	.22	.31	.02	-.17
3 PRACOWI (5 Items)			.40	-.21	-.30	-.04	.23	.49*
4 Systematische Exploration ¹				-.45	-.89**	-.64**	.47	.81**
5 Versuch-und-Irrtum ¹					.15	.02	-.14	-.13
6 Rigide Exploration ¹						.61**	-.54*	-.86**
7 Informationssuche ¹							-.34	-.65**
8 Flow_End								.65**
9 Performanz								

Anmerkungen. * $p < .05$, ** $p < .01$; zweiseitig. ¹ Relative Häufigkeit der Explorationsstrategie über den gesamten Messzeitraum (25 Minuten).

Entsprechend der Erwartungen stand das systematische Vorgehen in einer positiven Beziehung zum Flow-Erleben. Der Korrelationskoeffizient wurde jedoch knapp nicht statistisch signifikant ($r = .47$, $p = .052$). Wie erwartet korrelierte die Nutzungshäufigkeit der suboptimalen Explorationsstrategie Rigide Exploration negativ mit dem Flow-Erleben ($r = -.54$, $p = .021$). Hier zeigte sich ein mittelstarker Zusammenhang. Personen, die wiederholt ineffektive Funktionen aufriefen, gaben gleichzeitig geringere Flow-Werte an. Die Häufigkeit von Versuch-und-Irrtum ($r = -.14$) und der Informationssuche ($r_s = -.34$) standen ebenfalls in einem negativen Zusammenhang mit dem Flow-Erleben. Die Koeffizienten wurden jedoch nicht statistisch signifikant (vgl. Tabelle 7).

Personen, die bei der Erkundung und Anwendung des Computerprogramms systematisch vorgegangen waren, erzielten gleichzeitig mehr Punkte bei der Bearbeitung der Aufgaben ($r = .81, p = .001$). Im Gegensatz zu diesem Ergebnis waren alle Arten suboptimaler Strategien (Versuch-und-Irrtum, Rigide Exploration, Informationssuche) negativ mit der Leistung assoziiert. Die rigide Exploration des Computerprogramms ging nicht nur mit einem geringeren Flow-Erleben, sondern auch mit einem geringeren Leistungsniveau einher ($r = -.86, p = .001$). Die Nutzung systematischer beziehungsweise rigider Strategien war zudem in der Stichprobe hoch negativ miteinander korreliert ($r = -.89, p = .001$). Auch die Häufigkeit der Kategorie Informationssuche stand in einem mittelstarken, negativen Zusammenhang mit der Performanz ($r = -.65, p = .003$).

Um den Effekt des Vorwissens zu überprüfen, wurden auch für die Variablen Statistikvorwissenstest, praktisches Computerwissen (PRACOWI) und Vertrautheit mit Computeranwendungen (VECA) Zusammenhänge mit der Leistung, der Strategienutzung sowie dem Flow-Erleben berechnet (siehe Tabelle 7). Es zeigt sich, dass in dieser Stichprobe insbesondere das Statistikvorwissen mit einem systematischen Explorationsverhalten einherging ($r = .60, p = .008$) und negativ mit allen suboptimalen Explorationsstrategien assoziiert war. Darüber hinaus zeigten sich mittlere, positive Zusammenhänge mit dem Flow-Erleben ($r = .62, p = .003$) sowie mit der Leistung ($r = .62, p = .006$). Auch bezüglich des PRACOWI gingen höhere Werte im Vorwissen mit einer höheren Leistung einher ($r = .49, p = .04$). Die Beziehung zwischen PRACOWI und systematischer Exploration wurde nicht signifikant ($r = .40, p = .10$). Der VECA wies insgesamt die niedrigsten Effekte des Zusammenhangs auf.

Tabelle A - 9 im Anhang (Abschnitt A 2.2.3) gibt die Partialkorrelationen der Variablen wieder, für die Bedingung, dass der Einfluss des Statistikvorwissens und des praktischen Computerwissens herauspartialisiert wurde. Auch unter Kontrolle des Vorwissens zeigt sich, dass die systematische Explorationsstrategie positiv mit der Leistung assoziiert war und negativ mit zwei der suboptimalen Explorationsstrategien (rigide Exploration, Informationssuche). Der Zusammenhang von systematischer Explorationsstrategie und Leistung fiel geringer aus als bei der Korrelation erster Ordnung (siehe oben): Beide Variablen wiesen einen schwachen, positiven Zusammenhang auf. Dies ist damit zu erklären, dass das Statistikvorwissen und die Leistung einen großen Anteil gemeinsamer Varianz besitzen, der nun herauspartialisiert wurde. Personen mit großem Vorwissen zeigten auch eher höhere Flow-Werte und waren gleichzeitig leistungsstärker. Die Häufigkeit rigider Exploration stand wie erwartet in negativem Zusammenhang mit der Leistung sowie mit dem Flow-Erleben. Letztere Beziehung erreichte jedoch keine statistische Signifikanz. Das Flow-Erleben korrelierte nicht statistisch signifikant, aber erwartungsgemäß positiv mit der Leistung.

Interpretationsmöglichkeiten für diese Befunde werden in der Diskussion zur Studie aufgegriffen (siehe Abschnitt 7.3). Die Ergebnisse unterstützen die in Hypothese 3 und 4 postulierten Annahmen insofern, dass ein systematisches Vorgehen sowohl stärker mit einem positiven Erlebenszustand während des Erlernens eines Computerprogramms einhergeht als auch mit einer besseren Leistung bei der Anwendung des Programms assoziiert ist. Unter den suboptimalen Explorationsstrategien zeigte insbesondere die rigide Exploration eine ungünstige Beziehung sowohl zur Leistung als auch zum Flow-Erleben.

7.3 Zusammenfassung und Diskussion

Im Folgenden werde ich zunächst auf die einzelnen Befunde zur Entwicklung des Kategoriensystems eingehen und diese interpretieren. Im Anschluss daran erfolgt eine zusammenfassende Diskussion im Hinblick auf Einschränkungen des entwickelten Beobachtungssystems. Im Forschungsausblick werden Hinweise zum praktischen Einsatz sowie zur Generalisierbarkeit des Kategoriensystems gegeben. Zusätzlich werde ich auf die Anwendungsziele des Kategoriensystems im Rahmen der beiden weiteren Studien dieser Arbeit eingehen.

Messgenauigkeit des Beobachtungssystems. Um ein interessierendes Konstrukt reliabel und valide erfassen zu können, muss ein Messinstrument bestimmten Gütekriterien genügen (vgl. Bortz & Döring, 2003; Kelava & Moosbrugger, 2007; Schermelleh-Engel & Werner, 2007). Das Beobachtungssystem zur Erfassung von Explorationsstrategien wies insgesamt eine gute Zuverlässigkeit auf. Allerdings ist eine reliable Messung des Explorationsverhaltens nicht in allen Kategorien zufriedenstellend gelungen. Mängel zeigten sich vor allem in Bezug auf die Strategie Versuch-und-Irrtum, bei der die Beobachterübereinstimmung besonders gering ausfiel. Die Kategorie wies sowohl mit der systematischen Exploration als auch mit der rigiden Exploration Überlappungen auf.

Kriteriumsvalidierung. Im Rahmen einer ersten Kriteriumsvalidierung wurde in Studie 1 überprüft, ob mittels des Kategoriensystems tatsächlich optimale bzw. suboptimale Explorationsverhaltensweisen (vgl. Dörner 1980) erfasst werden. Entsprechend den Annahmen zeigte sich, dass die systematische Explorationsstrategie mit einem höheren Flow-Erleben sowie einer höheren Leistung einherging. Vor allem die Kategorien Systematische Exploration und Rigide Exploration wiesen starke Beziehungen zur Leistung auf. Personen, die häufiger systematisch explorierten, erzielten beim Lösen der Aufgaben mehr Punkte. Gleichzeitig hatten sie eher das Gefühl von der Aufgabe absorbiert zu sein und den Verlauf des Prozesses unter Kontrolle zu haben. Demgegenüber führte die rigide Exploration dazu, dass die Teilnehmer nur wenige Aufgaben in der vorgegebenen Zeit richtig lösen konnten. Darüber hinaus

stand das aufgabenrelevante Statistikvorwissen ebenfalls in Beziehung zu einer systematischen Exploration des Programms. Personen mit größerem Vorwissen nutzten häufiger ein zielführendes, optimales Vorgehen. Diese Befunde unterstützen die Annahme eines validen Messinstrumentes mit dessen Hilfe leistungs- und motivationsrelevantes Verhalten am Computer erfasst werden kann.

Generalisierbarkeit der Befunde. Ein zentraler Kritikpunkt ergibt sich hinsichtlich der Stichprobenziehung: Die Daten in Studie 1 wurden anhand einer Gelegenheitsstichprobe erhoben. Darüber hinaus konnten ausschließlich weibliche Personen in die Entwicklung und Erprobung des Kategoriensystems einbezogen werden. Die Aussagekraft der Ergebnisse kann begrenzt sein, wenn das Nutzerverhalten geschlechtsspezifisch geprägt ist (vgl. Studie 2). Dies könnte zur Folge haben, dass in dem neu entwickelten Beobachtungssystem nicht alle exploratorischen Verhaltensweisen am Computer erfasst wurden, die theoretisch beobachtet werden können. Im Vergleich zu Ergebnissen früherer Studien (van der Linden et al., 2001, 2003) zeigten die (weiblichen) Teilnehmer in dieser Studie jedoch eine große Variabilität des Explorationsvorgehens. Die Kodierungen des Explorationsverhaltens streuen über alle auf der Basis von van der Linden et al. (2001) vorgeschlagenen Kategorien. Dies belegen die unabhängigen Kodierungen der beiden Beobachter. Eine Ausnahme bildete die Kategorie Versuch-und-Irrtum: Hier wich die Beobachterübereinstimmung deutlich von denjenigen der Vorgängerstudien ab. Dies könnte der Homogenität der Stichprobe anzulasten sein.

Die Stichprobengröße ist neben der Art des statistischen Tests, der Größe des Effekts sowie dem Signifikanzniveau und der Teststärke ein entscheidender Faktor für das Ergebnis des Signifikanztests (Rost, 2007, S. 95). Obwohl die Stichprobe in Studie 1 nur 18 Personen umfasste, wurden mehrere erwartete Zusammenhangskoeffizienten statistisch signifikant. Das Ergebnis kann als Hinweis für die Stichprobenauswahl in Folgestudien genutzt werden. Zusätzlich zum Signifikanzniveau wurde stets auch die Effektstärke mit in Betracht gezogen, um die praktische Bedeutung des Effekts abschätzen zu können.

Fazit und Forschungsausblick. Auf der Grundlage des Explorationsverhaltens von 18 weiblichen Studierenden am Computer ist es gelungen ein reliables und valides Messinstrument zur Klassifizierung des Explorationsverhaltens am Computer zu entwickeln. Zur Erhöhung der Anwendungsobjektivität enthält das Kategoriensystem Beispiele für charakteristische Indikatoren der jeweiligen Explorationsstrategie. Die genannten Indikatoren bilden jedoch keine erschöpfende Menge aller Verhaltensweisen, die mit Hilfe des Kategoriensystems klassifiziert werden müssen. Die eindeutige Zuordnung ist dementsprechend von der präzisen (operationalen) Definition der Kategorien abhängig. Im Hinblick auf die operationale Definition der Kategorie Versuch-und-Irrtum war diese eindeutige Zuordnung aus den oben ge-

nannten Gründen nicht erfolgreich. Hier weist das Kategoriensystem trotz Revision weiterhin Einschränkungen auf.

Das Kategoriensystem wurde im Folgenden an zwei weiteren Datensätzen angewendet (siehe Studien 2 und 3; Kapitel 8 und 9). Die beobachteten Explorationsstrategien wurden mit motivationalen und kognitiven Variablen in Bezug gesetzt. Die im nächsten Kapitel dargestellte Studie 2 hatte zum Ziel geschlechtsspezifische Unterschiede im Explorationsverhalten am Computer zu erforschen sowie Erklärungsansätze für mögliche Leistungsunterschiede zwischen den Geschlechtern beim Erlernen von Computerprogrammen erforschen.

8. Geschlechtsunterschiede auf der Basis des kognitiv-motivationalen Prozessmodells (Studie 2)

In Studie 2 wurden mehrere Forschungsfragen thematisiert: Die Entwicklung des Kategoriensystems zur Strategiesystematik in Studie 1 basierte auf den Daten einer Stichprobe, die sich ausschließlich aus weiblichen Teilnehmern zusammensetzte. Im Sinne einer Kreuzvalidierung sollte das Beobachtungssystem in Studie 2 zusätzlich an einer heterogenen Stichprobe aus männlichen und weiblichen Probanden geprüft werden. Im Mittelpunkt stand dabei die Frage, ob das neu entwickelte Instrument erlaubt, interindividuelle Differenzen im Explorationsverhalten bei männlichen und bei weiblichen Computernutzern gleichermaßen abzubilden.

Das primäre Forschungsanliegen von Studie 2 bestand darin zu untersuchen, ob Geschlechtsunterschiede in der Leistung am Computer mit Hilfe des kognitiv-motivationalen Prozessmodells (Vollmeyer & Rheinberg, 1998, 2000, 2006) erklärt werden können. Es wurde überprüft, inwieweit Geschlechtsdifferenzen in der Eingangsmotivation sowie in den Mediatoren (Flow und Strategiesystematik) zu Leistungsunterschieden beitragen. Aufgrund der theoretischen Befunde (vgl. Kapitel 5 und 6) wurde erwartet, dass eine günstige Eingangsmotivation, ein höheres Flow-Erleben und eine optimale Strategiesystematik bei Männern zu einer höheren Leistung führen. Um bereits vor der Untersuchung bestehende Geschlechtsunterschiede in der Computererfahrung und den Statistikkenntnissen kontrollieren zu können, wurden diese Kenntnisse auch in Studie 2 miterhoben.

Für Studie 2 wurde eine Population ausgewählt, in der männliche und weibliche Personen zu etwa gleichen Anteilen vertreten sind.¹⁵ Das Geschlechterverhältnis im Fach Wirtschaftswissenschaften (Betriebswirtschaftslehre, Volkswirtschaftslehre, Wirtschaftspädagogik) an der Goethe-Universität Frankfurt/Main erwies sich vor der Datenerhebung im Sommer 2005 als annähernd ausgewogen. Zudem stellte das Teilgebiet Statistik einen elementaren Bestandteil des Curriculums in wirtschaftswissenschaftlichen Studiengängen dar. Die Population der Studierenden der Wirtschaftswissenschaften (Diplom), die studienrelevante Computersoftware nutzen, erfüllte damit im Hinblick auf das Geschlechterverhältnis sowie die erforderlichen theoretischen Statistikkenntnisse die zentralen Anforderungen an die Studienteilnehmer.

¹⁵ Der aktuellen Studierendenstatistik der Goethe Universität Frankfurt/Main zufolge waren am 10. November 2008 1108 Männer und 807 Frauen im Fach Wirtschaftswissenschaften eingeschrieben. Zum Vergleich: Das Fach Psychologie studierten zum gleichen Zeitpunkt 557 Frauen und 141 Männer. Leider waren die Angaben für das SoSe 2005 zum Zeitpunkt des Zugriffs nicht mehr verfügbar. Quelle: <https://qis.server.uni-frankfurt.de/stat/SQLServlet>, Zugriff am 10. November 2008.

Die Datenerhebung erfolgte in einem Projekt zur geschlechtsspezifischen Nutzung computergestützter Technologien im Studium (Projektleitung: Prof. Dr. Regina Vollmeyer, Prof. Dr. Margarete Imhof), welches durch die Förderung kleinerer Gender- und Frauenforschungsprojekte an der Goethe-Universität Frankfurt ermöglicht wurde (vgl. Vollmeyer, Imhof & Beierlein, 2006). Zusammenfassend wurden in Studie 2 die folgenden Fragen untersucht (s. Abschnitt 6.2):

1. Bildet das Kategoriensystem interindividuelle Unterschiede im Explorationsverhalten bei männlichen und weiblichen Computernutzern ab?
2. Zeigen Männer und Frauen Unterschiede in Vorerfahrungen mit dem Computer sowie in der Eingangsmotivation vor Beginn des Explorationsprozesses?
3. Nutzen Männer und Frauen beim Erlernen eines neuen Computerprogramms unterschiedliche Explorationsstrategien?
4. Führen Unterschiede in Flow-Erleben und Strategiesystematik zu Leistungsdifferenzen zwischen den Geschlechtern?

Zur besseren Nachvollziehbarkeit der in Studie 2 getesteten Annahmen sind die bereits in Abschnitt 6.2 dargestellten Hypothesen an dieser Stelle erneut aufgelistet:

Hypothese 5: Frauen haben eine geringere Computervorerfahrung als Männer, d. h. Frauen nutzen weniger häufig Computerprogramme wie Statistiksoftware, Tabellenkalkulationsprogramme oder Datenbanken.

Hypothese 6: Männer und Frauen unterscheiden sich in der Eingangsmotivation dahingehend, dass Frauen eine ungünstigere Motivation vor Beginn der Aufgabenbearbeitung aufweisen: Sie fühlen sich weniger herausgefordert, haben eine geringere Erfolgswahrscheinlichkeit, sind weniger interessiert und haben gleichzeitig eine höhere Misserfolgsbefürchtung.

Hypothese 7: Männer nutzen häufiger eine systematische Explorationsstrategie beim Erlernen neuer Programme als Frauen (Hypothese 7a). Männer geben ein stärkeres Flow-Erleben an als Frauen (Hypothese 7b).

Hypothese 8: Männer erreichen bei der Exploration von SPSS eine höhere Leistung (Punktzahl bei der Aufgabenbearbeitung) als Frauen.

Hypothese 9: Die Eingangsmotivation wirkt sich positiv auf das Flow-Erleben sowie auf die Häufigkeit systematischer Exploration aus. Beide Variablen haben wiederum einen positiven Effekt auf die Performanz (*Mediationshypothese*; s. Abb. 9, Kapitel 6.2).

8.1 Methode

8.1.1 Stichprobe

Insgesamt nahmen 40 Studierende (22 männliche, 18 weibliche) der Betriebswirtschaftslehre bzw. der Volkswirtschaftslehre der Goethe-Universität Frankfurt an der Studie teil. Es wurden ausschließlich Teilnehmer geworben, die nach eigenen Angaben Novizen im Umgang mit SPSS waren (vgl. Studie 1, Abschnitt 7.1.1). Der Altersdurchschnitt der Teilnehmer lag bei $M = 22.28$ Jahren ($SD = 1.43$). Die durchschnittliche Fachsemesteranzahl betrug in der Stichprobe $M = 3.33$ ($SD = 1.20$). Wie bereits in Studie 1 der Fall, handelte es sich auch in Studie 2 um eine Gelegenheitsstichprobe: Die Teilnehmer wurden in einer verpflichtenden Statistikvorlesung bzw. in studentischen Tutorien geworben. Als Hauptziel der Studie wurde die Erforschung der Motivation während des Erlernens eines studienrelevanten Statistikprogramms genannt. Studierende, die Interesse an der Studie signalisierten, wurden von der Versuchsleiterin per E-Mail zur Teilnahme eingeladen. Angestrebt wurde ein ausgeglichener Anteil von männlichen und weiblichen Teilnehmern.

Da es sich bei der Teilnahme der angehenden Wirtschaftswissenschaftler nicht um einen obligatorischen Teil ihres Studiums handelte, wurden mehrere Anreize zur Mitarbeit gesetzt. Insbesondere wurde auf die Möglichkeit hingewiesen, ein neues, studienrelevantes Computerprogramm zu erlernen. Des Weiteren wurde den Studierenden eine Teilnahme an einer SPSS-Schulung angeboten, die auch schriftlich bescheinigt wurde. Die Studierenden wurden darauf hingewiesen, dass sie die Bescheinigung zukünftigen Bewerbungen um Praktika oder Nebentätigkeiten beilegen könnten. Darüber hinaus wurde den Studierenden zugesichert, dass die in der Studie verwendeten SPSS-Aufgaben auf den Vorlesungsstoff abgestimmt seien. Dies gäbe ihnen Gelegenheit, ihre theoretischen Statistikkenntnisse vor der Vorlesungsklausur zu überprüfen. Außerdem wurde jedem Teilnehmer eine finanzielle Belohnung von sieben Euro in Aussicht gestellt.

Aufgrund technischer Probleme konnten nicht alle Datensätze der 40 Teilnehmer in die Analysen einbezogen werden: Die Datensätze von sechs Personen mussten vollständig ausgeschlossen werden (listenweiser Fallausschluss), da die Videoaufzeichnungen der SPSS-Sitzungen fehlerhaft waren. Die Programmstörungen beim Ablauf des Screenshotprogramms, die zum Datenausfall führten, wurden auf Inkompatibilitäten an den eingesetzten PCs bzw. Laptops zurückgeführt. Der Datenausfall war dabei unabhängig von den persönlichen Charakteristika des Teilnehmers. Auswirkungen der Programmstörungen auf die Prozessvariablen (Motivation, Strategienutzung) konnten jedoch nicht ausgeschlossen werden.

Die Daten einer weiteren Person mussten ebenfalls vollständig aus den Analysen ausgeschlossen werden, da Lösungsschritte des zuvor getesteten Teilnehmers vom Programm aufge-

zeichnet worden waren und diese dem nächstfolgenden Teilnehmer während der Bearbeitung der Aufgaben sichtbar wurden. Folglich konnte die Lösung der Aufgaben mit SPSS nicht eindeutig auf die Leistung des nachfolgenden Teilnehmers zurückgeführt werden. Auch hier war der Datenausfall einem technischen Problem geschuldet. Insgesamt mussten somit die Datensätze von sieben Personen ausgeschlossen werden (ca. 18%). Dieser Personenanteil stellt einen beträchtlichen Teil der Stichprobe dar und führte zu einer erheblichen Verringerung des Stichprobenumfangs. Es verblieb ein Stichprobenumfang von $N = 33$ von 18 männlichen und 15 weiblichen Teilnehmern für weitere Analysen.

8.1.2 Vorgehen

Die Datenerhebung erfolgte im Sommersemester 2005. Die Versuchsanordnung war ein quasi-experimentelles Design, bei dem das Verhalten von zwei nicht-randomisierten natürlichen Gruppen (Männer, Frauen) im Mittel miteinander verglichen wurde. Aufbau und Ablauf sind in Abbildung 14 grafisch dargestellt und entsprachen weitgehend dem Vorgehen in Studie 1.

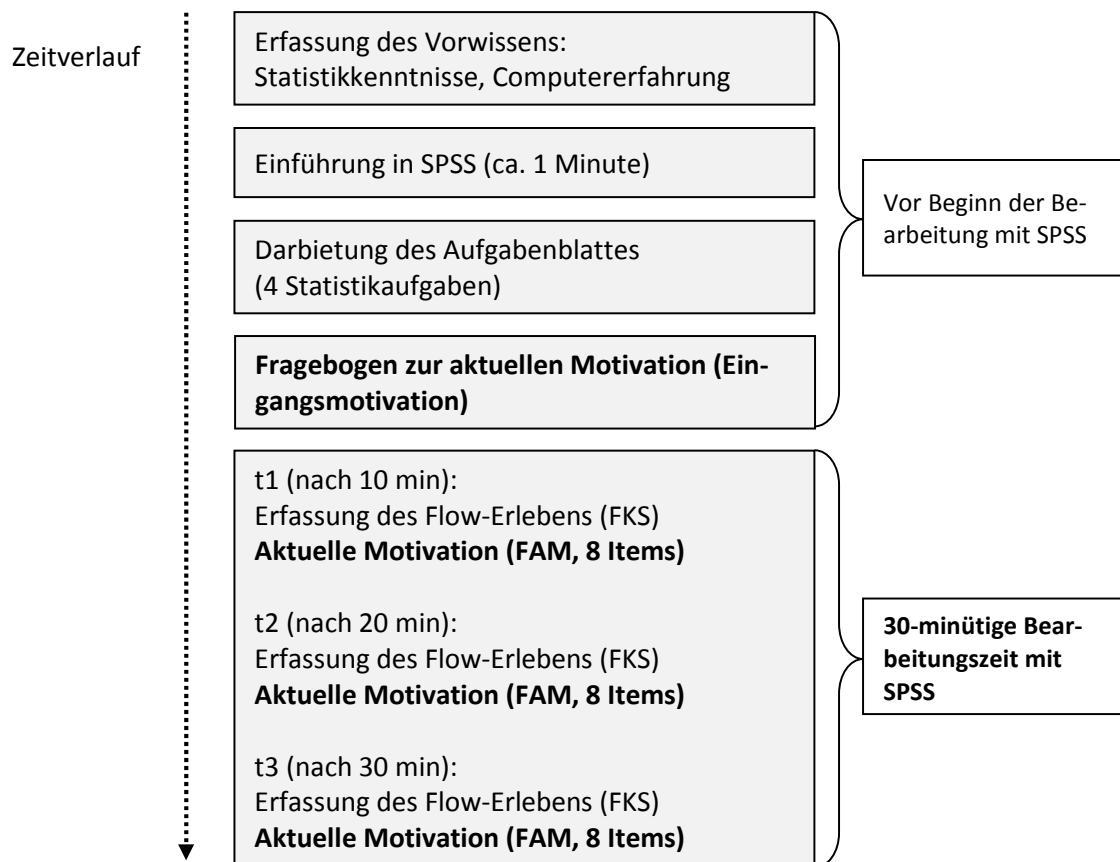


Abb. 14. Versuchsplan von Studie 2 im Zeitverlauf mit Angabe der erfassten Konstrukte. Modifikationen im Vergleich zu Studie 1 sind fettgedruckt.

Im Folgenden werden lediglich Abweichungen von der bereits in Studie 1 beschriebenen Methodik aufgezeigt (vgl. Kap. 7.1.2). Anders als in Studie 1 hatten die Teilnehmer 30 Minuten statt 25 Minuten Zeit zur Exploration des Programms SPSS 12.0. Diese Veränderung des Versuchsdesigns ermöglichte es, einen dritten Messzeitpunkt direkt vor Beendigung der SPSS-Sitzung einzuführen. Ziel war dabei, möglichst kurz vor Ende der Bearbeitungszeit (d.h. nach ca. 30 min) die aktuelle Motivation und das Flow-Erleben der Teilnehmer zu erfassen. Im Unterschied zu Studie 1 beteiligten sich in Studie 2 weibliche und männliche Studierende. In einzelnen Fällen nahmen drei Personen gleichzeitig an der Studie teil. In diesen Fällen wurde darauf geachtet, die Personen so auf die beiden Laborräume zu verteilen, dass sich Personen gleichen Geschlechts im gleichen Laborraum befanden. Diese Aufteilung war in der überwiegenden Anzahl von Fällen realisierbar.

8.1.3 Materialien

8.1.3.1 Fragebogen zu Statistikkenntnissen und Computernutzung

Erfassung der Statistikkenntnisse. Die Erfassung des Statistikvorwissens stimmte zum großen Teil mit dem in Studie 1 gewählten Vorgehen überein (siehe Abschnitt 7.1.3.1). Erfasst wurde zunächst, ob die Teilnehmer die verpflichtende(n) Statistik-Grundlagenveranstaltung(en) sowie die dazugehörigen Tutorien besuchten. Darüber hinaus gaben die Teilnehmer an, auf welche statistischen Inhaltsgebiete sie ihr Wissen anwenden konnten. Die Selbsteinschätzung der Statistikkenntnisse erfolgte erneut auf einer 5-stufigen Likertskala mit den Endpolen „sehr gut“ (1) und „mangelhaft“ (5).

Auch in Studie 2 wurde zusätzlich ein Mehrfach-Antwort-Test mit insgesamt neun Items entworfen. Die Items wurden so zusammengestellt, dass deren Inhalt mit dem spezifischen theoretischen Kenntnisstand in der Population der Studierenden der Wirtschaftswissenschaften übereinstimmte. Als inhaltliche Referenz diente hierzu die von den Teilnehmern besuchte Statistikvorlesung im Fachbereich Wirtschaftswissenschaften, in der die Teilnehmer auch rekrutiert worden waren. Die Wahl der statistischen Termini und Indizes erfolgte auf der Grundlage des Skripts zur Statistikvorlesung. Sechs der neun Items entsprachen den Items aus Studie 1. Hinzu kamen drei Items, in denen Wissen zum Skalenniveau von Messwerten, zur Kovarianz zweier Zufallsvariablen sowie zur Bedeutung des Pearson-Produkt-Moment-Korrelationskoeffizienten (siehe Anhang A 3.1.1) abgefragt wurden.

Drei der neun Items (Nr. 1 „Median“, Nr. 6 „Skalenniveau“, Nr. 7 „Kovarianz“) wurden aufgrund ihrer unzureichenden Trennschärfen zwischen $r_{it} = 0$ und $.13$ bei der Berechnung des Summenscores des Statistikvorwissenstests ausgeschlossen (siehe Tabelle A - 8 im Anhang; vgl. Bühner, 2006; Kelava & Moosbrugger, 2007). Die interne Konsistenz der Skala betrug nach Elimi-

nierung dieser drei Items Cronbachs $\alpha = .64$ (6 Items) anstelle von $\alpha = .57$ (9 Items). Die Schwierigkeiten der sechs Items variierten zwischen $P = 50\%$ und $P = 75\%$. Dieses Spektrum beinhaltete dabei überwiegend mittelschwere Items. In den sechs verbleibenden Testitems konnten insgesamt 22 Punkte erreicht werden.

Angaben zur Computererfahrung. Abweichend von dem Vorgehen in Studie 1 wurde im Rahmen von Studie 2 die Computererfahrung lediglich über drei Items erfasst. In Studie 2 wurde auf einen Datensatz aus dem Jahr 2005 zurückgegriffen, da der Datensatz Daten von männlichen sowie weiblichen Teilnehmern beinhaltete. In der Studie war jedoch kein standardisiertes Verfahren zur Erfassung der Computerbildung eingesetzt worden (vgl. Naumann et al., 2001; Richter, Naumann & Groeben, 2001). Die Operationalisierung der Computervorerfahrung erfolgte über die selbstberichtete Häufigkeit der Beschäftigung mit dem Computerprogramm. Die Messung war auf diejenigen Computerprogramme begrenzt, die von den Projektleiterinnen dem Statistikprogramm SPSS als inhaltlich und funktional ähnlich eingestuft wurden. Die Teilnehmer wurden aufgefordert, die Häufigkeit anzugeben, mit der sie in den letzten sechs Monaten mit folgenden Computeranwendungen gearbeitet hatten: Statistikprogramme, Tabellenkalkulation und Datenbanken. Ihre Computervorerfahrung mit den genannten Programmgruppen konnten die Teilnehmer auf einer Skala mit den Stufen 0 = „nie“, 1 = „bis 10-mal“, 2 = „bis 20-mal“, 3 = „häufiger als 20-mal“ einstufen. Damit sollte eine möglichst valide Messung des zeitlichen Umfangs der Beschäftigung mit den jeweiligen Programmgruppen in einem zuvor definierten Zeitraum erreicht werden. Die Antworten für die drei Bereiche wurden zu einem Summenwert addiert (Maximalwert: 12). Die Itemschwierigkeiten variierten im niedrigen bis mittleren Bereich ($P = 2\%$ bis 46%). Die Itemtrennschärfen für die Nutzung von Statistikprogrammen, Tabellenkalkulation und Datenbanken lagen bei $r_{it} = .36, .36$ bzw. $.41$, wenn alle drei Programme in einer Skala zusammengefasst wurden.

8.1.3.2 Fragebogen zur aktuellen Motivation (FAM)

Messung aktueller Motivation in Lern- und Problemlösesituationen. Boekaerts (1986, S. 231) empfiehlt, die aktuelle Motivation in dem Moment zu messen, in dem der Lernende mit der Lernaufgabe konfrontiert wurde. Vor Beginn des Explorierens des Computerprogramms beantworteten die Teilnehmer den Fragebogen zur aktuellen Motivation (FAM; Rheinberg et al., 2001). Der FAM diene als *State*-Maß der aktuellen Lernmotivation und wurde in der vorliegenden Untersuchung genutzt, um ex ante, d. h. vor Beginn des Lernprozesses, die motivationsrelevanten Erlebens- und Handlungskomponenten zu erfassen (vgl. Vollmeyer & Rheinberg, 2003, S. 283). Der Theorie zufolge werden diese durch ein problemorientiertes Explorieren angeregt (vgl. Winther, 2006, S. 178). Die Items des FAM wurden leicht verändert, um den Inhalt an die Aufgabenstellung

anzupassen. So wurde das Original-FAM-Item „Bei der Aufgabe mag ich die Rolle des Wissenschaftlers, der Zusammenhänge entdeckt“ folgendermaßen umformuliert: „Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen“ (siehe Abschnitt A 1.3 im Anhang).

Beschreibung der Skalen. Die von Rheinberg et al. (2001) berichteten Ergebnisse einer Faktorenanalyse der 18 Fragebogenitems ergaben vier unabhängige Faktoren. Diese umfassen affektive und kognitive Bewertungen der aktuellen Lern- bzw. Problemlösesituation: Erfolgswahrscheinlichkeit, Interesse, Herausforderung und Misserfolgsbefürchtung. Zwei der vier Faktoren stehen stärker in Bezug mit der Wertkomponente der Motivation (Interesse, Herausforderung). Die anderen beiden Motivationsaspekte beziehen sich stärker auf die Erwartungskomponente der Motivation (Erfolgswahrscheinlichkeit, Misserfolgsbefürchtung). Der Faktor *Erfolgswahrscheinlichkeit* misst mit vier Items (davon zwei invertiert) die subjektive Einschätzung darüber, wie gut die Person bei einer Aufgabe abschneiden wird (Rheinberg et al., 2001). Aus dem Abgleich der eingeschätzten Problemlösefähigkeiten und den wahrgenommenen Anforderungen der Aufgabe ergibt sich „der Grad an Zuversicht, mit der man die anstehende Aufgabe in Angriff nimmt“ (Vollmeyer & Rheinberg, 2003, S. 284). Die vier Items des Faktors *Herausforderung* erfassen, inwieweit die Aufgabe kognitiv-situational als leistungsthematisch aufgefasst wird und sich die Person um ein gutes Abschneiden bei der Aufgabe bemüht (Vollmeyer & Rheinberg, 2003; Winther, 2006). Der Faktor *Misserfolgsbefürchtung* thematisiert mit Hilfe von fünf Items die Furcht vor Misserfolg in einer konkreten Lernsituation (Rheinberg et al., 2001). Thematisiert wird das Ausmaß der Betroffenheit im Falle eines Misserfolgs und nicht die Erwartung eines Misserfolgs. Die Misserfolgsbefürchtung ist damit nicht als Gegenpol des Faktors Erfolgswahrscheinlichkeit zu verstehen, wie Vollmeyer und Rheinberg (2003) betonen. Die fünf Items der Skala *Interesse* spiegeln den Person-Gegenstands-Bezug im Sinne der Interessentheorie (vgl. Krapp, 1998, 2005) wider. Bei diesem Faktor steht nicht die Einschätzung der eigenen Tüchtigkeit im Vordergrund. Stattdessen ist der Anreiz für die Auseinandersetzung mit dem Problem bzw. dem Lerngegenstand in der Bewertung des Lerngegenstands selbst begründet. Wird die Auseinandersetzung mit dem Lerngegenstand als wichtig erlebt, drückt sich dies in einem hohen Interesse am Lerngegenstand aus (vgl. Vollmeyer & Rheinberg, 2003). Beispielitems für die Skalen des FAM sind die folgenden: „Ich glaube der Schwierigkeit der Aufgabe gewachsen zu sein“ (Erfolgswahrscheinlichkeit), „Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde“ (Herausforderung), „Solche Aufgaben würde ich auch in meiner Freizeit bearbeiten“ (Interesse), „Ich fürchte mich ein wenig davor, dass ich mich hier blamieren könnte“ (Misserfolgsbefürchtung) (siehe Abschnitt A 1.3 im Anhang).

Antwortmodus der Skalen. Der Antwortmodus der FAM-Skalen ist eine 7-stufige Rating-skala mit den Endpolen „trifft nicht zu“ (1) bis „trifft zu“ (7). Da die einzelnen Skalen theoretisch unabhängige Aspekte der Motivation erfassen, empfehlen die Autoren für jede Skala einen eigenen Skalenmittelwert zu bilden. Die interne Konsistenz der einzelnen Skalen wird im Durchschnitt mindestens als zufriedenstellend beschrieben (Range der internen Konsistenzen in verschiedenen Skalen und Stichproben: Cronbachs α von .66 bis .90; siehe Rheinberg et al., 2001).

Messung der aktuellen Motivation während der Bearbeitung. Um die aktuelle Motivation im Explorationsprozess zu mehreren Messzeitpunkten erheben zu können („On-line Erfassung“, vgl. Vollmeyer & Rheinberg, 2003, S. 289), wurde eine Kurzform des FAM zur Operationalisierung des Konstrukts entwickelt. Die Auswahl der Items erfolgte anhand der von Rheinberg et al. (2001) vorgelegten Faktorladungsmatrix: Es wurden jeweils die zwei Items der vier Skalen des FAM in die Kurzversion aufgenommen, welche die höchste Faktorladung aufwiesen. Dazu zählten Item 9 und 12 für Faktor 1 „Misserfolgsbefürchtung“ (Ladungen: .85, .84), Item 2 und 3 für Faktor 2 „Erfolgswahrscheinlichkeit“ (Ladungen: -.78, .74), Item 8 und 10 für Faktor 3 „Herausforderung“ (Ladungen: .79, .71), sowie Item 17 und 4 für Faktor 4 „Interesse“ (Ladungen: .75, .68). Die ausgewählten Items sind inklusive ihrer Zuordnung zu den FAM-Skalen in Abschnitt 1.5 des Anhangs aufgeführt. Die Kurzversion des FAM wurde den Teilnehmern zeitgleich mit den Items der Flow-Kurz-Skala (s. u.) vorgelegt, um den Arbeitsprozess möglichst wenig zu stören.

Durch die Verkürzung des Fragebogens von ursprünglich vier bis fünf Items auf zwei Items pro Skala waren Einschränkungen in der Höhe der Reliabilität der jeweiligen Skala zu erwarten (vgl. Kelava & Moosbrugger, 2007). Aus diesem Grund wurde auf eine Auswertung auf der Ebene der einzelnen Skalen verzichtet. Stattdessen wurden die acht ausgewählten Items als Indikatoren eines übergeordneten Faktors der aktuellen Motivation betrachtet (vgl. Vollmeyer & Rheinberg, 2006). Im Sinne von Heckhausens Konzept der „Netto-Hoffnung“ (Heckhausen, 1963; vgl. Brunstein & Heckhausen, 2006, S. 150) wurde ein Differenzwert der acht Itemwerte pro Messzeitpunkt gebildet (im Folgenden als „Netto-Motivation“ bezeichnet): Die Mittelwerte in den Skalen Erfolgswahrscheinlichkeit, Herausforderung und Interesse wurden addiert und der Mittelwert in der Skala Misserfolgsbefürchtung wurde subtrahiert.

Item- und skalenanalytische Prüfung. Zunächst werden die Befunde für die Skalen der *Eingangsmotivation* unmittelbar vor Beginn der Aufgabenbearbeitung berichtet (siehe auch Tabelle A – 11 im Anhang). Die einzelnen Skalen des FAM wurden separaten item- und skalenanalytischen Prüfungen unterzogen. Detaillierte Angaben zu Itemtrennschärfen, Itemschwierigkeiten sowie der internen Konsistenzen der vier Skalen der Eingangsmotivation sind Tabelle A-9 im Anhang zu entnehmen. Für die internen Konsistenzen der Skalen ergaben sich die folgenden Kennwerte:

Cronbachs $\alpha = .66$ für die Skala „Erfolgswahrscheinlichkeit“ (4 Items), $\alpha = .69$ für die Skala „Herausforderung“ (4 Items), $\alpha = .79$ für die Skala „Misserfolgsbefürchtung“ (5 Items) und $\alpha = .66$ für die Skala „Interesse“ (5 Items). Diese Konsistenzkoeffizienten lagen innerhalb der von Rheinberg et al. (2001) berichteten Spannweite der Reliabilitäten. Die Schwierigkeitsindizes der Items verteilten sich im mittleren und hohen Bereich ($P = 53$ bis 92). Werte um $P = 50$ werden dabei als günstig eingestuft (Mummendey, 1995, S. 73). Hohe Schwierigkeitskoeffizienten zeigten sich vor allem in der Skala Herausforderung ($P = 77$ bis 92). Die Items waren demzufolge leicht im Sinne des Konstrukts zu beantworten. Die Trennschärfen der Mehrheit der Items konzentrierten sich im mittleren bis hohen Bereich ($r_{it} = .28$ bis $.73$). Vor allem die Items der Skala Misserfolgsbefürchtung erwiesen sich als geeignet, zwischen Personen mit hohen versus niedrigen Ausprägungen zu differenzieren ($r_{it} = .50$ bis $.73$).

Die Items der Kurzversion des FAM, welche von den Teilnehmern während der Aufgabenbearbeitung beantwortet wurden, wurden separaten Item- und Skalenanalysen für die drei Messzeitpunkte unterzogen. Die Ergebnisse dieser Prüfungen sind in den Tabellen A-15, A-16 und A-17 in Abschnitt A 3.2.1 im Anhang dargestellt. Die Reliabilität der vier Kurzskalen mit jeweils zwei Items konnte nicht über die interne Konsistenz bestimmt werden (vgl. McDonald, 1999; Hancock & Mueller, 2006). Um Hinweise auf die Zuverlässigkeit der Kurzskalen zu erhalten, wurde deshalb ein Retestrelabilitätskoeffizient r_{tt} für jede der vier Skalen berechnet. Der Retestrelabilitätskoeffizient wird auch als Maß für die Stabilität eines Merkmals aufgefasst (Bühner, 2006, S. 155). Zur Berechnung wurden die bivariaten Korrelationen für jede der vier Kurzskalen zwischen den drei Messzeitpunkten ($t1$ bis $t3$) berechnet und gemittelt. Die Retestrelabilität für die beiden Items der Skala Interesse lag mit $r_{tt} = .78$ am höchsten. Die beiden Skalen Herausforderung und Erfolgswahrscheinlichkeit erreichten Stabilitäten von $r_{tt} = .74$ bzw. $.69$. Am niedrigsten fiel die Retestrelabilität der Skala Misserfolgsbefürchtung aus ($r_{tt} = .45$). Die Misserfolgskomponente der Motivation lag damit erheblich unter den Stabilitäten der drei weiteren FAM-Kurzskalen. Die *part-whole* korrigierte Itemtrennschärfe wurde für die Items nicht berechnet, da diese lediglich die Korrelation zwischen den beiden Items wiedergegeben hätte.

8.1.3.3 Flow-Kurz-Skala (FKS)

Auch in Studie 2 wurde die bereits in Abschnitt 4.2.3.3 ausführlich beschriebene Flow-Kurz-Skala von Rheinberg et al. (2003) zur Erfassung des Flow-Erlebens eingesetzt. Die Trennschärfekoeffizienten wurden separat für die beiden Subskalen „Glatter Verlauf“ (6 Items) und „Absorbiertheit“ (4 Items) zu drei Messzeitpunkten ($t1$ bis $t3$) berechnet. Die Trennschärfekoeffizienten lagen zwischen $r_{it} = .33$ und $.43$ ($t1$), $r_{it} = .39$ und $.81$ ($t2$) und $r_{it} = .21$ und $.68$ ($t3$). Nur bei

einem Item zeigte sich eine besonders niedrige Trennschärfe von unter .30 (Item 1, Messzeitpunkt t3; siehe Tabellen A – 12 bis A - 14 im Anhang).

Die 10 Items der Skala erreichten eine zufrieden stellende interne Konsistenz von Cronbachs $\alpha = .76$ (Messzeitpunkt t1, $N = 33$), $\alpha = .84$ (Messzeitpunkt t2, $n = 28$) bzw. $\alpha = .87$ (Messzeitpunkt t3, $n = 21$). Der geringere Stichprobenumfang zu Messzeitpunkt t2 und t3 geht dabei auf die frühzeitige Beendigung der Aufgabe durch einige Teilnehmer zurück. Die beiden Subskalen „Glatter Verlauf“ und „Absorbiertheit“ korrelierten mittelstark zu den letzten beiden für die weiteren Analysen relevanten Messzeitpunkten mit $r = .47$ ($p < .05$, $n = 28$) bzw. $r = .58$ ($p < .01$, $n = 21$). Aus diesem Grund wurde von einer ausreichenden Überlappung der Varianz ausgegangen und ein Summenwert über die beiden Subskalen berechnet. Dies entspricht dem Vorgehen in Studie 1.

8.1.3.4 Problemlöseaufgabe, Performanzmaß und Erfassung der Strategiesystematik

Problemlöseaufgabe. Die Problemlöseaufgabe bestand auch in Studie 2 im Explorieren des Programms SPSS 12.0 anhand von vier Statistikaufgaben (vgl. Studie 1, Abschnitt 4.2.3.4). Tabelle 8 bietet eine Übersicht über die jeweiligen Anforderungen in den vier Aufgaben.

Tabelle 8

Übersicht über die im Arbeitsblatt enthaltenen Statistikaufgaben („Leistung_Arbeitsblatt“; $N = 33$)

	Aufgabe auf dem Arbeitsblatt ¹	Max.	%
1	Angabe von <i>Mittelwert</i> und <i>Standardabweichung</i> einer Variablen in der Gesamtstichprobe sowie getrennt für Frauen und Männer ($N = 33$)	6 Punkte	100
2	Angabe des <i>Korrelationskoeffizienten</i> , des Signifikanzniveaus sowie Interpretation des Ergebnisses ($n = 27$)	4 Punkte	82
3	Interpretieren eines <i>Boxplots</i> (Median Frauen und Männer, Ausreißer und Extremwerte, Verteilung) ($n = 22$)	4 Punkte	67
4	Erstellen eines gruppierten Balkendiagramms ($n = 13$)	3 Punkte	40
	1 Punkt: zu den Daten passende Grafik		
	2 Punkte: Balkendiagramm		
	3 Punkte: Gruppiertes Balkendiagramm		

Anmerkung. Max. = mit der maximal zu erreichenden Punktzahl pro Aufgabe auf dem Arbeitsblatt; % = relativer Anteil der Personen, welche die jeweilige Aufgabe in Angriff genommen haben. ¹Der genaue Wortlaut der Aufgabenstellung ist im Anhang Teil A 3.1.2 nachzulesen.

Auch in Studie 2 bearbeiteten die Teilnehmer die Aufgaben anhand eines eigens konstruierten Datensatzes, in dem mehrere Variablen enthalten waren (z. B. „Anfangsgehalt“, „Gehalt nach 10-jähriger Betriebszugehörigkeit“, „Studienfach“). Um das Interesse der Teilnehmer an den Aufgaben zu erhöhen, wurden diese so gewählt, dass sie für den Bereich Wirtschaftswissenschaften eine augenscheinliche inhaltliche Relevanz besaßen. Das bedeutet, die Aufgaben umfassten Beispiele aus dem Bereich der Wirtschaftswissenschaften. In *Aufgabe 1* sollten für unterschiedliche

Teilstichproben Mittelwerte und Standardabweichungen in einer Variablen berechnet werden. In *Aufgabe 2* sollte eine bivariate Korrelation durchgeführt und das Ergebnis interpretiert werden. *Aufgabe 3* beinhaltete das Erstellen von Boxplots getrennt für Männer und Frauen. Die Teilnehmer sollten darüber hinaus angeben, wie das Ergebnis zu interpretieren sei. *Aufgabe 4* erforderte das möglichst exakte Reproduzieren eines gruppierten Balkendiagramms, welches auf dem Arbeitsblatt abgedruckt war (siehe Arbeitsblatt in Abschnitt A 3.1.2 im Anhang zu dieser Arbeit). Die Verwendung der statistischen Symbole, Indizes und Begriffe wurde auf die Schreibnormen der Wirtschaftswissenschaften abgestimmt. Die Teilnehmer wurden aufgefordert, die Aufgaben in der vorgegebenen Reihenfolge mit Hilfe von SPSS zu lösen. Die Ergebnisse wurden von ihnen auf dem Arbeitsblatt notiert bzw. im SPSS-Ausgabefenster gespeichert (Bsp. Aufgabe 4: Gruppiertes Balkendiagramm; siehe auch Tabelle 8).

Die Statistikaufgaben waren auf den theoretischen Kenntnisstand der Teilnehmer abgestimmt. Als Referenz diente hierbei der Stand der Statistikvorlesung im Fachbereich Wirtschaftswissenschaften, welche die Teilnehmer im laufenden Semester besuchten. Die mit SPSS zu bearbeitenden Aufgaben erforderten sowohl die korrekte Produktion des Ergebnisses mit Hilfe des Computerprogramms als auch die Interpretation der im Ausgabefenster angezeigten Ergebnisse. Die Reihenfolge der Aufgaben war so gewählt, dass der Schwierigkeitsgrad möglichst kontinuierlich von Aufgabe zu Aufgabe anstieg.

Leistungsindikator. Anders als in Studie 1 wurden zwei Skalen gebildet, um die Performanz bei der Bearbeitung der Aufgaben mit SPSS zu ermitteln: „Leistung_Arbeitsblatt“ und „Leistung_SPSS“. Der erste Summenwert („Leistung_Arbeitsblatt“) bezog sich in erster Linie auf die Identifikation korrekter Analyseergebnisse (z.B. das Niederschreiben des Korrelationskoeffizienten) sowie auf die Interpretationsleistung (z.B. der inhaltlichen Bedeutung eines Korrelationskoeffizienten). Für die Skala „Leistung_Arbeitsblatt“ wurden die Punkte summiert, die ein Teilnehmer für das Notieren korrekter Kennwerte bzw. einer richtigen Interpretation dieser Kennwerte erhielt.

Um die reine Anwendungsleistung des Programms von der Interpretationsweise (annäherungsweise) separieren zu können, wurde der zweite Summenwert berechnet („Leistung_SPSS“). Diese Skala bezog sich auf das korrekte Auffinden und Bedienen von Programmfunktionen (z.B. sich ein Boxplot in SPSS anzeigen lassen). Eine Ausnahme bildet die Aufgabe 4 („Erstellen eines Gruppendiagramms“): Die in dieser Aufgabe erreichten Punkte wurden in beide Skalen einbezogen, da keine Aufgabe des Arbeitsblattes in den Summenwerten unberücksichtigt bleiben sollte. Der Maximalpunktwert lag in der Skala „Leistung_Arbeitsblatt“ bei 17 Punkten (siehe Tabelle A-18 im Anhang). In der Skala „Leistung_SPSS“ konnte eine Maximalpunktzahl von 9 erreicht werden (siehe Tabelle A-19 im Anhang).

Um einen größeren Datenverlust zu vermeiden, erhielten Teilnehmer, welche eine Aufgabe nicht in Angriff genommen hatten, Null Punkte auf die jeweilige Aufgabe (vgl. Studie 1).¹⁶ Die Skala „Leistung_Arbeitsblatt“ erreichte eine interne Konsistenz von Cronbachs $\alpha = .65$. Die Trennschärfen der Items lagen im mittleren Bereich zwischen $r_{it} = .41$ und $.53$. Die Itemschwierigkeiten variierten nach Inangriffnahmekorrektur zwischen $P = 39$ und 100 . Auffallend war dabei, dass alle Personen, welche Aufgabe 4 bearbeitet haben, diese auch korrekt lösen konnten. Die interne Konsistenz der Skala „Leistung_SPSS“ betrug Cronbachs $\alpha = .77$. Die Schwierigkeitsindizes der Items ($P = 72$ bis 100 ; nach Inangriffnahmekorrektur) lagen insgesamt höher als in der Skala „Leistung_Arbeitsblatt“. Die Trennschärfen betrugen $r_{it} = .49$ bis $.82$. Das korrekte Bedienen des Programms SPSS („Leistung_SPSS“) erreichte damit günstigere Reliabilitätswerte. Darüber hinaus erwiesen sich die Aufgabenanforderungen in der Skala „Leistung_SPSS“ als leichter zu bewältigen im Vergleich mit denjenigen in der Skala „Leistung_Arbeitsblatt“. Die Tabellen A-18 und A-19 im Anhang geben jeweils Mittelwert, Standardabweichung, Trennschärfe, Schwierigkeit sowie Schiefe und Exzess der Verteilungen in den einzelnen Items der beiden Skalen wider. Hinweise auf eine Abweichung von der Normalverteilung auf Itemebene zeigten sich nicht.

Explorationsstrategien. Das im Rahmen von Studie 1 entwickelte Kategoriensystem wurde erneut zur Quantifizierung der Häufigkeitsanteile der Explorationsstrategien verwendet. Das Explorationsverhalten der Teilnehmer am Computer wurde auch hier mittels des Screenshotprogramms *ScreenVirtuoso* aufgezeichnet. Das Explorationsverhalten wurde in dieser Studie jedoch ausschließlich von einem Rater kodiert. Wie bereits in Studie 1 konnte auch in Studie 2 die Anzahl der Beobachtungseinheiten leicht variieren: Für computerbedingte Verzögerungen beim Ausführen der Befehle wurden den Teilnehmern kompensatorisch zusätzliche Zeiteinheiten zur Aufgabenbearbeitung gewährt. Die absoluten Häufigkeiten der einzelnen Strategien (Systematische Exploration, Rigide Exploration, Versuch-und-Irrtum, Informationssuche) sowie der Kategorie „Unterbrechung“ wurden erneut an der individuellen Anzahl der Beobachtungseinheiten (in der Regel 60 Einheiten) relativiert. Für den gesamten Beobachtungszeitraum (ca. 30 min) wurde für jede Strategie ein Summenwert gebildet, der die relative Häufigkeit der Strategie im Vergleich zur Häufigkeit der anderen Strategien abbildete. Darüber hinaus wurde die relative Häufigkeit für jede Strategie separat für jeden der drei Beobachtungszeiträume (erste 10 Minuten, zweite 10 Minuten, dritte 10 Minuten) bestimmt.

¹⁶ Ein Nachteil dieses Vorgehens liegt darin, dass Personen „künstlich“ Werte zugewiesen werden und sich dies auch auf die Höhe der Reliabilität auswirken kann. Dies sollte bei der Interpretation der Kennwerte berücksichtigt werden (vgl. Studie 1).

8.1.4 Wahl und Diskussion der statistischen Analyseverfahren

Prüfung auf Mittelwertunterschiede. Die in den Hypothesen erwarteten geschlechtsspezifischen Unterschiede wurden mittels einer Varianzanalyse (ANOVA) mit Messwiederholung überprüft (Bortz, 2005, S. 336 ff.). Ein Nachteil der Varianzanalyse mit Messwiederholung besteht darin, dass konstante Residualwerte für die Varianzen der Abhängigen Variablen angenommen werden und lediglich Gruppenunterschiede in den Mittelwerten ausgewertet werden (vgl. Cohen, Cohen, West & Aiken, 2003, s. 578 ff.). Aussagen über individuelle Veränderungen im Zeitverlauf sind auf dieser Grundlage nicht möglich (vgl. Preacher, Wicham, MacCallum & Briggs, 2008).

Überprüfung von Mediationseffekten. In Hypothese 9 wird ein Modell mit multiplen Mediatoren postuliert. Eine multiple Mediation liegt dann vor, wenn der kausale Einfluss einer Variablen auf eine andere durch mehr als eine weitere Mediatorvariable vermittelt wird (Hoyle & Kenny, 1999, S. 195; Iacobucci, Saldanha & Deng, 2007). Um einen Mediatoreffekt nachweisen zu können, müssen nach Baron und Kenny (1986; vgl. auch Rost, 2007, S. 143) folgende Bedingungen gelten: Erstens, Prädiktor und Kriterium müssen statistisch signifikant miteinander assoziiert sein. Zweitens, Prädiktor und Mediator müssen statistisch signifikant korreliert sein. Drittens, die Mediatorvariable muss einen signifikanten Einfluss auf die Kriteriumsvariable ausüben. Viertens, der Effekt der Prädiktorvariable auf das Kriterium muss bei Hinzunahme des Mediators in das Modell signifikant reduziert werden.

Die statistische Überprüfung von Mediationsmodellen erfolgt in der Regel auf der Basis von Regressionsanalysen (vgl. Baron & Kenny, 1986; Hoyle & Kenny, 1999; McKinnon, Lockwood, Hoffman, West & Sheets, 2002). Mediationseffekte können jedoch auch auf Basis von Strukturgleichungsmodellen überprüft werden. Im Strukturgleichungsmodell ist ein erster statistischer Hinweis auf einen Mediatoreffekt dann gegeben, wenn sowohl der Einfluss des Prädiktors auf den Mediator als auch der Einfluss des Mediators auf das Kriterium statistisch signifikant werden. Wie groß der geschätzte Mediationseffekt, d.h. der indirekte Effekt des Prädiktors auf das Kriterium ist, kann über die Ergebnisse der Effektzerlegung nachvollzogen werden (Kline, 2005). Darüber hinaus ist entscheidend, ob sich die Größe des direkten Pfadkoeffizienten infolge der Einführung des Mediators statistisch signifikant verringert hat (Iacobucci et al., 2007, S. 152).

Die Befunde mehrerer Simulationsstudien weisen auf die Überlegenheit von Strukturgleichungsmodellen gegenüber regressionsanalytischen Verfahren bei der Analyse von Mediatoreffekten hin (vgl. Iacobucci et al., 2007; Kenny, Kashy & Bolger, 1998). Diese Überlegenheit zeigte sich beispielsweise in zwar geringen, jedoch systematischen Differenzen der Standardfehler der Parameterschätzung insbesondere bei kleineren Stichproben von $N = 30$ (Iacobucci et al., 2007, S. 144). Vor diesem Hintergrund wurde in Studie 2 Strukturgleichungsanalysen der Vorzug gegeben.

Statistische Voraussetzungsprüfungen. Wie bereits in Abschnitt 7.2.1 zu Studie 1 beschrieben, wurde auch in Studie 2 routinemäßig überprüft, ob die Verteilungen der zentralen Untersuchungsvariablen die mathematisch-statistischen Voraussetzungen für die eingesetzten statistischen Analyseverfahren erfüllen (z.B. Normalverteilung). Zur Vorbereitung der statistischen Mittelwertvergleiche wurde auch stets die Gleichheit der Varianzen in den interessierenden Untersuchungsvariablen getestet (vgl. Bortz, 2005, S. 284f.). Spezifische Voraussetzungen gelten auch für eine ANOVA mit Messwiederholung: Nach Bortz (2005, S. 354) ist es erforderlich, dass „die Varianzen unter den einzelnen Faktorstufen und die Korrelationen zwischen den Faktorstufen homogen sein“ müssen.

Strukturgleichungsanalysen auf der Grundlage von Maximum-Likelihood-Algorithmen (vgl. Hypothese 9) machen in der Regel eine multivariate Normalverteilung der Daten erforderlich (Kline, 2005; Schermelleh-Engel et al., 2003). Bei Verletzung dieser Voraussetzung können erhöhte χ^2 -Werte die Folge sein (Bühner, 2006, S. 285). Der Mardia-Test auf multivariate Schiefe und Kurtosis wird als Maß zur Überprüfung der multivariaten Normalverteilung herangezogen (vgl. McDonald & Ho, 2002). In AMOS 17.0 ist der Mardia-Test auf multivariate Kurtosis standardmäßig implementiert (vgl. Bühner, 2006, S. 251).

Software für die Datenanalyse. Den Pfadanalysen mit *Mplus* (Muthén & Muthén, 1998–2007) wurde eine Überprüfung der Stichprobenverteilung mittels AMOS 17.0 (Arbuckle, 2008; siehe auch Byrne, 2001) vorgeschaltet. Zur Auswertung der Daten wurden die Statistikprogramme SPSS 17.0 (SPSS Inc, 1993-2007) sowie *Mplus* 5.0 verwendet. *Mplus* wurde dabei ausschließlich zur Berechnung des Pfadmodells eingesetzt. Das Programm ist syntaxbasiert und bietet u.a. den Vorteil, dass auch bei fehlenden Werten Standardfehler für die Schätzung der standardisierten Pfadkoeffizienten sowie für die indirekten Effekte ausgegeben werden. Standardisierte Lösungen sind in der Regel einfacher zu vergleichen (Kline, 2005, S. 127). Die Schätzung der indirekten Effekte ist z. B. bei der Prüfung von Mediatoreffekten interessant (Reinecke, 2005, S. 50f.).

8.1.5 Umgang mit fehlenden Werten

Wie bereits in Studie 1 wurden einzelne fehlende Werte mit Hilfe der Prozedur von Sijsma und van der Ark (2003) ersetzt, um einen vollständigen Datensatz zu erhalten. Die Prozedur wurde in Abschnitt 7.1.4 ausführlich beschrieben.

8.2 Ergebnisse

8.2.1 Vorbereitende Analysen zur empirischen Untersuchung

Fehlende Werte und Missingskorrektur. Einzelne fehlende Angaben traten im Fragebogen zur aktuellen Motivation (1 von 594 Werten) sowie in der Flow-Kurz-Skala (3 von 990 Werten) auf. Der Anteil fehlender Werte betrug dabei in beiden Skalen weniger als 0.01% und ist damit als marginal zu bezeichnen. Ein systematischer Datenausfall ergab sich zu Messzeitpunkt t2 bei fünf Teilnehmern und bei Messzeitpunkt t3 bei insgesamt 10 Teilnehmern: Diese Teilnehmer hatten die Bearbeitung des Aufgabenblattes bereits vorzeitig, d. h. vor Ende der 30-minütigen Bearbeitungszeit, abgeschlossen. In Übereinstimmung mit dem in Studie 1 gewählten Vorgehen wurden diese fehlenden Werte nicht ersetzt (siehe Abschnitt 7.2.1). Die Datengrundlage für die statistischen Auswertungen zu den beiden Messzeitpunkten reduzierte sich entsprechend um die Angaben der fünf (t2) bzw. zehn (t3) Personen. Um das Flow-Erleben bei allen Teilnehmern mit in die Analysen einbeziehen zu können, wurde auch hier die Variable „flow_end“ geschaffen (vgl. Studie 1; Abschnitt 7.2.1), welche die Werte aller Teilnehmer zum jeweils letzten Messzeitpunkt vor Ende der individuellen Aufgabebearbeitung enthielt.

Statistische Voraussetzungsprüfungen. In Tabelle A – 18 im Anhang (Abschnitt A 3.2.2) sind die Ergebnisse des Kolmogorov-Smirnov-Tests auf Abweichung von der Normalverteilung inklusive der Mittelwerte, Standardabweichung, Schiefe und Exzess der Skalenmittelwerte der zentralen Untersuchungsvariablen aufgeführt. Der Test ergab statistisch signifikante Abweichungen für die Selbstbeurteilung der eigenen Statistikkenntnisse ($Z = 1.64$, Asymptotische Signifikanz = .01). Schiefe und Exzess lieferten jedoch keine Hinweise auf Abweichungen von der Normalverteilung. Alle Werte blieben unter den von West et al. (1995) genannten kritischen Grenzen von 7 für den Exzess und 2 für die Schiefe. Die Werte in der Variablen „Häufigkeit der Computernutzung“ wies eine rechtsschiefe Verteilung auf ($Z = 1.63$, Asymptotische Signifikanz = .01). Das bedeutet, dass die Mehrheit der Teilnehmer (wie erwartet) eine geringe Erfahrung in der Anwendung von Tabellenkalkulations- und Statistikprogrammen bzw. Datenbanken hatte. Für alle anderen Variablen galten die Voraussetzungen auf (univariate) Normalverteilung als erfüllt.

Statistische Drittvariablenkontrolle. 97% der Teilnehmer ($n = 32$) gaben an, regelmäßig die Statistik-I-Grundlagenveranstaltung sowie das dazugehörige Tutorium ihres Fachbereichs zu besuchen. Gut die Hälfte der Teilnehmer hörte (zusätzlich) die Statistik-II-Veranstaltung (55%; $n = 18$). Weniger als ein Fünftel der Teilnehmer (18% bzw. $n = 6$) berichtete, sich bereits im Schulunterricht mit Statistik auseinandergesetzt zu haben. Ungefähr 79% der Befragten ($n = 26$) gaben an, ihre theoretischen Statistikkenntnisse auf Aufgaben zur Deskriptiven Statistik anwenden zu kön-

nen. 88% ($n = 29$) sahen sich in der Lage, diese gar auf die grafische Darstellung der statistischen Auswertungen übertragen zu können. Die Teilnehmer schätzten ihre eigenen Statistikkenntnisse auf der 5-stufigen Skala im Mittel als durchschnittlich ein ($M = 2.88$, $SD = 0.82$). Die Abfrage der Nutzungshäufigkeiten für die drei Programmgruppen Statistikprogramme, Tabellenkalkulationsprogramme und Datenbanken ergab, dass die große Mehrheit der Teilnehmer (64% bzw. 97%) nie bzw. weniger als 10 mal in den letzten sechs Monaten mit Statistikprogrammen und Datenbanken gearbeitet hatte. Demgegenüber lag der Durchschnitt bei der Nutzung von Tabellenkalkulationsprogrammen zwischen den Kategorien (1) „bis 10 mal“ und (3) „bis 20 mal“. Ungefähr 94% hatten nach eigenen Angaben bereits Erfahrungen mit Tabellenkalkulationsprogrammen gesammelt.

Es wurde angenommen, dass das Vorwissen der Teilnehmer in Statistik sowie die Computererfahrung einen Einfluss auf den Explorationsprozess des Computerprogramms SPSS haben könne. Aus diesem Grund wurde überprüft, ob Zusammenhänge zwischen der Vorerfahrung bzw. dem Vorwissen eines Teilnehmers und seiner Leistung beim Erlernen von SPSS bestehen. Im Falle positiver Beziehungen zwischen den Variablen Vorwissen und Leistung, müssten erstere als Kovariate in die Analysen aufgenommen werden, um die Differenzen zwischen den Teilnehmern vor Beginn der Untersuchung berücksichtigen zu können.

Tabelle A - 19 im Anhang gibt eine Übersicht über die Produkt-Moment-Korrelationen sämtlicher Variablen der Untersuchung in der Gesamtstichprobe. Vorwissen bzw. Vorerfahrungen standen in der Gesamtstichprobe in keiner statistisch signifikanten Beziehung zur Leistung – weder die Häufigkeit der Computernutzung ($r = .20$, $p = .26$) noch die Statistikkenntnisse (Anzahl besuchter Statistikveranstaltungen: $r = .29$, $p = .11$; Selbstbeurteilung: $r = .15$, $p = .39$; Statistikvorwissenstest [6 Items]: $r = .24$, $p = .19$) kovariierten mit der Leistung bei der Aufgabenlösung auf dem Arbeitsblatt. Dabei ist zu beachten, dass die Verteilung im Item „Selbstbeurteilung der Statistikkenntnisse“ signifikant von der Normalverteilung abwich. Produkt-Moment-Korrelationen gelten jedoch als robust gegenüber dieser Voraussetzungsverletzung (Havlicek & Peterson, 1977). Da die Korrelationsanalysen keinen Zusammenhang zwischen der Computererfahrung und der Leistung sowie dem Statistikvorwissen und der Leistung ergeben hatten, wurden diese nicht als Kovariaten in Analysen zu Leistungsunterschieden aufgenommen.

8.2.2 Hypothese 5: Geschlechtsunterschiede in Computernutzung und Statistikwissen

Gemäß Hypothese 5 haben Frauen eine geringere Computererfahrung als Männer, d. h. Frauen nutzen weniger häufig Computerprogramme wie Statistiksoftware, Tabellenkalkulationsprogramme oder Datenbanken.

Deskriptive Statistiken. Erfasst wurde die selbst berichtete Häufigkeit, mit der die Teilnehmer in den letzten sechs Monaten mit diesen Computeranwendungen gearbeitet hatten. Die männlichen Teilnehmer gaben an, durchschnittlich 10- bis 20-mal ($M = 1.67, SD = 0.91$) im letzten halben Jahr vor der Untersuchung mit Tabellenkalkulationsprogrammen gearbeitet zu haben. Die weiblichen Teilnehmer hatten im gleichen Zeitraum im Mittel „bis 10-mal“ mit diesen Programmen gearbeitet ($M = 1.07, SD = 0.59$). Männer ($M = 0.11, SD = 0.32$) und Frauen ($M = 0, SD = 0$) hatten in den sechs Monaten vor der Untersuchung nie bis selten Statistikprogramme genutzt (exklusive SPSS). Auch bei der Häufigkeit der Datenbank-Nutzung lagen Männer ($M = 0.50, SD = 0.79$) und Frauen ($M = 0.27, SD = 0.46$) ihren eigenen Angaben zufolge in dem Bereich seltener Nutzung. Der empirische Mittelwert lag bei beiden Geschlechtern folglich zwischen den Kategorien „nie“ bis „10mal“. Die weiblichen und männlichen Teilnehmer hatten demzufolge eine geringe bis gar keine Vorerfahrung in der Nutzung von Statistikprogrammen und Datenbanken.

Varianzanalytische Hypothesenprüfung. Zur Überprüfung der Hypothese wurden die Nutzungshäufigkeiten der einzelnen Computeranwendungen in einer Variablen („Häufigkeit der Computernutzung“) zusammengefasst. Die erwarteten Geschlechtsunterschiede wurden über eine Multivariate Varianzanalyse (MANOVA) mit dem Faktor Geschlecht und den vier Abhängigen Variablen Häufigkeit der Computernutzung, besuchte Statistikveranstaltungen, Selbstbeurteilung der Statistikkenntnisse und Statistikvorwissenstest getestet (vgl. Bortz, 2005; Stelzl, 2005). Der Globaltest der MANOVA ergab keine statistisch signifikanten Geschlechtsunterschiede in den Abhängigen Variablen (Hotelling-Spur: $F [4, 28] = 1.56, p = .21$). Das bedeutet, dass das Geschlecht keinen Einfluss auf das Statistikvorwissen und die Computererfahrung hatte, wenn beide Aspekte des Vorwissens simultan betrachtet wurden.

Im Anschluss an den Globaltest wurde überprüft, ob sich Männer und Frauen in einzelnen Variablen statistisch signifikant unterschieden. Um eine Alpha-Fehler-Inflation zu vermeiden, wurde zuvor eine Bonferroni-Korrektur vorgenommen (Stelzl, 2005, S. 123). Das Signifikanzniveau der vier Einzeltests wurde a priori auf $p = .0125$ festgelegt. Tabelle 9 fasst die Angaben der Teilnehmer bezüglich der Häufigkeit der individuellen Computernutzung sowie zu ihren Vorkenntnissen im Bereich theoretische Statistikkenntnisse zusammen. Darüber hinaus sind Ergebnisse des Hypothesentests aufgeführt. Zur Beurteilung der praktischen Bedeutsamkeit möglicher Differenzen wurde zusätzlich die Effektstärke über Cohens d berechnet (Cohen, 1992).

Tabelle 9

Ergebnisse der Multivariaten Varianzanalyse (MANOVA) zur Häufigkeit spezifischer Computernutzung und Statistikkenntnisse getrennt nach Frauen und Männern

Variable	Männer (n = 18)		Frauen (n = 15)		MANOVA ¹ (df ₁ = 1, df ₂ = 31)		Effektstärke
	M	SD	M	SD	F	P	d
Computernutzung ² (Häufigkeit)	0.76	0.5	0.44	0.3	4.62	.04	0.77
Besuchte Statistikveranstaltungen (Anzahl)	3.11	1.3	3.40	1.1	0.54	.47	-0.24
Statistikkenntnisse ³ (Selbstbeurteilung)	3.00	0.8	2.73	0.8	0.86	.36	0.33
Statistikvorwissenstest ⁴	2.64	0.7	2.32	0.5	2.15	.15	0.52

Anmerkungen. *d* = Cohens *d* (Effektstärke). ¹Die Varianzen in der Variablen „Computernutzung“ erwiesen sich als ungleich (Levene-Test: $F [1, 31] = 5.83, p = .02$), so dass die Anzahl der Freiheitsgrade auf $df = 29$ korrigiert wurde. ²Durchschnittliche Häufigkeit der Nutzung von Statistik- und Tabellenkalkulationsprogrammen sowie von Datenbanken in den letzten sechs Monaten vor der Untersuchung; Skalierung: 0 = „nie“, 1 = „bis 10-mal“, 2 = „bis 20-mal“, 3 = „häufiger als 20-mal“. ³ Antwortmöglichkeiten von 1 = „mangelhaft“ bis 5 = „sehr gut“. ⁴ Mittelwert über alle Items.

Zusammenfassung. Die von Männern und Frauen berichteten Häufigkeiten ihrer Computeranwendung entsprechen der erwarteten Tendenz: Männer gaben an, in den letzten sechs Monaten häufiger mit den abgefragten Computerprogrammen gearbeitet zu haben als Frauen. Die Geschlechtsunterschiede in der Häufigkeit der Computernutzung wurden jedoch nach Bonferroni-Korrektur nicht statistisch signifikant. Neben der Computererfahrung wurde in der Studie das Statistikvorwissen erfasst. Der *F*-Test resultierte nicht in statistisch signifikanten Geschlechtsunterschieden in der Leistung im Statistiktest ($F [1, 31] = 2.15, p = .15$; siehe Tabelle 9). Männer und Frauen unterschieden sich demnach nicht in ihrem Statistikvorwissen. Im Gegensatz zu Studie 1 korrelierte das Statistikvorwissen zudem nicht mit der Leistung beim Erlernen von SPSS (vgl. Abschnitt 8.2.1). Das Ergebnis eines Signifikanztests wird jedoch durch die Stichprobengröße beeinflusst (Sedlmeier & Renkewitz, 2007; Sedlmeier & Gigerenzer, 1989, 1997). Aus diesem Grund wurden neben den exakten *p*-Werten zusätzlich die Effektstärken berechnet (vgl. Wilkinson & Task Force on Statistical Inference, 1999). Tabelle 9 gibt Hinweise darauf, dass zumindest bei der Häufigkeit der Computernutzung starke Geschlechtseffekte beobachtbar sind ($d = 0.77$).

8.2.3 Hypothese 6: Geschlechtsunterschiede in der Eingangsmotivation

Gemäß Hypothese 6 sollten sich Geschlechtsunterschiede in der Eingangsmotivation manifestieren. Es wurde erwartet, dass Frauen eine ungünstigere Eingangsmotivation im Vergleich zu Männern aufweisen: Sie sollten weniger herausgefordert und weniger interessiert sein, sollten die Erfolgswahrscheinlichkeit beim Lösen der Aufgaben niedriger bewerten und gleichzeitig eine höhere Misserfolgsbefürchtung haben.

Deskriptive Statistiken. In Tabelle 10 sind die Mittelwerte und Standardabweichungen für Männer und Frauen in allen Skalen der Eingangsmotivation dargestellt. Die Teilnehmer fühlten sich in Relation zum theoretischen Skalenmittelwert von der Aufgabe leicht überdurchschnittlich herausgefordert ($M = 5.05$, $SD = 0.97$). Die Eingangsmotivation lag in der Gesamtstichprobe sowohl in der Skala Interesse als auch in der Skala Erfolgswahrscheinlichkeit im mittleren Bereich ($M = 4.36$, $SD = 0.94$ bzw. $M = 4.35$, $SD = 0.95$). Die Misserfolgsbefürchtung war dagegen unterdurchschnittlich ausgeprägt ($M = 3.44$, $SD = 1.14$).

Tabelle 10

Ergebnisse der MANOVA für den Faktor Geschlecht und die vier Abhängigen Variablen (Erfolgswahrscheinlichkeit, Herausforderung, Interesse, Misserfolgsbefürchtung)

Variable	Männer ($n = 18$)		Frauen ($n = 15$)		MANOVA ($df_1 = 1$, $df_2 = 31$)		Effekt- stärke
	M	SD	M	SD	F	p	d
Eingangsmotivation (18 Items)							
Interesse	4.46	0.95	4.25	0.96	0.37	.55	0.22
Erfolgswahrscheinlichkeit	4.51	0.88	3.94	0.98	3.17	.09	0.61
Herausforderung	4.93	0.96	5.18	0.99	0.55	.46	-0.26
Misserfolgsbefürchtung	3.26	1.19	3.65	1.07	1.00	.33	-0.34

Anmerkungen. M = Skalenmittelwerte; SD = Standardabweichungen; d = Cohens d (Effektstärke).

Varianzanalytische Hypothesenprüfung. Die in der Hypothese postulierten Unterschiede in der Eingangsmotivation wurden in einem ersten Schritt mit einer MANOVA geprüft. In diese wurden alle Skalen des FAM als Abhängige Variablen einbezogen. Im Anschluss an den Globaltest der MANOVA wurden Unterschiede in einzelnen Skalen überprüft. Zuvor wurden die Varianzen in den Gruppen auf Homogenität getestet. Das Testergebnis ergab keine signifikanten Abweichungen von der Varianzhomogenität. Das Signifikanzniveau für die einzelnen Vergleiche wurde a priori auf $p = .0125$ festgelegt (vgl. Abschnitt 8.2.2).

Entgegen der Hypothese zeigen die Ergebnisse, dass sich die Höhe der Eingangsmotivation bei Frauen und Männern vor Beginn der Arbeit mit SPSS nicht statistisch signifikant unterschied (Hotelling-Spur: $F [4, 28] = 0.48$, $p = .11$). Der gleiche Befund zeigte sich auch bei separater Testung der einzelnen Skalenwerte: Wie Tabelle 10 zeigt, spiegeln die Gruppenmittelwerte zwar die erwarteten Tendenzen wider. Jedoch erreichte keiner der vier Mittelwertvergleiche statistische Signifikanz.

Effektstärken. Nach Cohen (1992) zeigen sich schwache Effekte zugunsten der Männer in den Skalen Interesse, Erfolgswahrscheinlichkeit und Misserfolgsbefürchtung. Männer berichteten

eine günstigere Motivation mit einem höheren Interesse. Frauen erreichten höhere Mittelwerte in den Items, die Misserfolgsbefürchtung abfragen. Mittelstarke Unterschiede zwischen den Geschlechtern zeigten sich hinsichtlich der Beurteilung der Erfolgswahrscheinlichkeit vor der Aufgabebearbeitung in SPSS. Männer schätzten dabei ihr erfolgreiches Abschneiden bei der Aufgabebearbeitung als wahrscheinlicher ein als dies bei Frauen der Fall war. Die Frauen fühlten sich demgegenüber etwas stärker herausgefordert als ihre männlichen Kommilitonen (s. Tabelle 10).

Zusammenfassung. Deskriptiv betrachtet unterschieden sich Männer und Frauen schwach bis mittelstark in den vier Skalen der Eingangsmotivation: Männer zeigten höhere Werte in den Skalen Interesse und Erfolgswahrscheinlichkeit. Frauen gaben dagegen leicht höhere Ausprägungen bei Misserfolgsbefürchtung und Herausforderung an. Bis auf die Werte in der Skala Herausforderung folgten die Tendenzen in der Stichprobe den in der Hypothese formulierten Erwartungen. Die Unterschiede wurden jedoch nicht statistisch signifikant. Die zusätzlich berichteten Effektstärken zeigen jedoch, dass die in der Stichprobe gemessenen Differenzen schwach bis mittelstark ausgeprägt sind. Die Ergebnisse basieren dabei auf einem relativ geringen Stichprobenumfang.

8.2.4 Hypothese 7a: Geschlechtsunterschiede in den Explorationsstrategien

Hypothese 7a zufolge wurde erwartet, dass Männer im Vergleich zu Frauen häufiger die systematische Explorationsstrategie anwenden. In Tabelle 11 sind die Häufigkeiten der Nutzung einzelner Explorationsstrategien für beide Geschlechter getrennt dargestellt.

Tabelle 11

Beobachtete relative Häufigkeit einzelner Explorationsstrategien bzw. der Kategorie „Unterbrechung“ in der Gesamtstichprobe sowie getrennt für Frauen und Männer

Variable	Gesamtstichprobe (N = 33)		Männer (n = 18)		Frauen (n = 15)		Effektstärke (Geschlecht) <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Systematische Exploration	41%	12	45%	12	35%	11	0.87
Versuch-und-Irrtum	5%	4	6%	5	3%	3	0.73
Rigide Exploration	20%	17	15%	15	28%	17	-0.81
Informationssuche	5%	7	3%	6	8%	8	-0.71
Unterbrechung	29%	12	32%	13	26%	11	0.50

Anmerkungen. *M* = Mittlere relative Häufigkeit der Strategienutzung in Prozent der individuellen Anzahl der Beobachtungseinheiten; *SD* = Standardabweichungen; *d* = Cohens *d*.

Nutzung der systematischen Strategie. Hypothesenkonform unterschieden sich die Geschlechter statistisch signifikant in der Häufigkeit der Nutzung der systematischen Strategie: Männer zeigten in 45% ($SD = 12$) der Beobachtungsabschnitte systematische Verhaltensweisen wohingegen dies bei Frauen nur zu 35% ($SD = 11$) der Fall war ($t [31] = -2.54, p = .016; d = 0.87$). Die Hypothese wird durch dieses Ergebnis empirisch unterstützt. Unabhängig von Geschlechtsunterschieden nahmen die systematischen Explorationsstrategien mit 41% ($SD = 12$) den größten Anteil an allen Strategien in der Gesamtstichprobe ein (s. Tabelle 11).

Nutzung unsystematischer Strategien. A priori lagen für die Nutzung unsystematischer Explorationsstrategien keine expliziten Annahmen zu Geschlechtsunterschieden vor. Differenzen in der Nutzungshäufigkeit der Explorationsstrategien Rigide Exploration, Versuch-und-Irrtum sowie Informationssuche konnten nicht mit Befunden aus Voruntersuchungen begründet werden. Bei der Überprüfung der Nutzungshäufigkeit dieser beiden Explorationsstrategien wurde daraufhin eine α -Fehler-Korrektur vorgenommen (Bortz, 1993, S. 248). Hierzu wurde das Signifikanzniveau bei $m = 3$ Gruppenvergleichen und einem Ausgangs-Signifikanzniveau von $p = .05$ in einem Einzeltest auf den Höchstwert von $p = .017$ festgesetzt (vgl. Bortz, 1993, S. 249; Stelzl, 2005).

Im Globaltest der Multivariaten Varianzanalyse zeigte sich ein statistisch signifikanter Haupteffekt des Geschlechts: Hypothesenkonform nutzten Frauen häufiger unsystematische Explorationsstrategien als Männer (Hotelling-Spur: $F [3, 29] = 3.86, p = .019$). Um die Unterschiede verorten zu können, wurden anschließend Einzeltests für die drei Explorationsstrategien durchgeführt. Der Geschlechtsunterschied in der Wahl unsystematischer Strategien verfehlte für die Variable Rigide Exploration die adjustierte Signifikanzgrenze ($F [1, 31] = 5.45, p = .026; d = -.81$). Das Ergebnis spiegelt jedoch eine Tendenz in die erwartete Richtung wider: Frauen zeigten eine stärkere Tendenz rigide zu explorieren als dies bei Männern der Fall war. Die Differenz im Ausmaß der Informationssuche erreichte keine statistische Signifikanz ($F [1, 31] = 3.74, p = .06, d = -.71$), wies jedoch gleichsam in die erwartete Richtung. Der Geschlechtseffekt in der Variablen Versuch-und-Irrtum wurde ebenfalls nicht statistisch signifikant ($F [1, 31] = 2.01, p = .167, d = .73$).

Unsystematische Explorationsstrategien wie Versuch-und-Irrtum und Informationssuche traten insgesamt betrachtet eher selten auf (s. Tabelle 11). Der Anteil der Zeit, in dem keine Strategienutzung zu beobachten war (Kategorie „Unterbrechung“), variierte bei Männern und Frauen zwischen $M = 32\%$ ($SD = 13$) und $M = 26\%$ ($SD = 11$). Diese Zeiteinheiten verbrachten die Teilnehmer mit der Niederschrift in SPSS produzierter Ergebnisse, dem Ausfüllen von Fragebögen oder dem Lesen neuer Aufgaben.

Strategiesystematik im Zeitverlauf. Interessant ist auch die Entwicklung der Strategiesystematik im zeitlichen Verlauf. Diese lässt sich für die Gesamtstichprobe ($N = 33$) folgenderma-

ßen beschreiben: In den *ersten 10 Minuten* nutzten die Teilnehmer in der überwiegenden Mehrheit der Beobachtungseinheiten systematische Explorationsstrategien ($M = 56\%$, $SD = 13$). Am zweithäufigsten, jedoch deutlich weniger häufig als bei der systematischen Strategie wurde rigide exploriert ($M = 19\%$, $SD = 17$). In den *zweiten 10 Minuten* zeigten die Teilnehmer in durchschnittlich 30% der Beobachtungseinheiten ($SD = 19$) ein systematisches Explorationsverhalten. Im Vergleich zum Messzeitpunkt t1 lag der Durchschnitt in Messzeitpunkt t2 für diese Strategie jedoch deutlich niedriger. Der Anteil der rigiden Exploration lag zu t2 bei 25% ($SD = 24$). Die gleiche Tendenz zeigte sich in den *letzten 10 Minuten* der Bearbeitungszeit: Die Teilnehmer explorierten hier am häufigsten auf systematische ($M = 29\%$, $SD = 15$) bzw. rigide Weise ($M = 22\%$, $SD = 25$). Bei den Häufigkeiten im Zeitverlauf ist der reduzierte Stichprobenumfang zu t2 und t3 zu beachten. Aufgrund des Datenausfalls vor allem bei den Männern wurde auf einen inferenzstatistischen Gruppenvergleich der Strategiesystematik über die Zeit verzichtet.

Zusammenfassung. Die untersuchte Hypothese konnte durch die empirischen Ergebnisse gestützt werden: Männer zeigten bei der Exploration eines neuen Computerprogramms demnach häufiger systematische Explorationsstrategien als Frauen. Männer und Frauen unterschieden sich deutlich, wenn gleichzeitig die Häufigkeit aller in der Studie erfassten unsystematischen Explorationsstrategien berücksichtigt wurde. Männer nutzten im Durchschnitt seltener unsystematische Explorationsstrategien als Frauen. Geschlechtsdifferenzen in spezifischen Explorationsstrategien zeigten sich in den Einzeltests der MANOVA nicht. Die systematische Explorationsstrategie nahm insgesamt und im Zeitverlauf stets den größten Anteil ein. Versuch-und-Irrtum sowie Informationssuche wurden erneut in geringerem Umfang beobachtet (vgl. Studie 1).

8.2.5 Hypothese 7b: Geschlechtsunterschiede im Flow-Erleben

In *Hypothese 7b* wurde erwartet, dass sich Männer und Frauen während des Aufgabenbearbeitungsprozesses in der Höhe des Flow-Erlebens unterscheiden. Männer sollten danach höhere Flow-Werte berichten als Frauen.

Das Flow-Erleben wurde in der Untersuchung zu drei Messzeitpunkten (t1 bis t3) erfasst. Wie oben beschrieben, beendeten einige Teilnehmer die Bearbeitung der Aufgaben vorzeitig, so dass zu Messzeitpunkt t2 und t3 ein systematischer Datenausfall festzustellen war. Um diesem Ausfall zu begegnen, wurde wie bereits in Studie 1 eine neue Variable gebildet: In der Variablen „Flow_end“ sind die Werte der Teilnehmer zum letzten Messzeitpunkt vor dem individuellen Bearbeitungsende zusammengefasst.

Tabelle 12 gibt eine Übersicht über die Deskriptiven Statistiken der Verteilungen für die drei Messzeitpunkte sowie die Variable Flow_end getrennt für Frauen und Männer. Zu allen drei

Messzeitpunkten lagen die Mittelwerte in der Gruppe der männlichen sowie der weiblichen Teilnehmer im mittleren Bereich der 7-stufigen Antwortskala: Die Teilnehmer fühlten sich weder unterfordert noch überfordert. Zum letzten Messzeitpunkt vor Bearbeitungsende gaben Männer im Durchschnitt leicht höhere Flow-Werte an als Frauen ($M_{\text{Männer}} = 4.53$, $SD = 0.9$ bzw. $M_{\text{Frauen}} = 4.07$, $SD = 1.2$). Dieser Geschlechtsunterschied wurde jedoch im t -Test nicht statistisch signifikant ($t[31] = -1.28$, $p = .21$). Das bedeutet, dass die Hypothese nicht empirisch unterstützt werden konnte.

Tabelle 12

Mittelwerte, Standardabweichungen separat für Frauen und Männer sowie Effektstärken für die Variable „Flow-Erleben zu verschiedenen Messzeitpunkten (t1 bis t3)

Variable	Männer		Frauen		Effektstärke <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Flow-Erleben t1 (Männer: $n = 18$; Frauen: $n = 15$)	4.07	0.8	3.67	0.9	0.48
Flow-Erleben t2 (Männer: $n = 13$; Frauen: $n = 15$)	4.43	1.0	3.96	1.0	0.47
Flow-Erleben t3 (Männer: $n = 7$; Frauen: $n = 14$)	3.97	1.0	3.99	1.2	-0.02
Flow_end ¹ (Männer: $n = 18$; Frauen: $n = 15$)	4.53	0.9	4.07	1.2	0.44

Anmerkungen. *M* = Skalenmittelwerte; *SD* = Standardabweichungen; *d* = Cohens *d* (Effektstärke). ¹„Flow_end“: Höhe des Flow-Erlebens zum letzten Messzeitpunkt vor dem individuellen Bearbeitungsende.

Wie Tabelle 12 zeigt, variierte die Größe des Effekts nur gering über die drei Messzeitpunkte. Lediglich zu Messzeitpunkt t3 zeigte sich kein Geschlechtseffekt. Dies ist darauf zurückzuführen, dass die leistungsstarken Personen bereits vor der letzten Flow-Messung die Aufgabenbearbeitung beendet hatten. Zu diesen leistungsstarken Personen zählten 92% Männer und 8% der Frauen.

Zusammenfassung. Insgesamt deutet die Effektstärkemessung auf mittlere Geschlechtseffekte hin (Flow_end: Cohens $d = .44$). Das Ergebnis des Signifikanztests fiel jedoch konservativ aus und die Nullhypothese konnte nicht verworfen werden. Auch in diesem Fall kann der geringe Stichprobenumfang dazu geführt haben, dass die Nullhypothese nicht zugunsten der Alternativhypothese abgelehnt werden konnte.

8.2.6 Hypothese 8: Geschlechtsunterschiede in der Leistung

In *Hypothese 8* wurde postuliert, dass Männer bei der Aufgabenbearbeitung in SPSS eine bessere Leistung erreichen als Frauen. Die Mittelwerte, Standardabweichungen und Effektstärken sind in Tabelle 13 dargestellt. Im Globaltest der MANOVA zeigte sich hypothesenkonform, dass

Männer bei der Anwendung des Programms sowie der Interpretation der Ergebnisse höhere Punktwerte erreichen als Frauen (Hotelling-Spur: $F [2, 32] = 6.07, p = .006$). Um zu prüfen, ob die Differenzen für beide Leistungsindikatoren gelten (Leistung_Arbeitsblatt, Leistung_SPSS), wurden zusätzlich zwei Einzeltests mit α -Fehler-Korrektur vorgenommen (adjustiertes p -Niveau: .025).

Tabelle 13

Mittelwerte, Standardabweichungen und Effektstärken für verschiedene Leistungsindikatoren für die Gesamtstichprobe bzw. getrennt für Frauen und Männer

Variable	Männer ($n = 18$)		Frauen ($n = 15$)		Effektstärke d
	M	SD	M	SD	
Leistung_SPSS (Berechnen bzw. Darstellen mit SPSS)	6.61	3.1	3.87	3.0	0.90
Leistung_Arbeitsblatt (Aufgaben auf dem Arbeitsblatt)	10.50	3.5	5.9	4.2	1.19
Punkte pro bearbeiteter Aufgabe (Quotient)	3.32	1.0	2.08	1.4	1.02
Zeit bis zur Aufgabenlösung (in Sek.) ¹					
Aufgabe 1 ($n = 27$)	533	530	864	332	- 0.75
Aufgabe 2 ($n = 24$)	198	91	313	264	- 0.58
Aufgabe 3 ($n = 21$)	346	156	320	138	0.18
Aufgabe 4 ($n = 14$)	307	113	319	173	- 0.08

Anmerkungen. d = Cohens d . ¹Veränderungen in der Datenbasis; diese sind der Aufgabenstellung geschuldet, welche die Bearbeitung der Aufgaben in der vorgegebenen Reihenfolge erforderte.

Im Hinblick auf die Leistung auf dem Arbeitsblatt zeigte sich, dass es Männer nicht nur besser vermochten, das Programm korrekt zu bedienen, sondern die SPSS-Ausgaben auch richtig zu interpretieren (Leistung_Arbeitsblatt: $F [1, 31] = 11.83, p = .002, d = 1.19$): Männer erzielten im Durchschnitt 10.5 ($SD = 3.5$) von 17 Punkten. Frauen erreichten hingegen durchschnittlich nur 5.9 Punkte ($SD = 4.2$; s. Tabelle 13 zur Variablen Leistung_Arbeitsblatt). Wurde die korrekte Bedienung des Programms alleine als Leistungsindikator berücksichtigt, zeigte sich der gleiche Effekt: Männer erreichten hier 6.6 Punkte ($SD = 3.1$) von maximal 9 Punkten. Frauen dagegen erzielten im Durchschnitt 3.87 Punkte ($SD = 3.04$). Der Unterschied in der Variablen „Leistung_SPSS“ wurde ebenfalls statistisch signifikant $F [1, 31] = 16.41, p = .017, d = .90$). Männer benötigten zudem weniger Zeit zur Lösung der Aufgabe 1 als die Frauen; es zeigten sich jedoch keine statistisch signifikanten Unterschiede im zeitlichen Umfang ($M_{\text{Männer}} = 533$ Sek., $SD = 530$ bzw. $M_{\text{Frauen}} = 864$ Sek., $SD = 332$; $t [24] = -1.69, p = .103$, siehe auch Tabelle 13).

Geschwindigkeits-Genauigkeits-Dilemma. Dass Frauen eine geringere Leistung zeigten als Männer kann auf unterschiedliche Ursachen zurückführbar sein: Zum Beispiel kann die niedri-

gere Anzahl an erreichten Punkten (Leistung_Arbeitsblatt) einer höheren Genauigkeit geschuldet sein. Eine sorgfältigere Bearbeitung der Aufgaben (Genauigkeit) führt zu einer geringen Anzahl an bearbeiteten Aufgaben (Geschwindigkeit), wenn die dafür verfügbare Zeit begrenzt ist. Unterschiede in der Prioritätensetzung (Geschwindigkeit vs. Genauigkeit) können sich auf die Anzahl insgesamt erreichter Punkte ausgewirkt haben.

Um diesem Geschwindigkeits-Genauigkeits-Dilemma (vgl. Moosbrugger & Goldhammer, 2006, S. 85) Rechnung zu tragen, wurde die Anzahl bearbeiteter Aufgaben mit der erreichten Punktzahl ins Verhältnis gesetzt. Dazu wurde ein Quotient gebildet mit der Anzahl erreichter Punkte im Zähler und der Anzahl der gelösten Aufgaben im Nenner. Frauen erreichten demzufolge $M = 2.08$ ($SD = 1.4$) Punkte pro bearbeiteter Aufgabe. Männer erzielten $M = 3.32$ ($SD = 1.0$) Punkte (s. Tabelle 13). Auch unter Berücksichtigung der Genauigkeit des Arbeitens zeigten demnach die Männer höhere Leistungen als die Frauen ($t [31] = -2.90$, $p = .007$; $d = 1.02$). Im Hinblick auf die Geschwindigkeitskomponente der Leistung ergab sich für die Gesamtstichprobe, dass die Teilnehmer in der 30-minütigen Bearbeitungszeit mit SPSS im Mittel drei der vier Aufgaben lösten ($M = 2.81$, $SD = 1.18$). 61% der Männer bearbeiteten alle vier vorgegebenen Aufgaben vollständig gegenüber 13% der Frauen. Ob die Aufgabenlösung korrekt war, wurde dabei nicht berücksichtigt.

Zusammenfassung. Die Hypothese, dass sich Männer und Frauen in der Leistung bei der Aufgabenbearbeitung in SPSS unterscheiden, wurde durch die Ergebnisse in der Stichprobe unterstützt. In beiden Leistungsindikatoren (Leistung_SPSS, Leistung_Arbeitsblatt) erreichten Frauen weniger Punkte. Ein Geschwindigkeits-Genauigkeits-*Trade-Off* konnte dabei nicht gefunden werden: Für die Vermutung, Frauen benötigten mehr Zeit für die Aufgaben, weil sie sorgfältiger arbeiteten, konnten keine empirischen Belege gefunden werden. Männer erreichten auch in dem Leistungsindikator, der Geschwindigkeit und Genauigkeit in Beziehung setzt, eine höhere Performanz.

8.2.7 Hypothese 9: Flow-Erleben und Strategiesystematik als Mediatoren

Bezugnehmend auf das kognitiv-motivationale Prozessmodell wurde in *Hypothese 9* erwartet, dass sich die Eingangsmotivation auf das Flow-Erleben sowie auf die Strategiesystematik während des Explorationsprozesses dergestalt auswirken, dass Frauen neben der Nutzung suboptimaler Explorationsstrategien auch eine ungünstigere Motivation während des Prozesses zeigen. Unterschiede in Flow-Erleben und Strategiesystematik sollten wiederum einen negativen Effekt auf die Leistung beim Lösen der Statistikaufgaben ausüben (*Mediationshypothese*).

Modellmodifikation. Da sich Männer und Frauen hinsichtlich ihrer Eingangsmotivation nicht signifikant voneinander unterschieden (siehe Abschnitt 8.2.3), wurde darauf verzichtet, das postulierte Mediationsmodell vollständig zu prüfen. Im Folgenden wird eine modifizierte Version

in Anlehnung an das kognitiv-motivationale Prozessmodell beschrieben. Das modifizierte Pfadmodell fasst die Annahmen zu Wirkungszusammenhängen zwischen den Variablen Geschlecht, Strategiesystematik, Flow-Erleben und deren Einfluss auf die Leistung zusammen. Das Modell ist in Abbildung 15 grafisch dargestellt.

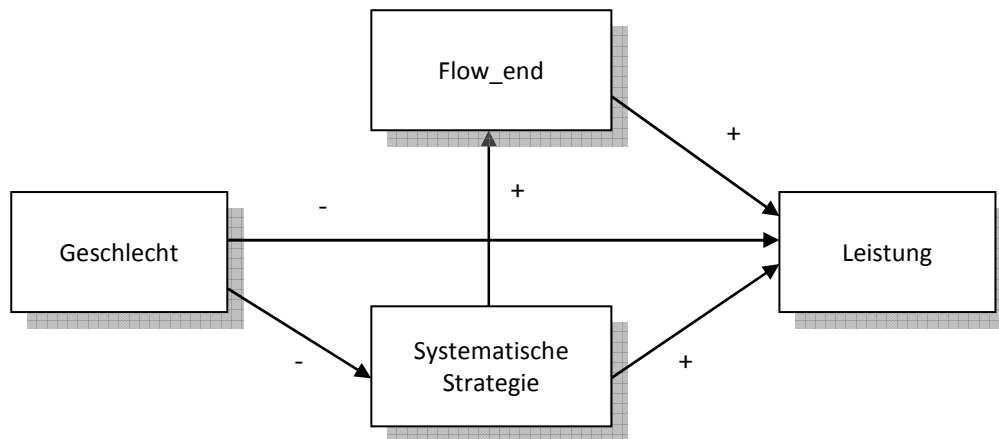


Abb. 15. Modifiziertes theoretisches Pfadmodell zur Vorhersage der Leistung mit den Variablen Geschlecht, Systematische Explorationsstrategie (alle Messzeitpunkte) und Flow-Erleben (Flow_End). Das Modell illustriert die erwartete Rolle von Strategienutzung und Flow-Erleben als Mediatoren zwischen Geschlecht und Leistung (Kodierung Geschlecht: Männer = -1, Frauen = 1). Pluszeichen spiegeln positive erwartete Pfadkoeffizienten und Minuszeichen negative erwartete Pfadkoeffizienten wider.

Als Strategiemaß wurde die Häufigkeit der Nutzung der systematischen Strategie gewählt. In der Nutzungshäufigkeit dieser Strategie hatten sich zuvor theoretisch begründbare Unterschiede zwischen den Geschlechtern gezeigt (siehe Abschnitt 8.2.4). Um die Modellkomplexität zu reduzieren, wurden die Häufigkeiten der systematischen Strategie über die drei Beobachtungszeiträume in einer Variablen zusammengefasst. Die Variable „Systematische Strategie“ beinhaltet die Summe der Häufigkeiten der Kategorie „Systematische Explorationsstrategie“ geteilt durch die individuelle Anzahl aller Beobachtungseinheiten. Das kurz vor Ende des individuellen Bearbeitungsprozesses erfasste Flow-Erleben („Flow_end“) wurde ebenfalls in das Modell aufgenommen.

In der modifizierten Version des Pfadmodells wurde postuliert, dass sich das (biologische) Geschlecht über die Häufigkeit der systematischen Strategie auf die Leistung auswirkt. Zusätzlich wurden Annahmen zur Beziehung zwischen den Prozessvariablen gemacht: Es wurde erwartet, dass die häufige Nutzung der systematischen Strategie das Flow-Erleben zum letzten Messzeitpunkt begünstigt. Es wurde weiterhin angenommen, dass beide Variablen einen eigenständigen Beitrag zur Erklärung der Varianz in der Leistung aufweisen.

Um auch jene direkten Einflüsse des Geschlechts abbilden zu können, die nicht durch die beiden Mediatorvariablen berücksichtigt werden, wurde zusätzlich ein direkter Pfad zwischen Geschlecht und Leistung aufgenommen. Für die Aufnahme dieses Pfads sprachen zwei Überlegungen: *Erstens* ist die Höhe eines Pfadkoeffizienten (vergleichbar zu Regressionsgewichten in der multiplen Regressionsanalyse) von weiteren Variablen respektive Pfaden im Modell abhängig. Demzufolge kann sich der Beitrag an Varianzaufklärung, der über die einzelnen Pfadkoeffizienten quantifizierbar ist, durch Hinzufügen von Variablen verändern. *Zweitens* verweisen Vollmeyer und Rheinberg (2000, 2006) in ihrem Modell auf die Bedeutung weiterer möglicher Prozessvariablen wie z. B. den motivationalen Zustand, die Persistenz und die Metakognition. Vor diesem Hintergrund erscheint die Annahme, dass sich Geschlechtsunterschiede auch in weiteren, im Modell nicht berücksichtigten Mediatorvariablen manifestieren können, als plausibel.

Statistische Modellprüfung. Die Überprüfung des postulierten Pfadmodells erfolgte mit Hilfe von *Mplus* 5.0 (Muthén & Muthén, 1998-2007). Der von *Mplus* analysierte Rohdatensatz beinhaltete keine fehlenden Werte. Im Folgenden werden Informationen zu den Voraussetzungen des Analyseverfahrens, zur Modellidentifikation sowie zur Modellschätzung zusammengefasst (vgl. APA Publication Manual, 2001; McDonald & Ho, 2002; Raykov & Marcoulides, 2000; Schumacker & Lomax, 2004). Es wird angenommen, dass alle endogenen Variablen (Häufigkeit systematischer Strategie, Flow_end, Performanz) auf Intervallskalenniveau gemessen wurden. Die exogene Variable Geschlecht ist nominalskaliert. Das Modell (s. Abb. 14) beinhaltete vier manifeste (beobachtete) und drei latente Variablen (*disturbances*), davon vier exogene und drei endogene. Die Skalierung der latenten Variablen wurde auf 1 festgesetzt. Somit standen $[4 (4+1)] / 2 = 10$ Stichprobenmomente neun zu schätzenden Parametern gegenüber. Das Modell ist damit überidentifiziert und die vorliegenden Informationen sind zur Schätzung der Parameter des spezifizierten Modells ausreichend (Kline, 2005, S. 105f.). Voraussetzung für die Maximum-Likelihood (ML)-Parameterschätzung ist die Multinormalverteilung der Verteilungen (Kline, 2005, S. 48 f.; Schumacker & Lomax, 2005, S. 230f.). Die Ergebnisse von Simulationsstudien zeigen jedoch, dass die ML-Parameterschätzungen relativ robust sind gegenüber Voraussetzungsverletzungen. Lediglich bei der Schätzung von Standardfehlern ist dies nicht der Fall (Hu & Bentler, 1995; West et al., 1995). Der Mardia-Tests auf multivariate Normalverteilung berechnet mit AMOS 17.0 ergab eine multivariate Kurtosis von -0.28 (*Critical Ratio*: -1.2). Dieser Wert liegt unter dem bei Bühner (2006, S. 285) angegebenen kritischen Wert von 1.96. Die Voraussetzung der multivariaten Normalverteilung galt somit als erfüllt, so dass als Schätzmethode der ML-Algorithmus gewählt wurde.

Die Deskriptiven Statistiken der einzelnen Variablen sind in Tabelle 14 nachzulesen. Die Tabelle gibt darüber hinaus die Interkorrelationen der Variablen sowie die Modellresiduen wieder.

Es zeigt sich, dass die Variablen des Modells nicht übermäßig miteinander korreliert sind. Die Pearson-Produkt-Korrelationen unterschreiten die von Kline (2005, S. 56) vorgeschlagene Grenze von $r = .85$. Verzerrungen der geschätzten Parameter aufgrund von Multikollinearität der Variablen konnten somit ausgeschlossen werden.

Tabelle 14

Deskriptive Statistiken und Interkorrelationen der Variablen im Pfadmodell sowie Diskrepanzen als Abweichung der empirischen von der durch das Modell geschätzten Kovarianzmatrix (N = 33)

Variablen	Deskriptive Statistiken				Korrelationen und Diskrepanzen			
	M	SD	Kurtosis	Schiefe	1	2	3	4
1 Geschlecht ¹	-	-	-	-	1.48	0	0	.01
2 Systematische Strategie ²	40.6	12.1	-.38	-.02	-.42*	1.06	0	-.01
3 Flow_End ³	4.3	1.05	-.13	-.45	-.22	.57**	18.99	-.04
4 Leistung_Arbeitsblatt	8.4	4.4	-.65	-.61	-.53**	.68**	.64**	0.25

Anmerkungen. Die Stichprobenkorrelationen sind oberhalb der Diagonale und die Diskrepanzen unterhalb der Diagonale wiedergegeben. In der Diagonalen sind die Kovarianzen abgetragen. ¹Für die kategoriale Variable Geschlecht wurden keine Deskriptiven Statistiken berechnet; die Variable wurde jedoch in Korrelationsanalysen einbezogen (Kodierung: Männer = -1, Frauen = 1). ²In der Variablen wurden die relativen Häufigkeiten der drei Beobachtungszeiträume zusammengefasst. ³Die Variable erfasst das Flow-Erleben zum letzten Messzeitpunkt vor dem individuellen Bearbeitungsende. * $p < .05$, ** $p < .01$.

Datengrundlage. Der Datenanalyse lag ein vergleichsweise geringer Stichprobenumfang ($N = 33$) zugrunde. Das Verhältnis von Probanden und zu schätzenden Parametern (9) betrug 3.7:1. Einige Autoren (Backhaus et al., 2006; Barrett, 2007; Kline, 2005) geben zu bedenken, dass ein zu geringer Stichprobenumfang (neben weiteren Einflussgrößen) einen ungünstigen Effekt auf die Stabilität bzw. auf die Zuverlässigkeit der Parameterschätzungen habe. Die Autoren nennen eine Stichprobengröße von $N > 100$ (Backhaus et al., 2006, S. 371) bzw. $N > 200$ (Barrett, 2007, S. 820) als Minimalanforderung für die Berechnung von Strukturgleichungsmodellen. Als Daumenregel führt Kline (2005, S. 110 f.; siehe auch Schumacker & Lomax, 2004, S. 50) an, dass die statistische Präzision der Ergebnisse einer Pfadanalyse ab einem Verhältnis von weniger als 5 Probanden pro geschätztem, freien Parameter „zweifelhaft“ werde. Muthén und Muthén (2002, zit. nach Schermelleh-Engel et al., 2003, S. 50) zufolge kann hierbei jedoch nicht von einer allgemeinen Daumenregel ausgegangen werden, die für alle Fälle gleichermaßen gelte.

Nach Schermelleh-Engel et al. (2003, S. 50) hängt die Angemessenheit der Parameterschätzmethode neben der Stichprobengröße z. B. auch von der Komplexität des Modells, der Anzahl der Indikatorvariablen, dem Anteil fehlender Werte sowie der Reliabilität der Variablen ab. McCallum, Browne und Sugarawa (1996, S. 144; vgl. auch „notwendige Bedingung“ nach Backhaus

et al., 2006, S. 367) geben als Richtlinie für die Maximum-Likelihood-Parameterschätzung an, dass die Stichprobengröße mindestens größer oder gleich der Anzahl beobachteter Variablen sein sollte. Für die Angemessenheit der Methode bei mittleren und kleinen Stichprobenumfängen sprechen auch die Befunde aus Simulationsstudien (z. B. Iacobucci et al., 2007; siehe auch Kap. 8.1.4), nach denen sich einfache Strukturgleichungsmodelle gegenüber regressionsanalytischen Berechnungen als überlegen erwiesen (z. B. zur Modellierung von Mediatoreffekten; s. Hoyle & Kenny, 1999). Iacobucci et al. (2007, S. 151) kritisieren demzufolge auch die Maßgabe einer Stichprobengröße von $N > 200$ als übermäßig konservativ. Vor diesem Hintergrund wurde die vorhandene Stichprobengröße für die Berechnung im Rahmen der vorliegenden Studie als ausreichend beurteilt.

Modellgüte. Die Beurteilung der Gesamtstruktur erfolgte anhand inferenzstatistischer und deskriptiver Indizes der globalen Modellgüte und Modellsparsamkeit (Barrett, 2007; Hu & Bentler, 1999; Kline, 2005; Schumacker & Lomax, 2004). Schermelleh-Engel et al. (2003) empfehlen aufgrund der Vor- und Nachteile einzelner Indizes sowie der Probleme allgemeingültiger Grenz- oder Schwellenwerte mehrere Maße zur Beurteilung eines Modells heranzuziehen.

Tabelle 15

Inferenzstatistische und deskriptive Maße zur Bewertung der globalen Modellgüte für das vorgeschlagene Modell

Modell	Index der Modellgüte				
	χ^2 (df; p)	RMSEA (90% CI)	SRMR	CFI	AIC ¹
Vorgegebene Modellstruktur	0.01 (1; .94)	.01 (.01, .12)	.01	1.00	18.01

Anmerkungen. df = Freiheitsgrade; p = Signifikanzniveau; CI = Konfidenzintervall. ¹ Das saturierte Modell lieferte hier einen Wert von AIC = 20 und das Independence-Modell erreichte einen AIC von 58; damit gilt das vorgeschlagene als das sparsamste der drei Modelle.

Eine Übersicht über ausgewählte Modellgüte-Indizes für die vorgegebene Modellstruktur findet sich in Tabelle 15. Nach den Richtlinien von Schermelleh-Engel et al. (2003, S. 52) spiegelt der empirische χ^2 -Wert von .01 eine gute Anpassung der empirischen Daten an das theoretische Modell wider. Der berechnete RMSEA-Wert überschreitet nicht die empfohlene Grenze von $RMSEA \leq .05$, was für eine gute Passung des Modells an die empirischen Daten spricht. Allerdings umschließt das Konfidenzintervall im oberen Bereich mögliche „wahre“ RMSEA-Werte von über $p = .10$. Die Hypothese, das Modell habe eine schlechte Passung, konnte nicht falsifiziert werden. Dieses Problem wird besonders für den Fall kleinerer Stichproben berichtet (Kline, 2005, S. 139). Kline (2005) rät hier zu einer Replikation der Modelltestung auf der Basis größerer Stichproben. Der SRMR (*Standardized Root Mean Square Residual*; Kline, 2005, S. 141) stellt ein skalenunabhängiges Maß für die *Badness-of-Fit* eines Modells dar und ist ein Maß für die absolute, mittlere

Abweichung zwischen den modellimplizierten von den empirischen Korrelationen. Da ein Wert von $<.05$ (Schermelleh-Engel et al., 2003) als gut bezeichnet werden kann, spricht auch dieses Kriterium für die angemessene Güte des getesteten Modells.

Für deskriptive Anpassungsmaße wie den *CFI* sind keine kritischen Werte definiert, da deren Verteilungsfunktionen unbekannt sind (Schermelleh-Engel et al., 2003). Hu und Bentler (1999) haben als Daumenregel vorgeschlagen, Modelle, deren *CFI* einen Wert von $.95$ überschreiten, als akzeptabel zu betrachten. Demzufolge legt dieser Modellgüte-Index (siehe Tabelle 15) ebenfalls einen sehr guten Modellfit nahe. Eine Inspektion der Teilstrukturen im Hinblick auf die standardisierten Residuen der Kovarianzen erbrachte keine Auffälligkeiten. Die größte Diskrepanz betrug $|.04|$ (Jöreskog & Sörbom, 1984, zit. nach Arbuckle, 2008; vgl. auch McDonald & Ho, 2002). Danach weicht die Anpassung der empirischen Daten an das theoretische Modell in einzelnen Teilstrukturen nicht bedeutsam voneinander ab.

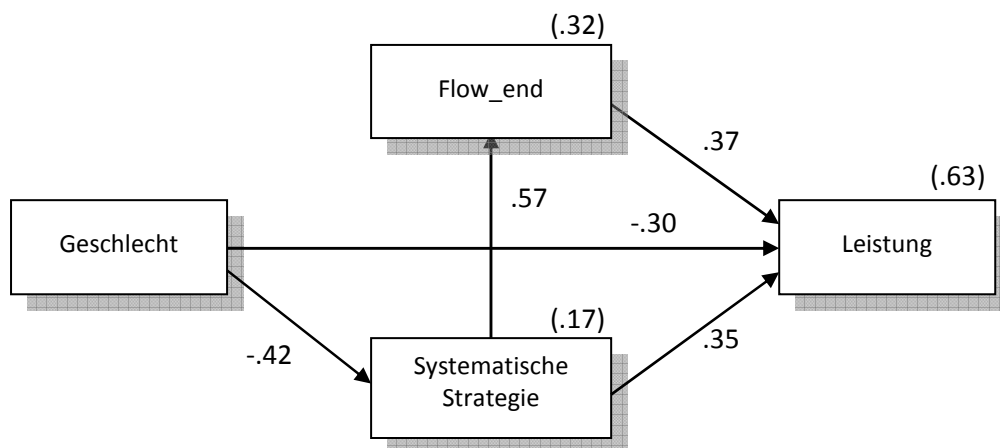


Abb. 16. Empirisches, rekursives Modell mit standardisierten Pfadkoeffizienten für den Zusammenhang der Variablen Geschlecht (Kodierung: Männer = -1, Frauen = 1), systematischer Explorationsstrategie und Flow-Erleben zur Vorhersage der Leistung. Alle Pfadkoeffizienten sind auf dem Niveau von $p <.05$ signifikant. Der Anteil der durch das Modell erklärten Varianz der jeweiligen endogenen Variablen (R^2) ist in Klammern dargestellt.

Signifikanzprüfung der direkten, indirekten und der Gesamteffekte. Abbildung 16 illustriert die Beziehungen der Variablen untereinander. Dargestellt sind die standardisierten Pfadkoeffizienten (St.) sowie der Anteil erklärter Varianz der endogenen Variablen des Modells (R^2 ; vgl. Kline, 2005, S. 127). Zur inferenzstatistischen Prüfung der Gesamteffekte bzw. der zusammengefassten indirekten Effekte wurde der Standardfehler des interessierenden Parameters mittels *Bootstrapping* (Byrne, 2001, S. 267 ff., S. 281 ff.; Kline, 2005, S. 130 f., s. auch West et al., 1995) auf der Basis von 500 Stichproben und des Maximum-Likelihood-Algorithmus in *Mplus* 5.0 geschätzt. Eine

detaillierte Übersicht über alle direkten, zusammengefassten indirekten sowie die Gesamteffekte des getesteten Pfadmodells bietet Tabelle 16.

Tabelle 16

Effekterlegung (direkte Effekte, indirekte Effekte, Gesamteffekte) für das überidentifizierte Pfadmodell zur Vorhersage der Leistung über die Variablen Geschlecht, Flow und Nutzung Systematischer Explorationsstrategie¹

Effekt	USt.	SE	USt./SE	P	St.
Geschlecht → Leistung					
Gesamt	-4.63	1.27	-3.64	.01	-.53
Gesamt indirekt	-2.04	0.86	2.38	.02	-.23
Spezifisch indirekt					
Geschlecht → Strategie → Leistung	-1.27	0.75	-1.70	.09	-.15
Geschlecht → Strategie → Flow → Leistung	-0.77	0.54	-1.43	.15	-.09
Direkt	-2.60	1.03	-2.53	.01	-.30
Geschlecht → Strategie → Flow	-0.49	0.24	-2.04	.04	-.24
Strategie → Flow → Leistung	0.75	0.43	1.77	.08	.21
Geschlecht → Strategie	-1.01	0.39	-2.57	.01	-.42
Strategie → Flow	0.48	0.15	3.17	.01	.57
Strategie → Leistung	1.26	0.61	2.08	.04	.35
Flow → Leistung	1.58	0.67	2.37	.02	.37

Anmerkungen. ¹ USt. = Unstandardisierter Pfadkoeffizient; SE = Standardfehler; St. = Standardisierter Pfadkoeffizient; p = Signifikanzniveau, 2-seitig.

Ich fasse zunächst die Ergebnisse zu den direkten Effekten im Modell zusammen. Wie erwartet nutzten Männer häufiger systematische Explorationsstrategien als Frauen ($St. = -.42, p = .01$). Die Nutzung systematischer Explorationsstrategien beeinflusste wiederum positiv das Flow-Erleben ($St. = .57, p = .01$). Flow-Erleben ($St. = .37, p = .02$) und systematische Explorationsstrategie ($St. = .35, p = .04$) trugen unabhängig voneinander zur Erklärung der Varianz in der Leistung bei. Ein zusätzlicher, bedeutsamer Anteil der Varianz in der Variablen Leistung wurde jedoch nicht durch die Mediatoren Flow-Erleben und Strategienutzung erklärt. Der direkte Pfad von Geschlecht auf Leistung wurde ebenfalls statistisch signifikant ($St. = -.30, p = .01$). Insgesamt wurden durch die direkten Pfade 63% der Varianz des Kriteriums Leistung aufgeklärt (siehe Angaben zu quadrierten multiplen Korrelationskoeffizienten in Abbildung 15). Lediglich 37% der Varianz gingen demnach auf die Wirkung anderer Einflussfaktoren zurück (vgl. Backhaus et al., 2006, S. 377).

In einem nächsten Schritt wurden die Pfadkoeffizienten für indirekte Pfade getestet. Die indirekten Pfade repräsentieren mögliche Mediatoreffekte mit einer oder mehreren Mediatorvariab-

len (Kline, 2005, S. 162; s. auch Iacobucci et al., 2007). Das Modell sieht zwei indirekte Pfade von Geschlecht auf Leistung vor: 1. über die Häufigkeit der systematischen Exploration; 2. über die systematische Exploration sowie über das Flow-Erleben. Beide separaten indirekten Pfade von Geschlecht zu Leistung verfehlten die statistische Signifikanzgrenze ($St. = -.15, p = .09$ bzw. $St. = -.09, p = .15$). Wenn die Effekte der beiden indirekten Pfade zusammengefasst wurden, zeigte sich ein statistisch signifikanter Gesamteffekt auf die Leistung ($St. = -.23, p = .02$; siehe „Gesamt indirekt“ in Tabelle 16). Die indirekten Pfade erklären jedoch insgesamt etwas weniger Varianz ($R^2 = |-.23|^2$) als der direkte Pfad von Geschlecht auf Leistung ($R^2 = |-.30|^2$).

Es zeigt sich, dass sich das Geschlecht über die Strategiesystematik auf die Höhe des Flow-Erlebens auswirkte: Der Mediationseffekt wurde statistisch signifikant ($St. = -.24, p = .04$). Ein direkter Pfad von Geschlecht zu Flow war im Modell nicht aufgenommen worden, da sich in dieser Variablen zuvor keine statistisch signifikanten Geschlechtsunterschiede gezeigt hatten (siehe Abschnitt 8.2.5).

Zusammenfassung. Die Ergebnisse zu Hypothese 9 lassen sich folgendermaßen zusammenfassen: Die niedrigere Leistung der weiblichen Teilnehmer lässt sich so erklären, dass sie über eine ungünstigere Strategiesystematik hinaus während des Arbeitens ein geringer ausgeprägtes Gefühl hatten, die eigenen Handlungsschritte unter Kontrolle zu haben und von der Aufgabe absorbiert zu sein (Flow-Erleben).

Die Ergebnisse zu einzelnen indirekten Effekten ergaben, dass sich das Geschlecht sowohl direkt als auch indirekt über die Strategiesystematik und das Flow-Erleben auf die Höhe der Leistung auswirkte. Das Geschlecht beeinflusste die Strategiesystematik, indem Männer häufiger systematische Strategien nutzten. Diese günstige Strategiesystematik bei den Männern hatte wiederum einen direkten positiven Effekt auf die Leistung: Das systematische Ausführen, Überwachen, Reflektieren und gegebenenfalls Korrigieren von Lösungsschritten beim Erlernen von SPSS führte in erster Linie die männlichen Probanden zum richtigen Lösungsweg und damit zur richtigen Aufgabenlösung. Der direkte Pfad von der Strategiesystematik auf das Flow-Erleben verdeutlicht, dass die Nutzung systematischer Strategien von einem positiven Flow-Erleben begleitet wurde bzw. dazu führte. Frauen kamen infolge ungünstigerer Strategienutzung weniger häufig in den Zustand leistungsförderlichen Flow-Erlebens. Das Flow-Erleben begünstigte zusätzlich eine höhere Leistung und trug mit einem eigenständigen Beitrag zur Performanz bei. Die Hinzunahme des Flow-Erlebens in das Modell führte dementsprechend zu einer höheren Varianzaufklärung in der Kriteriumsvariablen Leistung.

Der indirekte Pfad von der Strategiesystematik über das Flow-Erleben zur Leistung verfehlte die statistische Signifikanzgrenze. Dies ist dem direkten Effekt der Strategiesystematik auf die

Leistung geschuldet. Die Häufigkeit systematischer Strategien beeinflusste auch direkt die Leistung, ohne dass dabei zwangsläufig eine günstige Motivation auftrat. Dass beide Variablen (Flow und Strategie) einen direkten signifikanten Effekt auf die Leistung hatten, deutet jedoch darauf hin, dass die Motivation unabhängig von der Strategiesystematik eine Rolle beim Zustandekommen der Leistung hatte. Der direkte Effekt von Geschlecht auf die Leistung weist darauf hin, dass sich Männer und Frauen in weiteren Variablen unterscheiden. Der nicht durch das Modell erklärte Anteil der Varianz in der Leistung kann beispielsweise auf metakognitive oder affektive Variablen zurückführbar sein, die jedoch in der Studie nicht miterfasst wurden.

8.2.8 Geschlechtsunterschiede im Verlauf der aktuellen Motivation

Dem kognitiv-motivationalen Prozessmodell zufolge hat die aktuelle Motivation im Lern- bzw. Problemlöseprozess vor als auch während der Ausführung der Lernhandlungen einen Einfluss auf das Lernergebnis (vgl. Vollmeyer & Rheinberg, 2006). Wie die Befunde zu Hypothese 6 zeigen, unterschieden sich Frauen und Männer im Hinblick auf ihre Eingangsmotivation nicht. Im Folgenden wurde getestet, ob sich Unterschiede in der aktuellen Motivation während der Aufgabenbearbeitung zeigten. Die aktuelle Motivation während der Aufgabenbearbeitung wurde in der Studie mit acht ausgewählten Items des FAM (Rheinberg et al., 2001) erfasst und als Netto-Motivation bezeichnet. Höhere Summenscores in dieser Variablen spiegeln dabei höhere Werte in den FAM-Skalen Interesse, Herausforderung und Erfolgswahrscheinlichkeit sowie niedrige Werte in der Skala Misserfolgsbefürchtung wider.

Multivariate Varianzanalyse. Eine MANOVA ergab statistisch signifikante Geschlechtsunterschiede (Globaltest: $F [4, 16] = 3.44, p = .03, d = 1.19$): Männer zeigten zu zwei der drei Messzeitpunkte (t_1 und t_2), zu denen die Teilnehmer die Aufgaben in SPSS bearbeiteten, eine höhere Netto-Motivation (bei adjustiertem Signifikanzniveau von $p = .0125$ und $m = 3$ Hypothesentests). Das Geschlecht hatte dabei einen starken Effekt. Die Ergebnisse der Einzeltests inklusive der Effektstärken sind in Tabelle 17 aufgeführt. Differenzen in der Netto-Motivation vor Beginn der Aufgabenbearbeitung wurden nicht statistisch signifikant ($F [4, 16] = 0.95, p = .34$). Dies war bereits aufgrund der Ergebnisse zu Hypothese 6 erwartet worden.

Tabelle 17

Mittelwerte, Standardabweichungen und Effektstärken für die Netto-Motivation (8 Items) vor und während der Aufgabenbearbeitung für die Gesamtstichprobe bzw. getrennt für Frauen und Männer

Variable	Männer (n = 18)		Frauen (n = 15)		MANOVA (df ₁ = 1, df ₂ = 31)		Effekt- stärke <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	
Netto-Motivation t0 (vor Aufgabenbearbeitung)	10.97	2.7	9.63	2.1	0.95	.34	0.57
Netto-Motivation t1	11.11	3.8	7.37	3.0	9.84	.01	1.07
Netto-Motivation t2	11.36	3.1	6.82	2.9	10.85	.01	1.51
Netto-Motivation t3	11.50	3.0	7.61	3.9	5.28	.03	1.12

Anmerkungen. *d* = Cohens *d*. ¹ Levene-Test: Messzeitpunkt t0: $F [1, 19] = 2.38, p = .14$; Messzeitpunkt t1: $F [1, 19] = 0.72, p = .41$; Messzeitpunkt t2: $F [1, 19] = 0.01, p = .98$; Messzeitpunkt t3: $F [1, 19] = 0.32, p = .58$. Box-M-Test: $F [10, 672] = 25.52, p = .06$. ¹ Messzeitpunkte nach Bearbeitungsbeginn; Reduzierter Stichprobenumfang; Veränderungen in der Datenbasis sind hierbei der Aufgabenstellung geschuldet; Messzeitpunkt t1: 15 Frauen, 18 Männer; Messzeitpunkt t2: 15 Frauen, 13 Männer; Messzeitpunkt t3: 14 Frauen, 7 Männer. Die Netto-Motivation ergibt sich aus der Summe der Items in den Skalen Herausforderung, Interesse und Erfolgswahrscheinlichkeit abzüglich des Wertes in der Skala Misserfolgsbefürchtung.

Produkt-Moment-Korrelationen. Die Interkorrelationen der Netto-Motivation vor und während der Aufgabenbearbeitung mit beiden Leistungsindikatoren sind in Tabelle 18 aufgeführt.

Tabelle 18

Produkt-Moment-Korrelationen der Netto-Motivation mit der Leistung in der Gesamtstichprobe ($N = 33$)

Variable	Deskriptive Statistik		Interkorrelationen				
	<i>M</i>	<i>SD</i>	2	3	4	5	6
1 Netto-Motivation t0 (vor Aufgabenbearbeitung)	10.64	3.0	.65**	.47*	.54*	.14	.22
2 Netto-Motivation t1	8.23	3.3		.73**	.71**	.40*	.40*
3 Netto-Motivation t2	8.33	3.6			.70**	.41*	.41*
4 Netto-Motivation t3	8.9	4.0				.39*	.55*
5 Leistung_Arbeitsblatt	6.61	3.1					.88*
6 Leistung_SPSS	10.50	3.5					

Anmerkungen. * $p < .05$, ** $p < .01$; zweiseitig.

Die Netto-Motivation vor Beginn der Aufgabenbearbeitung (Netto-Motivation t0) wies in der Gesamtstichprobe keine statistisch signifikante Beziehung zur Leistung im Arbeitsblatt auf ($r =$

.22, $p = .23$). Im Gegensatz hierzu stand die Netto-Motivation zum ersten Messzeitpunkt (t1) bereits in positiver Beziehung zur Leistung ($r = .40, p = .02$). Es stellte sich die Frage, ob sich bereits in den ersten 10 Minuten nach Beginn der Bearbeitung der Aufgaben in SPSS Unterschiede in der aktuellen Motivation ergeben, welche mit den Leistungsunterschieden zwischen Frauen und Männern kovariieren.

Um diese Frage nach Veränderungen über die Zeit zu klären, wurde der Verlauf der Netto-Motivation (Abhängige Variable) vor und kurz nach Beginn der Aufgabenbearbeitung für beide Geschlechter (Gruppierungsvariable) mittels einer Varianzanalyse mit Messwiederholung ausgewertet. Als Einflussfaktoren zählten neben dem Messzeitpunkt auch interindividuelle Unterschiede in der Gruppenzugehörigkeit sowie Interaktionen zwischen dem Messzeitpunkt und der Gruppenzugehörigkeit. Den Gruppierungsfaktor stellte dabei das Geschlecht dar. Der Messwiederholungsfaktor hatte ebenfalls zwei Stufen: Die Netto-Motivation wurde über die beiden Messzeitpunkte t0 (vor Beginn der Bearbeitung) und t1 (nach 10 min) verglichen. Für die Aufnahme des ersten Messzeitpunkts nach Bearbeitungsbeginn in diese Analyse sprach neben inhaltlichen Erwägungen auch das vollständige Vorliegen der Daten zum Messzeitpunkt. Diese Bedingung war auch für den Messzeitpunkt t1 erfüllt. Um die beiden Messzeitpunkte vergleichbar zu machen, wurden für die Eingangsmotivation lediglich die Werte in den acht Items zu einem Wert der Netto-Motivation zusammengefasst, welche auch während des Prozesses erfasst wurden.

Varianzanalyse mit Messwiederholung. Die Ergebnisse der Varianzanalyse mit Messwiederholung sind in Tabelle 19 zusammengefasst.

Tabelle 19

Ergebnisse der Varianzanalyse mit Messwiederholung für die Faktoren Geschlecht und Messzeitpunkt sowie die Abhängige Variable Netto-Motivation zu den Messzeitpunkten t0 und t1 (N = 33)^{1 2}

Quelle	F	p	η^2_{partiell}
Innerhalb von Personen			
Messzeitpunkt	4.79	.04	.13
Messzeitpunkt x Geschlecht	6.12	.02	.17
Zwischen Personen			
Geschlecht	7.41	.01	.19

Anmerkungen. η^2_{partiell} = Effektstärke; Freiheitsgrade: $df_1 = 1; df_2 = 31$. ¹ *Levene-Test:* Messzeitpunkt t0: $F [1, 31] = 0.76, p = .39$; Messzeitpunkt t1: $F [1, 31] = 0.16, p = .69$. *Box-M-Test:* $F [3, 899337] = 0.43, p = .73^2$ Der Intercept („Konstante“) ist in der Tabelle nicht aufgeführt.

Für das Geschlecht zeigte sich ein statistisch signifikanter Haupteffekt: Danach lag die Netto-Motivation der Männer zu beiden Messzeitpunkten signifikant über derjenigen der Frauen ($F [1, 31] = 7.41, p = .01, \eta^2_{\text{partiell}} = .19$). Der Messzeitpunkt alleine trug ebenfalls zur Varianzaufklärung bei: Demzufolge waren die Teilnehmer im Durchschnitt vor dem Arbeiten mit SPSS stärker motiviert als nach den ersten 10 Minuten. Die Kombination der beiden Faktoren (Messzeitpunkt, Geschlecht) ergab einen signifikanten Interaktionseffekt: Die beiden Variablen interagierten insofern, dass die simultane Betrachtung von Messzeitpunkt und Geschlecht ca. 17% ($\eta^2_{\text{partiell}} = .17$) der Varianz in der Variablen Netto-Motivation aufklärt ($F [1, 31] = 6.12, p = .02$). Der Interaktionsterm leistete damit einen beachtlichen Erklärungsbeitrag.

Inferenzstatistische Prüfung der Mittelwertdifferenz in den Gruppen. Im Anschluss an die Varianzanalyse wurden t -Tests für verbundene Stichproben durchgeführt. Mit diesem Vorgehen sollte festgestellt werden, ob die Werte in *beiden Gruppen* (Männer und Frauen) in der Variablen „Netto-Motivation“ über die Zeit signifikant systematisch variierten oder ob dies *nur für eine Gruppe* zutreffend war. Die inferenzstatistische Prüfung der Mittelwertdifferenz in der aktuellen Motivation für die beiden Messzeitpunkte wurde in der Gruppe der Frauen signifikant ($M_{t_0} = 9.63, SD_{t_0} = 2.1; M_{t_1} = 7.33, SD_{t_1} = 3.0; t [14] = 3.54, p = .004$). Im Gegensatz dazu zeigten sich bei den männlichen Teilnehmern keine statistisch signifikanten Veränderungen in der Höhe der Netto-Motivation vor Bearbeitungsbeginn und nach 10 Minuten Bearbeitungszeit ($M_{t_0} = 10.97, SD_{t_0} = 2.7; M_{t_1} = 11.11, SD_{t_1} = 3.9; t [17] = -.20, p = .85$). Dieses Ergebnis verdeutlicht, dass die aktuelle Motivation in der Gruppe der Männer über die beiden Messzeitpunkte hinweg konstant geblieben war. Dahingegen nahm die Netto-Motivation der Frauen über die Zeit ab. Die Verläufe sind in Abbildung 17 grafisch dargestellt.

Zusammenfassung. Der gegenläufige Linienverlauf in Abbildung 16 verdeutlicht, dass Frauen vor Beginn eine höhere Netto-Motivation berichteten als zum ersten Messzeitpunkt während des Prozesses. Das bedeutet, dass die Frauen nachdem sie einige Zeit mit dem Aufgabenmaterial und SPSS gearbeitet hatten, ihre Erfolgsaussichten als niedriger einstufen, weniger interessiert waren und eine geringere Herausforderung erlebten. Die aktuelle Motivation der Männer lag der Tendenz nach bereits vor Beginn der Bearbeitung leicht über derjenigen der Frauen und blieb über die Messzeitpunkte hinweg weitgehend stabil. Nach Bortz (2005, S. 301) kann hier von einem statistisch signifikanten, disordinalen Interaktionseffekt gesprochen werden. Die Höhe der aktuellen Motivation während des Bearbeitungsprozesses stand in positiver Beziehung zur Leistung. Danach korrelierte die niedrigere aktuelle Motivation der Frauen mit einer niedrigen Performanz. Aufgrund der geringen Stichprobengröße und des Datenausfalls zu den Messzeitpunkten t_2 und t_3

konnte weder der weitere Verlauf der aktuellen Motivation noch die Höhe in den spezifischen FAM-Skalen näher untersucht werden.

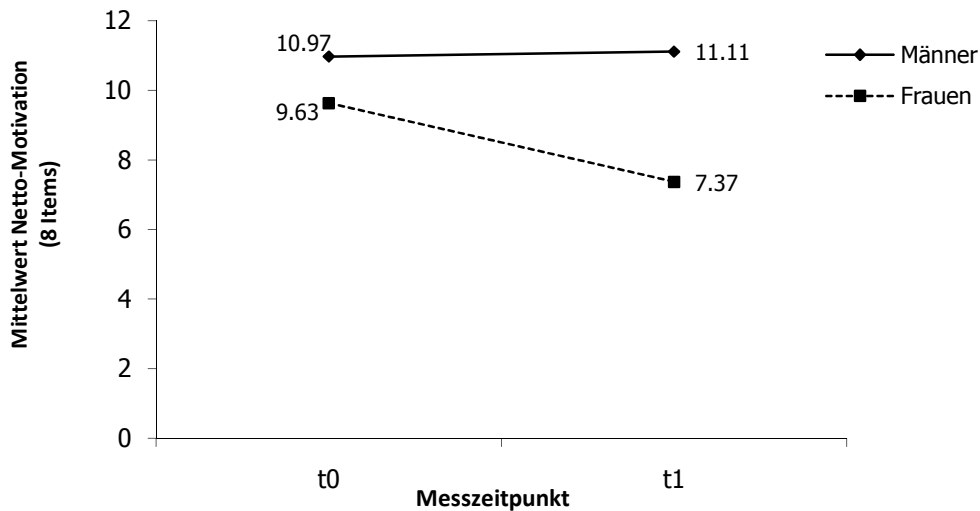


Abb. 17. Höhe der Netto-Motivation zu Messzeitpunkt t0 (vor Beginn) und t1 (10 Min nach Bearbeitungsbeginn) getrennt für Frauen ($n = 15$) und Männer ($n = 18$).

8.3 Zusammenfassung und Diskussion

Zusammenfassung und Diskussion der Ergebnisse. Erwartungsgemäß unterschieden sich die Geschlechter in einer Vielzahl der erhobenen Untersuchungsvariablen: Für den Zeitraum vor der Studienteilnahme gaben die männlichen Teilnehmer an, den Computer häufiger genutzt zu haben als dies bei den weiblichen Teilnehmern der Fall war. Jedoch zeigten Frauen und Männer keine Unterschiede in der Eingangsmotivation vor Beginn des Arbeitens mit SPSS. Nach dem Lesen der Instruktion der Aufgabe schätzten Frauen demzufolge ihre Erfolgswahrscheinlichkeit gleich hoch ein wie ihre männlichen Altersgenossen. Darüber hinaus fühlten sie sich ebenso interessiert und in gleichem Maße herausgefordert. Überdies gaben die Geschlechter keine gravierenden Unterschiede in der Misserfolgsbefürchtung beim Erlernen des neuen Computerprogramms an. Der Vorsprung der männlichen Teilnehmer hinsichtlich der Computererfahrung führte demnach nicht automatisch zu einer geringeren Eingangsmotivation bei den Frauen.

Unabhängig von dem erwarteten Geschlechtseffekt spiegeln die Ergebnisse die von Rheinberg et al. (2001) berichtete Tendenz wider, dass sich das Lösen einer Problemlöseaufgabe vor allem auf die Herausforderung auswirkt. Im Durchschnitt bewerteten die Teilnehmer die Aufgaben in Relation zu ihren subjektiven Fähigkeitseinschätzungen vor Beginn der Aufgabenbearbeitung weder als zu schwierig noch als zu leicht. Sie fühlten sich herausgefordert ohne gleichzeitig ein

starkes Gefühl der Bedrohung durch einen Misserfolg zu erleben. Im Hinblick auf die Gesamtstichprobe lässt sich damit auf eine günstige Eingangsmotivation schließen.

Kurz nach Bearbeitungsbeginn wurden jedoch signifikante Geschlechtsunterschiede deutlich: Die aktuelle Motivation der Frauen entwickelte sich bereits in den ersten 10 min nach Bearbeitungsbeginn ungünstig. Sie sank deutlich unter das Anfangsniveau in dieser Gruppe. In der Gruppe der Männer hingegen lag sie über die Messzeitpunkte hinweg konstant im leicht überdurchschnittlichen Bereich. Diese divergente Entwicklung führte zu dem dargestellten Interaktionseffekt zwischen Messzeitpunkt und Geschlecht. Tendenziell zeigten die Frauen auch hinsichtlich des Flow-Erlebens ungünstigere Werte. Der Geschlechtseffekt wurde in dieser Variablen jedoch nicht signifikant. Flow wird als ein Indikator für positive Befindlichkeit während des Arbeitsprozesses sowie eines Aufgehens in der Tätigkeit betrachtet (z.B. Rheinberg, 2006). In Übereinstimmung mit früheren empirischen Untersuchungen erwies sich Flow als positiv mit der Leistung korreliert.

Frauen erreichten im Vergleich zu ihren männlichen Altersgenossen ein geringeres Leistungsniveau beim Lösen der Statistikaufgaben mit SPSS. Sie benötigten bereits für die Lösung der Aufgabe 1 mehr Zeit als ihre männlichen Altersgenossen, lösten insgesamt weniger Aufgaben und erreichten folglich weniger Punkte. Dieser Befund konnte nicht mit einem ungünstigeren *Trade-off* von Genauigkeit und Geschwindigkeit in der Arbeitsweise der Frauen erklärt werden: Männer arbeiteten nicht weniger sorgfältig bei der Lösung der einzelnen Aufgabe. Bedeutsame Geschlechtsunterschiede zeigten sich jedoch hinsichtlich der Nutzung spezifischer Explorationsstrategien: Die männlichen Probanden nutzten von Anfang an häufiger die systematische Strategie als die Frauen dies taten. Rigide Explorationsweisen wurden dagegen bei Frauen häufiger beobachtet.

Die Ergebnisse der Studie demonstrieren zudem, dass die ungünstige Strategiesystematik nicht nur zu einer niedrigeren Leistung bei den Frauen führte, sondern auch das damit einhergehende Flow-Erleben beeinträchtigte. Aufgrund der vorliegenden Befunde kann geschlussfolgert werden, dass die niedrigeren Messwerte der Frauen in der aktuellen Motivation eine Folge der ungünstigen Strategiesystematik darstellen. Wie oben beschrieben entdeckten Frauen durch die Nutzung suboptimaler Strategien weniger häufig relevante Funktionen des Programms und erlebten infolgedessen weniger Erfolge. Diese Erfahrung sollte in erster Linie einen negativen Effekt auf die Einschätzung eigener Erfolgswahrscheinlichkeit sowie auf das Gefühl der Herausforderung haben. Die ungünstigere Einschätzung der Erfolgswahrscheinlichkeit eigener Handlungen und das Gefühl mangelnder Kontrolle über den Prozess (Stichwort: Flow-Erleben) kann die Frauen wiederum zusätzlich dazu veranlassen, weniger Funktionen auszuprobieren. Gleichzeitig verließen sie einmal eingeschlagene Wege im Programm nicht. Dies hatte den Nachteil, dass sie neue Aufgaben nicht bewältigen konnten. Zusammenfassend lässt sich bemerken, dass die Berücksichtigung moti-

vationaler Erlebensaspekte während des Prozesses zusätzliche Erklärungsbeiträge zu Leistungsdifferenzen der Geschlechter liefern konnte.

Methodenkritische Überlegungen. Als kritisch erwies sich insbesondere der geringe *Stichprobenumfang*, welcher sich über die einzelnen Messzeitpunkte hinweg zusätzlich reduzierte. Die Aufgabenbearbeitung wurde von einzelnen Teilnehmern frühzeitig, d.h. vor Ablauf der 30 Minuten, beendet und führte so zu einem teils beträchtlichen Datenverlust. Dieser Umstand lässt sich auf zwei Bedingungen zurückführen: Erstens sind die Aufgabenschwierigkeiten im Hinblick auf die Fähigkeiten einzelner leistungsstarker Teilnehmer als zu niedrig zu bewerten. Zweitens hat sich die Anzahl der Aufgaben für den gewählten Zeitraum als zu gering erwiesen. Eine potenzielle Lösung für dieses Problem ist das Ersetzen der Aufgaben durch mehr und anspruchsvollere Aufgaben. Alternativ könnten weitere Aufgaben hinzugefügt werden, so dass auch bei leistungsstarken Personen der Arbeitsprozess über die 30 Minuten hinweg beobachtet und Veränderungen analysiert werden können. Durch den geringen Stichprobenumfang waren komplexere Analysen zu Wirkungszusammenhängen der während des Arbeitens erhobenen Variablen eingeschränkt.

Der *Datenverlust zu den Messzeitpunkten t2 und t3* verursachte Schwierigkeiten bei der Interpretation der Ergebnisse vor allem im Hinblick auf das Flow-Erleben sowie die Höhe des motivationalen Zustands. Zum letzten Messzeitpunkt fehlten die Daten der leistungsstarken Personen, d. h. insbesondere diejenigen der männlichen Teilnehmer. Durchschnittswerte waren streng genommen mit den korrespondierenden Werten des Messzeitpunkts t1 nicht vergleichbar. Diesem Problem wurde bei der Analyse von Verläufen in der Varianzanalyse mit Messwiederholung Rechnung getragen. In der ANOVA mit Messwiederholung wurden aus diesem Grund lediglich Veränderungen der Motivation vor Beginn der Bearbeitung in Bezug auf den ersten Messzeitpunkt nach Bearbeitungsbeginn berücksichtigt. Zu diesen beiden Zeitpunkten konnte auf einen vollständigen Datensatz zurückgegriffen werden. Eine Veränderungsmessung über den gesamten Bearbeitungszeitraum war damit nicht möglich, so dass keine Aussagen über den weiteren Verlauf der aktuellen Motivation gemacht werden können.

Die *Bildung von Summenwerten für die aktuelle Motivation* führte dazu, dass Veränderungen in der Höhe einzelner Motivationsaspekte (z.B. Erfolgswahrscheinlichkeit und Herausforderung) nicht berücksichtigt werden konnten. Diese Einschränkung war weitgehend dem geringen Stichprobenumfang ($N = 33$) geschuldet; eine detailliertere Analyse komplexer Zusammenhänge ohne die Bildung eines Differenzwertes „Netto-Motivation“ hätte einer größeren Datenbasis bedurft.

Grenzen der Untersuchung. Eine wesentliche Einschränkung von Studie 2 ist, dass im Pfadmodell keine *Beziehungen zwischen den Prozessvariablen* (Flow, aktuelle Motivation und Strategien) modelliert wurden. Das empirisch überprüfte Pfadmodell wurde auf der Grundlage des gewählten

Versuchsdesigns konstruiert, wobei aufgrund der Beschränkungen nur das Flow-Erleben zum letzten Messzeitpunkt vor Bearbeitungsende berücksichtigt werden konnte. Diese zeitliche Abfolge der Messung einzelner Variablen führte zu dem Schluss, dass die Nutzung systematischer Strategien einen Einfluss auf das Flow-Erleben hat und sich beide Variablen auf die Leistung in SPSS auswirken. Wechselwirkungen der Variablen im Prozess wurden nicht berücksichtigt. Es lässt sich keine Aussage dazu treffen, ob das Flow-Erleben zu unterschiedlichen Zeitpunkten während des Arbeitens wiederum einen Beitrag zur Erklärung der Strategiesystematik leistete. Die längsschnittliche Auswertung der Daten macht jedoch einen größeren Stichprobenumfang und vorzugsweise vollständige Datensätze zu allen Messzeitpunkten erforderlich. Eine weitere Begrenzung stellt in der vorliegenden Studie die *Konzentration auf die Analyse nur einer einzigen Strategie* dar (Systematische Strategie). Die Beziehungen der suboptimalen Explorationsstrategien zu den Prozessvariablen und zur Leistung wurden nicht näher untersucht.

Aus zeitökonomischen Gründen, die sich aus der Laufzeit des Projekts ergaben, wurden alle Untersuchungsteilnehmer im Rahmen eines Semesters geworben und getestet. Daraus resultierte eine vergleichsweise geringe *Stichprobengröße* (vgl. Rost, 2007, S. 94). Wie bereits oben angemerkt wurde, schmälert die Art der Teilnehmerrekrutierung (Stichwort: Gelegenheitsstichprobe) die Verallgemeinerbarkeit der Ergebnisse. Des Weiteren kann eine kleine Stichprobe dazu führen, dass Effekte „übersehen“ werden, da sie nicht durch statistische Signifikanz abgesichert wurden. Wenn zwei unabhängige Gruppen auf Mittelwertsunterschiede geprüft werden, empfiehlt Rost (2007, S. 94; vgl. auch Stelzl, 2005) einen Mindestumfang von $n = 35$ Personen pro Gruppe. Diese Anzahl wurde in der vorliegenden Studie unterschritten. Dass Mittelwertsunterschiede dennoch die statistische Signifikanz erreichten, ist z.B. auf die Stärke des Effekts zurückführbar.

Die im Rahmen von Studie 3 getätigten Analysen dienen dazu, offene Fragen und Grenzen von Studie 2 zu adressieren und auf der Grundlage eines umfangreicheren Datensatzes untersuchen zu können. Studie 3 bezieht sich dabei insbesondere auf die wechselseitige Wirkung der Prozessvariablen Flow-Erleben und Strategiesystematik sowie auf die Nutzung unterschiedlicher, systematischer und unsystematischer Strategien im zeitlichen Verlauf.

9. Entwicklung von Motivation und Strategiesystematik im Prozessverlauf (Studie 3)

Studie 3 hatte zum Ziel, individuelle Veränderungen im Flow-Erleben, der aktuellen Motivation, der Strategiesystematik und deren Auswirkung auf das Erlernen eines Computerprogramms im zeitlichen Verlauf zu untersuchen (vgl. Forschungsfrage 3, Abschnitt 6.3). Zur Untersuchung dieser Frage wurde ein längsschnittliches Design gewählt. Darüber hinaus sollte an einem unabhängigen Datensatz getestet werden, ob sich die Befunde zu Geschlechtsunterschieden und zur Beziehung der Untersuchungsvariablen aus Studie 2 replizieren.

In Studie 2 war erwartet worden, dass der Effekt des Geschlechts auf die Leistung über Unterschiede in der Eingangsmotivation vermittelt wird. Die Eingangsmotivation sollte wiederum einen positiven Einfluss auf Flow-Erleben und die Strategiesystematik aufweisen. Der Effekt der Eingangsmotivation konnte in Studie 2 jedoch nicht empirisch untermauert werden: Es zeigten sich keine statistisch signifikanten Geschlechtsdifferenzen in der Höhe des Interesses, der Erfolgswahrscheinlichkeit, der Herausforderung und der Misserfolgsbefürchtung. Aus diesem Grund wurde das in Studie 2 postulierte Modell (vgl. Kapitel 8.2.7) modifiziert, in dem die Eingangsmotivation nicht in die Pfadanalyse einbezogen wurde. Stattdessen wurde ein direkter Effekt des Geschlechts auf die Leistung sowie auf das Flow-Erleben angenommen.

Um die Replizierbarkeit dieses (nicht erwarteten) Befunds zu testen, wurde in Studie 3 erneut ein Effekt des Geschlechts auf die Eingangsmotivation postuliert. Es wurden zudem erneut das Flow-Erleben sowie die systematische Explorationsstrategie als Mediatoren einbezogen, um die Modelle aus Studie 2 und Studie 3 miteinander vergleichen zu können. Demgegenüber wurde die aktuelle Motivation, die in Studie 3 ebenfalls erfasst wurde, nicht im Pfadmodell berücksichtigt. Sie wurde jedoch in einem eigenständigen Test über mehrere Messzeitpunkte hinweg verglichen. Die Beziehungsmuster der Variablen (Geschlecht, Prozessvariablen, Leistung) wurden analysiert, indem Daten aus drei Messzeitpunkten in Beziehung gesetzt wurden und deren Auswirkung auf die Leistung überprüft wurde.

Studie 3 hatte des Weiteren zum Ziel, Personen danach zu klassifizieren, wie häufig sie die unterschiedlichen Explorationsstrategien nutzen. Es wurde ein exploratorisches Vorgehen gewählt: Die konstituierenden Merkmale der Cluster wurden a posteriori datengeleitet bestimmt und nicht theoretisch hergeleitet. Zum Zeitpunkt der Untersuchung lagen noch keine theoretischen Annahmen oder empirischen Hinweise auf Muster in der Strategienutzung vor.

Aufgrund der Erfahrungen aus Studie 2 wurden in Studie 3 einzelne methodische Veränderungen vorgenommen. In Studie 2 hatte sich die hohe Zahl der *Drop outs* während der Untersuchung ungünstig auf die Erfassung insbesondere der Prozessvariablen ausgewirkt. Teilnehmer,

die bei der Lösung der Aufgaben besonders erfolgreich gewesen waren, hatten die Aufgabenbearbeitung vor dem letzten Messzeitpunkt beendet. Dies hatte dazu geführt, dass für diese Personen zu diesem Messzeitpunkt keine Daten zu Motivation und Strategiesystematik zur Verfügung standen. Um einen erneuten Datenverlust dieser Art zu vermeiden, wurde die Aufgabenstellung in Studie 3 leicht modifiziert.

Die Wahl der Stichprobe fußte auf den bereits für Studie 2 getätigten Überlegungen (siehe Kap. 8.1.1). Studie 3 beinhaltet im Vergleich zu den vorangegangenen Studien komplexere statistische Analysen (z. B. Pfadanalysen über mehrere Messzeitpunkte), die im Allgemeinen einen größeren Stichprobenumfang voraussetzen. Gleichzeitig lagen aus Studie 2 erste Hinweise auf die in der Population zu erwartenden Effektstärken zu Geschlechtsunterschieden in den Explorationsstrategien und Motivation vor. Vor diesem Hintergrund ergab sich die angestrebte Stichprobengröße für Studie 3 erstens aus den Erfordernissen der verwendeten statistischen Verfahren (z. B. den Strukturgleichungsanalysen auf der Grundlage von Maximum-Likelihood-Schätzungen). Zweitens wurde der optimale Stichprobenumfang auf der Basis der Effektgrößen der vorangegangenen Untersuchung (Studie 2) für die Folgestudie berechnet (vgl. Cohen, 1992; Rost, 2007, S. 219 ff.). Mit Hilfe des Programms G*Power 3.0 (Faul, Erdfelder, Lang & Buchner, 2007) wurde auf der Grundlage der Effektstärke des Mittelwertunterschieds von Männern und Frauen in der Leistung ($d = 1.19$, $\alpha_{\text{err prob}} = .05$, Teststärke = .80) eine minimale Stichprobengröße von $n = 17$ pro Gruppe für diese Mittelwertsunterschiede ermittelt. Ziel war es, diese minimal erforderliche Stichprobengröße in Studie 3 zu überschreiten. Der optimale Stichprobenumfang für das zu prüfende Pfadmodell leitete sich aus den in der Literatur zu Strukturgleichungsanalysen gegebenen Empfehlungen ab (z.B. Kline, 2005).

Zusammenfassend wurden in Studie 3 die folgenden Fragen erforscht (vgl. auch Abschnitt 6.3 und 6.4):

1. Beeinflussen sich die Mediatoren im Explorationsprozess gegenseitig? Wenn ja, welche Wirkmechanismen können als plausibel gelten?
2. Welchen individuellen Beitrag leisten die Prozessvariablen zur Aufklärung der Geschlechtsunterschiede in der Abhängigen Variablen Leistung? Hat das Flow-Erleben einen direkten und/oder indirekten Effekt auf die Leistung bei der Exploration des Statistikprogramms?
3. Verändert sich die Häufigkeit der Nutzung einer spezifischen Explorationsstrategie von Messzeitpunkt zu Messzeitpunkt oder bleibt der Einsatz der jeweiligen Strategie über die Zeit hinweg stabil?
4. Können anhand charakteristischer Muster der Explorationsstrategie-Nutzung Gruppen von Computernutzern sinnvoll voneinander abgegrenzt werden?

Zur besseren Nachvollziehbarkeit der in Studie 3 getesteten Annahmen sind die bereits in Abschnitt 6.3 und 6.4 dargestellten Hypothesen an dieser Stelle erneut aufgelistet:

Hypothese 10: Es wird erwartet, dass der Einfluss des Geschlechts auf die Leistung partiell über die Eingangsmotivation, die Häufigkeit der systematischen Explorationsstrategie sowie die Höhe des Flow-Erlebens zu mehreren Messzeitpunkten vermittelt wird. Die Häufigkeit der systematischen Explorationsstrategie wirkt sich günstig auf das Flow-Erleben aus. Der Effekt des Flow-Erlebens auf die Leistung wird während des Prozesses über die Strategiesystematik vermittelt. Am Ende des Lernprozesses trägt das Flow-Erleben neben der systematischen Explorationsstrategie mit einem inkrementellen Beitrag zur Erklärung der Leistung bei.

Hypothese 11: Personen lassen sich auf der Grundlage der Nutzungshäufigkeit einzelner Explorationsstrategien in Cluster klassifizieren.

9.1 Methode

9.1.1 Stichprobe

Die Stichprobe setzte sich aus Studierenden der Wirtschaftswissenschaften ($n = 80$) und der Psychologie ($n = 17$) zusammen. Insgesamt nahmen damit $N = 97$ Personen an der Untersuchung teil. Die Teilnehmer wurden auch in dieser Studie im Rahmen von verpflichtenden Statistikveranstaltungen in beiden Studiengängen geworben. Der Anteil der Frauen war im Fach Psychologie mit knapp 77% ($n = 13$) deutlich größer als unter den Wirtschaftswissenschaftsstudierenden mit 43% ($n = 32$) weiblichen Teilnehmern. Der Altersdurchschnitt lag in der Gesamtstichprobe bei 22.6 Jahren ($SD = 4.2$). Im Durchschnitt waren die Teilnehmer aktuell im zweiten Fachsemester eingeschrieben ($M = 2.5$, $SD = 2.4$). Die Anreize zur Teilnahme waren weitgehend identisch mit denjenigen in Studie 2 (s. Abschnitt 8.1.1). Den Psychologie-Studierenden wurden zusätzlich Versuchspersonenstunden gutgeschrieben. Darüber hinaus wurde ein Essen mit einer Führungskraft eines großen deutschen Kreditinstituts unter den Teilnehmern verlost. Dieser Anreiz sollte vor allem die Teilnahmebereitschaft in der Gruppe der Wirtschaftswissenschaftler erhöhen.

Wie bereits in Studie 2 (siehe Kap. 8.1.1) der Fall mussten auch in Studie 3 Datensätze von insgesamt fünf Personen ausgeschlossen werden (listenweiser Ausschluss). In einem Fall musste der Versuch aufgrund körperlichen Unwohlseins einer Teilnehmerin abgebrochen werden. Die Datensätze von vier weiteren Teilnehmern konnten ebenfalls nicht in der Untersuchung

berücksichtigt werden: In diesen Fällen traten während des Versuchs technische Inkompatibilitätsprobleme auf. Die Folge waren fehlerhafte Videoaufzeichnungen der Bildschirmoberfläche. Das unter diesen Umständen gewonnene Datenmaterial war somit lückenhaft und nicht vollständig auswertbar.

Darüber hinaus traten bereits während der Versuchsdurchführung technische Probleme auf, die von der Versuchsleiterin in Anwesenheit des Versuchsteilnehmers behoben werden mussten. Infolgedessen wurde der Bearbeitungsprozess unterbrochen, was die Vergleichbarkeit der Untersuchungsbedingungen im Sinne der Standardisierung erschwert. Die Störungen können darüber hinaus zu Beeinträchtigungen beim Explorieren des verwendeten Computerprogramms geführt haben. Der Datenausfall aufgrund der geschilderten Problematiken belief sich auf ca. 5%. Insgesamt konnten die Datensätze von $N = 92$ Personen, davon 45 weiblich und 47 männlich, in die Analysen einbezogen werden.

9.1.2 Vorgehen

Die Durchführung der Untersuchung entsprach der in Studie 2 gewählten Methode. Der Ablauf der Untersuchung ist in Abbildung 18 grafisch dargestellt.

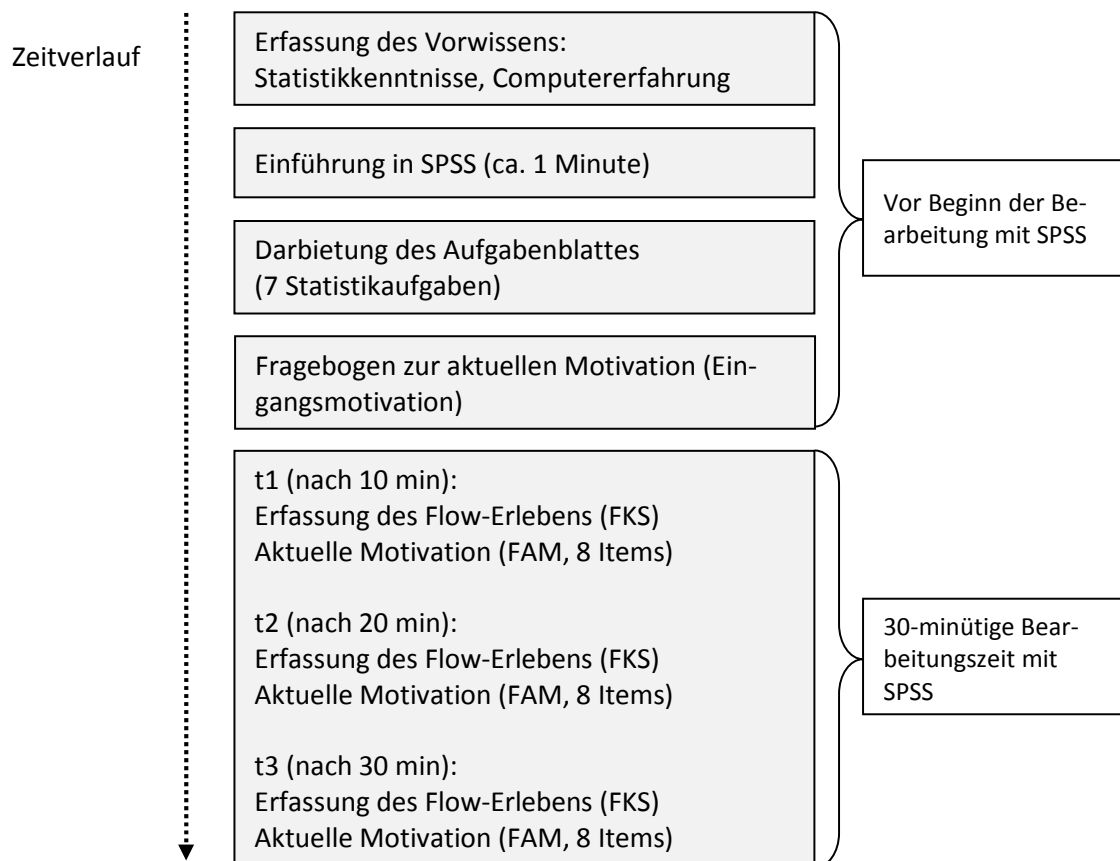


Abb. 18. Versuchsplan von Studie 3 im Zeitverlauf mit Angabe der erfassten Konstrukte.

Auch hier wurde ein quasi-experimentelles Design implementiert, bei dem zwei natürliche Gruppen (Frauen und Männer) miteinander verglichen wurden. Die Datenerhebung erfolgte im Wintersemester 2007/2008 sowie im Sommersemester 2008 an der Goethe-Universität in Frankfurt. Es wurde eine aktuelle Programmversion von SPSS (SPSS 17; SPSS Inc, 1993-2007) verwendet. Aufbauend auf Studie 2 wurde das Vorgehen folgendermaßen modifiziert: Das Arbeitsblatt umfasste sieben Aufgaben, von denen die Teilnehmer möglichst viele innerhalb von 30 Minuten erfolgreich bearbeiten sollten. Die Aufgabenanzahl wurde gegenüber den vorangegangenen Studien erhöht, um zu gewährleisten, dass die Teilnehmer den vollständigen Beobachtungszeitraum mit der Aufgabe befasst waren. Auf diese Weise konnte ihre aktuelle Motivation bzw. das Flow-Erleben dreimal im Prozess erfasst werden.

9.1.3 Materialien

Die eingesetzten Fragebögen sowie die verwendeten Materialien sind im Folgenden zusammenfassend beschrieben.

9.1.3.1 Fragebogen zu Statistikkenntnissen

In Anlehnung an die vorangehenden Studien wurden die Statistikkenntnisse über Selbsteinschätzungen sowie über einen Multiple-Choice-Test erfasst. Im ersten Teil des Fragebogens wurden wie bereits in Studie 1 und 2 die Anzahl der besuchten Statistik-Grundlagenveranstaltung(en) sowie die dazugehörigen Tutorien abgefragt. Die Teilnehmer gaben des Weiteren an, auf welche statistischen Inhaltsgebiete sie ihr Wissen anwenden konnten. Vor dem Statistikvorwissenstest stuften die Teilnehmer ihre Statistikkenntnisse auf der 5-stufigen Likertskala ein (vgl. Studie 1, Abschnitt 7.1.3.1 und 8.1.3.1).

Die sechs Multiple-Choice-Aufgaben dienten dazu, die Kenntnisse der Teilnehmer direkt zu erfassen. Pro Aufgabe konnten maximal vier Punkte erreicht werden. Die Verteilung der Werte in den einzelnen Aufgaben (s. Tabelle A - 22 im Anhang) wies eine Schiefe von -2.03 (Item 1) bis 0.62 (Item 3) auf. Der Exzess reichte von -.08 (Item 3) bis 3.65 (Item 1). Alle Trennschärfen lagen unter der empfohlenen Untergrenze von .40 (vgl. Kelava & Moosbrugger, 2007) und waren damit als niedrig einzustufen (siehe Tabelle A – 22 im Anhang). Die Höhe des Exzesses sowie der Schiefe in Item 1 deuteten nach Bühner (2006, S. 285) auf eine leichte Abweichung von der Normalverteilung hin. Der hohe Schwierigkeitsindex von Item 1 ($P = 86\%$) und die niedrige Trennschärfe ($r_{it} = .10$) ließen darauf schließen, dass das Item von der Mehrheit der Teilnehmer richtig beantwortet werden konnte und nicht zwischen den Teilnehmern differenzierte. Die Verteilung in dem Item konzentrierte sich demzufolge im Bereich hoher Punktwerte. Das Item stellte damit eine gelungene „Eisbrecher“-Frage dar (vgl. Mummendey, 1995, S. 67). Item 3 besaß

eine noch niedrigere Trennschärfe ($r_{it} = .01$). Der Schwierigkeitsindex der Items reichte von $P = 46\%$ bis $P = 86\%$ und liegt damit insgesamt im mittleren Bereich. Die interne Konsistenz der Skala betrug lediglich $\alpha = .37$. Um die Reliabilität der Skala zu erhöhen wurden Item 1 und Item 3 aus der Skala ausgeschlossen. Dies erhöhte die interne Konsistenz der vier verbleibenden Items auf $\alpha = .47$. Diese lag jedoch weiterhin unter der von Schermelleh-Engel und Werner (2007) empfohlenen Grenze. Insgesamt konnten auf der Grundlage der vier verbleibenden Items maximal 16 Punkte erreicht werden.

9.1.3.2 Fragebogen zur Computerbildung (INCOBI)

Zur Erfassung der Computerbildung wurde auf die bereits in Studie 1 verwendeten Skalen des Inventars zur Computerbildung (INCOBI; Richter et al., 1999) zurückgegriffen: Die 10 Items des VECA wurden erneut zur Kontrolle der Vertrautheit mit spezifischen Gruppen von Computerprogrammen genutzt (vgl. Studie 1, Abschnitt 7.1.3.2). Aus der Skala zur Erfassung des praktischen Computerwissens (PRACOWI) wurde Item 6 der Studie 1 (im Original: Item 2 des PRACOWI) eliminiert, da es einen sehr hohen Schwierigkeitsindex aufwies. Die verbleibenden fünf Items erreichten erneut eine niedrige interne Konsistenz von $\alpha = .42$. Wie bereits in Studie 1 erkennbar, wirkte sich die Kürzung des Instruments nachteilig auf die interne Konsistenz der Skala aus. Weitere Kennwerte der PRACOWI-Items sind Tabelle A - 23 in Anhang A 4.2.1 zu entnehmen.

Die psychometrischen Eigenschaften der Skala zur Vertrautheit mit Computeranwendungen (VECA) waren erneut zufriedenstellend (vgl. Tabelle A - 24 im Anhang). Die 10 Items der Skala erreichten eine interne Konsistenz von $\alpha = .74$. Angesichts der Heterogenität der Skala wurde diese als akzeptabel eingestuft. Die Variabilität der Itemschwierigkeiten in der Stichprobe war hoch. Die Kennwerte verteilten sich über die gesamte Bandbreite der gewählten Antwortskala (von $P = 8\%$ bis $P = 97\%$). Das bedeutet, dass sowohl leichte, mittlere sowie schwierige Items in der Skala vorhanden waren. Das Item zur Vertrautheit mit Statistikprogrammen wies wie erwartet den niedrigsten Schwierigkeitsindex auf: Nur wenige Teilnehmer hatten bereits Erfahrung mit dieser Gruppe von Programmen. Hingegen nutzten sowohl Personen mit hoher als auch niedriger Computererfahrung das Internet. Dies spiegelte sich in der niedrigen Trennschärfe des entsprechenden Items wider. Die Trennschärfen variierten in der Skala zwischen $r_{it} = .23$ (Item 8 „Internet“) und $.62$ (Item 1 „Textverarbeitungsprogramme“). Dieses Ergebnis stimmte mit den Angaben der Teilnehmer in Studie 1 weitgehend überein.

9.1.3.3 Fragebogen zur aktuellen Motivation (FAM)

Zur Messung der Motivation vor Beginn der Bearbeitung wurde erneut der Fragebogen zur aktuellen Motivation (FAM) von Rheinberg et al. (2001) eingesetzt.

Eingangsmotivation. An dieser Stelle werden zusammenfassend die psychometrischen Eigenschaften der vier Subskalen des Fragebogens (18 Items insgesamt) berichtet. Eine detaillierte Übersicht der Kennwerte findet sich in Tabelle A - 25 im Anhang. Die Items der Subskala Herausforderung erreichten mit $\alpha = .61$ die niedrigste interne Konsistenz verglichen mit den anderen drei Subskalen. Die höchste interne Konsistenz wies die Subskala Misserfolgsbefürchtung auf ($\alpha = .85$). Die Schwierigkeitsindizes lagen in den Subskalen Misserfolgsbefürchtung, Erfolgswahrscheinlichkeit und Interesse im mittleren Bereich (P zwischen 50 und 81%). In der Subskala Herausforderung konzentrierten sich die Werte demgegenüber im oberen Bereich ($P = 85\%$ bis 99%). Das bedeutet, dass die Items der Subskala Herausforderung in der Stichprobe besonders leicht im Sinne des Konstrukts zu beantworten waren. Die Trennschärfen variierten über alle Subskalen hinweg zwischen $r_{it} = .31$ und $.77$. Schiefe und Exzess der Items überschritten nicht die von Bühner (2006, S. 285) angeführten Grenze von 2.0 bzw. 7.0, so dass eine kritische Abweichung von der Normalverteilung in den Items unwahrscheinlich war. Die Befunde sind mit denjenigen in Studie 2 vergleichbar.

Aktuelle Motivation während der Bearbeitung. In Anlehnung an Studie 2 dienten acht ausgewählte Items des FAM zur On-line-Messung der aktuellen Motivation während des Explorationsprozess. Die acht Items, welche insgesamt vier Skalen abbildeten, wurden den Teilnehmern gemeinsam mit den FKS-Items dreimalig zur Einschätzung vorgelegt. Die Kennwerte der Items sind separat für die einzelnen Subskalen im Anhang in den Tabellen A - 29 bis A - 31 nachzulesen. Im Folgenden werden die wichtigsten item- und skalenanalytischen Ergebnisse zusammengefasst.

Die beiden Items der Subskala „Herausforderung“ erreichten konsistent über die drei Messzeitpunkte hinweg die höchsten Schwierigkeitsindizes ($P = 78\%$ bis 98%) verglichen mit den drei weiteren Subskalen. Das bedeutet, dass die Teilnehmer diesen Items besonders leicht zustimmen konnten und dass auch während der Bearbeitung das Leistungsmotiv der Teilnehmer angeregt wurde. Die Schwierigkeitsindices der restlichen Items konzentrierten sich im mittleren Bereich. Die Verteilungen in den acht Items wiesen hinsichtlich Schiefe und Exzess keine kritischen Werte (über 2.0 bzw. 7.0) auf. Auf die Berechnung von Cronbachs α und der Trennschärfe der Items wurde verzichtet, da sich die Subskalen lediglich aus je zwei Items zusammensetzten. Um die Komplexität der Analysen bzw. des zu überprüfenden Modells zu testen, wurde aus den acht Items ein Differenzwert (Netto-Motivationswert) berechnet (vgl. Abschnitt 8.1.3.2).

6.1.3.4 Flow-Kurz-Skala (FKS)

In Übereinstimmung mit Studie 1 und 2 wurde die Flow-Kurz-Skala (Rheinberg et al., 2003) zu drei Messzeitpunkten während der Bearbeitung eingesetzt. Da sich die beiden Subskalen Glatter Verlauf und Absorbiertheit zu allen drei Messzeitpunkten erneut als mittel bis stark korreliert erwiesen hatten ($r = .49$ bis $.73$), wurden die 10 Items der Subskalen zu einem Gesamtwert zusammengefasst. Die interne Konsistenz der Gesamtskala lag zwischen $\alpha = .84$ und $.92$. Wie bereits in Studie 1 und 2 ergaben sich für die Subskala „Glatter Verlauf“ insgesamt höhere interne Konsistenzen als für die Subskala „Absorbiertheit“. Demgegenüber wies die Subskala „Absorbiertheit“ konsistent höhere Schwierigkeitsindices auf. Eine Ausnahme bildete Item 10, das weniger leicht zu beantworten war (P zwischen 64% und 69%). Im Vergleich zu den vorangegangenen Studien erreichten die Schwierigkeitsindices in Studie 3 im Durchschnitt höhere Werte. Die Trennschärfen der Items lagen getrennt für die beiden Subskalen mehrheitlich im hohen Bereich über $r_{it} = .50$ (vgl. Bühner, 2006, S. 140). Schiefe und Exzess der Items überschritten nicht die kritischen Grenzen (vgl. Bühner, 2006, S. 285). Die entsprechenden Kennwerte für die beiden Subskalen sind den Tabellen A – 26 bis A - 28 im Anhang zu entnehmen.

9.1.3.5 Problemlöseaufgabe, Performanzmaß und Erfassung der Strategiesystematik

Problemlöseaufgabe. Das Vorgehen in Studie 3 entsprach dem Design, das für die anderen beiden Studien gewählt worden war. Die Problemlöseaufgabe zur Exploration von SPSS wurde jedoch leicht verändert: Die Anzahl der Aufgaben wurde gegenüber Studie 1 und 2 auf sieben erweitert. Damit sollte sichergestellt werden, dass alle Teilnehmer 30 Minuten mit dem Programm beschäftigt waren und zu drei Messzeitpunkten befragt werden konnten.

Die *Aufgaben 1, 2, 4* und *5* entsprachen den bereits in Studie 1 beschriebenen Aufgaben (vgl. Abschnitt 7.1.3.4). *Aufgabe 3* beinhaltete das Erstellen eines Streudiagramms der bivariaten Verteilung zweier Variablen. In *Aufgabe 7* mussten Fälle aus dem Datensatz ausgewählt werden, um gruppenspezifische Korrelationskoeffizienten berechnen zu können (s. Arbeitsblatt im Anhang zu dieser Arbeit in Abschnitt A 4.1.2 sowie Tabelle 20). *Aufgabe 3* und *5* galten als gelöst, wenn die entsprechende Grafik im Ausgabefenster des Computerprogramms erschien. Die Reihenfolge der Aufgaben war so gewählt, dass diese im Schwierigkeitsgrad zunahmen. Tabelle 20 gibt die Anforderungen der jeweiligen Aufgaben wieder. Im Anhang sind die Item- und Skalenstatistiken für diese Variable tabellarisch dargestellt (s. Tabellen A – 32 und A - 33 im Anhang).

Tabelle 20

Übersicht über spezifische Anforderungen der im Arbeitsblatt enthaltenen statistischen Aufgaben (Aufg.; N = 92)

Aufg.	Aufgaben auf dem Arbeitsblatt ¹	Max.	%
1	Angabe von <i>Mittelwert</i> und <i>Standardabweichung</i> für die Gesamtstichprobe, sowie getrennt für Frauen und Männer (N = 92)	6	100
2	Angabe des Korrelationskoeffizienten, des Signifikanzniveaus, sowie Interpretation des Ergebnisses (n = 68)	4	74
3	Erstellen eines Streudiagramms in SPSS (n = 64)	2	70
	0 Punkte: keine Grafik		
	1 Punkt: zu den Daten passende Grafik		
	2 Punkte: korrektes Streudiagramm		
4	Interpretieren eines Boxplots (Median Frauen und Männer, Ausreißer und Extremwerte, Verteilung) (n = 50)	4	54
5	Erstellen eines gruppierten Balkendiagramms (n = 40)	3	43
	0 Punkte: keine Grafik		
	1 Punkt: zu den Daten passende Grafik		
	2 Punkte: Balkendiagramm		
	3 Punkte: Gruppierendes Balkendiagramm		
6	Erstellen einer Häufigkeitstabelle (n = 17)	4	18
	1 Punkt: korrekte Tabelle		
	1 Punkt: korrekte Tabelle, korrekte Variable		
	2 Punkte: Angabe von absoluter und relativer Häufigkeit aus einer Häufigkeitstabelle		
7	Fälle auswählen: Angabe getrennter Korrelationskoeffizienten und p-Werte für Männer und Frauen und Interpretation des Ergebnisses (4 Kennwerte plus Interpretation; n = 9)	5	10

Anmerkung. Max. = mit der maximal zu erreichenden Punktzahl pro Aufgabe auf dem Arbeitsblatt; % = relativer Anteil der Personen, welche die jeweilige Aufgabe in Angriff genommen haben. ¹ Der genaue Wortlaut der Aufgabenstellung ist im Anhang in Abschnitt A 4.1.2 nachzulesen.

Leistungsindikator. Um die Performanz bei der Bearbeitung der Aufgaben zu ermitteln, wurden wie bereits in Studie 2 zwei Summenwerte gebildet: „Leistung_Arbeitsblatt“ und „Leistung_SPSS“. In Abschnitt 8.1.3.4 sind diese beiden Indikatoren näher beschrieben. In „Leistung_Arbeitsblatt“ konnten bei richtiger Lösung aller Aufgaben 28 Punkte erreicht werden. In der Skala „Leistung_SPSS“ betrug die maximal zu erzielende Punktzahl 14.

Die Tabellen A - 32 und A - 33 im Anhang geben Mittelwert, Standardabweichung, Trennschärfe, Schwierigkeit sowie Schiefe und Exzess der Verteilungen in den einzelnen Items wider. Es zeigte sich, dass die Aufgaben, die das korrekte Bedienen des Programms erforderten („Leistung_SPSS“), eine leicht höhere interne Konsistenz erzielten ($\alpha = .78$) als die Kombination aus korrektem Bedienen des Programms plus statistischer Auswertung der Ergebnisse („Leis-

tung_Aufgabenblatt“; $\alpha = .69$). Dies sprach dafür, dass die Aufgaben des erstgenannten Indikators mehr gemeinsame Varianz teilten und sich ähnlicher waren in ihren Anforderungen als diejenigen der Skala „Leistung_Arbeitsblatt“. Die Trennschärfe in Item 7 des Indikators „Leistung_Arbeitsblatt“ fiel mit $r_{it} = .14$ besonders niedrig aus. Dies konnte auf die niedrige Anzahl an Teilnehmern (10%) zurückgeführt werden, die die Aufgabe bearbeiteten. Item 1 wies ebenfalls eine niedrige Trennschärfe auf: Alle Teilnehmer hatten das Item bearbeitet und 79,3% erzielten dabei mindestens einen Punkt. Auf diese Weise differenzierte die Aufgabe nicht zwischen leistungsschwachen und leistungsstarken Personen.

Nach Inangriffnahmekorrektur variierten die Schwierigkeiten der Aufgaben auf dem Arbeitsblatt etwas stärker als in der Variablen „Leistung_SPSS“. Die Itemschwierigkeiten der Skala „Leistung_Arbeitsblatt“ lagen jedoch insgesamt im günstigen, mittleren Bereich (siehe Tabelle A - 32 im Anhang). Aufgrund der Aufgabenstellung wurden die Aufgaben 6 und 7 lediglich von $n = 17$ bzw. $n = 9$ Teilnehmern bearbeitet. Die Punkteverteilungen in diesen beiden Items erwiesen sich als linkssteil und schmalgipflig ($\text{Exzess}_{\text{Item 6 Arbeitsblatt}} = 6.01$; $\text{Exzess}_{\text{Item 7 Arbeitsblatt}} = 26.61$; $\text{Schiefe}_{\text{Item 6 Arbeitsblatt}} = 2.73$; $\text{Schiefe}_{\text{Item 7 Arbeitsblatt}} = 5.21$; weitere Kennwerte siehe Tabellen A - 32 und A - 33 im Anhang): Nur wenige Personen konnten die Anforderungen der Aufgaben 6 und 7 gut bewältigen.

Erfassung der Explorationsstrategien. Das Explorationsverhalten der Teilnehmer wurde erneut mit dem in Studie 1 entwickelten Beobachtungsinstrument erfasst. Die Aufzeichnung des Nutzerexplorationsverhaltens erfolgte mit der Screenscapturesoftware *ScreenVirtuoso*. Wie bereits in Studie 2 wurde die Klassifikation des interessierenden Explorationsverhaltens lediglich von einer Raterin vorgenommen. Interindividuelle Schwankungen in der Anzahl der Beobachtungseinheiten waren auf die bereits unter 7.1.3.5 angeführten Bedingungen zurückzuführen. Die Daten wurden nach dem in Studie 1 und 2 bereits beschriebenen Prinzip zu Summenwerten aggregiert. Um die Strategiesystematik über die Zeit zu erforschen, wurden in Studie 3 zusätzlich relative Häufigkeiten der Nutzung spezifischer Strategien für die drei Beobachtungszeiträume (bzw. Messzeitpunkte) gebildet.

9.1.4 Wahl und Diskussion der statistischen Analyseverfahren

Überprüfung komplexer Variablenzusammenhänge. Die theoretisch angenommenen Beziehungen der Variablen im Explorationsprozess wurden mittels Strukturgleichungsanalysen mit manifesten Variablen (Pfadanalyse) überprüft. Das hypothetische Modell wurde an einem Rohdatensatz mit *Mplus* 5.0 (Muthén & Muthén, 1998-2007) getestet. Weitere Hinweise zu den statistischen Analysen sind in den Abschnitten 8.1.4 sowie 8.2.7 nachzulesen.

Überprüfung von multivariaten Gruppenunterschieden. Mit Hilfe von varianzanalytischen Methoden (ANOVA/MANOVA) wurde überprüft, ob interindividuelle Unterschiede in den Abhängigen Variablen auf Ausprägungen in den Unabhängigen Variablen zurückführbar sind (vgl. Backhaus et al., 2006; Bortz, 2005). Dabei wurde zunächst die globale Nullhypothese getestet, dass keine Unterschiede in den Mittelwerten zwischen den Gruppen bestehen. Die multivariaten Signifikanztests der MANOVA erfolgten über das Prüfverfahren nach Hotelling-Spur (vgl. Bortz, 2005; Diehl & Staufenbiel, 2007). Um eventuelle statistisch signifikante Gruppendifferenzen lokalisieren und quantifizieren zu können, wurden im Anschluss an die globalen Signifikanzprüfungen, paarweise Einzelvergleiche mittels Post-Hoc-Tests (LSD) durchgeführt. Diese Tests berücksichtigen die drohende Kumulierung des α -Fehlers bei mehreren Signifikanztests auch bei Zutreffen partieller Nullhypothesen (Stelzl, 2005, S. 123 f.). So garantieren die nachträglichen Einzelvergleiche, „dass die Wahrscheinlichkeit eines α -Fehlers für jeden beliebigen, a-posteriori durchgeführten Einzelvergleich nicht größer ist als das Signifikanzniveau für den Overall-Test der Varianzanalyse“ (Bortz, 1993, S. 250).

9.1.5 Umgang mit fehlenden Werten

In Anlehnung an das Vorgehen in den Studien 1 und 2 wurde auch hier eine Imputation einzelner, zufällig fehlender Werte nach der Prozedur von Sijsma und van der Ark (2003; vgl. Abschnitt 7.1.4 dieser Arbeit) vorgenommen. Die Parameterschätzung des Strukturgleichungsmodells in Hypothese 10 erfolgte auf der Grundlage der *Full information maximum likelihood*(FIML)-Methode (auch *Direct ML*; vgl. Allison, 2003; Reinecke, 2005). Die FIML-Methode stellt ein modellbasiertes Verfahren zum Umgang mit fehlenden Werten dar, bei dem „eine Schätzung der Populationsparameter und ihrer Standardfehler auf der Basis der beobachteten Daten ... vorgenommen“ wird (Lüdtke et al., 2007, S. 112). Neben der multivariaten Normalverteilung der Daten nennt Reinecke (2005) als eine weitere wichtige Voraussetzung für die Anwendung der *Direct ML*-Methode die Art des Ausfallprozesses: „Eine direkte ML-Schätzung aus Daten mit fehlenden Werten setzt voraus, dass der Ausfallprozess ignoriert werden kann und die MAR [d. h. Missing At Random]-Annahme gilt“ (S. 290). Gelten die Voraussetzungen als erfüllt, dann bewertet Allison (2003) das *Direct ML*-Verfahren als „the best method for handling missing data for most SEM applications“ (S. 550).

Ob die Voraussetzungen für das FIML-Verfahren gegeben sind, wird im Folgenden diskutiert: In Studie 2 war ein Design bedingter, systematischer Ausfall von Messwerten zum Messzeitpunkt 3 festzustellen. Einzelne Personen beendeten die Aufgabenbearbeitung bereits vor dem dritten Messzeitpunkt, so dass die Höhe der aktuellen Motivation und des Flow-Erlebens nach Ablauf der 30 min nicht mehr erfasst werden konnten. Die Wahrscheinlichkeit des Datenausfalls

steht hier in Zusammenhang mit einer weiteren, beobachteten Variable, nämlich der Anzahl der vollständig bearbeiteten Aufgaben. Der Datenausfall in dieser Variablen ist jedoch von der Ausprägung in den kritischen Variablen (aktuelle Motivation/Flow-Erleben) unabhängig. Infolgedessen wurden in der vorliegenden Studie Datenausfälle dieser Art als MAR betrachtet (Reinecke, 2005, S. 286). Diese Annahme wird durch die Ausführungen von Little (1995, S. 1114) unterstützt, der systematische Ausfälle bei bestimmten Untersuchungsdesigns in längsschnittlichen Studien als „planned drop-out“ und damit als „vorwiegend zufällig“ einstuft. Die Voraussetzungen galten somit als erfüllt, so dass die Strukturgleichungsanalysen auf der Basis des FIML-Algorithmus durchgeführt wurden.

9.2 Ergebnisse

9.2.1 Vorbereitende Analysen zur empirischen Untersuchung

Der Datenauswertung in den Hypothesen (siehe Abschnitte 9.2.2 ff.) wurden auch in Studie 3 einige statistische Analysen vorgeschaltet. Erstens wurde erforscht, inwiefern sich die in Studie 2 empirisch belegten Geschlechtsunterschiede in zentralen Untersuchungsvariablen replizieren. Empirische Studien lassen darauf schließen, dass Geschlechtsunterschiede in unterschiedlichen Kohorten variieren können (z. B. Hyde, 2005, 2007). Insbesondere für die Nutzung studienrelevanter Computeranwendungen liefern neuere Studien Hinweise darauf, dass sich Differenzen zwischen Männern und Frauen über die Zeit hinweg verringert haben (z. B. Imhof et al., 2007). Demzufolge könnten die beiden Stichproben aus dem Jahr 2005 und aus den Jahren 2007/2008 aus unterschiedlichen Kohorten stammen. Vor diesem Hintergrund wurden multiple Gruppenvergleiche in den relevanten Variablen durchgeführt. Die Überprüfung der Mittelwertdifferenzen diente darüber hinaus dazu, die Vergleichbarkeit der beiden Stichproben zu testen (siehe Abschnitt 9.2.1.1).

Zweitens dienten die in Abschnitt 9.2.1.2 dargestellten Ergebnisse der Vorbereitung der Pfadanalysen. Dieser Teil der Arbeit bietet eine Übersicht über die Häufigkeitsverteilungen der einzelnen Explorationsstrategien separat für die drei Beobachtungsabschnitte der Studie 3. Anders als in den vorangegangenen Studien hatte Studie 3 zum Ziel, die Wechselwirkung der Prozessvariablen und deren Einfluss auf die Performanz im Rahmen eines komplexen Pfadmodells zu untersuchen. Die Entscheidung, welche Explorationsstrategie in das Pfadmodell einbezogen werden sollte, basierte auf der empirischen Relevanz der jeweiligen Strategie. Die Relevanz wurde über die relative Häufigkeit gemessen.

Drittens wurden zur Vorbereitung auf die multivariaten Analysen des Pfadmodells die bivariaten Zusammenhänge der relevanten Variablen untersucht. Hierzu wurden die Produkt-

Moment-Korrelationskoeffizienten über alle Variablen zu allen drei Messzeitpunkten berechnet. Die Ergebnisse dieser Analysen sind in Abschnitt 9.2.1.3 zusammengefasst.

Fehlende Werte und Missingskorrektur. Der Anteil der fehlenden Angaben bzw. Werte erreichte für die jeweilige Skala folgende Ausmaße: Vertrautheit mit Computeranwendungen (2; < 1%), Statistikvorwissen (3; < 1%), Eingangsmotivation (1; < 1%), Fragebogen zur aktuellen Motivation während der Bearbeitung plus Flow-Kurz-Skala (Messzeitpunkt 1: 9; < 1%; Messzeitpunkt 2: 9; < 1%; Messzeitpunkt 3: 10; < 1%). Die fehlenden Angaben verteilten sich dabei über verschiedene Personen. Insgesamt lag der Anteil an allen Werten der entsprechenden Skala stets deutlich unter 1%. Das Auftreten von Verteilungsverzerrungen in Folge der Imputation (vgl. Abschnitt 9.1.4) wurde damit als unwahrscheinlich bewertet. Zufällig fehlende Werte wurden bei den $N = 92$ Teilnehmern über die *Two-Way*-Imputation-Prozedur von Sjitsma und van der Ark (2003) ersetzt.

Zwei Teilnehmer beendeten die Bearbeitung des Aufgabenblattes vor Ablauf der 30-minütigen Frist, so dass die Motivationsmessung zum letzten Messzeitpunkt nicht durchgeführt werden konnte. Die fehlenden Werte dieser Teilnehmer wurden nicht ersetzt, da der Ausfall auf einem Datenerhebungsfehler basierte, der auf das gewählte Untersuchungsdesign zurückführbar war (vgl. Reinecke, 2005, S. 283 f.). In Strukturgleichungsanalysen wurde der Datenausfall berücksichtigt, in dem die Parameter unter Anwendung des FIML-Verfahren geschätzt wurden (siehe Abschnitt 9.1.5).

Statistische Drittvariablenkontrolle. 91% der Teilnehmer ($n = 84$) gaben an, regelmäßig die Statistik-I-Grundlagenveranstaltung ihres Fachbereichs zu besuchen. 87% ($n = 80$) nahmen zusätzlich an den entsprechenden Übungen im Tutorium teil. Im Vergleich zu Studie 2 hörte hier nur eine Minderheit von 13% ($n = 12$) überdies regelmäßig die Statistik-II-Veranstaltung. Etwa die Hälfte der Teilnehmer (45% bzw. $n = 45$) gab an, dass statistische Themen bereits im Schulunterricht behandelt worden seien. Wie bereits in Studie 2 konnten ca. 80% der Befragten ($n = 73$) ihre theoretischen Statistikkenntnisse nach eigener Einschätzung auf Aufgaben zur Deskriptiven Statistik und auf die grafische Darstellung der statistischen Auswertungen anwenden. Auf der 5-stufigen Skala zur Einschätzung der eigenen Statistikkenntnisse bewerteten die Teilnehmer ihre Statistikkenntnisse als durchschnittlich ($M = 2.85$, $SD = 0.88$). Die Teilnehmer in Studie 3 unterschieden sich bezüglich der Einschätzung ihrer eigenen Statistikkenntnisse demzufolge nicht statistisch signifikant von den Teilnehmern in Studie 2 ($t [123] = -0.18$, $p = .86$).

Statistische Voraussetzungsprüfungen. Tabelle A – 34 (siehe Abschnitt A 4.2.2) im Anhang gibt eine Übersicht über Mittelwerte, Standardabweichung, Schiefe und Exzess der Skalenmittelwerte der zentralen Untersuchungsvariablen. Darüber hinaus ist das Ergebnis des Kolmogorov-Smirnov-Tests auf Abweichung von der Normalverteilung aufgeführt. Statistisch signi-

signifikante Abweichungen ergaben sich für die Misserfolgsbefürchtung als Teil der Eingangsmotivation ($Z = 0.01, p = .04$) sowie für die Häufigkeiten in den beiden Explorationsstrategien „Versuch-und-Irrtum“ ($Z = 2.04, p = .01$) und „Informationssuche“ ($Z = 2.06, p = .01$). Bei der Misserfolgsbefürchtung handelte es sich um eine zweigipflige Verteilung: Ein Teil der Werte konzentrierte sich im unteren Bereich, ein zweiter Teil im oberen Bereich. Die Häufigkeitsverteilungen in den beiden Explorationsstrategien erwiesen sich als stark linkssteil. Das bedeutet, dass viele Personen diese (suboptimalen) Explorationsstrategien nicht oder selten genutzt hatten. Wenn Schiefe und Exzess der Variablen zur Bewertung hinzugezogen wurden, erreichte jedoch nur die Variable „Versuch-und-Irrtum“ kritische Werte (Schiefe = 3.08, Exzess = 14.4). Für alle anderen Variablen galten die Voraussetzungen als erfüllt.

9.2.1.1 Replizierbarkeit von Geschlechtsunterschieden

Studie 2 ergab statistisch signifikante Geschlechtsdifferenzen in der Leistung: Frauen erreichten eine statistisch signifikant niedrigere Punktzahl bei der Aufgabenlösung im Vergleich zu den Männern. Die Leistungsunterschiede konnten auf Unterschiede in Motivation und Strategiesystematik zurückgeführt werden. Frauen wiesen zudem systematisch ungünstigere Messwerte in der Computerbildung auf.

Die vergleichenden Analysen wurden nicht auf der Basis von normativen, sondern auf der Basis von z -standardisierten Werten durchgeführt (Fischer & Milfont, 2010). Bei der z -Standardisierung wird die „individuelle Leistung [oder der individuelle Messwert; Anm. der Verf.] zuvor an der Gesamtleistung [bzw. am Niveau der Messwerte; Anm. der Verf.] des Kollektivs relativiert“ (Bortz, 1993, S. 45). Eine z -Transformation wird dann empfohlen, wenn Messwerte von Personen aus unterschiedlichen Kollektiven miteinander verglichen werden (Diehl & Kohr, 1994). Diese Empfehlung ist auf Personen aus verschiedenen Universitätsjahrgängen anwendbar. Darüber hinaus war die Vergleichbarkeit der normativen Werte zwischen den beiden Stichproben eingeschränkt, da der Messung in Studie 2 und Studie 3 zum Teil unterschiedliche Instrumente zugrunde gelegt wurden. Um die Verteilungen in den beiden Stichproben (Studie 2 und Studie 3) miteinander vergleichen zu können, wurden die Messwerte folglich innerhalb der jeweiligen Stichprobe z -standardisiert.

In Tabelle 21 wurden die Messwerte von Männern und Frauen in den zentralen Untersuchungsvariablen der beiden Stichproben gegenübergestellt. Die in der Tabelle aufgeführten normativen Messwerte sollen dem Leser lediglich einen Überblick über die Ausprägungen in den zentralen Untersuchungsvariablen bieten.

9. Studie 3: Entwicklung von Motivation und Strategiesystematik im Prozessverlauf

Tabelle 21

Mittelwerte (*M*) und Standardabweichung (*SD*) in den zentralen Untersuchungsvariablen auf der Basis normativer, d.h. nicht χ -standardisierter, Werte getrennt nach Stichprobe und Geschlecht. Kursiv gedruckte Items bzw. Skalen weisen auf leicht unterschiedliche Operationalisierungen in den beiden Stichproben hin

Item/Skala	Studie 2		Studie 3	
	Männlich (<i>n</i> = 18)	Weiblich (<i>n</i> = 15)	Männlich (<i>n</i> = 47)	Weiblich (<i>n</i> = 45)
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
Vorwissen/Vorerfahrung				
<i>Statistikprogramme</i> ¹	0.11 (0.32)	.00 (0.24)	0.38 (0.64)	0.24 (0.53)
<i>Tabellenkalkulationsprogramme</i> ¹	1.67 (0.91)	1.07 (1.39)	1.87 (1.08)	1.22 (0.97)
<i>Datenbanken</i> ¹	0.50 (0.79)	0.27 (0.46)	1.26 (1.09)	1.31 (1.29)
Statistikkenntnisse (Selbstbeurt.)	3.00 (0.84)	2.73 (0.80)	3.04 (0.93)	2.64 (0.77)
<i>Statistikvorwissenstest</i> ²	2.64 (0.72)	2.32 (0.46)	2.78 (0.51)	2.33 (0.53)
Eingangsmotivation				
Interesse	4.46 (0.95)	4.25 (0.95)	4.34 (0.98)	4.36 (1.08)
Erfolgswahrscheinlichkeit	4.51 (0.88)	3.94 (0.98)	4.06 (1.04)	3.79 (1.23)
Herausforderung	4.93 (0.96)	5.18 (0.99)	5.24 (0.84)	5.54 (0.92)
Misserfolgsbefürchtung	3.26 (1.20)	3.65 (1.07)	3.49 (1.35)	3.84 (1.36)
Explorationsstrategien ³				
Systematische Exploration	45% (12)	35% (11)	41% (12)	41% (12)
Versuch-und-Irrtum	6% (5)	3% (3)	4% (6)	4% (4)
Rigide Exploration	15% (15)	28% (17)	21% (14)	22% (15)
Informationssuche	3% (6)	8% (8)	6% (8)	5% (7)
Unterbrechung	32% (13)	26% (15)	28% (10)	28% (11)
Netto-Motivation (letzter MZP)	11.00 (3.63)	8.20 (4.43)	10.45 (3.80)	8.99 (4.57)
Flow-Erleben (letzter MZP)	4.53 (0.93)	4.07 (1.15)	4.60 (1.64)	4.39 (1.43)
<i>Leistung_Arbeitsblatt</i> ⁴	10.50 (3.47)	5.90 (4.22)	11.57 (6.59)	9.53 (7.16)

Anmerkungen. ¹ In Studie 2 wurde die Erfahrung im Umgang mit dieser Programmfamilie über die Häufigkeit bisheriger Nutzung mittels folgender Kategorien erfasst: 0 = nie, 1 = bis 10mal, 2 = bis 20-mal, 3 = mehr als 20-mal. In Studie 3 wurde mittels des INCOBI (Skala VECA) die Vertrautheit mit den angegebenen Programmfamilien im Vergleich zu anderen Studierenden mittels der folgenden Antwortkategorien erfragt: 0 = weit unterdurchschnittlich, 1 = unterdurchschnittlich, 2 = durchschnittlich, 3 = überdurchschnittlich, 4 = weit überdurchschnittlich. ² Angegeben sind die im Durchschnitt erreichten Punkte auf Itemebene mit maximal 4 Punkten. Dabei betrug die maximale Gesamtpunktzahl der 6 Items des Statistikvorwissenstests in Studie 2 betrug die 22 Punkte. In Studie 3 konnten bei 4 Aufgaben maximal 16 Punkte erreicht werden. ³ Häufigkeiten über alle Messzeitpunkte hinweg. ⁴ In Studie 2 konnten maximal 17 Punkte auf alle 4 Aufgaben des Arbeitsblattes erreicht werden. In Studie 3 betrug die Maximalpunktzahl bei den 7 Aufgaben 28 Punkte.

Auf der Grundlage der χ -Werte konnte überprüft werden, ob die Verteilungen der Messwerte für Männer und Frauen in den beiden Stichproben voneinander abweichen. Die χ -Transformation beinhaltet jedoch den Nachteil, dass keine Aussagen darüber gemacht werden können, ob die beiden Gesamtstichproben sich im Mittel in den interessierenden Variablen vo-

neinander unterscheiden. Dieser Nachteil ist jedoch für die Überprüfung der Geschlechtsunterschiede zwischen den beiden Studien nicht relevant und wurde deshalb in Kauf genommen.

Folglich wurden in allen durchgeführten Analysen Haupteffekte des Faktors Geschlecht auf statistische Signifikanz getestet. Darüber hinaus wurde die partielle Effektstärke des Interaktionsterms der beiden Faktoren (Geschlecht, Stichprobe) untersucht. Der Einfluss des Faktors Stichprobe alleine wurde nicht untersucht, da die Mittelwerte in beiden Stichproben infolge der z-Standardisierung Null betragen. Ein signifikanter Haupteffekt des Geschlechts würde dafür sprechen, dass das Geschlecht über beide Stichproben hinweg einen bedeutsamen Einfluss auf die Ausprägung in den Abhängigen Variablen hat. Ein signifikanter Interaktionseffekt wiese darauf hin, dass sich mögliche Geschlechtsunterschiede nicht einheitlich über die Stichproben verteilen, sondern dass sich die Größe des Geschlechtseffekts zwischen den Stichproben unterscheidet.

Replikation von Leistungsunterschieden. In einem ersten Schritt wurde mittels *t*-Test für unabhängige Stichproben getestet, ob sich Männer und Frauen der Studie 3 statistisch signifikant in der Leistung bei der Aufgabenbearbeitung mit SPSS (Leistung_Arbeitsblatt) unterscheiden. Frauen erreichten (nach z -Standardisierung) einen Mittelwert von $M = -0.15$ ($SD = 1.0$), Männer hingegen von $M = 0.14$ ($SD = 0.95$). Der Signifikanztest ergab keinen statistischen Unterschied zwischen den Geschlechtern hinsichtlich der Leistung ($t [90] = 0.16$, $p = .16$, $d = 0.30$). Jedoch erreichten Männer etwas höhere Werte als Frauen. Die Effektstärke deutet auf einen geringen Effekt des Geschlechts auf die Leistung hin. Trotz ähnlicher Tendenzen weicht dieser Befund deutlich von den in Studie 2 gefundenen Effektstärken ab. In dieser Stichprobe hatte das Geschlecht mit $d = 1.20$ einen starken Effekt auf die Leistung (Cohen, 1992). Die Geschlechtsunterschiede variierten damit systematisch zwischen den Stichproben, insofern, dass sich in Studie 3 ein deutlich geringerer Effekt zeigte als in Studie 2.

Multivariate Varianzanalysen zum Einfluss von Geschlecht und Stichprobe. Um die unterschiedlichen Effektstärken in der Leistung zwischen Studie 2 und Studie 3 genauer eruieren zu können, wurde erforscht, in welchen (weiteren) Variablen sich Männer und Frauen der Studie 2 und Studie 3 unterschieden. Hierzu wurden Mittelwertvergleiche mit Hilfe von univariaten und multivariaten Varianzanalysen (MANOVA) mit den zwei Faktoren Geschlecht (männlich, weiblich) und Stichprobe (Studie 2, Studie 3) über die z -standardisierten Variablen vorgenommen. Aus der Kombination der Faktorstufen ergaben sich vier Gruppen, deren Mittelwerte miteinander verglichen wurden.

In die MANOVA wurden unterschiedliche Abhängige Variablen, z. B. die Leistung, das Flow-Erleben, die Eingangsmotivation und die Strategiesystematik, einbezogen. Um die Anzahl der durchgeführten Varianzanalysen zu reduzieren, wurden die Abhängigen Variablen auf der Grundlage theoretischer und praktischer Überlegungen gruppiert. Zum Beispiel wurden alle Va-

riablen, die zur Erfassung des Vorwissens dienen (Computererfahrung, Statistikvorwissen), simultan in eine MANOVA aufgenommen. Dabei gingen nur solche Variablen in die Berechnungen ein, die in beiden Stichproben (Studie 2, Studie 3) mit vergleichbaren Messinstrumenten erfasst worden waren. Die Prüfung der Unterschiede erfolgte zunächst allgemein über den Overall-Test der MANOVA bzw. mittels einer ANOVA. Im Falle signifikanter Haupteffekte bzw. Interaktionen wird im Folgenden zusätzlich berichtet, in welchen Abhängigen Variablen die Differenzen lokalisiert waren.

Im Anschluss an den Globaltest wurde überprüft, in welchen der Variablen sich Männer und Frauen der Studien 2 und 3 statistisch signifikant unterschieden. Um eine Alpha-Fehler-Inflation zu vermeiden, wurde zuvor eine Bonferroni-Korrektur vorgenommen (Stelzl, 2005, S. 123). Hierzu wurde das konventionelle Signifikanzniveau von $p = .05$ durch die Anzahl der Einzeltests geteilt (siehe Abschnitt 8.2.2). Beispielsweise führte das bei fünf Einzeltests zu einem Signifikanzniveau von $p = .01$. Im Anschluss an die Einzeltests wurde mittels Post-Hoc-Tests überprüft, zwischen welchen der vier Gruppen statistisch signifikante Unterschiede bestanden und gegebenenfalls welche Richtung die Unterschiede zeigten.

Leistung. Mittels einer univariaten Varianzanalyse wurden zu Beginn der Analyse Unterschiede in der Variablen „Leistung_Arbeitsblatt“ in den vier Gruppen (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3) getestet. Der Overall-Test ergab signifikante Unterschiede für den Faktor Geschlecht ($F [1, 121] = 11.57, p < .01, \eta^2 = .09$). Die Interaktion zwischen Geschlecht und Stichprobe verfehlte knapp die Signifikanzgrenze ($F [1, 121] = 3.61, p = .06, \eta^2 = .03$).

Wie Abbildung 19 illustriert, unterschieden sich die Männer in Studie 2 signifikant sowohl von den weiblichen Probanden in Studie 2 als auch von denjenigen in Studie 3. Auch in Studie 3 erreichten die männlichen Teilnehmer leicht bessere Leistungswerte als die Frauen, jedoch zeigte sich hier zwischen den Geschlechtern kein signifikanter Unterschied (wie bereits im t -Test, s.o.). Demgegenüber erzielten die Männer der Studie 3 signifikant höhere Leistungen als die Frauen in Studie 2. Die Leistung zwischen den beiden Gruppen mit männlichen Teilnehmern in Studie 2 und Studie 3 differierte nicht statistisch signifikant. Die niedrige Leistung der Frauen in Studie 2 führte dazu, dass die Männer insgesamt statistisch signifikant bessere Leistungen erzielten als die Frauen.

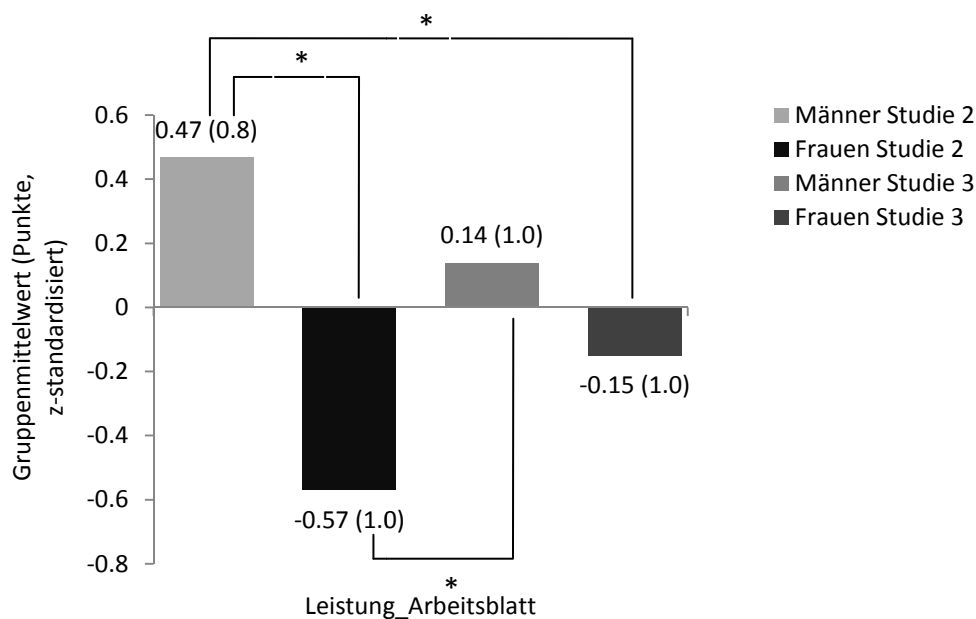


Abb. 19. Mittelwerte (und Standardabweichungen) für den Indikator „Leistung_Arbeitsblatt“ getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3). Die Werte wurden innerhalb der Gruppen zuvor z-standardisiert. Signifikante Gruppenunterschiede sind mit einem Asterisk (*) gekennzeichnet.

Computererfahrung und Statistikvorwissen. Zunächst wurde das Vorwissen (Statistikenkenntnisse, Computererfahrung) in den vier Gruppen hinsichtlich der Nutzung von Tabellenkalkulations-, Statistikprogrammen sowie von Datenbanken verglichen. Die Selbstbeurteilung der Statistikenkenntnisse sowie das Ergebnis des Statistikvorwissenstests wurden zusätzlich in die Analysen einbezogen. Kritisch muss dabei angemerkt werden, dass die Operationalisierung zwischen den Studien leicht differierte (vgl. Abschnitte 8.1.3.1, 9.1.3.1, 9.1.3.2).

Tabelle 22

Ergebnisse der multivariaten Varianzanalyse (MANOVA) des Faktors Geschlecht für verschiedene Indikatoren der Computererfahrung und für das Statistikvorwissen in beiden Stichproben ($N = 125$)¹

Abhängige Variable	F	P	η^2
Nutzungshäufigkeit Statistikprogramme	2.92	.09	.02
Nutzungshäufigkeit Tabellenkalkulation	11.74	.01	.09
Nutzungshäufigkeit Datenbanken	0.57	.45	.01
Selbstbeurteilung Statistikenkenntnisse	3.78	.05	.03
Statistikvorwissenstest	7.83	.01	.06

Anmerkungen. p = Signifikanzniveau; η^2 = partielle Effektstärke; Freiheitsgraden: $df_1 = 1$, $df_2 = 121$. ¹Ergebnis des Box-M-Tests auf Gleichheit der Kovarianzmatrizen: $F(30, 9985) = 1.01$, $p = .45$. ²Der Levene-Test ergab signifikante Abweichungen von der Gleichheit der Fehlervarianzen in den Gruppen für die beiden Variablen „Häufigkeit Statistikprogramme“ ($F[3, 121] = 7.90, p < .01$) und „Häufigkeit Tabellenkalkulation“ ($F[3, 121] = 4.12, p < .01$).

Die MANOVA ergab einen signifikanten globalen Haupteffekt für den Faktor Geschlecht auf alle Vorwissensvariablen ($F [5, 117] = 4.16, p < .01, \eta^2 = .15$). Die Interaktion aus Stichprobe und Geschlecht konnte demgegenüber nicht zur Varianzaufklärung in den Abhängigen Variablen beitragen ($F [5, 117] = 0.23, p = .95$). Auf der Grundlage dieses Befunds sind ausschließlich die Ergebnisse der fünf Gruppenvergleiche für den Faktor Geschlecht in Tabelle 22 dargestellt. Wie die Tabelle zeigt, ging der globale Haupteffekt des Geschlechts nach Bonferroni-Korrektur auf statistisch signifikante Unterschiede in der Nutzungshäufigkeit von Tabellenkalkulationsprogrammen ($F [1, 121] = 11.74, p = .01, \eta^2 = .09$) sowie im Statistikvorwissen ($F [1, 121] = 7.83, p = .01, \eta^2 = .06$) zurück.

Um zu überprüfen, wie stark sich die Geschlechter in der Nutzung von Tabellenkalkulation und im Statistikvorwissenstest unterschieden, wurden im Anschluss t -Tests getrennt für die beiden Stichproben über die Werte in der Variablen Tabellenkalkulation und Statistikvorwissen durchgeführt. Das Ergebnis des Levene-Tests legte bedeutsame Gruppenunterschiede für die Variable „Nutzungshäufigkeit von Tabellenkalkulation“ hinsichtlich deren Fehlervarianzen nahe (siehe Tabelle 22). Vor dem Hintergrund dieser Voraussetzungsverletzung wurden zusätzlich non-parametrische Einzelvergleiche mittels Mann-Whitney-U-Tests vorgenommen (Bortz, 2005). Die Ergebnisse wichen jedoch nicht von denjenigen der parametrischen Tests ab. Im Folgenden werden deshalb ausschließlich die parametrischen Testergebnisse (t -Tests) berichtet.

In Abbildung 20 sind die Ergebnisse für die vier Gruppen gegenübergestellt. In Studie 3 zeigten sich signifikante Unterschiede zwischen Männern ($M = 0.30, SD = 1$) und Frauen ($M = -0.31, SD = .91$) in der Nutzung von Tabellenkalkulationsprogrammen ($t [90] = 3, p < .01$; Cohens $d = .64$) sowie im Statistikvorwissen ($t [90] = 2.98, p < .01$; Männer: $M = 0.29, SD = 0.89$; Frauen: $M = -0.30, SD = 1.03$; Cohens $d = .61$). Frauen wiesen dabei eine geringere Vorerfahrung auf bzw. erreichten weniger Punkte im Statistikvorwissenstest. In Studie 2 zeigten sich ebenfalls signifikante Geschlechtsunterschiede im Vorwissen zugunsten der Männer. Die Unterschiede wurden jedoch nur im Bereich der Nutzung von Tabellenkalkulationsprogrammen statistisch signifikant. Der Effekt hatte eine mittlere Stärke ($t [31] = 2.20, p < .05$; Männer: $M = 0.33, SD = 1.1$; Frauen: $M = -0.40, SD = 0.72$, Cohens $d = 0.79$; bzw. Statistikvorwissen: $t [31] = 1.47, p = .15$; Männer: $M = 0.23, SD = 1.1$; Frauen: $M = -0.27, SD = 0.74$; Cohens $d = 0.53$).¹⁷

¹⁷ Leichte Unterschiede der in Kapitel 4 und Kapitel 6 berichteten Effektstärken (Cohens d) sind Rundungsergebnissen geschuldet. In Kapitel 4 wurden die Effektstärken für Unterschiede von normativen, d.h. nicht ipsativierten Werten berechnet. In Kapitel wurden die z -Werte aus zwei Stichproben miteinander verglichen.

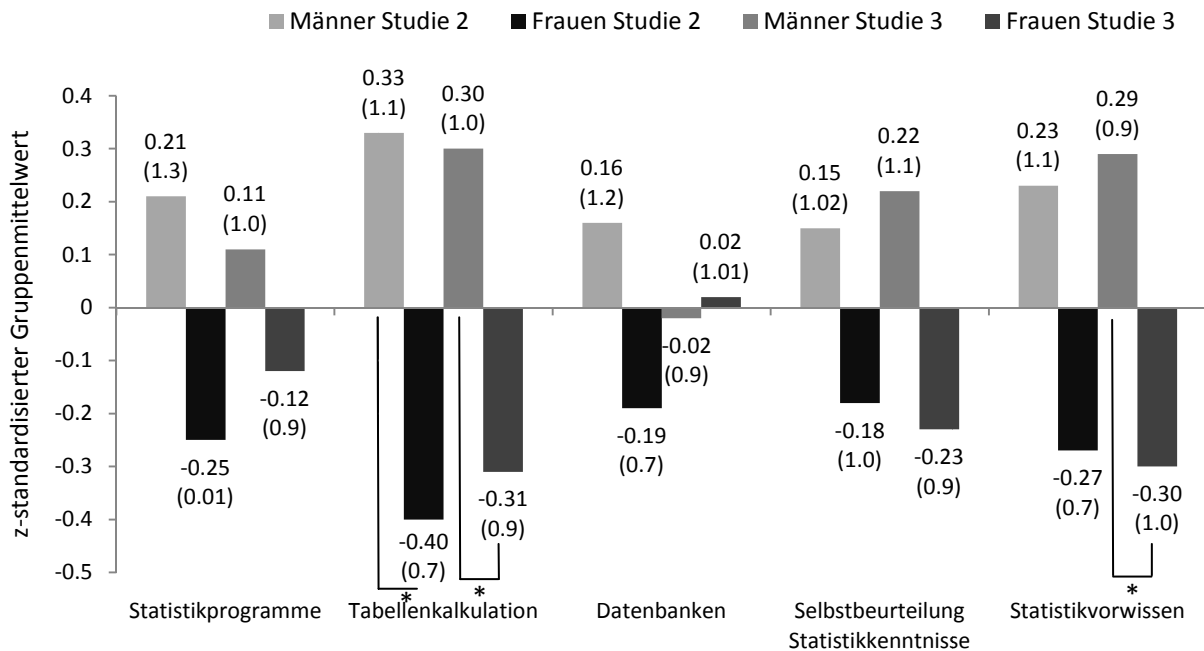


Abb. 20. Mittelwerte (und Standardabweichungen) für verschiedene Indikatoren der Computererfahrung und für das Statistikvorwissen getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3). Die Werte wurden innerhalb der Gruppen zuvor z-standardisiert. Signifikante Gruppenunterschiede sind mit einem Asterisk (*) gekennzeichnet.

Es stellte sich die Frage, ob die in Studie 2 gefundenen geschlechtsspezifischen Leistungseinbußen beim Lösen der Statistikaufgaben durch Unterschiede im Vorwissen mitbedingt sein können. Die Höhe des Statistikvorwissens korrelierte in beiden Stichproben niedrig und nicht signifikant mit der Leistung (Studie 2: $r = .24, p = .18, N = 33$; Studie 3: $r = .16, p = .13, N = 92$, vgl. Studie 2, Abschnitt 8.2.1). Demgegenüber zeigten sich in der Stichprobe aus Studie 3 signifikante, positive Beziehungen zwischen der Häufigkeit der Nutzung von Tabellenkalkulationsprogrammen und der Leistung ($r = .26, p < .01, N = 92$). Demgegenüber wurde der Korrelationskoeffizient in Studie 2 nicht signifikant und der Effekt fiel geringer aus ($r = .16, p = .13, N = 33$). Diese beiden Korrelationskoeffizienten ($r = .26$ bzw. $r = .16$) wurden nach Fisher- χ -Transformation miteinander verglichen: Die inferenzstatistische Überprüfung ergab, dass der Korrelationskoeffizient in Studie 2 nicht signifikant von demjenigen in Studie 3 abwich ($\chi = 0.04, p = .69$). Das Ergebnis kann so interpretiert werden, dass in beiden Stichproben eine geringfügige, positive Beziehung zwischen den beiden Variablen bestand, die jedoch aufgrund der Stichprobengröße lediglich in Studie 3 statistisch signifikant wurde. Die größere Vorerfahrung der Männer im Umgang mit Tabellenkalkulationsprogrammen könnte sich beim Erlernen des Statistikprogramms förderlich auf die Leistung der Männer ausgewirkt haben.

Eingangsmotivation. Im nächsten Schritt wurde überprüft, ob Männer und Frauen in beiden Stichproben hinsichtlich der Höhe ihrer Eingangsmotivation vergleichbar sind. Abbildung 20 stellt die Mittelwerte der vier Subskalen der Eingangsmotivation getrennt für vier Untersuchungsgruppen dar.

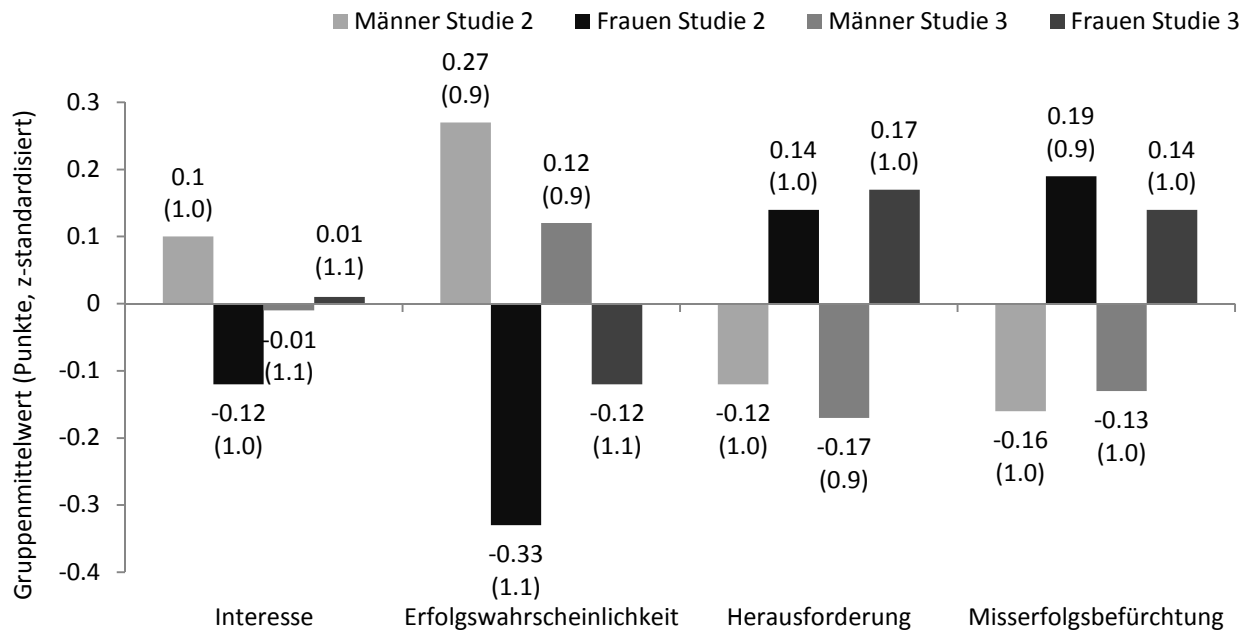


Abb. 21. Mittelwerte (und Standardabweichungen) für die vier Subskalen der Eingangsmotivation getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3). Die Werte wurden innerhalb der Gruppen z-standardisiert. Signifikante Gruppenunterschiede sind mit einem Asterisk (*) gekennzeichnet.

Im Overall-Tests zeigte sich weder ein signifikanter Haupteffekt des Geschlechts (Geschlecht: $F [4, 118] = 1.71, p = .15$) noch der Interaktion der Faktoren Geschlecht und Stichprobe auf die Motivation vor Beginn der Bearbeitung ($F [4, 118] = 0.22, p = .93$). Die vier Gruppen wichen demnach nicht systematisch in den vier Aspekte der Eingangsmotivation voneinander ab. Folglich können Geschlechtsdifferenzen in der Leistung über die Stichproben hinweg nicht über diese Variable erklärt werden. Die Deskriptiven Statistiken gibt Abbildung 21 wieder.

Explorationsstrategien. Im nächsten Schritt wurde überprüft, ob sich zwischen den Stichproben und den Geschlechtern Unterschiede im Verhalten während des Bearbeitungsprozesses zeigten. Alle vier Gruppen wurden hinsichtlich der Häufigkeiten der Explorationsstrategien sowie der Kategorie Unterbrechung miteinander verglichen.

Die Overall-Test der MANOVA ergab, dass das Geschlecht über die Stichproben hinweg einen statistisch signifikanten Haupteffekt auf die Nutzung der Explorationsstrategien hatte ($F [5, 117] = 3.55, p = .01$). Da a priori keine Hypothesen zur Richtung (insbesondere der suboptima-

len) Explorationsstrategien vorlagen, wurde erneut eine Bonferroni-Korrektur des Signifikanzniveaus vorgenommen. Bei fünf Einzeltests im Anschluss an die MANOVA wurde ein adjustiertes Signifikanzniveau von $p = .01$ für den Effekt in den einzelnen Strategien bzw. der Kategorie Unterbrechung festgelegt. Tabelle 23 fasst die Ergebnisse der Analysen zusammen.

Tabelle 23

Ergebnisse der multivariaten Varianzanalyse (MANOVA) für die Nutzung der Explorationsstrategien mit F-Tests für den Faktor Geschlecht sowie den Interaktionsterm aus Geschlecht und Stichprobe (N = 125) ^{1 2}

Quelle	AV	MANOVA		
		F	p	η^2 <i>partiell</i>
Geschlecht				
	Systematische Exploration	4.42	.04	.04
	Versuch-und-Irrtum	2.10	.15	.02
	Rigide Exploration	4.22	.04	.03
	Informationssuche	2.04	.16	.02
	Unterbrechung	1.04	.31	.01
Interaktion Geschlecht x Stichprobe				
	Systematische Exploration	3.96	.04	.03
	Versuch-und-Irrtum	0.90	.34	.01
	Rigide Exploration	3.05	.08	.03
	Informationssuche	3.21	.08	.03
	Unterbrechung	1.36	.25	.01

Anmerkungen. $df_1 = 1$, $df_2 = 121$; η^2 = partielle Effektstärke. ¹ Ergebnis des Box-M-Tests auf Gleichheit der Kovarianzmatrizen: $F(30, 10132) = 34.11$, $p = .40$. ² Der Levene-Test ergab für keine der Abhängigen Variablen signifikante Abweichungen von der Gleichheit der Fehlervarianzen.

Die Effekte der einzelnen Explorationsstrategien wurden nicht statistisch signifikant: Alle empirischen p -Werte überschritten den zuvor festgesetzten kritischen p -Wert (s. Tabelle 23). Danach hat das Geschlecht insgesamt einen Einfluss auf die Kombination der Nutzung aller Strategien, wirkt sich jedoch nicht statistisch signifikant auf die Nutzung spezifischer Strategien aus. Ein ähnliches Ergebnis der MANOVA zeigte sich für den Interaktionseffekt von Geschlecht und Stichprobe: Die Nutzung aller Explorationsstrategien variierte auch mit der Interaktion von Geschlecht und Stichprobe (Overall-Test: $F[5, 117] = 3.07$, $p = .01$). Wie Tabelle 23 zeigt, lässt sich das Ergebnis insbesondere auf Gruppendifferenzen im Ausmaß systematischer Exploration zurückführen. Anders als bei der simultanen Überprüfung des Einflusses der Interaktion von Geschlecht und Stichprobe ergaben sich bei der univariaten Testung im Anschluss (einfaktorielle ANOVA) keine statistisch signifikanten Interaktionseffekte auf die Strategiesystematik für die Variablen.

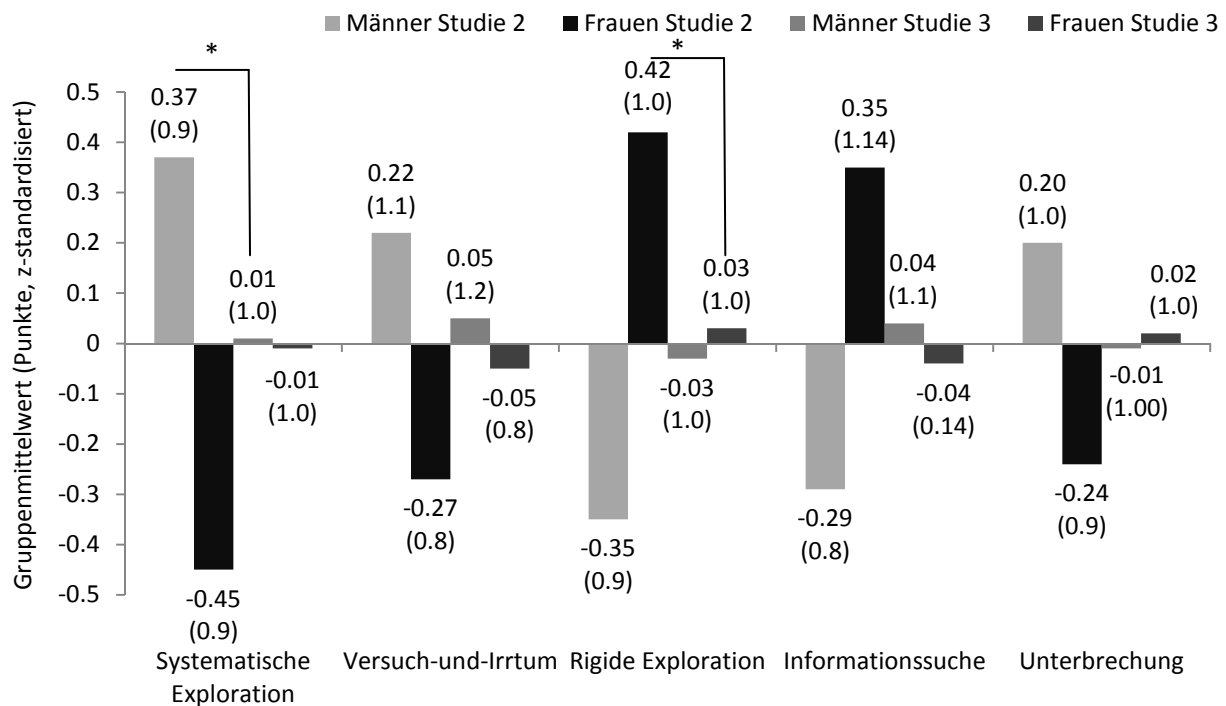


Abb. 22. Mittlere relative Häufigkeiten (und Standardabweichungen) für die vier Explorationsstrategien getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3). Die Werte wurden innerhalb der Gruppen z-standardisiert.

In Abbildung 22 sind die Gruppenmittelwerte in den einzelnen Variablen noch einmal veranschaulicht. Trotz der fehlenden statistischen Signifikanz in den Einzeltests lassen sich die Ergebnisse der MANOVA der Tendenz nach folgendermaßen beschreiben: Frauen der im Jahre 2005 erhobenen Stichprobe (Studie 2) nutzten weniger häufig systematische und häufiger unsystematische Explorationsstrategien als Männer in Studie 2. In der Stichprobe, welche im Jahr 2007/2008 erhoben worden war, nutzten Männer und Frauen die Explorationsstrategien in vergleichbarem Umfang. Dieses Ergebnis verdeutlicht inhaltlich den signifikanten Interaktionseffekt. Der Interaktionseffekt von Geschlecht und Stichprobe auf die Strategiesystematik wirkte sich ausschließlich auf Unterschiede in der Kombination von relevanten Explorationsstrategien aus und nicht auf die Nutzung einzelner Explorationsstrategien. Wurden die Geschlechter über die Stichproben hinweg jedoch zusammengefasst (Haupteffekt Geschlecht), nutzten Frauen im Vergleich zu Männern tendenziell stärker unsystematische Strategien und weniger die systematische Explorationsstrategie.

Aktuelle Motivation und Flow. Mittelwertunterschiede in Flow-Erleben und aktueller Motivation (während der Bearbeitung) wurden lediglich für den letzten Messzeitpunkt vor dem individuellen Bearbeitungsende überprüft. Als Indikator für die aktuelle Motivation diente die „Netto-Motivation“, d. h. der Differenzwert der acht aus dem FAM ausgewählten Items (siehe

Kapitel 8.1.3.2 und 9.1.3.3). Dabei wurden die Werte der sechs ausgewählten Items in den FAM-Subskalen Herausforderung, Interesse und Erfolgswahrscheinlichkeit addiert und die Werte in den beiden Items zur Misserfolgsbefürchtung subtrahiert. Ähnlich wie bereits bei der Variablen Flow wurde auch für die aktuelle Motivation eine zusätzliche Variable „Mot_End“ gebildet. Sie fasst die Netto-Motivation der Teilnehmer zum letzten Zeitpunkt vor dem individuellen Bearbeitungsende zusammen. Gruppenunterschiede wurden mittels einer MANOVA geprüft. Die Ergebnisse der anschließenden Einzeltests sind in Tabelle 24 dargestellt.

Tabelle 24

Ergebnisse der multivariaten Varianzanalyse (MANOVA) für die aktuelle Motivation sowie das Flow-Erleben (jeweils zum letzten Messzeitpunkt: Flow_end, Mot_end) sowie die Faktoren Geschlecht und Stichprobe inklusive der partiellen Effektstärke (η^2)^{1 2}

Quelle	AV	MANOVA		
		F	p	η^2 partiell
Geschlecht				
	Netto-Motivation	8.78	.01	.07
	Flow-Erleben	2.33	.13	.02
Geschlecht x Stichprobe				
	Netto-Motivation	1.39	.24	.01
	Flow-Erleben	0.69	.41	.01

Anmerkungen. $df_1 = 1$, $df_2 = 119$, η^2 = partielle Effektstärke. ¹ Ergebnis des Box-M-Tests auf Gleichheit der Kovarianzmatrizen: $F(9, 22268) = 18.04$, $p = .04$: Die Kovarianzmatrizen der Abhängigen Variablen über die Gruppen sind nicht gleich. ² Der Levene-Test ergab für keine der Abhängigen Variablen signifikante Abweichungen von der Gleichheit der Fehlervarianzen.

Das Geschlecht hatte über die Stichproben hinweg einen statistisch signifikanten Overall-Effekt auf die motivationalen Variablen ($F[2, 118] = 4.51$, $p = .01$): Dieser globale Effekt fußte jedoch allein auf Geschlechtsdifferenzen in der aktuellen Motivation zum letzten Messzeitpunkt ($F[1, 119] = 8.78$, $p = .01$; kritisches Signifikanzniveau bei 2 Tests nach Bonferroni-Korrektur: $p = .025$). Frauen hatten danach über die Stichproben hinweg eine niedrigere Netto-Motivation zum letzten Messzeitpunkt der Aufgabenbearbeitung als Männer. Ein Haupteffekt des Geschlechts auf das Flow-Erleben ließ sich dagegen nicht belegen ($F[1, 119] = 2.33$, $p = .13$). Der Interaktionseffekt aus Stichprobe und Geschlecht wurde für die Netto-Motivation und das Flow-Erleben nicht statistisch signifikant ($F[2, 118] = 0.69$, $p = .50$).

Um die Geschlechtsunterschiede in der aktuellen Motivation in den Stichproben zu lokalisieren, wurde zusätzlich eine Univariate Varianzanalyse für die Variable „Mot_end“ über die vier Gruppen berechnet. Im Folgenden werden ausschließlich die Ergebnisse der Einzelvergleiche im Rahmen der ANOVA (LSD) dargestellt, da nur diese als Ergänzung der zuvor berechneten MA-

NOVA von Interesse waren. Zusammenfassend zeigten die Post-hoc-Tests der ANOVA, dass in Studie 3 keine statistisch signifikanten Geschlechtsunterschiede in der aktuellen Motivation vorlagen ($\text{Diff} = 0.35$, $SE = 0.2$, $p = 0.9$). Die Mittelwertsunterschiede gingen jedoch in die erwartete Richtung. Abbildung 23 gibt die Mittelwerte und Standardabweichungen für die motivationalen Variablen „Netto-Motivation“ und „Flow-Erleben“ für die vier Gruppen an.

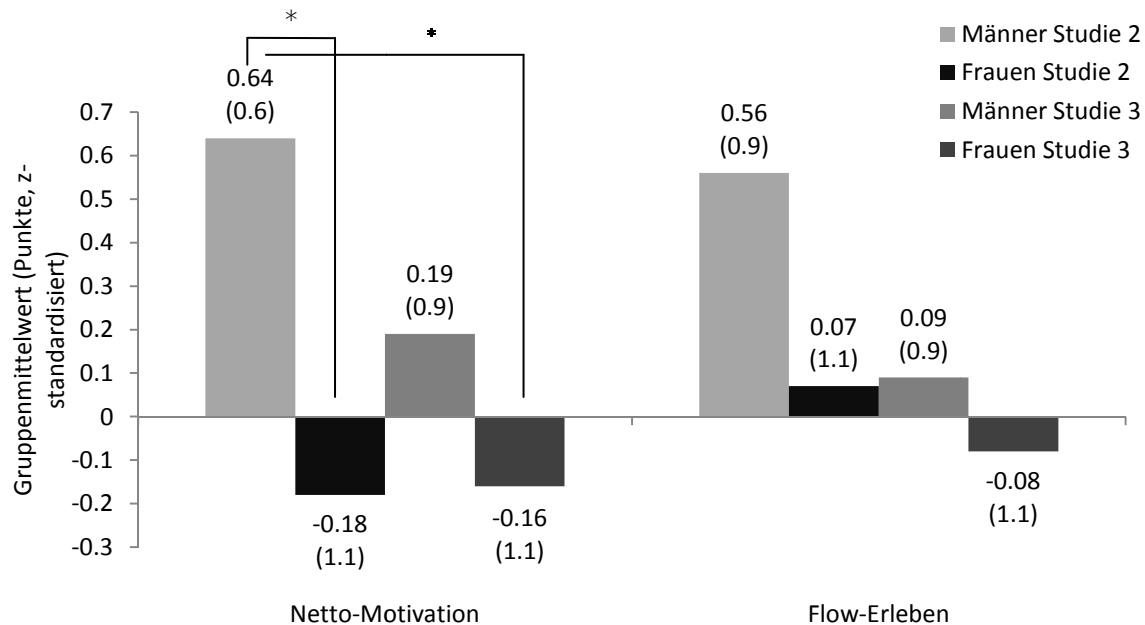


Abb. 23. Mittelwerte (und Standardabweichungen) in der Netto-Motivation sowie dem Flow-Erleben (jeweils zum letzten Messzeitpunkt) getrennt nach den vier Gruppen der MANOVA (Männer Stichprobe 2, Frauen Stichprobe 2, Männer Stichprobe 3, Frauen Stichprobe 3). Die Werte wurden innerhalb der Gruppen z-standardisiert. Signifikante Gruppenunterschiede sind mit einem Asterisk (*) gekennzeichnet.

Demgegenüber unterschieden sich Männer und Frauen der Studie 2 signifikant in ihrer Netto-Motivation ($\text{Diff} = -0.81$, $SE = 0.36$, $p = .017$): Frauen wiesen hier eine deutlich niedrigere Netto-Motivation ($M = -0.18$, $SD = 1.1$) auf als ihre männlichen Kommilitonen ($M = 0.64$, $SD = 0.6$). Das bedeutet, dass die Männer aus Studie 2 sowohl im Vergleich zu Frauen der Studie 2 höhere Werte der aktuellen Motivation erreichten als auch im Vergleich zu Frauen der Stichprobe 3 ($\text{Diff} = 0.80$, $SE = 0.27$, $p = .01$). Die Ergebnisse lassen darauf schließen, dass sich Frauen und Männer in Stichprobe 2 und 3 folgendermaßen unterschieden: Waren die Männer in Studie 2 noch deutlich motivierter als die Frauen, zeigte sich diese starke Differenz in Stichprobe 3 so nicht mehr. Hier waren Männer und Frauen in etwa gleich motiviert. Die Netto-Motivation der Männer in Studie 3 näherte sich dabei dem Niveau der Netto-Motivation der Frauen an.

Die aktuelle Motivation zum letzten Messzeitpunkt stand sowohl in Studie 2 als auch in Studie 3 signifikant mit der Leistung in Verbindung ($r_{\text{Studie 2}} = .56$, $p = .001$ bzw. $r_{\text{Studie 3}} = .48$, $p = .001$). Dementsprechend gingen ungünstigere Werte in der aktuellen (Netto-)Motivation bei

Frauen der Stichprobe 2 mit einer geringeren Leistung einher. In Studie 3 erzielten die Frauen in der aktuellen Motivation keine niedrigeren Werte mehr als die Männer, so dass die Höhe der aktuellen Motivation nicht zu Geschlechtsunterschieden in der Leistung beitrug.

Zusammenfassung. Der in Studie 2 nachgewiesene starke Geschlechtseffekt auf die Leistung replizierte sich in Studie 3 nicht. Hier zeigten sich lediglich schwache Effekte und die Leistung der Frauen war beim Explorieren des Computerprogramms nicht mehr statistisch signifikant geringer als diejenige der Männer. In Studie 3 näherte sich das Performanzniveau von Männern und Frauen insofern an, als dass Frauen eine Verbesserung in der Leistung zeigten. Mit Hilfe mehrerer weitere Analyseschritte wurde erforscht, worauf die Effektdifferenzen des Geschlechts in Bezug auf die Leistung zurückgeführt werden können. Ziel war es, Variablen zu identifizieren, die es erlauben, die Stichproben voneinander zu unterscheiden. Sollten sich zum Beispiel Frauen in Studie 2 und 3 in wesentlichen Prädiktoren der Leistung unterscheiden, könnte sich dies auch auf geschlechtsspezifische Leistungsdifferenzen auswirken.

In den Prozessvariablen wurden Stichprobenunterschiede innerhalb der Geschlechter hinsichtlich der aktuellen Motivation und der Strategiesystematik deutlich: Frauen nutzten in Studie 3 deutlich häufiger die systematische Strategie als ihre weiblichen Kommilitonen in Studie 2. Gleichzeitig reduzierte sich der Anteil dieser Explorationsstrategie bei den Männern der Studie 3 leicht. Demgegenüber lag der Anteil der rigiden Explorationsstrategie bei den Männern in der dritten Studie höher als in der zweiten. Bezogen auf die Motivation fühlten sich auch in die Studie 3 die Frauen tendenziell etwas mehr herausgefordert als die Männer. Die (schwachen) Unterschiede in der Eingangsmotivation wurden jedoch nicht statistisch signifikant. Darüber hinaus egalisierte sich das Niveau der aktuellen (Netto-)Motivation bei Männern und Frauen in Studie 3 weitgehend. In Studie 3 zeigte sich in der aktuellen Motivation nur noch ein geringer Geschlechtseffekt ($d = 0.35$ im Vergleich zu $d = 0.93$ in Studie 2), der nicht statistisch signifikant wurde. Bezüglich der Strategiesystematik reduzierte sich die Effektstärke des Geschlechts sogar auf $d = 0.01$ in Studie 3 (im Vergleich zu $d = -0.87$ in Studie 2). Beide Variablen (aktuelle Motivation und systematische Strategie) stellten in Studie 2 bedeutsame Prädiktoren der Leistung dar.

9.2.1.2 Häufigkeiten der Explorationsstrategien

In dem in Hypothese 11 postulierten Pfadmodell wurden verschiedene Zeitpunkte zur Erfassung der Strategiesystematik berücksichtigt. Im Rahmen der explorativen Datenanalyse wurde überprüft, ob die Nutzungshäufigkeit einzelner Explorationsstrategien zwischen den drei Messzeitpunkten (erstes, zweites und drittes Intervall von 10 Minuten) systematisch variierte. Dabei sollte erforscht werden, ob und welche Explorationsstrategien in bestimmten Arbeitsphasen bei den Teilnehmern besonders häufig zu beobachten waren.

Es wurden fünf separate Varianzanalysen mit Messwiederholung getrennt für die fünf Kategorien des Beobachtungssystems und die drei Messzeitpunkte durchgeführt. In den Varianzanalysen wurde der Effekt des Messzeitpunkts auf die relative Häufigkeit einer der vier Explorationsstrategien bzw. der relativen Häufigkeit in der Kategorie Unterbrechung erforscht. Als Kovariate wurde das praktische Computerwissen in die Analysen einbezogen, um den Einfluss der Vorkenntnisse zu kontrollieren. Das praktische Computerwissen korrelierte beispielsweise positiv mit der Häufigkeit der systematischen Explorationsstrategie (aggregiert über die Messzeitpunkte; $r = .25, p = .02$; vgl. Kang & Yoon, 2008). Dabei sollte untersucht werden, ob die Wechselwirkung aus praktischem Computerwissen und Messzeitpunkt einen zusätzlichen Einfluss auf das Explorationsverhalten ausübt. Vermutet wurde, dass das praktische Computerwissen insbesondere zu Beginn des Explorationsprozesses einen positiven Effekt auf die Strategiesystematik ausübt. Personen, die ein größeres Wissen besaßen, sollten dies auf das neue Programm transferieren können.

Tabelle 25

Ergebnisse der fünf Varianzanalysen mit Messwiederholung für den Faktor Messzeitpunkt sowie jeweils für die relative Nutzungshäufigkeit der einzelnen Explorationsstrategien sowie der Kategorie Unterbrechung zu den drei Messzeitpunkten¹

Abhängige Variable	MZP 1	MZP 2	MZP 3	ANOVA ²				
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>F</i>	<i>p</i>	<i>df₁</i>	<i>df₂</i>	η^2 <small>partiell</small>
Systematische Exploration	55 (12)	33 (16)	34 (19)	36.54	.01	2	178	.29
Versuch-und-Irrtum	3 (5)	4 (7)	5 (8)	2.66	.08	1.8	158	.03
Rigide Exploration	19 (15)	21 (19)	24 (20)	7.33	.01	2	178	.08
Informationssuche	8 (12)	5 (9)	5 (9)	0.33	.70	1.8	163	.01
Unterbrechung	16 (15)	32 (16)	32 (15)	15.99	.01	2	178	.15

Anmerkungen. MZP = Messzeitpunkt; *M* = Mittlere relative Nutzungshäufigkeit; *SD* = Standardabweichung der mittleren relativen Nutzungshäufigkeit. ¹ Zur Berechnung der mittleren prozentualen Häufigkeit wurde zunächst die Häufigkeit der jeweiligen Strategie an der Häufigkeit aller Strategien eines ausgewählten Zeitabschnitts für jede Person relativiert. Über alle Personen hinweg wurde dann für jeden Zeitpunkt ein mittlerer Häufigkeitswert pro Strategie berechnet. ² Ergebnisse der F-Tests unter Annahme der Sphärität, außer für die Abhängigen Variablen Versuch-und-Irrtum und Informationssuche. Hier wurde die Korrektur nach Greenhouse-Geisser vorgenommen.

Tabelle 25 gibt die mittlere relative Häufigkeit der einzelnen Explorationsstrategie sowie der Kategorie Unterbrechung getrennt für die drei Messzeitpunkte wieder. Darüber hinaus werden die Ergebnisse der fünf Varianzanalysen für den Faktor Messzeitpunkt (3 Messzeitpunkte) und die Abhängige Variable (Häufigkeit Explorationsstrategie bzw. Kategorie Unterbrechung) berichtet. Den Ergebnissen der Varianzanalysen zufolge hatte der Messzeitpunkt nach Bonferoni-Korrektur einen statistisch signifikanten Haupteffekt auf die Nutzungshäufigkeit der systematischen sowie der rigiden Explorationsstrategie ($F [2, 178] = 36.54, p = .01, \eta^2 = .29$ bzw. $F [2, 178] = 7.33, p = .01, \eta^2 = .08$). Darüber hinaus variierte die Häufigkeit von Unterbrechungen

während des Explorationsprozesses systematisch mit dem Messzeitpunkt ($F [2, 178] = 15.99, p = .01, \eta^2 = .15$). In Abbildung 24 sind die Verläufe der mittleren prozentualen Häufigkeiten noch einmal grafisch dargestellt.

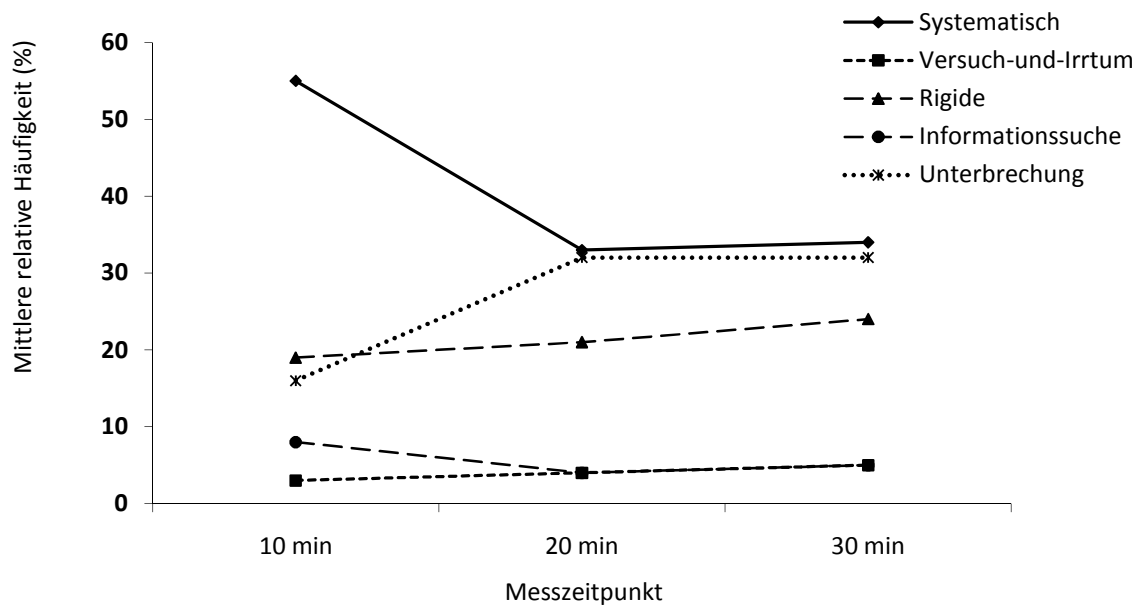


Abb. 24. Mittlere relative Häufigkeit (%) der vier Explorationsstrategien sowie der Kategorie Unterbrechung zu den drei Messzeitpunkten in der Gesamtstichprobe ($N > 90$).

Post-Hoc Vergleiche zwischen einzelnen Messzeitpunkten. Um die Unterschiede zwischen den Messzeitpunkten genauer spezifizieren zu können, wurden für die drei Abhängigen Variablen (systematische Explorationsstrategie, rigide Explorationsstrategie, Kategorie Unterbrechung) paarweise Post-hoc-Vergleiche mit Bonferroni-Korrektur. Die post-hoc Vergleiche ergaben, dass sich die Häufigkeit der *systematischen Explorationsstrategie* zu Messzeitpunkt 1 statistisch signifikant von derjenigen zu den Messzeitpunkten 2 (Diff = 4.26, $p = .01$) und 3 (Diff = 4.07, $p = .01$) unterschied. Die Häufigkeit der systematischen Exploration nahm zum zweiten Messzeitpunkt im Vergleich zum ersten Messzeitpunkt statistisch signifikant ab und blieb dann vom zweiten bis dritten Messzeitpunkt stabil (Diff = -0.19, $p = 1.0$). Die Häufigkeit in der Variablen *Rigide Exploration* stieg demgegenüber zwischen den Messzeitpunkten 1 und 3 signifikant an (Diff = 1.13, $p = .02$). Der Unterschied zwischen den Messzeitpunkten 1 und 2 sowie 2 und 3 wurde dabei nicht signifikant (Diff = -0.41, $p = .78$ bzw. Diff = -0.72, $p = .19$). Die Häufigkeit nahm demzufolge über die Messzeitpunkte hinweg kontinuierlich leicht zu. In der Variablen *Unterbrechung* unterschieden sich die Häufigkeiten zu Messzeitpunkt 1 systematisch von denen zu Messzeitpunkt 2 (Diff = -3.16, $p = .01$) und zu Messzeitpunkt 3 (-3.31, $p = .01$). Dagegen zeigten sich keine systematischen Unterschiede zwischen den beiden letzten Messzeitpunkten 2 und 3 (Diff =

0.15, $p = 1.00$). Zum ersten Messzeitpunkt war die Zahl der Unterbrechungen sehr niedrig. Danach erhöhte sich der Anteil der Unterbrechungen an allen Kategorien vom ersten zum zweiten Messzeitpunkt und stabilisierte sich dann auf diesem Niveau.

Interaktionseffekt von Vorwissen und Messzeitpunkt. Im folgenden Abschnitt sind lediglich die Ergebnisse für die systematische und die rigide Explorationsstrategie aufgeführt. Diese beiden Variablen standen mit zentralen Untersuchungsvariablen in bedeutsamem Zusammenhang (z.B. aktuelle Motivation, Leistung). Das praktische Computerwissen übte in Interaktion mit dem Messzeitpunkt einen statistisch signifikanten Einfluss auf die relative Häufigkeit der *systematischen Strategie* aus ($F [2, 31] = 4.05, p = .019, \eta^2 = .04$). Im Laufe der Aufgabenbearbeitung verstärkte sich der Zusammenhang zwischen systematischem Explorieren und praktischem Computerwissen.

Demgegenüber hatte das praktische Computerwissen in Wechselwirkung mit dem Messzeitpunkt einen statistisch signifikanten Effekt auf die Nutzung der *rigiden Explorationsstrategie* ($F [2, 24] = 3.53, p = .031, \eta^2 = .04$). Der negative Zusammenhang zwischen dem praktischen Computerwissen und der Häufigkeit dieser unsystematischen Explorationsstrategie nahm von Messzeitpunkt 1 zu Messzeitpunkt 3 zu: Personen mit hohem praktischem Computerwissen zeigten zum letzten Messzeitpunkt eine geringere Tendenz zu rigider Exploration als Personen mit niedriger Computererfahrung.

Zusammenfassung. Die systematische Exploration stellte über alle Messzeitpunkte hinweg die am häufigsten beobachtete Explorationsstrategie dar (siehe Abbildung 23 sowie Tabelle 25). Systematisches Explorationsverhalten wurde dabei insbesondere zum ersten Messzeitpunkt beobachtet. Der Anteil dieser Explorationsstrategie erreichte hier den höchsten Wert aller Beobachtungen (55%). Demgegenüber wiesen die Explorationsstrategien Versuch-und-Irrtum sowie Informationssuche konstant eine relativ niedrige Frequenz im Explorationsprozess auf. Der jeweilige Anteil dieser Kategorien lag pro Messzeitpunkt stets unter 10%. Die rigide Exploration erreichte über die Messzeitpunkte hinweg einen Anteil von ca. 20% der Beobachtungseinheiten. Während des Explorationsprozesses war ein leichter Anstieg rigider Explorationsweisen zu verzeichnen. Die Kategorie „Unterbrechung“ umfasste zu Messzeitpunkt 1 einen Anteil von 16% der Beobachtungseinheiten. Dieser Anteil steigerte sich zum Messzeitpunkt 2 und blieb dann konstant. Rein deskriptiv verdeutlicht das Ergebnis, dass sich die Explorationsstrategien über die Zeit in unterschiedliche Richtungen entwickeln. Anders als die unsystematischen Explorationsstrategien zeigte sich bei der systematischen Explorationsstrategie eine deutliche Abnahme über die Zeit. In Interaktion mit dem Messzeitpunkt hatte das Vorwissen einen bedeutsamen Effekt auf die Strategiesystematik: Über die Zeit hinweg verstärkte sich die positive Korrelation mit der

systematischen Explorationsstrategie und die negative Korrelation mit der rigiden Explorationsstrategie.

9.2.1.3 Korrelationen von Strategiesystematik, Flow und Leistung

Die Basis für komplexere Zusammenhänge zwischen den Variablen im Rahmen des Pfadmodells (vgl. Hypothese 10, Abschnitt 9.2.2) bilden die bivariaten Interkorrelationen der zentralen Untersuchungsvariablen. Diese sind in Tabelle A – 35 (Abschnitt A 4.2.3) unter Berücksichtigung der drei Messzeitpunkte (t1 bis t3) aufgeführt. Die Deskriptiven Statistiken der Skalenwerte sind ebenfalls im Anhang (Tabelle A – 34, Abschnitt A 4.2.2) nachzulesen. Da im Pfadmodell nur die systematische Explorationsstrategie berücksichtigt wird, wurden zusätzlich die Korrelationen der unsystematischen Explorationsstrategien zu den drei Messzeitpunkten mit dem Flow-Erleben und der Leistung berechnet. Die Korrelationen werden in Tabelle A – 36 im Anhang (Abschnitt A 4.2.3) präsentiert.

Für die unsystematischen Explorationsstrategien ergaben sich unterschiedliche Beziehungen zu Flow und zur Leistung (s. Tabelle A – 36): Die *rigide* Explorationsstrategie stand zu allen Messzeitpunkten in einem schwachen bis mittelstarken, negativen Zusammenhang mit dem Flow-Erleben sowie in einer starken, negativen Beziehung zur Leistung. Die Explorationsstrategie *Versuch-und-Irrtum* wies nur zu Messzeitpunkt 2 statistisch signifikante, negative Beziehungen zur Leistung und zum Flow-Erleben auf. Die *Informationssuche* korrelierte zum Ende des Explorationsprozesses schwach und negativ mit Leistung und Flow-Erleben. Das Kategoriensystem umfasste zusätzlich die Kategorie „Unterbrechung“, deren Beziehung zur Leistung über die Messzeitpunkte hinweg variierte (t1: $r = .51, p = .01$; $r = .39, p = .01$; $r = .15, p = .15$).

9.2.2 Hypothese 10: Wirkungsbeziehungen der Variablen im Zeitverlauf

In Hypothese 10 (s. Abschnitt 6.3) wurde ein Pfadmodell postuliert, das die Beziehungen zwischen dem Flow-Erleben und der Strategiesystematik sowie deren Einfluss über die Zeit auf die Leistung zusammenfassend beschreibt. In dem theoretischen Pfadmodell sollte der Effekt des Geschlechts auf die Leistung über Unterschiede in der Eingangsmotivation vermittelt sein. Es wurde erwartet, dass diese Unterschiede wiederum einen Einfluss auf das Flow-Erleben und die Häufigkeit der systematischen Strategie während des Lernprozesses ausüben.

Modellmodifikation. Vor der empirischen Überprüfung des theoretischen Modells wurde das Modell leicht modifiziert. Die Variable „Systematische Explorationsstrategie t1“ hatte in der bivariaten Analyse lediglich einen schwachen Zusammenhang mit der Leistung sowie dem Flow-Erleben gezeigt. Aufgrund dieses Befunds habe ich entschieden, diese Variable nicht in das Pfadmodell aufzunehmen. Stattdessen wurden, wie bereits in Abschnitt 9.2.1.3 dargelegt, die Da-

ten dieser beiden Messzeitpunkte zu einer Variablen „Systematische Explorationsstrategie t1/t2“ aggregiert und in das Modell aufgenommen. Dieses Vorgehen hatte den Vorteil, dass die Modellstruktur vereinfacht werden konnte. Für die Variable Flow wurden ausschließlich die letzten beiden Messzeitpunkte (t2 und t3) mit der Strategiesystematik in Bezug gesetzt. Wie Tabelle A – 35 im Anhang zeigt, korrelierte das Geschlecht (wie bereits in Studie 2 der Fall) nicht mit der Eingangsmotivation. Daraufhin wurden die vier Subskalen der Eingangsmotivation erneut nicht zur Erklärung potenzieller Geschlechtseffekte einbezogen. Die Variable Eingangsmotivation wurde aus dem Modell entfernt. Das modifizierte Pfadmodell ist in Abbildung 25 illustriert.

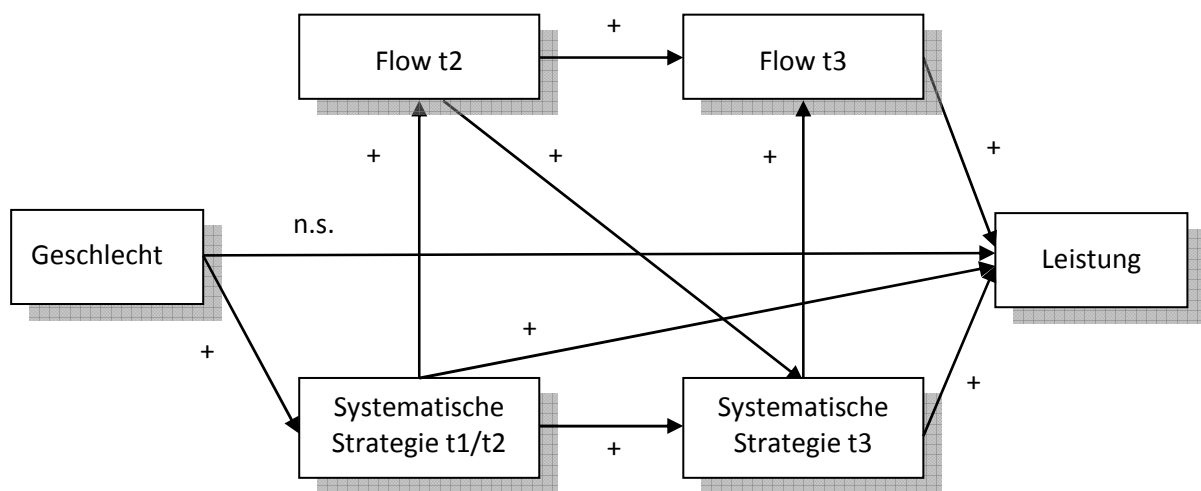


Abb. 25. Modifiziertes theoretisches Pfadmodell zur Vorhersage der Leistung mit den Variablen Geschlecht sowie den Prozessvariablen Systematische Explorationsstrategie und Flow-Erleben über mehrere Messzeitpunkte (Kodierung Geschlecht: Männer = -1, Frauen = 1). Die Strategienutzung der Messzeitpunkte t1 und t2 wurde dabei zu einem Wert zusammengefasst. Das Modell illustriert die erwartete Wechselbeziehung der Prozessvariablen. Pluszeichen spiegeln positive erwartete Pfadkoeffizienten und Minuszeichen negative erwartete Pfadkoeffizienten wider.

Statistische Modellprüfung. Die linearen Strukturbeziehungen der Modellvariablen wurden mit Hilfe von *Mplus* 5.0 (Muthén & Muthén, 1998-2007) auf der Basis des Maximum-Likelihood-Algorithmus überprüft. Nach Modifikation beinhaltete das zu prüfende Modell sechs manifeste (beobachtete) und vier latente Variablen (*disturbances*), davon sechs exogene und vier endogene. Die Skalierung der latenten Variablen wurde dabei auf 1 festgesetzt. Mit $[6 (6+1)] / 2 = 21$ Stichprobenmomenten und 16 zu schätzenden Parametern galt das Modell als identifiziert.

Zur Prüfung der Voraussetzungen für die Parameterschätzungen wurde die Abweichung der Verteilungen in den Variablen von der theoretischen multivariaten Normalverteilung getestet. Zu Messzeitpunkt 3 hatte sich ein systematischer Datenausfall ergeben, da zwei Teilnehmer die Bearbeitung in SPSS vorzeitig beendet hatten. Diese beiden Teilnehmer zeigten eine besonders hohe Leistung, in dem sie die sieben Statistikaufgaben vor Ablauf der 30 min lösen konnten.

Aufgrund dieser fehlenden Messwerte zu Messzeitpunkt 3 konnte der Mardia-Test auf multivariate Schiefe und Kurtosis nicht mittels AMOS durchgeführt werden. Es wurde lediglich eine statistische und eine grafische Prüfung der univariaten Verteilungen vorgenommen. Einschränkend muss dazu bemerkt werden, dass diese als notwendige, jedoch nicht als hinreichende Bedingung für die Annahme einer multivariaten Normalverteilung gilt (Kline, 2005). Die Stichprobengröße wurde Bezug nehmend auf McCallum et al. (1996) sowie Iacobucci et al. (2007) als ausreichend bewertet (vgl. Abschnitt 8.1.4).

In Tabelle 26 sind die Deskriptiven Statistiken sowie die Residuen als Abweichung der empirischen von der durch das Modell geschätzten Kovarianzmatrix aufgeführt. Die Interkorrelationen der ins Modell aufgenommenen Variablen sind ebenfalls dargestellt. Alle endogenen Variablen wurden als intervallskaliert angenommen. Die univariate Kurtosis sowie die univariate Schiefe lagen in der Stichprobe zwischen -.97 und -.06 bzw. -.52 und -.06. Nach West et al. (1995) gilt ein Kurtosis-Koeffizient von nahe bzw. über 7.0 als bedenklich. Eine Schiefe ab einer Höhe von 2.0 wird von den Autoren ebenfalls als kritisch betrachtet. Die empirischen Werte unterschritten jedoch die kritischen Grenzen.

Tabelle 26

Mittelwerte (M) und Standardabweichungen (SD), univariate Kurtosis und Schiefe, Interkorrelationen der Modellvariablen in Hypothese 10, Kovarianzen sowie standardisierte Residuen (N ≥ 90)

Variablen	Deskriptive Statistiken				Korrelationen und Residuen					
	M	SD	Kurtosis	Schiefe	1	2	3	4	5	6
1 System. Strategie t1/t2	44.13	11.0	-.22	-.16	120.44	0	0	0	0	0
2 System. Strategie t3	33.83	18.7	-.74	-.07	.46**	347.45	0	0	0	0
3 Flow t2	4.36	1.2	-.39	-.35	.32**	.38**	1.42	0	0	0
4 Flow t3	4.47	1.3	-.06	-.52	.25*	.41**	.84**	1.66	0	0
5 Leistung_Arbeitsblatt	10.57	6.9	-.97	-.06	.64**	.69**	.46**	.43**	47.22	0
6 Geschlecht ¹	-	-	-	-	.02	.02	.01	.09	.15	1.0

Anmerkungen. System. = Systematische. Die Stichprobenkorrelationen sind unterhalb der Diagonale und die Residuen oberhalb der Diagonale wiedergegeben. In der Diagonalen sind die Kovarianzen abgetragen. ¹ Für die kategoriale Variable Geschlecht wurden keine Deskriptiven Statistiken berechnet; die Variable wurde jedoch in Korrelationsanalysen einbezogen (Kodierung: Männer = 1, Frauen = - 1). * $p < .05$, ** $p < .01$.

Ergebnisse der Parameterschätzung und Modellgüte. Die Beurteilung der Gesamtstruktur erfolgte auf der Basis ausgewählter deskriptiver und inferenzstatistischer Indizes globaler Modellgüte und Modellsparsamkeit (z. B. Schermelleh-Engel et al., 2003). Die in Tabelle 27 aufgeführten Indizes geben Aufschluss darüber, inwieweit sich die theoretisch angenommene Modellstruktur empirisch bestätigen ließ. Die Höhe der Indices sowie die Empfehlungen nach

Schermelleh-Engel et al. (2003, S. 52) unterstützen die Annahme einer guten Passung des theoretischen Modells und der empirischen Daten.

Tabelle 27

Inferenzstatistische und deskriptive Maße zur Bewertung der globalen Modellgüte für das vorgeschlagene Modell

Modell	Index der Modellgüte				
	$\chi^2 (df; p)$	RMSEA (CI)	SRMR	CFI	AIC ¹
Vorgegebene Modellstruktur	3.35 (5; .65)	.001 (.001, .12)	.01	1.00	47.06

Anmerkungen. df = Freiheitsgrade; p = Signifikanzniveau; CI = Konfidenzintervall. ¹Das saturierte Modell lieferte hier einen Wert von AIC = 54 und das Independence-Modell erreichte einen AIC von 269. Damit gilt das im theoretisch spezifizierte als das sparsamste der drei Modelle.

Der χ^2 -Wert von 3.35 unterschritt mit $df = 5$ deutlich die Grenze von 2. Der empirische RMSEA-Wert fiel unter die empfohlene Grenze von $RMSEA \leq .05$; dieses Ergebnis spricht für einen guten Modellfit. Die obere Grenze des Konfidenzintervalls beinhaltet jedoch mögliche „wahre“ RMSEA-Werte von über $p = .10$. Die deskriptiven Anpassungsmaße CFI und AGFI mit Werten von 1.00 bzw. .99 sprechen wiederum für eine gute Anpassung des Modells an die empirischen Daten (vgl. Hu & Bentler, 1999). Allerdings berücksichtigen diese Indices nicht die Modellkomplexität. Der SRMR-Wert lag unter der von Schermelleh-Engel et al. (2003) genannten Grenze von .05 und impliziert damit ebenfalls einen guten Modellfit. Eine Inspektion der Teilstrukturen, die über die standardisierten Residuen der Kovarianzen vorgenommen wurde, ergab keine Auffälligkeiten (siehe Tabelle 26). Lokale Abweichungen zwischen der vorgegebenen und der empirischen Modellstruktur wurden nicht festgestellt (Jöreskog & Sörbom, 1984, zit. nach Arbuckle, 2008).

Signifikanzprüfung der direkten und indirekten Effekte. Abbildung 26 zeigt die empirisch ermittelten direkten Beziehungen der Modellvariablen. Details zu den Pfadkoeffizienten der direkten sowie der indirekten Pfade sind in Tabelle 28 aufgeführt.¹⁸

Das *Geschlecht* übte keinen statistisch signifikanten Einfluss auf die Häufigkeit der systematischen Explorationsstrategie aus. Darüber hinaus ließ sich keine Assoziation des Geschlechts mit der Höhe des Flow-Erlebens nachweisen. Der direkte Pfad zwischen Geschlecht und Leistung verfehlte ebenfalls die Signifikanzgrenze. Das Geschlecht wirkte sich somit weder direkt noch vermittelt über die im Modell berücksichtigten Mediatoren auf die Leistung aus.

¹⁸ Die standardisierten Pfadkoeffizienten im Text wurden zur Vereinfachung des Leseflusses mit der Abkürzung „St.“ gekennzeichnet.

Erwartungsgemäß hatte die gemittelte Häufigkeit der *systematischen Exploration* zu Messzeitpunkt *t1* und *t2* einen statistisch signifikanten Effekt auf das Flow-Erleben zu Messzeitpunkt *t2*, die Häufigkeit der systematischen Exploration zu Messzeitpunkt *t3* sowie auf die Leistung. Der Pfadkoeffizient von der systematischen Explorationsstrategie (*t1/t2*) zur Leistung spiegelt eine Varianzaufklärung von ca. 16% in der Leistung wider. Der direkte, mittelstarke Effekt (vgl. Cohen, 1988, zitiert nach Kline, 2005, S. 122) der Strategiesystematik *t1/t2* auf die Strategiesystematik zu Messzeitpunkt *t3* ($St. = .38$) wurde ergänzt durch einen indirekten Pfad: Ein Teil des Effekts der systematischen Explorationsstrategie zu *t1/t2* auf die Strategiesystematik in *t3* wurde über das Flow-Erleben zu Messzeitpunkt *t2* vermittelt ($St. = .08$; siehe Tabelle 28).

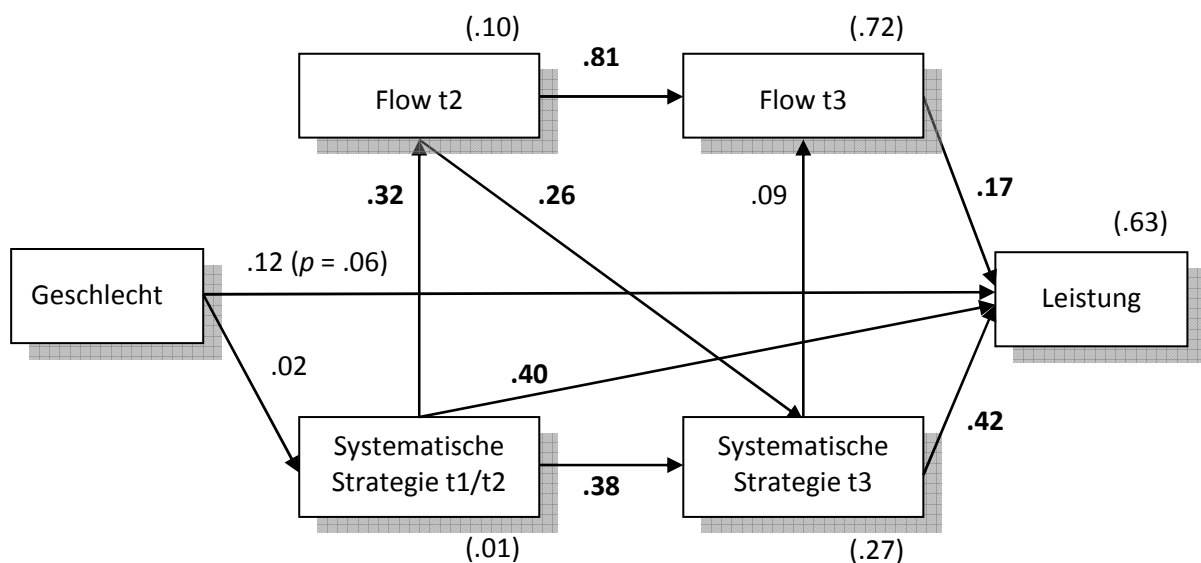


Abb. 26. Empirisches, rekursives Modell mit standardisierten Pfadkoeffizienten mit den Variablen Geschlecht (Kodierung: Männer = -1, Frauen = 1), Systematischer Explorationsstrategie und Flow-Erleben zur Vorhersage der Leistung. Die Strategiesystematik der Messzeitpunkte *t1* und *t2* wurde dabei zu einem Wert zusammengefasst. In Klammern ist der Anteil der aufgeklärten Varianz angegeben (R^2). Alle fettgedruckten Pfadkoeffizienten sind auf dem Niveau von $p < .05$ signifikant ($N = 92$).

Demgegenüber wurde das Flow-Erleben zu Messzeitpunkt *t3* stark von dem *Flow-Erleben* zu Messzeitpunkt *t1* und *t2* beeinflusst ($St. = .81$, $p = .01$). Der positive Effekt des Flow-Erlebens über die Messzeitpunkte hinweg erbrachte zusätzlich einen (indirekten) Effekt auf die Leistung ($St. = .14$, $p = .04$). Darüber hinaus hatte das Flow-Erleben in den ersten 20 Minuten (*t2*) einen statistisch signifikanten Einfluss auf die Strategiesystematik (*t3*) in den letzten 30 Minuten.

Die Häufigkeit der *systematischen Strategie im letzten Drittel* des Explorationsprozesses hatte einen mittelstarken Effekt auf die Leistung ($St. = .42$, $p = .27$). Dieser wurde nicht über das Flow-Erleben vermittelt, das am Ende des Arbeitsprozesses erfasst wurde. Der indirekte Effekt der systematischen Strategie zu *t3* auf die Leistung wurde nicht signifikant ($St. = .02$, $p = .27$). Auf

das Flow-Erleben zum letzten Messzeitpunkt (t3) hatte die Strategiesystematik entgegen der Erwartungen keinen statistisch signifikanten Einfluss ($St. = .09, p = .16$).

Das *Flow-Erleben zum letzten Messzeitpunkt (t3)* wirkte sich auch auf die Leistung aus ($St. = .17, p = .04$). Obwohl der Effekt der Strategiesystematik zu t3 auf die Leistung mehr als doppelt so groß war, klärte das Flow-Erleben einen zusätzlichen Varianzanteil in der Leistung auf, der nicht auf die Strategiesystematik zu Messzeitpunkt t1, t2 und t3 zurückgeführt werden konnte. Darüber hinaus übte das Flow-Erleben zu Messzeitpunkt t2 über zwei weitere Pfade, nämlich über die Strategiesystematik ($St. = .11, p = .01$) sowie über das Flow-Erleben zu t3 ($St. = .14; p = .04$), einen statistisch signifikanten Einfluss auf die Leistung aus. Insgesamt konnten mit Hilfe des Modells 63% der Varianz in der Leistung erklärt werden ($R^2 = .63$; vgl. Abbildung 24).

Tabelle 28

Effektzerlegung (direkte Effekte, indirekte Effekte) für das überidentifizierte Pfadmodell zur Vorhersage der Leistung über die Variablen Geschlecht, Flow und Nutzung Systematischer Explorationsstrategie (N = 92)

Effekt	USt.	SE	USt./SE	p	St.
Direkte Effekte					
Geschlecht → Strategie t1/t2	0.18	1.09	0.17	.87	.02
Geschlecht → Leistung	0.83	0.45	1.85	.06	.12
Strategie t1/t2 → Flow t2	0.03	0.01	2.94	.03	.32
Strategie t1/t2 → Strategie t3	0.64	0.17	3.75	.01	.38
Strategie t1/t2 → Leistung	0.25	0.04	6.06	.01	.40
Strategie t3 → Flow t3	0.01	0.01	1.41	.16	.09
Strategie t3 → Leistung	0.15	0.03	6.16	.01	.42
Flow t2 → Flow t3	0.88	0.07	12.95	.01	.81
Flow t3 → Leistung	0.89	0.42	2.10	.04	.17
Indirekte Effekte					
Geschlecht → Strategie t1/t2 → Strategie t3 → Leistung	0.02	0.11	0.16	.87	.01
Flow t2 → Flow t3 → Leistung	0.78	0.38	2.06	.04	.14
Flow t2 → Strategie t3 → Leistung	0.62	0.25	2.46	.01	.11
Strategie t1/t2 → Flow t2 → Strategie t3	0.14	0.07	2.10	.04	.08
Strategie t3 → Flow t3 → Leistung	0.01	0.01	1.11	.27	.02

Anmerkungen. Strategie = Systematische Explorationsstrategie; t1, t2, t3 = Messzeitpunkte 1 bis 3; USt. = Unstandardisierter Pfadkoeffizient; SE = Standardfehler; St. = Standardisierter Pfadkoeffizient; p = Signifikanzniveau Unstandardisierter Koeffizient, 2-seitig).

Zusammenfassung. Es wurde ein hypothetisches Modell getestet mit dem Ziel, die Wirkungsbeziehungen zwischen den Variablen Geschlecht, Flow-Erleben, Strategiesystematik

und Leistung beim Erlernen eines Computerprogramms erforschen und quantifizieren zu können. Insgesamt zeigte das theoretische Pfadmodell eine gute globale Anpassung an die empirischen Daten, wenn mehrere Gütekriterien gleichzeitig zur Beurteilung des Modellfits herangezogen wurden. In Teilstrukturen des Modells ergaben sich jedoch Abweichungen von den theoretischen Annahmen.

Bezüglich der *direkten Effekte* lassen sich die Befunde folgendermaßen zusammenfassen: Entgegen den Erwartungen hatte das Geschlecht weder direkt noch indirekt einen statistisch signifikanten Einfluss auf die Leistung. Im theoretischen Pfadmodell (Hypothese 10) wurde postuliert, dass sich mögliche Geschlechtsunterschiede in der Leistung über die Ausprägungen in den Prozessvariablen Flow-Erleben und Strategiesystematik erklären lassen. Entgegen der Erwartungen hatte das Geschlecht keinen Einfluss auf die Strategiesystematik zum ersten und zweiten Messzeitpunkt (t1/t2): Männer und Frauen explorierten das Computerprogramm in gleichem Ausmaß in systematischer Weise.

Entgegen der in Hypothese 10 formulierten Zusammenhänge stand die Häufigkeit der systematischen Strategien nur innerhalb der ersten 20 Minuten (Messzeitpunkte t1 und t2) in einer positiven Beziehung mit dem Flow-Erleben. In den letzten 10 Minuten der Bearbeitungszeit trat Flow-Erleben (t3) unabhängig von der Strategiesystematik (t3) auf. In der bivariaten Korrelationsanalyse hatte sich zuvor ein mittelstarker, positiver Zusammenhang zwischen den beiden Variablen gezeigt ($r = .41, p = .01$). Im Pfadmodell zeigte sich, dass ein Großteil der Varianz des Flow-Erlebens zu Messzeitpunkt t3 (64%) durch die Höhe des Flow-Erlebens in den vorangegangenen Bearbeitungsphasen (t1/t2) erklärt wurde.

Die Strategiesystematik zu t3 wurde ebenfalls von der Strategiesystematik zu t1/t2 beeinflusst. Allerdings fiel dieser Effekt deutlich geringer aus im Vergleich zur Flow-Messung. Die Strategiesystematik in den ersten 20 Minuten (t1/t2) erklärte nur ca. 14% der Strategiesystematik in den letzten 10 Minuten (t3). Personen, die in den ersten 20 Minuten systematisch explorierten, taten dies nicht zwangsläufig auch in den letzten 10 Minuten. Die Strategiesystematik trug zu allen Messzeitpunkten (t1/t2, t3) mit jeweils 16% Varianzaufklärung zur Vorhersage von Leistungsunterschieden bei.

Zusätzlich zu den direkten Pfaden wurde auch der *Einfluss indirekter Effekte* auf die Leistung berücksichtigt. Hier zeigte sich, dass ein Teil des Effekts von Flow (t2) auf die Leistung über die Strategiesystematik zu t3 vermittelt wurde. Zum letzten Messzeitpunkt (t3) wurde der indirekte Pfad von der Strategiesystematik über das Flow-Erleben zur Leistung nicht signifikant. Das Flow-Erleben (t2) wirkte sich sowohl über das Flow-Erleben zu t3 als auch über die Strategiesystematik zu t3 indirekt auf die Leistung aus.

9.2.3 Hypothese 11: Identifizierbarkeit von Personenclustern

Gemäß Hypothese 11 wurde erwartet, dass Personen mittels einer Clusteranalyse danach klassifiziert werden können, welche Kombinationen von Explorationsstrategien bei ihnen besonders häufig gemeinsam auftreten. Um eine erste Validierung der Typenbildung vorzunehmen, wurden im Anschluss an die Clusteranalyse Mittelwertvergleiche in mehreren Abhängigen Variablen vorgenommen. Es wurde überprüft, ob sich die Cluster zum Beispiel hinsichtlich der aktuellen Motivation, des Flow-Erlebens und der Leistung unterscheiden.

Identifikation der Cluster von Personengruppen. Zunächst wurde eine Clusteranalyse zur Identifikation der Anzahl voneinander abgrenzbarer Gruppen durchgeführt. Die Teilnehmer wurden dazu hinsichtlich der individuellen Häufigkeiten in den Kategorien Systematische Exploration, Versuch-und-Irrtum, Rigide Exploration sowie Informationssuche miteinander verglichen (vgl. Wiedenbeck & Züll, 2001, S. 2). Die Kategorie Unterbrechung wurde ausgelassen, da es sich bei ihr nicht um die Beschreibung einer Explorationsstrategie handelt.

Die Bestimmung der Clusterzentroide erfolgte auf der Basis des hierarchisch-agglomerativen Ward-Verfahrens (vgl. Spiel, 1998). Das clusteranalytische Vorgehen nach Ward hat nach Backhaus et al. (2006, S. 522 ff.) den Vorteil, dass möglichst homogene Cluster gebildet werden und damit die „Binnenvarianz der Gruppierungen minimal wächst“ (Wiedenbeck & Züll, 2001, S. 9). Das Auftreten von Ausreißern wird auf diese Weise vermieden. Im Zuge des Prozesses werden diejenigen Gruppen bzw. Objekte fusioniert, welche den Zuwachs der Gesamtvarianz innerhalb des Clusters am geringsten erhöhen (Langfeldt & Prücher, 2001). Um die Zunahme der Varianz innerhalb der Gruppe bestimmen zu können, wurde die Fehlerquadratsumme als Heterogenitätsmaß herangezogen (Backhaus et al., 2006, S. 522). Als Maß für die Unähnlichkeiten fungierte die quadrierte euklidische Distanz.

Die Clusteranalyse ergab den folgenden Verlauf der Fehlerquadratsummen für die letzten sechs Fusionierungsstufen (Ein-Cluster-Lösung bis Sieben-Cluster-Lösung): 364, 250, 179, 143, 120, 101, 91. Nach dem dritten Fusionsschritt fällt der Zuwachs der Fehlerquadratsumme deutlich geringer aus. Dies wurde als „Ellbogen“ der Verteilung im Sinne des *Scree*-Tests interpretiert (vgl. Bacher, 1994, S. 87; Backhaus et al., 2006, S. 536). Dieses Ergebnis sprach für eine Drei-Cluster-Lösung. Das Ergebnis der Clusteranalyse ist in Tabelle 29 noch einmal tabellarisch dargestellt.

Tabelle 29

Ergebnisse der hierarchisch-agglomerativen Clusteranalyse zur Gruppierung auf der Grundlage der Explorationsstrategienutzung (Anzahl der Fälle: $N = 92$)

Agglomerationsschritt (nur die letzten sieben Schritte)	Zusammengefasste Cluster	Zunahme der Heterogenität
85	2 und 7	91
86	1 und 12	101
87	2 und 5	120
88	1 und 6	143
89	2 und 15	179
90	1 und 3	250
91	1 und 2	364

Um zu überprüfen, ob sich die drei Cluster im Hinblick auf alle berücksichtigten Abhängigen Variablen (vier Explorationsstrategien) signifikant voneinander unterscheiden (vgl. Bortz, 2005, S. 585 ff.), wurden multivariate Mittelwertvergleiche vorgenommen (MANOVA). In Tabelle 30 sind die Ergebnisse der Clusteranalyse sowie der Varianzanalyse zusammenfassend dargestellt. Die Ergebnisse der MANOVA zeigen für die Clusterzugehörigkeit einen statistisch signifikanten globalen Effekt ($F [8, 170] = 46.69, p = .001, \eta^2 = .69$). Dieses Ergebnis weist darauf hin, dass distinkte Gruppen identifiziert werden konnten.

Tabelle 30

Ergebnisse der Clusteranalyse für die durchschnittliche Nutzung unterschiedlicher Explorationsstrategien (in % der Häufigkeit aller Strategien)

Variable	Gruppe 1 ($n_1 = 54$)		Gruppe 2 ($n_2 = 22$)		Gruppe 3 ($n_3 = 16$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Systematische Exploration	48.25 ^a	7.40	27.21 ^b	7.71	37.22 ^c	8.85
Versuch-und-Irrtum	2.92 ^a	3.11	8.68 ^b	7.74	1.88 ^a	1.59
Rigide Exploration	14.68 ^a	9.12	40.05 ^b	13.00	17.70 ^a	8.92
Informationssuche	2.60 ^a	3.04	3.51 ^a	4.52	17.82 ^b	5.95

Anmerkungen. *M* = Mittlere relative Häufigkeit (%); *SD* = Standardabweichung. Die hochgestellten Buchstaben verdeutlichen statistisch signifikante Unterschiede zwischen den Gruppen, welche in paarweisen a-posteriori Einzelvergleichen (LSD) gefunden wurden. Ungleiche Buchstaben zeigen dabei statistisch signifikante Mittelwertsunterschiede zwischen zwei Clustern in der Variablen auf.

Die Varianzen in den drei Clustern haben sich bezüglich der Variablen Versuch-und-Irrtum sowie Informationssuche nicht als homogen erwiesen (Levene-Statistik Versuch-und-Irrtum: $F [2, 89] = 8.32, p = .001$; Informationssuche: $F [2, 89] = 5.72, p = .005$). Da diese Voraussetzungsverletzung bei gleichzeitigem Vorliegen ungleicher Gruppengrößen für die Ergebnisse

einer Varianzanalyse bedeutsam sein kann (Bortz, 1993, S. 263), wurde der Gruppenvergleich zusätzlich über ein nonparametrisches Testverfahren vorgenommen. Der Kruskal-Wallis-Test als verteilungsfreies Pendant zur Varianzanalyse (Bortz et al., 2008) ergab in allen Variablen zu den Explorationsstrategien ebenfalls signifikante Unterschiede ($20.91 \leq \chi^2 \leq 52.34$; alle $df = 2$, $p \leq .001$). Einzelvergleiche auf Rangdatenniveau wurden mittels Mann-Whitney-U-Tests vorgenommen. Die Ergebnisse der nonparametrischen Einzelvergleiche wichen jedoch nicht von denjenigen der parametrischen Post-Hoc-Tests (LSD) ab, so dass an dieser Stelle nicht näher auf die unterschiedlichen Tests eingegangen wird. Die Ergebnisse der Einzelvergleiche sind im nächsten Abschnitt zusammenfassend dargestellt. Auf der Grundlage dieser Befunde fußt die im nächsten Abschnitt beschriebene und inhaltlich interpretierte Clusterlösung.

Beschreibung der Clusterlösung. Die drei Cluster wurden folgendermaßen benannt: „Systematische Explorierer“ (Cluster 1), „Unsystematische Explorierer“ (Cluster 2), „Informationssucher“ (Cluster 3). Männer und Frauen waren in allen Clustern ungefähr gleich häufig vertreten. Die charakteristischen Merkmale der Gruppen sind deskriptiv in Tabelle 31 dargestellt.

Tabelle 31

Zusammenfassende Interpretation der Clusterlösung über Charakteristiken der Nutzung verschiedener Explorationsstrategien (+++ = häufige Nutzung, ++ durchschnittliche Nutzung, + seltene Nutzung)

Clusterbezeichnung	Beobachtungshäufigkeit der einzelner Explorationsstrategien			
	Systematisch	Versuch-und-Irrtum	Rigide	Informationssuche
Systematische Explorierer (n = 54)	+++	+	++	+
Unsystematische Explorierer (n = 22)	+	++	+++	+
Informationssucher (n = 16)	++	+	++	+++

Die Cluster lassen sich wie folgt beschreiben:

Die *systematischen Explorierer* (Cluster 1; n = 54, davon 52% weiblich) zeichneten sich im Vergleich zu Personen in den beiden anderen Clustern dadurch aus, dass sie besonders häufig eine günstige, d. h. systematische Explorationsstrategie eingesetzt haben (siehe Tabelle 30). Die häufige Nutzung der systematischen Strategie war ein Charakteristikum dieser Gruppe, die sich dadurch statistisch signifikant von den anderen beiden Clustern unterschied. Gleichzeitig wurde in dieser Gruppe die rigide Explorationsstrategie im Vergleich zu anderen Gruppen besonders selten beobachtet. In der Kategorie „Versuch-und-Irrtum“ ergaben sich bei den „Systematischen Explorierern“ und den „Informationssuchern“ (Cluster 3) ähnliche, niedrige Häufigkeitswerte. Jedoch unterschieden sich die beiden Gruppen dahingehend, dass die „Systematischen Explorie-

rer“ eher selten von der Hilfe-Funktion in SPSS Gebrauch machten und diese weniger häufig nutzten, um Informationen über das Programm zu suchen. Mit Personen aus Cluster 2 („Unsystematische Explorierer“, siehe unten) zeigten die „systematischen Explorierer“ keinerlei Übereinstimmung in den Explorationsstrategien. Im Gegensatz zu den „unsystematischen Explorierer“ gingen „systematische Explorierer“ stärker zielgerichtet vor. Ineffektive Handlungen wurden von ihnen nicht mehrmals hintereinander wiederholt. Die überwiegende Mehrheit der Studienteilnehmer wurde diesem ersten Cluster zugeordnet.

Die *unsystematischen Explorierer* in Cluster 2 ($n = 22$, davon 45% weiblich) zeichnen sich gegenüber den anderen beiden Personenclustern durch eine besonders geringe Nutzungshäufigkeit systematischer Strategien aus. Stattdessen zeigen Personen in diesem Cluster die höchsten Häufigkeiten bei der Nutzung von Versuch-und-Irrtum sowie der Nutzung rigider Explorationsstrategien. Die „unsystematischen Explorierer“ waren im Vergleich zu den Informationssuchern jedoch weniger damit beschäftigt, Informationen über Programmfunktionen einzuholen. Sie neigten stärker dazu, einmal im System beschrittene Wege mehrmals zu wiederholen, auch wenn sich diese als ineffektiv erwiesen hatten. Ihre Verhaltensweisen variierten deutlich weniger als bei den „Systematischen Explorierern“ (Cluster 1). Zudem waren die Handlungsschritte der unsystematischen Explorierer mittels Fremdbeobachtung nicht vorhersagbar und nicht als planvoll zu erkennen.

Informationssucher in Cluster 3 ($n = 16$, davon 44% weiblich) unterschieden sich von den anderen Gruppen in erster Linie durch eine besonders häufige Nutzung der Hilfefunktion in SPSS. Die Gruppenmitglieder verbrachten einen großen Teil des Aufgabenbearbeitungsprozesses mit dem Sammeln von Informationen über das Computerprogramm. Unsystematische Explorationsstrategien (übermäßige Informationssuche, rigide Exploration, Versuch-und-Irrtum) hielten sich mit dem Anteil systematischer Explorationsstrategien in etwa die Waage (37.40% vs. 37.22%). Ein Vorgehen nach Versuch-und-Irrtum wurde bei Personen dieses Clusters jedoch selten beobachtet.

Gruppenspezifische Unterschiede in Flow-Erleben, aktueller Motivation und Leistung. In einer ersten Validierung der Clusterzuordnungen wurde überprüft, ob sich die unterschiedliche Nutzung sämtlicher beobachteter Explorationsstrategien (repräsentiert in den Personenclustern) auf motivationale Variablen (Flow-Erleben und aktuelle Motivation zum letzten Messzeitpunkt) sowie auf die Leistung auswirkt.

Zu diesem Zweck wurde eine MANOVA berechnet, die signifikant ausfiel ($F [6, 170] = 12.30, p = .001, \eta^2 = .30$). Es wurden drei Abhängige Variablen (aktuelle Netto-Motivation, Flow, Leistung) aufgenommen. Das praktische Computerwissen (PRACOWI), das zuvor in Korrelationsanalysen einen statistisch signifikanten Zusammenhang zur Leistung zeigte, wurde dabei als

Kovariate in die MANOVA aufgenommen. Der Einfluss dieser Kovariate erwies sich im Globaltest als statistisch signifikant ($F [3, 86] = 3.14, p < .029, \eta^2 = .10$): Die Höhe des Flow-Erlebens ($F = [1, 88] = 4.55, p = .04, \eta^2 = .05$) und der Leistung ($F = [1, 88] = 4.96, p = .03, \eta^2 = .05$) wird durch die Kovariate „Praktisches Computerwissen“ mit beeinflusst. Auf die aktuelle Motivation hatte die Kovariate keinen statistisch signifikanten Effekt ($F = [1, 88] = 0.96, p = .33, \eta^2 = .01$). Die Ergebnisse sind in Tabelle 32 aufgeführt.

Tabelle 32

Mittelwerte (M) und Standardabweichungen (SD) in den drei Gruppen¹ der Clusteranalyse für die Variablen Flow-Erleben und aktuelle (Netto-)Motivation sowie für die Variable Leistung

Variable	Gruppe 1 ($n_1 = 54$)		Gruppe 2 ($n_2 = 22$)		Gruppe 3 ($n_3 = 16$)	
	M	SD	M	SD	M	SD
Flow (letzter Messzeitpunkt)	4.85 ^a	1.1	3.69 ^b	1.3	4.41 ^a	1.5
Aktuelle (Netto-)Motivation (letzter Messzeitpunkt)	10.65 ^a	4.1	8.70 ^{a,b}	4.2	8.06 ^b	4.1
Leistung_Arbeitsblatt	14.06 ^a	5.4	3.61 ^b	5.1	8.38 ^c	5.5

Anmerkungen. ¹Interpretation der Clusterlösung: Gruppe 1 „Systematisch Explorierende“; Gruppe 2 „Unsystematisch Explorierende“; Gruppe 3 „Informationssucher“. Die hochgestellten Buchstaben verdeutlichen statistisch signifikante Unterschiede zwischen den Gruppen, welche in paarweisen a-posteriori Einzelvergleichen (LSD) gefunden wurden. Ungleiche Buchstaben zeigen dabei Mittelwertunterschiede zwischen zwei Clustern in der jeweiligen Variablen auf. Die Voraussetzungen für diese Berechnungen waren zuvor überprüft worden und können als erfüllt gelten.

Über den signifikanten Einfluss des praktischen Computerwissens hinaus zeigte die Gruppenzugehörigkeit nach Bonferroni-Korrektur in zwei der drei Abhängigen Variablen einen bedeutsamen Haupteffekt: Die Clusterzugehörigkeit („Systematische Explorierer“, „Unsystematische Explorierer“, „Informationssucher“) erklärte ca. 12% der Gesamtvarianz in der Variablen Flow-Erleben ($F [2, 88] = 5.80, p = .01, \eta^2 = .12$). In der Variablen Performanz („Leistung_Arbeitsblatt“) kann fast die Hälfte der Gesamtvarianz (40%) auf die unterschiedliche Nutzung der Explorationsstrategien, zusammengefasst in den Clustern, zurückgeführt werden ($F [2, 88] = 28.72, p = .01; \eta^2 = .40$).¹⁹ Der Haupteffekt der aktuellen Motivation zum letzten Messzeitpunkt war bei Kontrolle des Vorwissens nicht signifikant ($F [2, 88] = 3.10, p = .05, \eta^2 = .07$).

¹⁹ Zum Vergleich: Im Rahmen einer multiplen Regressionsanalyse wurden die vier Explorationsstrategien Systematische Exploration, Rigide Exploration, Versuch-und-Irrtum, Informationssuche als Prädiktoren zur Vorhersage der Leistung simultan in die Regressionsgleichung einbezogen. Dieses Vorgehen diente dem Zweck, den von allen Variablen gemeinsam erklärten Varianzanteil im Kriterium zu quantifizieren. Im Gegensatz zur Clusteranalyse handelte es sich hierbei um einen Ansatz mit Variablenfokus anstelle eines Personenfokus (vgl. von Eye & Bogat, 2006).

Für die Prädiktoren Systematische Exploration ($\beta = .18, p = .04$), Rigide Exploration ($\beta = -.70, p < .01$) sowie die Informationssuche ($\beta = -.21, p < .01$) ergaben sich signifikante Beta-Gewichte. Insgesamt konnten mit Hilfe des Modells 76 % der Varianz im Kriterium erklärt werden ($R^2_{\text{kor}} = .76$). Im Vergleich zum personenzentrierten Ansatz der Clusteranalyse konnte mit dem Vorgehen ein zusätzlicher Beitrag zur Varianzaufklärung von 26 % erreicht werden. Jedoch muss die partielle Multikollinearität der Prädiktoren (Toleranz, Variance Inflation Factor VIF) beachtet

Über das praktische Computervorwissen hinaus ist danach die Kombination eingesetzter Explorationsstrategien beim Erlernen eines Computerprogramms ausschlaggebend für die erreichte Leistung. Welche Explorationsstrategien die Lernenden nutzen, hat darüber hinaus einen bedeutsamen Einfluss auf die Höhe des Flow-Erlebens während des Prozesses. Dieser Einfluss wird unabhängig vom computerbezogenen Vorwissen wirksam. Post-Hoc-Tests (LSD; ohne Kovariate) machten statistisch signifikante Unterschiede zwischen den Clustern in den Abhängigen Variablen sichtbar - allerdings zeigten sich diese nicht konsistent zwischen allen Gruppen (s. Tabelle 32).

In der *Leistung* zeigten sich Unterschiede zwischen den „Systematischen Explorierern“ (Cluster 1) gegenüber den „Unsystematischen Explorierern“ (Diff = 10.45, $p = .001$) sowie gegenüber den „Informationssuchern“. „Systematische Explorierer“ erzielten dabei erwartungsgemäß und in Übereinstimmung mit früheren Befunden die höchste Leistung. „Unsystematische Explorierer“ und „Informationssucher“ unterschieden sich ebenfalls statistisch signifikant in der Leistung (Diff = -4.76, $p = .029$): Die „Informationssucher“ erreichten dabei bessere Leistungen als die „Unsystematischen Explorierer“ (s. Tabelle 32).

Auch bezüglich des *Flow-Erlebens* wurden bedeutsame mittlere Differenzen zwischen den Clustern gemessen: Die „Systematischen Explorierer“ gaben signifikant höhere Flow-Werte an als die „Unsystematischen Explorierer“ (Diff = -1.17, $p = .001$). Die „Informationssucher“ unterschieden sich dabei nicht statistisch signifikant von der letztgenannten Gruppe; das Ausmaß ihres Flow-Erlebens war mit dem der „systematischen Explorierer“ vergleichbar (s. Tabelle 32).

Da bezüglich der *aktuellen (Netto-)Motivation* kein statistisch bedeutsamer Overall-Effekt zu verzeichnen war, wird auf diese Differenz nicht näher eingegangen. Eine zusätzliche Prüfung auf Gruppenunterschiede in der Eingangsmotivation vor der Aufgabenbearbeitung ergab ebenfalls für keine der vier Aspekte statistisch signifikante Differenzen (Interesse: $F [2, 89] = 0.05$, $p = .948$; Erfolgswahrscheinlichkeit: $F [2, 89] = 0.11$, $p = .92$; Herausforderung: $F [2, 89] = 0.62$, $p = .54$; Misserfolgsbefürchtung: $F [2, 89] = 0.01$, $p = .99$).

Zusammenfassung. In der Clusteranalyse ist es gelungen, Personen nach deren Nutzung unterschiedlicher Explorationsstrategien voneinander abzugrenzen. Danach konnten die Teilnehmer in drei distinkte Gruppen klassifiziert werden: „Systematische Explorierer“, „Unsystematische Explorierer“ und „Informationssucher“. Es zeigten sich nicht nur qualitative Diffe-

werden, die zu Verzerrungen der Parameterschätzungen führen kann: Die Systematische Exploration ($T = .34$, $VIF = 2.96$) und die Rigide Exploration ($T = .40$, $VIF = 2.52$) erwiesen sich als korrelierte Prädiktoren. Beide Werte (T , VIF) sollten nahe null liegen, überschreiten jedoch nicht den kritischen Wert von 10 (vgl. Diehl & Staufenberg, 2007, S. 439). Die Ergebnisse der Regressionsanalyse sind vorsichtig zu interpretieren. Darüber hinaus liefert das Ergebnis der Regressionsanalyse keine Hinweise darauf, welche Profile von Ausprägungen in den einzelnen Variablen zu einer günstigen bzw. ungünstigen aktuellen Motivation bzw. zu einer Beeinträchtigung des Flow-Erlebens führen. Für Interventionen, die an personenspezifischen Defiziten beim Erlernen von Computerprogrammen ansetzen sollen, ist diese Analyse weniger informativ.

renzen in der Strategiesystematik. In zwei der drei getesteten Abhängigen Variablen führte die Clusterbildung auch zu quantitativen Unterschieden: Beispielsweise wirkte sich eine häufige Nutzung der systematischen Explorationsstrategie bei gleichzeitig seltener Nutzung aller weiteren Explorationsstrategien günstig auf das Flow-Erleben und die Leistung aus: „Systematische Explorierer“ hatten danach am stärksten das Gefühl, die Aufgabe unter Kontrolle zu haben und darin aufzugehen.

Anders war dies bei den „Unsystematischen Explorierern“. Personen, welche dieser Gruppe zugeordnet wurden, gaben die niedrigsten Flow-Werte an im Vergleich aller Gruppen und ließen sich damit sowohl von den „Systematischen Explorierern“ in Cluster 1 als auch von „Informationssuchern“ in Cluster 3 abgrenzen. Die unsystematische Exploration, insbesondere die wiederholte Durchführung nicht zielführender Explorationsverhaltensweisen, hatte negative Auswirkungen auf die Leistung. „Unsystematische Explorierer“ lösten dementsprechend am wenigsten Aufgaben in SPSS bzw. erreichten dabei die geringste Punktezahl. Das Ergebnis kann damit erklärt werden, dass sich die Nutzung suboptimaler Strategien ungünstig auf das Flow-Erleben auswirkt und das mangelnde Flow-Erleben wiederum einen ungünstigen Effekt auf die Nutzung systematischer Strategien hatte (vgl. Hypothese 10).

Im Gegensatz zum Flow-Erleben und der Leistung konnten die Cluster keine Unterschiede in den vier Aspekten der Eingangsmotivation aufklären: Vor dem Bearbeitungsprozess unterschieden sich die Nutzergruppen nicht in ihrer Eingangsmotivation. Das gleiche Ergebnis zeigte sich im Hinblick auf die aktuelle (Netto-)Motivation während der Aufgabenbearbeitung. Hier zeigten sich ebenfalls keine Unterschiede zwischen den Clustern. Frauen und Männer verteilten sich ungefähr gleichermaßen über die Cluster. Es zeigte sich folglich keine geschlechtsspezifische Häufung in einzelnen Clustern. Das bedeutet, dass auch mittels des personenzentrierten Ansatzes der Clusteranalyse (vgl. von Eye & Bogat, 2006) keine Unterschiede zwischen den Geschlechtern im Sinne von Computernutzertypen aufgedeckt wurden.

9.3 Zusammenfassung und Diskussion

Aufbauend auf den Befunden der beiden vorangegangenen Studien wurden in Studie 3 Veränderungen in den Prozessvariablen (Flow, aktuelle Motivation, Strategiesystematik) näher erforscht. Die Analysen sollten dazu dienen, Geschlechtsunterschiede zu eruieren und zu erklären. Um die statistischen Voraussetzungen hierfür zu schaffen, wurde in Studie 3 ein größerer Stichprobenumfang realisiert. Ein weiterer Forschungsschwerpunkt lag auf den Wirkungsbeziehungen der Prozessvariablen untereinander (vgl. Bachmann, 2009). Um die Beziehungen der Variablen untersuchen zu können, wurde ein komplexes Pfadmodell postuliert und empirisch mit-

tels Strukturgleichungsanalysen getestet. Die variablenzentrierten Analyseansätze wurden in der Studie durch eine personenzentrierte Clusteranalyse ergänzt (vgl. von Eye & Bogat, 2006; vgl. auch Bacher, 1994; Wiedenbeck & Züll, 2001).

Geschlechtsunterschiede im Rahmen des kognitiv-motivationalen Prozessmodells. Entgegen der Erwartungen waren Geschlechtseffekte in Studie 3 nur gering ausgeprägt. Die männlichen Teilnehmer erzielten zwar weiterhin in beiden *Leistungsindikatoren* (Leistung_Arbeitsblatt, Leistung_SPSS) höhere Werte als die weiblichen Teilnehmer. Die Differenz erreichte jedoch keine statistische Signifikanz. Das bedeutet, dass die Frauen im Mittel ebenso viele Statistikaufgaben erfolgreich am Computer lösen konnten wie die Männer. Demzufolge ist es den Frauen und den Männern gleichermaßen gelungen, sich das neue Computerprogramm anzueignen und zielführend anzuwenden. Ähnliche Ergebnisse zeigten sich bezüglich der *Motivation* der Teilnehmer: Wie bereits in Studie 2 wiesen die Frauen keine ungünstigere Eingangsmotivation auf als die Männer. Hypothesenkonträr gaben die Frauen sogar tendenziell eine höhere Herausforderung an als die Männer. Auch während des Bearbeitungsprozesses berichteten Frauen ähnliche Ausprägungen in der *aktuellen Motivation* und im *Flow-Erleben* wie die Männer. Dieser Befund weicht von demjenigen in Studie 2 insofern ab, dass in Studie 2 statistisch signifikante Geschlechtsunterschiede in der aktuellen Motivation beobachtet wurden: Hier wurde eine höhere Netto-Motivation der Männer festgestellt. In Studie 3 zeigten sich zwischen Männern und Frauen auch hinsichtlich der *Explorationsstrategien* keine Unterschiede: Frauen nutzten im Vergleich zu Männern nahezu gleich häufig systematische bzw. unsystematische Explorationsstrategien (rigide Exploration, Versuch-und-Irrtum, Informationssuche). Dieses Ergebnis weicht von dem entsprechenden Befund in Studie 2 ab; hier hatten sich starke Geschlechtseffekte in Nutzung der systematischen und der rigiden Explorationsstrategie zuungunsten der weiblichen Teilnehmer gezeigt. Bei den varianzanalytischen Auswertungen über die beiden Stichproben muss beachtet werden, dass die Zellenbesetzungen stark voneinander abwichen (Bortz, 2005). Die Analyse sollten an weiteren, größeren Stichproben wiederholt werden.

Zusammenhang von Vorwissen und Leistung. Inwiefern das Vorwissen und die Vorerfahrung zur Leistung beitragen, wurde gesondert analysiert. Zunächst wurde überprüft, ob sich in Studie 2 beobachtete Geschlechtsunterschiede im Vorwissen in Studie 3 replizieren. Hierzu wurden die Ergebnisse beider Stichproben miteinander verglichen. Allerdings wurde das Vorwissen in den Studien 2 und 3 teilweise unterschiedlich operationalisiert. Es wurden für den Vergleich nur die Vorwissensvariablen ausgewählt, die in beiden Studien einbezogen worden waren.

Sowohl in Studie 2 als auch in Studie 3 zeigten sich im Vorwissen signifikante Unterschiede zwischen den Geschlechtern (siehe oben): Frauen gaben allgemein eine signifikant geringere Computererfahrung an als Männer (insbesondere bei Tabellenkalkulationsprogrammen).

Darüber hinaus ergaben beide Studien, dass die Computererfahrung positive Korrelationen mit der Leistung aufwies. Dieser Befund wirft jedoch Fragen auf, da sich in beiden Studien Geschlechtsunterschiede in der Computervorerfahrung zeigten, diese sich in Studie 3 jedoch nicht in unterschiedlichen Leistungen am Computer manifestierten. Ähnliche Beziehungen zwischen Vorwissen und Leistung zeigten sich auch für die PRACOWI-Items, die in Studie 3 nicht aber in Studie 2 eingesetzt wurden. Der PRACOWI wies positive Beziehungen zur Leistung auf ($r = .29$) und korrelierte auch negativ mit dem Geschlecht ($r = -.28$; s. Tabelle A-35 im Anhang). Danach konnten Frauen tendenziell weniger Aufgaben lösen als Männer. Diese Kompetenz hatte dabei einen positiven Einfluss auf den tatsächlichen Umgang mit dem Computerprogramm SPSS. Interessant bleibt deshalb der Befund, dass sich weibliche und männliche Teilnehmer der Studie 3 weiterhin im Vorwissen unterschieden, dies jedoch in Studie 3 mit geringeren Geschlechtseffekten in der Leistung einherging.

Variablenbeziehungen im Pfadmodell. Das in Hypothese 10 postulierte Pfadmodell erzielte in Studie 3 einen guten Modellfit. Bis auf die erwarteten Geschlechtsunterschiede (siehe oben) wurden die modelltheoretischen Annahmen mehrheitlich empirisch unterstützt. Die Nutzung der systematischen Strategie hatte positive Auswirkungen auf das Flow-Erleben sowie auf die Leistung. Lediglich die Vorhersage des Flow-Erlebens kurz vor Bearbeitungsende wurde nicht statistisch signifikant. Die Nutzung systematischer Strategien war in den ersten 20 Minuten von dem Gefühl begleitet, den Prozess unter Kontrolle zu haben und bei jedem Schritt zu wissen, was zu tun ist. Personen, die systematisch voringen, planten ihre Handlungen und evaluierten nach dem Handlungsvollzug das Ergebnis. Falls es erforderlich war, adaptierten sie ihr Verhalten im Anschluss an diese Bewertung. Dieser strukturierte Handlungsregulationskreislauf begünstigte nicht nur die erfolgreiche Aufgabenlösung, sondern auch das Flow-Erleben (vgl. Csikszentmihalyi, 1990/2008; Dörner, 1980, 2006; Schiefele & Rheinberg, 1997). Beide Variablen – die Strategiesystematik und das Flow-Erleben – trugen zur Varianzerklärung in der Leistung bei.

Explorationsstrategien im Zeitverlauf. Studie 3 diente des Weiteren dazu, die Nutzung der Explorationsstrategien zu mehreren Messzeitpunkten zu untersuchen. Veränderungen und Stabilitäten der einzelnen Explorationsstrategien im Zeitverlauf können sowohl im Pfadmodell als auch in der Varianzanalyse mit Messwiederholung nachvollzogen werden. In der Varianzanalyse zeigte sich, dass sich der Anteil der systematischen Strategie vom ersten zum zweiten Messzeitpunkt deutlich verringerte. Die relativen Häufigkeiten der unsystematischen Explorationsstrategien blieben dagegen stabil. Wurden jedoch die ersten beiden Messzeitpunkte (t1 und t2) zusammengefasst, so blieb die Häufigkeit der systematischen Strategie über den Lernprozess hinweg relativ konstant. Dies bedeutet, dass Personen, die in den ersten 20 Minuten systematisch explorierten, dies auch in den letzten 10 Minuten vor Bearbeitungsende taten. Allerdings variierte

die Häufigkeit der systematischen Explorationsstrategie im Vergleich zum Flow-Erleben stärker über die Zeit hinweg. Beide Befunde legen den Schluss nahe, dass sich bei den systematischen Explorierern am Ende des Bearbeitungsprozesses die Beobachtungshäufigkeiten in anderen Kategorien erhöhten. Dies kann insbesondere für die Kategorie Unterbrechung gelten, da die systematischen Explorierer in den ersten 20 Minuten Aufgaben erfolgreich bearbeiten konnten und immer häufiger mit dem Notieren der Ergebnisse auf dem Arbeitsblatt befasst waren. Dies wurde vom Beobachter als „Unterbrechung“ kodiert.

Der Zusammenhang einzelner Explorationsstrategien mit dem praktischen Computerwissen wurde ebenfalls im Rahmen der Studie analysiert. Über die Messzeitpunkte hinweg verstärkte sich der Zusammenhang zwischen systematischem Explorieren und praktischem Computerwissen. Inhaltlich bedeutet das Ergebnis, dass zu Beginn des Explorationsprozesses sowohl Personen mit niedrigem als auch mit hohem praktischem Computervorwissen systematisch explorierten. Zu Messzeitpunkt 2 und 3 wurde die Nutzung der systematischen Explorationsstrategie häufiger von denjenigen Personen gezeigt, die bereits vor der Studienteilnahme eine höhere Kompetenz erworben hatten, praktisch mit dem Computer zu arbeiten. Eine gegenteilige Beziehung zeigte sich zwischen der Computererfahrung und der rigiden Explorationsstrategie: Die negative Beziehung zwischen praktischem Computerwissen und der Nutzung der rigiden Explorationsstrategie verstärkte sich über die Messzeitpunkte hinweg. Folglich werden Personen mit hohem praktischem Computerwissen im Laufe des Prozesses seltener ineffektive Handlungen wiederholen. Im Gegenzug lernen diese Personen neue Handlungsrouitinen hinzu und systematisches Explorationsverhalten kann häufiger beobachtet werden.

Clusteranalyse. In der Clusteranalyse konnten drei Personengruppen anhand ihrer Explorationsstrategien voneinander abgegrenzt werden. Die erste Gruppe war gekennzeichnet durch eine häufige Nutzung der systematischen Explorationsstrategie, wohingegen die zweite Gruppe überwiegend unsystematische Explorationsstrategien wie die rigide Exploration und Versuch-und-Irrtum-Explorationsweisen zeigte. Die dritte Gruppe unterschied sich von den ersten beiden dahingehend, dass ihr Explorationsverhalten durch eine häufige Informationssuche gekennzeichnet war. Wurde neben der Nutzung der systematischen und der rigiden Explorationsstrategie vorwiegend auf die Informationssuche im Programm zurückgegriffen, wirkte sich diese Kombination ungünstig auf die Leistung aus, nicht jedoch auf das Flow-Erleben. Die günstigste Motivation und die höchste Leistung erzielten erwartungsgemäß die „systematischen Explorierer“. Im Unterschied dazu waren die „unsystematischen Explorierer“ statistisch signifikant geringer motiviert und zeigten sich bei der Aufgabenlösung weniger erfolgreich. Männer und Frauen verteilten sich ungefähr gleich über die drei Cluster. Das Ergebnis der Clusteranalyse legt den Schluss nahe, dass die als unsystematisch eingestuften Explorationsstrategien unterschiedliche

Effekte auf die Motivation und die Leistung beim Erlernen eines Computerprogramms haben können.

10. Gesamtdiskussion

In der vorliegenden Arbeit wurde der aktive, selbstgesteuerte Lernprozess beim Explorieren eines unbekanntes Computerprogramms empirisch untersucht. Frühere Studien liefern bereits Hinweise auf den Einfluss des Vorwissens, des Alters (Kang & Yoon, 2008) sowie der Belastung kognitiver Ressourcen (van der Linden et al., 2003) auf das erfolgreiche Erlernen einer Computeranwendung. Geschlechtsspezifische Unterschiede bei der Exploration einer neuen Software wurden bisher weitgehend vernachlässigt. Dies ist insofern von Nachteil, als dass mehrere Forschungsarbeiten aufzeigen, welche große Bedeutung die geschlechtsspezifische Sozialisation auf die Computernutzung ausübt (z. B. Cooper, 2006; Dickhäuser & Stiensmeier-Pelster 2002a, 2002b, 2003).

Um diese Forschungslücke zu schließen, thematisierte die vorliegende Untersuchung potenzielle Unterschiede zwischen Männern und Frauen im Umgang mit einer ihnen unbekanntes Computersoftware. Zur Erklärung von Leistungsunterschieden am Computer wurden Modelle aus dem Bereich des selbstregulierten Lernens sowie des Problemlösens herangezogen (z.B. Dörner, 2006; Dörner & Schaub, 1994; Rheinberg et al., 2000; Wirth & Leutner, 2008). Als Rahmenmodell zur empirischen Prüfung der Hypothesen diente dabei das kognitiv-motivationale Prozessmodell (Vollmeyer & Rheinberg, 1998, 2006), in dem Leistungsunterschiede auf Unterschiede in motivationalen und kognitiven Prozessvariablen zurückgeführt werden. Die besondere Rolle, welche die Motivation im Lernprozess spielt, wurde bereits in mehreren theoretischen und empirischen Werken herausgearbeitet. In Verhaltensstudien zum Erlernen von Computerprogrammen hat die Lernmotivation bisher jedoch eine untergeordnete Rolle gespielt (vgl. Kang & Yoon, 2008; Koch et al., 2008; Roy et al., 2003; Sieverding & Koch, 2009; Trudel & Payne, 1995, 1996; van der Linden et al., 2001, 2003; für Studien zum Fehlermanagement-Training beim Computerlernen siehe Keith & Frese, 2008). Als motivationale Prozessvariablen wurden in der vorliegenden Arbeit die aktuelle Motivation und das Flow-Erleben berücksichtigt (z.B. Csikszentmihalyi, 1990; Rheinberg, 2006; Rheinberg et al., 2001; Vollmeyer & Rheinberg, 2003). Die Strategiesystematik wurde als kognitive Prozessvariable in die Untersuchungen aufgenommen (van der Linden et al., 2001; Kang & Yoon, 2008; Vollmeyer et al., 1996).

Insgesamt widmete sich die vorliegende Arbeit vier Forschungsfragen. *Forschungsfrage 1* thematisierte das Problem, wie Explorationsstrategien beim Erlernen des Statistikprogramms SPSS reliabel und valide erfasst werden können. Im Rahmen von *Forschungsfrage 2* wurde eruiert, inwiefern Geschlechtsunterschiede in der Leistung auf Variablen des kognitiv-motivationalen Prozessmodells (z.B. Strategiesystematik, Flow-Erleben) zurückgeführt werden können. *For-*

schungsfrage 3 richtete sich auf die wechselseitige Beziehung der motivationalen und kognitiven Variablen während des Lernprozesses sowie auf deren Beziehung zur Leistung. In *Forschungsfrage 4* wurde erforscht, ob sich Personen danach klassifizieren lassen, welche Explorationsstrategien sie überwiegend bzw. in Kombination mit anderen Explorationsstrategien nutzen. In diesem Zusammenhang wurde explorativ untersucht, inwiefern Frauen in einzelnen Clustern stärker vertreten sind als Männer.

Es wurden drei Studien durchgeführt, in denen die vier genannten Forschungsfragen empirisch untersucht wurden. In allen Studien wurde ein lernerzentriertes Szenario gewählt, bei dem die Studienteilnehmer das Statistikprogramm SPSS anhand vorgegebener Aufgaben aktiv, selbstreguliert explorierten (vgl. Bell & Kozlowski, 2008; Green & Gilhooly, 1990; Keith et al., 2010; Zimmerman, 2008; Wirth & Leutner, 2008).

In den folgenden Abschnitten 10.1.1 bis 10.1.4 werden die Befunde zu den vier Forschungsfragen getrennt voneinander diskutiert. Zu Beginn eines jeden Kapitels werden die zentralen Ergebnisse zu der jeweiligen Forschungsfrage zusammengefasst. Im Anschluss an die theoretische Einordnung der Befunde (Forschungsfragen 2 bis 4) erfolgt eine kritische Betrachtung der Ergebnisse. In diesem Zusammenhang wird auch ein Ausblick auf zukünftige Forschungen gegeben, die sich aus Befunden und Grenzen der vorliegenden Arbeit ableiten lassen. Das darauffolgende Kapitel (10.2) fasst die methodenkritischen Überlegungen zu den drei Studien zusammen. Im abschließenden Kapitel der Arbeit (Abschnitt 10.3) wird die praktische Bedeutung der Forschungsergebnisse reflektiert. Interventionsmöglichkeiten zur Förderung des selbstregulierten Erlernens eines Computerprogramms stehen dabei im Mittelpunkt.

10.1. Ergebnisdiskussion und Forschungsausblick

10.1.1 Zur handlungsnahen Erfassung von Explorationsstrategien

(Forschungsfrage 1)

Zusammenfassung zentraler Ergebnisse. Ein wesentliches Ziel der vorliegenden Arbeit war es, die Strategiesystematik handlungsnah und non-reaktiv im Explorationsprozess zu erfassen (z.B. Artelt, 2000; Spörer & Brunstein, 2006). In Studie 1 ist es gelungen, ein reliables und valides Messinstrument zu entwickeln. In Übereinstimmung mit früheren Studien (Green & Gilhooly, 1990; van der Linden et al., 2001, 2003; vgl. auch Kang & Yoon, 2008) erlaubte das neu entwickelte Beobachtungssystem, verschiedene Explorationsstrategien beim Erkunden von SPSS trennscharf zu erfassen. Die in den drei Studien berechneten relativen Häufigkeiten der einzelnen Explorationsstrategien entsprachen mehrheitlich den Anteilen, die in früheren Studien berichtetet

wurden. Lediglich die Explorationsstrategie Versuch-und-Irrtum wies in der vorliegenden Arbeit zum Teil niedrigere Prävalenzraten auf.

Die Zuverlässigkeit des Messinstruments konnte für folgende Explorationsstrategien belegt werden: Systematische Exploration, Rigide Exploration und Informationssuche. Die Kategorie „Versuch-und-Irrtum“ ließ sich demgegenüber nur unzureichend von der systematischen sowie der rigiden Exploration abgrenzen. Die Verhaltensweisen, die als „Versuch-und-Irrtum“-Explorationsstrategie eingestuft wurden, variierten in hohem Maße zwischen den Beobachtern. Dies resultierte in einer geringen Beobachterübereinstimmung. Lediglich im Hinblick auf diese Kategorie genügte das Kategoriensystem damit nicht den wissenschaftlichen Anforderungen (vgl. Wirtz & Caspar, 2002; Greve & Wentura, 1997).

Anhand verschiedener Validitätskriterien wurde geprüft, ob mit Hilfe des Kategoriensystems tatsächlich optimale und suboptimale Explorationsverhaltensweisen erfasst werden. Auf der Grundlage bisheriger Studien wurden als Kriterien die Leistung in Form der erfolgreichen Lösung der Statistikaufgaben sowie das Flow-Erleben herangezogen (Kang & Yoon, 2008; van der Linden et al., 2001, 2008; Vollmeyer & Rheinberg, 2006). Das Kategoriensystem wies mehrheitlich die erwarteten Beziehungen zu Außenkriterien auf. In allen drei Studien korrelierte die systematische Exploration erwartungsgemäß positiv mit der Leistung (Debowski et al., 2001; Dörner, 2006; Green & Gilhooly, 1990; Pintrich & De Groot, 1990; Schiefele & Rheinberg, 1997; Vollmeyer et al., 1996): Computernutzer, die das Statistikprogramm zielgerichtet explorierten und systematisch mit verschiedenen Bedingungen experimentierten, erreichten stets höhere Leistungen beim Erkunden und Anwenden des Computerprogramms. Darüber hinaus zeigte sich in den Studien, dass Personen, die häufig systematisch explorierten, auch höhere Flow-Werte angaben (siehe auch Abschnitt 10.1.3 zur Wechselbeziehung der Prozessvariablen). Dieses Ergebnis stimmt mit Befunden aus früheren Studien überein (vgl. Engeser et al., 2005; Rheinberg, 2006; Rheinberg & Schiefele, 1997). Ähnliche Ergebnisse zeigten sich für die aktuelle Motivation: Eine höhere aktuelle Motivation während der Bearbeitung ging mit einem höheren Anteil systematischer Exploration einher. Personen mit einer höheren Computerbildung explorierten häufiger systematisch als Personen mit einer niedrigeren Computerbildung (vgl. Kang & Yoon, 2008; Richter et al., 2010). Eine Ausnahme bildeten die Ergebnisse in Studie 2, in der sich kein Zusammenhang zeigte. In Studie 1 zeigten sich zusätzlich positive Beziehungen der systematischen Exploration zum Statistikvorwissen. Diese wurden in Studie 2 und 3 jedoch nicht repliziert. Im Gegensatz zu Ghani und Deshpande (1994) war die Nutzung von Versuch-und-Irrtum in der vorliegenden Arbeit nicht positiv mit dem Flow-Erleben assoziiert.

Wie aufgrund theoretischer Überlegungen erwartet, standen die beiden unsystematischen Explorationsstrategien Rigide Exploration und Informationssuche in einer negativen Beziehung

mit der Leistung (vgl. Green & Gilhooly, 1990; van der Linden et al., 2001, 2003). Im Gegensatz zur systematischen Exploration gaben Personen, die häufiger rigide explorierten oder mehr Zeit für die Informationssuche aufwendeten, niedrigere Flow-Werte und eine geringere aktuelle Motivation an. Die Häufigkeiten in der Kategorie Versuch-und-Irrtum wiesen in keiner der drei Studien einen statistisch signifikanten Bezug zu Vorwissen, Motivation oder zur Leistung auf. Dieser Befund widerspricht den Ergebnissen, die von Kang & Yoon (2008), van der Linden et al. (2001, 2003) sowie Trudel und Payne (1995, 1996) berichtet wurden. Die Studien dieser Autoren haben ergeben, dass eine Exploration nach Versuch-und-Irrtum bzw. ein stärkeres Ausmaß thematischen Vagabundierens (Dörner, 1980, 2006) zu einer niedrigeren Leistung führt. Die Explorationsstrategien waren untereinander in charakteristischer Weise und erwartungsgemäß assoziiert: Die systematische Explorationsstrategie ging mit einer niedrigen Häufigkeit unsystematischer Explorationsstrategien einher. So korrelierte die Rigide Exploration mit der Systematischen Exploration in den Studien konsistent negativ.

Kritische Würdigung und Forschungsausblick. Die Ergebnisse der vorliegenden Arbeit weisen darauf hin, dass insbesondere die *Operationalisierung der Explorationsstrategie Versuch-und-Irrtum* kritisch eingeschätzt werden muss. Auch nach Modifikation des Messinstruments lag die Beobachterübereinstimmung in dieser Kategorie weiterhin außerhalb des akzeptablen Bereichs. Zudem korrelierte die Explorationsstrategie nicht wie erwartet mit den ausgewählten Kriteriumsvariablen. Die Ergebnisse lassen sich auf verschiedene Ursachen zurückführen: 1) Die mangelnde Disjunktheit der Kategorien, 2) die Prävalenz der Explorationsstrategie Versuch-und-Irrtum in der Stichprobe, 3) die Länge des Beobachtungsintervalls.

Die reliable Kodierung von Verhaltensspuren setzt voraus, dass das Beobachtungsinstrument ausreichend trennscharfe bzw. disjunkte Kategorien beinhaltet (vgl. Bakeman, 2000; Wirtz & Caspar, 2003). Gelingt das nicht, können unterschiedliche Verhaltensweisen nicht zufriedenstellend voneinander abgegrenzt werden. Anhand von zwei Beispielen soll verdeutlicht werden, wo sich Kategorien im neuen Kategoriensystem überlappt haben. Im Kategoriensystem wurde die Explorationsstrategie Versuch-und-Irrtum folgendermaßen von der rigiden Exploration abgegrenzt: „Wenn eine Variable immer wieder in die gleiche Box gesetzt wird, obwohl dieser Weg früher bereits zu einem Misserfolg geführt hat ... dann ist dies eine Handlung, die als rigide Exploration einzustufen ist“ (s. Abschnitt A 1.6, Tabelle A – 1). Dieser Indikator unsystematischer Exploration wurde als Merkmal sowohl in der Kategorie „Versuch-und-Irrtum“ als auch in der Kategorie „Rigide Exploration“ aufgeführt. Allein die Häufigkeit, mit der diese Handlung wiederholt wurde, sollte eine Abgrenzung gewährleisten. Das bedeutet, dass der Beobachter sich erinnern muss, ob er die Verhaltensweise bereits beobachtet hat und wie oft. Die Gedächtnisleistung kann damit einen maßgeblichen Einfluss auf die Kodierung haben. Ein zweites Beispiel

betrifft die Reaktion auf Fehlermeldungen des Programms: Dass Fehlermeldungen nicht ausreichend beachtet werden, war sowohl bei der rigiden Exploration als auch bei der Exploration nach Versuch-und-Irrtum ein Bestimmungsmerkmal (vgl. van der Linden et al., 2001). Bei der Modifikation des Kategoriensystems sollten die genannten Unschärfen genauer überprüft und korrigiert werden.

Verglichen mit bisherigen Untersuchungen war die Prävalenz, d. h. der Anteil der als Versuch-und-Irrtum eingestuften Verhaltensweisen in den drei Studien eher klein (vgl. z. B. Excel-Aufgabe von van der Linden et al., 2001, S. 195; van der Linden et al., 2003, S. 490). Die Aufgabenanforderungen, welche die Studienteilnehmer zu bewältigen hatten, variierten zwischen der vorliegenden Arbeit und früheren Untersuchungen. Es ist denkbar, dass unterschiedliche Aufgabenanforderungen die Nutzung einzelner Explorationsstrategien (z.B. Versuch-und-Irrtum) beeinflussen haben. Im Gegensatz zu Versuch-und-Irrtum war jedoch der Anteil der systematischen Exploration mit demjenigen in anderen Untersuchungen vergleichbar (van der Linden et al., 2001, S. 195).

Darüber hinaus ist denkbar, dass Verhaltensweisen, die dem Vorgehen nach Versuch-und-Irrtum entsprachen, generell kurze Beobachtungszeiträume umfassten. In der vorliegenden Arbeit wurde immer nur eine Kategorie pro Zeiteinheit vergeben. Die Zeiteinheit wurde in der Studie auf 30 Sekunden festgelegt. Falls die Explorationsstrategie Versuch-und-Irrtum generell nur kurze Zeiträume einnimmt (d. h., weniger als 15 Sekunden), würde dieses Verhalten weniger häufig kodiert werden als andere Explorationsstrategien. Die geringere Prävalenz der Kategorie wäre somit auf einen Methodenartefakt zurückführbar (vgl. z. B. Rost, 2007, S. 102). Diese Vermutung wird durch die Tatsache unterfüttert, dass im Vergleich zum Vorgehen von van der Linden et al. (2003) in der vorliegenden Arbeit ein längeres Beobachtungsintervall gewählt wurde (30 Sek. statt 20 Sek.). Für alle anderen Kategorien des Messinstruments erwies sich allerdings das gewählte Intervall als angemessen.

Insgesamt lassen die Befunde der Arbeit vermuten, dass die Explorationsstrategie Versuch-und-Irrtum von den Computernutzern aus unterschiedlichen Gründen genutzt wird: Einerseits zeigen solche Personen ein „Trial-and-Error“-Verhalten, welche die Funktionen des Computers gerne spielerisch entdecken. Andererseits kann die Explorationsstrategie Versuch-und-Irrtum auch ein Ausdruck dafür sein, dass sich Personen von der Situation überfordert fühlen und eine niedrige Kompetenz im Umgang mit dem Computer haben (Kang & Yoon, 2008). Im letzten Falle würde es sich eher um eine „Emergency-Reaktion“ der Computernutzer als um ein spielerisches Ausprobieren handeln (vgl. Doerner, 2006; Schaub, 2006). In diesem Fall haben die Computernutzer, die mit Hilfe von Versuch-und-Irrtum explorieren, weniger das Gefühl, den Prozess unter Kontrolle zu haben. Dieses Gefühl ist jedoch eine wichtige Komponente des

Flow-Erlebens (Csikszentmihalyi, 1975; Engeser & Vollmeyer, 2006; Rheinberg, 2006). Welche Personen- oder Situationsvariablen beeinflussen, ob Versuch-und-Irrtum als spielerische Interaktion oder als Notfallreaktion beim Erlernen eines Computers genutzt wird? Die Suche nach Moderatorvariablen könnte eine interessante Fragestellung zukünftiger Studien sein.

Neben der inhaltlichen Abgrenzung der einzelnen Explorationsstrategien voneinander bedarf die *Operationalisierung der Strategiesystematik* insgesamt einer kritischen Reflexion. Dies geschieht im Folgenden anhand des Beispiels der systematischen Explorationsstrategie. Die Beobachtungseinheiten, in denen ein Teilnehmer systematisch explorierte, wurden in einem Summenwert zusammengefasst. Die Häufigkeit unterschiedlicher, systematischer Explorationssschritte innerhalb einer Beobachtungseinheit wurde nicht berücksichtigt. Dass die Geschwindigkeit der Handlungsausführung vernachlässigt wurde, kann die prädiktive Validität der Messung der Explorationsstrategien beeinträchtigt haben: Eine Person, die innerhalb von 30 Minuten durchgängig systematische Strategien anwendete, dabei jedoch sehr langsam vorging, konnte nicht die gleiche Punktzahl beim Aufgabenlösen erzielen, wie eine Person, welche die gleichen systematischen Handlungssequenzen in deutlich kürzerer Zeit ausführte. Die erste Person konnte trotz optimaler Strategienutzung nicht die gleiche Anzahl an Aufgaben in der vorgegebenen Zeit bearbeiten. Der empirische Zusammenhang zwischen den Variablen „Systematische Exploration“ und „Leistung“ könnte infolgedessen geringer sein, als er tatsächlich ist.

Zusätzlich zu den Explorationsstrategien wurde die *Kategorie Unterbrechung* in das Beobachtungssystem aufgenommen. Diese Kategorie ermöglichte es, solche Phasen im Kategoriensystem zu kennzeichnen, in denen keine Bildschirmaktivitäten zu beobachten waren. Die Ursache der jeweiligen Unterbrechung wurde im Kategoriensystem nicht vermerkt. Zu den möglichen Ursachen zählen Arbeitspausen während des Ausfüllens eines Fragebogens, Unterbrechungen aufgrund von „Denkpausen“ und Reflexion über die nächsten Schritte, das Notieren von Ergebnissen sowie die Interpretation der im Output-Fenster angezeigten Inhalte durch den Teilnehmer. Darüber hinaus unterbrachen Teilnehmer ihre Arbeit, wenn sie ratlos waren und wenn die Anforderungen der Aufgabe ihre Fähigkeiten überstiegen. Das bedeutet, dass eine Vielzahl heterogener Ereignisse im Laufe des Explorationsprozesses als „Unterbrechung“ klassifiziert wurde. Dies erschwert die Interpretation der Befunde. In zukünftigen Untersuchungen sollte auch der Grund der Unterbrechung kodiert werden. Möglichkeiten hierzu bieten zum Beispiel Szenarien, in denen die Untersuchungsleiter nähere Informationen zu den Gründen einer Unterbrechung erhalten. Darüber hinaus können Videoaufzeichnungen Aufschluss darüber geben, was der Teilnehmer in den Phasen einer Unterbrechung tut.

Eine weitere kritische Würdigung betrifft die *Generalisierbarkeit der Befunde*. Das Kategoriensystem wurde an einer Stichprobe entwickelt, die ausschließlich aus weiblichen Personen be-

stand. Die Nachteile dieser Stichprobenselektion könnten sich z.B. in der Reliabilität der Explorationsstrategie „Versuch-und-Irrtum“ widerspiegeln. Die Ergebnisse aus Studie 2 und 3 unterstützen weitgehend die Annahme, dass das Messinstrument auf heterogene, gemischtgeschlechtliche Stichproben verallgemeinert werden kann. Dort zeigte sich, dass auch bei männlichen Studienteilnehmern verschiedene Explorationsstrategien differenziert werden konnten. Überdies stand die Häufigkeit der jeweiligen Explorationsstrategie auch bei den männlichen Studienteilnehmern in charakteristischer Beziehung zur Motivation und zur Leistung. Diese Beobachtungen lassen das Fazit zu, dass trotz der Homogenität der Stichprobe ein Kategoriensystem entwickelt werden konnte, welches eine erschöpfende Menge von Kategorien zur Beschreibung des Verhaltens in unterschiedlichen Stichproben beinhaltet.

Ein anderer Aspekt der Generalisierbarkeit betrifft die Frage, ob die Befunde über das Programm SPSS hinaus Gültigkeit besitzen. Das Kategoriensystem wurde speziell zu dem Zweck entwickelt, die Strategiesystematik der Computernutzer beim Explorieren von SPSS zu erfassen. Problemlösestrategien weisen häufig einen bereichsspezifischen Bezug auf (vgl. Boekaerts, 1999; Perels et al., 2005). Die Problemlösestrategien, von denen die Kategorien des Beobachtungssystems abgeleitet wurden, sind jedoch auf verschiedene Problemlösesituationen transferierbar (vgl. Dörner und Schaub, 1994; Schaub, 2006). Vor diesem Hintergrund wird vermutet, dass sich die Strategien von Computernutzern bei einer Software mit ähnlicher Programmstruktur hinreichend ähneln. Dies betrifft nicht nur Software gleicher Programmfamilien wie zum Beispiel Excel von Microsoft und Calc von OpenOffice als zwei Varianten von Tabellenkalkulationsprogrammen. Auch über Programmfamilien hinweg können sich Ähnlichkeiten zeigen (z.B. besitzt sowohl Microsoft Word als auch Microsoft Excel eine Menüleiste, von der aus die Software gesteuert wird). Ob sich die beim Erkunden von SPSS erlernten und geübten Explorationsstrategien auf ähnliche Probleme übertragen lassen, sollte in zukünftigen Studien erforscht werden.

Aufgrund theoretischer Überlegungen wurde die Strategiesystematik in dieser Arbeit über die Auswertung von *Verhaltensspuren* erfasst (vgl. Artelt, 2000; Ericsson & Simon, 1993; Hadwin et al., 2007; Leow & Morgan-Short, 2004). Das neu entwickelte Kategoriensystem erlaubt, die Verhaltensspuren inhaltlich zu klassifizieren. Zimmerman (2008, S. 171) bilanziert mit Hinblick auf die Vor- und Nachteile unterschiedlicher Messverfahren selbstregulierten Lernens: „Collectively, these results indicate that self-reports are often incongruent with trace measures of self-regulatory processes when studied in a specialized learning environment“. Winne und Perry (2000; vgl. auch Hadwin et al., 2007) empfehlen daher, die Verhaltensspuren in Verbindung mit anderen Daten zu interpretieren, die mit anderen Erfassungsmethoden gewonnen wurden. Hierzu zählen zum Beispiel Selbstberichte der Computernutzer sowie think aloud-Protokolle. Die Kombination mehrerer Erfassungsmethoden erlaube eine validere Messung der interessierenden

Konstrukte. In der Konstruktionsphase eines neuen Kategoriensystems können insbesondere think aloud-Protokolle oder Interviews mit stimulated recall dazu beitragen, Probleme bei der Operationalisierung der Strategiesystematik aufzudecken. In zukünftigen Studien zum Erlernen von Computerprogrammen sollte diesen Empfehlungen von Winne und Perry (2000) entsprechend Rechnung getragen werden.

10.1.2 Zur Erklärung von Geschlechtsunterschieden mittels des kognitiv-motivationalen Prozessmodells (Forschungsfrage 2)

Zusammenfassung zentraler Ergebnisse. In Studie 2 und Studie 3 der vorliegenden Arbeit wurde überprüft, inwiefern Geschlechtsunterschiede beim Erlernen eines Statistikprogramms mit Hilfe des kognitiv-motivationalen Prozessmodells erklärt werden können. Entgegen der Erwartungen zeigten sich in beiden Studien keine Geschlechtsunterschiede in der Eingangsmotivation: Männer und Frauen gaben vor Beginn des Aufgabenbearbeitungsprozesses ähnliche Werte hinsichtlich der Herausforderung, des Interesses, der Erfolgswahrscheinlichkeit und der Misserfolgsbefürchtung an. Dies steht in Widerspruch zu Studienergebnissen, die Geschlechtsdifferenzen in der computerspezifischen Motivation (z.B. Computer Self-Efficacy; Computer-ängstlichkeit) erbrachten (Baloglu & Çevik, 2008; Meelissen & Drent, 2008; Shotick & Stephens, 2006). Es zeigten sich lediglich leichte Niveauunterschiede in einzelnen Aspekten der Eingangsmotivation: In beiden Studien berichteten Frauen der Tendenz nach eine höhere Misserfolgsbefürchtung und fühlten sich, entgegen der Erwartungen, etwas stärker herausgefordert als Männer. Die schwachen Geschlechtsunterschiede erreichten jedoch keine statistische Signifikanz. Aufgrund dieser Ergebnisse wurde die Eingangsmotivation bei der empirischen Prüfung der Pfadmodelle nicht als erklärende Variable einbezogen.

Hinsichtlich der aktuellen Motivation sowie der Strategiesystematik erbrachten die beiden Studien inkonsistente Befunde: In Studie 2 nutzten Frauen hypothesenkonform häufiger die un-systematischen Explorationsstrategien und seltener die systematische Explorationsstrategie. Bei Männern zeigte sich ein umgekehrtes Verhältnis. Der Befund korrespondiert mit den Beobachtungen von Beckwith et al. (2006), wonach Frauen Handlungen beim Erlernen eines Computerprogramms häufiger wiederholen. Überdies wiesen Frauen im Vergleich zu Männern in Studie 2 eine ungünstigere aktuelle Motivation auf. Im Flow-Erleben erreichten die Männer in der Studie leicht höhere Werte, jedoch wurde der Geschlechtsunterschied nicht wie erwartet statistisch signifikant. Die Männer konnten allerdings signifikant mehr Aufgaben mit dem Statistikprogramm lösen. In Studie 2 zeigte sich zudem, dass der Einfluss des Geschlechts partiell über die Prozess-

variablen auf die Leistung vermittelt wurde. In Studie 3 ergaben sich für das Flow-Erleben, die aktuelle Motivation und die Leistung keine systematischen Geschlechtsunterschiede.

Um die inkonsistenten Ergebnisse zwischen Studie 2 und 3 erklären zu können, wurde erforscht, in welchen Variablen des kognitiv-motivationalen Prozessmodells sich die Männer und die Frauen in Studie 2 von denjenigen in Studie 3 unterschieden. Der Vergleich ergab, dass die Stichprobenzugehörigkeit die Beziehung zwischen Geschlecht und Strategiesystematik moderierte: In Studie 3 nutzten Frauen und Männer alle Explorationsstrategien (systematisch und unsystematisch) in etwa gleichem Umfang. Anhand der normativen Werte in Tabelle 21 wurde deutlich, dass Frauen in Studie 3 im Vergleich zu Studie 2 die systematische Explorationsstrategie stärker nutzten. Bei Männern zeigte sich eine umgekehrte Tendenz: Sie explorierten in Studie 3 etwas seltener systematisch. In beiden Studien stand das Ausmaß systematischer Exploration in einem positiven Zusammenhang mit der Leistung. Dass sich die Geschlechter in Studie 2 auch in der Leistung am Computer unterschieden, stimmt mit den Befunden bisheriger Studien überein (z.B. Imhof et al., 2007; Roy et al., 2003). Dass sich in Studie 3 Männer und Frauen weder in der Strategiesystematik noch in der Leistung unterschieden, lässt den Schluss zu, dass die Veränderungen in der Strategiesystematik dazu beigetragen haben, dass in Studie 3 zwischen den Geschlechtern geringere Leistungsunterschiede beobachtet wurden. Die Ergebnisse verdeutlichen, dass insbesondere die Erfassung der Prozessvariablen Strategiesystematik mittels des neu konstruierten Beobachtungsinstruments erfolgreich dazu beitragen konnte, die Inkonsistenzen zwischen beiden Studien nachzuvollziehen.

Als Kontrollvariablen wurden in Studie 2 und 3 das Statistikvorwissen sowie die Computerbildung berücksichtigt. Obwohl die Computerbildung in den beiden Studien teilweise unterschiedlich operationalisiert wurde, zeigten sich vergleichbare Ergebnisse. Nach eigenen Angaben hatten die Frauen in Studie 2 und 3 im Vorfeld der Untersuchung die abgefragten Softwareprogramme weniger häufig genutzt als die Männer dies getan hatten. Dieses Ergebnis stützte sich insbesondere auf die Nutzungshäufigkeit von Tabellenkalkulationsprogrammen. Wie bereits in früheren Studien zeigte sich, dass die Computerbildung mit der Strategiesystematik bei der Erkundung eines Computerprogramms zusammenhängt (vgl. Kang & Yoon, 2008): In Problemlöse- und Lernsituationen treten rigide Explorationsstrategien häufig dann auf, wenn der Problemraum als zu groß empfunden wird, das bereichsspezifische Vorwissen gering ist und Selbstregulationskompetenzen im Bereich der Überwachung (Monitoring) des eigenen Verhaltens unzureichend sind (van der Linden et al., 2001, 2003). Dass Frauen in Studie 2 häufiger die rigide Explorationsstrategie nutzten, lässt sich folglich auf deren geringere Computererfahrung zurückführen (vgl. auch Richter, Naumann & Horz, 2001). Im Widerspruch zu dieser Annahme steht jedoch das Ergebnis in Studie 3: Hier zeigten sich statistisch signifikante Geschlechtsunterschiede in der

Computerbildung, jedoch hatten diese keine Auswirkung auf die Häufigkeit, mit der Männer und Frauen rigide explorierten. In beiden Studien konnten Personen, die zuvor häufiger Programme der ausgewählten Softwarefamilien genutzt hatten, mehr Aufgaben korrekt lösen. Im Gegensatz zur Computerbildung korrelierte das Statistikvorwissen nicht mit der Leistung.

Theoretische Einordnung der Ergebnisse. Die Ergebnisse in Studie 3 stimmen nicht mit den zuvor formulierten Annahmen zu Geschlechtsunterschieden beim Erlernen des Statistikprogramms überein. Dennoch stehen sie nicht im Widerspruch zu den Postulaten des *kognitiv-motivationalen Prozessmodells* (Vollmeyer & Rheinberg, 1998, 2000, 2006). Da sich Frauen und Männer weder in der Eingangsmotivation noch in den Prozessvariablen voneinander unterscheiden, sollten sich dem Modell gemäß auch keine Leistungsunterschiede zwischen den Geschlechtern zeigen. Dass die Eingangsmotivation bei Männern und Frauen ein ähnliches Ausmaß hatte, Frauen aber gleichzeitig statistisch signifikant geringere Leistungen erbrachten, entspricht jedoch nicht den theoretischen Vorhersagen auf der Basis des Prozessmodells. Die Prozessvariablen (systematische Explorationsstrategie/Flow) wiesen dagegen die theoretisch erwarteten Beziehungen zur Leistung auf: Hohe Flow-Werte und eine häufige systematische Exploration gingen mit einer hohen Leistung einher. Dass sich in der jüngeren der beiden Studien keine substantiellen Geschlechtsunterschiede in den Untersuchungsvariablen zeigten, kann aus dem Blickwinkel der *Gender Similarities*-Hypothese von Hyde (2005, 2007) interpretiert werden. In der *Gender Similarities*-Hypothese wird hervorgehoben, dass die Ähnlichkeiten zwischen den Geschlechtern in psychologischen Merkmalen und kognitiven Fähigkeiten die Unterschiede zwischen ihnen übertreffen (Hyde & Linn, 2007).

Kritische Würdigung und Forschungsausblick. Zur Erklärung der inkonsistenten Befunde kann einmal auf die oben angeführte *Gender Similarities*-Hypothese rekurriert werden. Darüber hinaus kann das Ergebnis durch *statistische Effekte* bedingt sein. Beide Erklärungsansätze werden im Folgenden diskutiert.

Aus der Perspektive der *Gender Similarities*-Hypothese korrespondieren die Ergebnisse der vorliegenden Arbeit mit denjenigen aktueller Studien zum Lern- und Leistungsverhalten. Diese ergaben, dass Geschlechtsunterschiede beispielsweise im Bereich der Problemlösefähigkeiten nur noch schwach ausgeprägt sind oder sich kontinuierlich verringern (z.B. Blossfeld et al., 2009; Hyde & Linn, 2006; Meece & Painter, 2008; PISA-Konsortium, 2007). Leistungsdifferenzen haben sich stark reduziert oder sind gar nicht mehr nachweisbar (vgl. Hyde & Durik, 2005). Geschlechtseffekte sind in der Regel schwach bis mittelstark (Hyde, 2005). In immer mehr Lern- und Leistungsbereichen scheinen die Ähnlichkeiten zwischen den Geschlechtern zu überwiegen.

Die Annahmen der *Gender Similarities*-Hypothese können auf die Entwicklung von Geschlechtsunterschieden in Motivation, Strategiesystematik und der Leistung beim Erlernen eines

Computers übertragen werden: Zwischen der Datenerhebung der beiden Studien 2 und 3 lag ein Zeitraum von ungefähr drei Jahren. Die Bedeutung des Computers ist im letzten Jahrzehnt auch in Schule und Studium weiter gewachsen (vgl. Blossfeld et al., 2009; Middendorf, 2002; Richter et al., 2010; Schaumburg, 2004). Diese Entwicklung hat dazu geführt, dass Frauen in pädagogischen Settings immer häufiger mit der Nutzung unterschiedlicher Computerprogramme konfrontiert werden. Im Gegensatz zu der im Jahr 2005 durchgeführten Studie 2 waren weibliche und männliche Studierende in Studie 3 gleichermaßen kompetent beim Erlernen und Anwenden eines studienrelevanten Statistikprogramms. Der Befund stimmt mit der Beobachtung von Imhof et al. (2007) überein, wonach sich in der Computernutzung von studienrelevanter Software keine oder nur noch schwache Geschlechtsunterschiede zeigen.

Durch die alltägliche Interaktion mit Computertechnologien könnten im Laufe der Zeit die Vertrautheit mit dem Computer, die Strategiesystematik und auch die computerbezogene Selbstwirksamkeit bei Frauen gestiegen sein. Infolgedessen sollten sich Geschlechtseffekte im Umgang mit dem Computer und in der Leistung über die Kohorten hinweg verringern. Diese Argumentation wird auch von Todman und Day (2006) unterstützt: „With the increasing penetration of computer-based technologies, it may be that masculinisation of the computer is not occurring to the same extent as previously documented“ (S. 865). Dass der Computer weniger stark als zuvor mit maskulinen Attributen assoziiert wird, kann auch in der vorliegenden Arbeit positive Auswirkungen auf Motivation und Strategiesystematik von Frauen beim Erlernen eines Computerprogramms gehabt haben.

Das Ergebnis in Studie 3 ist maßgeblich darauf zurückzuführen, dass Frauen in Studie 3 stärker als ihre Kommilitoninnen in Studie 2 systematisch explorierten. Diese Veränderung führte wiederum zu einer besseren Leistung der Frauen. Die häufige Anwendung der (kognitiven) systematischen Exploration hatte zudem einen positiven Effekt auf die Motivation während des Explorationsprozesses. Dies entspricht der Annahme, dass Geschlechtseffekte in kognitiven Kompetenzen kaum mehr nachweisbar sind. Interessant ist jedoch, dass sich Frauen und Männer in den vorliegenden Studien - zumindest was den Selbstbericht über die eigene Computerbildung betrifft - weiterhin in charakteristischer Weise unterscheiden (vgl. Meece & Painter, 2008).

Alternativ zu der oben angeführten inhaltlichen Erklärung kann die Ursache für die abweichenden Befunde in Studie 2 und 3 auch rein statistischer Natur sein. Der in Studie 2 gemessene Geschlechtseffekt könnte eine extreme Ausprägung des Populationsparameters darstellen, so dass der *Regressionseffekt zur Mitte* in Studie 3 zu einer schwächeren Ausprägung des Geschlechtseffekts geführt hat (vgl. Rost, 2007, S. 103). Die Männer und Frauen der Studie 2 könnten hinsichtlich ihrer Ausprägung in den entscheidenden Untersuchungsvariablen eine Extremgruppe gebildet haben. In diesem Fall hätten sich die Mittelwerte in der Population bei Frauen

und Männern über die Studien hinweg nicht bedeutsam geändert. Lediglich ein Methodenartefakt hätte zu den beobachteten Unterschieden zwischen den Stichproben geführt. Dieser Methodenartefakt kann dabei durch den geringen Stichprobenumfang in Studie 2 ($N = 33$) begünstigt worden sein. Kleinere Stichproben bergen das Risiko einer verzerrten Schätzung (vgl. Bortz, 2005). Die Befunde in Studie 3 hingegen basierten in etwa auf dem dreifachen Stichprobenumfang. Eine wiederholte Ziehung unterschiedlicher Stichproben kann Hinweise darauf geben, inwiefern sich Geschlechtsunterschiede aus Studie 2 an einer größeren Stichprobe replizieren.

Einer kritischen Würdigung bedarf auch der Befund, dass sich die Geschlechter weder in Studie 2 noch in Studie 3 in der *Eingangsmotivation* unterschieden. Dies könnte darauf zurückführbar sein, dass die Anregungsbedingungen vor der Aufgabenbearbeitung die tatsächlichen Aufgabenbedingungen nicht vorwegnehmen konnten. Da die große Mehrheit der Teilnehmer zuvor noch nicht mit dem Statistikprogramm gearbeitet hatte, war es für sie besonders schwierig, ihre Erfolgswahrscheinlichkeit und ihre Misserfolgsbefürchtung valide und reliabel einzuschätzen. Imhof et al. (2007) fassen diese Vermutung folgendermaßen zusammen: „To complete a job with a new software package ... may be hard to imagine for some, so that it is unclear if a person's judgments of his or her capability has a sound point of reference” (S. 2826). Erst im Verlauf der Interaktion mit SPSS erhielten die Teilnehmer Hinweise darauf, ob sie die Aufgabenanforderungen bewältigen können würden. Die Frage, ob sich Männer und Frauen tatsächlich in ihrer Eingangsmotivation beim Erlernen eines neuen Computerprogramms unterscheiden, konnte somit nicht abschließend geklärt werden.

In der vorliegenden Arbeit wurden motivationale und kognitive Variablen herangezogen, um potenzielle Geschlechtsunterschiede beim Computerlernen erklären zu können. *Metakognitive Selbstregulationsprozesse* wie das Monitoring wurden weder in Studie 2 noch in Studie 3 explizit untersucht (vgl. Hasselhorn & Labuhn, 2008; Winne & Hadwin, 2008). Dennoch liefern die Ergebnisse Hinweise darauf, dass Geschlechtsunterschiede in metakognitiven Kompetenzen zu Leistungsunterschieden beitragen. Van der Linden et al. (2001, 2003) weisen auf den Zusammenhang zwischen metakognitiven Kompetenzen, Vorwissen, Strategiesystematik und Leistung hin. Sowohl in Studie 2 als auch in Studie 3 gaben Frauen geringere Vorerfahrungen mit den Computerprogrammen an. Frauen in Studie 2 explorierten gleichzeitig deutlich häufiger rigide als Männer. Die Befunde lassen darauf schließen, dass sich die Frauen in Studie 2 mit der Aufgabe überfordert fühlten, weniger Rückmeldungen des Programms wahrnahmen bzw. diese nicht nutzen konnten, um das eigene Verhalten den Anforderungen der Aufgabe entsprechend flexibel anzupassen. In zukünftigen Forschungsstudien sollte die Beziehung zwischen Geschlecht und der metakognitiven Selbstregulationskompetenzen beim Erlernen eines Computerprogramms jedoch genauer eruiert werden (vgl. Meece & Painter, 2008; Littleton et al., 1998; Roy et al., 2003).

10.1.3 Zur Entwicklung von Motivation und Strategiesystematik im Prozessverlauf (Forschungsfrage 3)

Zusammenfassung zentraler Ergebnisse. Forschungsfrage 3 richtete sich auf die Beziehung der Prozessvariablen Flow und Strategiesystematik untereinander und auf deren Erklärungsbeitrag zur Leistung. In diesem Zusammenhang wurde auch erforscht, wie sich die Strategiesystematik über die Zeit entwickelt. Im Folgenden gehe ich zunächst auf die Entwicklung der Strategiesystematik im Laufe des Bearbeitungsprozesses ein. Anschließend diskutiere ich die Beziehungen zwischen Flow, Strategiesystematik und Leistung im Zeitverlauf.

Die einzelnen Explorationsstrategien unterlagen unterschiedlichen Entwicklungstrends. Zu allen drei Messzeitpunkten in Studie 3 nahm die *systematische Explorationsstrategie* den größten Anteil an allen Explorationsstrategien ein (vgl. van der Linden et al., 2001, 2003). Wie die Varianzanalyse mit Messwiederholung ergab, reduzierte sich der Anteil der systematischen Explorationsstrategie jedoch vom ersten zum zweiten Messzeitpunkt deutlich. Danach blieb er bis zum Ende des Bearbeitungsprozesses relativ konstant.

Der Befund lässt sich folgendermaßen erklären: In den ersten 10 Minuten waren sowohl Teilnehmer mit einer niedrigen als auch mit einer hohen Computerbildung damit beschäftigt, die Benutzeroberfläche von SPSS gleichermaßen systematisch zu explorieren. Dieses Ergebnis kann darauf zurückgeführt werden, dass die Teilnehmer in der Einführung in SPSS gezielt auf die Menüleiste hingewiesen wurden. Überdies ähnelt die Benutzeroberfläche von SPSS derjenigen vieler Standardanwendungsprogramme (z.B. den Textverarbeitungsprogrammen von Microsoft oder OpenOffice). Viele Teilnehmer nutzten ihre Vertrautheit mit diesen Computerprogrammen: Sie transferierten ihre Kenntnisse auf das neue Programm und stellten Analogien her. Die systematische Exploration ging in dieser (ersten) Explorationsphase mit einem positiven Gefühl der Kontrolle und der Passung von Fähigkeit und Anforderungen einher. Dies stimmt mit den theoretischen Erwartungen und empirischen Ergebnissen überein (z.B. Csikszentmihalyi, 1975, 2000; Engeser et al., 2005; Rheinberg, 2006; Schiefele & Rheinberg, 1997). Im Gegensatz zur systematischen Exploration war der Anteil der Kategorie *Unterbrechung* in dieser Phase vergleichsweise niedrig. Die Häufigkeit der Unterbrechungen korrelierte jedoch stark und positiv mit der Leistung. Der Befund lässt sich so interpretieren, dass Unterbrechungen in dieser Phase hauptsächlich genutzt wurden, um Ergebnisse zu notieren.

Im Prozessverlauf differenzierte sich das Explorationsverhalten zwischen leistungsschwachen und leistungsstarken Teilnehmern aus. Die leistungsschwachen Teilnehmer verwendeten im auch solche Explorationsstrategien, die weniger systematisch und damit weniger zielführend waren (z.B. Rigide Exploration, Versuch-und-Irrtum; vgl. Dörner, 1980, 2006; Schaub, 2006). Dies

zeigt sich daran, dass sich der Zusammenhang zwischen Computerbildung und der Häufigkeit der systematischen Exploration über die Messzeitpunkte hinweg verstärkte. Je höher die Computerbildung, desto häufiger zeigten Teilnehmer zu den letzten beiden Messzeitpunkten systematisches Explorationsverhalten. Bei der *rigiden Explorationsstrategie* zeigte sich eine gegenteilige Entwicklung: Je niedriger die Computerbildung vor Beginn der Studienteilnahme, desto häufiger wurde die rigide Explorationsstrategie bei Teilnehmern beobachtet. Die rigide Explorationsstrategie erreichte stets den höchsten Anteil unter den unsystematischen Explorationsstrategien. Ihr Anteil blieb im Gegensatz zur systematischen Exploration über die Zeit hinweg stabil. Die Anteile der anderen, unsystematischen Explorationsstrategien (Versuch-und-Irrtum, Informationssuche) variierten ebenfalls kaum über die Zeit.

Der Anteil der Unterbrechungen stieg über die Zeit deutlich an. Diese Beobachtungen führen zu folgenden Schlussfolgerungen: Leistungsstarke Teilnehmer hatten sich zu Beginn des Prozesses bereits zügig mit der Funktionsweise des Programms vertraut gemacht. In der nächsten Phase unterbrachen die leistungsstarken Teilnehmer ihre Arbeit häufiger, um die Aufgaben auf dem Arbeitsblatt zu bearbeiten. Im Kategoriensystem wurde dies als „Unterbrechung“ klassifiziert. Eine weitere Erklärung für den wachsenden Anteil der Kategorie Unterbrechung zu Messzeitpunkt t2 und t3 ist die folgende: Teilnehmer, welche die Aufgaben nicht bewältigen konnten, wurden zunehmend inaktiv. Infolgedessen waren keine Handlungen mehr auf der Bildschirmoberfläche erkennbar. Es ist denkbar, dass beide Erklärungen gleichzeitig zutreffen: Während der Endphase können sowohl leistungsstarke als auch leistungsschwache Teilnehmer ihre Arbeit aus unterschiedlichen Gründen ungefähr gleichhäufig unterbrochen haben. Für diese Annahme spricht, dass die Anzahl der Unterbrechungen über die Messzeitpunkte hinweg immer schwächer mit der Leistung korrelierte.

Im *Pfadmodell* in Studie 3 wurde die Entwicklung der Strategiesystematik mit der Entwicklung des Flow-Erlebens in Verbindung gesetzt. Darüber hinaus wurde erforscht, ob beide Prozessvariablen zur Erklärung der Leistung beitragen. Die Leistung wurde am besten durch die Häufigkeit der systematischen Exploration zu den Messzeitpunkten t1/t2 und t3 vorhergesagt. Jedoch hatte auch das Flow-Erleben zu Messzeitpunkt t3 einen schwachen, aber ebenfalls günstigen Einfluss auf die Leistung. Das Flow-Erleben leistete in beiden Pfadmodellen (in Studie 2 und Studie 3) dabei einen inkrementellen Beitrag, der über den Effekt der Strategiesystematik hinausging. Insgesamt zeigte sich in den Studien 2 und 3, dass sich die systematische Exploration bei der Bearbeitung der Statistikaufgaben positiv auf das Flow-Erleben auswirkte. Der günstige Effekt der Strategiesystematik trat in Studie 3 insbesondere in den ersten 20 Minuten des Explorationsprozesses auf. In den letzten 10 Minuten zeigte sich dagegen keine signifikante Korrelation zwischen Strategiesystematik und Flow-Erleben.

Die Ergebnisse in Studie 3 ergaben zudem eine mittlere bis hohe *Stabilität der beiden Prozessvariablen* über die Messzeitpunkte hinweg. Personen, die bereits in den ersten 20 Minuten ein starkes Gefühl von Absorbiertheit und Glatttem Verlauf hatten, berichteten dies auch nach den letzten 10 Minuten der Bearbeitungszeit. Der Effekt von Flow wurde so über die Messzeitpunkte hinweg „weitergegeben“. Ein ähnliches Ergebnis zeigte sich für die Strategiesystematik, wobei der Effekt nur schwach bis mittelstark war. Personen, die in den ersten 20 Minuten systematisch explorierten, setzten dieses Verhalten auch eher zum letzten Messzeitpunkt fort. Ähnliche Befunde zur Konsistenz der Strategiesystematik werden von Vollmeyer und Rheinberg (2006) berichtet. Der Effekt wurde dabei partiell über das Flow-Erleben vermittelt: Das systematische Explorieren war begleitet von einem positiven Gefühl, völlig in der Aufgabe aufzugehen. Erste Erfolge, die sich insbesondere bei der Nutzung der systematischen Strategie einstellten, führten bei Männern und Frauen zu der Einschätzung, die Aufgabenanforderungen bewältigen zu können und die Aufgabe unter Kontrolle zu haben. Zum letzten Messzeitpunkt ging die systematische Exploration jedoch nicht mit einem höheren Flow-Erleben einher. Dieser Befund, der im Widerspruch zu den im theoretischen Pfadmodell postulierten Variablenbeziehungen steht, lässt sich folgendermaßen erklären: Durch die hohe Stabilität des Flow-Erlebens über die Zeit, wurde ein Großteil der Varianz des Flow-Erlebens zu Messzeitpunkt t3 bereits durch das Flow-Erleben zu Messzeitpunkt t1/t2 erklärt. Der Einfluss der Strategiesystematik auf das Flow-Erleben war infolgedessen gering ausgeprägt und überschritt nicht die statistische Signifikanzgrenze.

Theoretische Einordnung der Ergebnisse. Die positive Beziehung zwischen Flow-Erleben und Strategiesystematik, die zu den Messzeitpunkten t1/t2 ermittelt wurde, stimmt mit den theoretischen Annahmen überein: Als wichtige Vorbedingung des Flow-Erlebens gilt die Passung von Anforderungen und Fähigkeiten (Engeser & Vollmeyer, 2005; Rheinberg, 2006). Personen, die häufiger systematisch explorierten, wiesen eine höhere Computerbildung auf. Aufgrund ihrer Fähigkeiten und Vorkenntnisse konnten diese Personen die Anforderungen am Computer erfolgreich bewältigen. Aus der Passung von Fähigkeiten und Anforderungen resultierte bei ihnen eine optimale, Flow-förderliche Beanspruchung (Csikszentmihalyi, 1975; Csikszentmihalyi & LeFevre, 1989). Sie fühlten sich stärker absorbiert von der Aufgabe und wiesen eine höhere Konzentration auf (vgl. Vollmeyer & Rheinberg, 2006).

Dass die systematische Explorationsstrategie im Pfadmodell eine Auswirkung auf das Flow-Erleben hatte, lässt sich auch mit einer weiteren Vorbedingung des Flow-Erlebens theoretisch begründen. Generell bieten Lernhandlungen günstige Voraussetzungen für das Erleben von Flow (Csikszentmihalyi & Schiefele, 1993). Nach Csikszentmihalyi (2002) gilt dies insbesondere dann, wenn der Lernende eindeutige Handlungsanforderungen erlebt und unverzügliche Rückmeldungen erhält. Dies war bei Personen, die systematisch explorierten, eher der Fall als bei Per-

sonen, die unsystematisch explorierten. Bei Ersteren manifestierte sich dies unter anderem darin, dass diese Personen auf (Fehler-)Rückmeldungen des Computerprogramms angemessen reagierten und ihr Verhalten erfolgreich variierten. Darüber hinaus konnten die Lernenden Rückschlüsse auf ihre Leistungen ziehen (vgl. Vollmeyer & Rheinberg, 2006). Die systematische Exploration kann durch die zielgerichtete Suche dazu beitragen, dass die Handlungsanforderungen als eindeutig erlebt werden. Gleichzeitig kann sie eine Folge dieser Wahrnehmung sein: Programme werden systematisch erkundet, weil die Handlungsanforderungen als eindeutig erlebt werden. In diesem Fall wären sowohl das Flow-Erleben als auch die systematische Explorationsstrategie Konsequenzen der Passung von Fähigkeiten und Anforderungen. Umgekehrt hatte auch das Flow-Erleben einen bedeutsamen, positiven Einfluss auf die Strategiesystematik. Dieser Effekt wird in der Literatur unter Berufung auf die Annahmen Csikszentmihalyi folgendermaßen erklärt:

A person who is experiencing flow is making full use of his or her capacities and should act at his or her highest level of performance. The experience of flow is mainly characterized by high and effortless concentration. Consequently, flow should facilitate optimal cognitive functioning. (Schiefele & Rheinberg, 1997, S. 276)

Neben der kognitiven Variablen „Strategiesystematik“ trug die motivationale Variable „Flow-Erleben“ auch zur Erklärung von Leistungsunterschieden bei. Dies stimmt mit den Implikationen des kognitiv-motivationalen Prozessmodells (Vollmeyer & Rheinberg, 1998, 2000, 2006) überein. In dem Modell werden unterschiedliche motivationale, kognitive und metakognitive Variablen als proximale Prädiktoren der Leistung postuliert. Auch in Prozessmodellen des selbst-regulierten Lernens werden diese Variablen als wichtige Determinanten des Lernerfolgs aufgeführt (vgl. Zimmerman, 2000a, 2000b, 2008; Wirth & Leutner, 2008). In diesem Zusammenhang ist anzumerken, dass nicht das vollständige Mediationsmodell als Grundlage der Analysen diente. Es wurde lediglich ein Ausschnitt des Modells empirisch getestet, bei dem die Variablen der Eingangsmotivation nicht berücksichtigt wurden. Folglich lassen die Befunde nur eingeschränkte Rückschlüsse auf das Prozessmodell von Vollmeyer und Rheinberg zu.

Kritische Würdigung und Forschungsausblick. Die Befunde werfen eine Reihe offener Fragen auf. Diese offenen Fragen beziehen sich auf die folgenden Themen: 1) Einschränkung der Analysen auf die systematische Explorationsstrategie, 2) alternative Erklärungen für die zeitliche Stabilität der Prozessvariablen Flow, 3) mangelnde Aussagemöglichkeiten zu individuellen Verläufen. Diese Themen werden im Folgenden diskutiert.

Um die Komplexität des Pfadmodells einzuschränken und auch eine stabile Parameterschätzung zu gewährleisten, wurde lediglich die *systematische Explorationsstrategie* in das Modell aufgenommen. Das bedeutet, dass die Aussagekraft des Modells auf die Wirkungsbeziehungen des

systematischen Explorationsverhaltens begrenzt ist. Offen bleibt die Frage, wie sich andere Explorationsstrategien (z.B. Rigide Exploration, Informationssuche) auf das Erleben während des Prozesses sowie auf die Leistung ausgewirkt hätten. Darüber hinaus wurden mögliche Wechselbeziehungen zwischen verschiedenen Explorationsstrategien vernachlässigt. In Hypothese 11 (siehe Kapitel 6.4 und 9.2.3) wurde dagegen der Beobachtung Rechnung getragen, dass Personen während des Prozesses nicht nur eine, sondern unterschiedliche Explorationsstrategien anwenden können. Bei der Analyse der Daten wurde jedoch nicht ein variablenzentrierter, längsschnittlicher, sondern ein auf die Person fokussierter Ansatz verfolgt (vgl. von Eye & Bogat, 2006).

Das empirische Pfadmodell belegt eine mittlere bis starke *zeitliche Stabilität des Flow-Erlebens*. Offen bleibt die Frage, ob dieser Effekt durch Methodenartefakte bedingt war oder substantieller Natur ist. Bei der erneuten Vorgabe der Flow-Kurz-Skala zu zwei (Studie 1) beziehungsweise zu drei Messzeitpunkten (Studie 2 und 3) können Gedächtniseffekte nicht ausgeschlossen werden: Die hohe Stabilität der Messwerte über die Zeit kann durch das kurze Zeitintervall zwischen den Messungen begünstigt worden sein (vgl. Vollmeyer & Imhof, 2007). Der Wiedererkennungseffekt kann dabei die Validität der Messung beeinträchtigt haben: Personen könnten sich daran erinnern, wie sie den Fragebogen zuvor beantwortet hatten. Um die zeitliche Stabilität des Flow-Erlebens näher zu erklären, sollte die Beziehung mit anderen Variablen des Modells noch einmal näher betrachtet werden: Das Flow-Erleben wurde in Studie 2 und 3 in der Mehrheit der Messzeitpunkte von der Strategiesystematik vorhergesagt. Für die Strategiesystematik zeigte sich eine stabile Tendenz über die Messzeitpunkte hinweg. Zusammenfassend deutet dieses Ergebnis darauf hin, dass das Flow-Erleben über die Zeit hinweg stabil bleibt, da es positiv mit dieser Explorationsstrategie assoziiert ist. Wie sensibel das Messinstrument bei mehrmaliger Vorgabe gegenüber tatsächlichen Veränderungen im Flow-Erleben ist, sollte jedoch in zukünftigen Studien mittels eines experimentellen Designs genauer geprüft werden. Auf diese Weise könnten Methodeneffekte von substantiellen, inhaltlichen Effekten separiert werden.

Eine letzte offene Frage betrifft die intraindividuellen Veränderungen über die Zeit: Welche Verläufe sind in der aktuellen Motivation, im Flow-Erleben, in der Strategiesystematik oder in der Leistung bei einzelnen Personen oder innerhalb der Gruppen zu beobachten? Mit Hilfe der in Studie 2 und Studie 3 eingesetzten varianzanalytischen Verfahren konnten keine *Aussagen über individuell unterschiedliche Veränderungsrate*n gemacht werden. Bei der Varianzanalyse wird lediglich die Stabilität bzw. Variabilität von Gruppenmittelwerten im Vergleich zueinander ausgewertet (Bortz, 2005). Zur Modellierung individueller Veränderungen im Zeitverlauf sind beispielsweise Analysen über *Latent growth curve models* nötig (vgl. Cohen et al., 2003, S. 578 ff.).

10.1.4 Zur Identifizierbarkeit von Personenclustern (Forschungsfrage 4)

Zusammenfassung zentraler Ergebnisse. Die Befunde der vorliegenden Arbeit unterstützen die Annahme, dass sich Personen danach klassifizieren lassen, welche Explorationsstrategien sie nutzen. Personen, die als „Systematische Explorierer“, „Unsystematische Explorierer“ oder „Informationssucher“ ähnelten sich darin, wie häufig sie einzelne Explorationsstrategien nutzten. Die Gruppenzugehörigkeit konnte zudem anhand von Außenkriterien validiert werden: Die Kombination von Explorationsstrategien, welche eine Person vorwiegend beim Erlernen von SPSS einsetzte, hatte einen differenziellen Einfluss auf die Höhe des Flow-Erlebens sowie auf die Leistung. In beiden dieser Außenkriterien erreichte das Cluster „Systematische Explorierer“ den höchsten Gruppenmittelwert, wohingegen das Cluster „Unsystematische Explorierer“ den niedrigsten Gruppenmittelwert aufwies. Die Kennwerte des Clusters „Informationssuche“ lagen für beide Variablen jeweils zwischen diesen beiden „Extremen“. Bezüglich der Eingangsmotivation zeigten sich innerhalb der Gruppen keine ähnlichen Merkmalsausprägungen.

Trudel und Payne (1995) unterschieden in ihrer Studie zum exploratorischen Erlernen einer Digitaluhr erfolgreiche und weniger erfolgreiche Lernende. Sie fanden heraus, dass erfolgreiche Lernende weniger häufig ineffektive Handlungen durchführten. Zur Klassifizierung der Lernenden in erfolgreiche und weniger erfolgreiche merken die Autoren an:

There is a weak, but uniform, tendency for the protocols ...of successful explorers to evince those exploration events that are indicative of a more thoughtful, reflective approach. However, it seems that the best predictor of learning success in this experiment is not the amount of ‚good‘ behavior during exploration, but the amount of ‚poor‘ behavior. (S. 325)

In der vorliegenden Arbeit lässt sich diese Beobachtung jedoch nur teilweise bestätigen. Frauen und Männer in Studie 2 unterschieden sich ausschließlich statistisch signifikant in der systematischen Explorationsstrategie. Wie die Ergebnisse der Clusteranalyse zeigen, unterschieden sich die drei Personengruppen ebenfalls systematisch in der Häufigkeit systematischer Exploration. Darüber hinaus zeigten sich bedeutsame Muster in der Nutzung der unsystematischen Explorationsstrategien: Im Vergleich zu allen weiteren Clustern erzielten die Personen im Cluster 3 eine mittlere Lernleistung. Gleichzeitig verbrachten sie deutlich mehr Zeit mit der Suche von Informationen als die systematischen Explorierer (vgl. Cluster 1). Das bedeutet, dass sich in der vorliegenden Arbeit erfolgreiche und weniger erfolgreiche Lernende nicht nur hinsichtlich der systematischen sondern auch hinsichtlich der unsystematischen Explorationsstrategien unterschieden.

Theoretische Einordnung der Ergebnisse. Insgesamt sprechen die Befunde für eine differenzierte Betrachtung der als „suboptimal“ beziehungsweise „unsystematisch“ (van der Linden et al., 2001, 2003; vgl. auch Green & Gilhooly, 1990) eingestuften Explorationsstrategien Rigide Exploration, Versuch-und-Irrtum sowie Informationssuche. Dies gilt insbesondere für die Kategorie Informationssuche. Die Informationssuche kann einerseits zu Einschränkungen in Bezug auf die Leistung führen, wenn eine Einkapselung in das Sammeln von Informationen den Computernutzer von der eigentlichen Aufgabenbearbeitung abhält (vgl. Dörner, 1980; Dörner & Schaub, 1994). Andererseits zeigten die „Informationssucher“ immer noch bessere Leistungen als die „Unsystematischen Explorierer“, die hauptsächlich rigide und mittels Versuch-und-Irrtum explorierten.

Eine übermäßige Informationssuche beeinträchtigte das Flow-Erleben in geringerem Maße als die rigide Exploration. Diese Beobachtung kann dadurch bedingt sein, dass mit der übermäßigen Informationssuche dennoch ein Gefühl der Kontrolle über den Prozess einhergeht. Die Person ist weiterhin der Überzeugung, im Dienste ihres Ziels zu handeln. Das Auffinden dieser Informationen kann als Erfolg des eigenen Handelns eingestuft werden und kann sich günstig auf das Flow-Erleben auswirken. Dies kann auch dann der Fall sein, wenn die erlangten Informationen nicht dem Aufgabenziel dienen. Gleichzeitig kann die hohe Involviertheit der Person in der Informationssuche die Zeitwahrnehmung - als Aspekt des Flow-Erlebens - beeinträchtigen. Die Zeit während der Informationssuche verging dann „wie im Flug“. Das spricht dafür, dass die „Informationssucher“ trotz geringerer Leistung die Aufgabenbearbeitung als positiv erlebten.

Kritische Würdigung und Forschungsausblick. Ein Kritikpunkt richtet sich auf die *Operationalisierung der Kategorie Informationssuche*, die das entscheidende Merkmal für die Zugehörigkeit zum dritten Cluster bildete („Informationssucher“; siehe Abschnitt 9.2.3). Diese Kategorie wies in Studie 1 eine sehr hohe Beobachterübereinstimmung auf (siehe Abschnitt 10.1.1). Das „Verharren“ des Computernutzers bei der Informationssuche im Hilfemenü wurde von beiden Ratern gleichermaßen als ein solches erkannt und entsprechend kodiert. In der Kategoriedefinition wurden jedoch folgende Fragen vernachlässigt: 1) Ab welchem Zeitraum wird eine Informationssuche im Hilfemenü als leistungsabträgliches „Verharren“ gewertet? 2) Wie lässt sich inhaltlich eine zielführende von einer nicht zielführenden Informationssuche unterscheiden (vgl. Dörner & Schaub, 1994; Funke, 2006b)? Die erste Frage betrifft dabei ein quantitatives und die zweite Frage ein qualitatives Bestimmungsmerkmal der Kategorie Informationssuche.

In der Version des Kategoriensystems, welche den Analysen dieser Arbeit zugrunde lag, wurden die beiden genannten Aspekte unzureichend beachtet. Das bedeutet, dass die beiden Rater mit einer hohen Messgenauigkeit erkennen konnten, dass es sich bei dem Verhalten eines Teilnehmers um eine Suche nach Informationen handelte. Es wurde jedoch nicht kodiert, ob die

Informationssuche des Teilnehmers tatsächlich als „übermäßig“ und „nicht zielführend“ gewertet werden kann (vgl. Schaub, 2006). Folglich kann die Informationssuche auf der Basis der gewählten Operationalisierung nicht eindeutig als eine unsystematische Explorationsstrategie eingestuft werden. Viel wahrscheinlicher ist es, dass die Informationssuche sowohl zielführend als auch wenig zielführend sein kann. Diese Ausführungen liefern damit eine weitere Erklärung, weshalb in der Clusteranalyse neben den „Systematischen Explorierern“ und den „Unsystematischen Explorierern“ ein weiteres Cluster identifiziert wurde. Die Clusterlösung sollte vor diesem Hintergrund vorsichtig interpretiert werden. In zukünftigen Studien sollten die Bestimmungsmerkmale der Kategorie genauer differenziert werden.

Die Ergebnisse der Clusteranalyse liefern trotz kritischer Aspekte interessante Implikationen für die pädagogisch-psychologische Praxis des Computerlernens sowie für zukünftige Forschungsstudien: Dass es gelungen ist, Gruppen von Computernutzern voneinander abzugrenzen, unterstreicht die Bedeutung interindividueller Unterschiede beim Computerlernen. Die Teilnehmer von Studie 3 wiesen bezüglich der Nutzung unterschiedlicher systematischer und unsystematischer Explorationsstrategien charakteristische Muster auf. Diese Muster (repräsentiert in den Personenclustern) standen in unterschiedlicher Weise in Zusammenhang mit der Leistung sowie dem Flow-Erleben beim Computerlernen.

Die Erkenntnisse zu interindividuellen Unterschieden können genutzt werden, um individuell zugeschnittene Trainingsprogramme zu entwickeln. Ziel der Interventionsmaßnahmen sollte dabei sein, die Häufigkeit unsystematischer Explorationsstrategien zu reduzieren und demgegenüber die Häufigkeit der systematischen Explorationsstrategie zu erhöhen. Um die Effizienz und die Effektivität des Trainingsprogramms zu steigern, könnte zunächst eine diagnostische Prüfung anhand einer Arbeitsprobe erfolgen. Nachdem der Trainingsteilnehmer einem der drei Cluster zugeordnet wurde, könnte sich eine an seinen Voraussetzungen orientierte Interventionsmaßnahme anschließen. Darüber hinaus gibt es Hinweise, dass die Explorationsstrategien mit der Computerbildung kovariieren (Kang & Yoon, 2008; vgl. auch Richter et al., 2010; Wild, 2000). Informationen zur Computerbildung könnten ebenfalls zur Klassifizierung der Trainingsteilnehmer einbezogen werden. Die praktischen Implikationen, die sich aus der vorliegenden Arbeit ergeben, sind in Abschnitt 10.3 näher beschrieben.

10.2 Studienübergreifende Überlegungen

In diesem Abschnitt werden die methodischen Stärken und Grenzen der Arbeit kritisch diskutiert. Diese sind für alle vier Forschungsfragen gültig.

Methodische Stärken der Arbeit. Im Gegensatz zu früheren Studien selbstregulierten Lernens (vgl. Artelt, 2000; Hadwin et al., 2007; Jamieson-Noel & Winne, 2003) wurden die eingesetzten Explorationsstrategien nicht über subjektive Selbstberichte erfasst. Stattdessen wurden in Anlehnung an van der Linden et al. (2001, 2003; siehe auch Debowski et al., 2001; Kang & Yoon) *Verhaltensmaße* entwickelt, mit denen das Auftreten verschiedener Explorationsstrategien quantifiziert werden konnte. Eine weitere Stärke ist die Verwendung eines *längsschnittlichen Designs*. Insbesondere in Studie 3 erlaubte dieses, Motivations- und Verhaltensunterschiede im Verlauf des Lernprozesses genauer zu bestimmen. Eine Stärke von Studie 3 stellt zudem der vergleichsweise *große Stichprobenumfang* dar. Frühere Studien, in welchen die Strategiesystematik und die Leistung beim Erlernen von Computerprogrammen untersucht wurden, wiesen zum Teil erheblich kleinere Stichprobenumfänge auf (Carroll et al., 1985: $N = 12$; Green & Gilhooly, 1990: $N = 10$; Kang & Yoon, 2008: $N = 60$; Trudel & Payne, 1995: $N \leq 48$; Trudel & Payne, 1996: $N \leq 32$; van der Linden et al., 2001: $N = 19$; 2003: $N = 68$).

Methodische Grenzen der Arbeit. Eine erste methodische Grenze betrifft die *Art der Stichprobenziehung* sowie den *Stichprobenumfang*. Der Charakter der in den drei Studien realisierten Gelegenheitsstichproben schränkt die Generalisierbarkeit der Befunde ein (vgl. Rost, 2007): Die Ziehung der Stichprobe, an der die Verhaltensdaten gewonnen wurden, erfolgte nicht zufällig. Folglich kann eine Verzerrung der Stichprobe durch Selbstselektionsprozesse zustande gekommen sein: Für Personen mit einer erhöhten Computerängstlichkeit (vgl. Chua et al., 1999) könnte die Wahrscheinlichkeit höher gewesen sein, dass sie sich gegen die Teilnahme an der Studie zum Erlernen eines Computerprogramms entschieden haben. Umgekehrt ist davon auszugehen, dass insbesondere hoch motivierte Personen an den Studien teilnahmen. Die Befunde der Studien sind vor diesem Hintergrund eventuell nicht auf die Population der Wirtschaftswissenschaften-Studierenden verallgemeinerbar. Obwohl der Stichprobenumfang in Studie 3 denjenigen in früheren Studien übertraf (siehe oben), ist Folgendes kritisch anzumerken: Die Prüfung des komplexen Zusammenhangs über mehrere Messzeitpunkte hinweg wurde auf einer relativ geringen Datenbasis vorgenommen. Eine Replikation der Modelltestung auf der Basis größerer Stichproben ist empfehlenswert (vgl. Kline, 2005).

Das *quasi-experimentelle Design der Studien 2 und 3* erschwert die kausale Interpretation der vorgefundenen Zusammenhänge zwischen Geschlecht, Motivation und Leistung. Da Störvariablen bzw. Drittvariablen im Rahmen des gewählten Designs nur ungenügend kontrolliert werden konnten (Stichwort: mangelnde Randomisierung), ist die eindeutige Rückführung der Ergebnisse in den Abhängigen Variablen auf das Geschlecht bzw. die mit dem Geschlecht korrelierten Motivationsunterschiede nicht einwandfrei möglich (vgl. Rost, 2007, S. 124; Tarling, 2009, S. 141). Es ist nicht vollkommen auszuschließen, dass die gemessenen Mittelwertsunterschiede zwischen

Männern und Frauen sowie die empirischen Variablenzusammenhänge durch Drittvariablen beeinflusst wurden. Hierzu zählen beispielsweise Selbstselektionseffekte, die durch die Wahl einer Gelegenheitsstichprobe bedingt sind (siehe oben).

Anders als bei Vollmeyer und Rheinberg (1998, 2006) wurde in der vorliegenden Arbeit nicht zwischen einer Lernphase (oder Explorationsphase) und einer Anwendungsphase unterschieden. Die Lernenden hatten die Aufgabe, das Computerprogramm anhand vorgegebener Aufgaben zu explorieren und diese Aufgaben mittels des Programms zu lösen. Die *mangelnde Trennbarkeit zwischen Wissenserwerb und Wissensanwendung* spiegelt sich auch in den Leistungsvariablen „Leistung_Arbeitsblatt“ und „Leistung_SPSS“ wider. Die Leistungspunkte, die bei der Aufgabenbearbeitung mit SPSS erzielt wurden, basieren sowohl auf einem erfolgreichen Wissenserwerb sowie einer erfolgreichen Wissensanwendung. Aufgrund dieser Konfundierung konnten beide Prozesse nicht getrennt voneinander analysiert werden. Zukünftige Forschungsarbeiten sollten sich der Frage widmen, wie der Lernende Wissen über die Funktionsweise eines Computerprogramms konstruiert und wie er das Wissen anwendet.

Um die aktuelle Motivation während des Lernprozesses zu erfassen, wurde eine Kurzversion des Fragebogens zur aktuellen Motivation (FAM; Rheinberg et al., 2001) entwickelt. Die *Verkürzung des Fragebogens zur aktuellen Motivation* ist kritisch zu bewerten. Die Messgenauigkeit der jeweils zwei Items pro Dimension konnte nur über die Retestrelabilität bestimmt werden. Problematisch an diesem Koeffizienten ist, dass sowohl Merkmalsfluktuationen (wahre Werte im Sinne der Klassischen Testtheorie) als auch Messfehler die Höhe dieses Koeffizienten beeinflussen. Darüber hinaus können Gedächtniseffekte die Stabilität der Werte beeinflussen und den Retestrelabilitäts-Koeffizienten künstlich erhöhen. Vollmeyer und Rheinberg (2003, S. 289) diskutierten dies in ihrer Studie. Eine wiederholte Fragebogenbearbeitung beeinflusst den Autoren zufolge den Lernprozess und die Reflexion des Fragebogens an sich. Konsistenzbestrebungen könnten die Folge sein. Um die Messgenauigkeit des Instruments zu erhöhen, wäre eine Selektion von mindestens drei Items pro Dimension für die neu gebildete Kurzversion des FAM vorteilhaft. Erst dann könnten alternative Reliabilitätsmaße sinnvoll berechnet werden. Dabei muss berücksichtigt werden, dass sich die Messung der aktuellen Motivation nicht störend auf Erleben und Verhalten der Teilnehmer bei der Aufgabenbearbeitung auswirkt.

Ein weiterer Kritikpunkt betrifft die *Bildung eines Differenzwertes für die aktuelle Motivation*. Die vier Aspekte der Eingangsmotivation (Erfolgswahrscheinlichkeit, Misserfolgsbefürchtung, Interesse, Herausforderung) wurden als unabhängige Dimensionen konzeptualisiert (Rheinberg et al., 2001; Vollmeyer & Rheinberg, 2003, 2006). In der vorliegenden Arbeit wurden sie nach dem von Heckhausen (1963) vorgeschlagenen Prinzip in einem Differenzwert für die Nettomotivation zusammengefasst. Das gewählte Vorgehen hat inhaltliche und messtheoretische

Schwächen: Ein Nachteil der Differenzwert-Bildung ist, dass keine differenzierten Aussagen zur Ausprägung sowie zu Veränderungen in der Höhe einzelner Motivationsaspekte (z.B. Erfolgswahrscheinlichkeit und Herausforderung) getroffen werden konnten. Eine Modellierung der aktuellen Motivation auf der Ebene latenter Variablen stellt dabei eine interessante Alternative für zukünftige Studien dar (z.B. Brown, 2006). Eine weitere Schwäche des gewählten Vorgehens ist, dass die Bildung von Differenzwerten zu einer geringen Reliabilität der Skalen führen kann.

Die *Kontrolle des Vorwissens und dessen Operationalisierung* muss ebenfalls kritisch diskutiert werden. Bezüglich des Vorwissens wurde in der vorliegenden Arbeit zwischen den Statistikkenntnissen und der Computerbildung der Teilnehmer unterschieden. Die Statistikkenntnisse wurden in allen drei Studien mittels eines ad hoc entwickelten Multiple-Choice-Statistiktests erfasst. Die Inhalte des Tests wurden an die Stichprobe adaptiert. Der Test wies über die Studien hinweg eine niedrige interne Konsistenz auf. Zudem korrelierte das Statistikwissen nur in der ersten Studie an den Psychologie-Studentinnen mit der Leistung. Eine eindeutige Schlussfolgerung, ob der Effekt durch die Unterschiedlichkeit der Stichproben oder durch Effekte des Messinstruments bedingt ist, konnte nicht eindeutig geklärt werden. Der Statistiktest sollte bezüglich dieser Fragen sorgfältig getestet und gegebenenfalls modifiziert werden. Die Computerbildung wurde in erster Linie über zwei Skalen des INCOBI (Richter et al., 1999) erfasst. Im Falle der Skala PRACOWI wurden einzelne Items nach inhaltlichen und statistischen Kriterien ausgewählt. Dies hatte zur Folge, dass die Güte des Instruments beeinträchtigt wurde und dass die Vergleichbarkeit mit der Originalversion der Skala nicht mehr gegeben war. Darüber hinaus wurden die beiden Skalen nicht in allen drei Studien zur Kontrolle der Computerbildung eingesetzt. Trotz der mangelnden Reliabilität des Instruments korrelierte die damit erfasste Computerbildung mit der Leistung. Dies unterstreicht die Bedeutung der ausgewählten Vorwissensvariablen zur Computerbildung. In zukünftigen Studien sollte darauf geachtet werden, bereits etablierte Messinstrumente einzusetzen, um die genannten Vorwissensvariablen reliabel und valide erfassen zu können.

10.3 Praktische Implikationen

Die vorliegende Arbeit verdeutlicht, dass sich Individuen hinsichtlich ihrer Defizite und Kompetenzen im selbstregulierten Erlernen eines Computerprogramms unterscheiden - auch wenn sich zwischen den Studien keine konsistenten Geschlechtsunterschiede zeigten. Das selbstregulierte Lernen und Problemlösen kann durch gezielte Maßnahmen gefördert werden (z.B. Cleary & Zimmerman, 2004; Dignath & Büttner, 2008). Interventions- oder Trainingsprogramme richten sich auf die Förderung unterschiedlicher Komponenten selbstregulierten Lernens (Boekaerts & Corno, 2005). Welche Komponenten dabei im Vordergrund stehen, ergibt sich aus den

Annahmen des Modells selbstregulierten Lernens, welches der Maßnahme zugrunde gelegt wird. Perels et al. (2005) zufolge sind Fördermaßnahmen dann besonders effektiv, wenn sie den Lernenden eine größere Verantwortung für den eigenen Lernprozess einräumen. Darüber hinaus ist entscheidend, dass Selbstregulationskompetenzen auf dem Gebiet gelernt werden, in dem sie (später) eingesetzt werden. Selbstregulationskompetenzen sollten von dem Lernenden auch auf andere ähnliche Lernaufgaben und Probleme transferiert werden können.

Aus den genannten Empfehlungen lassen sich folgende Implikationen für das selbstregulierte Erlernen eines Computerprogramms ableiten: Computernutzer sollten ihren eigenen Lernprozess aktiv gestalten können. Eine aktive Beteiligung des Lernenden beim Erwerb von Selbstregulationskompetenzen im Umgang mit dem Computer wird als effektiv und motivierend angesehen. Eine Möglichkeit hierfür bieten solche Ansätze, welche dem Lernenden ermöglichen, neue Computersoftware unter Anleitung zu erlernen (vgl. Bell & Kozlowski, 2008; Carroll & Mack, 1983; Keith et al., 2010). Dieser Ansatz wurde auch in der folgenden Arbeit verfolgt. Darüber hinaus sollte die Strategiesystematik, die ein Computernutzer bei der Erkundung eines spezifischen Computerprogramms erlernt, auch bei der Anwendung eines weiteren Computerprogramms hilfreich sein. Das bedeutet, dass der Lernende deklaratives und prozedurales Wissen erwerben sollte, das er auf eine Transferaufgabe übertragen kann. Dies soll den Lernenden befähigen, zukünftige Anforderungen und Aufgaben am Computer selbst meistern zu können.

Förderung kognitiver Strategien. Interventionsmaßnahmen können das aktive, selbstgesteuerte Erlernen neuer Computeranwendungen fördern, in dem der angemessene Einsatz kognitiver Strategien erlernt wird (Boekaerts, 1997). Zu diesen kognitiven Strategien zählen auch die Explorationsstrategien. Eine Möglichkeit, den Einsatz dieser Strategien zu fördern, besteht darin, Computernutzer auf eine optimale, systematische Explorationsstrategie hinzuweisen und Gelegenheit zu bieten, diese zu üben (Vollmeyer et al., 1996; Vollmeyer & Rheinberg, 1999).

Computernutzer sollten in dem Strategietraining jedoch nicht nur lernen, welche Strategien sie anwenden können, wie Dignath und Büttner (2008) betonen:

Strategy training should integrate information about *how* to use several strategies, and about the *conditions* under which these strategies are most useful Besides the instruction of several types of strategies, students should acquire knowledge about how, when, why, and where to apply these strategies. (S. 236)

Eine Strategie kann unter bestimmten Bedingungen optimal sein und unter anderen dagegen weniger zielführend (Wirth & Leutner, 2008). Insofern sind bei der Planung von Interventionen die Lernsituation und die genauen Aufgabenanforderungen stets mit zu berücksichtigen. Beispielsweise könnte neben der systematischen Explorationsstrategie unter bestimmten Beding-

ungen auch eine zielgerichtete Informationssuche nützlich sein. Wenn der Computernutzer zu Beginn des Explorationsprozesses erste Erkenntnisse über die Funktionen des Programms gesammelt hat, kann er seine Erkundungen fortsetzen, in dem er gezielt Hypothesen zur Funktionsweise des Programms testet.

Förderung der Motivation. Wie die vorliegende Arbeit gezeigt hat, tragen motivationale Variablen wie das Flow-Erleben zu einem erfolgreichen Erlernen und Anwenden eines neuen Computerprogramms bei (vgl. Ghani & Deshpande; Rheinberg & Trapp, 2006; Rheinberg & Vollmeyer, 2003). Das Arbeiten mit dem Computer erfüllt viele Voraussetzungen, die das „optimale“ Gefühl des Aufgehens in einer glatt laufenden Tätigkeit ermöglichen (Rheinberg, 2006; Rheinberg & Trapp, 2006): In der Freizeit können Computernutzer aus einer Vielzahl von Aufgaben selbstbestimmt wählen, was sie am Computer tun möchten. Dabei haben sie häufig die Möglichkeit, das Schwierigkeitsniveau der Aufgabe so zu bestimmen, dass es mit ihren Fähigkeiten übereinstimmt. Im Studium und im Beruf bietet der Computer zahlreiche Lernmöglichkeiten. Programme, die dem Computernutzer ein unmittelbares, eindeutiges Feedback zur Verfügung stellen, fördern dabei das Kompetenzerleben. Auf diese Weise bieten auch alltägliche Aufgaben – wie das Bearbeiten von Statistikaufgaben am Computer - die Chance, als Flow-förderlich erlebt zu werden:

Is enjoyment of work unique to people doing creative tasks, or can everyone experience it if some set of favorable conditions is met? If everyone can experience such enjoyment, then boring everyday tasks might also be turned into enjoyable and meaningful activities. (Csikszentmihalyi, 1975, S. xii, Preface).

Beim Erlernen eines neuen Computerprogramms treffen Computernutzer häufig auf Probleme, Hindernisse oder Fehler, wenn sie eine konkrete Aufgabe mit der neuen Software bearbeiten. Erfahrene Computernutzer lösen aufkommende Probleme, in dem sie die Bedingungen, welche zu dem Problem geführt haben, systematisch variieren und explorieren (van der Linden et al., 2001; Vollmeyer et al., 1996). Dagegen kann die rigide Explorationsstrategie als ein Ausdruck einer erlernten Hilflosigkeit des Computernutzers verstanden werden (vgl. Dickhäuser & Stiensmeier-Pelster, 2002). Mit Hilfe von Reattributionstrainingsprogramme kann das mit der erlernten Hilflosigkeit einher gehende ungünstige Attributionsmuster verändert werden und Handlungsspielräume aufgezeigt werden (Dweck, 1975). In *Error-Management*-Programmen werden Computernutzer explizit aufgefordert, Fehler bei der Computernutzung zu begehen und aus diesen zu lernen (Keith & Frese, 2008).

Interventionsmaßnahmen können auch darauf abzielen, die computerbezogene Selbstwirksamkeitserwartung günstig zu beeinflussen (Chua et al., 1999; Deng et al., 2004; Meelissen &

Drent, 2008). Computernutzer, die ihre eigenen Computerfähigkeiten positiv einschätzen, zeigen ein höheres Durchhaltevermögen, wenn sie mit Hindernissen bei der Problemlösung am Computer konfrontiert sind. Eine höhere Selbstwirksamkeitserwartung beeinflusst außerdem die Strategiesystematik beim Lernen positiv und geht mit einer höheren Leistung einher (Bandura, 1997). Personen, die eine hohe Selbstwirksamkeitserwartung berichten, schätzen häufig ihren eigenen Lernprozess genauer ein und setzen effektivere Selbstregulationsstrategien ein (Pajares, 2002). Wie van der Linden et al. (2003) betonen, wird bei der rigiden Exploration der eigene Lernprozess durch den Lernenden nur unzureichend überwacht und reflektiert. Der Computernutzer verharnt trotz negativer Rückmeldung bei Lösungswegen, die nicht effektiv sind. Das eigene Lernverhalten wird den Erfordernissen der Lernaufgabe nicht in angemessener Weise angepasst (Zimmerman, 2000a, b). Vor diesem Hintergrund erscheinen Maßnahmen, welche die computerbezogene Selbstwirksamkeit fördern, insbesondere für Personen hilfreich, welche zu einer rigiden Exploration neigen.

Förderung der Metakognition. Neben kognitiven Strategien betonen Winne und Hadwin (2008) in ihrem Modell des selbstregulierten Lernens die Rolle metakognitiver Komponenten. Interventionsmaßnahmen auf der Basis dieses Modells sind vornehmlich darauf gerichtet, metakognitive Strategien zu Planung, Monitoring und Evaluation zu fördern (Boekaerts, 1999; Zimmerman, 2008). Mangelnde metakognitive Kompetenzen zeigen sich am Computer unter anderem darin, dass ineffektive Handlungsschritte häufig wiederholt werden und Fehlermeldungen unzureichend beachtet werden (van der Linden et al., 2001; vgl. auch Dörner, 1980). In der vorliegenden Arbeit nahm die rigide Exploration den größten Anteil unter allen unsystematischen Explorationsstrategien ein. Um rigide Explorationsweisen zu reduzieren, empfehlen Trudel und Payne (1996) Interventionen, die den Reflexionsprozess während der Interaktion mit der Software anregen. Die Reflexion des eigenen Handelns am Computer ist eine wichtige Voraussetzung dafür, das „innere, mentale Modell“, welches sich der Computernutzer von der Funktionsweise des Programms gebildet hat, an der Realität zu überprüfen (Schaub, 2006). Der Computernutzer wird so in die Lage versetzt, Handlungsfehler zu erkennen, sein Verhalten den äußeren Anforderungen anzupassen und neue, effektive Lösungswege zu suchen (Zimmerman, 2002, 2008). Auch auf der Grundlage des von Dörner (2006) vorgeschlagenen Phasenmodells der Handlungsregulation könnten Computernutzer zu einem systematischen Vorgehen angeregt werden: Vor Beginn der Aufgabenlösung sollten Ziele formuliert werden, Hypothesen über die Funktionsweise des Computerprogramms gebildet werden und nach der Durchführung einzelner Befehle sollten die Computernutzer ihre Handlungsergebnisse evaluieren.

Fördermaßnahmen in unterschiedlichen Lernsettings. Die Förderung des selbstregulierten Erlernens eines neuen Computerprogramms sollte Dignath und Büttner (2008) zufolge in

Schule und Studium in die Lehrveranstaltungen integriert sein. Dabei sollte Schülerinnen und Schülern im Unterricht die Kompetenz vermittelt werden, gängige Probleme am Computer selbstständig lösen zu können (Schaumburg, 2004). Da es nicht immer möglich ist, Schülerinnen und Schüler in jede spezifische Software einzuführen, sollten allgemeine Selbstregulationsfähigkeiten erlernt werden. Diese erlauben es den Schülerinnen und Schülern, sich Programmkenntnisse selbst anzueignen. Initiativen, bei denen Schülerinnen und Schüler für mehrere Jahre mit einem Laptop ausgestattet wurden, erbrachten positive Ergebnisse (Schaumburg, 2004): Das Computer- und Internetwissen der Mädchen konnte in einem Modellversuch deutlich gesteigert werden. Die Kluft zwischen Mädchen und Jungen im deklarativen Computerwissen konnte deutlich reduziert werden. Auf die computerbezogene Selbstwirksamkeitserwartung der Mädchen wirkte sich die Computerausstattung jedoch nicht aus.

Computerbezogene Selbstwirksamkeitserwartungen von Mädchen und Jungen können im Rahmen des Schulunterrichts durch stellvertretende Erfahrung verstärkt werden (Dignath & Büttner, 2008; Pajares, 2002; Zimmerman, 2002, 2008). Gleiches gilt für den Umgang mit dem Computer in universitären Lernsettings. Die Beobachtung eines Modells kann Schülern und Studierenden helfen, eine angemessene Strategiesystematik zu erlernen. Gleichzeitig sollte die Lehrperson weitere Handlungskompetenzen vermitteln. Dazu zählen Kompetenzen, wie sich Lernende spezifische Ziele bei der Nutzung des Computers setzen können und wie sie auf Fehlermeldungen des Computers zu reagieren (vgl. Boekaerts & Corno, 2005).

Die in diesem abschließenden Kapitel beschriebenen Interventionsansätze stellen einige Möglichkeiten dar, um Mädchen und Jungen beziehungsweise Frauen und Männer zu befähigen, sich Computerprogramme selbstständig zu erschließen. Insbesondere kognitive Kompetenzen (z. B. die Strategiesystematik) und metakognitive Kompetenzen tragen dazu bei, dass Computerprogramme erfolgreich und selbstständig erlernt werden können. Gleichzeitig ist die kompetente Strategienutzung häufig begleitet von Flow-Zuständen während der Beschäftigung mit dem Computer. Zusammenfassend stellen Kompetenzen zum selbstregulierten Lernen die entscheidenden Voraussetzungen dafür dar, dass sich Frauen und Männer gleichermaßen Computertechnologien zu ihren eigenen Gunsten aneignen können und so an den gesellschaftlichen Errungenschaften teilhaben können (vgl. Bebel, 1929/1994).

11. Literaturverzeichnis

- Aellig, S. (2004). *Über den Sinn des Unsinn. Flow-Erleben und Wohlbefinden als Anreize für autotelische Tätigkeiten*. Münster: Waxmann.
- Ainley, M., Hidi, S. & Berndorff, D. (2002). Interest, learning and the psychological processes that mediate their relationship. *Journal of Educational Psychology*, 94, 545-561.
- Allison, P. D. (2003). Missing data techniques for structural equation models. *Journal of Abnormal Psychology*, 112, 545-557.
- Amelang, M. & Zielinski, W. (2002). *Psychologische Diagnostik und Intervention* (3. Aufl.). Berlin: Springer.
- American Psychological Association (Ed.). (2001). *Publication Manual of the American Psychological Association*. Washington, D.C.: APA.
- Anderson, J. R. (2007). *Kognitive Psychologie. Eine Einführung* (6. Aufl.). Heidelberg: Spektrum.
- Aronson, J. A. & Steele, J. C. (2005). Stereotypes and the fragility of academic competence, motivation, and self-concept. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 436-456). London: Guilford Press.
- Artelt, C. (1999). Lernstrategien und Lernerfolg - Eine handlungsnaher Studie. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31, 86-96.
- Artelt, C. (2000). Wie prädiktiv sind retrospektive Selbstberichte über den Gebrauch von Lernstrategien für strategisches Lernen? *Zeitschrift für Pädagogische Psychologie*, 14, 72-84.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64, 359-372.
- Bacher, J. (1994). *Clusteranalyse. Anwendungsorientierte Einführung*. München: Oldenbourg.
- Bachmann, G. (2009). *Zielorientierungen und aktuelle Motivation: eine Integration im Kontext des selbstregulierten Lernens*. Digitale Dissertation, Universität Frankfurt a. M..
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (10. Aufl., Kapitel 6: Strukturgleichungsmodelle, S. 337-424). Berlin, Heidelberg: Springer.
- Bakeman, R. (2000). Behavioral Observations and Coding. In H. T. Reis & C. K. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 138-159). New York: Cambridge Uni-

- versity Press.
- Baloğlu, M. & Çevik, V. (2008). Multivariate effects of gender, ownership, and the frequency of use on computer anxiety among high school students. *Computers in Human Behavior*, 24, 2639-2648.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122-147.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, 42, 815–824.
- Bebel, A. (1929/1994). *Die Frau und der Sozialismus* (Neusatz der 1929 erschienenen Jubiläumsausgabe, 3. Aufl.). Bonn: J. H. W. Dietz Nachfolger.
- Beckmann, J. & Heckhausen, H. (2006). Motivation durch Erwartung und Anreiz. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (3. Aufl.) (S. 105-142). Heidelberg: Springer.
- Bielaczyc, K., Pirolli, P. & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge. *Cognition and Instruction*, 13, 221-252.
- Blossfeld, H.-P., Bos, W., Hannover, B., Lenzen, D., Müller-Böling, D., Prenzel, M. & Wößmann, L. (2009). Geschlechterdifferenzen im Bildungssystem. Verein der Bayerischen Wirtschaft e.V. (Hrsg.), *Jahresgutachten 2009 des Aktionsrats Bildung* (S. 95-109). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Boekaerts, M. (1986). The measurement of state and trait motivational orientation: Refining our measures. In J. H. L. van den Bercken, E. E. J. De Bruyn & T. C. M. Bergen (Eds.), *Achievement and task motivation* (pp. 229-245). Berwyn: Swets North-America.
- Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist*, 1, 100-112.
- Boekaerts, M. (1997). Self-regulated learning. A new concept embraced by researchers, policy

- makers, educators, teachers, and students. *Learning and Instruction*, 7, 161-186.
- Boekaerts, M. (1999). Self-regulated learning: where we are today. *International Journal of Educational Research*, 31, 445-457.
- Boekaerts, M. (2002). The on-line motivation questionnaire: A self-report instrument to assess students' context sensitivity. In P. R. Pintrich & M. L. Maehr (Eds.), *Advances in Motivation and Achievement, Vol. 12: New Directions in Measures and Methods* (pp. 77-120). New York: JAI/Elsevier Science.
- Boekaerts, M. & Corno, L. (2005). Self-regulation in the classroom: A perspective on Assessment and Intervention. *Applied Psychology: An International Review*, 54, 199-231.
- Boekaerts, M., Maes, S. & Karoly, P. (2005). Self-regulation across domains of applied psychology: Is there an emerging consensus? *Applied Psychology: An International Review*, 54, 149-154.
- Bortz, J. (1993). *Statistik für Sozialwissenschaftler* (4. Aufl.). Berlin: Springer.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.). Heidelberg: Springer.
- Bortz, J. & Döring, N. (2003). *Forschungsmethoden und Evaluation* (3. Aufl.). Berlin: Springer.
- Bortz, J., Lienert, G. A. & Boehnke, K. (2008). *Verteilungsfreie Methoden in der Biostatistik* (3. Aufl.). Heidelberg: Springer.
- Brosius, F. (2008). *SPSS 16. Das mitp-Standardwerk*. Heidelberg: Redline.
- Brosnan, M. (1998) *Technophobia: the psychological impact of information technology*. London: Routledge.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.
- Bruner, J.S. (1961). The Act of Discovery. *Harvard Educational Review*, 31, 21-32.
- Brunstein, J. C. & Heckhausen, H. (2006). Leistungsmotivation. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (3. Aufl.) (S. 143-191). Berlin: Springer.
- Buchner, A. (1995). Theories of complex problem solving. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European Perspective* (pp. 27-63). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (3. Aufl.). München: Pearson.
- Bungard, W. & Lück, H. E. (1974). *Forschungsartefakte und nicht-reaktive Messverfahren*. Stuttgart: Teubner.

- Bussey, K. & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106*, 676-713.
- Bussey, K. & Bandura, A. (2004). Social cognitive theory of gender development and functioning. In A. H. Eagly, A. Beall & R. Sternberg (Eds.), *The psychology of gender* (2nd ed., pp. 92-119) New York: Guilford.
- Byrne, B. (2001). *Structural equation modeling with AMOS. Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Caplan, P. A., Crawford, M., Hyde, J. S. & Richardson J. T. E. (Eds.). (1997). *Gender differences in human cognition*. New York: Oxford University Press.
- Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G. M., Barbaranelli, C. & Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology*, *100*, 525-534.
- Carroll, J. M. & Mack, R. (1983). Actively learning to use a word processor. In W. Cooper (Ed.), *Cognitive aspects of skilled typewriting* (pp. 259-281). New York, NY: Springer.
- Carroll, J. M., Mack, R. L., Lewis, C. H., Grischkowsky, N. L. & Robertson, S. R. (1985). Exploring exploring a word processor. *Human Computer Interaction*, *1*, 283-307.
- Carver, C. S. (2004). Self-regulation of action and affect. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 13-39). New York: Guilford Press.
- Carver, C. S. & Scheier, M. F. (1981). *Attention and self-regulation: a control theory approach to human behavior*. New York: Springer.
- Carver, C. S. & Scheier, M. F. (2000). On the structure of behavioral self-regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation research* (pp. 41-84). San Diego: Academic Press.
- Cassidy, S. & Eachus, P. (2002). Developing the computer user self-efficacy (CUSE) scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. *Journal of Educational Computing Research*, *26*, 169-189.
- Chua, S. L., Chen, D. T. & Wong, A. F. L. (1999). Computer anxiety and its correlates: A meta-analysis. *Computers in Human Behavior*, *15*, 609-623.
- Cleary, T. J. & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-

- based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools*, 41, 537-550.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, N.J.: Lawrence Erlbaum.
- Compeau, D. R. & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 19, 189-211.
- Cooper, J. & Weaver, K. D. (2003). *Gender and computers: Understanding the digital divide*. Mahwah, NJ: Erlbaum.
- Cooper, J. (2006). The digital divide: the special case of gender. *Journal of Computer Assisted Learning*, 22, 320–334.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98-104.
- Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. San Francisco, CA: Jossey-Bass.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the Psychology of Discovery and Invention*. New York: Harper Collins.
- Csikszentmihalyi, M. (2000). *Das Flow-Erlebnis: jenseits von Angst und Langeweile: im Tun aufgehen* (8. Aufl.). Stuttgart : Klett-Cotta.
- Csikszentmihalyi, M. (2002). *Flow: Das Geheimnis des Glücks* (10. Aufl.). Stuttgart: Klett-Cotta.
- Csikszentmihalyi, M. & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, 56, 815 - 822.
- Csikszentmihalyi, M. & Schiefele, U. (1993). Die Qualität des Erlebens und der Prozeß des Lernens. *Zeitschrift für Pädagogik*, 2, 207-221.
- Debowski, S., Wood, R. & Bandura, A. (2001). Impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic search. *Journal of Applied Psychology*, 86, 1129-1141.

- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deng, X., Doll, W. J. & Truong, D. (2004). Computer self-efficacy in an ongoing use context. *Behaviour and Information Technology*, 23, 395–412.
- Deutsche Gesellschaft für Psychologie (Hrsg.). (2007). *Richtlinien zur Manuskriptgestaltung* (3., überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.
- Dickhäuser, O. & Stiensmeier-Pelster, J. (2000). Geschlechtsunterschiede im Lern- und Leistungsverhalten am Computer: Ein theoretischer Rahmen. In F. Försterling, J. Stiensmeier-Pelster & L.-M. Silny (Hrsg.), *Kognitive und emotionale Aspekte der Motivation* (S. 53-76). Göttingen: Hogrefe.
- Dickhäuser, O. & Stiensmeier-Pelster, J. (2002a). Gender differences in computer work - Evidence for the model of achievement-related choices. *Contemporary Educational Psychology*, 27, 486-496.
- Dickhäuser, O. & Stiensmeier-Pelster, J. (2002b). Erlernte Hilflosigkeit am Computer? Geschlechtsunterschiede in computerspezifischen Attributionen. *Psychologie in Erziehung und Unterricht*, 49, 44-55.
- Dickhäuser, O. & Stiensmeier-Pelster, J. (2003). Gender differences in choice of computer courses: Applying an expectancy-value model. *Social Psychology of Education*, 6, 173-189.
- Diehl, J. M. & Kohr, H. U. (1994). *Deskriptive Statistik*. Eschborn bei Frankfurt: Klotz.
- Diehl, J. M. & Staufenbiel, T. (2007). *Statistik mit SPSS, Versionen 14 und 15*. Eschborn: Verlag Dietmar Klotz.
- Diekmann, A. (2007). *Empirische Sozialforschung* (18. Aufl.). Reinbek: Rowohlt.
- Dignath, C. C. & Büttner, G. (2008). Components of fostering self-regulated learning among students. a meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3, 231-264.
- Dörner, D. (1980). On the difficulties people have in dealing with complexity. *Simulation & Games*, 11, 87-106.
- Dörner, D. (2006). *Die Logik des Misslingens* (5. Aufl.). Reinbek: Rowohlt.
- Dörner, D. & Schaub, H. (1994). Errors in planning and decision-making and the nature of human information processing. *Applied Psychology. An International Review*, 43, 433-453.
- Durndell, A., Haag, Z. & Laithwaite, H. (2000). Computer self efficacy and gender: A cross

- cultural study of Scotland and Romania. *Personality and Individual Differences*, 28, 1037-1044.
- Dutke, S. (1994). Error handling: Visualizations in the human-computer interface and exploratory learning. *Applied Psychology: An International Review*, 43, 521-541.
- Dweck, C. S. & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.
- Eagly, A. H., Beall, A. E. & Sternberg, R. J. (Eds.). (2004). *The psychology of gender* (2nd ed.). New York: The Guilford Press.
- Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly*, 18, 585-609.
- Eccles, J. S. (2005). Subjective task value and the Eccles et al. Model of Achievement-Related Choices. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105-121). London: Guilford Press.
- Eccles, J. S., Adler, T. F. & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, 46, 26-43.
- Eccles, J. S. & Wigfield, A. (1995). In the mind of the achiever: The structure of adolescents' academic achievement related-beliefs and self-perceptions. *Personality and Social Psychology Bulletin*, 21, 215-225.
- Ellis, L., Hershberger, S., Field, E., Wersinger, S., Pellis, S., Geary, D., Palmer, C., Hoyenga, K., Hetsroni, A. & Karadi, K. (2008). *Sex differences: Summarizing more than a century of scientific research*. Mahwah, NJ: Erlbaum.
- Engeser, S., Rheinberg, F., Vollmeyer, R. & Bischoff, J. (2005). Motivation, Flow-Erleben und Lernleistung in universitären Lernsettings. *Zeitschrift für Pädagogische Psychologie*, 19, 159-172.
- Engeser, S., Rheinberg, F. (2008). Flow, moderators of challenge-skill-balance, and performance. *Motivation and Emotion*, 32, 158-172.
- Engeser, S. & Vollmeyer, R. (2005). Tätigkeitsanreize und Flow-Erleben. In R. Vollmeyer & J. C. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendungen* (S. 59 – 71). Stuttgart: Kohlhammer.
- Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis. Verbal reports as data* (revised ed.).

- Cambridge, MA: Bradford books/MIT Press.
- Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Fischer, R. & Milfont, T. L. (2010). Standardization in psychological research. *International Journal of Psychological Research*, 3, 89-97.
- Fleiss, J. L. & Cohen, J. (1973). The equivalence of weighted kappa and intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33, 613–619.
- Frensch, P. A. & Funke, J. (Eds.). (1995). *Complex problem solving: The European perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Freud, S. (1952). *Triebe und Triebchicksale* (Gesammelte Werke, Bd. X). Frankfurt: Fischer.
- Funke, J. (2006a). Lösen komplexer Probleme. In J. Funke & P. Frensch (Hrsg.), *Handbuch der Allgemeinen Psychologie - Kognition* (S. 439-445). Göttingen: Hogrefe.
- Funke, J. (2006b). Komplexes Problemlösen. In J. Funke (Hrsg.), *Denken und Problemlösen* (Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition, Band 8, S. 375-446). Göttingen: Hogrefe.
- Funke, J. & Frensch, P. A. (2007). Complex problem solving: The European perspective - 10 years after. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 25-47). New York: Lawrence Erlbaum.
- Funke, J. & Zumbach, J. (2006). Problemlösen. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 206-220). Göttingen: Hogrefe.
- Ghani, J. A. & Deshpande, S. P. (1994). Task characteristics and the experience of optimal flow in human-computer interaction. *Journal of Psychology*, 128, 381-391.
- Gold, A. (2003). Lernen. In S. Preiser (Hrsg.), *Pädagogische Psychologie. Psychologische Grundlagen von Erziehung und Unterricht* (S. 99-124). Weinheim: Juventa.
- Gold, A. (2008). Lehrstrategien. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 245-255). Göttingen: Hogrefe.
- Graham, J. W., Cumsille, P. E. & Elek-Fisk, E. (2003). Methods for handling missing data. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology* (Bd. 2, S. 87–114). New York: John Wiley & Sons.

- Green, A. J. K. & Gilhooly, K. J. (1990). Individual differences and effective learning procedures: The case of statistical computing. *International Journal of Man-Machine Studies*, 33, 97-119.
- Greeno, J. G. (1978). Nature of problem solving abilities. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 5). Hillsdale, NJ: Erlbaum.
- Greeno, J. G. & Simon, H. A. (1988). Problem solving and reasoning. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (2nd ed., pp. 589-672). New York: Wiley.
- Greif, S. (1990). Exploratorisches Lernen in der Mensch-Computer Interaktion. In F. Frei & I. Udris (Hrsg.), *Das Bild der Arbeit* (S. 143-157). Bern: Huber.
- Greif, S. (1996). Computer systems as learning environments. In J. Valsiner & H.G. Voss (Eds.), *The structure of learning processes* (pp. 144-174). Norwood, New Jersey: Ablex Publishing Comp.
- Greve, W. & Wentura, D. (1997). *Wissenschaftliche Beobachtung. Eine Einführung* (2. korrigierte Aufl.). Weinheim: PVU Beltz.
- Hadwin, A. F., Nesbit, J. C., Code, J., Jamieson-Noel, D. L. & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2, 107-124.
- Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum, Associates, Inc. Publishers.
- Halpern, D. F. (2004). A cognitive-process taxonomy for sex differences in cognitive abilities. *Current Directions in Psychological Science*, 13, 135-139.
- Hancock, G. R. & Mueller, R. O. (Eds.). (2006). *Structural equation modeling: A second course* (pp. 69-115). Greenwich, CT: Information Age Publishing.
- Hannover, B. & Bettge, S. (1993). *Mädchen und Technik*. Göttingen: Hogrefe.
- Hannover, B. & Kessels, U. (2008). Geschlechtsunterschiede beim Lernen. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie (Reihe Handbuch der Psychologie)* (S. 116-125). Göttingen: Hogrefe.
- Hasselhorn, M. & Labuhn, A. S. (2008). Metakognition und selbstreguliertes Lernen. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch Pädagogische Psychologie* (S. 28 – 37). Göttingen: Hogrefe.
- Havlicek, L. L. & Peterson, N. L. (1977). Effect of the violation of assumptions upon signify-

- cance levels of the Pearson r. *Psychological Bulletin*, 84, 373-377.
- Heckhausen, H. & Rheinberg, F. (1980). Lernmotivation im Unterricht, erneut betrachtet. *Unterrichtswissenschaft*, 8, 7-47.
- Heckhausen, H. (1963). *Hoffnung und Furcht in der Leistungsmotivation*. Meisenheim: Hain.
- Heckhausen, H. (1977). Motivation: Kognitionspsychologische Aufspaltung eines summarischen Konstrukts. *Psychologische Rundschau*, 13, 175-189.
- Hidi, S. & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st Century. *Review of Educational Research*, 70, 151-179.
- Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41, 111-127
- Hidi, S., Renninger, K. A. & Krapp, A. (2004). Interest, a motivational construct that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion and cognition: Integrative perspectives on intellectual functioning and development* (pp. 89-115). Mahwah, NJ: Erlbaum.
- Hirsch, E. (2002). *Frauen und Computer-Bildung in der Informationsgesellschaft : Informations- und Kommunikationstechnologie als Herausforderung frauenspezifischer Bildungsarbeit*. Klagenfurt: IFF.
- Hollnagel, E. (1993). *Human reliability analysis: Context and control*. London: Academic Press.
- Hoyle, R. H. & Kenny, D. A. (1999). Sample size, reliability, and tests of statistical mediation. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (S. 195-222). Thousand Oaks, CA: Sage.
- Hsu, L. M. & Field, R. (2003). Interrater agreement measures: Comments on Kappa, Cohen's Kappa, Scott's π , and Aickin's α . *Understanding Statistics*, 2, 205-219.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Huber, O. (2005). *Das psychologische Experiment. Eine Einführung* (4. Aufl.). Huber: Bern.
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60, 581-592.
- Hyde, J. S. (2007). New directions in the study of gender similarities and differences. *Current Directions in Psychological Science*, 16, 259-263.
- Hyde, J. S. & Durik, A. M. (2005). Gender, competence, and motivation. In A. Elliot & C. Dweck (Eds.), *Handbook of competence and motivation* (pp. 375-391). New York: Guilford.
- Hyde, J. S. & Linn, M. C. (2006). Gender similarities in mathematics and science. *Science*, 314,

599-600.

- Iacobucci, D., Saldanha, N. & Deng, X. (2007). A meditation on mediation: Evidence that structural equation models perform better than regressions. *Journal of Consumer Psychology*, 17, 139- 53.
- Imhof, M., Vollmeyer, R. & Beierlein, C. (2007). Computer use and the gender gap: The issue of access, use, and performance. *Computers in Human Behavior*, 23, 2823-2837.
- Jackson, L. A., Ervin, K. S., Gardner, P. D. & Schmitt, N. (2001). Gender and the Internet: Women communicating and men searching. *Sex Roles*, 44, 363-380.
- Jacobs, J. E., & Eccles, J. S. (2000). Parents, task values, and real-life achievement-related choices. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 408–433). San Diego: Academic Press.
- Jamieson-Noel, D. L. & Winne, P. H. (2003). Comparing self-reports to traces of studying behavior as representations of students' studying and achievement. *Zeitschrift für Pädagogische Psychologie*, 17, 159-171.
- Jones, M. G., Brader-Araje, L., Carboni, L., Carter, G., Rua, M., Banilower, E. & Hatch, H. (2000). Tool time: Gender and students' use of tools, control, and authority. *Journal of Research in Science Teaching*, 37, 760-783.
- Jonkisz, E. & Moosbrugger, H. (2007). Planung und Entwicklung von psychologischen Tests und Fragebogen. Unter Mitarbeit von H. Brandt. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 27-71). Heidelberg: Springer.
- Kang, N. E. & Yoon, W. C. (2008). Age- and experience-related user behavior differences in the use of complicated electronic devices. *International Journal of Human-Computer Studies*, 66, 425-437.
- Karoly, P. (1993). Mechanisms of self-regulation: a systems view. *Annual Review of Psychology*, 44, 23-52.
- Keith, N. & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology*, 93, 59-69.
- Keith, N., Richter, T. & Naumann, J. (2010). Active/exploratory training promotes transfer even in learners with low motivation and cognitive ability. *Applied Psychology: An International Review*, 59, 97-123.
- Kelava, A. & Moosbrugger, H. (2007). Deskriptivstatistische Evaluation von Items (Itemanaly-

- se) und Testwertverteilungen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 73-98). Heidelberg: Springer.
- Kenny, D. A., Kashy, D. A. & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 223–265). New York: McGraw-Hill.
- Kessels, U., Rau, M. & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, 74, 761-780.
- Klauer, K. C. (1996). Urteilerübereinstimmung bei dichotomen Kategoriensystemen. *Diagnostica*, 42, 101–118.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.
- Koch, S. C., Müller, S. M. & Sieverding, M. (2008). Women and computers. Effects of stereotype threat on attribution of failure. *Computers & Education*, 51, 1795-1803.
- Krampen, G. (2000). *Handlungstheoretische Persönlichkeitspsychologie: Konzeptuelle und empirische Beiträge zur Konstrukterhellung* (2. Aufl.). Göttingen: Hogrefe.
- Krapp, A. (1992). Das Interessenkonstrukt. Bestimmungsmerkmale der Interessenhandlung und des individuellen Interesses aus der Sicht einer Person-Gegenstands-Konzeption. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung* (S. 297-329). Münster: Aschendorff.
- Krapp, A. (2005). Psychologische Bedürfnisse und Interesse. Theoretische Überlegungen und praktische Schlussfolgerungen. In R. Vollmeyer & J. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendungen* (S. 23-38). Stuttgart: Kohlhammer.
- Kukla, A. (1974). Performance as a function of resultant achievement motivation (perceived ability) and perceived difficulty. *Journal of Research in Personality*, 7, 374-383.
- Landis, J. R. & Koch, G. G. The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Langens, T., Schmalt, H.-D. & Sokolowski, K. (2005). Motivmessung: Grundlagen und Anwendungen. In R. Vollmeyer & J. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendung* (3. Aufl., S. 72-91). Stuttgart: Kohlhammer.
- Leow, R. & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, 26, 35-57.

- Lepper, M. R. & Malone, T. W. (1987). Intrinsic motivation and instructional effectiveness in computer-based education. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning and instruction* (pp. 255-286). Hillsdale, NJ: Erlbaum.
- Lewin, K. (1926). Untersuchungen zur Handlungs- und Affekt-Psychologie. II.: Vorsatz, Wille und Bedürfnis. *Psychologische Forschung*, 7, 330-385.
- Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.
- Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, 2, 34-46.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: PVU Beltz.
- Lind, G. & Sandmann, A. (2003). Lernstrategien und Domänenwissen. *Zeitschrift für Psychologie*, 211, 171-192.
- Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, 90, 1112- 1121.
- Littleton, K., Light, P., Joiner, R., Messer, D. & Barnes, P. (1998). Gender, task scenarios and children's computer-based problem solving. *Educational Psychology*, 18, 327-340.
- Locke, E. A. & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.
- Lorenz, K. (1963). *Das sogenannte Böse. Zur Naturgeschichte der Aggression*. Wien: Borotha-Schoeler.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, 85, 103-177.
- Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal*, 29, 861-878.
- Lytton, H. & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin*, 109, 267-296.
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G. & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7, 83-104.
- Massimini, F. & Carli, M. (1988). The systematic assessment of flow in daily experience. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 266-287). Cambridge: University Press.
- Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: W. H. Freeman and Company.

- McCallum, R. C., Browne, M. W. & Sugarawa, H. M. (1996). Power analysis and determination of sample size for covariance structure modelling. *Psychological Methods*, 1, 130-149.
- McClelland, D. C., Atkinson, J. W., Clark, R. A. & Lowell, E. L. (1953). *The achievement motive*. Princeton, NJ: Van Nostrand.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- McDonald, R. P. & Ho, R. M. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, 7, 64-82.
- Meece, J. L. & Painter, J. (2008). Gender, self-regulation, and motivation. In D. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning. Theory, research, and applications* (pp. 339-368). New York: Lawrence Erlbaum Associates.
- Meelissen, M. R. M. & Drent, M. (2008). Gender differences in computer attitudes: Does the school matter? *Computers in Human Behavior*, 24, 969-985.
- Mercier, E. M., Barron, B. & O'Connor, K. M. (2006). Images of self and others as computer users: The role of gender and experience. *Journal of Computer Assisted Learning*, 22, 335-348.
- Messersmith, E. E., Garrett, J. L., Davis-Kean, P. E., Malanchuk, O. & Eccles, J. S. (2008). Career development from adolescence through emerging adulthood: Insights from information technology occupations. *Journal of Adolescent Research*, 23, 206-227.
- Middendorf, E. (2002). *Computernutzung und Neue Medien im Studium*. Bonn, Germany: Bundesministerium für Bildung und Forschung.
- Miller, G. A., Galanter, E. & Pribram, K. A. (1960). *Plans and the structure of behavior*. New York: Holt, Rhinehart & Winston.
- Moosbrugger, H. & Goldhammer, F. (2006). Aufmerksamkeitsdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 83-102). Berlin: Springer.
- Moosbrugger, H. & Kelava, A. (Hrsg.). (2007). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer
- Moschner, B. (2006). Pädagogische Psychologie und Geschlechterforschung. In G. Steins (Hrsg.), *Handbuch Psychologie und Geschlechterforschung* (S. 175-186). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Muis, K. R., Winne, P. H. & Jamieson-Noel, D. L. (2007). Using a multitrait multimethod ana-

- lysis to examine conceptual similarities of three self-regulated learning inventories. *British Journal of Educational Psychology*, 77, 177-195.
- Mummendey, H. D. (1995). *Die Fragebogen-Methode*. Göttingen: Hogrefe.
- Naumann, J., Richter, T. & Groeben, N. (2001). Validierung des Inventars zur Computerbildung (INCOBI) anhand eines Vergleichs von Anwendungsexperten und Anwendungsnovizen. *Zeitschrift für Pädagogische Psychologie*, 15, 219-232.
- Neber, H. (1987). *Angewandte Problemlösepsychologie*. Münster: Aschendorf.
- Neber, H. (1993). Generierung von Wissen über ein Computerprogramm durch exploratives und problemorientiertes Lernen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, 206-224.
- Nelson, L. & Cooper, J. (1997). Gender differences in children's reactions to success and failure with computers. *Computers in Human Behavior*, 13, 247-267.
- Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, N.J., Prentice-Hall.
- Nisbett, R. E. & Ross, L. D. (1980). *Human inference: strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 7, 231-259.
- O'Hara, K. & Payne, S. J. (1998) The effects of operator cost of planfulness of problem solving and learning. *Cognitive Psychology*, 35, 34-70.
- Öllinger, M. & Knoblich, G. (2006). Das Lösen einfacher Probleme. In J. Funke & P. A. Frensch (Eds.), *Handbuch der Allgemeinen Psychologie* (S. 431-438). Göttingen: Hogrefe.
- Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.). *Advances in motivation and achievement* (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.
- Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice*, 41, 116-125.
- Pajares, F. & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology*, 24, 124-139.
- Perels, F., Gürtler, T. & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. *Learning and Instruction*, 15, 123-139.

- Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research*, 21, 381-391.
- Pintrich, P. R. & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
- PISA-Konsortium Deutschland (Hrsg.). (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie der dritten internationalen Vergleichsstudie*. Münster: Waxmann.
- Poissant, H., Poëllhuber, B. & Falardeau, M. (1994). Résolution de problèmes, autorégulation et apprentissage. *Canadian Journal of Education*, 19, 30-44.
- Preacher, K. J., Wichman, A. L., MacCallum, R. C. & Briggs, N. E. (2008). *Latent growth curve modeling*. Thousand Oaks, CA: Sage Publications.
- Prenzel, M., Krapp, A. & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik*, 32, 163-173.
- Prümper, J. (1991). Die Inter-Rater-Reliabilität von Fehlerbeobachtungen im Feld. In M. Frese & D. Zapf (Hrsg.), *Fehler bei der Arbeit mit dem Computer - Ergebnisse von Beobachtungen und Befragungen im Bürobereich* (S. 47-59). Bern: Huber
- Raykov, T. & Marcoulides, G. A. (2006). *A first course in structural equation modeling*. Mahwah, NJ: Lawrence Earlbaum Associates.
- Reinecke, J. (2005). *Strukturgleichungsmodelle in den Sozialwissenschaften*. München/Wien: Oldenbourg.
- Renninger, K.A., Hidi, S. & Krapp, A. (Hrsg.). (1992). *The role of interest in learning and development*. Hillsdale, NJ: Erlbaum.
- Rheinberg, F. (1989). *Zweck und Tätigkeit*. Göttingen: Hogrefe.
- Rheinberg, F. (2004). *Motivation* (5. Aufl.). Stuttgart: Kohlhammer.
- Rheinberg, F. (2006). Intrinsische Motivation und Flow-Erleben. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (3. Aufl., S. 331-354). Berlin: Springer.
- Rheinberg, F. & Manig, Y. (2003). Was macht Spaß am Graffiti-Sprayen? Eine induktive Anreizanalyse. *Report Psychologie*, 4, 222-234.
- Rheinberg, F., Manig, Y., Kliegl, R., Engeser, S. & Vollmeyer, R. (2007). Flow bei der Arbeit, doch Glück in der Freizeit. Zielausrichtung, Flow und Glücksgefühle. *Zeitschrift für Arbeits- und Organisationspsychologie*, 51, 105-115.
- Rheinberg, F. & Tramp, N. (2006). Anreizanalyse intensiver Nutzung von Computern in der

- Freizeit. *Zeitschrift für Psychologie*, 214, 97-102.
- Rheinberg, F. & Vollmeyer, R. (2003). Flow-Erleben in einem Computerspiel unter experimentell variierten Bedingungen. *Zeitschrift für Psychologie*, 211, 161-170.
- Rheinberg, F., Vollmeyer, R. & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, 47, 57-66.
- Rheinberg, F., Vollmeyer, R. & Engeser, S. (2003). Die Erfassung des Flow-Erlebens. In J. Stiensmeier-Pelster & F. Rheinberg (Hrsg.), *Diagnostik von Motivation und Selbstkonzept* (Tests und Trends N.F. Bd. 2, S. 261-279). Göttingen: Hogrefe.
- Rheinberg, F., Vollmeyer, R. & Rollett, W. (2000). Motivation and action in self-regulated learning. In M. Boekaerts, P. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 503-529). San Diego: Academic Press.
- Richter, T., Naumann, J. & Groeben, N. (1999). *Das Inventar zur Computerbildung (INCOBI): Ein Instrument zur Erfassung der Computer Literacy und computerbezogenen Einstellungen*. Köln: Universität Köln.
- Richter, T., Naumann, J. & Groeben, N. (2001). Das Inventar zur Computerbildung (INCOBI): Ein Instrument zur Erfassung von Computer Literacy und computerbezogenen Einstellungen bei Studierenden der Geistes- und Sozialwissenschaften. *Psychologie in Erziehung und Unterricht*, 48, 1-13.
- Richter, T., Naumann, J. & Horz, H. (2001). Computer Literacy, computerbezogene Einstellungen und Computernutzung bei männlichen und weiblichen Studierenden. In H. Oberquelle, R. Oppermann & J. Krause (Hrsg.), *Mensch & Computer 2001: 1. Fachübergreifende Konferenz* (Berichte des German Chapter of the ACM, Bd. 55, S. 71-80). Stuttgart: Teubner.
- Richter, T., Naumann, J. & Horz, H. (2010). Eine revidierte Fassung des Inventars zur Computerbildung (INCOBI-R). *Zeitschrift für Pädagogische Psychologie*, 24, 23-37.
- Rost, D. H. (2007). *Interpretation und Bewertung pädagogisch-psychologischer Studien* (2., überarbeitete u. erweiterte Auflage). Weinheim: Beltz UTB.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80.
- Rotter, J. B. (1989). Internal versus external control of reinforcement: A case history of a variable. *American Psychologist*, 45, 489-493.

- Roy, M., Taylor, R. & Chi, M. T. H. (2003). Searching for information on-line and off-line: Gender differences among middle school students. *Journal of Educational Computing Research*, 29, 229-252.
- Rozell, E. J. & Gardner, W. L. (2000). Cognitive, motivation, and affective processes. Associated with computer-related performance: A path analysis. *Computers in Human Behavior*, 16, 199-222.
- Sansone, C. & Harackiewicz, J. M. (Eds.). (2000). *Intrinsic and extrinsic motivation: The search for optimal motivation and performance*. San Diego, CA: Academic Press.
- Schafer, J. L. & Graham, J. W. (2002). Missing Data: Our view of the state of the art. *Psychological Methods*, 7, 147-177.
- Schafer, J. L. & Olson, M. K. (1998). Multiple imputation for multivariate missing data problems: A data analyst's perspective. *Multivariate Behavioral Research*, 33, 545-571.
- Schaub, H. (2006). Störungen und Fehler beim Denken und Problemlösen. In J. Funke (Ed.), *Denken und Problemlösen* (Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition, Band 8) (S. 447-482). Göttingen: Hogrefe.
- Schaumburg, H. (2004). Laptops in der Schule. Ein Weg zur Überwindung des Digital Divide zwischen Jungen und Mädchen? *Zeitschrift für Medienpsychologie*, 16, 142-154.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research-Online*, 8, 23-74.
- Schermelleh-Engel, K. & Werner, C. S. (2007). Reliabilität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 113-133). Berlin: Springer.
- Schiefele, U. (1996). *Motivation und Lernen mit Texten*. Göttingen: Hogrefe.
- Schiefele, U. (2008). Lernmotivation und Interesse. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Psychologie, Bd. 10: Pädagogische Psychologie* (S. 38-49). Göttingen: Hogrefe.
- Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, 120-148.
- Schiefele, U. & Rheinberg, F. (1997). Motivation and knowledge acquisition: Searching for mediating processes. In P. Pintrich & M. L. Maehr (Eds.), *Advances in motivation and achievement*. (pp. 251-301). Greenwich, CT: JAI Press.

- Schiefele, U. & Streblow, L. (2005). Intrinsische Motivation - Theorien und Befunde. In R. Vollmeyer & J. C. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendung* (S. 39-58). Stuttgart: Kohlhammer.
- Schiefele, U., Streblow, L., Ermgassen, U. & Moschner, B. (2003). Lernmotivation und Lernstrategien als Bedingungen der Studienleistung: Ergebnisse einer Längsschnittstudie. *Zeitschrift für Pädagogische Psychologie*, 17, 185-198.
- Schmidt-Weigand, F., Hänze, M. & Wodzinski, R. (2009). Complex problem solving and worked examples: The role of prompting strategic behavior and fading-in solution steps. *Zeitschrift für pädagogische Psychologie*, 23, 129-138.
- Schmitz, B. (2000). Auf der Suche nach dem verlorenen Individuum: Vier Theoreme zur Aggregation von Prozessen. *Psychologische Rundschau*, 51, 83-92.
- Schmitz, B. (2001). Self-Monitoring zur Unterstützung des Transfers einer Schulung in Selbstregulation für Studierende: Eine prozessanalytische Untersuchung. *Zeitschrift für Pädagogische Psychologie*, 15, 181-197.
- Schnell, R., Hill, P. & Esser, E. (2008). *Methoden der empirischen Sozialforschung* (8. Aufl.). München: Oldenbourg.
- Schooler, J. W., Ohlsson, S. & Brooks, K. (1993). Thoughts beyond words. When language overshadows insight. *Journal of Experimental Psychology: General*, 122, 166-183.
- Schumacker, R. E. & Lomax, R. G. (2004). *A beginner's guide to structural equation modelling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schunk, D. H. & Pajares, F. (2004). Self-efficacy in education revisited: Empirical and applied evidence. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited* (pp. 115-138). Greenwich, CT: Information Age.
- Schunk, D. H., & Zimmerman, B. J. (Eds.). (1994). *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale, NJ: Erlbaum.
- Sedlmeier, P. & Gigerenzer, G. (1989). Do studies of power have an effect on the power of studies? *Psychological Bulletin*, 105, 309-316.
- Sedlmeier, P. & Gigerenzer, G. (1997). Intuitions about sample size: The empirical law of large numbers. *Journal of Behavioral Decision Making*, 10, 33-51.
- Sedlmeier, P. & Renkewitz, F. (2007). *Forschungsmethoden und Statistik in der Psychologie*. München: Pearson Education.

- Shapka, J. D. & Ferrari, M. (2003). Computer-related attitudes and actions of teacher candidates. *Computers in Human Behavior*, *19*, 319-334.
- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*, 407-441.
- Shotick, J. & Stephens, P. R. (2006). Gender inequities of self-efficacy on task-specific computer applications in business. *Journal of Education for Business*, *81*, 269-273.
- Sieverding, M. & Koch, S. C. (2009). (Self-) Evaluation of computer competence: How gender matters. *Computers & Education*, *52*, 696-701.
- Sjitsma, K. & van der Ark, L. (2003). Investigation and treatment of missing item scores in test and questionnaire data. *Multivariate Behavioral Research*, *38*, 505-528.
- Skinner, N. & Brewer, N. (2002). The dynamics of threat and challenge appraisals prior to a stressful achievement event. *Journal of Personality and Social Psychology*, *83*, 678-692.
- Smith, C. A. & Ellsworth, P. C. (1987). Patterns of appraisal and emotion related to taking an exam. *Journal of Personality and Social Psychology*, *52*, 475-488.
- Smith, J. L., Morgan, C. L. & White, P. H. (2005). Investigating a measure of computer technology domain identification: A tool for understanding gender-differences and stereotypes. *Educational and Psychological Measurement*, *65*, 336-355 .
- Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4-28.
- Spiel, C. (1998). Four methodological approaches to the study of stability and change in development. *Methods of Psychological Research Online*, *3*, 7-22.
- Spörer, N. & Brunstein, J. C. (2006). Erfassung selbstregulierten Lernens mit Selbstberichtsverfahren: Ein Überblick zum Stand der Forschung. *Zeitschrift für Pädagogische Psychologie*, *20*, 147-160.
- Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-811.
- Stelzl, I. (2005). *Fehler und Fallen der Statistik – für Psychologen, Pädagogen und Sozialwissenschaftler* (Reprint). Münster: Waxmann
- Sternberg, R. J. (1996). *Cognitive psychology*. Fort Worth, TX: Harcourt Brace College Publishers.
- Stout, J. G., Dasgupta, N., Hunsinger, M. & McManus, M. (in press). STEMing the tide: Using ingroup experts to inoculate women's self-concept and professional goals in science,

- technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, 100, 255-270.
- Straka, G. A. (2006). Lernstrategien in Modellen selbst gesteuerten Lernens. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 390 - 404). Göttingen: Hogrefe.
- Tarling, R. (2009). *Statistical Modelling for Social Researchers: Principles and Practice*. London: Routledge.
- Todman, J. & Day, K. (2006). Computer anxiety: the role of psychological gender. *Computers in Human Behavior*, 22, 856-869.
- Todtenhaupt, E. (2007). *Vorgehensweisen von Novizen bei der Aufgabenbearbeitung mit SPSS*. Unveröffentlichte Diplomarbeit, Johann Wolfgang Goethe-Universität, Frankfurt am Main.
- Trudel, C. I. & Payne, S. J. (1995). Reflection and goal management in exploratory learning. *International Journal of Human Computer Studies*, 42, 307-339.
- Trudel, C. I. & Payne, S. J. (1996). Self-monitoring during exploration of an interactive device. *International Journal of Human-Computer Studies*, 45, 723-747.
- Tuovinen, J. E. & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, 91, 334-341.
- Van der Linden, D., Frese, M. & Sonnentag, S. (2003). Fatigue and exploration in a complex computer program: Systematic versus unsystematic behavior. *Human Factors*, 45, 483-494.
- Van der Linden, D., Sonnentag, S., Frese, M. & van Dyck, C. (2001). Exploration strategies, performance, and error consequences when learning a complex computer task. *Behaviour and Information Technology*, 20, 189-198.
- Van Oostendorp, H. & De Mul, S. (1999). Learning by exploration: Thinking aloud while exploring an information system. *Instructional Science*, 27, 269-284.
- Vollmeyer, R. (2005). Ein Ordnungsschema zur Integration verschiedener Motivationskomponenten. In R. Vollmeyer & J. C. Brunstein (Hrsg.), *Motivation und ihre Anwendung* (S. 9-19). Stuttgart: Kohlhammer.
- Vollmeyer, R., Burns, B. D. & Holyoak, K. J. (1996). The impact of goal on strategy use and the acquisition of problem structure. *Cognitive Science*, 20, 75-100.
- Vollmeyer, R. & Funke, J. (1999). Personen- und Aufgabenmerkmale beim Komplexen Problemlösen. *Psychologische Rundschau*, 50, 213-219.

- Vollmeyer, R. & Imhof, M. (2007). Are there gender differences in computer performance, and if so, can motivation explain them? *Zeitschrift für Pädagogische Psychologie*, 21, 251-261.
- Vollmeyer, R. & Rheinberg, F. (1998). Motivationale Einflüsse auf Erwerb und Anwendung von Wissen in einem computersimulierten System. *Zeitschrift für Pädagogische Psychologie*, 12, 11-23.
- Vollmeyer, R. & Rheinberg, F. (1999). Motivation and metacognition when learning a complex system. *European Journal of Psychology of Education*, 14, 541-554.
- Vollmeyer, R. & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction*, 10, 293-309.
- Vollmeyer, R. & Rheinberg, F. (2003). Aktuelle Motivation und Motivation im Verlauf. In J. Stiensmeier-Pelster & F. Rheinberg (Hrsg.), *Diagnostik von Motivation und Selbstkonzept* (S. 281-295). Göttingen: Hogrefe.
- Vollmeyer, R. & Rheinberg, F. (2006). Motivational effects on self-regulated learning with different tasks. *Educational Psychology Review*, 18, 239-253.
- Von Eye, A. & Bogat, G. A. (2006). Person-oriented and variable-oriented research: Concepts, results, and development. *Merrill-Palmer Quarterly*, 52, 390-420.
- Weiner, B. (1986). *An attributional theory of emotion and motivation*. New York: Springer-Verlag.
- Weiner, B. (2010). The development of an attribution-based theory of motivation: A history of ideas. *Educational Psychologist*, 45, 28-36.
- West, S. G., Finch, J. F. & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modelling: Concepts, issues and applications* (S. 56-75). Thousand Oaks, CA: Sage.
- Whitley, B. E., Jr. (1997). Gender differences in computer-related attitudes and behavior. A meta-analysis. *Computers in Human Behavior*, 13, 1-22.
- Wiedenbeck, M. & Züll, C. (2001). *Klassifikation mit Clusteranalyse: Grundlegende Techniken hierarchischer und K-means-Verfahren*. ZUMA, How-to-Reihe, Nr. 10.
- Wild, K.-P. (2000). *Lernstrategien im Studium. Strukturen und Bedingungen*. Münster: Waxmann.
- Wilkinson, L. and the Taks Force on Statistical Interference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604.
- Winne, P. (1995). Self-regulation is ubiquitous but its forms vary with knowledge. *Educational*

- Psychologist*, 30, 223-228.
- Winne, P. H. & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297-314). Mahwah, NJ: Lawrence Erlbaum Associates.
- Winne, P. H. & Perry, N. E. (2000). Measuring self-regulated learning. In P. Pintrich, M. Boekaerts & M. Seidner (Eds.), *Handbook of self-regulation* (pp. 531-566). Orlando, FL: Academic Press.
- Winther, E. (2006). *Motivation in Lernprozessen. Konzepte in der Unterrichtspraxis von Wirtschaftsgymnasien*. Wiesbaden: Deutscher Universitäts-Verlag.
- Wirth, J. & Leutner, D. (2008). Self-regulated learning as a competence. Implications of theoretical models for assessment methods. *Zeitschrift für Psychologie/Journal of Psychology*, 216, 102-110.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- Wolters, C. A. (2003). Regulation of Motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist*, 38, 189-205.
- Wood, R. E., Kakebeeke, B. M., Debowski, S. & Frese, M. (2000). The impact of enactive exploration on intrinsic motivation, strategy, and performance in electronic search. *Applied Psychology: An International Review*, 49, 263-283.
- Young, B. (2000). Gender differences in student attitudes toward computers. *Journal of Research in Computing in Education*, 33, 204-217.
- Zimmerman, B. J. (2000a). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA: Academic Press.
- Zimmerman, B. J. (2000b). Self-regulatory cycles of learning. In G. A. Straka (Ed.), *Conceptions of self-directed learning* (pp. 221-234). Münster: Waxmann.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: an overview. *Theory into Practice*, 41, 64-70.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45, 166-183.

- Zimmerman, B. J. & Kitsantas, A. (2005). The hidden dimension of personal competence: self-regulated learning and practice. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 204-222). New York: Guilford Press.
- Zimmerman, B. J. & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, 80, 284-290.
- Zimmerman, B. J. & Schunk, D. H. (2008). Motivation: An essential dimension of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 1-30). Mahwah, NJ: Erlbaum.

Software:

- Arbuckle, J. L. (2008). *AMOS 17* [Computer software]. Chicago, IL: SPSS Inc.
- Buchner, A., Faul, F. & Erdfelder, E. (1998). *G*Power 3.0.2* [Shareware]: Zugriff am 15.05.2007 unter <http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower3/>
- Fox Magic Software. (2006). *Screen Virtuoso Pro 2.50* [Computer software]. Miami: Author.
- Muthén, L. K. & Muthén, B. (1998-2007). *Mplus* (Version 5.0) [Computer software]. Los Angeles: Muthén & Muthén.
- Preacher, K. J. & Leonardelli, G. J. (2006). *Calculation for the Sobel test. An interactive calculation tool for mediation tests* [Online calculation tool]:
Zugriff am 20.05.2009 unter <http://people.ku.edu/~preacher/sobel/sobel.htm>.
- SPSS Inc. (2003). *SPSS 12.0* [Computer software]. Chicago/IL: SPSS.
- SPSS Inc. (1993-2007). *SPSS 17.0* [Computer software]. Chicago/IL: SPSS.
- SXR Software. (2006). *StatWin Pro 7.5.5* [Computer software]. Perm, Russia: Author.

Elektronische Quellen:

- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (2000). *Fähigkeit zum selbstregulierten Lernen als fächerübergreifende Kompetenz*. Zugriff am 30.3.2009 unter <http://www.mpib-berlin.mpg.de/pisa/CCCdt.pdf>.
- Beckwith, L., Kissinger, C., Burnett, M., Widenbeck, S., Lawrance, J., Blackwell, A. & Cook, C. (2006). *Tinkering and Gender in End-User Programmers' Debugging*. In Proceedings of ACM Conference on Human Factors in Computing Systems (CHI'06).
Zugriff am 12.06.2007 unter <ftp://ftp.cs.orst.edu/pub/burnett/chi06-genderTinker.pdf>

OECD (2010), *PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I)*.

Zugriff am 15.03.2011 unter

<http://browse.oecdbookshop.org/oecd/pdfs/browseit/9810071E.pdf>

Richter, T., Naumann, J. & Horz, H. (2001). Computer Literacy, computerbezogene Einstellungen und Computernutzung bei männlichen und weiblichen Studierenden. In H. Oberquelle, R. Oppermann & J. Krause (Hrsg.), *Mensch & Computer 2001: 1. Fachübergreifende Konferenz* (Berichte des German Chapter of the ACM, Bd. 55, S. 71-80). Stuttgart: Teubner.

Zugriff am 05.10.2008 unter

<http://mc.informatik.uni-hamburg.de/konferenzbaende/mc2001/index.html>

Anhang

Tabellenverzeichnis des Anhangs

	Seite	
Tabelle A – 1	Kategoriensystem mit Konstruktdefinitionen und Indikatoren zur Erfassung der Explorationsstrategien beim Erlernen des Statistikprogramms SPSS	A – 11
Tabelle A – 2	Ergebnis der Item- und Skalenanalyse des Fragebogens zu Statistikkenntnissen mit Cronbachs α	A – 24
Tabelle A – 3	Ergebnis der Item- und Skalenanalyse der ausgewählten Items der Skala Praktisches Computerwissen PRACOWI mit Cronbachs α	A – 25
Tabelle A – 4	Ergebnis der Item- und Skalenanalyse der ausgewählten Items der Skala Vertrautheit mit Computeranwendungen VECA mit Cronbachs α	A – 26
Tabelle A – 5	Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den ersten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit	A – 27
Tabelle A – 6	Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den zweiten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit	A – 28
Tabelle A – 7	Ergebnis der Item- und Skalenanalyse für die Leistung bei der Bearbeitung der fünf Aufgaben des Arbeitsblattes (Variable: „Leistung_Arbeitsblatt“)	A – 29
Tabelle A – 8	Deskriptive Statistiken für die Verteilung der Skalensummenwerte inklusive der Ergebnisse des Kolmogorov-Smirnov-Tests auf Normalverteilung	A – 30
Tabelle A – 9	Partialkorrelationen der Variablen in der Stichprobe unter Konstanzhaltung des Vorwissens (PRACOWI, Statistikvorwissenstest)	A – 31
Tabelle A – 10	Ergebnis der Item- und Skalenanalyse des Fragebogens zu Statistikkenntnissen	A – 37
Tabelle A – 11	Ergebnis der Item- und Skalenanalyse des Fragebogens zur aktuellen Motivation (Eingangsmotivation) sowie Cronbachs α getrennt für die vier Subskalen (I = Interesse, E = Erfolgswahrscheinlichkeit, H = Herausforderung, M = Misserfolgsbefürchtung)	A – 38
Tabelle A – 12	Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den ersten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit	A – 39
Tabelle A – 13	Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den zweiten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit	A – 40
Tabelle A – 14	Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den dritten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit	A – 41

	Seite	
Tabelle A – 15	Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum ersten Messzeitpunkt	A – 42
Tabelle A – 16	Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum zweiten Messzeitpunkt	A – 43
Tabelle A – 17	Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum dritten Messzeitpunkt	A – 44
Tabelle A – 18	Ergebnis der Item- und Skalenanalyse für die Leistung bei der Bearbeitung der vier Aufgaben des Arbeitsblattes (Variable: „Leistung_Arbeitsblatt“)	A – 45
Tabelle A – 19	Ergebnis der Item- und Skalenanalyse für die Leistung beim Lösen der vier Aufgaben in SPSS (Variable: „Leistung_SPSS“)	A – 46
Tabelle A – 20	Deskriptive Statistiken für die Verteilung der Skalensummenwerte inklusive der Ergebnisse des Kolmogorov-Smirnov-Tests auf Normalverteilung	A – 47
Tabelle A – 21	Interkorrelationen der Variablen in der Gesamtstichprobe	A – 48
Tabelle A – 22	Ergebnis der Item- und Skalenanalyse des Fragebogens zu Statistikkenntnissen mit Cronbachs α	A – 57
Tabelle A – 23	Ergebnis der Item- und Skalenanalyse der ausgewählten Items der Skala Praktisches Computerwissen PRACOWI mit Cronbachs α	A – 58
Tabelle A – 24	Ergebnis der Item- und Skalenanalyse der ausgewählten Items der Skala Vertrautheit mit Computeranwendungen VECA mit Cronbachs	A – 59
Tabelle A – 25	Ergebnis der Item- und Skalenanalyse des Fragebogens zur aktuellen Motivation (Eingangsmotivation) sowie Cronbachs α getrennt für die vier Subskalen (I = Interesse, E = Erfolgswahrscheinlichkeit, H = Herausforderung, M = Misserfolgsbefürchtung).	A – 60
Tabelle A – 26	Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den ersten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit	A – 61
Tabelle A – 27	Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den zweiten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit	A – 62
Tabelle A – 28	Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den dritten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit	A – 63
Tabelle A – 29	Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum ersten Messzeitpunkt	A – 64

	Seite
Tabelle A – 30 Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum zweiten Messzeitpunkt	A – 65
Tabelle A – 31 Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum dritten Messzeitpunkt	A – 66
Tabelle A – 32 Ergebnis der Item- und Skalenanalyse für die Leistung bei der Bearbeitung der sieben Aufgaben des Arbeitsblattes (Variable: „Leistung_Arbeitsblatt“) sowie Cronbachs α zur Bestimmung der Reliabilität	A – 67
Tabelle A – 33 Ergebnis der Item- und Skalenanalyse für die Leistung beim Lösen der sieben Aufgaben in SPSS (Variable: „Leistung_SPSS“)	A – 68
Tabelle A – 34 Deskriptive Statistiken für die Verteilung der Skalensummenwerte inklusive der Ergebnisse des Kolmogorov-Smirnov-Tests auf Normalverteilung	A – 69
Tabelle A – 35 Interkorrelationen zwischen dem Vorwissen, dem Flow-Erleben, der Häufigkeit der Nutzung der systematischen Explorationsstrategie und der Leistung	A – 71
Tabelle A – 36 Interkorrelationen zwischen den unsystematischen Explorationsstrategien, dem Flow-Erleben und der Leistung zu den drei Messzeitpunkten der Untersuchung	A – 72

A 1. Untersuchungsmaterialien der Studien 1 bis 3

In diesem Kapitel sind die Materialien dargestellt, die in allen drei Studien in unveränderter Form eingesetzt wurden. Hierzu zählen die Instruktion der Teilnehmer (A 1.1), die Einführung in das Statistikprogramm SPSS (A 1.2) und die Flow-Kurz-Skala (A 1.4). Der Fragebogen zur aktuellen Motivation wurde in seiner Langversion (A 1.3) sowie einer Kurzversion (A 1.5), welche von den Teilnehmern während des Bearbeitungsprozesses beantwortet wurde, ausschließlich in Studie 2 und 3 eingesetzt. Das in Studie 1 entwickelte Kategoriensystem, das auch in den Studien 2 und 3 benutzt wurde, ist in Abschnitt A 1.6 beschrieben.

A 1.1 Begrüßung der Teilnehmer und Aufgabeninstruktion

Im Folgenden ist die Instruktion der Teilnehmer aus Studie 2 exemplarisch für alle drei Studien wiedergegeben. Die Instruktion variierte kaum zwischen den Studien. Variationen sind fettgedruckt und in Klammern erläutert.

„Wir freuen uns, dass Sie die Zeit gefunden und sich bereit erklärt haben, an diesem Versuch zu multimedialem Lernen teilzunehmen. Mein Name ist ... Die Studie findet im Rahmen einer Dissertation im Institut für Psychologie an der Universität Frankfurt statt. In der Studie beschäftigen wir uns mit Motivation und selbstregulierten Lernen bei der Nutzung von Computeranwendungen. Für diese Studie haben wir uns das Statistik-Programm SPSS als Anwendung herausgesucht. Ihre Aufgabe ist es, in den **30 Minuten** (Studie 1: 25 Minuten), die Ihnen zur Verfügung stehen, möglichst viele der Aufgaben auf dem Arbeitsblatt zu lösen. Das heißt, Sie sollen die Aufgaben so schnell wie möglich bearbeiten.

Wichtig: Bitte gehen Sie der Reihenfolge nach vor und gehen Sie erst zur nächsten Aufgabe über, wenn Sie für die vorherige Aufgabe ein Ergebnis gefunden haben. Bevor Sie mit einer neuen Aufgabe beginnen, tragen Sie bitte die Uhrzeit im dafür vorgesehenen Feld oberhalb der Aufgabe auf dem Arbeitsblatt ein. Die Uhrzeit zeigt Ihnen der Computer an. Ihr Vorgehen bei der Aufgabenlösung wird auf dem Rechner aufgezeichnet und später von uns analysiert. Zusätzlich erheben wir Ihre Motivation während der Bearbeitung. Dafür unterbrechen wir Ihre Arbeit kurz zum Ausfüllen der Fragebögen. Die Daten werden anonym ausgewertet und nur für unser Forschungsvorhaben verwendet.

Bei Unklarheiten zu den Aufgabenstellungen können Sie vor Beginn der Aufgabenbearbeitung Rückfragen stellen. Zur Aufgabenlösung und Fragen zu SPSS können wir keine Hilfe geben, weil dies Ihre Ergebnisse und die der Studie beeinflussen könnte. Wichtig für

unsere späteren Analysen ist es, dass Sie das Ausgabefenster von SPSS nicht schließen, sondern nur minimieren. Wenn Sie sich Notizen machen möchten, finden Sie hierfür neben Ihrem Computerarbeitsplatz Papier und Kugelschreiber.

Ich schildere Ihnen nun den genauen Ablauf der Studie:

1. Zunächst bitte ich Sie einen Fragebogen zu Ihrer Computererfahrung und zu Ihren Statistikenntnissen zu beantworten.
2. Im Anschluss erhalten Sie eine Kurzeinführung in die Benutzeroberfläche von SPSS 12.0.
3. Ich gebe Ihnen dann das Arbeitsblatt mit den Aufgaben, die Sie in SPSS bearbeiten sollen. Sie haben Zeit, sich die Aufgaben durchzulesen und Verständnisfragen zu stellen.
4. Vor Beginn der Aufgabenbearbeitung erhalten Sie einen Fragebogen zur Motivation, den ich Sie bitte auszufüllen. (In Studie 1 wurde hierauf verzichtet)
5. Im Anschluss daran haben Sie **30 Minuten** (Studie 1: 25 Minuten) Zeit, die Aufgaben auf dem Arbeitsblatt mit SPSS zu bearbeiten.
6. Es ist wichtig, dass Sie die Aufgaben der Reihe nach bearbeiten und jedes Mal die Zeit zu Beginn der neuen Aufgabe auf dem Arbeitsblatt eintragen.
7. Während der Bearbeitungszeit werde ich Sie **dreimal** (Studie 1: zweimal) unterbrechen, um Sie einen weiteren Fragebogen ausfüllen zu lassen. Immer wenn Sie einen Fragebogen ausfüllen oder etwas auf dem Aufgabenblatt notieren, schieben Sie bitte den Mauszeiger auf die Uhr, die oben in der Bildschirmmitte eingeblendet ist.
8. Nach dem Ausfüllen können Sie direkt weiterarbeiten.
9. Nach Ablauf der Bearbeitungszeit ist die Studie beendet. Ich gebe Ihnen dann gerne Rückmeldung zu den Aufgaben und Hinweise auf die Ziele unserer Untersuchung.”

A 1.2 Einführung in das Statistikprogramm SPSS

Die Teilnehmer wurden von der Untersuchungsleiterin in die Benutzeroberfläche des Programms SPSS mit eingeführt. Die folgende Instruktion wurde dabei in allen drei Studien verwendet:

„Jetzt möchte ich Ihnen eine kurze Einführung in das Programm SPSS geben: Was wir jetzt auf dem Bildschirm sehen ist die Benutzeroberfläche von SPSS mit dem von uns für dieses Experiment zusammengestellten Datensatz. Am unteren Rand der Tabelle kann zwischen zwei Datenblättern gewählt werden: ein Blatt enthält die Liste der Variablen (wie Nummer des/der Befragten, von denen wir Daten erhoben haben, Geschlecht etc.). Die Variablen sind untereinander aufgelistet.

Das zweite Blatt - das eigentlich entscheidende für die Bearbeitung der Aufgaben – zeigt die Werte der einzelnen Personen in den einzelnen Variablen (Spalten). D.h. in der linken Spalte sind die Personen untereinander aufgeführt; rechts daneben sind in den dazugehörigen Zeilen die den Personen zugeordneten Daten dargestellt; die Variablen (z.B. Zahlengedächtnis) sind in den Spalten aufgeführt. Am oberen Rand des Tabellenblatts kann man in der Menüleiste verschiedene Optionen wählen (z.B. Datei, Bearbeiten, Ansicht, Transformieren, Analysieren, Grafiken, Extras, Fenster, Hilfe). Wenn man die einzelnen Optionen anklickt, dann kann man zwischen verschiedenen Operationen wählen, die man durchführen bzw. berechnen kann.

Auch gibt es zwei verschiedenen Fenster, das eben beschriebene SPSS-Datenfenster und das Ausgabenfenster. Das Ausgabefenster ist momentan noch leer, aber sobald Sie eine Aufgabe ausrechnen lassen, wird hier das Ergebnis angezeigt. Dieses Fenster bitte nicht schließen, sondern immer nur minimieren. Die aktuelle Ausgabe finden Sie immer am Ende des Ausgabenfensters. Soweit unsere kurze Einführung in das Programm. Leider können wir nicht mehr Hinweise geben, da das die Aufgabenbearbeitung und die Ergebnisse beeinflussen könnte.“

A 1.3 Fragebogen zur aktuellen Motivation (FAM)

Vor Beginn der Bearbeitung des Aufgabenblattes wurde den Teilnehmern der Fragebogen zur aktuellen Motivation (Rheinberg et al., 2001) vorgelegt.

		trifft						trifft
		nicht zu						zu
1.	Ich mag solche Rätsel und Knobeleyen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	Wahrscheinlich werde ich die Aufgabe nicht schaffen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.	Ich fühle mich unter Druck, bei der Aufgabe gut abschneiden zu müssen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.	Die Aufgabe ist eine richtige Herausforderung für mich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.	Nach dem Lesen der Instruktion erscheint mir die Aufgabe sehr interessant.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8.	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.	Bei Aufgaben wie dieser brauche ich keine Belohnung, sie machen mir auch so viel Spaß.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12.	Es ist mir etwas peinlich, hier zu versagen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13.	Ich glaube, das kann jeder schaffen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14.	Ich glaube, ich schaffe diese Aufgabe nicht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15.	Wenn ich die Aufgabe schaffe, werde ich schon ein wenig stolz auf meine Tüchtigkeit sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.	Wenn ich an die Aufgabe denke, bin ich etwas beunruhigt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17.	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18.	Die konkreten Leistungsanforderungen hier lähmen mich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A 1.4 Flow-Kurz-Skala (FKS)

Im Prozess der Aufgabenbearbeitung beantworteten die Teilnehmer zwei- bzw. dreimal die folgenden Items der Flow-Kurz-Skala (Rheinberg et al., 2003).

	trifft nicht zu							trifft zu
1. Ich fühle mich optimal beansprucht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Ich merke gar nicht, wie die Zeit vergeht.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Ich habe keine Mühe, mich zu konzentrieren.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Mein Kopf ist völlig klar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Ich bin ganz vertieft in das was ich gerade mache.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Die richtigen Gedanken/Bewegungen kommen wie von selbst.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Ich weiß bei jedem Schritt, was ich zu tun habe.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Ich bin völlig selbstvergessen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Für mich persönlich sind die jetzigen Anforderungen ...								
zu gering	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	zu hoch

A 1.5 Kurzversion des Fragebogens zur aktuellen Motivation (FAM-K)

In Studie 2 und 3 wurde neben dem Flow-Erleben auch die aktuelle Motivation zu mehreren Zeitpunkten während des Aufgabenbearbeitungsprozesses erfasst. Die folgenden Items wurden aus dem Fragebogen zur aktuellen Motivation von Rheinberg et al. (2001) entnommen.

	trifft nicht zu						trifft zu
1. Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Wahrscheinlich werde ich die Aufgabe nicht schaffen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Es ist mir etwas peinlich, hier zu versagen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A 1.6 Beobachtungssystem zur Erfassung der Explorationsstrategien

Die folgende Tabelle gibt eine Übersicht über das im Rahmen von Studie 1 entwickelte Kategoriensystem zur Erfassung von Explorationsstrategien. Die Tabelle umfasst die Definition der vier Explorationsstrategien sowie der Kategorie "Unterbrechung". Eine erschöpfende Beschreibung der Kategorie-Indikatoren wurde als zu umfangreich eingestuft; aus diesem Grund werden in der Tabelle ausgewählte, repräsentative Indikatoren für die jeweilige Explorationsstrategie aufgeführt (vgl. van der Linden et al., 2001, 2003).

Tabelle A – 1

Kategoriensystem mit Konstruktdefinitionen und Indikatoren zur Erfassung der Explorationsstrategien beim Erlernen des Statistikprogramms SPSS

Systematische Exploration (Kategoriebezeichnung)

Konstruktdefinition:

Die systematische Exploration des Statistikprogramms ist daran erkennbar, dass der Computernutzer intentional/zielgerichtet vorgeht und Hypothesen zu testen, Ergebnisse der Prüfung zu evaluieren und seine (weiteren) Handlungen zu planen scheint. Der nächste Handlungsschritt ist durch den Beobachter nachvollziehbar und vorhersehbar. Die Programmfunktionen werden in einem strukturierten und kohärenten Weg ausprobiert.

Ausgewählte Indikatoren:

- **Systematisches Absuchen der Menüleiste:** Wenn eine neue Aufgabe begonnen wird, wird die Menüleiste systematisch nach der relevanten Funktion von links nach rechts oder direkt unter Analysieren beginnend abgesucht
 - **Systematisches Absuchen der Variablenansicht:** Bei der Bearbeitung der ersten Aufgabe des Aufgabenblattes schaut sich der Computernutzer die Variablenlabels in der Variablenansicht an, um danach die richtigen Variablen auswählen zu können
 - **Erkennbare Orientierung an „Namen“ und Bezeichnungen von Funktionen:** Z.B. werden bei Aufgabe zu Mittelwerten „Deskriptive Statistik“ oder Mittelwerte angeklickt. Wenn die Aufgabe das Erstellen einer Grafik beinhaltet, dann wählt der Computernutzer direkt den Menüpunkt „Grafik“ und sucht hier nach der entsprechenden Funktion
 - **Eine Handlung folgt aus der vorangegangenen Handlung:** Z.B. ändert ein Computernutzer die Settings für eine Option (z.B. Häkchen setzen, Variable auswählen) und evaluieren das Ergebnis der Veränderungen. Nach jeder (erfolgreich) gelösten Aufgabe, geht der Computernutzer zur Menüleiste zurück und sucht nach der nächsten aufgabenrelevanten Funktion.
 - **Der Computernutzer variiert systematisch die Einstellungen beim Analysieren, Berechnen oder Erstellen einer Grafik in SPSS:** Beispielsweise startet der Computernutzer mit dem Erstellen eines einfachen Balkendiagramms und lässt sich dies im Output anzeigen. Dann geht er zurück und verändert die zunächst getroffenen Einstellungen (z.B. wechselt von einfachem zu gruppiertem Balkendiagramm).
-

Systematische Exploration (Kategoriebezeichnung) (Forts.)

- **Wenn das Programm einen Fehler rückmeldet oder der Computernutzer einen Fehler registriert, dann korrigiert er seine Handlungen:** z.B. wenn ein Mittelwert mit einer falschen Variablen berechnet wurde, dann ruft der Computernutzer im nächsten Schritt das gleiche Fenster erneut auf und wählt eine andere Variable aus. Der Computernutzer geht nicht immer wieder neue Wege, um das Problem zu lösen, sondern sucht nach Fehlern in seinem Vorgehen, sucht die Quelle des Fehlers und versucht diese abzuschalten (z.B. wenn „ok“-Button nicht aktiviert ist und somit nicht anklickbar, verlässt der Computernutzer nicht das Fenster, sondern sucht nach fehlenden Eingaben, bis das Programm den Button aktiviert.
- **Ergebniskontrolle nach Befehlsausführung:** Wenn ein Befehl ausgeführt wurde, d.h. eine Funktion in Analysieren oder Grafiken aufgerufen und mit dem Klick auf „ok“ beendet wurde, navigiert der Computernutzer zum Ausgabefenster und sucht dort nach dem Ergebnis der Aktion. Wenn das Anzeigefenster ein neues Ergebnisses durch Blinken anzeigt, wechselt der Computernutzer direkt zum Ausgabefenster.
- **Nicht effektive Handlungsweisen werden vom Computernutzer höchstens zweimal durchgeführt** (z.B. Markieren von Spalten wie dies in Excel möglich ist). Diese Handlungsweisen werden schnell verworfen und andere Wege ausprobiert.
- **Es ist erkennbar, dass der Computernutzer erfolgreiche Bedienungsweisen im Programm in seinem Gedächtnis speichert:** Das erworbene Wissen um Lösungswege wird für die Nutzung der neuen Aufgabe genutzt (Transfer). Beispiel: Nachdem die Person Mittelwerte über „Analysieren/Mittelwerte vergleichen“ berechnet hat, geht sie direkt zur Menüleiste, wählt Analysieren und sucht nach der Funktion für Korrelationen. Beim Erstellen eines Boxplots ist erkennbar, dass der Computernutzer die Funktion für Balkendiagramm schnell in SPSS auffindet.

Versuch-und-Irrtum-Exploration (Kategoriebezeichnung)

Konstruktdefinition:

Der Computernutzer exploriert das System ohne erkennbare Anzeichen, dass er Hypothesen testet, plant oder Ergebnisse evaluiert. Die Wahl des nächsten Handlungsschrittes ist komplett unvorhersehbar auf der Grundlage vorheriger Mausclicks und Navigationen.

Ausgewählte Indikatoren:

- Der Computernutzer drückt zufällig Buttons oder ruft Menüoptionen auf ohne Bezug zum Aufgabenziel (erkennbares Ziel).
 - **Die Interaktion des Computernutzers mit dem Computer ist durch einen schnellen, häufigen Wechsel in der Navigation gekennzeichnet.** Neue Bedingungen werden ausprobiert (z.B. eine Variable erst in eine Box, dann in die andere gesetzt wird bzw. Häkchen angeklickt werden und dann wieder ausgeschaltet werden), jedoch ohne erkennbare Hypothese dahinter.
 - **Die Versuch-und-Irrtum-Exploration lässt sich von der Rigiden Exploration abgrenzen:** Wenn eine Variable immer wieder in die gleiche Box gesetzt wird, obwohl dieser Weg früher bereits zu einem Misserfolg geführt hat (oder bereits früher aufgegeben wurde), dann ist dies eine Handlung, die als Rigide Exploration einzustufen ist.
 - **Fehlermeldungen werden vom Computernutzer nicht oder nur in geringem Umfang beachtet:** Über Fehlermeldungen wird schnell hinweggegangen und eine neue Handlung
-

Versuch-und-Irrtum-Exploration (Kategoriebezeichnung) (Forts.)

wird begonnen, ohne das Zustandekommen der Fehlermeldung genauer zu erforschen.

- Das unsystematische Handeln des Computernutzers nach Versuch-und-Irrtum zeigt sich beispielsweise in einem schnellen Verschieben von Variablen in Auswahlboxen, die vorher nicht aufgerufen wurden. Der Beobachter merkt angesichts der Kürze der Zeit, dass die Person nicht über den Inhalt der Funktion nachgedacht haben kann.
-

Rigide Exploration (Kategoriebezeichnung)

Konstruktdefinition:

Der Computernutzer wiederholt häufig ineffektive Handlungen (d.h. Handlungen, die nicht zu einem Ergebnis geführt haben) mehr als zweimal hintereinander. Es wird erkennbar, dass die Person an „falschen Ideen“ oder „falschen Hypothesen“ festhält. Er testet diese mehrmals und scheint dabei nicht zu erkennen, dass sein Vorgehen nicht zum Ziel führt.

Ausgewählte Indikatoren:

- **Ineffektive Handlungen ohne verwertbares Ergebnis werden mehr als zweimal ausgeführt:** Beispielsweise versucht der Computernutzer mehr als zweimal auf den „ok“-Button zu klicken, ohne dass dieser aktiviert ist. Der Computernutzer ruft nach gescheitertem Versuch erneut die gleiche Funktion auf und ändert sein Vorgehen dabei nicht, z.B. versucht die Person wiederholt, Spalten zu markieren, ohne dass dies eine Auswirkung hat.
 - **Wenn der Computernutzer zum dritten Mal das gleiche Fenster aufruft (z.B. Variable berechnen) und er dann die gleiche Eingabe wählt, dann ist dies ein Anzeichen rigider Exploration.** Wenn jedoch nach der zweimaligen gleichen Eingabe, eine neue erfolgt, dass ist dies eher ein Hinweis auf systematische Exploration oder Versuch-und-Irrtum.
 - **Ein weiteres Kennzeichen rigider Exploration ist das vermehrte Anrufen salienter Programmfunktionen** (d.h. Funktionen, die sich auf der „Startseite“ des Programms befinden): Beispielsweise werden häufig *Shortcuts* unterhalb der Menüleiste angeklickt, kein Klick auf „Optionen“ im Menüpunkt Korrelationen.
 - **Der Computernutzer durchsucht bei derselben Aufgabe mehr als zweimal die Menüleiste, klickt jedoch nach dieser erneuten Recherche keine (neue) Unterfunktion an.** Beispielsweise werden die gleichen Reiter der Menüleiste mehr als zweimal hoch- und herunterfahren, ohne dass eine Funktion ausprobiert wird.
 - **Wird nach einer Fehlermeldung der gleiche Weg wieder beschritten, dann wurden Fehlermeldungen nicht beachtet.** Dies gilt ebenfalls als Anzeichen rigider Exploration.
 - **Nach der Durchführung von Berechnungen und dem Erstellen von Grafiken schaut der Computernutzer nicht in das Ausgabefenster, um neue Ergebnisse zur Kenntnis zu nehmen und zu evaluieren.** Andererseits zeigt sich die rigide Exploration auch darin, dass das Ausgabefenster mehrmals angeschaut wurde (auch zeitlich versetzt), obwohl Einstellungen nicht verändert wurden, d.h. kein neuer Output erstellt wurde (es sei denn, Ergebnis wurde noch nicht gefunden, dann: systematische Exploration).
 - **Das gleiche Ergebnis wird erneut erzeugt.** Dieses stellt kein exploratives Verhalten dar und wird hier als rigide gewertet. Beispielsweise wird der Befehl zur Durchführung eines Mittelwertvergleichs und das Ergebnis wird erneut im Output-Fenster angezeigt. Dabei gilt bereits die erste Wiederholung des gleichen Ergebnisses als rigide Exploration.
-

Rigide Exploration (Kategoriebezeichnung) (Forts.)

- Das gleiche Fenster wird wiederholt aufgerufen und es wurden nur minimale Veränderungen vorgenommen, z.B. Häkchen gesetzt und wieder weggenommen, wobei keine Konsequenzen dieser Veränderungen nicht getestet werden.

(Übermäßige) Informationssuche (Kategoriebezeichnung)

Konstruktdefinition:

Der Computernutzer verharrt in der Suche nach Informationen im Hilfemenü.

Ausgewählte Indikatoren:

- Der Computernutzer ruft wiederholt die Hilfefunktion und/oder das Lernprogramm auf (alle Funktionen unter „?“ oder „Hilfe“)
- Der Computernutzer verbringt „viel Zeit“ damit, alle angezeigten Informationen zu lesen, unabhängig davon, ob sie für die Lösung des Problems relevant sind.
- Im Hilfemenü „vagabundiert“ der Computernutzer von einem Thema zum anderen. Dabei verarbeitet der Informationen nicht systematisch, d.h. probiert Funktionen aus, über die er gerade gelesen hat.
- In Abgrenzung zur (übermäßigen) Informationssuche handelt es sich um systematische Informationssuche, wenn nach relevanten Begriffen gesucht wird, nur einmal das Fenster aufgerufen wird und der Computernutzer parallel ausprobiert, was der Computernutzer im Hilfemenü gelesen hat.

Unterbrechung (Kategoriebezeichnung)

Konstruktdefinition:

Unterbrechungen werden als solche definiert, bei denen Computernutzer mindestens 10 sec keine erkennbare Bildschirmaktivität zeigen.

Ausgewählte Indikatoren:

- **Aufgabenbedingte Unterbrechungen des Handlungsflusses sind auf Interventionen des Untersuchungsleiters zurückzuführen.** Diese kommen zustande, wenn der Computernutzer Aufgabenlösungen auf dem Aufgabenblatt notiert, die nächste zu bearbeitende Aufgabe liest oder einen Fragebogen ausfüllt.
 - **Programmbedingte Unterbrechungen gehen auf Fehlfunktionen des Programms zurück.** Störungen des Programmablaufs resultieren in zeitlichen Verzögerungen. Der Computernutzer kann in dieser Zeit keine Eingaben machen und keine Befehle ausführen (z.B. bei Absturz des Betriebssystems).
-

A 2. Anhang zu Studie 1

In den folgenden Abschnitten sind die in Studie 1 eingesetzten Materialien dargestellt (A 2.1) sowie Ergebnisse (zusätzlicher) statistischer Analysen aufgelistet (A 2.2).

A 2.1 Untersuchungsmaterialien

A 2.1.1 Fragebogen zu Computervorwissen und Statistikkenntnissen

Im Folgenden ist der in Studie 1 eingesetzte Fragebogen zu Computervorwissen (PRACOWI und VECA entnommen aus dem INCOBI; vgl. Richter et al., 2001) und Statistikkenntnissen (ad hoc entwickelt) dargestellt. Variationen sind fettgedruckt und in Klammern erläutert.

Datum: _____

Teilnehmer/-in (nummer): _____

Geschlecht weiblich männlich

Fachsemester **Psychologie**: _____

(für Wirtschaftswissenschafts-Studierende:

Fachsemester Wirtschaftswissenschaften: _____ **ich studiere WiPäd im Nebenfach)**

Berufsziel/-wunsch: _____

Alter: _____ Jahre

Teil I

1. Welche Statistik-Veranstaltungen haben Sie in der Vergangenheit besucht bzw. besuchen Sie gerade?

- Statistik 1 - Veranstaltung Statistik 1 - Tutorium Statistik/Stochastik im Schulunterricht
- Statistik 2 - Veranstaltung Statistik 2 - Tutorium Sonstige: _____

2. Bitte schätzen Sie auf der folgenden Skala Ihre theoretischen Kenntnisse in Statistik ein.

- sehr gut mangelhaft

3. Bitte geben Sie Ihre Note in der letzten Statistik-Klausur an (wenn vorhanden und wenn Sie möchten): _____

4. Bitte kreuzen Sie an, auf welche Inhalte Sie Ihr statistisches Wissen anwenden können:

- Deskriptive Statistik (z.B. Bestimmung der Maße zentraler Tendenz und von Streuungsmaßen)
- Inferenzstatistische Überprüfung von Hypothesen (z.B. Signifikanztests, t -Test)
- grafische Darstellung der statistischen Auswertungen (z.B. Histogramm anfertigen)

Bitte kreuzen Sie bei den folgenden Fragen zum Thema Statistik die richtige(n) Lösung(en) an. Beachten Sie, dass bei einer Frage mindestens eine Aussage richtig ist.

1. Der Median entspricht dem

- 2. Quartil
- 75 % - Perzentil
- 25 % - Perzentil
- 50 % - Perzentil

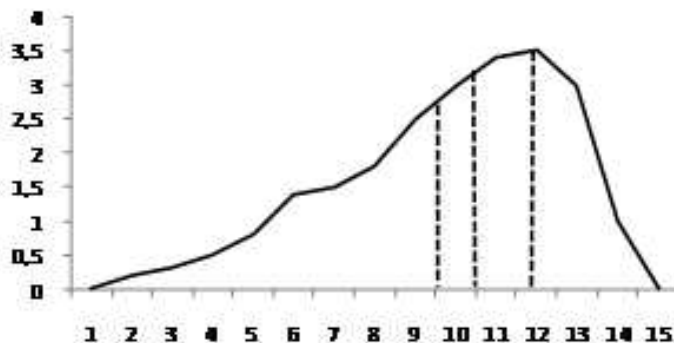
2. Varianz und Standardabweichung einer Verteilung zeigen an, ...

- wie stark die Werte der Verteilung um den Modalwert streuen.
- wie stark die Werte der Verteilung um den Median streuen.
- wie stark die Werte der Verteilung um den arithmetischen Mittelwert streuen.
- keine der drei vorangegangenen Antwortalternativen ist richtig.

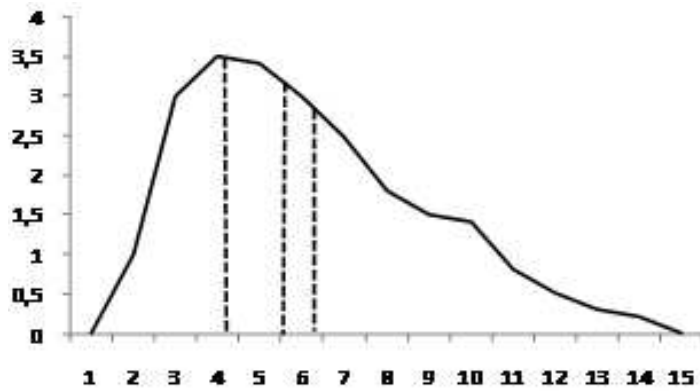
3. Empfindlich gegenüber Ausreißern unter den Beobachtungen ist....

- der Modalwert.
- das Arithmetische Mittel.
- der Median.
- die Streuung.

4. Bitte zeichnen Sie in die Verteilung Median, Modalwert und das Arithmetische Mittel ein und geben Sie an, ob es sich um eine links- oder rechtsschiefe Verteilung und um eine uni- oder bimodale Verteilung handelt.



Die oben dargestellte Verteilung ist _____-modal und _____-schief.



Die oben dargestellte Verteilung ist _____-modal und _____-schief.

5. Eine Verteilung kann mit Hilfe ...

- von Liniendiagrammen veranschaulicht werden.
- von Häufigkeitstabellen veranschaulicht werden
- des Mittelwerts veranschaulicht werden.
- von Kreisdiagrammen veranschaulicht werden.

6. Bei einer empirischen Häufigkeitsverteilung gelten folgende Annahmen:

- Wenn der Median größer ist als das arithmetische Mittel, dann weist dies auf eine linkssteile Verteilung hin.
- Wenn der Median größer ist als das arithmetische Mittel, dann weist dies auf eine rechtssteile Verteilung hin.
- Die zentrale Tendenz einer Verteilung, wo einige extreme Werte auftreten, beschreibt man am besten mit dem Modalwert.
- Wenn Median, Modalwert und Arithmetisches Mittel auf einen Wert fallen, ist die Verteilung symmetrisch.

7. Wenn von einem Ergebnis gesagt wird, es sei „statistisch signifikant“, dann bedeutet dies, dass...

- es praktisch relevant ist.
- seine Zufallswahrscheinlichkeit höchstens 5% beträgt.
- es wahrscheinlich richtig ist.
- es für 5% aller Personen zutrifft.

8. Wodurch ist eine symmetrische Verteilung gekennzeichnet?

- Die Flächen des Bereichs (Median – x) und (Median + x) sind annäherungsweise gleich.
- Der Median, Modalwert und der arithmetische Mittelwert fallen alle auf einen Wert.
- Die Verteilungsbreite ist links und rechts vom Mittelwert annähernd gleich.
- Sie muss unimodal (eingipflig) sein.

9. Für den Produkt-Moment-Korrelationskoeffizienten r gilt:

- $r = 1$; perfekter positiver linearer Zusammenhang
- $r = -1$; im Streudiagramm liegen alle Punkte exakt auf einer fallenden Geraden
- $-1 \leq r \leq 1$ (Normierung)
- $r = 0$; es besteht weder ein linearer noch ein kurvilinear (nicht linearer) Zusammenhang zwischen den Variablen.

Teil II

Im folgenden Teil des Fragebogens geht es um Ihre Vertrautheit mit verschiedenen Computeranwendungen. Sie sollen sich selbst daraufhin einschätzen, wie vertraut Sie im Umgang mit einzelnen Computeranwendungen sind. Bitte beurteilen Sie für jede der aufgelisteten Computeranwendungen jeweils, wie vertraut Sie im Vergleich zu anderen Studierenden mit der jeweiligen Anwendung sind.

	Weit überdurchschnittlich ++	Überdurchschnittlich +	Durchschnittlich 0	Unterdurchschnittlich -	Weit unterdurchschnittlich --
1. Textverarbeitungsprogramme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Multimedia-Anwendungen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Programmiersprachen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Tabellenkalkulation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Statistik-Programme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. E-Mail	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Datenbanken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Internet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Computerspiele	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Grafikprogramme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Teil III

Im nächsten Fragebogenteil geht es um praktisches Computerwissen, d.h. um solches Wissen, das für den Umgang mit dem Computer unmittelbar relevant sein kann. Im Folgenden finden Sie Problembeschreibungen, mit denen man bei der täglichen Arbeit mit dem Computer konfrontiert sein oder zu tun haben kann.

Ihre Aufgabe ist es, für jede geschilderte Problemsituation die eine Handlungsalternative anzukreuzen, die Ihrer Einschätzung nach die beste Möglichkeit darstellt, mit dem Problem umzugehen.

Da die Fragen sich zum Teil auf bestimmte Betriebssysteme beziehen, geben Sie bitte an, mit welchem Betriebssystem/welchen Betriebssystemen Sie persönlich arbeiten und bezüglich welcher Betriebssysteme Sie Kenntnisse besitzen (z.B. weil Sie früher einmal damit gearbeitet haben).

- | | | |
|-----------------------------|--|--------------------------|
| Ich arbeite mit... | <input type="checkbox"/> Windows '95 oder höher (98, NT, ME, XP, Vista...) | <input type="checkbox"/> |
| (Mehrfachnennungen möglich) | <input type="checkbox"/> Windows 3.x | <input type="checkbox"/> |
| | <input type="checkbox"/> OS/2 | <input type="checkbox"/> |
| | <input type="checkbox"/> DOS | <input type="checkbox"/> |
| | <input type="checkbox"/> Unix | <input type="checkbox"/> |
| | <input type="checkbox"/> Apple Macintosh | <input type="checkbox"/> |
| Ich habe Kenntnisse in... | <input type="checkbox"/> Windows '95 oder höher (98, NT, ME, XP, Vista...) | <input type="checkbox"/> |
| (Mehrfachnennungen möglich) | <input type="checkbox"/> Windows 3.x | <input type="checkbox"/> |
| | <input type="checkbox"/> OS/2 | <input type="checkbox"/> |
| | <input type="checkbox"/> DOS | <input type="checkbox"/> |
| | <input type="checkbox"/> Unix | <input type="checkbox"/> |
| | <input type="checkbox"/> Apple Macintosh | <input type="checkbox"/> |

1. Sie haben in Word Änderungen an einem Textdokument vorgenommen, und möchten sowohl die geänderte Datei speichern als auch die ursprüngliche Version beibehalten. Was tun Sie?

- (a) Ich rufe in der Textverarbeitung den Menüpunkt "Versionsvergleich" auf.
- (b) Ich verschiebe die Datei vor dem Speichern in ein anderes Verzeichnis.
- (c) Ich speichere die geänderte Datei unter einem neuen Namen.
- (d) Ich wähle in Word den Menüpunkt "Änderungen in einer neuen Datei speichern".
- weiß ich nicht*

2. Sie wissen, dass ein bestimmtes Programm auf Ihrem Computer installiert ist, Sie können es aber nicht auf die gewohnte Weise starten. Wie können Sie dieses Problem beheben?

- (a) Ich installiere das Programm neu.
- (b) Ich suche das Programmverzeichnis im Explorer/Dateimanager und rufe die *.exe-Datei des Programms mit einem Doppelklick auf.
- (c) Programme, die nicht "normal" aufrufbar sind, sind in Verbindung mit dem benutzten Betriebssystem nicht lauffähig. Ich besorge eine neue Version des Programms für mein Betriebssystem.
- (d) Ich füge unter "Systemsteuerung" das Programm ins "Programmverzeichnis" ein.
- weiß ich nicht*

3. Sie bekommen ein als ZIP-Archiv gepacktes Textdokument. Wie verfahren Sie damit?

- (a) Ich öffne das Dokument wie gewohnt in meinem Textverarbeitungsprogramm.
- (b) Ich entpacke das Programm über den Explorer/Dateimanager.
- (c) Ich ziehe die Datei mit der Maus auf den Desktop, dabei wird sie automatisch entpackt.
- (d) Ich ändere die Erweiterung der Datei von *.zip in die Word-Erweiterung *.doc. Danach lässt sie sich in Word öffnen.
- weiß ich nicht*

4. Sie haben über das Internet eine als selbstextrahierendes Archiv gepackte Textdatei auf Ihren Computer geladen. Nun möchten Sie diese lesen. Was tun Sie?

- (a) Die Datei entpackt sich nach dem Laden selbst auf der Festplatte. Danach kann ich sie in der Textverarbeitung wie gewohnt aufrufen und lesen.
 - (b) Ich öffne die Datei in meinem Textverarbeitungsprogramm über den Menüpunkt "Extras – Archiv entpacken".
 - (c) Ich öffne das Dokument wie gewohnt in meinem Textverarbeitungsprogramm.
 - (d) Ich entpacke die Datei per Doppelklick im Explorer/Dateimanager und öffne sie dann im Textverarbeitungsprogramm.
- weiß ich nicht*
-

5. Sie wollen den Computer starten. Es erscheint kurz nach dem Einschalten die Meldung "Disk not ready. Insert bootable disk and press any key" bzw. "This disk can't boot. It was formatted without the /s (system-) option". Was tun Sie?

- (a) Offensichtlich versagt die Festplatte. Ich muss den Computer zum Händler bringen.
 - (b) Anscheinend sind die Einstellungen im Setup fehlerhaft. Ich rufe das System-Setup auf und ändere die Einstellungen für die Festplatte.
 - (c) Ich habe eine Diskette im Diskettenlaufwerk vergessen. Ich entferne die Diskette aus dem Laufwerk und drücke die "Enter"-Taste.
 - (d) Ich führe einen Warmstart des Computers durch.
- weiß ich nicht*
-

6. Ihr Computer ist abgestürzt und Sie wollen ihn möglichst "schonend" neu starten. Was tun Sie?

- (a) Ich drücke den "reset"-Knopf.
 - (b) Ich drücke die Tastenkombination "Strg"+"Alt"+"Entf".
 - (c) Ich drücke die Tastenkombination "Ende"+"Enter".
 - (d) Ich schalte den Computer aus und wieder ein.
- weiß ich nicht*
-

7. Ihre Maus ist ausgefallen und Sie wollen das Programm, das Sie geöffnet haben, beenden. Was tun Sie?

- (a) Ich beende das Programm, indem ich die Tastenkombination "Strg" + "Ende" (bei englischen/amerikanischen Tastaturen "Ctr" + "End") drücke. Alternativ kann das Programm mit der Tastenkombination "Alt" + "F3" beendet werden.
 - (b) Ich beende das Programm, indem ich die Taste "Strg" (englisch/amerikanische Tastatur: "Ctr") gedrückt halte, und dabei die Tastenkombination "Ende" + "Enter" ("End" + "Enter") drücke. Alternativ kann das Programm mit der Tastenkombination "Alt" + "F6" beendet werden.
 - (c) Ich beende das Programm, indem ich gleichzeitig "Shift" und "Ende" (bzw. "End") drücke. Alternativ kann das Programm mit der Tastenkombination "Alt" + "F5" beendet werden.
 - (d) Ich beende das Programm, indem ich die "Alt"-Taste gedrückt halte und dabei nacheinander die Tasten "D" und "B" (bei englischsprachigen Programmen "F" und "X") drücke. Alternativ kann das Programm mit der Tastenkombination "Alt" + "F4" beendet werden.
- weiß ich nicht*
-

A 2.1.2 Aufgabenblatt

Im Folgenden ist das Arbeitsblatt zu Studie 1 wiedergegeben. Die statistischen Indizes, die auf dem Arbeitsblatt aufgeführt sind, wurden über die Entwicklung zweier unterschiedlicher Fragebogenversionen für die jeweilige Stichprobe (Psychologie vs. Wirtschaftspädagogik-Studierende) angepasst. Variationen sind durch Fettdruck gekennzeichnet und in Klammern erläutert.

Arbeitsblatt

- Bearbeiten Sie die 5 Aufgaben bitte der Reihe nach.
- Bitte schließen Sie das Ausgabefenster nicht, sondern minimieren Sie es.
- Bitte tragen Sie vor Beginn einer neuen Aufgabe, die Uhrzeit ein.
- Bitte positionieren Sie den Mauszeiger an der Uhr während Sie Lösungen auf dem Blatt eintragen oder Fragebögen ausfüllen.

Aufgabe 1 (Uhrzeit bei Beginn: _____)

Bestimmen Sie den Mittelwert (\bar{X} bzw. M) und die Standardabweichung (s ; Fragebogenversion Psychologie-Studierende: SD) für die Variable „Wort- und Redegewandtheit“ in der gesamten Stichprobe. Tragen Sie die Ergebnisse bitte auf dem Arbeitsblatt ein.

Mittelwert und Standardabweichung:

M	s (Fragebogenversion Psychologie-Studierende: SD)
_____	_____

Aufgabe 2 (Uhrzeit bei Beginn: _____)

In einem kleinen Unternehmen mit 36 Mitarbeiterinnen und Mitarbeitern möchte sich die Personalverantwortliche einen Überblick über die Gedächtnisfähigkeiten der Mitarbeiterinnen und Mitarbeiter verschaffen.

- Erstellen Sie hierzu eine Häufigkeitstabelle für die Variable „Zahlengedächtnis“!
- Die Personalverantwortliche beschließt, allen Mitarbeiter/-innen mit einem Wert unter 5 Punkten im Zahlengedächtnistest ein Angebot für ein Gedächtnistraining zu machen.

Wie vielen Personen macht sie dieses Angebot?

Überlegen Sie und tragen Sie das Ergebnis hier ein:

	Prozent der Mitarbeiter/-innen	Anzahl der Personen mit weniger als 5 Punkten
Zahlengedächtnistest	_____	_____

Aufgabe 3 (Uhrzeit bei Beginn: _____)

Die Personalverantwortliche in dem kleinen Unternehmen möchte jetzt gerne wissen, ob es zwischen der „Kooperationsfähigkeit“ der Mitarbeiter/-innen und dem „Teamgeist“ einen positiven Zusammenhang gibt.

a. Berechnen Sie die **Korrelation** zwischen den beiden Variablen, geben Sie die relevanten Kennwerte an und interpretieren Sie das Ergebnis.

Ergebnis und Interpretation:

b. Jetzt will sich die Personalverantwortliche die Beziehung der Variablen „Teamgeist“ und „Kooperationsfähigkeit“ noch einmal **grafisch** darstellen lassen. Erstellen Sie ein einfaches **Streudiagramm** für die beiden Variablen. (Lösung in SPSS)

Aufgabe 4 (Uhrzeit bei Beginn: _____)

Versuchen Sie für die Variable „Wort- und Redegewandtheit“ die vorliegende Abbildung möglichst genau zu reproduzieren, d.h. sich die folgende Abbildung in SPSS anzeigen zu lassen. (Lösung in SPSS. Die Farbe spielt dabei keine Rolle!)

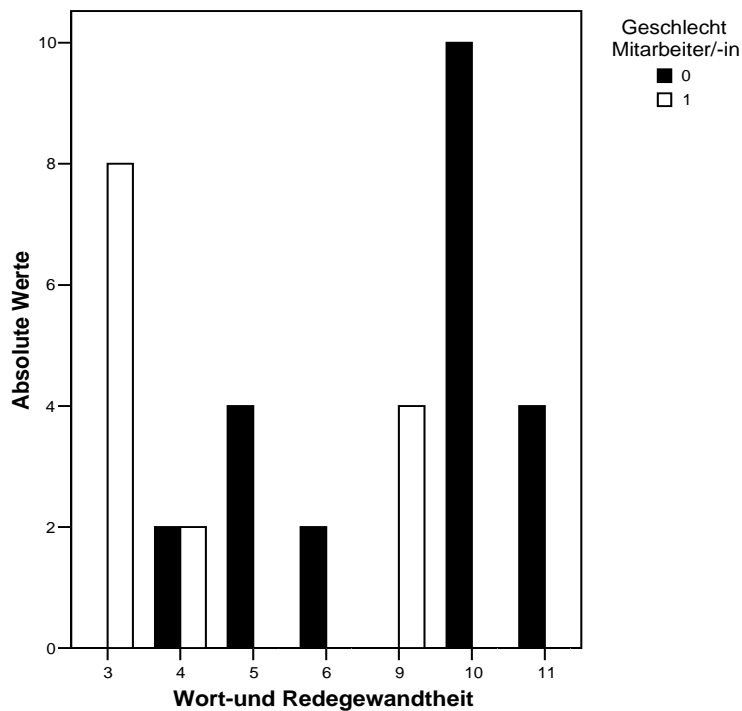


Abbildung zu Aufgabe 4.

Aufgabe 5 (Uhrzeit bei Beginn: _____)

Eine Sozialwissenschaftlerin hat Daten zur Überzeugungskraft erhoben. Sie möchte die zentralen Kennwerte für die Verteilung der Variablen „Überzeugungskraft“ berechnen.

Beschreiben Sie die Verteilung, indem Sie den Mittelwert (\bar{X} bzw. M) und den maximalen Wert in der Verteilung (Maximum) angeben. Als Streuungsmaße geben Sie bitte die Standardabweichung, die Varianz und den Range (Spannweite) in der Variablen an.

Mittelwert: _____

Maximum: _____

Standardabweichung: _____

Varianz: _____

Range: _____

A 2.2 Ergebnisse statistischer Analysen

In Kapitel A 2.2.1 sind die Ergebnisse der Item- und Skalenanalysen für die in Studie 1 verwendeten Verfahren separat tabellarisch dargestellt. Die Ergebnisse der Voraussetzungsprüfungen für die statistischen Analysen sind in Kapitel A 2.2.2 aufgeführt. Kapitel A 2.2.3 bietet einen Überblick über die Ergebnisse zusätzlicher statistischer Analysen.

A 2.2.1 Item- und Skalenanalysen

Tabelle A – 2

Ergebnis der Item- und Skalenanalyse des Fragebogens zu Statistikkenntnissen mit Cronbachs $\alpha = .42$ (alle 9 Items) bzw. $.62$ (ohne Items 4, 6 und 8; $N = 18$)

Item-Nr.	Itemwortlaut (abgekürzt) ¹	<i>M</i>	<i>SD</i>	<i>P</i> ²	<i>r</i> _{it}	Schiefe	Exzess
1	Der Median entspricht dem	3.33	0.5	83	.41	0.77	-1.59
2	Varianz und Standardabweichung einer Verteilung zeigen an...	3.78	0.6	95	.38	-2.71	5.98
3	Empfindlich gegenüber Ausreißern unter den Beobachtungen ist...	2.56	0.9	64	.68	-0.94	2.46
4	<i>Bitte zeichnen Sie in die Verteilung Median, Modalwert und Arithmetisches Mittel...</i>	6.22	2.2	62	.11	0.64	-0.97
5	Eine Verteilung kann mit Hilfe von... veranschaulicht werden.	2.61	1.1	65	.12	-0.01	-1.27
6	<i>Bei einer empirischen Häufigkeitsverteilung gelten folgende Annahmen...</i>	3.11	0.8	78	.04	-0.91	1.04
7	Wenn von einem Ergebnis gesagt wird, es sei „statistisch signifikant“ bedeutet dies...	3.00	1.0	75	.40	-0.36	-1.37
8	<i>Wodurch ist eine symmetrische Verteilung gekennzeichnet?</i>	2.33	1.1	58	-.19	0.48	-0.95
9	Für den Produkt-Moment-Korrelationskoeffizienten <i>r</i> gilt...	1.94	0.7	49	.20	0.09	-0.90

Anmerkungen. *M* = Itemmittelwert; *SD* = Standardabweichung des Items; *P* = Schwierigkeit in %; *r*_{it} = Trennschärfe. ¹Die aus der Skala eliminierten Items (nach Item- und Skalenanalyse) sind kursiv gesetzt. ²In dem Multiple-Choice-Test mit mindestens einer richtigen Antwort bei insgesamt 4 Antwortalternativen pro Aufgabe konnten zwischen 0 und 4 Punkten erreicht werden. Eine Ausnahme bildet Item 4 mit maximal 10 Punkten. Punkte wurden sowohl für das Ankreuzen richtigen Antworten als auch für das Freilassen falscher Antworten vergeben. Auf diesem Weg sollte eine Ratekorrektur erfolgen. Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 3

Ergebnis der Item- und Skalenanalyse der ausgewählten Items der Skala Praktisches Computerwissen PRACO-WI mit Cronbachs $\alpha = .52$ (alle 6 Items) bzw. $.57$ (5 Items; $N = 18$)¹

Item-Nr.	Itemwortlaut (abgekürzt) ²	M^3	SD	P^4	r_{it}^5
1	Sie wissen, dass ein bestimmtes Programm auf Ihrem Computer installiert ist...	0.24	0.4	28	.02
2	Sie bekommen ein als ZIP-Archiv gepacktes Textdokument...	0.44	0.5	44	.29
3	Sie haben über das Internet eine als selbstextrahierendes Archiv gepackte Textdatei auf Ihren Computer geladen...	0.28	0.5	28	.50
4	Sie wollen den Computer starten...	0.56	0.5	56	.18
5	Ihr Computer ist abgestürzt...	0.58	0.5	56	.42
6	Ihre Maus ist ausgefallen ...	0.06	0.2	6	.35

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ Schiefe und Exzess wurden nicht berechnet, da dichotome Antwortkategorien vorlagen („richtig“, „falsch“). ² Die aus der Skala eliminierten Items (nach Item- und Skalenanalyse) sind kursiv gesetzt. ³ In dem Multiple-Choice-Test mit genau einer richtigen Antwort bei insgesamt 4 Antwortalternativen konnte maximal 1 Punkt pro Aufgabe erreicht werden. ⁴ Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ⁵ Die Korrelation zwischen dichotomen Items und einem intervallskalierten Merkmal machen eine punktbiseriale Korrelation erforderlich (Bühner, 2006, S. 97f.; Diehl & Kohr, 1994, S. 261). Eine punktbiseriale Korrelation erhält man, in dem „in die Gleichung für die Produkt-Moment-Korrelation ... für das dichotome Merkmal die Werte 0 und 1 eingesetzt werden“ (Bortz, 2005, S. 224). Die Trennschärfe der dichotomen Items mit den beiden Ausprägungen 1 = „Aufgabe gelöst“ und 0 = „Aufgabe nicht gelöst“, wurde demzufolge auf der Grundlage des Pearson-Produkt-Moment-Korrelationskoeffizienten bestimmt.

Tabelle A – 4

Ergebnis der Item- und Skalenanalyse der ausgewählten Items der Skala Vertrautheit mit Computeranwendungen VECA mit Cronbachs $\alpha = .77$ (alle 10 Items; $N = 18$)

Item-Nr.	Itemwortlaut (abgekürzt)	M^1	SD	P^2	r_{it}	Schiefe	Exzess
1	Textverarbeitungsprogramme	1.94	0.9	49	.70	0.12	0.62
2	Multimedia-Anwendungen	1.76	0.9	44	.75	-0.16	-0.56
3	Programmiersprachen	0.50	0.8	13	.44	2.05	5.20
4	Tabellenkalkulation	1.28	1.1	32	.43	0.02	-1.38
5	Statistik-Programme	0.39	0.6	10	.12	1.36	1.13
6	E-Mail	2.56	1.0	64	.12	0.67	-1.04
7	Datenbanken	1.38	1.1	35	.48	0.16	-1.36
8	Internet	2.50	0.7	63	-.32	0	0.12
9	Computerspiele	1.06	1.1	26	.69	1.23	2.18
10	Grafikprogramme	1.06	1.2	26	.83	1.00	0.38

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 5-stufiger Antwortmodus von „weit unterdurchschnittlich“ (0) bis „weit überdurchschnittlich“ (4). ² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A - 5

Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den ersten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit (N = 18)

Item-Nr.	Skala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Glatter Verlauf ($\alpha = .72$)							
2	Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	4.22	1.7	70	.48	-0.31	-0.76
4	Ich habe keine Mühe, mich zu konzentrieren.	4.78	1.9	78	.32	-0.98	-0.10
5	Mein Kopf ist völlig klar.	4.67	1.6	78	.57	-0.53	-0.10
7	Die richtigen Gedanken/Bewegungen kommen wie von selbst.	3.83	1.8	64	.61	0.01	-1.28
8	Ich weiß bei jedem Schritt, was ich zu tun habe.	2.44	1.2	41	.42	1.07	0.64
9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	3.17	1.3	53	.38	0.57	-0.04
Absorbiertheit ($\alpha = .53$)							
1	Ich fühle mich optimal beansprucht.	4.44	1.7	74	.63	-0.45	0.19
3	Ich merke gar nicht, wie die Zeit vergeht.	5.22	1.6	87	-.07	-0.86	-0.20
6	Ich bin ganz vertieft in das was ich gerade mache.	5.33	1.3	89	.70	-0.52	-0.73
10	Ich bin völlig selbstvergessen.	3.61	1.2	60	.24	0.87	2.94

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³ Getrennt für die beiden Subskalen.

Tabelle A – 6

Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den zweiten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit ($N = 14$)

Item-Nr.	Skala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Glatter Verlauf ($\alpha = .81$)							
2	Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	4.07	1.2	68	.41	-0.46	-0.55
4	Ich habe keine Mühe, mich zu konzentrieren.	4.79	1.3	80	.52	-1.93	5.07
5	Mein Kopf ist völlig klar.	4.29	1.3	71	.66	-1.08	1.73
7	Die richtigen Gedanken/Bewegungen kommen wie von selbst.	3.50	1.2	58	.67	0.0	-1.62
8	Ich weiß bei jedem Schritt, was ich zu tun habe.	2.57	1.1	43	.61	-0.21	-1.12
9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	2.93	1.4	49	.57	-0.26	-1.51
Absorbiertheit ($\alpha = .58$)							
1	Ich fühle mich optimal beansprucht.	4.50	1.3	75	.09	0.0	-0.14
3	Ich merke gar nicht, wie die Zeit vergeht.	4.21	1.6	70	.44	0.41	-0.35
6	Ich bin ganz vertieft in das was ich gerade mache.	5.07	0.8	85	.52	0.80	1.16
10	Ich bin völlig selbstvergessen.	3.50	1.1	58	.60	0.0	-1.18

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7).

² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

³ Getrennt für die beiden Subskalen.

Tabelle A – 7

Ergebnis der Item- und Skalenanalyse für die Leistung bei der Bearbeitung der fünf Aufgaben des Arbeitsblattes
(Variable: „Leistung_Arbeitsblatt“; Cronbachs $\alpha = .82$; $N = 18$)

Item -Nr.	Aufgabe mit Itemwortlaut (abgekürzt) ¹	Max.	<i>M</i>	<i>SD</i>	<i>P</i> ²	<i>r</i> _{it}	Schiefe	Exzess
1	Angabe von Mittelwert und Standardabweichung für die Gesamtstichprobe, sowie getrennt für Frauen und Männer (<i>n</i> = 18)	2	1.3	1.0	57	.55	-0.77	-1.59
2	Erstellen einer Häufigkeitstabelle (<i>n</i> = 14) 1 Punkt: korrekte Tabelle 1 Punkt: korrekte Tabelle, korrekte Variable Angabe von absoluter und relativer Häufigkeit aus einer Häufigkeitstabelle (2 Punkte)	4	1.6	1.2	50	.54	-0.15	-1.52
3	Berechnung einer Korrelation (<i>n</i> = 12) 1 Punkt: Korrelationskoeffizient 1 Punkt: Signifikanzniveau 2 Punkt: Interpretation des Ergebnisses Erstellen eines Streudiagramms 0 Punkte: keine Grafik 1 Punkt: zu den Daten passende Grafik 2 Punkte: korrektes Streudiagramm	6	2.8	2.3	70	.80	-0.09	-1.58
4	Erstellen eines gruppierten Balkendiagramms (<i>n</i> = 12) 1 Punkt: zu den Daten passende Grafik 2 Punkte: Balkendiagramm 3 Punkte: Gruppiertes Balkendiagramm	3	1.2	1.3	60	.86	0.25	-1.84
5	Berechnung von Deskriptiven Statistiken (<i>n</i> = 8) (Mittelwert, Maximum, Standardabweichung, Varianz, Range)	5	1.9	2.4	88	.60	0.50	-1.82

Anmerkungen. Max. = Maximale Punktzahl; *M* = Itemmittelwert; *SD* = Standardabweichung des Items; *P* = Schwierigkeit in %; *r*_{it} = Trennschärfe. ¹ *n* gibt die Anzahl der Personen wieder, welche die Aufgabe bearbeitet haben. ² Mit Inangriffnahme-Korrektur; d.h. einbezogen wurden nur die Werte von Teilnehmern, welche die Aufgabe bearbeitet haben. Höhere Werte geben eine niedrigere Schwierigkeit wider.

A 2.2.2 Voraussetzungsprüfungen

Tabelle A – 8 fasst die Ergebnisse der statistischen Voraussetzungsprüfungen der in Studie 1 verwendeten Verfahren zusammen. Zusätzlich sind die deskriptiven Kennwerte angegeben.

Tabelle A – 8

Deskriptive Statistiken für die Verteilung der Skalensummenwerte inklusive der Ergebnisse des Kolmogorov-Smirnov-Tests auf Normalverteilung (N = 18)

Skala	M	SD	Schiefe	Exzess	Z	Asymp.
Subjektive Einschätzung Statistikkenntnisse ¹	2.89	0.8	0.22	-1.52	1.04	.23
Statistikvorwissenstest (6 Items) ¹	2.89	0.5	-0.55	2.11	0.73	.66
VECA ¹	1.44	0.5	0.48	-0.89	0.73	.66
PRACOWI (5 Items) ¹	0.38	0.3	0.76	-0.04	0.84	.48
Flow_end ¹	3.06	0.8	-0.06	-0.49	0.52	.95
Explorationsstrategie ²						
Systematische Exploration	44	21	0.05	-1.29	0.51	.96
Versuch-und-Irrtum	7	8	1.16	0.59	0.98	.30
Rigide Exploration	23	17	0.51	-0.72	0.51	.96
Informationssuche	6	9	1.65	1.37	1.37	.05
Leistung (Arbeitsblatt)	8.86	6.6	0.01	-1.23	0.56	.91

Anmerkungen. Z = Prüfgröße Kolmogorov –Smirnov-Tests; Asymp. = Asymptotische Signifikanz. ¹Es wird der Skalennittelwert auf Itemebene berichtet (d.h. die Summe der Itemwerte geteilt durch die Anzahl der Items). ²Dargestellt ist jeweils die mittlere, relative Häufigkeit (%) der Explorationsstrategie über den gesamten Messzeitraum von 25 Minuten.

A 2.2.3 Ergebnisse ergänzender statistischer Analysen zu den Hypothesen

Tabelle A – 9 stellt eine Ergänzung zu den in Abschnitt 7.2.3 (zu Hypothese 2) berichteten Analyseergebnisse dar. Die Tabelle gibt die Partialkorrelationen der Variablen wieder, für die Bedingung, dass der Einfluss des Vorwissens konstant gehalten wird.

Tabelle A – 9

Partialkorrelationen der Variablen in der Stichprobe (N = 18) unter Konstanthaltung des Vorwissens (PRACOWI, Statistikvorwissenstest)

Variable	Interkorrelationen				
	2	3	4	5	6
1 Systematische Exploration ¹	-.23	-.83**	-.66**	.14	.66**
2 Versuch-und-Irrtum ¹		-.20	-.13	.24	.30
3 Rigide Exploration ¹			.59*	-.48	-.80**
4 Informationssuche ¹				-.21	-.70**
5 Flow_End					.45
6 Performanz					

Anmerkungen. * $p < .05$, ** $p < .01$; zweiseitig. ¹Relative Häufigkeit der Explorationsstrategie über den gesamten Messzeitraum (25 Minuten).

A 3. Anhang zu Studie 2

Im Abschnitt A 3.1 sind die Untersuchungsmaterialien, die ausschließlich in Studie 2 verwendet wurden, aufgeführt. Der darauf folgende Abschnitt A 3.2 fasst in tabellarischer Form die Ergebnisse der Item- und Skalenanalysen sowie der statistischen Voraussetzungsprüfungen zusammen.

A 3.1 Untersuchungsmaterialien

A 3.1.1 Fragebogen zu Computervorwissen und Statistikkenntnissen

Teil I

1. Wie häufig haben Sie in den letzten 6 Monaten mit einer der folgenden Computeranwendungen gearbeitet?

Statistikprogramm
(z.B. SPSS, Stata, Eview): nie bis 10 mal bis 20 mal häufiger als 20 mal

Tabellenkalkulation
(z.B. Excel): nie bis 10 mal bis 20 mal häufiger als 20 mal

Datenbank
(z.B. Access): nie bis 10 mal bis 20 mal häufiger als 20 mal

2. Welche Veranstaltungen zu Statistik haben Sie in der Vergangenheit besucht?

- Statistik 1 - Vorlesung Statistik 1 - Tutorium Statistik im Schulunterricht
 Statistik 2 - Vorlesung Statistik 2 - Tutorium Sonstige: _____

3. Beschäftigen Sie sich über den Inhalt der Lehrveranstaltungen hinaus mit statistischen Problemen und Aufgaben?

- nein ja

4. Haben Sie in den letzten 6 Monaten eine Statistik-Lernsoftware benutzt?

- nein ja - Welche Statistik-Lernsoftware haben Sie benutzt?

5. Bitte schätzen Sie auf der folgenden Skala Ihre theoretischen Kenntnisse in Statistik ein. Orientieren Sie sich bei Ihrer Einschätzung an Ihren Leistungen in (Übungs-)Klausuren, Tests etc.

sehr gute Kenntnisse mangelhafte Kenntnisse

6. Bitte geben Sie an, auf welche Inhalte Sie Ihr statistisches Wissen anwenden können:

- Deskriptive Statistik (z.B. Bestimmung der Maße zentraler Tendenz und von Streuungsmaßen)
- Inferenzstatistische Überprüfung von Hypothesen
- grafische Darstellung der statistischen Auswertungen (z.B. Histogramm anfertigen)

Teil II (Statistikvorwissenstest)

Bitte kreuzen Sie bei den folgenden Fragen zum Thema Statistik die richtige(n) Lösung(en) an. Beachten Sie bitte, dass bei einer Aufgabe mindestens eine Aussage zutrifft.

7. Der Median entspricht dem

- 2. Quartil.
- 75 % Quantil.
- 25 % Quantil.
- 50 % Quantil.

8. Varianz und Standardabweichung einer Verteilung zeigen an, ...

- wie stark die Werte der Verteilung um den Modalwert streuen.
- wie stark die Werte der Verteilung um den Median streuen.
- wie stark die Werte der Verteilung um den arithmetischen Mittelwert streuen.
- keine der drei vorangegangenen Antwortalternativen ist richtig.

9. Empfindlich gegenüber Ausreißern unter den Beobachtungen ist....

- der Modus (Modalwert).
- das Arithmetische Mittel.
- der Median.
- die Streuung.

10. Wenn von einem Ergebnis gesagt wird, es sei „statistisch signifikant“, dann bedeutet dies, dass...

- es praktisch relevant ist.
- seine Zufallswahrscheinlichkeit höchstens 5% beträgt.
- es wahrscheinlich richtig ist.
- es für 5% aller Personen zutrifft.

11. Für den Produkt-Moment-Korrelationskoeffizienten r gilt:

- $r = 1$; perfekter positiver linearer Zusammenhang.
- $r = -1$; alle Produkte liegen exakt auf einer fallenden Geraden.

- $-1 \leq r \leq 1$ (Normierung)
- $r = 0$; es besteht weder ein linearer noch ein kurvilinearere Zusammenhang zwischen den Variablen.

12. Welches Skalenniveau ist mindestens erforderlich, damit folgende Aussagen zutreffen? Bitte kreuzen Sie rechts das Skalenniveau an, das für die jeweilige Aussage zutrifft.

	Nominal	Ordinal	Intervall (metrisch)	Verhältnis
Ein Messwert ist größer als ein anderer.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ein Messwert ist doppelt so groß.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Differenz zweier Messwerte ist gleich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zwei Messwerte sind ungleich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
$(x_1 - x_2) < (x_3 - x_4)$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ein Messwert liegt zwischen zwei anderen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

13. Zwei Zufallsvariablen X und Y sind korreliert mit einer Kovarianz von $\text{Cov}(X, Y) = 7$. Ihre Varianzen betragen $\sigma^2_x = 16$ und $\sigma^2_y = 25$. Dann gilt:

- $\text{Var}(X + Y) = 63$
- $\text{Var}(X + Y) = 41$
- $\text{Var}(X - Y) = 63$
- $\rho_{xy} = .35$

14. Bei einer empirischen Häufigkeitsverteilung gelten folgende Annahmen:

- Wenn $\text{Median } X_{0.5} < \bar{X}$ (Arithmetischer Mittelwert), dann weist dies auf eine linkssteile Verteilung hin.
- Ein Schiefekoeffizient von $\gamma_1 > 0$ weist auf eine rechtssteile Verteilung hin.
- Die Verteilung ist symmetrisch, wenn $| X_{0.5} - X_p |$ ungefähr gleich $| X_{0.5} - X_{1-p} |$.
- Für den Schiefekoeffizienten (γ_1) gilt definitionsgemäß: $\gamma_1 \geq 0$.

15. Ein Wissenschaftler hat in einer Studie die Ausprägungen von zwei (empirischen) Merkmalen X und Y erhoben; r_{xy} bezeichne den Korrelationskoeffizienten bei n Beobachtungspaaren. Welche der folgenden Annahmen kann der Wissenschaftler aufstellen?

- Aus $r_{xy} = 0$ folgt, dass zwischen den Merkmalen kein linearer Zusammenhang besteht.
- Ein Korrelationskoeffizient von $r_{xy} = -.90$ deutet auf einen starken linearen Zusammenhang von X und Y hin.
- Das Bestimmtheitsmaß von $R^2 = .81$ für $r_{xy} = -.90$ bedeutet, dass 81 % der Streuung der y-Werte aufgrund der Streuung der x-Werte erklärbar sind.
- Der Wert $r_{xy} = .85$ bedeutet, dass in 85 % der Beobachtungen $x_i = y_i$ gilt.

A 3.1.2 Aufgabenblatt

Arbeitsblatt

WICHTIG:

- Bitte bearbeiten Sie alle Aufgaben mit SPSS; alle Aufgaben sind mit SPSS lösbar.
- Bearbeiten Sie die 6 Aufgaben bitte der Reihe nach.
- Bitte tragen Sie vor dem Beginn einer Aufgabe die Uhrzeit auf dem Arbeitsblatt ein.
- Schließen Sie das Ausgabefenster nicht, sondern minimieren sie es.
- Bitte positionieren Sie den Mauszeiger an der Uhr während Sie Lösungen auf dem Blatt eintragen oder Fragebögen ausfüllen.

Aufgabe 1 (Uhrzeit bei Beginn: _____)

Bestimmen Sie den Mittelwert (\bar{X}) und die Standardabweichung (s) für die Variable „Anfangsgehalt“ in der gesamten Stichprobe und getrennt für die weiblichen und für die männlichen Personen. Tragen Sie die Ergebnisse bitte auf dem Arbeitsblatt ein:

	\bar{X}	s
<u>Gesamtstichprobe</u> Mittelwert und Streuung:	_____	_____
<u>Weibliche Personen</u> Mittelwert und Streuung:	_____	_____
<u>Männliche Personen</u> Mittelwert und Streuung:	_____	_____

Aufgabe 2 (Uhrzeit bei Beginn: _____)

Ein Wirtschaftsinstitut erforscht die Entwicklung der Gehälter vom Anfangsgehalt beim Einstieg in den Beruf bis zu dem Zeitpunkt, zu dem die Person 10 Jahre im Unternehmen gearbeitet hat. Frage: Gibt es einen Zusammenhang zwischen dem Anfangsgehalt und dem Gehalt, das ein/-e Arbeitnehmer/-in nach 10 Jahren Betriebszugehörigkeit verdient?

Bitte berechnen Sie die Korrelation zwischen den beiden Variablen, geben Sie die relevanten Kennwerte an und interpretieren Sie das Ergebnis.

Ergebnis und Interpretation

Aufgabe 3 (Uhrzeit bei Beginn: _____)

Das Statistische Bundesamt möchte ermitteln, ob sich Männer und Frauen in der Variablen „Anfangsgehalt“ unterscheiden. Bitte veranschaulichen Sie die Verteilungen der Anfangsgehälter für Männer und Frauen mit Hilfe von Box-Plots und interpretieren Sie das Ergebnis.

Ergebnis und Interpretation

Aufgabe 4 (Uhrzeit bei Beginn: _____)

Versuchen Sie die vorliegende Abbildung für die Variable „Studienfach“ möglichst genau zu reproduzieren. (Wichtig: Lösung in SPSS. Die Farbe spielt dabei keine Rolle!)

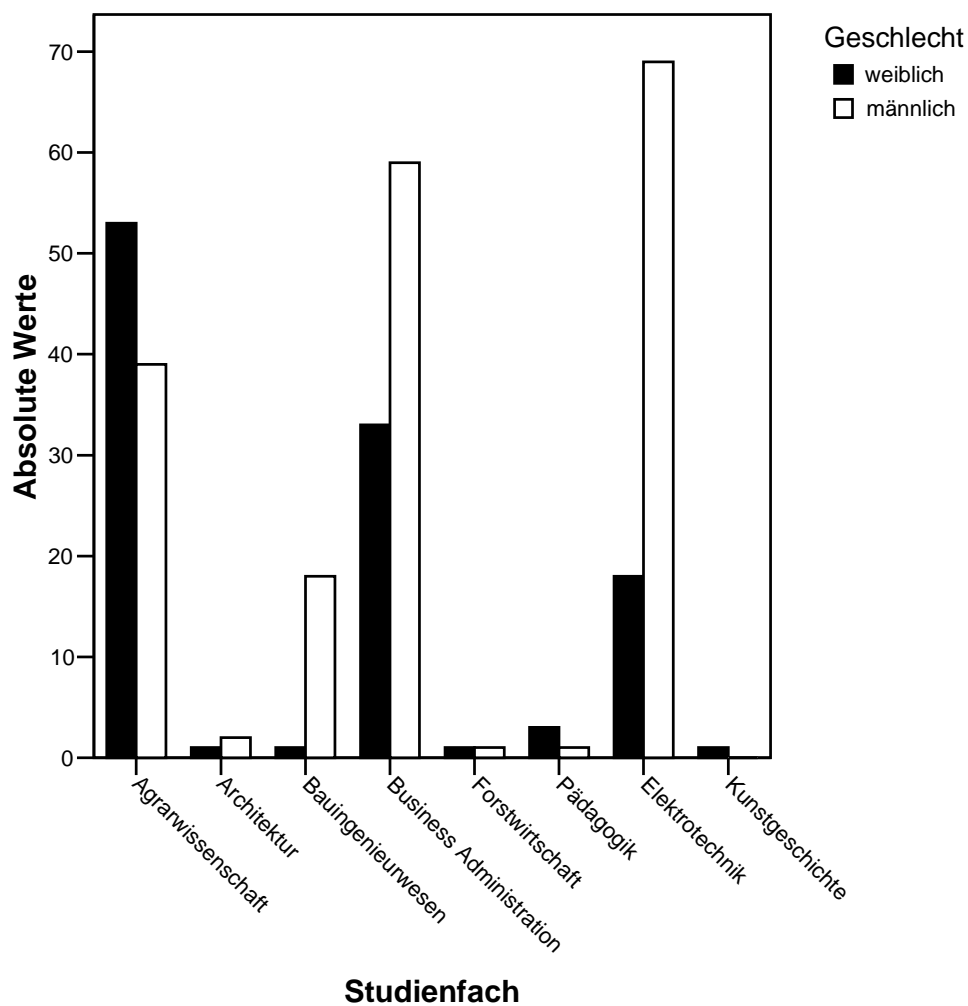


Abbildung zu Aufgabe 4

A 3.2 Ergebnisse statistischer Analysen

Im folgenden Kapitel (A 3.2.1) sind die Ergebnisse der Item- und Skalenanalysen für die in Studie 2 verwendeten Analysen separat in tabellarischer Form dargestellt.

A 3.2.1 Item- und Skalenanalysen

Tabelle A – 10

Ergebnis der Item- und Skalenanalyse des Fragebogens zu Statistikkenntnissen (Cronbachs $\alpha = .57$ für die 9 Items und $\alpha = .64$ für 6 Items ; $N = 33$)

Item-Nr.	Itemwortlaut (abgekürzt) ¹	<i>M</i>	<i>SD</i>	<i>P</i> ²	<i>r</i> _{it}	Schiefe	Exzess
1	<i>Der Median entspricht dem...</i> ³	4.00	0	100	0	-	-
2	Streuung und Standardabweichung einer Verteilung zeigen an...	2.91	1.1	73	.58	-0.46	-0.41
3	Empfindlich gegenüber Ausreißern unter den Beobachtungen ist...	3.00	0.9	75	.40	-1.23	1.52
4	Wenn von einem Ergebnis gesagt wird, es sei „statistisch signifikant“, dann bedeutet dies...	2.00	1.1	50	.22	0.34	0.63
5	Für den Produkt-Moment-Korrelationskoeffizienten gilt, ...	2.21	1.2	55	.37	-0.20	-0.71
6	<i>Welches Skalenniveau ist mindestens erforderlich, damit folgende Aussagen zutreffen?</i> (max. 6 Punkte)	1.33	1.1	33	.13	0.54	-0.30
7	<i>Zwei Zufallszahlen X und Y sind korreliert mit einer Kovarianz (...). Dann gilt...</i>	2.12	1.2	53	.13	-0.25	0.08
8	Bei einer empirischen Häufigkeitsverteilung gelten folgende Annahmen...	2.55	1.2	65	.25	-0.44	-0.74
9	Annahmen zu Korrelationskoeffizienten...	2.30	0.8	58	.27	0.01	1.15

Anmerkungen. *M* = Itemmittelwert; *SD* = Standardabweichung des Items; *P* = Schwierigkeit in %; *r*_{it} = Trennschärfe. ¹ Die aus der Skala eliminierten Items (nach Item- und Skalenanalyse) sind kursiv gesetzt. ² In dem Multiple-Choice-Test mit mindestens einer richtigen Antwort bei insgesamt 4 Antwortalternativen konnten zwischen 0 und 4 Punkten erreicht werden. Eine Ausnahme bildet Item 6 mit maximal 6 Punkten. Punkte wurden sowohl für das Ankreuzen richtiger Antworten als auch für das Nicht-Ankreuzen falscher Antworten vergeben. Auf diesem Weg sollte eine Ratekorrektur erfolgen. Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³ Für Item 1 konnten Schiefe und Exzess nicht berechnet werden, da die Varianz in dem Item Null betrug.

Tabelle A – 11

Ergebnis der Item- und Skalenanalyse des Fragebogens zur aktuellen Motivation (Eingangsmotivation) sowie Cronbachs α getrennt für die vier Subskalen (I = Interesse, E = Erfolgswahrscheinlichkeit, H = Herausforderung, M = Misserfolgsbefürchtung; N = 33). Die Invertierung eines Items ist durch ein Minus in Klammern gekennzeichnet

Item -Nr.	Skala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Skala Interesse ($\alpha = .66$)							
1	Ich mag solche Rätsel und Knobeleyen. (I)	4.73	1.2	78	.39	-0.60	-0.58
4	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen. (I)	4.73	1.4	78	.49	-0.67	-0.58
7	Nach dem Lesen der Instruktion erscheint mir die Aufgabe sehr interessant. (I)	4.75	1.4	78	.41	-0.67	0.58
11	Bei Aufgaben wie dieser brauche ich keine Belohnung, sie machen mir auch so viel Spaß. (I)	4.36	1.5	73	.45	-0.43	-0.38
17	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten. (I)	3.27	1.7	55	.36	0.14	-0.83
Skala Erfolgswahrscheinlichkeit ($\alpha = .66$)							
2	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein. (E)	4.30	1.1	72	.46	0.04	0.17
3	Wahrscheinlich werde ich die Aufgabe nicht schaffen. (E-)	4.61	1.2	77	.50	-0.15	-0.77
13	Ich glaube, das kann jeder schaffen. (E)	3.24	1.6	53	.28	0.19	-0.96
14	Ich glaube, ich schaffe diese Aufgabe nicht. (E-)	4.85	1.4	82	.61	-0.28	-0.67
Skala Herausforderung ($\alpha = .69$)							
6	Die Aufgabe ist eine richtige Herausforderung für mich. (H)	4.55	1.5	77	.47	-0.39	-0.10
8	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde. (H)	5.30	1.4	88	.56	-0.80	0.23
10	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen. (H)	5.52	1.1	92	.58	-0.58	0.03
15	Wenn ich die Aufgabe schaffe, werde ich schon ein wenig stolz auf meine Tüchtigkeit sein. (H)	4.82	1.4	80	.31	-0.61	-0.20
Skala Misserfolgsbefürchtung ($\alpha = .79$)							
5	Ich fühle mich unter Druck, bei der Aufgabe gut abschneiden zu müssen. (M)	3.64	1.8	60	.61	-0.31	-1.40
9	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte. (M)	3.64	1.8	60	.73	-0.02	-0.94
12	Es ist mir etwas peinlich, hier zu versagen. (M)	3.36	1.7	57	.67	-0.12	-0.99
16	Wenn ich an die Aufgabe denke, bin ich etwas beunruhigt. (M)	3.64	1.3	60	.50	0.01	-0.85
18	Die konkreten Leistungsanforderungen hier lähmen mich. (M)	2.91	1.1	48	.33	-0.11	-0.54

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³ Getrennt für die vier Subskalen.

Tabelle A – 12

Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den ersten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit (N = 33)

Item-Nr.	Skala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Glatter Verlauf ($\alpha = .85$)							
2	Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	3.64	1.5	61	.80	0.04	-1.16
4	Ich habe keine Mühe, mich zu konzentrieren.	4.45	1.5	74	.39	-0.11	-0.97
5	Mein Kopf ist völlig klar.	4.27	1.5	71	.55	-0.10	-1.00
7	Die richtigen Gedanken/Bewegungen kommen wie von selbst.	3.15	1.5	53	.67	0.15	-0.92
8	Ich weiß bei jedem Schritt, was ich zu tun habe.	2.36	1.3	39	.66	1.26	0.91
9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	2.88	1.7	48	.75	0.72	-0.71
Absorbiertheit ($\alpha = .59$)							
1	Ich fühle mich optimal beansprucht.	4.52	1.7	75	.33	-0.52	-0.58
3	Ich merke gar nicht, wie die Zeit vergeht.	4.27	1.9	71	.43	0.01	-1.35
6	Ich bin ganz vertieft in das was ich gerade mache.	5.12	1.1	85	.35	-0.52	1.03
10	Ich bin völlig selbstvergessen.	3.35	1.5	56	.41	0.09	-0.78

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ²Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³Getrennt für die beiden Subskalen.

Tabelle A – 13

Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den zweiten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit ($N = 28$)

Item-Nr.	Skala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Glatter Verlauf ($\alpha = .83$)							
2	Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	3.96	1.8	66	.81	-0.03	-1.07
4	Ich habe keine Mühe, mich zu konzentrieren.	4.82	1.6	80	.52	-0.52	-0.26
5	Mein Kopf ist völlig klar.	4.46	1.5	74	.61	-0.10	-1.05
7	Die richtigen Gedanken/Bewegungen kommen wie von selbst.	3.36	1.6	56	.68	0.63	-0.55
8	Ich weiß bei jedem Schritt, was ich zu tun habe.	2.68	1.4	45	.60	0.62	-0.23
9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	3.18	1.6	53	.43	0.83	0.06
Absorbiertheit ($\alpha = .72$)							
1	Ich fühle mich optimal beansprucht.	4.64	1.7	77	.39	-0.19	-1.37
3	Ich merke gar nicht, wie die Zeit vergeht.	4.93	1.7	82	.64	-0.57	-0.92
6	Ich bin ganz vertieft in das was ich gerade mache.	5.43	1.1	91	.63	-0.14	-0.78
10	Ich bin völlig selbstvergessen.	3.44	1.6	57	.47	0.32	-0.69

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³ Getrennt für die beiden Subskalen.

Tabelle A – 14

Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den dritten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit ($N = 21$)

Item-Nr.	Skala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Glatter Verlauf ($\alpha = .89$)							
2	Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	3.95	1.7	66	.86	-0.05	-0.82
4	Ich habe keine Mühe, mich zu konzentrieren.	4.24	1.7	71	.61	0.17	-1.41
5	Mein Kopf ist völlig klar.	4.10	1.8	68	.78	0.23	-1.01
7	Die richtigen Gedanken/Bewegungen kommen wie von selbst.	2.81	1.4	47	.73	0.51	-0.01
8	Ich weiß bei jedem Schritt, was ich zu tun habe.	2.71	1.5	45	.67	0.35	-0.75
9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	2.86	1.5	48	.64	0.27	-0.78
Absorbiertheit ($\alpha = .63$)							
1	Ich fühle mich optimal beansprucht.	4.67	1.7	78	.21	-0.11	-1.02
3	Ich merke gar nicht, wie die Zeit vergeht.	4.86	1.9	81	.50	-0.54	-0.92
6	Ich bin ganz vertieft in das was ich gerade mache.	5.10	1.2	85	.61	-0.20	-0.38
10	Ich bin völlig selbstvergessen.	3.46	1.8	58	.42	-0.06	-1.33

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³ Getrennt für die beiden Subskalen.

Tabelle A – 15

Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum ersten Messzeitpunkt (N = 33)¹. Die Invertierung eines Items ist durch ein Minus in Klammern gekennzeichnet

Item-Nr.	Skala und Wortlaut der zugehörigen Items	M^2	SD	P^3	Schiefe	Exzess
Interesse						
13	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen. (I)	4.73	1.6	79	-0.72	-0.02
18	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten. (I)	3.33	1.9	56	0.40	-1.01
Erfolgswahrscheinlichkeit						
11	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein. (E)	3.64	1.7	61	-0.04	-1.25
12	Wahrscheinlich werde ich die Aufgabe nicht schaffen. (E-)	4.45	1.8	74	-0.27	-0.93
Herausforderung						
14	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde. (H)	4.67	1.8	78	-0.81	-0.51
16	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen. (H)	5.39	1.0	90	-1.29	2.91
Misserfolgsbefürchtung						
15	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte. (M)	4.55	1.9	76	-0.09	-1.26
17	Es ist mir etwas peinlich, hier zu versagen. (M)	4.67	1.8	78	-0.25	-1.13

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %. ¹ Cronbachs α wurde nicht berechnet, da dieser Koeffizient für zwei Items pro Skala nicht definiert ist (vgl. McDonald, 1999; Hancock & Mueller, 2006). ² 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ³ Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 16

Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum zweiten Messzeitpunkt (N = 28)¹. Die Invertierung eines Items ist durch ein Minus in Klammern gekennzeichnet

Item -Nr.	Skala und Wortlaut der zugehörigen Items	M^2	SD	P^3	Schiefe	Exzess
Interesse						
13	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen. (I)	4.39	1.8	73	-0.52	-0.81
18	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten. (I)	3.25	1.8	54	0.41	-0.62
Erfolgswahrscheinlichkeit						
11	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein. (E)	3.18	1.6	53	0.82	-0.16
12	Wahrscheinlich werde ich die Aufgabe nicht schaffen. (E-)	4.30	1.9	72	-0.23	-1.20
Herausforderung						
14	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde. (H)	4.68	1.7	78	-0.86	-0.27
16	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen. (H)	5.54	1.1	92	-1.07	2.60
Misserfolgsbefürchtung						
15	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte. (M)	4.89	1.7	82	-0.30	-1.22
17	Es ist mir etwas peinlich, hier zu versagen. (M)	4.96	1.7	83	-0.41	-1.01

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; ¹ Cronbachs α wurde nicht berechnet, da dieser Koeffizient für zwei Items pro Skala nicht definiert ist (vgl. McDonald, 1999; Hancock & Mueller, 2006). ² 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ³ Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 17

Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum dritten Messzeitpunkt ($N = 21$)¹. Die Invertierung eines Items ist durch ein Minus in Klammern gekennzeichnet

Item-Nr.	Skala und Wortlaut der zugehörigen Items	M^2	SD	P^3	Schiefe	Exzess
Interesse						
13	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen. (I)	4.57	1.7	76	-0.46	-0.17
18	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten. (I)	3.05	1.8	51	0.52	-0.65
Erfolgswahrscheinlichkeit						
11	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein. (E)	3.14	1.6	52	0.76	0.26
12	Wahrscheinlich werde ich die Aufgabe nicht schaffen. (E-)	3.90	2.0	65	-0.25	-1.14
Herausforderung						
14	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde. (H)	4.48	1.6	75	-1.15	0.59
16	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen. (H)	4.95	1.6	83	-0.98	0.74
Misserfolgsbefürchtung						
15	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte. (M)	4.95	1.7	83	-0.19	-1.07
17	Es ist mir etwas peinlich, hier zu versagen. (M)	4.95	1.8	83	-0.68	-0.45

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %. ¹ Cronbachs α wurde nicht berechnet, da dieser Koeffizient für zwei Items pro Skala nicht definiert ist (vgl. McDonald, 1999; Hancock & Mueller, 2006). ² 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ³ Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 18

Ergebnis der Item- und Skalenanalyse für die Leistung bei der Bearbeitung der vier Aufgaben des Arbeitsblattes (Variable: „Leistung_Arbeitsblatt; $N = 33$; Cronbachs $\alpha = .65$)

Item-Nr.	Aufgabe mit Itemwortlaut (abgekürzt) ¹	Max.	M	SD	P^2	r_{it}	Schiefe	Exzess
1	Angabe von Mittelwert und Standardabweichung für die Gesamtstichprobe, sowie getrennt für Frauen und Männer ($N = 33$)	6	4.45	2.3	74	.53	-1.13	-0.37
2	Angabe des Korrelationskoeffizienten, des Signifikanzniveaus, sowie Interpretation des Ergebnisses ($n = 27$)	4	1.48	1.3	45	.44	0.15	-1.21
3	Interpretieren eines Boxplots (Median Frauen und Männer, Ausreißer und Extremwerte, Verteilung; $n = 22$)	4	1.03	1.0	39	.41	0.32	-1.33
4	Erstellen eines gruppierten Balkendiagramms ($n = 13$) 1 Punkt: zu den Daten passende Grafik 2 Punkte: Balkendiagramm 3 Punkte: Gruppiertes Balkendiagramm	3	1.42	1.5	100	.51	0.06	-2.01

Anmerkungen. Max. = Maximale Punktzahl, M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ n gibt die Anzahl der Personen wider, die die Aufgabe bearbeitet haben. ² Mit Inangriffnahme-Korrektur; d.h. einbezogen wurden nur die Werte von Teilnehmern, welche die Aufgabe bearbeitet haben. Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 19

Ergebnis der Item- und Skalenanalyse für die Leistung beim Lösen der vier Aufgaben in SPSS (Variable: „Leistung_SPSS“; $N = 33$; Cronbachs $\alpha = .77$)

Item-Nr.	Aufgabe mit Itemwortlaut (abgekürzt) ¹	Max.	M	SD	P^2	r_{it}	Schiefe	Exzess
1	Produktion des Ergebnisses (M , SD) für die Gesamtstichprobe, sowie getrennt für Frauen und für Männer ($n = 33$)	3	2.15	1.3	72	0.49	-0.91	-1.03
2	Produktion einer korrekten Korrelations-tabelle in SPSS ($n = 27$)	1	0.70	0.5	85	0.58	-0.90	-1.27
3	Erstellen eines Boxplots in SPSS ($n = 22$) 0 Punkte: keine Grafik 1 Punkt: zu den Daten passende Grafik 2 Punkte: korrekter Boxplot	2	1.10	0.9	84	0.82	-0.19	-1.93
4	Erstellen eines gruppierten Balkendiagramms ($n = 13$) 0 Punkte: keine Grafik 1 Punkt: zu den Daten passende Grafik 2 Punkte: Balkendiagramm 3 Punkte: Gruppiertes Balkendiagramm	3	1.42	1.4	100	0.66	0.06	-2.01

Anmerkungen. Max. = Maximale Punktzahl, M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ n gibt die Anzahl der Personen wider, die die Aufgabe bearbeitet haben. ² Mit Inangriffnahme-Korrektur; d.h. einbezogen wurden nur die Werte von Teilnehmern, welche die Aufgabe bearbeitet haben. Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

A 3.2.2 Voraussetzungsprüfungen und weitere statistische Analyseergebnisse

Tabelle A – 20 fasst die Ergebnisse der statistischen Voraussetzungsprüfungen für die zentralen Untersuchungsvariablen in Studie 2 zusammen.

Tabelle A – 20

Deskriptive Statistiken für die Verteilung der Skalensummenwerte inklusive der Ergebnisse des Kolmogorov-Smirnov-Tests auf Normalverteilung (N = 33)

Skala	M	SD	Schiefe	Exzess	Z	Asymp.
Häufigkeit Computernutzung ¹	1.85	1.3	1.07	0.34	1.63	.01
Statistikselbstbeurteilung (1 Item)	2.88	0.8	-0.49	0.46	1.64	.01
Statistikvorwissenstest (6 Items) ¹	2.49	0.6	-0.36	0.23	0.57	.90
FAM (Eingangsmotivation) ¹						
Interesse	4.36	1.0	-0.01	-0.52	0.49	.97
Erfolgswahrscheinlichkeit	4.25	1.0	-0.27	0.07	0.51	.96
Herausforderung	5.05	1.0	-0.39	-0.88	0.78	.58
Misserfolgsbefürchtung	3.44	1.1	-0.67	-1.08	1.18	.13
Aktuelle Motivation ^{2,3}						
Netto-Motivation t0 (8 Items)	10.36	2.5	0.42	-0.75	0.87	.44
Netto-Motivation t1 (8 Items)	9.41	3.9	-0.37	0.18	0.43	.99
Netto-Motivation t2 (8 Items)	9.32	3.8	0.02	-0.62	0.52	.95
Netto-Motivation t3 (8 Items)	8.90	4.0	-0.50	0.66	0.59	.88
Flow ^{1,3}						
Messzeitpunkt t1 (10 Items)	3.89	0.8	0.11	-0.25	0.64	.80
Messzeitpunkt t2 (10 Items)	4.18	1.0	0.01	-0.47	0.52	.95
Messzeitpunkt t3 (10 Items)	3.98	1.1	-0.06	0.16	0.41	.97
Flow_end (10 Items)	4.32	1.0	-0.47	0.05	0.48	.98
Explorationsstrategie ⁴						
Systematische Exploration	41	12	-0.02	-0.24	0.38	.99
Versuch-und-Irrtum	5	5	0.69	-0.50	0.87	.43
Rigide Exploration	20	17	0.67	-0.66	0.71	.70
Informationssuche	5	7	1.25	0.83	1.48	.70
Leistung						
Arbeitsblatt	8.41	4.4	-0.43	-1.27	0.83	.49
SPSS	5.36	3.4	-0.64	-0.55	1.02	.25

Anmerkungen. Z = Prüfgröße Kolmogorov – Smirnov-Tests; Asymp. = Asymptotische Signifikanz. ¹ Es wird der Skalensummenwert auf Itemebene berichtet. ² Die Netto-Motivation wurde auf der Grundlage von jeweils zwei Items der FAM-Skalen Herausforderung, Interesse, Erfolgswahrscheinlichkeit und Misserfolgsbefürchtung gebildet. Aus den beiden Items pro Skala wurde ein Mittelwert gebildet. Die Netto-Motivation wurde berechnet, in dem die Mittelwerte der Skalen Herausforderung, Erfolgswahrscheinlichkeit und Interesse addiert wurden und der Mittelwert der beiden Items zur Misserfolgsbefürchtung subtrahiert wurde. ³ Der Stichprobenumfang variierte über die Messzeitpunkte: N = 33 (t0 und t1), n = 28 (t2), n = 21 (t3); vgl. Kapitel 5.3.1. ⁴ Dargestellt ist jeweils die mittlere, relative Häufigkeit (%) der Explorationsstrategie über den gesamten Messzeitraum von 30 Minuten.

A.3.2.3 Ergebnisse ergänzender statischer Analysen

In Tabelle A – 21 sind die Interkorrelationen der zentralen Untersuchungsvariablen dargestellt (vgl. Kapitel 8.2.1).

Tabelle A – 21

Interkorrelationen der Variablen in der Gesamtstichprobe ($N = 33$)

Variablen	2	3	4	5	6	7	8	9	10	11	12	13	14
Vorwissen/Vorerfahrung													
1 Computernutzung (Häufigkeit)	.12	.35*	.23	.01	-.30	.19	.32	.17	-.12	-.16	-.10	.19	.20
2 Statistiktest		.15	.22	.16	.21	.26	.04	.21	.14	-.03	-.11	.16	.24
Eingangsmotivation													
3 Interesse			.32	.37*	.10	.10	.20	.03	-.13	.01	.26	-.12	-.04
4 Erfolgswahrscheinlichkeit				-.09	-.26	.12	.19	.27	.27	-.14	.01	-.05	.09
5 Herausforderung					.30	.37*	.21	.28	-.16	-.18	.01	.19	.24
6 Misserfolgsbefürchtung						.07	-.44*	-.05	-.04	.16	.12	-.21	-.16
Motivationaler Zustand und Flow													
7 Flow-Erleben (letzter MZP ¹ vor Bearbeitungsende)							.57**	.57**	-.21	-.61**	-.12	.57**	.64**
8 Motivationaler Zustand (letzter MZP vor Bearbeitungsende)								.44*	-.04	-.63**	-.24	.59**	.66**
Explorationsstrategien													
9 Systematisch									-.19	-.79**	-.45**	.69**	.68**
10 Versuch-und-Irrtum										.17	-.16	-.13	-.03
11 Rigide											.23*	-.84**	-.74**
12 Informationssuche												-.50**	-.55**
Leistung													
13 Leistung_SPSS													.88**
14 Leistung_Arbeitsblatt													

Anmerkungen. * $p < .05$, ** $p < .01$; ¹MZP = Messzeitpunkt.

A 4. Anhang zu Studie 3

A 4.1 Untersuchungsmaterialien

Abschnitt A 4.1 beinhaltet die im Rahmen von Studie eingesetzten Untersuchungsmaterialien (Vorwissenstest/Arbeitsblatt). Der darauffolgende Abschnitt 4.2 präsentiert Ergebnisse statistischer Analysen, die zusätzlich zu den Hypothesenprüfungen durchgeführt wurden.

A 4.1.1 Fragebogen zu Computervorwissen und Statistikkenntnissen

Angaben zur Person

Studienfach: Wirtschaftswissenschaften Wirtschaftspädagogik

Angestrebter Abschluss: Bachelor Master Diplom

Fachsemester: _____ Alter: _____ Jahre

Geschlecht: weiblich männlich

Statistikvorkenntnisse

1. Welche Statistik-Veranstaltungen haben Sie in der Vergangenheit besucht bzw. besuchen Sie gerade?

- Statistik 1 – Vorlesung (für Diplom bzw. Bachelor/Master) Statistik 1 - Tutorium Statistik/Stochastik im Schulunterricht
- Statistik 2 - Veranstaltung Statistik 2 - Tutorium Sonstige: _____

2. Bitte schätzen Sie auf der folgenden Skala Ihre theoretischen Kenntnisse in Statistik ein.

- sehr gut mangelhaft

3. Bitte kreuzen Sie an, auf welche Inhalte Sie Ihr statistisches Wissen anwenden können:

- Deskriptive Statistik (z.B. Bestimmung der Lage- und Streuungsmaße)
- Inferenzstatistische Überprüfung von Hypothesen (z.B. Signifikanztests, t-Test)
- grafische Darstellung der statistischen Auswertungen (z.B. Histogramm anfertigen)

Bitte kreuzen Sie bei den folgenden Fragen zum Thema Statistik die richtige(n) Lösung(en) an. Beachten Sie, dass pro Frage mehrere Aussagen richtig sein können (mindestens eine).

1. Varianz und Standardabweichung einer Verteilung zeigen an, ...

- wie stark die Werte der Verteilung um den Modalwert streuen.
- wie stark die Werte der Verteilung um den Median streuen.
- wie stark die Werte der Verteilung um den arithmetischen Mittelwert streuen.
- keine der drei vorangegangenen Antwortalternativen ist richtig.

2. Empfindlich gegenüber Ausreißern unter den Beobachtungen ist....

- der Modalwert (bzw. Modus).
- das Arithmetische Mittel.
- der Median.
- die Streuung.

3. Wenn von einem Ergebnis gesagt wird, es sei „statistisch signifikant“, dann bedeutet dies, dass...

- es praktisch relevant ist.
- seine Zufallswahrscheinlichkeit höchstens 5% beträgt.
- es wahrscheinlich richtig ist.
- es für 5% aller Personen zutrifft.

4. Für den Produkt-Moment-Korrelationskoeffizienten r_{xy} gilt:

- $r_{xy} = 1$; perfekter positiver linearer Zusammenhang
- $r_{xy} = 0$; es besteht weder ein linearer noch ein kurvilinear (nicht linearer) Zusammenhang zwischen den Variablen.
- $r_{xy} = -1$; im Streudiagramm liegen alle Punkte exakt auf einer fallenden Geraden
- $-1 \leq r_{xy} \leq 1$: r nimmt ausschließlich Werte zwischen 1.0 und -1.0 an.

5. Für die Darstellung der Daten einer Verteilung mittels „Boxplots“ gelten üblicherweise bestimmte Regeln. Bitte kreuzen Sie die richtige(n) Aussage(n) an:

- Das Arithmetische Mittel in der Box vermittelt durch seine Lage innerhalb der Box einen Eindruck von der Schiefe der den Daten zugrunde liegenden Verteilung.
- Die Bereiche oberhalb und unterhalb der Box umfassen zusammen 50 % der Daten.
- Ausreißerwerte werden im Boxplot extra gekennzeichnet.
- Ein Boxplot ist eine schematische grafische Darstellung der Quartile, des Medians, sowie des Minimums und des Maximums der Messwerte eines metrischen Merkmals

6. Bei einem Histogramm....

- werden die Messwerte einer Häufigkeitsverteilung grafisch dargestellt.
- werden auf der Abszisse (X-Achse) die absoluten, relativen oder prozentualen Häufigkeiten einer Verteilung abgetragen.
- wird stets die Häufigkeit einzelner Variablen oder Punktwerte abgetragen.
- ist die Balkenlänge/-höhe proportional zur Anzahl der Beobachtungen in einer Klasse von Messwerten.

Im folgenden Teil des Fragebogens geht es um Ihre Vertrautheit mit verschiedenen Computeranwendungen. Sie sollen sich selbst daraufhin einschätzen, wie vertraut Sie im Umgang mit einzelnen Computeranwendungen sind. Bitte beurteilen Sie für jede der aufgelisteten Computeranwendungen jeweils, wie vertraut Sie im Vergleich zu anderen Studierenden mit der jeweiligen Anwendung sind.

	Weit überdurch- schnittlich ++	Überdurch- schnittlich +	Durch- schnittlich 0	Unter- durch- schnittlich -	Weit unterdurch- schnittlich --
1. Textverarbeitungsprogramme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Multimedia-Anwendungen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Programmiersprachen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Tabellenkalkulation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Statistik-Programme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. E-Mail	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Datenbanken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Internet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Computerspiele	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Grafikprogramme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Jetzt geht es um Ihr praktisches Computerwissen, d.h. um solches Wissen, das für den Umgang mit dem Computer unmittelbar relevant sein kann. Im Folgenden finden Sie Problem-beschreibungen, mit denen man bei der täglichen Arbeit mit dem Computer konfrontiert sein oder zu tun haben kann.

Ihre Aufgabe ist es, für jede geschilderte Problemsituation die eine Handlungsalternative anzukreuzen, die Ihrer Einschätzung nach die beste Möglichkeit darstellt, mit dem Problem umzugehen.

Da die Fragen sich zum Teil auf bestimmte Betriebssysteme beziehen, geben Sie bitte an, von welchem Betriebssystem Sie Kenntnisse besitzen (z.B. weil Sie damit arbeiten/damit gearbeitet haben).

- Ich arbeite mit...
(Mehrfachnennungen möglich)
- Windows '95 oder höher (98, NT, ME, XP, Vista...)
 Windows 3.x
 OS/2
 DOS
 Unix
 Apple Macintosh
- Ich habe Kenntnisse in...
(Mehrfachnennungen möglich)
- Windows '95 oder höher (98, NT, ME, XP, Vista...)
 Windows 3.x
 OS/2
 DOS
 Unix
 Apple Macintosh

8. Sie wissen, dass ein bestimmtes Programm auf Ihrem Computer installiert ist, Sie können es aber nicht auf die gewohnte Weise starten. Wie können Sie dieses Problem beheben?

- (e) Ich installiere das Programm neu.
- (f) Ich suche das Programmverzeichnis im Explorer/Dateimanager und rufe die *.exe-Datei des Programms mit einem Doppelklick auf.
- (g) Programme, die nicht "normal" aufrufbar sind, sind in Verbindung mit dem benutzten Betriebssystem nicht lauffähig. Ich besorge eine neue Version des Programms für mein Betriebssystem.
- (h) Ich füge unter "Systemsteuerung" das Programm ins "Programmverzeichnis" ein.
- weiß ich nicht*

9. Sie bekommen ein als ZIP-Archiv gepacktes Textdokument. Wie verfahren Sie damit?

- (e) Ich öffne das Dokument wie gewohnt in meinem Textverarbeitungsprogramm.
- (f) Ich entpacke das Programm über den Explorer/Dateimanager.
- (g) Ich ziehe die Datei mit der Maus auf den Desktop, dabei wird sie automatisch entpackt.
- (h) Ich ändere die Erweiterung der Datei von *.zip in die Word-Erweiterung *.doc. Danach lässt sie sich in Word öffnen.
- weiß ich nicht*
-

10. Sie haben über das Internet eine als selbstextrahierendes Archiv gepackte Textdatei auf Ihren Computer geladen. Nun möchten Sie diese lesen. Was tun Sie?

- (e) Die Datei entpackt sich nach dem Laden selbst auf der Festplatte. Danach kann ich sie in der Textverarbeitung wie gewohnt aufrufen und lesen.
- (f) Ich öffne die Datei in meinem Textverarbeitungsprogramm über den Menüpunkt "Extras – Archiv entpacken".
- (g) Ich öffne das Dokument wie gewohnt in meinem Textverarbeitungsprogramm.
- (h) Ich entpacke die Datei per Doppelklick im Explorer/Dateimanager und öffne sie dann im Textverarbeitungsprogramm.

weiß ich nicht

11. Sie wollen den Computer starten. Es erscheint kurz nach dem Einschalten die Meldung "Disk not ready. Insert bootable disk and press any key" bzw. "This disk can't boot. It was formatted without the /s (system-) option". Was tun Sie?

- (e) Offensichtlich versagt die Festplatte. Ich muss den Computer zum Händler bringen.
- (f) Anscheinend sind die Einstellungen im Setup fehlerhaft. Ich rufe das System-Setup auf und ändere die Einstellungen für die Festplatte.
- (g) Ich habe eine Diskette im Diskettenlaufwerk vergessen. Ich entferne die Diskette aus dem Laufwerk und drücke die "Enter"-Taste.
- (h) Ich führe einen Warmstart des Computers durch.

weiß ich nicht

12. Ihr Computer ist abgestürzt und Sie wollen ihn möglichst "schonend" neu starten. Was tun Sie?

- (e) Ich drücke den "reset"-Knopf.
- (f) Ich drücke die Tastenkombination "Strg"+"Alt"+"Entf".
- (g) Ich drücke die Tastenkombination "Ende"+"Enter".
- (h) Ich schalte den Computer aus und wieder ein.

weiß ich nicht

A 4.1.2 Aufgabenblatt

Arbeitsblatt

WICHTIG:

- Bitte bearbeiten Sie alle Aufgaben mit SPSS; alle Aufgaben sind mit SPSS lösbar.
- Bearbeiten Sie die 6 Aufgaben bitte der Reihe nach.
- Tragen Sie vor dem Beginn einer Aufgabe die Uhrzeit auf dem Arbeitsblatt ein.
- Schließen Sie das Ausgabefenster nicht, sondern minimieren sie es.
- Bitte positionieren Sie den Mauszeiger an der Uhr während Sie Lösungen auf dem Blatt eintragen oder Fragebögen ausfüllen.

Aufgabe 1 (Uhrzeit bei Beginn: _____)

Bestimmen Sie den Mittelwert (\bar{X}) und die Standardabweichung (s; mittlere Abweichung vom Mittelwert einer Verteilung) für die Variable „Anfangsgehalt“ in der gesamten Stichprobe und getrennt für die weiblichen und für die männlichen Personen.

Tragen Sie die Ergebnisse bitte auf dem Arbeitsblatt ein.

	\bar{X}	s
<u>Gesamtstichprobe</u> Mittelwert und Streuung:		
<u>Weibliche Personen</u> _____ Mit- telwert und Streuung:		
<u>Männliche Personen</u> _____ Mit- telwert und Streuung:		

Aufgabe 2 (Uhrzeit bei Beginn: _____)

Ein Wirtschaftsinstitut erforscht die Entwicklung der Gehälter vom Anfangsgehalt beim Einstieg in den Beruf bis zu dem Zeitpunkt, zu dem die Person 10 Jahre im Unternehmen gearbeitet hat. **Frage:** Gibt es einen Zusammenhang zwischen dem „Anfangsgehalt“ und dem „Gehalt, das eine Person nach 10 Jahren Betriebszugehörigkeit“ verdient?

Bitte berechnen Sie die Korrelation zwischen den beiden Variablen, geben Sie die relevanten Kennwerte an und interpretieren Sie das Ergebnis.

Ergebnis und Interpretation

Aufgabe 3 (Uhrzeit bei Beginn: _____)

Erstellen Sie nun ein **Streudiagramm** für die beiden Variablen „Anfangsgehalt“ und „Gehalt nach 10-jähriger Betriebszugehörigkeit“. (Wichtig: Lösung in SPSS)

Aufgabe 4 (Uhrzeit bei Beginn: _____)

Das Statistische Bundesamt möchte ermitteln, ob sich Männer und Frauen in der Variablen „Anfangsgehalt“ unterscheiden. Bitte veranschaulichen Sie die Verteilungen der Anfangsgehälter für Männer und Frauen mit Hilfe von **Box-Plots** und interpretieren Sie das Ergebnis.

Ergebnis und Interpretation

Aufgabe 5 (Uhrzeit bei Beginn: _____)

Versuchen Sie die vorliegende Abbildung für die Variable „Studienfach“ möglichst genau zu reproduzieren. (Wichtig: Lösung in SPSS. Die Farbe spielt dabei keine Rolle!)

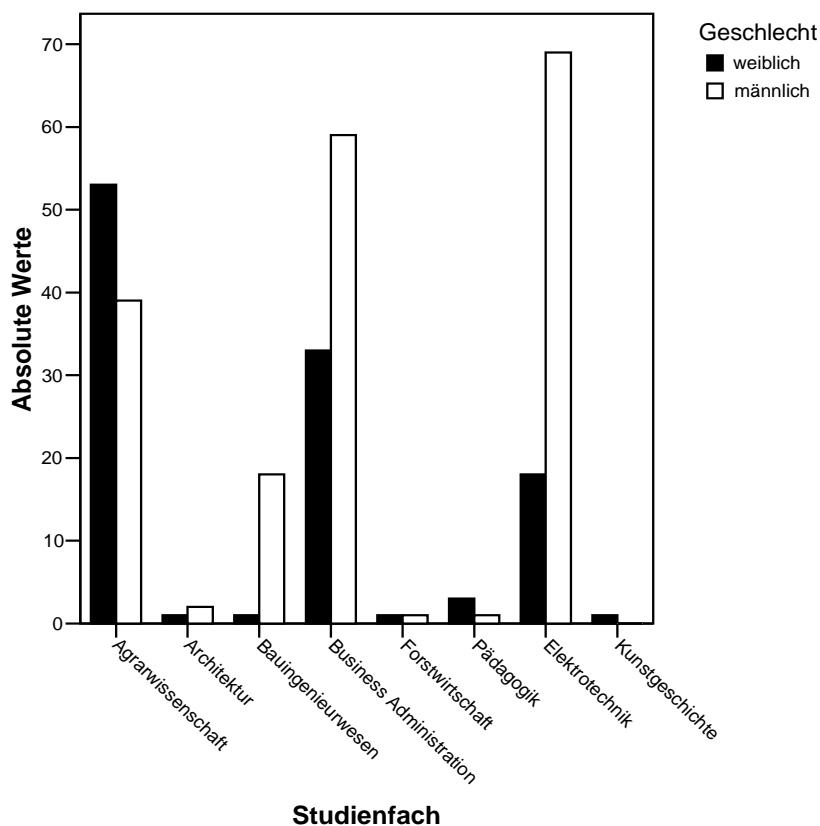


Abbildung zu Aufgabe 5

Aufgabe 6 (Uhrzeit bei Beginn: _____)

Die Personalleitung der Firma hat die Gedächtnisfähigkeiten aller Mitarbeiter und Mitarbeiterinnen getestet.

- a) Erstellen Sie eine **Häufigkeitstabelle** für die Variable „Gedächtnisfähigkeit“.
(Wichtig: Lösung in SPSS)

- b) Die Personalleitung bietet ausgewählten Mitarbeitern und Mitarbeiterinnen ein Gedächtnistraining an. Alle Mitarbeiter und Mitarbeiterinnen mit einem Wert unter 5 Punkten werden zu dem Training eingeladen.

Wie viele Mitarbeiter und Mitarbeiterinnen werden zu dem Gedächtnistraining eingeladen? Entnehmen Sie diese Information der Häufigkeitstabelle und schreiben Sie die entsprechenden Zahlen in die Lücken:

Anzahl von Mitarbeitern/-innen
mit weniger als 5 Punkten:

Prozentsatz von Mitarbeiter/innen
mit weniger als 5 Punkten:

Aufgabe 7 (Uhrzeit bei Beginn: _____)

Im Datensatz finden Sie die Werte aller weiblichen und männlichen Mitarbeiter/-innen der Firma in den entsprechenden Variablen. Sie wollen sich nun ausschließlich die Werte der männlichen Personen ansehen. Suchen Sie in SPSS die Funktion, mit der Sie **Fälle auswählen** können. Wählen Sie dann **nur die männlichen Personen** für Ihre weiteren Berechnungen aus.

Wenn Sie die männlichen Personen ausgewählt haben, berechnen Sie erneut die **Korrelation** für die Variablen „Anfangsgehalt“ und „Gehalt nach 10-jähriger Betriebszugehörigkeit“. **Tragen Sie das Ergebnis hier ein und interpretieren Sie es:**

Ergebnis und Interpretation

A 4.2 Ergebnisse statistischer Analysen

Kapitel A 4.2.1 gibt in mehreren Tabellen eine Übersicht über die Ergebnisse der Item- und Skalenanalysen der zentralen Untersuchungsvariablen in Studie 3. Kapitel A 4.2.2 bietet detaillierte Ergebnisse der Voraussetzungsprüfungen. Die Resultate weiterer statistischer Analysen sind in Kapitel A 4.2.3 aufgeführt.

A 4.2.1 Item- und Skalenanalysen

Tabelle A – 22

Ergebnis der Item- und Skalenanalyse des Fragebogens zu Statistikkenntnissen mit Cronbachs $\alpha = .37$ (6 Items; $N = 92$) bzw. $.47$ (4 Items; ohne Item 1 und Item 3)

Item-Nr.	Itemwortlaut (abgekürzt)	<i>M</i>	<i>SD</i>	<i>P</i> ²	<i>r</i> _{it}	Schiefe	Exzess
1	<i>Varianz und Standardabweichung einer Verteilung zeigen an...</i>	3.45	1.0	86	.10	-2.03	3.65
2	Empfindlich gegenüber Ausreißern unter den Beobachtungen ist...	2.86	0.9	72	.22	-0.78	0.26
3	<i>Wenn von einem Ergebnis gesagt wird, es sei „statistisch signifikant“, dann bedeutet dies...</i>	2.38	1.0	60	.01	0.62	-0.76
4	Für den Produkt-Moment-Korrelationskoeffizienten gilt, ...	1.83	0.9	46	.31	-0.19	-0.73
5	Für die Darstellung der Daten mittels „Box plots“ gelten üblicherweise bestimmte Regeln...	2.48	1.0	62	.26	-0.23	-0.28
6	Bei einem Histogramm...	3.08	0.9	77	.17	-0.67	-0.23

Anmerkungen. *M* = Itemmittelwert; *SD* = Standardabweichung des Items; *P* = Schwierigkeit in %; *r*_{it} = Trennschärfe. ¹ Die aus der Skala eliminierten Items (nach Item- und Skalenanalyse) sind kursiv gesetzt. ² In dem Multiple-Choice-Test mit mindestens einer richtigen Antwort bei insgesamt 4 Antwortalternativen konnten zwischen 0 und 4 Punkten pro Aufgabe erreicht werden. Punkte wurden sowohl für das Ankreuzen richtigen Antworten als auch für das Freilassen falscher Antworten vergeben. Auf diesem Weg sollte eine Ratekorrektur erfolgen. Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 23

Ergebnis der Item- und Skalenanalyse der ausgewählten Items der Skala Praktisches Computerwissen PRACOWI mit Cronbachs $\alpha = .42$ (alle 5 Items; $N = 92$)¹

Item-Nr.	Itemwortlaut (abgekürzt) ²	M^3	SD	P^4	r_{it}
1	Sie wissen, dass ein bestimmtes Programm auf Ihrem Computer installiert ist...	0.41	0.7	41	.17
2	Sie bekommen ein als ZIP-Archiv gepacktes Textdokument...	0.48	0.5	48	.18
3	Sie haben über das Internet eine als selbstextrahierendes Archiv gepackte Textdatei auf Ihren Computer geladen...	0.22	0.4	22	.42
4	Sie wollen den Computer starten...	0.65	0.6	65	.16
5	Ihr Computer ist abgestürzt...	0.62	0.5	62	.23

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ Schiefe und Exzess wurden nicht berechnet, da es nur zwei Antwortkategorien gab („richtig“, „falsch“). ² Die aus der Skala eliminierten Items (nach Item- und Skalenanalyse) sind kursiv gesetzt. ³ In dem Multiple-Choice-Test mit genau einer richtigen Antwort bei insgesamt 4 Antwortalternativen konnte maximal 1 Punkt pro Aufgabe erreicht werden. ⁴ Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 24

Ergebnis der Item- und Skalenanalyse der ausgewählten Items der Skala Vertrautheit mit Computeranwendungen VECA mit Cronbachs $\alpha = .74$ (alle 10 Items; $N = 92$)

Item-Nr.	Itemwortlaut (abgekürzt) ¹	<i>M</i>	<i>SD</i>	<i>P</i> ²	<i>r</i> _{it}	Schiefe	Exzess
1	Textverarbeitungsprogramme	2.61	0.99	65	.62	-0.45	-0.28
2	Multimedia-Anwendungen	2.43	1.28	61	.50	-0.38	-0.94
3	Programmiersprachen	0.50	0.81	13	.48	1.55	1.58
4	Tabellenkalkulation	1.55	1.07	39	.44	0.21	-0.87
5	Statistik-Programme	0.32	0.59	8	.32	1.74	1.95
6	E-Mail	3.84	0.40	96	.29	-2.37	5.08
7	Datenbanken	1.28	1.19	32	.39	0.71	-0.28
8	Internet	3.87	0.37	97	.23	-2.88	8.19
9	Computerspiele	0.80	0.93	20	.27	1.16	1.01
10	Grafikprogramme	1.05	1.03	26	.53	0.81	0.08

Anmerkungen. *M* = Itemmittelwert; *SD* = Standardabweichung des Items; *P* = Schwierigkeit in %; *r*_{it} = Trennschärfe. ¹ 5-stufiger Antwortmodus von „weit unterdurchschnittlich“ (0) bis „weit überdurchschnittlich“ (4).
² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 25

Ergebnis der Item- und Skalenanalyse des Fragebogens zur aktuellen Motivation (Eingangsmotivation) sowie Cronbachs α getrennt für die vier Subskalen (I = Interesse, E = Erfolgswahrscheinlichkeit, H = Herausforderung, M = Misserfolgsbefürchtung; N = 92). Die Invertierung eines Items ist durch ein Minus in Klammern gekennzeichnet

Item-Nr.	Subskala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Interesse ($\alpha = .73$)							
1	Ich mag solche Rätsel und Knobeleyen. (I)	4.69	1.5	76	.52	-0.06	-1.09
4	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen. (I)	4.57	1.4	76	.49	-0.44	-0.42
7	Nach dem Lesen der Instruktion erscheint mir die Aufgabe sehr interessant. (I)	4.84	1.3	81	.48	-0.51	-0.15
11	Bei Aufgaben wie dieser brauche ich keine Belohnung, sie machen mir auch so viel Spaß. (I)	4.77	1.4	80	.58	-0.15	-0.60
17	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten. (I)	2.98	1.7	50	.43	0.50	-0.73
Erfolgswahrscheinlichkeit ($\alpha = .75$)							
2	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein. (E)	4.10	1.4	69	.52	0.07	-0.34
3	Wahrscheinlich werde ich die Aufgabe nicht schaffen. (E-)	4.11	1.5	68	.64	-0.17	-0.89
13	Ich glaube, das kann jeder schaffen. (E)	3.08	1.5	51	.41	0.49	-0.46
14	Ich glaube, ich schaffe diese Aufgabe nicht. (E-)	4.42	1.6	74	.63	-0.10	-0.84
Herausforderung ($\alpha = .61$)							
6	Die Aufgabe ist eine richtige Herausforderung für mich. (H)	5.08	1.2	85	.31	-0.19	-0.48
8	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde. (H)	5.55	1.5	93	.41	-1.22	1.07
10	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen. (H)	5.92	1.0	99	.41	-1.10	1.49
15	Wenn ich die Aufgabe schaffe, werde ich schon ein wenig stolz auf meine Tüchtigkeit sein. (H)	5.00	1.5	83	.47	-0.78	0.44
Misserfolgsbefürchtung ($\alpha = .85$)							
5	Ich fühle mich unter Druck, bei der Aufgabe gut abschneiden zu müssen. (M)	3.92	1.7	65	.63	0.05	-1.05
9	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte. (M)	3.92	1.9	65	.79	-0.08	-1.26
12	Es ist mir etwas peinlich, hier zu versagen. (M)	3.90	1.8	65	.77	-0.14	-1.07
16	Wenn ich an die Aufgabe denke, bin ich etwas beunruhigt. (M)	3.87	1.6	64	.60	-0.05	-0.83
18	Die konkreten Leistungsanforderungen hier lähmen mich. (M)	2.67	1.6	45	.55	0.73	-0.48

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³ Getrennt für die vier Subskalen.

Tabelle A – 26

Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den ersten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen *Glatter Verlauf* und *Absorbiertheit* (N = 92)

Item-Nr.	Subskala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Glatter Verlauf ($\alpha = .82$)							
2	Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	3.65	1.8	61	.79	0.24	-1.00
4	Ich habe keine Mühe, mich zu konzentrieren.	4.83	1.7	80	.55	-0.61	-0.83
5	Mein Kopf ist völlig klar.	4.76	1.6	79	.55	-0.44	-0.97
7	Die richtigen Gedanken/Bewegungen kommen wie von selbst.	3.50	1.6	58	.70	0.15	-0.80
8	Ich weiß bei jedem Schritt, was ich zu tun habe.	2.60	1.5	43	.66	0.74	-0.24
9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	2.99	1.6	50	.77	0.52	-0.47
Absorbiertheit ($\alpha = .62$)							
1	Ich fühle mich optimal beansprucht.	4.57	1.8	76	.45	-0.54	-0.73
3	Ich merke gar nicht, wie die Zeit vergeht.	5.07	1.8	84	.39	-0.79	-0.45
6	Ich bin ganz vertieft in das was ich gerade mache.	5.53	1.2	92	.47	-1.11	1.07
10	Ich bin völlig selbstvergessen.	3.85	1.4	64	.32	-0.15	-0.36

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe.¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7).² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.³ Getrennt für die beiden Subskalen.

Tabelle A – 27

Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den zweiten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen Glatter Verlauf und Absorbiertheit ($N = 92$)

Item-Nr.	Skala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Glatter Verlauf ($\alpha = .91$)							
2	Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	4.04	1.7	67	.84	-0.15	-1.06
4	Ich habe keine Mühe, mich zu konzentrieren.	4.68	1.7	78	.60	-0.39	-1.04
5	Mein Kopf ist völlig klar.	4.47	1.7	74	.73	-0.23	-0.97
7	Die richtigen Gedanken/Bewegungen kommen wie von selbst.	3.66	1.8	61	.77	-0.05	-1.23
8	Ich weiß bei jedem Schritt, was ich zu tun habe.	3.02	1.6	50	.74	0.42	-0.55
9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	3.43	1.7	57	.77	0.20	-0.97
Absorbiertheit ($\alpha = .74$)							
1	Ich fühle mich optimal beansprucht.	5.00	1.7	83	.53	-0.97	0.24
3	Ich merke gar nicht, wie die Zeit vergeht.	4.87	2.0	81	.59	-0.60	-1.00
6	Ich bin ganz vertieft in das was ich gerade mache.	5.51	1.5	92	.60	-1.09	0.77
10	Ich bin völlig selbstvergessen.	3.92	1.6	65	.44	-0.23	-0.47

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³ Getrennt für die beiden Subskalen.

Tabelle A – 28

Ergebnis der Item- und Skalenanalyse für die Flow-Kurz-Skala für den dritten Messzeitpunkt sowie Cronbachs α getrennt für die beiden Subskalen *Glatter Verlauf* und *Absorbiertheit* ($N = 90$)

Item-Nr.	Subskala und Wortlaut der zugehörigen Items	M^1	SD	P^2	r_{it}^3	Schiefe	Exzess
Glatter Verlauf ($\alpha = .92$)							
2	Meine Gedanken bzw. Aktivitäten laufen flüssig und glatt.	4.33	1.8	72	.85	-0.16	-0.96
4	Ich habe keine Mühe, mich zu konzentrieren.	4.60	1.8	77	.72	-0.44	-0.89
5	Mein Kopf ist völlig klar.	4.67	1.8	78	.77	-0.42	-0.94
7	Die richtigen Gedanken/Bewegungen kommen wie von selbst.	3.78	1.7	63	.79	-0.05	-0.98
8	Ich weiß bei jedem Schritt, was ich zu tun habe.	3.16	1.6	53	.73	0.24	-1.00
9	Ich habe das Gefühl, den Ablauf unter Kontrolle zu haben.	3.54	1.6	59	.79	-0.07	-1.06
Absorbiertheit ($\alpha = .78$)							
1	Ich fühle mich optimal beansprucht.	4.94	1.6	82	.59	-0.86	0.26
3	Ich merke gar nicht, wie die Zeit vergeht.	5.07	2.0	84	.57	-0.74	-0.84
6	Ich bin ganz vertieft in das was ich gerade mache.	5.53	1.6	92	.67	-1.38	1.44
10	Ich bin völlig selbstvergessen.	4.16	1.7	69	.52	-0.11	-0.57

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %; r_{it} = Trennschärfe. ¹ 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ² Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung. ³ Getrennt für die beiden Subskalen.

Tabelle A – 29

Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum ersten Messzeitpunkt¹. Die Invertierung eines Items ist durch ein Minus in Klammern gekennzeichnet

Item-Nr.	Subskala und Wortlaut der zugehörigen Items	M^2	SD	P^3	Schiefe	Exzess
Interesse						
13	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen. (I)	4.66	1.6	78	-0.44	-0.64
18	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten. (I)	3.18	1.9	53	0.48	-1.03
Erfolgswahrscheinlichkeit						
11	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein. (E)	3.57	1.6	60	0.24	-0.76
12	Wahrscheinlich werde ich die Aufgabe nicht schaffen. (E-)	2.89	2.0	48	0.76	-0.68
Herausforderung						
14	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde. (H)	4.74	1.8	79	-0.65	-0.69
16	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen. (H)	5.90	1.1	98	-1.05	1.43
Misserfolgsbefürchtung						
15	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte. (M)	3.60	2.0	60	0.28	-1.25
17	Es ist mir etwas peinlich, hier zu versagen. (M)	3.68	2.0	61	0.16	-1.34

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %. ¹ Cronbachs α wurde nicht berechnet, da dieser Koeffizient für zwei Items pro Skala nicht definiert ist. ² 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ³ Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 30

Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum zweiten Messzeitpunkt¹. Die Invertierung eines Items ist durch ein Minus in Klammern gekennzeichnet

Item-Nr.	Subskala und Wortlaut der zugehörigen Items	M^2	SD	P^3	Schiefe	Exzess
Interesse						
13	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen. (I)	4.51	1.8	75	-0.38	-0.89
18	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten. (I)	3.16	1.9	53	0.43	-1.10
Erfolgswahrscheinlichkeit						
11	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein. (E)	3.83	1.7	64	-0.09	-0.83
12	Wahrscheinlich werde ich die Aufgabe nicht schaffen. (E-)	3.48	2.2	58	0.21	-1.49
Herausforderung						
14	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde. (H)	4.67	2.0	78	-0.51	-1.03
16	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen. (H)	5.88	1.2	98	-1.24	1.64
Misserfolgsbefürchtung						
15	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte. (M)	3.43	2.1	57	0.39	-1.28
17	Es ist mir etwas peinlich, hier zu versagen. (M)	3.47	2.1	58	0.29	-1.38

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %. ¹ Cronbachs α wurde nicht berechnet, da dieser Koeffizient für zwei Items pro Skala nicht definiert ist. ² 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ³ Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 31

Ergebnis der Item- und Skalenanalyse für die acht ausgewählten Items des Fragebogens zur aktuellen Motivation während des Bearbeitungsprozesses zum dritten Messzeitpunkt ¹. Die Invertierung eines Items ist durch ein Minus in Klammern gekennzeichnet

Item-Nr.	Subskala und Wortlaut der zugehörigen Items	M^2	SD	P^3	Schiefe	Exzess
Interesse						
13	Bei der Aufgabe mag ich es, das statistische Anwendungsprogramm selbst zu erforschen. (I)	4.54	1.8	76	-0.38	-0.89
18	Eine solche Aufgabe würde ich auch in meiner Freizeit bearbeiten. (I)	3.36	2.0	56	0.37	-1.07
Erfolgswahrscheinlichkeit						
11	Ich glaube, der Schwierigkeit dieser Aufgabe gewachsen zu sein. (E)	3.87	1.7	64	-0.14	-0.96
12	Wahrscheinlich werde ich die Aufgabe nicht schaffen. (E-)	3.51	2.1	59	0.14	-1.46
Herausforderung						
14	Ich bin sehr gespannt darauf, wie gut ich hier abschneiden werde. (H)	4.72	2.0	79	-0.55	-1.03
16	Ich bin fest entschlossen, mich bei dieser Aufgabe voll anzustrengen. (H)	5.81	1.3	97	-1.37	1.66
Misserfolgsbefürchtung						
15	Ich fürchte mich ein wenig davor, dass ich mich blamieren könnte. (M)	3.30	2.0	55	0.47	-1.17
17	Es ist mir etwas peinlich, hier zu versagen. (M)	3.34	2.1	56	0.40	-1.31

Anmerkungen. M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit in %. ¹ Cronbachs α wurde nicht berechnet, da dieser Koeffizient für zwei Items pro Skala nicht definiert ist. ² 7-stufiger, endpunktbenannter Antwortmodus von „trifft nicht zu“ (1) bis „trifft zu“ (7). ³ Höhere Werte des Schwierigkeitsindex sprechen für eine geringere Schwierigkeit der Itembeantwortung.

Tabelle A – 32

Ergebnis der Item- und Skalenanalyse für die Leistung bei der Bearbeitung der sieben Aufgaben des Arbeitsblattes (Variable: „Leistung_Arbeitsblatt“) sowie Cronbachs α zur Bestimmung der Reliabilität ($\alpha = .69$; $N = 92$)

Item-Nr.	Aufgabe mit Itemwortlaut (abgekürzt) ¹	Max.	M	SD	P ²	r _{it}	Schiefe	Exzess
1	Angabe von Mittelwert und Standardabweichung für die Gesamtstichprobe, sowie getrennt für Frauen und Männer ($N = 92$)	6	4.14	2.5	69	.38	-0.81	-1.18
2	Angabe des Korrelationskoeffizienten, des Signifikanzniveaus, sowie Interpretation des Ergebnisses ($n = 68$)	4	1.17	1.0	37	.61	0.43	-0.50
3	Erstellen eines Streudiagramms in SPSS ($n = 64$) 0 Punkte: keine Grafik 1 Punkt: zu den Daten passende Grafik 2 Punkte: korrektes Streudiagramm	2	1.09	1.0	77	.68	-0.18	-1.96
4	Interpretieren eines Boxplots (Median Frauen und Männer, Ausreißer und Extremwerte, Verteilung) ($n = 50$)	4	0.63	0.9	28	.57	1.05	-0.41
5	Erstellen eines gruppierten Balkendiagramms ($n = 40$) 0 Punkte: keine Grafik 1 Punkt: zu den Daten passende Grafik 2 Punkte: Balkendiagramm 3 Punkte: Gruppiertes Balkendiagramm	3	0.70	1.2	54	.56	1.23	-0.30
6	Erstellen einer Häufigkeitstabelle ($n = 17$) 1 Punkt: korrekte Tabelle 1 Punkt: korrekte Tabelle, korrekte Variable 2 Punkte: Angabe von absoluter und relativer Häufigkeit aus einer Häufigkeitstabelle	4	0.20	0.6	47	.38	2.73	6.01
7	Fälle auswählen: Angabe getrennter Korrelationskoeffizienten und p-Werte für Männer und Frauen und Interpretation des Ergebnisses (4 Kennwerte plus Interpretation; $n = 9$)	5	0.17	0.8	36	.14	5.21	26.61

Anmerkungen. Max. = Maximale Punktzahl; M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit; r_{it} = Trennschärfe. ¹ n gibt die Anzahl der Personen wider, die die Aufgabe bearbeitet haben. ² Mit Inangriffnahme-Korrektur; d.h. einbezogen wurden nur die Werte von Teilnehmern, welche die Aufgabe bearbeitet haben.

Tabelle A – 33

Ergebnis der Item- und Skalenanalyse für die Leistung beim Lösen der sieben Aufgaben in SPSS (Variable: „Leistung_SPSS“; $\alpha = .78$; $N = 92$)

Item-Nr.	Aufgabe mit Itemwortlaut (abgekürzt) ¹	Max.	M	SD	P ²	r _{it}	Schiefe	Exzess
1	Produktion des Ergebnisses (M, SD) für die Gesamtstichprobe, sowie getrennt für Frauen und für Männer (N = 92)	3	0.73	0.4	32	.37	-1.02	-0.71
2	Produktion einer korrekten Korrelationstabelle in SPSS (n = 68)	1	0.77	0.4	97	.55	-1.32	-0.23
3	Erstellen eines Streudiagramms in SPSS (n = 64) 0 Punkte: keine Grafik 1 Punkt: zu den Daten passende Grafik 2 Punkte: korrektes Streudiagramm	2	1.09	1.0	75	.73	-0.18	-1.96
4	Erstellen eines Boxplots in SPSS (n = 50) 0 Punkte: keine Grafik 1 Punkt: zu den Daten passende Grafik 2 Punkte: korrekter Boxplot	2	0.78	0.9	69	.69	0.45	-1.70
5	Erstellen eines gruppierten Balkendiagramms (n = 40) 0 Punkte: keine Grafik 1 Punkt: zu den Daten passende Grafik 2 Punkte: Balkendiagramm 3 Punkte: Gruppiertes Balkendiagramm	3	0.70	1.2	54	.68	1.23	-0.30
6	Erstellen einer Häufigkeitstabelle (n = 17) 0 Punkte: keine Grafik 1 Punkt: korrekte Tabelle, falsche Variable 2 Punkte: korrekte Tabelle und Variable	2	0.20	0.6	20	.43	2.73	6.01
7	Neue Korrelationstabelle (n = 9)	1	0.05	0.2	56	.36	4.00	14.29

Anmerkungen. Max. = Maximale Punktzahl; M = Itemmittelwert; SD = Standardabweichung des Items; P = Schwierigkeit; r_{it} = Trennschärfe. ¹ n gibt die Anzahl der Personen wider, welche die Aufgabe bearbeitet haben. ² Mit Inangriffnahme-Korrektur; d.h. einbezogen wurden nur die Werte von Teilnehmern, welche die Aufgabe bearbeitet haben.

A 4.2.2 Voraussetzungsprüfungen

Tabelle A – 34

Deskriptive Statistiken für die Verteilung der Skalensummenwerte inklusive der Ergebnisse des Kolmogorov-Smirnov-Tests auf Normalverteilung ($N \geq 90$)

Skala ¹	<i>M</i>	<i>SD</i>	Schiefe	Exzess	<i>Z</i>	Asymp.
Statistikvorwissen ²	2.56	0.6	-0.23	0.12	1.05	.22
VECA ²	1.83	0.5	0.16	0.25	0.67	.37
PRACOWI ²	2.38	1.5	0.27	-0.20	1.37	.05
FAM (Eingangsmotivation) ²						
Interesse	4.35	1.0	-0.16	-0.70	0.77	.59
Erfolgswahrscheinlichkeit	3.93	1.1	-0.19	-0.46	0.87	.43
Herausforderung	5.39	0.9	-0.90	1.27	1.31	.06
Misserfolgsbefürchtung	3.66	1.4	-0.01	-0.96	0.91	.04
Aktuelle Motivation (Mot_end)	9.73	4.24	-0.06	-0.58	0.59	.88
Flow ²						
Messzeitpunkt t1	4.24	1.0	-0.23	0.25	0.68	.74
Messzeitpunkt t2	4.36	1.2	-0.35	-0.39	0.64	.81
Messzeitpunkt t3	4.47	1.3	-0.52	0.01	0.60	.87
Flow_end	4.50	1.3	-0.55	-0.01	0.60	.87
Explorationsstrategie						
Systematische Exploration	41.30	12	-0.29	-0.17	0.67	.77
Versuch-und-Irrtum	4.12	5	3.08	14.41	2.04	.01
Rigide Exploration	21.27	15	0.72	-0.22	0.96	.32
Informationssuche	5.5	7	0.52	-0.28	2.06	.01
Leistung						
Arbeitsblatt	10.57	7	-0.06	-0.97	1.03	.24
SPSS	5.77	4	0.25	-1.08	1.06	.22

Anmerkungen. *Z* = Prüfgröße Kolmogorov –Smirnov-Tests; Asymp. = Asymptotische Signifikanz. ¹ Außer für die Skala PRACOWI werden die Ergebnisse für die einzelnen Skalensummenwerte berichtet. Für die Skala PRACOWI wird der aggregierte Wert über alle Items berichtet. ² Skalennittelwert auf Itemebene.

A 4.2.3 Ergebnisse ergänzender statistischer Analysen

In Tabelle A – 35 sind die bivariaten Korrelationen der Untersuchungsvariablen in Studie 3 aufgeführt. Die detaillierte Übersicht zu den Interkorrelationen von Flow-Erleben, Explorationsstrategien und der Leistung zu drei Messzeitpunkten sind separat in Tabelle A – 36 dargestellt.

Tabelle A – 35

Interkorrelationen zwischen dem Vorwissen, dem Flow-Erleben, der Häufigkeit der Nutzung der systematischen Explorationsstrategie und der Leistung (N ≥ 90)

Variablen	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 Geschlecht	.01	-.12	.17	.13	-.40**	-.28**	-.30**	-.02	-.01	-.09	.01	-.03	-.02	-.02	-.09	-.15
Eingangsmotivation																
2 Interesse		.34**	.34**	-.06	.07	.16	.24*	.35**	.33**	.40**	.02	.02	-.02	.03	-.02	-.03
3 Erfolgswahrscheinlichkeit			.02	-.44	-.09	.08	.34**	.13	.28**	.22*	.05	.19	.12	.17	.14	.12
4 Herausforderung				.12	-.09	-.06	-.22*	.24*	.20	.20	-.10	-.09	-.05	-.12	-.04	-.03
5 Misserfolgsbefürchtung					.07	.04	-.14	-.30**	-.30**	-.21*	.01	-.19	-.06	-.14	-.13	-.08
Vorwissen/Vorerfahrung																
6 Statistikvorwissenstest ¹						.26*	.19	.07	.02	.06	-.04	.07	.04	.03	.15	.25*
7 PRACOWI ²							.35**	.23*	.35**	.26*	-.04	.25*	.26*	.16	.27**	.29**
8 VECA ³								.07	.17	.19	.03	.24*	.30**	.20	.31**	.31**
Flow-Erleben																
9 Flow_t1									.77**	.72**	.08	.30**	.23*	.27**	.32**	.32**
10 Flow_t2										.84**	.06	.38**	.38**	.32**	.46**	.46**
11 Flow_t3											.11	.26*	.41**	.25*	.42**	.43**
Explorationsstrategie																
12 Systematische Exploration_t1												.17	.23*	.69*	.25*	.25*
13 Systematische Exploration_t2													.45**	.83**	.67**	.68**
14 Systematische Exploration_t3														.46**	.71**	.67**
15 Systematische Exploration_t1/t2															.64**	.64**
Leistung																
16 Leistung_SPSS																.97**
17 Leistung_Arbeitsblatt																

Anmerkungen. * $p < .05$, ** $p < .01$; t1, t2, t3 = Messzeitpunkte 1, 2 und 3; t1/t2: Messzeitpunkte 1 und 2. ¹4 Items. ²5 Items. ³10 Items; Kodierung Geschlecht: Männer = -1, Frauen = 1.

Tabelle A – 36

Interkorrelationen zwischen den unsystematischen Explorationsstrategien, dem Flow-Erleben und der Leistung zu den drei Messzeitpunkten der Untersuchung ($N \geq 90$)

Variablen	2	3	4	5	6	7	8	9	10	11	12	13
1 Leistung_Arbeitsblatt	.32**	.46**	.43**	-.64**	-.73**	-.65**	-.04	-.28**	-.01	-.08	-.17	-.34**
2 Flow_t1		.77**	.72**	-.25**	-.30**	-.30**	-.20	-.18	.05	-.18	-.22	-.12
3 Flow_t2			.84**	-.29**	-.48**	-.44**	-.16	-.28**	.13	.03	-.26*	-.19
4 Flow_t3				-.27**	-.39**	-.48**	-.07	-.07	.12	-.01	-.18	-.31**
5 Rigide Exploration_t1					.47**	.41**	-.04	.34**	-.03	-.21*	.05	.21*
6 Rigide Exploration_t2						.67**	.07	.28**	.01	-.11	.05	.13
7 Rigide Exploration_t3							.05	.26*	.07	-.10	-.01	.18
8 Versuch-und-Irrtum_t1								.29**	.34**	-.08	.03	-.07
9 Versuch-und-Irrtum_t2									.18	-.22*	-.01	-.13
10 Versuch-und-Irrtum_t3										-.02	.04	-.18
11 Informationssuche_t1											.14	.21*
12 Informationssuche_t2												.30**
13 Informationssuche_t3												

Anmerkungen. * $p < .05$, ** $p < .01$; t1, t2, t3 = Messzeitpunkte 1, 2 und 3.