

**Der Einfluss von leistungsunabhängigem Lob
auf Arbeitsprozess und Problemlöseleistung:
Eine experimentelle Untersuchung**

Diplomarbeit im Fachbereich Psychologie

Universität Konstanz

vorgelegt von Christian Brehm

März 2003

Mein Dank gilt

In erster Linie Frau Prof. Dr. Sonnentag für ihre hervorragende Betreuung über die gesamte Zeit trotz der räumlichen Distanz, meinen Versuchsleitern Julius und Betti auch für ihre Hilfe bei den Transkriptionen, meiner Freundin Kathi für die Zweitkategorisierung und die Korrekturen, Herrn Dr. Wegge dafür, dass ich meine Ergebnisse im Forschungskolloquium vorstellen durfte, Claudi für ihre hilfreichen Anregungen und zuletzt dem Mittagsstammtisch in der Uni, der es immer wieder aufs neue geschafft hat, mich auf andere Gedanken zu bringen.

Inhaltsverzeichnis

1	Zusammenfassung.....	15
2	Einleitung	18
3	Theorie und Hypothesen.....	20
3.1	Übersicht	20
3.2	Bisherige Forschung zur kurzfristigen Auswirkung von positivem Feedback auf die Arbeitsleistung	20
3.2.1	Günstige Wirkung von positivem Feedback	21
3.2.2	Ungünstige Wirkung von positivem Feedback	21
3.3	Die „Feedback Intervention Theory“ (Kluger & DeNisi, 1996).....	22
3.3.1	Die fünf Basisannahmen	23
3.3.2	Weitere Herausforderungen für eine umfassende Feedback InterventionTheory	27
3.3.2.1	Feedback Interventionen können sich negativ auf Lernen auswirken	27
3.3.2.2	Affekt - ausgelöst durch eine Feedback Intervention – kann sich auf die Leistung auswirken	27
3.3.3	Regulierung der Aufgabenleistung über drei Prozesse	31
3.3.4	Der Metaaufgabenprozess	33
3.4	Aufgabenspezifische Selbstwirksamkeit als Variable	38
3.4.1	Selbstwirksamkeit und Aufgabenstandards.....	38
3.4.2	Selbstwirksamkeit und Art der Gedanken.....	39

3.4.3	Selbstwirksamkeit als Moderator	40
3.5	Aufgabenleistung	41
3.5.1	Aufgabenkompetenz	41
3.5.2	Lob und Aufgabenleistung	41
3.5.3	Art der Gedanken als Mediator	42
3.6	Alle Hypothesen im Überblick	44
4	Methode	47
4.1	Übersicht	47
4.2	Stichprobe	47
4.3	Untersuchungsdesign und Durchführung	48
4.4	Manipulation der unabhängigen Variablen	50
4.5	Operationalisierung und Erfassung der abhängigen Variablen	51
4.5.1	Aufgabenleistung	51
4.5.2	Spezifische und unspezifische persönliche Aufgabenstandards ..	56
4.5.3	Ort der Aufmerksamkeit	57
4.5.4	Freude und Aktivierung	62
4.6	Erfassung der aufgabenspezifischen Selbstwirksamkeit	63
4.7	Auswertung der Daten	64
4.7.1	Auswertung der verbalen Daten	64
4.7.2	Auswertung der Fragebogendaten	66
5	Ergebnisse	69
5.1	Übersicht	69

5.2	Manipulationchecks	69	
5.2.1	Bekanntheitsgrad der Aufgabe	69	
5.2.2	Manipulation der unabhängigen Variable	70	
5.2.3	Qualität der Manipulation	70	
5.2.4	Ausschluss von Datensätzen	73	
5.3	Testen der Hypothesen	73	
5.3.1	Der Einfluss von Lob auf Freude (H1) und Aktivierung (H2)	73	
5.3.2	Der Einfluss von Lob auf den Ort der Aufmerksamkeit (H3)	74	
5.3.3	Der Einfluss von Lob auf den persönlichen Aufgabenstandard (H4)	77	
5.3.4	Der Zusammenhang von Freude (H5) und Aktivierung (H6)	mit der Art der Gedanken	78
5.3.5	Der Einfluss von Lob auf die Veränderung des persönlichen Aufgabenstandards bei Personen mit relativ hoher aufgabenspezifischer Selbstwirksamkeit (H7)	78	
5.3.6	Der Zusammenhang von aufgabenspezifischer Selbstwirksamkeit mit der Art der Gedanken (H8)	80	
5.3.7	Aufgabenspezifische Selbstwirksamkeit als Moderator	zwischen Lob und der Art der Gedanken (H9)	82
5.3.8	Die Auswirkung von Lob auf die Arbeitsleistung (H10)	86	
5.3.9	Der Ort der Aufmerksamkeit als Mediator (H11)	86	
5.4	Weiterführende Analysen	90	
5.4.1	Glaubwürdigkeit des Lobs	90	
5.4.2	Explorative Datenanalyse	90	
5.4.3	Geschlecht	93	
5.4.3.1	Haupteffekte	94	
5.4.3.2	Hypothesen	97	
5.4.3.3	Geschlechtsspezifische Zusammenhänge	108	
5.4.4	Vorhersage der Leistung	111	

6	Diskussion.....	114
6.1	Übersicht.....	114
6.2	Zusammenfassung der Ergebnisse.....	114
6.2.1	Hypothesen und explorative Ergebnisse.....	114
6.2.2	Geschlecht	117
6.2.3	Vorhersage der Leistung.....	120
6.3	Mögliche Ursachen und Erklärungen für die Ergebnisse	123
6.4	Methodenkritik: Stärken und Schwächen der Studie.....	134
6.5	Grenzen der Studie und Vorschläge für die weitere Forschung ...	139
6.6	Praktische Relevanz der Studie.....	143
6.7	Abschluss (Konklusion).....	144
7	Literaturverzeichnis.....	145
8	Anhang	158

Tabellenverzeichnis

Tabelle 1: Mittelwerte, Standardabweichungen und Extremwerte aller Leistungsindikatoren, nach Geschlecht sortiert.....	53
Tabelle 2: Faktorenanalyse mit den Ladungen der aufgabenbezogenen Aufmerksamkeitsitems und den unspezifischen Aufgabenzielen (AZ) auf drei Faktoren	60
Tabelle 3: Faktorladungen aller selbstbezogenen Aufmerksamkeitsitems.....	61
Tabelle 4: Korrelationen der Aktivierungs- und Freudeskalen	63
Tabelle 5: Mittelwerte, Standardabweichung und Interkorrelationen der abhängigen Variablen Ort der Aufmerksamkeit	68
Tabelle 6: Vergleich der beiden Versuchsleiter innerhalb der Experimentalbedingung, vor und nach Ausschluss von Datensätzen aufgrund der Manipulationchecks	72
Tabelle 7: Einfaktorielle Varianzanalyse mit Messwiederholung zur Überprüfung des Einflusses von Lob auf die Veränderung der Freude und der Aktivierung (H1 und H2)	74
Tabelle 8a: Einfaktorielle Varianzanalyse mit Messwiederholung zur Überprüfung des Einflusses von Lob auf den Ort der Aufmerksamkeit (Denkprotokoll; Fragebogen Tabelle 8b, H3)	76
Tabelle 9: Einfaktorielle Varianzanalyse mit Messwiederholung zur Überprüfung des Einflusses von Lob auf die Veränderung des persönlichen, spezifischen Aufgabenstandards (H4).....	77
Tabelle 10: Einfaktorielle Varianzanalyse mit Messwiederholung zur Überprüfung des Einflusses von Lob auf die Veränderung des persönlichen, spezifischen Aufgabenstandards bei Personen mit relativ hoher aufgabenspezifischer Selbstwirksamkeit vorher (H7)	79

Tabelle 11: Korrelationsanalyse zur Überprüfung des Zusammenhangs von aufgabenspezifischer Selbstwirksamkeit (vorher und nachher) mit dem Ort der Aufmerksamkeit (H8)	80
Tabelle 12: Unterschiede zwischen hoher und niedriger aufgabenspezifischer Selbstwirksamkeit nachher hinsichtlich der Variablen des Orts der Aufmerksamkeit (H8).....	82
Tabelle 13: Korrelationen zwischen der Bedingung und dem Ort der Aufmerksamkeit vor und nach der Manipulation bei Personen mit hoher versus niedriger aufgabenspezifischer Selbstwirksamkeit vorher und Überprüfung der Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher (H9).....	85
Tabelle 14: Einfaktorielle Varianzanalyse und Kruskal-Wallis-Test zur Überprüfung des Einflusses von Lob auf die Arbeitsleistung (Hypothese 10).....	86
Tabelle 15: Korrelationen zwischen dem Ort der Aufmerksamkeit und den Leistungsmaßen	88
Tabelle 16: Korrelationen zwischen dem Ort der Aufmerksamkeit und der Bedingung und den Leistungsmaßen und der Bedingung	89
Tabelle 17: Korrelation der Freude und der Aktivierung mit dem Ort der Aufmerksamkeit	91
Tabelle 18: Korrelation der Freude, der Aktivierung, der aufgabenspezifischen Selbstwirksamkeit und den drei persönlichen Aufgabenstandards mit den Leistungsmaßen	93
Tabelle 19: Einfaktorielle Varianzanalyse und Kruskal-Wallis-Test zur Überprüfung des Einflusses des Geschlechts auf den Ort der Aufmerksamkeit	95
Tabelle 20: Überprüfung von Hypothese 9, getrennt nach Geschlecht	99
Tabelle 21: Test von Hypothese 11 bei den Frauen	108
Tabelle 22: Korrelation der Freude und der Aktivierung mit dem Ort der Aufmerksamkeit, getrennt nach Geschlecht	109

Abbildungsverzeichnis

Abbildung 1: Vereinfachte Darstellung möglicher Auswirkung von Feedback je nach Inhalt des Feedbacks in Anlehnung an die Feedback Intervention Theory (Kluger & DeNisi, 1996).....	33
Abbildung 2: Schematische Darstellung aller Hypothesen (außer H4 und H7).....	46
Abbildung 3: Die Auswirkung der Bedingung auf die Häufigkeit von positiven, selbstbezogenen Gedanken (in Prozent) nach der Manipulation in Abhängigkeit von der aufgabenspezifischen Selbstwirksamkeit vorher (H9).....	84
Abbildung 4: Veränderung der aufgabenspezifischen Selbstwirksamkeit bei den weiblichen Versuchspersonen in Abhängigkeit von der Bedingung.....	96
Abbildung 5: Die <i>signifikante</i> Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der negativen selbstbezogenen Gedanken nach der Manipulation bei den Frauen (Hypothese 9).....	100
Abbildung 6: Die <i>nicht signifikante</i> Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der negativen selbstbezogenen Gedanken nach der Manipulation bei den Männern (Hypothese 9).....	101
Abbildung 7: Die <i>signifikante</i> Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der Leistungsbewertung bei den Frauen (Hypothese 9).....	102
Abbildung 8: Die <i>nicht signifikante</i> Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der Leistungsbewertung bei den Männern (Hypothese 9).....	103
Abbildung 9: Die <i>nicht signifikante</i> Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der sonstigen Gedanken nach der Manipulation bei den Frauen (Hypothese 9).....	104

Abbildung 10: Die <i>tendenziell signifikante</i> Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der sonstigen Gedanken nach der Manipulation bei den Männern (Hypothese 9).....	105
Abbildung 11: Die Auswirkung der Bedingung auf die Anzahl Züge in Abhängigkeit vom Geschlecht (Hypothese 10)	106
Abbildung 12: Die Auswirkung der Bedingung auf die Relative Problemlöseleistung in Abhängigkeit vom Geschlecht (Hypothese 10).....	107
Abbildung 13: Die Auswirkung der Bedingung auf die Effizienz in in Abhängigkeit vom Geschlecht (Hypothese 10)	107
Abbildung 14: Vorschlag eines Modells zur Auswirkung von gleichgeschlechtlichem Lob auf die Problemlöseleistung, basierend auf Ideen der Feedback Intervention Theory (Kluger & DeNisi, 1996).	142

Variablenverzeichnis

Code	Bedeutung	Skala	Quelle	Reliabilität (α)
<u>Fragebogen vor der Bearbeitung</u>				
vas	Vorher Aufgabenstandard	keine	Pilotstudie (n =10)	Einzelitems: Fehler, Züge, Zeit
vsw	Vorher Selbstwirksamkeit (SW)	„Aufgabenspezifische SW“	Thompson & Perlini, 1998	4 Items (.70)
vem	Vorher Freude	„Freude“	Mano, 1991	4 Items (.75)
va ve	Vorher Aktivierung Vorher „elation“	Eine Skala „Aktivierung“	Mano, 1991	6 Items (.69)
<u>Denkprotokoll während der Bearbeitung</u>				
1.1	Aufgabenbezogene Aufmerksamkeit auf Aufgabenbearbeitung	keine, Kategorie „Aufgabenbearbeitung“	selbst gebildet	Interraterübereinstimmung (.88)
1.2	Aufgabenbezogene Aufmerksamkeit auf Aufgabenschwierigkeit	keine, Kategorie „Aufgabenschwierigkeit“	selbst gebildet	Interraterübereinstimmung (.57)
2	Sonstige Aufmerksamkeit	keine, Kategorie „Sonstige Gedanken“	selbst gebildet	Interraterübereinstimmung (.75)
3.2*	Selbstbezogene Aufmerksamkeit auf die pos. Selbstbewertung	Keine, Kategorie „Positive Selbstbewertung“	selbst gebildet	Interraterübereinstimmung (.81)
3.3	Selbstbezogene Aufmerksamkeit auf die neg. Selbstbewertung	Keine, Kategorie „Negative Selbstbewertung“	selbst gebildet	Interraterübereinstimmung (.84)

Anmerkung: * Kategorie 3.1 „Aufmerksamkeit auf die Leistungsbewertung“ fiel wegen mangelnder Statements weg

Code	Bedeutung	Skala	Quelle	Reliabilität (α)
<u>Fragebogen nach der Bearbeitung</u>				
em	Nachher Emotion	„Emotion“	Mano, 1991	4 Items (.85)
ae	Nachher Aktivierung Nachher „elation“	Eine Skala „Aktivierung“	Mano, 1991	6 Items (.68)
as	Nachher Aufgabenstandard	keine	Pilotstudie (n =10)	Einzelitems
nsw	Nachher Selbstwirksamkeit (SW)	„Aufgabenspezifische SW“	Thompson & Perlini, 1998	4 Items (.75)
on	Aufgabenbezogene Aufmerksamkeit	„On-task attention“ <i>Neu gebildet:</i> Nachdenken, Fehlervermeidung, Geschwindigkeit	Kanfer & Ackerman, 1989	3 Items Nachdenken (.80) 2 Items Fehlervermeidung (.82) 2 Items Geschwindigkeit (.76)
LB	Selbstbezogene Aufmerksamkeit auf die Leistungsbewertung	„Leistungsbewertung“	Kanfer & Ackerman, 1989	3 Items (.92)
Pos	Selbstbezogene Aufmerksamkeit auf pos. Selbsteinschätzung	„Positive Selbsteinschätzung“	Kanfer & Ackerman, 1989	3 Items (.64)
Neg	Selbstbezogene Aufmerksamkeit auf neg. Selbsteinschätzung	„Negative Selbsteinschätzung“	Kanfer & Ackerman, 1989	2 Items (.84)
bek	„bekannt“: 2 Kontrollitems	keine	Pilotstudie (n =10)	2 Einzelitems
lob	9 Kontrollitems	keine	Pilotstudie (n =10)	9 Einzelitems
az	3 Aufgabenziele (Zeitziel, AZ Züge, AZ Fehler)	keine	Pilotstudie (n =10)	3 Einzelitems

1 Zusammenfassung

Problemstellung. Feedback wirkt sich manchmal positiv, manchmal neutral auf die Arbeitsleistung und die vermittelnden Prozesse aus. Besonders kontraintuitiv erschien mir die Beobachtung, dass sich positives Feedback auch negativ auf die Arbeitsleistung auswirken kann. Eine Voraussetzung für eine negative Auswirkung von Lob sollte nach der Feedback Intervention Theory (Kluger & DeNisi, 1996) die kurzfristige Verschiebung des Aufmerksamkeitsfokus auf das Selbst der Person sein und eine damit verbundene Neuverteilung der Aufmerksamkeitskapazität. Die Fragestellung der vorliegenden Arbeit war deshalb: Wie wirkt sich Lob auf die Person aus, die gelobt wird? Was passiert mit den Gedanken, den Gefühlen, der Aufmerksamkeit und der Handlung einer Person nach einem Lob?

Theorie und Hypothesen. Ausgehend von der Feedback Intervention Theory (Kluger & DeNisi, 1996) wurden 11 Hypothesen abgeleitet. Die Hauptannahme war, dass Lob sich negativ auf die Aufgabenleistung einer kognitiv anspruchsvollen Aufgabe auswirken sollte. Der vermittelnde Prozess dieser Wirkung sollte die Aufmerksamkeitsverteilung sein. Einfluss auf diese Verteilung sollte das Lob auf verschiedene Weise haben. Erstens sollten gelobte Personen mehr Aufmerksamkeit ihrem Selbst widmen (direkter Einfluss des Lobs). Zweitens sollten gelobte Personen sich mehr freuen und deshalb mehr Aufmerksamkeit sonstigen Dingen schenken. Drittens sollten gelobte Personen aktiver sein und mehr Aufmerksamkeit der Aufgabe schenken. Die Wirkung des Lobs auf die Aufmerksamkeitsverteilung sollte von der aufgabenspezifischen Selbstwirksamkeit moderiert werden.

Methode. Mit Hilfe eines Experiments mit Messwiederholung wurden die Hypothesen im Doppelblindversuch getestet. 67 Studierende der Uni Konstanz wurden während der Bearbeitung der „Turm von Hanoi“ Aufgabe mit 6 Scheiben unabhängig von ihrer Leistung nach 2 Minuten gelobt. Dabei wurden die Männer nur von einem männ-

lichen Versuchsleiter, die Frauen nur von einem weiblichen Versuchsleiter gelobt. In der Kontrollbedingung, wurde den Versuchspersonen ein neutrales Statement gegeben. Die Versuchspersonen mussten vorher und nachher einen Fragebogen über ihre Aktivierung, Freude, Aufgabenstandards, aufgabenspezifische Selbstwirksamkeit, Ort der Aufmerksamkeit (nur nach der Bearbeitung) und Kontrollvariablen über den Bekanntheitsgrad der Aufgabe und die Glaubwürdigkeit des Lobs ausfüllen. Während der gesamten Bearbeitungszeit mussten die Versuchspersonen laut denken. Die Denkprotokolle wurden nach dem Ort der Aufmerksamkeit kategorisiert.

Ergebnisse. Die Hauptergebnisse waren, dass sich Lob nicht direkt auf die Freude, die Aktivierung und den Ort der Aufmerksamkeit auswirkte. Die einzige direkte Auswirkung von Lob auf die Aufmerksamkeit war, dass die gelobten Versuchspersonen ihre Leistung weniger stark mit der Leistung anderer verglichen. Indirekt wirkte sich Lob über die aufgabenspezifische Selbstwirksamkeit und das Geschlecht auf die Aufmerksamkeitsverteilung aus. Personen, die sich die Aufgabe zutrauten äußerten mehr positive selbstbezogene Gedanken nach Lob als Personen, die sich die Aufgabe nicht zutrauten. Frauen, die sich die Aufgabe zutrauten und Männer, die sich die Aufgabe nicht zutrauten profitierten von gleichgeschlechtlichem Lob: Männer machten sich tendenziell weniger sonstige Gedanken und Frauen machten sich signifikant weniger negative selbstbezogene Gedanken. Hinderlich für die Bearbeitung der Aufgabe war gleichgeschlechtliches Lob für Männer, die sich die Aufgabe zutrauten: sie machten sich tendenziell mehr sonstige Gedanken. Frauen, die sich die Aufgabe vor der Bearbeitung nicht zutrauten, verglichen sich nach Lob seltener mit anderen, um ihre Leistung zu bewerten. Unabhängig von Lob zeigte sich, dass je mehr (weniger) Freude die Versuchspersonen empfanden, desto weniger (mehr) sonstige Gedanken hatten sie. Die Aktivierung hing nicht mit den aufgabenbezogenen Gedanken zusammen. Lob wirkte sich nur bei den Frauen positiv auf die Leistung aus. Der Ort der Aufmerksamkeit – vor allem die negativen selbstbezogenen Gedanken – eignete sich gut zur Vorhersage der Leistung. Bei den Frauen eignete sich die Bedingung am besten, bei den Männern die Aufmerksamkeit auf die Geschwindigkeit.

Diskussion. Anders als erwartet wirkte sich Lob nicht auf die Aufmerksamkeitsverteilung aus. Dies könnte mit der fehlenden Selbstrelevanz des (künstlichen) Lobs zusammenhängen. Unabhängig von der Aufmerksamkeitsverteilung wirkte sich Lob positiv auf die Leistung der Frauen und nicht auf die Leistung der Männer aus. Es wird vermutet, dass Lob von Frauen eher angenommen wird als von Männern und die verbesserte Arbeitsatmosphäre (zum Beispiel weniger Leistungsdruck) zur Leistungsverbesserung beitrug. Lob wirkte sich indirekt über die aufgabenspezifische Selbstwirksamkeit und das Geschlecht auf die Aufmerksamkeitsverteilung aus. Dieses Ergebnis wird über die unterschiedliche Attribution des Lobs bei Männern und Frauen in Abhängigkeit von der aufgabenspezifischen Selbstwirksamkeit erklärt.

2 Einleitung

„Net broddelt isch gnuag g`lobt“
(Nicht gemeckert ist genug gelobt)

Dieses Sprichwort aus dem schwäbischen Raum ist vermutlich nicht nur kennzeichnend für die „Lobkultur“ der Schwaben, sondern für ganz Deutschland. Vielleicht ist das fehlende Lob historisch begründbar: Um Deutschland innerhalb von 50 Jahren wirtschaftlich auf das heutige Niveau zu bringen, war Arbeit notwendig. Jede Arbeitskraft wurde gebraucht, alle haben gelernt, was es heißt, zu arbeiten ohne Dank und Anerkennung von außen dafür zu bekommen. Diese Einstellung der Nachkriegsgeneration – hart zu arbeiten ohne Lob oder Dank von irgendjemand zu bekommen – hat sich vielleicht bis heute hartnäckig gehalten. Im Ausland wird sowohl der typische Deutsche als auch deutsche Unternehmen heute noch als fleißig und zuverlässig angesehen. Diese Charakterisierung oder Stereotypisierung lässt sich aber auch in deutschen Sprichwörtern („Erst die Arbeit, dann das Vergnügen“ oder spezifisch für die Schwaben „Schaffe, schaffe, Häusle baue...“) wiederfinden. In Amerika wird sicher nicht weniger erfolgreich gewirtschaftet und gearbeitet: Jedoch herrscht in amerikanischen Unternehmen eine andere Lobkultur vor. Amerikanische Manager werden trainiert, ihre Mitarbeiter zu loben um ihnen das Gefühl von Wertschätzung zu vermitteln. Das entspricht dem amerikanischen Selbstbild: Wer stolz auf sein Land und auf sich selbst ist, der will auch das verdiente Lob und die verdiente Anerkennung. Wirtschaftlicher Erfolg durch Lob, dafür könnte Amerika einstehen.

Deutschland könnte - im Vergleich mit Amerika - ein Exempel für wirtschaftlichen Erfolg durch harte Arbeit sein, ohne die dabei verdiente Anerkennung in Form von Lob erhalten zu haben. Was sich so negativ anhört, wird in der Feedback Inter-

vention Theory von Kluger und DeNisi (1996) verteidigt: Kluger und deNisi würden vielleicht aufgrund ihrer Theorie vermuten, dass das Geheimnis des erfolgreichen Deutschlands gerade die fehlende Lobkultur war! Lob – als ein Beispiel für positives Feedback – kann unter bestimmten Voraussetzungen zu schlechterer Leistung führen als kein Lob. Kluger und DeNisi versuchen mit Hilfe einer vorläufigen „Feedback Intervention Theory“ (Kluger & DeNisi, 1996) empirisch bestätigte negative Effekte von Feedback zu erklären. Wenn durch Lob die Aufmerksamkeit des Empfängers zu sehr auf sich selbst gerichtet ist, führe dies zu schlechterer Leistung. In dem vorliegenden Experiment soll es genau darum gehen: Wie wirkt sich Lob auf die Person aus, die gelobt wird? Wie verändern sich ihre Gedanken, ihre Gefühle, ihre Aufmerksamkeit nach einem Lob? Im nun folgenden Theorieteil versuche ich, meine Hypothesen und die zugrundeliegende Feedback Intervention Theory (Kluger & DeNisi, 1996) darzulegen.

3 Theorie und Hypothesen

3.1 Übersicht

In diesem Kapitel versuche ich zuerst kurz darzulegen, warum positives Feedback günstige als auch ungünstige Auswirkungen auf die Leistung haben kann. Die anschließende Beschreibung der integrierenden Feedback Intervention Theory (Kluger & DeNisi, 1996) gliedert sich in folgende Abschnitte: zuerst werde ich die 5 Basisannahmen erklären; dann werde ich darstellen, welche Aspekte für die weitere Entwicklung der Theorie wichtig waren und stelle dann die drei Prozesse vor, die den Kern der Theorie bilden und die Aufgabenleistung regulieren sollen. Aus diesen drei Prozessen greife ich den für meine Arbeit wichtigen Metaaufgabenprozess heraus und beschreibe die vier ihn ausmachenden Mechanismen.

Danach erweitere ich die Theorie um die Variable aufgabenspezifische Selbstwirksamkeit: inwieweit wirkt sie sich auf das Setzen von Aufgabenstandards und die Verteilung der Aufmerksamkeit der Personen aus, inwieweit moderiert sie den Einfluss von Lob auf die Aufmerksamkeitsverteilung? Anschließend gehe ich auf den Zusammenhang zwischen der Aufgabenkompetenz einer Person und dem kognitiven Anspruch einer Aufgabe ein, um die negative Auswirkung von Lob auf die Aufgabenleistung plausibler zu machen. Abschließend wird die Aufmerksamkeitsverteilung als Kern der Feedback Intervention Theory (Kluger & DeNisi, 1996) verantwortlich gemacht für die Auswirkung von Lob auf die Arbeitsleistung. Am Ende des Kapitels findet man eine Übersicht über alle Hypothesen und ein Schaubild über die Zusammenhänge.

3.2 Bisherige Forschung zur kurzfristigen Auswirkung von positivem Feedback auf die Arbeitsleistung

Unter Feedback versteht man alle Arten von Informationen über die Aufgabenleistung einer Person (Kluger & DeNisi, 1996).

3.2.1 Günstige Wirkung von positivem Feedback

Oftmals wurde die günstige Wirkung von positivem Feedback (zum Beispiel Lob) auf die Arbeitsleistung betont (zum Beispiel Ammons, 1956; Pritchard, Jones, Roth, Stuebing, & Ekeberg, 1988). Die Arbeitsleistung werde aus verschiedenen Gründen verbessert: Feather, (1966, 1968) fand heraus, dass Lob während des Arbeitsprozesses als Form der Erfolgsmeldung zu höherer Erfolgserwartung bei der weiteren Arbeit führt. Lob kann zu gesteigerter Selbstwirksamkeit führen (Bandura, 1977). Die erhöhte Selbstwirksamkeit führt direkt zu besserer Arbeitsleistung, unabhängig von persönlichen Fähigkeiten und unabhängig von der vorgegebenen Zielschwierigkeit (Bandura, 1986). Indirekt wirkt sich hohe Selbstwirksamkeit über die Zielsetzung und Aufgabenstrategien auf die Arbeitsleistung aus: Menschen mit hoher Selbstwirksamkeit stecken sich höhere, persönliche Ziele und fühlen sich ihren hohen (auch extern gesetzten) Zielen mehr verpflichtet. Dies führt zu vermehrter Anstrengung und Ausdauer bei der Zielverfolgung. Darüber hinaus wählen Personen mit höherer Selbstwirksamkeit auch die effektiveren Aufgabenstrategien aus. Bessere Aufgabenstrategien führen vor allem bei komplexen Aufgaben zu besseren Leistungen (Latham & Locke, 1991). Lob erhöht die intrinsische Motivation (Koestner, Zuckerman, & Koestner, 1987) und Lob kann dazu führen, dass man sich hohe Fähigkeiten zuschreibt (günstige Selbstattribution, Weiner, Heckhausen & Meyer, 1972). Günstige Selbstattributionen führen wiederum zu einer Erhöhung der Selbstwirksamkeit mit den genannten positiven Auswirkungen (Latham & Locke, 1991).

3.2.2 Ungünstige Wirkung von positivem Feedback

Baumeister, Hutton und Cairns (1990) argumentieren, dass die positiven Auswirkungen der in 3.2.1 genannten Mechanismen (Erfolgserwartung, Selbstwirksamkeit, intrinsische Motivation und günstige Selbstattribution) durch vermehrte Anstrengung und Ausdauer bei der Bewältigung der Aufgabe hervorgerufen werde. Bei komplexen Aufgaben führe vermehrte Anstrengung und Ausdauer jedoch nicht zwangsläufig zu

besseren Ergebnissen. Baumeister (1984) stellte beispielsweise fest, dass Belohnung anspruchsvolle Leistung störte und verschlechterte. Unter bestimmten Bedingungen führe Lob zu negativen Emotionen und zu der Selbstwahrnehmung, niedrige Fähigkeiten zu besitzen (Ungünstige Selbstattribution, Meyer, 1992). Eine Bedingung ist, dass der Lobempfänger den Eindruck hat, der Lobgeber halte nicht viel von den Fähigkeiten des Lobempfängers und lobt nur als Hilfestellung, Ermunterung (Meyer, 1992).

3.3 Die „Feedback Intervention Theory“ (Kluger & DeNisi, 1996)

Feedback und die Wirkung von Feedback auf die Arbeitsleistung wurde in vielen Studien untersucht (vgl. Metaanalyse von Kluger & DeNisi, 1996). Bei einer Feedback Intervention wird das Feedback von einer anderen Person geliefert (Kluger & DeNisi, 1996). Kluger und DeNisi (1996) berechneten in einer Metaanalyse zur Auswirkung von Feedback Interventionen die Effektgröße d (Cohen) für die brauchbaren Studien (zuerst 131 Studien, 12652 Versuchspersonen mit 607 Effektstärken, nach Ausschluss noch 407 Effektstärken) und stellten fest, dass sich neben den vielen positiven Effekten ein Drittel negative Effekte zeigten. Bei einigen Studien zeigte sich kein Effekt des Feedbacks auf die Arbeitsleistung. Um die unterschiedlichen Effekte von Feedback zu erklären, stellten die Autoren eine vorläufige Feedback Intervention Theory vor. Diese Theorie will erklären, wie sich Feedback Interventionen verschiedenster Art auf die Arbeitsleistung auswirken. Kluger und DeNisi (1996) testeten ihre Theorie mittels einer meta-analytischen Moderatoranalyse und fanden heraus, dass sich auch positives Feedback (zum Beispiel Lob) negativ auf die Arbeitsleistung auswirken kann. Dieses Ergebnis wurde von dem Modell vorhergesagt. Über die richtige Vorhersage der Auswirkung einer Feedback Intervention auf die Arbeitsleistung hinaus konnte keine Aussage über die Prozesse gemacht werden, die sich innerhalb des Feedback Empfängers nach Erhalt eines Feedbacks abspielen. Deshalb können die Ergebnisse der meta-analytischen Moderatoranalyse von Kluger und DeNisi (1996) nicht als Bestätigung ihrer Feedback Intervention Theory gesehen werden. Im Folgenden lege ich die für mein Experiment zur Auswirkung von Lob zentralen theoretischen Überlegungen der

Feedback Intervention Theory dar und leite überprüfbare Hypothesen ab. Das Konzept der aufgabenspezifischen Selbstwirksamkeit (Bandura, 1982) versuche ich in das theoretische Gerüst zu integrieren.

3.3.1 Die fünf Basisannahmen

Wichtig für das Verständnis der Feedback Intervention Theory sind fünf Basisannahmen. Diese sind voneinander abhängig und bauen aufeinander auf:

1. Das Verhalten einer Person wird reguliert durch den Vergleich eines erhaltenen Feedbacks mit den persönlichen Zielen (oder Standards) der Person.
2. Diese Ziele (oder Standards) sind hierarchisch organisiert. Die Person hat Oberziele, Ziele und Unterziele.
3. Eine Person hat nur eine begrenzte Aufmerksamkeitskapazität. Damit sich eine Abweichung zwischen Feedback und den Zielen einer Person tatsächlich verhaltensregulierend auswirkt, muss diese Abweichung Aufmerksamkeit erhalten, sprich bemerkt werden.
4. Normalerweise wird die Aufmerksamkeit einer Person der Aufgabe und der Erfüllung der Aufgabenziele gewidmet.
5. Eine Feedback Intervention verändert den Ort der Aufmerksamkeit. Deshalb wirkt sich eine Feedback Intervention auf das Verhalten aus.

Vergleich des Feedbacks mit persönlichen Zielen (Basisannahme 1). In Übereinstimmung mit der Goal Setting Theory (Latham & Locke, 1991) und der Control Theory (Carver & Scheier, 1981) sehen Kluger und DeNisi (1996) das Verhalten einer Person als zielgerichtet. Damit die Person weiß, wie gut ihre bisherige Leistung war, vergleicht sie das Feedback mit ihrem persönlichen Ziel oder Standard. Das Resultat dieser inneren Leistungsbewertung der Person heißt „Feedbacksign“; je nach Ziel und Feedback fällt das Feedbacksign negativ oder positiv aus (Bandura, 1991; Mikulincer, 1994; Podsakoff & Farr, 1989; Locke & Latham, 1990): negativ fällt es aus, wenn das

Feedback schlechter ausfällt als das persönliche Ziel der Person; positiv fällt es aus, wenn das Feedback das Ziel der Person übersteigt. Die Frage ist nun, wie die Person auf das Feedbacksign reagiert. Die Person will die entstandene Diskrepanz reduzieren. Sie kann entweder verändern (das Ziel oder ihr Verhalten) oder eliminieren (das Ziel oder das Feedback). Das wiederholte Vergleichen des Feedbacks mit den Zielen und das reagieren auf eine Abweichung wird „Negative Feedbackschleife“ genannt (Carver & Scheier, 1981). Kluger und DeNisi (1996) unterscheiden 4 mögliche Strategien, die jeweils unterschiedliche Auswirkungen auf die Leistung der Person haben:

- Ist das Feedbacksign negativ, versucht die Person oft durch *erhöhte Anstrengung das Ziel zu erreichen*. Durch das Setzen von hohen, spezifischen Zielen (goal-setting, Locke & Latham, 1990) wird die Wahrscheinlichkeit dieser Strategie erhöht. Voraussetzung dafür ist eine hohe Zielbindung und hohe Erfolgserwartung.
- Falls die Person Zweifel daran hat, ihr Ziel durch weiteres Handeln zu erreichen, *gibt sie es auf*. Diese Strategie tritt bei wiederholtem, extrem negativem Feedbacksign gehäuft auf (Bandura, 1991; Miculincer, 1988b).
- Falls es nicht möglich ist, das Ziel oder den persönlichen Standard aufzugeben, *wird er verändert*. Bei negativem Feedbacksign könnte der Standard gesenkt werden, bei positivem erhöht (Lewin, Dembo, Festinger, & Sears, 1944).
- Bei einem negativem Feedbacksign wird das Feedback *eher zurückgewiesen* als bei einem positiven (Ilgen, 1979).

Die Feedback Intervention Theory baut diese vier Mechanismen in ihr Konzept ein. Das entspricht dem ersten Basisargument: „Verhalten wird reguliert durch den Vergleich des Feedbacks mit den Zielen der Person“. Will man die Auswirkung einer Feedback Intervention auf die Aufgabenmotivation und damit der Leistung vorhersagen, muss man wissen, mit welcher Strategie die Person reagiert. Die Vorhersage der Strategie ist jedoch nicht ganz einfach: man muss das persönliche Ziel der Versuchsperson und das Feedback kennen und die Wahrscheinlichkeit, das eines von beiden verändert werden kann. Um das Feedbacksign nach einer Feedback Intervention noch besser vorherzusagen, muss beachtet werden, dass Personen *mehr als einen Standard*

haben. Um vorherzusagen, wie eine Person aufgrund eines erhaltenen Feedbacks ihre Leistung einschätzt, sollte man wissen, welche verschiedenen Standards als Vergleich herangezogen werden:

- eine Norm (Wie gut sind die anderen?) (Kluger & DeNisi, 1996)
- vorherige Erwartung (Ilgen, 1971; Kluger & DeNisi, 1994; Kluger 2001)
- vorangegangene Leistung (Habe ich Fortschritte gemacht in Richtung meines Ziels?) (Carver & Scheier, 1990)
- Leistung in anderen Gruppen (Kluger & DeNisi, 1996)
- ideales Ziel (Lewin et al., 1944)

Kluger und DeNisi (1996) vermuten, dass verschiedene Feedback-Standard Diskrepanzen gewichtet und aufsummiert in eine affektive Gesamtbewertung der Feedback Intervention eingehen. Das Resultat ist ein Gesamtfeedbacksign aller Diskrepanzen. Weitere Probleme bei der Vorhersage des Gesamtfeedbacksigns: Bei einer neuen Aufgabe ohne vorgegebene Ziele kann es sein, dass die Feedback Intervention sich nicht auf die Anstrengungsleistung der Person auswirkt, weil sie keinen klaren Standard hat, der als Vergleichswert dienen könnte (Erez, 1977). Durch die Feedback Intervention wird eventuell ein neuer Standard gesetzt (Kluger & DeNisi, 1996). Falls die Feedback Intervention natürlich auftretendem Feedback widerspricht, ist es zusätzlich schwierig, das Feedbacksign vorherzusagen (Albright, Levy, & Williams, 1992). Jedoch ist eine objektiv positive Feedback Intervention eine gute Schätzung des subjektiv wahrgenommenen Feedbacksigns (Kluger et al., 1994). Das bedeutet, dass auch ohne Wissen der Standards oder der Ziele, die die Person als Vergleich heranzieht, positives Feedback (zum Beispiel ein Lob) als positiv erlebt wird und die Person ihre Arbeitsleistung positiv bewertet. Um der Herausforderung der multiplen Standards gerecht zu werden, stellten Kluger und DeNisi Basisannahme zwei auf.

Hierarchie von Zielen (Basisannahme 2). Kluger und DeNisi (1996) gehen davon aus, dass die multiplen Ziele der Person hierarchisch organisiert sind. Negative Feedbackschleifen, die durch eine Feedback-Standard Diskrepanz ausgelöst wurden, um die Diskrepanz zu verringern, sind dementsprechend auch hierarchisch organisiert. Oben in der Hierarchie stehen Selbstziele, ganz unten in der Hierarchie stehen körper-

liche Handlungsziele (Kluger & DeNisi, 1996). Empirische Unterstützung erhält die Annahme einer Zielhierarchie über Ergebnisse in der Forschung zur action-identification-theory (Vallacher & Wegner, 1987). Dieselbe Handlung (lesen) kann beispielsweise als „Wörter lesen“ (Handlungsziel) angesehen werden oder als „in meine wissenschaftliche Karriere investieren“ (Selbstziel).

Aufmerksamkeit (Basisannahme 3). Aufmerksamkeitszuteilung wird von Kluger und DeNisi (1996) als Wahrscheinlichkeitsprozess angesehen. Am wahrscheinlichsten ist, dass die Aufmerksamkeit auf einen bestimmten Bereich innerhalb der Hierarchie fokussiert ist. Denkbar ist auch, dass die Aufmerksamkeit simultan auf mehrere Hierarchieebenen gelenkt wird oder schnell zwischen verschiedenen Ebenen und verschiedenen Standards hin und her wechselt (Carver & Scheier, 1981).

Aufgabenebene als normaler Ort der Aufmerksamkeit (Basisannahme 4). In vielen Situationen bleibt die Aufmerksamkeit unterhalb der Selbstebene (Wicklund, 1975) und oberhalb der Aufgabendetailebene (Carver & Scheier, 1981). Der genaue Ort der Aufmerksamkeit variiert mit der Vertrautheit der Aufgabe (Vallacher & Wegner, 1987). Normalerweise ist die Aufgabe selbst im Brennpunkt der Aufmerksamkeit (Kluger & DeNisi, 1996).

Auswirkung einer Feedback Intervention auf den Ort der Aufmerksamkeit (Basisannahme 5). Feedback Interventionen haben meistens ernsthafte Konsequenzen für den Feedbackempfänger. Deshalb erhält eine Feedback Intervention beträchtliche Aufmerksamkeit und hat aus diesem Grund das Potential, die Aufmerksamkeit auf andere Hierarchieebenen als die „normale“ zu lenken (Kluger & DeNisi, 1996).

3.3.2 Weitere Herausforderungen für eine umfassende Feedback Intervention Theory

Zwei weitere Aspekte wurden von Kluger und DeNisi bei der weiteren Entwicklung der Feedback Intervention Theory (drei Prozesse, die sich nach einer Feedback Intervention abspielen und auf den 5 Basisannahmen aufbauen, Abschnitt 3.3.3) miteinbezogen.

- 1 Feedback Interventionen können sich negativ auf Lernen auswirken (Carroll & Kay, 1988).
- 2 Affekt - ausgelöst durch eine Feedback Intervention (Kluger, Lewinsohn, & Aiello, 1994) - kann sich auf die Leistung auswirken.

3.3.2.1 Feedback Interventionen können sich negativ auf Lernen auswirken

Feedback Interventionen, die das Ergebnis der Aufgabenausführung als Inhalt haben, können das Lernen von komplexen Aufgaben verschlechtern und damit die nachfolgende Aufgabenleistung (Azuma & Cronbach, 1966). Carroll und Kay (1988) stellten fest, dass sich zuviel Trainingsinformation – geliefert in Form von Feedback von und auf dem Computer während der Lernphase – eher hinderlich auf das Lernen eines neuen Wortprozessors auswirkte. Darüber hinaus gibt es weitere ungünstige Wirkungen von positivem Feedback (vgl. Abschnitt 3.2.2), die von einer umfassenden Theorie erklärt werden müssen.

3.3.2.2 Affekt - ausgelöst durch eine Feedback Intervention – kann sich auf die Leistung auswirken

Feedback Interventionen können Einfluss auf die Emotionen der Person haben, die das Feedback bekommt (Kluger et al., 1994). Eine affektive Erfahrung setzt sich zusammen aus zwei dominanten Dimensionen (Mano, 1991): pleasantness (oder Stimmung) und arousal (oder Aktivierung). Feedback Interventionen, die ein positives Feedback-

sign (Feedback übertrifft die persönlichen Ziele und Standards) haben, lösen positive Stimmung aus (Freude), Feedback Interventionen mit einem negativen Feedbacksign lösen negative Stimmung aus. Feedback Interventionen mit einem extrem negativen oder extrem positiven Feedbacksign (große Abweichung des Feedbacks von den persönlichen Zielen oder Standards) führen zu einer erhöhten Aktivierung (Kluger et al., 1994). Ich gehe davon aus, dass das Lob positiver ist als die persönliche Einschätzung der Versuchspersonen (Hypothese 1). Das Lob soll darüber hinaus die persönliche Einschätzung der Versuchspersonen substantiell übertreffen (Hypothese 2).

Hypothese 1: Die Personen, die gelobt werden, empfinden mehr Freude als die Personen, die nicht gelobt werden.

Hypothese 2: Die Personen, die gelobt werden, sind aktiver als die Personen, die nicht gelobt werden.

Die **Stimmung**, positiv wie negativ, wirkt sich auf die kognitive Verarbeitung einer Aufgabe und die Aufgabenleistung aus (Übersicht von Isen, 1987, 1999; Abele, 1995; Jerusalem & Pekrun, 1999; Forgas, 2002). Gute Stimmung hat eine Reihe positiver Auswirkungen auf die Leistung. Isen, Daubman und Nowicki (1987) stellten beispielsweise eine *bessere Problemlöseleistung* gut gestimmter Personen verglichen mit Personen in neutraler Stimmung fest (Kerzenaufgabe, Duncker, 1945). Bei Assoziations- tests („Remote Associate Test“ von Mednick, Mednick, & Mednick, 1964) erwiesen sich gut gestimmte Personen als *kreativer*. Sie fanden mehr ungewöhnliche Assoziationen als Personen in neutraler Stimmung (Isen et al., 1987). Murray, Sujana, Hirt und Sujana (1990) stellten fest, dass gut gestimmte Versuchspersonen (im Vergleich zu neutral gestimmten) mehr Gemeinsamkeiten zwischen verschiedenen Stimuli sehen und – nach entsprechender Instruktion – auch mehr Unterschiede benennen können. Sie interpretieren dieses Ergebnis als *höhere kognitive Flexibilität* gut gestimmter Personen. Isen und Means (1983) fanden heraus, dass in komplexen Entscheidungssituationen (Entscheidung für ein von sechs Autos) gut gestimmte Personen durch Vereinfachung (Reduzierung der Merkmalsdimensionen) *schneller auf die Lösung* kamen als neutral gestimmte Personen.

Gute Stimmung hat aber auch negative Effekte auf die kognitive Verarbeitung. Gut gestimmte Versuchspersonen werden eher mit *qualitativ schwachen Argumenten* überzeugt als schlecht gestimmte Versuchspersonen. Schlecht gestimmte Versuchspersonen sind nur mit starken Argumenten zu überzeugen (Bless, Mackie, & Schwarz, 1992). Mackie und Worth (1989) gehen von einer *Aufmerksamkeitsdiffusion* bei positiv gestimmten Personen aus: Obwohl ihre Versuchspersonen hoch motiviert waren (hoher versus geringer Geldanreiz bei genauem lesen von Argumenten; Versuchspersonen strengten sich bei hohem Geldanreiz mehr an), wurden sie nur bei guter Stimmung von schwachen Argumenten überzeugt. Bless et al. (1992) sehen diese Ergebnisse als Beleg dafür, dass Denkstile stimmungsabhängig sind: gut gestimmte Personen verarbeiten Informationen eher *heuristisch-intuitiv*, während schlecht gelaunte Personen Informationen eher *systematisch-analytisch* verarbeiten. Bodenhauser, Kramer und Süsser (1994) fanden heraus, dass Personen mit guter Stimmung bei der Beurteilung von Personen *stärker von Stereotypen* beeinflusst werden als Personen mit neutraler Stimmung. Weary, Marsh, Gleicher, und Edwards (1993) kamen zu dem Ergebnis, dass schlecht gestimmte Personen bei der Personenbeurteilung bevorzugt *diagnostische Informationen* benützen (im Vergleich zu neutral gestimmten Personen). Isen und Daubman (1984) konnten zeigen, dass Versuchspersonen in guter Stimmung bei der Enkodierung von Objekten breitere, weniger detaillierte Kategorien heranziehen als Versuchspersonen in neutraler Stimmung. Dies führte dazu, dass auch eher *untypische Exemplare einer Kategorie zugeordnet* wurden.

Schlecht gestimmte Versuchspersonen geben *bessere Schätzungen* bei einer Kovariationsaufgabe ab als gut gestimmte (Sinclair & Mark, 1992). Bei schlechter Stimmung erfolgte eine *stärkere Fokussierung auf Details* der zur Verfügung stehenden Informationen (Sinclair, 1988). Der Einfluss von guter wie schlechter Stimmung auf den Denkstil sollte als Voreinstellung („default“) betrachtet werden, der durch Verhaltensziele beeinflusst werden kann (Martin, Ward, Achee, & Wyer, 1993). Otto und Schmitz (1993) stellten fest, dass die negative Stimmung nur bei der Bearbeitung einer Aufgabe, bei der analytisches Denken gefordert wurde, zunahm; nicht bei einer kreativen Aufgabe. Otto (1999) geht davon aus, dass eine Affect-Task Interaktion

(Isen, Daubman, & Gorgiolione, 1987) besteht, die Auswirkungen in beide Richtungen zulässt.

Forgas, Ciarrochi und Moylan (2000b) konnten zeigen, dass gut gestimmte („happy“) Personen (zum Beispiel auf die Frage „Soll Australien eine Republik werden?“) *kreativere, originellere, selbstrelevantere* (Ist der Inhalt, das Thema eines selbst ausgedachten Arguments persönlicher Natur?) und *mehr Argumente* produzieren als traurige Personen. Traurige Personen produzieren dafür *qualitativ bessere, überzeugendere Argumente*. Diese (und frühere) empirische Ergebnisse werden in der neueren Forschung (Bless, 2000; Fiedler, 2000; Fiedler, 2001; Fiedler & Bless, 2001; Higgins, 2001; alle zitiert in Forgas, 2002) als Belege für zwei eng an Affekt geknüpfte, komplementäre Verarbeitungsstile interpretiert, die für jede Aufgabe zur Bewältigung benötigt werden. Positiver Affekt soll *Assimilation*, negativer Affekt *Akkommodation* (beide Begriffe ursprünglich aus Piagets' Entwicklungstheorie) begünstigen. Assimilation meint hier einen Verarbeitungsstil, der schema- und routinenbasiert ist und unter Verwendung schon vorhandener Wissensstrukturen Informationen aus der Umwelt aktiv verarbeitet und umwandelt (top-down Verarbeitungsstil). Akkommodation meint hier einen Verarbeitungsstil, der auf die Erfordernisse der Situation fokussiert ist. Deshalb werden Reizen von außen besonders viel Aufmerksamkeit geschenkt und systematisch-erschöpfend verarbeitet und gespeichert (bottom-up Verarbeitungsstil).

Aktivierung wirkt sich auch auf die Aufgabenleistung aus. Nach Eysenck (1992) wird durch hohe Erregung die Aufgabenleistung bei einfachen Vigilanzaufgaben verbessert, bei Kreativitätsaufgaben verschlechtert. Nach Easterbrooks (1959) cue-utilization Hypothese erhöht Aktivierung die Aufmerksamkeit auf zentrale Hinweisreize und reduziert die Aufmerksamkeit auf periphere Hinweisreize. Hohe Aktivierung sei schlecht bei Aufgaben, die häufigen Wechsel zwischen fokalen und peripheren Hinweisreizen erfordern. Matthews, Davies und Lees (1990) fanden dagegen heraus, dass höhere Aktivierung (selbst eingeschätzt) konsistent positiv assoziiert ist mit effizienterer Leistung bei unterschiedlich anspruchsvollen Aufmerksamkeitsaufgaben. Die Beziehung zwischen Aktivierung und Leistung wird von Matthews et al. (1990) als weitgehend linear beschrieben. Meine Hypothesen zum Zusammenhang von Aktivierung

und Stimmung auf die kognitive Verarbeitung erläutere ich im Kontext des Metaaufgabenprozesses, Abschnitt 3.3.4 unter Punkt 4 („Bestimmte affektive Prozesse werden ausgelöst“).

3.3.3 Regulierung der Aufgabenleistung über drei Prozesse

Die Aufgabenleistung wird über drei Prozesse reguliert, die hierarchisch angeordnet sind. Diese bilden den Kern der Feedback Intervention Theory (Kluger & DeNisi, 1996) und fundieren auf den 5 Basisannahmen. Über die 5 Basisannahmen hinaus integrierten die Autoren Affekte in ihre Theorie. Diese entfalten innerhalb des „Metaaufgabenprozesses“ ihre Wirkung (siehe Abschnitt 3.3.4). Mit Hilfe dieser Prozesse erklären die Autoren mögliche Auswirkungen einer Feedback Intervention (auch negative oder affektbedingte, vgl. Abschnitt 3.3.2).

1. Der **Metaaufgabenprozess** (mit Selbstzielen), ganz oben in der Hierarchie
2. Der **Aufgabenmotivationsprozess** (mit Aufgabenzielen), in der Mitte der Hierarchie
3. Der **Aufgabenlernprozess** (mit Lernzielen über Aufgabendetails), ganz unten in der Hierarchie.

Um vorherzusagen, wie sich eine bestimmte Feedback Intervention auf die Leistung auswirkt, ist von zentraler Bedeutung, ob die Aufmerksamkeit des Feedback-Empfängers auf die Metaebene gelenkt wird (auf das Selbst und Selbstziele, Klassifikation des Selbst, Higgins, 1989) oder auf eine der unteren Ebenen (Aufgabenlevel). Bleibt die Aufmerksamkeit auf dem Aufgabenlevel, sollte das Feedback positive Auswirkungen auf die Leistung haben, weil beispielsweise Lernprozesse angeregt werden oder die Aufgabenmotivation steigt. Wird die Aufmerksamkeit auf das Selbst gelenkt, sollte das Feedback negative Auswirkungen auf die Leistung haben, weil aufgrund zusätzlicher, unnötiger affektiver und kognitiver Prozesse der Aufgabe we-

niger Aufmerksamkeit geschenkt wird. Die zentralen Annahmen bestätigten sich für bestimmte Arten von Feedback Interventionen in einer von Kluger und DeNisi (1996) durchgeführten, meta-analytischen Moderatoranalyse mit 470 Effektstärken. Die Autoren stellten fest, dass bei Lob, bei mündlichem Feedback und bei selbstwertbedrohendem oder entmutigendem Feedback (zum Beispiel Tadel) die Arbeitsleistung eher zurückgeht. Die Autoren vermuten, dass durch diese Feedback Interventionen die Aufmerksamkeit auf die Selbstebene gelenkt wird (Abbildung 1).

Hypothese 3: Die Personen, die gelobt werden, schenken dem Selbst mehr Aufmerksamkeit und haben deshalb mehr selbstbezogene und sonstige Gedanken und weniger aufgabenbezogene Gedanken als Personen, die nicht gelobt werden.

Bei Feedback über Leistungsverbesserung (zum Beispiel „Ihre Verkaufszahlen sind in diesem Monat um 5% höher als im letzten Monat“) oder korrektivem Feedback (zum Beispiel „Sie sollten ihren Daumen beim Schreibmaschinenschreiben mit benutzen“) stellten Kluger und DeNisi (1996) einen leistungssteigernden Effekt fest. Die Autoren führen diese Effekte auf ihre innerhalb der Feedback Intervention Theory postulierten Mechanismen und Prozesse zurück. In der weiterführenden Beschreibung der Prozesse der Feedback Intervention Theory gehe ich nur noch auf den für mein Experiment relevanten Metaaufgabenprozess ein.

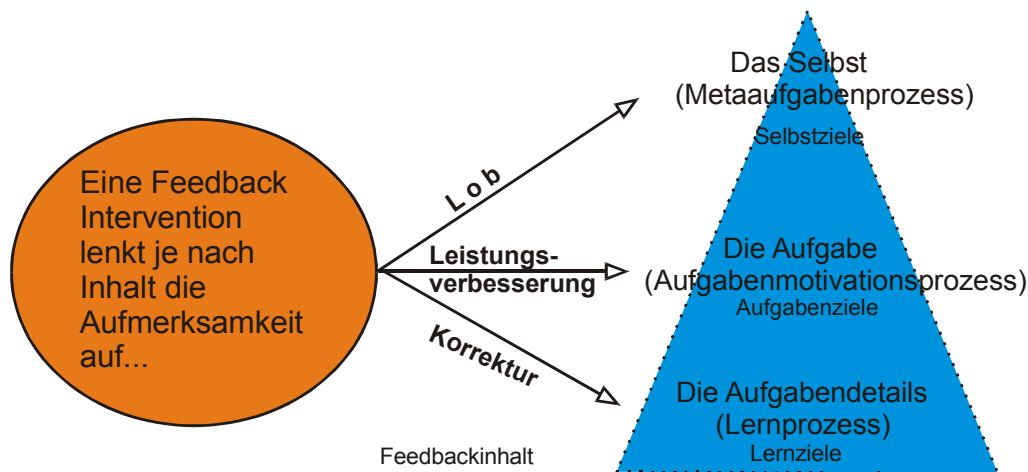


Abbildung 1: Vereinfachte Darstellung möglicher Auswirkung von Feedback je nach Inhalt des Feedbacks in Anlehnung an die Feedback Intervention Theory (Kluger & DeNisi, 1996).

3.3.4 Der Metaaufgabenprozess

Der durch Aufmerksamkeit auf das Selbst (beispielsweise ausgelöst durch Lob der Person) folgende Metaaufgabenprozess beinhaltet die Aktivierung von vier voneinander unabhängigen Mechanismen. Jeder von ihnen soll sich auf die Leistung auswirken.

- 1 **Die Motivation, selbstbezogene Diskrepanzen zu lösen, wird aktiviert**
- 2 **Die Selbstaufmerksamkeit wird erhöht**
- 3 **Die kognitiven Ressourcen werden verringert**
- 4 **Bestimmte affektive Prozesse werden ausgelöst**

Die Motivation, selbstbezogene Diskrepanzen zu lösen, wird aktiviert. Damit eine selbstbezogene Diskrepanz entstehen kann, müssen Selbstziele vorhanden sein, mit denen das selbstbezogene Feedback verglichen wird. Selbstziele können eingeteilt werden in Idealselbst, tatsächliches Selbst (actual self oder Selbstkonzept) und Sollselbst (Higgins, 1989). Das Idealselbst ist das Selbst, das eine Person zu sein wünscht,

das Selbstkonzept ist das Selbst, das eine Person zu sein glaubt und das Sollselbst ist das Selbst, was eine Person glaubt, sein zu müssen (wahrgenommene Verpflichtungen, Erwartungen von außen, Sollstandard). Das Sollselbst und das Idealselbst werden als persönliche Standards oder Selbstführer angesehen (Higgins, 1997). Kluger und DeNisi (1998) greifen diese Klassifikation auf und versuchen, die Auswirkung von selbstbezogenem Feedback auf den Aufgabenstandard und die Leistung über die jeweilige Diskrepanz zu einem der Selbsttypen vorherzusagen. Wird die Aufmerksamkeit der Person auf das Selbst gelenkt (zum Beispiel durch Lob), werden mindestens zwei verschiedene Diskrepanzen salient:

- Die Diskrepanz zwischen Selbstkonzept (zum Beispiel „Ich bin schlecht geeignet für diese Aufgabe“) und Sollselbst (zum Beispiel „Ich sollte eigentlich gut geeignet sein für diese Aufgabe“): Diese Diskrepanz lenkt die Aufmerksamkeit auf mögliche „Verluste“, wenn man sein Soll nicht erreicht. Durch ein Lob während der Aufgabe („Du eignest dich gut für diese Sorte Aufgaben“) könnte das Sollselbstziel als schon erreicht angesehen werden. Als Folge könnte einer der Sollselbstziel relevanten Aufgabenstandards (zum Beispiel „Ich will möglichst wenig Fehler machen“) reduziert werden und die Leistung sich verschlechtern.
- Die Diskrepanz zwischen Selbstkonzept und Idealselbst (zum Beispiel „Ich will am besten von allen Versuchspersonen geeignet sein für diese Aufgabe“): Diese Diskrepanz lenkt die Aufmerksamkeit auf mögliche Gewinne. Durch ein Lob könnte das Ideal-Selbstziel als noch näher gerückt erscheinen. Als Folge könnte einer der Aufgabenstandards erhöht werden, die für das Ideal-Selbstziel relevant sind und die Leistung sich nachfolgend verbessern.

Die Annäherung an persönlich gesetzte Aufgabenstandards können gleichzeitig als Indikator für das Erreichen oder nicht Erreichen eines aktuellen Selbstzieles sowie eines Aufgabenzieles dienen. Falls die zu bearbeitende Aufgabe den Selbstzielen zu wenig dienlich ist, wird eine selbstzielrelevante Aufgabe aktiviert (Kluger & DeNisi, 1996). Die Diskrepanz wird direkt auf der Selbstebene reduziert. Dadurch ist weniger oder gar keine Aufmerksamkeit für die eigentliche Aufgabe zur Verfügung. Die eigentliche Aufgabe wird abgebrochen beziehungsweise mit verringerter Anstrengung fortgeführt (Kluger & DeNisi, 1996).

Falls Lob die Aufmerksamkeit auf das Selbst lenkt und die Versuchsperson die Aufgabe und bestimmte Aufgabenziele als selbstzielrelevant empfindet, ist sie motiviert, ihren ursprünglichen Aufgabenstandard (vor dem Lob gesetzt) zu verändern (vgl. oben beschriebene Mechanismen). Deshalb sollte sich nur in der Lobgruppe – wenn Lob die Aufmerksamkeit kurz auf das Selbst und Selbstziele lenkt (H3) – eine Veränderung der spezifischen, persönlichen und selbstzielrelevanten Aufgabenstandards zeigen. Der veränderte Aufgabenstandard der Lobgruppe könnte als indirekter Beleg der Zielhierarchie (vgl. Basisannahme zwei und Abbildung eins) interpretiert werden.

Hypothese 4: Die Personen, die gelobt werden, verändern eher ihren persönlichen, spezifischen und selbstzielrelevanten Aufgabenstandard als die Personen, die nicht gelobt werden.

In welcher Weise eine Person ihren Aufgabenstandard verändert, hängt meiner Meinung nach stark mit der aufgabenspezifischen Selbstwirksamkeit (Bandura, 1986) zusammen. Hypothese sieben geht auf diesen Zusammenhang ein (vgl. Abschnitt 3.4.1).

Die Selbstaufmerksamkeit wird erhöht. Selbstaufmerksamkeit ist ein Zustand, bei dem die Person Objekt ihrer eigenen Aufmerksamkeit wird. Dies führt zu einem Vergleich ihres Verhaltens mit internen Standards. Die Aufgabenleistung wird nach ihrer Nützlichkeit für höhere Selbstziele eingeschätzt (zum Beispiel einen guten Eindruck machen; Vallacher & Wegner, 1987). Erhöhte Selbstaufmerksamkeit verschlechtert die Leistung bei einer kognitiv fordernden Aufgabe und verbessert die Leistung bei einer weniger kognitiv fordernden Aufgabe. Dieser Effekt wurde auch bei aufgabenuabhängigem Lob erzielt (Baumeister, Hutton, & Cairns, 1990).

Die kognitiven Ressourcen werden verringert. Dadurch, dass die Aufmerksamkeit nicht mehr auf die Aufgabe fokussiert ist und abschweift, kommt es zu Leistungseinbussen aufgrund neuer Verteilung kognitiver Ressourcen (unabhängig von dem Er-

gebnis des Selbstprozesses). Die Aufteilung der begrenzten Ressource „Aufmerksamkeit“ erfolgt auf den Metaaufgabenprozess und den Aufgabenprozess, deshalb verringert sich die Aufgabenleistung vor allem bei nicht-automatisierten Aufgaben (Kanfer & Ackerman, 1989).

Bestimmte affektive Prozesse werden ausgelöst. Affekte entscheiden mit darüber, auf welche Art und Weise unsere kognitiven Ressourcen (hier: Aufmerksamkeit) genutzt werden (Kluger & DeNisi, 1996). Kluger und DeNisi (1998) postulieren als Reaktion auf eine Feedback-Standard Diskrepanz zwei parallel laufende Prozesse: Einen asymmetrischen und einen symmetrischen Prozess hinsichtlich des (Aufgaben)-Standards. Der asymmetrische Prozess besagt, dass je positiver das Feedbacksign ist, desto größer ist das Ausmaß an Freude (Kluger, Lewinson, & Aiello, 1994) und desto größer ist das Ausmaß an sonstigen Gedanken (Kluger, 2001). Der symmetrische Prozess besagt, dass je größer die absolute Diskrepanz ist zwischen dem Standard und dem Feedback (unabhängig von der Richtung), desto höher ist die Aktivierung (Kluger et al., 1994) und desto größer ist das Ausmaß an aufgabenbezogenen Gedanken (Kluger, 2001).

Klugers Prozesshypothesen (2001) bestätigten sich nur teilweise: Je positiver die Diskrepanz, desto größer die Freude. Die Größe der Diskrepanz zwischen Feedback und Standard – operationalisiert als Differenz zwischen erwarteter und erhaltener Schulnote – löste Aktivierung aus. Der weitere Prozessverlauf bleibt jedoch weitgehend unklar: Die subjektiv angegebene Menge an *sonstigen Gedanken* (Fragebogen) korrelierte signifikant positiv mit den aufgabenbezogenen Gedanken, mit der Aktivierung und nicht mit dem Ausmaß an Freude. Die *aufgabenbezogenen Gedanken* korrelierten signifikant negativ mit dem Ausmaß an Freude und signifikant positiv mit der Höhe der Aktivierung. Zwei Dinge sind mir aufgefallen: die hohe signifikante positive Korrelation der aufgabenbezogenen Gedanken mit den sonstigen Gedanken und die subjektive Art der Messung: die Versuchspersonen sollten angeben, wie oft sie in den letzten paar Minuten an etwas bestimmtes gedacht haben (Fragebogen mit 35 Items, Häufigkeit auf einer 0 – 10 Skala angeben). Die beiden Skalen wurden dann mittels einer Faktorenanalyse gebildet. Um klar zwischen aufgabenbezogenen und sonstigen

Gedanken zu unterscheiden und die Einschätzung der Menge der jeweiligen Art der Gedanken nicht der Versuchsperson selbst zu überlassen ist meiner Meinung nach eine objektive Vorgehensweise (zum Beispiel „Lauter Denken“, Ericsson & Simon, 1984) zu empfehlen. Ob die Aktivierung oder die Freude einen stärkeren positiven (oder negativen) Zusammenhang zu den jeweiligen Gedanken – objektiv gemessen – haben, lässt sich nur empirisch explorativ überprüfen. Über die explorative Datenanalyse hinaus formuliere ich zwei scheinbar widersprüchliche Hypothesen (angeregt durch Kluger, 2001):

Hypothese 5: Je größer das Ausmaß an Freude ist, desto mehr sonstige Gedanken werden geäußert.

Hypothese 6: Je aktiverter eine Person ist, desto mehr aufgabenbezogene Gedanken äußert sie.

Geht man davon aus, dass sich Lob genauso stark auf die Aktivierung wie auf die Freude auswirkt, erscheinen die Hypothesen als Widerspruch. Dann würden sich die beiden Prozesse in ihrer Wirkung auf die unterschiedlichen Gedanken aufheben. Unter der Annahme, dass sich Lob jedoch in unterschiedlich starkem Masse auf die Aktivierung und die Freude auswirkt, und diese beiden Mechanismen unabhängig voneinander sind und damit unabhängig voneinander messbar sind (Mano, 1991), kann man die Zusammenhänge zu der Menge aufgabenbezogener und sonstiger Gedanken empirisch überprüfen.

Lob führt aller Wahrscheinlichkeit nach nur zu einer moderaten Aktivierung und diese – in Einklang mit Easterbrooks' cue utilization Hypothese – sollte den von Easterbrook (1959) genannten Effekt bei steigender Aktivierung erkennen lassen:“(Ansteigende Aktivierung)...will reduce the proportion of irrelevant cues employed“ (Easterbrook, 1959). Der Aufmerksamkeitsfokus sollte sich auf die Aufgabe richten. Deshalb formuliere ich einen positiven Zusammenhang zwischen den aufgabenbezogenen Gedanken und der Aktivierung (vgl. auch Matthews, Davies, & Lees (1990): Linearer positiver Zusammenhang von Aktivierung und Leistung).

Einen positiven Zusammenhang der sonstigen Gedanken zu dem Ausmaß an Freude formuliere ich aufgrund der genannten empirischen Ergebnisse zum Einfluss guter Stimmung (siehe 3.3.2.2; vor allem Mackie & Worth (1989): Verbreiterung beziehungsweise Diffusion des Aufmerksamkeitsfokus; Isen et al. (1987) und Isen (1987): Größerer kognitiver Kontext ist aktiviert, mehr ungewöhnliche Wortassoziationen).

3.4 Aufgabenspezifische Selbstwirksamkeit als Variable

Selbstwirksamkeit ist die persönliche Einschätzung, wie gut man die Handlungen ausführen kann, die nötig sind, um eine kommende Situation zu meistern (Bandura, 1982). Aufgabenspezifische Selbstwirksamkeit ist die subjektive Einschätzung, wie gut man die Handlungen ausführen kann, die nötig sind, um eine kommende Aufgabe zu meistern. Ein Beispiel-Item zur Erfassung von aufgabenspezifischer Selbstwirksamkeit, übernommen von Thompson und Perlini (1998), ist: „Ich glaube, dass ich besser bin als die meisten Personen bei dieser Aufgabe“. Die aufgabenspezifische Selbstwirksamkeit könnte sich auf das Setzen von Aufgabenstandards nach Lob auswirken und auf den Ort der Aufmerksamkeit. Darüber hinaus könnte die Auswirkung von Lob auf den Ort der Aufmerksamkeit durch die aufgabenspezifische Selbstwirksamkeit moderiert werden.

3.4.1 Selbstwirksamkeit und Aufgabenstandards

Wie wirkt sich die Selbstwirksamkeit einer Person auf das Setzen von Aufgabenstandards aus? Nach Bandura (1997) visualisieren Personen mit hoher Selbstwirksamkeit Erfolgsszenarios die als positive Führer für ihre Leistung dienen. Geht man davon aus, dass Personen, die gelobt werden, motiviert sind, die durch das Lob ausgelöste Diskrepanz zu einem Selbstziel zu verringern und Personen mit hoher Selbstwirk-

samkeit sich zutrauen, ein höheres Aufgabenziel zu bewältigen um damit ihrem Selbstziel näher zu rücken, formuliere ich folgende Hypothese:

Hypothese 7: Personen, die gelobt werden und über hohe Selbstwirksamkeit verfügen, setzen ihren persönlichen, spezifischen und selbstzielrelevanten Aufgabenstandard eher nach oben als Personen, die nicht gelobt werden und über hohe Selbstwirksamkeit verfügen.

3.4.2 Selbstwirksamkeit und Art der Gedanken

Kluger und DeNisi (1996) sehen Selbstwirksamkeit (nach Bandura, 1986) als einen Meta-Aufgaben-Faktor, der die Person befähigt, schneller die Aufmerksamkeit von der Selbstebene weg zurück auf die Aufgabe zu lenken. Majo und Matsumi (1996) untersuchten dysphorische und nicht-dysphorischen Studierende. Ergebnis war, dass jene Studierende, die ihre Problemlösefähigkeit als gering ansahen, signifikant mehr emotionsfokussierte und aufgabenirrelevante Kommentare und weniger problemfokussierte Kommentare beim lauten Denken von sich gaben. Dysphorische Studierende schätzten ihre Problemlösefähigkeit signifikant geringer ein als nicht dysphorische Studierende, deshalb bleibt unklar, in welchem Ausmaß die Kommentare des Denkprotokolls von der Variabel dysphorisch bzw. nicht-dysphorisch oder der Variable Problemlösefähigkeit beeinflusst wurden.

Bell und Kozlowski (2002) korrelierten selbsteingeschätzte, aufgabenspezifische Selbstwirksamkeit mit dem erworbenen deklarativen und strategischem Wissen über eine bearbeitete Simulationsaufgabe (Tactical navy decision making TANDEM; Dwyer, Hall, Volpe, Cannon-Bowers, & Salas, 1992) und erhielten einen hochsignifikanten positiven Zusammenhang. Geht man davon aus, dass dieses aufgabenbezogene Wissen beim lauten Denken geäußert wird, hätten Personen mit hoher aufgabenspezifischer Selbstwirksamkeit mehr aufgabenbezogene Gedanken.

Blankstein, Toner und Flett (1989) korrelierten selbsteingeschätzte Erfolgserwartung mit der Problemlöseleistung bei einfachen und schwierigen Anagrammaufgaben, negative selbstbezogene, aufgabenirrelevante Gedanken (CIQ, Cognitive Interferenz Questionnaire, Sarason & Stoops, 1978), und retrospektiv aufgelistete Gedanken (Instruktion und Kategorisierung nach Petty & Cacioppo, 1977). Bei selbsteingeschätzter positiver Erfolgserwartung hatten die Versuchspersonen signifikant weniger negative, selbstbezogene, aufgabenirrelevante Gedanken und signifikant mehr potentiell lösungsförderliche, aufgabenbezogene Gedanken. Nach Bandura (1997) erklärt die Selbstwirksamkeitseinschätzung die meiste Varianz innerhalb der eingeschätzten Ergebniserwartung. Aufgabenspezifische Selbstwirksamkeit sollte daher zu gleichen Ergebnissen führen wie Erfolgserwartung oder Ergebniserwartung. In Übereinstimmung mit den aufgelisteten Ergebnissen formuliere ich folgende Hypothese:

Hypothese 8: Je niedriger die aufgabenspezifische Selbstwirksamkeit, desto mehr selbstbezogene und sonstige Gedanken und desto weniger aufgabenbezogene Gedanken hat die Person.

3.4.3 Selbstwirksamkeit als Moderator

In Hypothese 8 postulierte ich einen direkten Zusammenhang zwischen Selbstwirksamkeit und der Art der Gedanken, die sich die Person macht. Lob sollte sich – falls sich die zuvor formulierten Hypothesen bestätigen – ebenfalls ursächlich auf die Art der Gedanken auswirken. Diese Auswirkung könnte durch niedrige Selbstwirksamkeit noch verstärkt und durch hohe Selbstwirksamkeit abgeschwächt werden. Deshalb formuliere ich folgende Moderatorhypothese:

Hypothese 9: Die Auswirkung von Lob auf die Art der Gedanken (Hypothese 3) wird von Selbstwirksamkeit moderiert. Personen mit niedriger Selbstwirksamkeit haben nach Lob mehr selbstbezogene und sonstige Gedanken und weniger aufgabenbezogene Gedanken als Personen ohne Lob. Personen mit hoher

Selbstwirksamkeit haben nach Lob weniger selbstbezogene und sonstigen Gedanken und mehr aufgabenbezogene Gedanken als Personen ohne Lob. Personen mit niedriger aufgabenspezifischer Selbstwirksamkeit haben nach Lob mehr selbstbezogene und sonstige Gedanken und weniger aufgabenbezogene Gedanken als Personen mit hoher aufgabenspezifischer Selbstwirksamkeit.

3.5 Aufgabenleistung

3.5.1 Aufgabenkompetenz

Kluger und DeNisi (1998) benützten den Begriff „task mastery“ (Aufgabenkompetenz), um den individuell unterschiedlichen kognitiven Anspruch einer Aufgabe deutlich zu machen. Aufgabenkompetenz bezeichnet die individuelle, subjektiv empfundene Schwierigkeit, eine Aufgabe zu lösen. Je schwieriger es für jemand ist, eine Aufgabe zu lösen, desto geringer ist die Aufgabenkompetenz der Person und desto kognitiv anspruchsvoller ist die Aufgabe für diese Person (weil die gesamte Aufmerksamkeit für die Aufgabe benötigt wird). Der kognitive Anspruch der Aufgabe wird von der vorhandenen Aufgabenkompetenz der Person bestimmt.

3.5.2 Lob und Aufgabenleistung

Wird die Aufmerksamkeit durch ein Lob auf die Selbstebene gerichtet und kognitive Ressourcen eingebüsst, sollte sich bei einer kognitiv anspruchsvollen Aufgabe positives Feedback negativ auf die Leistung auswirken. Baumeister et al. (1990) konnten zeigen, dass sich auch aufgabenirrelevantes Lob (zum Beispiel ein Kompliment) bei einer anspruchsvollen Aufgabe negativ auf die Leistung auswirkt. Baumeister et al. (1990) erklären diesen Effekt durch die erhöhte Selbstaufmerksamkeit, ausgelöst durch das Lob. Falls eine Person nur eine begrenzte Aufmerksamkeitskapazität zur

Verfügung hat (Dritte Basisannahme, Kluger & DeNisi, 1996) und durch Lob die Aufmerksamkeit von der Aufgabe abschweift (Fünfte Basisannahme, Kluger & DeNisi, 1996), kognitiv anspruchsvolle Aufgaben jedoch die volle Aufmerksamkeit zur guten Bewältigung benötigen, sollte Lob die Leistung bei kognitiv anspruchsvollen Aufgaben eher beeinträchtigen als fördern.

Hypothese 10: Personen, die während der Aufgabenbearbeitung gelobt werden, bringen schlechtere Leistung bei einer kognitiv anspruchsvollen Aufgabe als Personen, die während der Aufgabenbearbeitung nicht gelobt werden.

3.5.3 Art der Gedanken als Mediator

Was sagt die empirische Forschung zu dem Zusammenhang zwischen Feedback und den entstehenden Gedanken (Kognitionen) sowie dem Zusammenhang zwischen den entstehenden Gedanken und der Aufgabenleistung? Mikulincer, Kedem und Segal (1989) untersuchten unter anderem off-task Kognitionen bei Personen, die negatives Feedback bekamen im Vergleich zu einer Kontrollgruppe ohne Feedback. Ergebnis war, dass je öfter negatives Feedback ausgesprochen wurde, desto mehr off-task Kognitionen wurden berichtet (subjektive Einschätzung mit dem CIQ, Cognitive Interference Questionnaire, Sarason, Sarason, Keefe, Hayes, & Shearin, 1986), unabhängig von gegebenen Hinweisreizen (keine, aufgabenrelevante, aufgabenirrelevante). Und je mehr off-task Kognitionen berichtet wurden, während die Versuchspersonen an dem Problem arbeiteten, desto ungenauer und desto ineffizienter wurde die Leistung bei der untersuchten Gedächtnisaufgabe und der visuellen Suchaufgabe. In einer weiteren Studie von Mikulincer (1989) mit den gleichen Aufgaben ergab sich keine Auswirkung der off-task cognitions auf die allgemeine Leistung. Nur bei der höchsten Schwierigkeitsstufe zeigte sich ein negativer Zusammenhang des „mind-wandering“ (eine Frage im CIQ, Sarason & Stoops, 1978) mit der Genauigkeit der Leistung.

Blankstein et al. (1989) korrelierten den CIQ (Sarason & Stoops, Version von 1978 mit „mind wandering“) mit retrospektiv aufgelisteten Gedanken (Instruktion von Petty & Cacioppo, 1977; kategorisiert in selbstbezogene, aufgabenbezogene und sonstige Gedanken, jeweils mit den Unterkategorien positiv, negativ, neutral) mit der Problemlöseleistung bei einer leichteren und einer schwierigen Anagrammaufgabe. Ergebnis war, dass je mehr positive, aufgabenbezogene Gedanken aufgelistet wurden, desto besser war die Problemlöseleistung. Je mehr negative selbstbezogene Gedanken aufgelistet wurden, desto schlechter war die Problemlöseleistung. Bei schweren Anagrammaufgaben war der Zusammenhang mit den jeweiligen Gedanken größer als mit leichten Anagrammaufgaben. Je mehr off-task Kognitionen (CIQ) auftraten, desto schlechter war die Problemlöseleistung. „Mind wandering“ hatte keinen Zusammenhang mit der Anzahl gelöster Anagrammaufgaben.

Sonntag (1998) konnte zeigen, dass Software Designer mit exzellenter Leistung insgesamt weniger aufgabenirrelevante Kognitionen bei der Bearbeitung des „Lift Kontroll Problems“ (Guindon, 1990) verbalisierten als Software Designer mit moderater Leistung. Bei genauerer Analyse der Verteilung und Folge aufgabenirrelevanter Kognitionen innerhalb eines Denkprotokolls zeigte sich, dass Software Designer mit moderater Leistung hochsignifikant länger bei ihren aufgabenirrelevanten Kognitionen verweilten, während Software Designer mit exzellenter Leistung nach einer aufgabenirrelevanten Kognition schneller an der Design Lösung weiterarbeiteten. Die signifikanten Zusammenhänge zwischen Feedback und den entstehenden Gedanken und die signifikanten Zusammenhänge zwischen den entstehenden Gedanken und der Arbeitsleistung deuten auf einen Mediatoreffekt hin. Die daraus abgeleitete Mediatorhypothese lautet:

Hypothese 11: Die Auswirkung von Lob auf die Arbeitsleistung bei einer kognitiv anspruchsvollen Aufgabe wird durch die Art der Gedanken (aufgabenbezogene; selbstbezogene; sonstige Gedanken) mediiert.

3.6 Alle Hypothesen im Überblick

H1: Die Personen, die gelobt werden, empfinden mehr Freude als die Personen, die nicht gelobt werden.

H2: Die Personen, die gelobt werden, sind aktiver als die Personen, die nicht gelobt werden.

H3: Die Personen, die gelobt werden, schenken dem Selbst mehr Aufmerksamkeit und haben deshalb mehr selbstbezogene und sonstige Gedanken und weniger aufgabenbezogene Gedanken als Personen, die nicht gelobt werden.

H4: Die Personen, die gelobt werden, verändern eher ihren persönlichen, spezifischen und selbstzielrelevanten Aufgabenstandard als die Personen, die nicht gelobt werden.

H5: Je größer das Ausmaß an Freude ist, desto mehr sonstige Gedanken werden geäußert

H6: Je aktiver eine Person ist, desto mehr aufgabenbezogene Gedanken äußert sie

H7: Personen, die gelobt werden und über hohe aufgabenspezifische Selbstwirksamkeit verfügen, setzen ihren persönlichen, spezifischen und selbstzielrelevanten Aufgabenstandard eher nach oben als Personen, die nicht gelobt werden und über hohe aufgabenspezifische Selbstwirksamkeit verfügen.

H8: Je niedriger die aufgabenspezifische Selbstwirksamkeit, desto mehr selbstbezogene und sonstigen Gedanken und desto weniger aufgabenbezogene Gedanken hat die Person.

H9: Die Auswirkung von Lob auf die Art der Gedanken (Hypothese 3) wird von aufgabenspezifischer Selbstwirksamkeit moderiert.

Personen mit niedriger aufgabenspezifischer Selbstwirksamkeit haben nach Lob mehr selbstbezogene und sonstige Gedanken und weniger aufgabenbezogene Gedanken als Personen ohne Lob.

Personen mit hoher aufgabenspezifischer Selbstwirksamkeit haben nach Lob weniger selbstbezogene und sonstigen Gedanken und mehr aufgabenbezogene Gedanken als Personen ohne Lob.

Personen mit niedriger aufgabenspezifischer Selbstwirksamkeit haben nach Lob mehr selbstbezogene und sonstige Gedanken und weniger aufgabenbezogene Gedanken als Personen mit hoher aufgabenspezifischer Selbstwirksamkeit

H10: Personen, die während der Aufgabenbearbeitung gelobt werden bringen schlechtere Leistung bei einer kognitiv anspruchsvollen Aufgabe als Personen, die während der Aufgabenbearbeitung nicht gelobt werden.

H11: Die Auswirkung von Lob auf die Arbeitsleistung bei einer kognitiv anspruchsvollen Aufgabe wird durch die Art der Gedanken (aufgabenbezogene; selbstbezogene; sonstige Gedanken) mediiert.

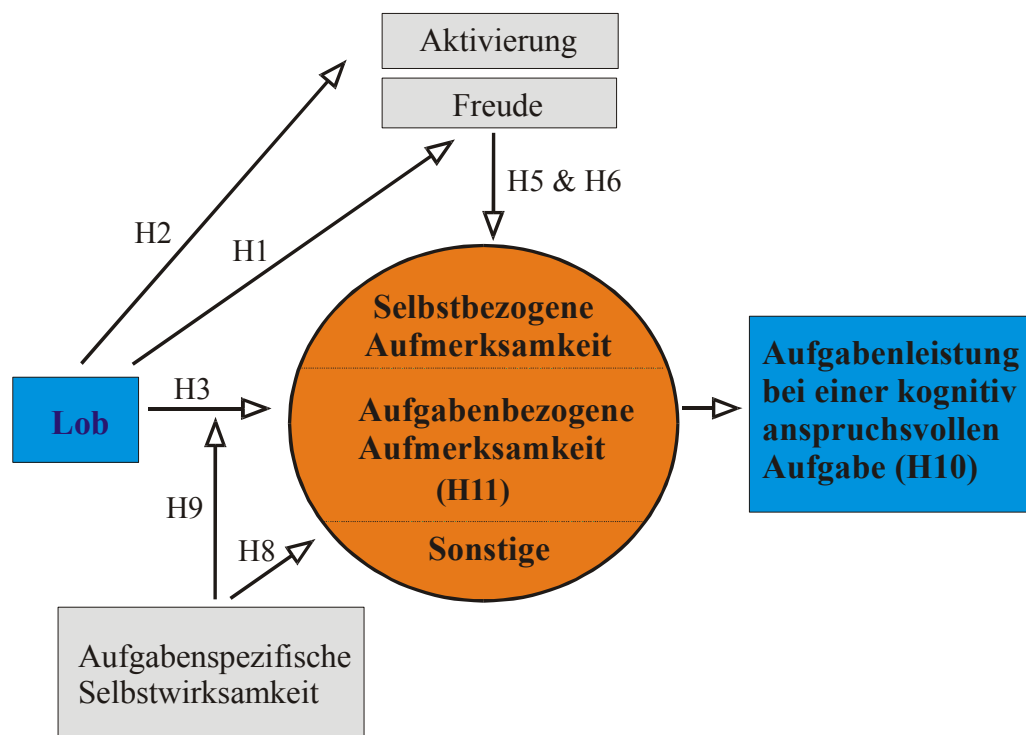


Abbildung 2: Schematische Darstellung aller Hypothesen (außer H4 und H7)

4 Methode

4.1 Übersicht

Nach der Beschreibung der Zusammensetzung der Stichprobe, des Untersuchungsdesigns, der Durchführung der Studie und der Manipulation der unabhängigen Variablen gehe ich auf die Operationalisierung und Erfassung der abhängigen Variablen Aufgabenleistung, Aufgabenstandards, Ort der Aufmerksamkeit, Freude und Aktivierung und die Variable aufgabenspezifische Selbstwirksamkeit ein. Zum Ende des Kapitels erkläre ich, wie die verbalen Daten aus dem Denkprotokoll und die Fragebogendaten ausgewertet wurden.

4.2 Stichprobe

An dem Experiment nahmen 67 Personen teil. Das Experiment wurde an der Universität Konstanz durchgeführt. Es handelte sich deshalb bei allen Personen um Studierende, die freiwillig gegen eine Belohnung von 5,- DM teilnahmen. Außer Psychologiestudierende und Studierende, die die zu lösende Aufgabe schon kannten, durfte jeder mitmachen. Psychologiestudierende wurden ausgeschlossen aufgrund der Befürchtung, sie könnten aufgrund der richtigen (oder falschen) Vermutung über Ziele der Studie nicht mehr unbefangen auf die geplanten Reize reagieren. Fast alle Studienfächer der Universität Konstanz waren vertreten: Wirtschaftswissenschaften mit 11 Personen, Jura und Verwaltung mit jeweils 10 Personen, Soziologie mit 8 Personen, 5 Deutsch-, 4 Mathematik-, 4 Informatik-, 3 Physik- und 3 Chemiestudierende. Die restlichen 9 Personen kamen aus den unterschiedlichsten Fachrichtungen (Anglistik, Bauingenieur, Biologie, Geschichte, Literatur, Politik und keine Angabe). Das Durchschnittsalter der Versuchspersonen betrug 22.4 Jahre mit einer Standardabweichung von 2.6. Die meisten Versuchspersonen waren im ersten Semester ihres Studienfachs (21 Personen, 31.3 %), im dritten Semester (17 Personen, 25.4 %) oder im fünften

Semester (11 Personen, 16.4 %). Die Hälfte der Versuchspersonen war weiblich (33 Personen).

4.3 Untersuchungsdesign und Durchführung

Als Untersuchungsdesign wurde das klassische Experiment (Experimental- und Kontrollgruppe) mit Messwiederholung gewählt. Die Durchführung erfolgte als Doppelblindversuch: weder die Versuchsleiter noch die Versuchspersonen wussten, um was es in dem Experiment ging.

Im Vorfeld führte ich eine Pilotstudie mit 10 Versuchspersonen (5 männlich, 5 weiblich) durch. Die Hauptziele der Pilotstudie waren: die Vorgaben in den Fragebögen sollten angemessen sein, es sollte genügend viel Zeit zur Aufgabenbearbeitung vorhanden sein, der subjektive Schwierigkeitsgrad der ausgewählten Aufgabe musste den Versuchspersonen angepasst werden und praktische Mängel des Designs sollten sichtbar gemacht werden. Die durch die Pilotstudie aufgezeigten Mängel des geplanten Experiments (wie: zu leichte Aufgabe, zu viel Zeit zur Bearbeitung, unklare Instruktion usw.) wurden dadurch identifiziert und vor der Durchführung des Experiments geändert. Die wichtigste Änderung betraf die problematische Manipulation der unabhängigen Variablen, operationalisiert als der Wortlaut des Lobs der Versuchsleiter. Die Schwierigkeit bestand darin, das Lob so zu formulieren, dass es trotz der experimentellen Situation glaubwürdig wirkte.

Zwei Versuchsleiter – Studierende der Universität Konstanz - führten das Experiment durch. Beiden wurde die Durchführung des Experiments erklärt. Die Theorie und Hypothesen wurden dabei ausgelassen, bei Fragen der Versuchspersonen über den Zweck des Experiments verwiesen die Versuchsleiter auf die Aufklärung am Ende des Experiments durch den Diplomanden (Doppelblindversuch). Den Ablauf probten die Versuchsleiter zweimal, bevor sie in einem Versuchsraum der Universität Konstanz die Versuchspersonen testeten. Jeder Versuchsleiter erhielt einen „Versuchsleiterfahrplan“, auf dem der genaue Handlungsablauf zu finden war (Anhang A1).

Die Rekrutierung der Versuchspersonen erfolgte an der gesamten Universität Konstanz, hauptsächlich jedoch im Eingangsbereich der Bibliothek. Die Personen wurden etwa auf folgende Art und Weise rekrutiert: „Hallo, hättest du vielleicht eine halbe Stunde Zeit für ein psychologisches Experiment im Rahmen meiner Diplomarbeit? Es geht dabei um eine Problemlöseaufgabe die der „Turm von Hanoi“ genannt wird. Kennst du diese Aufgabe? Das ist kein Intelligenztest. Am Ende des Experiments bekommst du 5,- DM für die Teilnahme und wirst von mir über Sinn und Zweck meiner Diplomarbeit informiert. Wenn du während der Bearbeitung der Aufgabe keine Lust mehr hast oder aus irgendwelchen anderen Gründen vorzeitig das Experiment abbrechen willst, ist das kein Problem. Alles weitere erfährst du dann von deinem Versuchsleiter wenn du mitmachst“.

Alle Versuchspersonen wurden einzeln getestet. Am Anfang des Experiments begrüßte der Versuchsleiter die Versuchsperson und stellte sich kurz vor. Gleich darauf gab er der Versuchsperson die Instruktion über den Ablauf des Experiments (Anhang A2). Darin wurde unter anderem auf die Anonymität der erhobenen Daten verwiesen, die freiwillige Teilnahme der Versuchsperson herausgestrichen und an die gewissenhafte und ehrliche Beantwortung der Fragen appelliert. Nach der Beantwortung von Verständnisfragen zur Instruktion wurde die Methode des „Lauten Denkens“ (Ericsson & Simon, 1984) anhand einer Wortfindungsaufgabe spielerisch illustriert. Die Versuchsperson versuchte so lange mit 10 anderen Buchstaben sinnvolle Wörter mit sechs Buchstaben zu legen, bis sie das Gefühl hatte, alle Gedanken leicht verbalisieren zu können. Danach füllte die Versuchsperson den Fragebogen „Fragen vor Turm von Hanoi“ (Anhang A3) aus. Bevor die Versuchspersonen jedoch endgültig mit der Bearbeitung des „Turm von Hanoi“ anfangen, wurden sie noch einmal auf die Regeln und Ziele der Aufgabe hingewiesen, und dass jetzt noch ein letztes mal die Möglichkeit zu Fragen bei Unklarheiten bestehe. Der Versuchsleiter hielt während der Aufgabenbearbeitung jeden Zug, jeden Regelverstoß und das erreichte Ergebnis jeder Versuchsperson auf der Tabelle „Leistungsmessung Turm von Hanoi“ fest (mit Begriffsdefinitionen im Anhang A5). Der Versuchsleiter saß in gebührendem Abstand zu der Versuchsperson, damit sie sich nicht gestört fühlte.

Alle geäußerten Gedanken wurden mit einem Tonbandgerät aufgezeichnet. Sagte die Versuchsperson 15 Sekunden lang nichts, wurde sie darum gebeten, alle Gedanken laut zu sagen („*Laut denken bitte*“ oder „*Bitte laut denken*“). Nach zwei Minuten lobte der Versuchsleiter die Versuchsperson ungeachtet ihrer Leistung in der Experimentalbedingung (siehe 4.4); in der Kontrollbedingung erhielt die Versuchsperson den neutralen Kommentar. Die Versuchsperson arbeitete danach noch vier Minuten an der Aufgabe weiter, dann beendete der Versuchsleiter die Aufgabenbearbeitung mit den Worten:

„Ich muss dich hier jetzt leider bei der Bearbeitung unterbrechen. Bevor du dich ein zweites Mal an die „Turm von Hanoi“ Aufgabe machst, bitte ich dich, diesen Fragebogen kurz auszufüllen.“

Zum Schluss musste jede Versuchsperson denselben Fragebogen wie am Anfang ausfüllen, ergänzt um leicht modifizierte Skalen zur Aufmerksamkeit von Kanfer und Ackerman (1989) und Manipulationchecks (Fragebögen nach „Turm von Hanoi“, Anhang A4a: Experimentalbedingung, A4b: Kontrollbedingung). Der Versuchsleiter dankte der Versuchsperson für die Teilnahme und bezahlte ihr 5,- DM. Wer wollte, konnte seine E-Mail-Adresse in eine Liste eintragen, um die Ergebnisse der Untersuchung zu erhalten. Die Versuchspersonen durften, wenn sie wollten, die „Turm von Hanoi“ Aufgabe an der unvollendeten Stelle weiterbearbeiten und bekamen bei Lösungsproblemen Hilfestellung. Am Ende wusste jeder, wie die Aufgabe funktioniert. Der Versuchsleiter schickte die Versuchsperson dann zum Diplomanden, der die Versuchsperson über den Zweck und den Inhalt der Untersuchung aufklärte und die Versuchsperson verabschiedete.

4.4 Manipulation der unabhängigen Variablen

Die unabhängige Variable in diesem Experiment war das Feedback des Versuchsleiters, unabhängig von der Güte der Leistung der Versuchspersonen. Das Feedback er-

folgte während der sechsminütigen Bearbeitungszeit nach zwei Minuten. Wichtig war dabei, dass die Versuchsleiter sich genau an die per Los verteilten Bedingungen (Experimentalgruppe wird gelobt, Kontrollgruppe bekommt neutrales Statement) hielten. Um Geschlechtseffekte auszuschalten und nur die Auswirkung des Lobs an sich zu registrieren, wurden alle männlichen Studierende von einem männlichen Versuchsleiter manipuliert und alle weiblichen Studierende von einer weiblichen Versuchsleiterin. Der Wortlaut des neutralen Kommentars in der Kontrollgruppe war:

„Du arbeitest jetzt seit einiger Zeit an der Aufgabe. Du kannst jetzt noch eine Weile daran arbeiten, dann machen wir etwas anderes.“

Die Versuchsleiter mussten darauf achten, dass sich dieser Satz wie eine sachliche, neutrale Information anhörte, ohne dass die Konnotation von Zeitdruck mitschwang. Bei der Lob-Bedingung wurde folgender Satz nach zwei Minuten gesagt:

*„Das sieht ja schon außerordentlich gut aus.
Du eignest Dich gut für diese Sorte Aufgaben!“*

Die Versuchsleiter sollten diese beiden Sätze möglichst natürlich klingen lassen. Die Versuchspersonen sollten das Gefühl bekommen, das Lob sei ernst gemeint, auf keinen Fall ironisch. Mittels eines Manipulationchecks am Ende des Experiments wurde überprüft, ob die Versuchspersonen sich tatsächlich gelobt fühlten (siehe 5.2).

4.5 Operationalisierung und Erfassung der abhängigen Variablen

4.5.1 Aufgabenleistung

Zur Messung der Aufgabenleistung habe ich die „Turm von Hanoi“ Aufgabe gewählt. Diese klassische Problemlöseaufgabe besteht je nach Schwierigkeitsgrad aus vier, fünf oder sechs pyramidenförmig angeordneten Bechern, Scheiben oder Ringen, je nach

Ausführung. Der kleinste Becher ist oben, der größte Becher unten. Die Aufgabe der Versuchsperson bestand darin, den ganzen Turm von einer Startplattform ausgehend komplett auf einer Zielplattform neu aufzubauen. Der Neuaufbau sollte möglichst schnell, mit möglichst wenig Fehlern und mit minimaler Anzahl Züge erfolgen (siehe Anhang A2). Die minimale Anzahl Züge zur Lösung der Aufgabe errechnet sich aus der Formel $(2^n) - 1$. „n“ steht für die Anzahl Becher, die verwendet werden. Bei 6 Bechern benötigt man minimal $(2^6) - 1 = 63$ Züge bis zur Problemlösung (neu aufgebauter Turm, entweder auf der Ziel- oder Hilfeplattform). Die Versuchsperson durfte dabei eine Zwischenplattform als Hilfe benutzen (Anhang A6). Auf eine Vorgabe der Zielplattform wurde verzichtet, was bedeutete, dass auch auf der Hilfeplattform der Turm neu aufgebaut werden durfte (vgl. Schöttke, 1990; zitiert in Schöttke, 2000). Zwei Regeln galt es beim Neuaufbau des Turms zu beachten: Man durfte immer nur den obersten Becher eines Turms bewegen und auf einen liegenden Becher immer nur einen kleineren Becher legen (Anhang A2).

Diese Aufgabe ist eine klassische Problemlöseaufgabe (Sequenzproblem) bei der weniger durch vermehrte Anstrengung als durch Nachdenken bessere Leistung erbracht wird. Die Vier-Scheiben-Version führt allein durch die Instruktion, möglichst schnell zu arbeiten, zu kürzeren Bearbeitungszeiten und weniger überflüssigen Schrittzahlen (Hacker, Schöppe, Grüllich & Wachter, 1999). Um die Aufgabe kognitiv anspruchsvoller zu machen, wählte ich die Sechs-Scheiben-Version aus. Neun Personen von 67 (13.4 Prozent) bewältigten während des Experiments die Aufgabe innerhalb der sechs Minuten Bearbeitungszeit. Die Aufgabe war somit ausreichend anspruchsvoll.

Alle im Experiment erzielten Werte der Leistungsindikatoren sind in Tabelle 1 deskriptiv und getrennt nach Geschlecht aufgelistet. Die Einzelindizes (Zeit, Züge, Fehler, „Absolut“ = absolute Problemlöseleistung) und die berechneten Indizes („Relativ“ = Absolut / Züge, „Effizienz“ = Relativ / Zeit, „Optimal“ = Effizienz / Fehler) werden im Folgenden beschrieben.

Tabelle 1: Mittelwerte, Standardabweichungen und Extremwerte aller Leistungsindikatoren, nach Geschlecht sortiert.

Leistungsindikatoren	Geschlecht ¹	Minimum	Maximum	M	SD
Zeit ² (s)	m	246	360	352	26
	w	141	360	349	41
Züge	m	20	110	50	24
	w	32	160	79**	28
Fehler ^{2 3}	m	0	5	0.6	1.1
	w	0	10	1.9**	2.1
Absolut ^{2 3} (Punkte)	m	100	500	326	124
	w	200	500	355	83
Relativ ^{2 3} (Punkte)	m	2.04	19.05	7.36	3.25
	w	2.41	9.09	4.89**	1.72
Effizienz ² (Punkte)	m	0.34	3.17	1.26	0.54
	w	0.40	2.66	0.88**	0.44
Optimal (Punkte)	m	0.13	3.17	1.07	0.63
	w	0.05	1.39	0.49**	0.40

Anmerkungen: ¹ Männer: n = 34; Frauen: n = 33

² nicht normalverteilt

³ Varianzen nicht homogen

** Unterschied Männer-Frauen signifikant: $p < .01$ (Varianzanalyse)

Als Leistungsindikatoren wurde die Schnelligkeit (Zeit bis zur Problemlösung), die Anzahl Züge nach sechs Minuten (bzw. bis zur Problemlösung), die Annäherung an die Problemlösung (Teilproblemlösung; auch Teiltürme, vgl. Reimer, 2001) sowie die Häufigkeit von Regelverstößen (= Fehler) bei der Sechs-Scheiben-Version gemessen. Jedes Indize kann für sich ausgewertet werden (Schöttke, 1995; zitiert in Schött-

ke, 2000). Da jedoch nur 9 Versuchspersonen von 67 (13.4 Prozent) innerhalb der 6 Minuten fertig wurden, fällt die Zeit als zwischen den Versuchspersonen unterscheidender Leistungsaspekt weg. Die Anzahl Züge und die Fehleranzahl nach sechs Minuten wurden von den Versuchsleitern auf einer Strichliste (Anhang A5) notiert. Die Annäherung an die Problemlösung wurde operationalisiert über das erfolgreiche Bewältigen von Teilproblemen in Form von Teiltürmen, die nach Ablauf der Zeit auf der Ziel- bzw. Hilfeplattform potentiell errichtet werden konnten. Reimer (2001) beschreibt ein Teilproblem als gelöst, wenn beispielsweise bei einem Vier-Scheiben-Turm die größte unterste Scheibe auf der Zielplattform liege. Bei der Sechs-Scheiben-Version sind auf der Ziel- oder Hilfeplattform folglich Vierer-, Fünfer- und Sechser-untertürme möglich, je nach dem, wie weit die Problemlösung fortgeschritten ist. Um die Annäherung an die Problemlösung zu quantifizieren, bekam eine Versuchsperson Punkte: wenn sie beispielsweise die Basis für einen Viererunterturm auf der Ziel- bzw. Hilfeplattform gelegt hatte, bekam sie 200 Punkte zugeteilt (hatten 16.4 Prozent von $n = 67$ Versuchspersonen). Ein Fünferunterturm gab 300 Punkte (26.9 Prozent) und ein Sechserunterturm 400 Punkte (38.8 Prozent). Wer fertig wurde bekam 500 Punkte (13.4 Prozent). Wer mit der Aufgabe nicht klar kam, bekam 100 Punkte (4.5 Prozent). Diesen Aspekt der Leistung nenne ich die **Absolute Problemlöseleistung**, weil hier das hierarchisch übergeordnete Ziel „Turm aufbauen“ am besten zum Ausdruck kommt. Je mehr Punkte die Versuchsperson erreichte, desto näher kam sie diesem Ziel.

Über die getrennte Auswertung der Problemlöseindizes hinaus gibt es auch Ansätze zur Berechnung zusammengesetzter Indizes (Goel & Grafman, 1995; zitiert in Schöttke, 2000). Für die Auswertung meiner Daten ist interessant, wie viele Züge jemand benötigt, um ein Teilproblem zu lösen. Je weniger Züge eine Versuchsperson braucht, um eine hohe absolute Problemlöseleistung zu erreichen, desto besser. Deshalb berechnete ich die **Relative Problemlöseleistung**:

$$\text{Relative Problemlöseleistung} = \text{Absolute Problemlöseleistung} / \text{Anzahl Züge}$$

Ein Subziel der Aufgabe lautete in der Instruktion, mit möglichst wenigen Zügen den Turm wieder aufzubauen. Ist der Wert für die Relative Problemlöseleistung groß, dann hat die Versuchsperson neben dem übergeordneten Ziel „Turm aufbauen“ dem Subziel „mit möglichst wenig Zügen“ Rechnung getragen.

Um die Schnelligkeit einzelner Versuchspersonen zu berücksichtigen, die vor Ablauf der 6 Minuten fertig waren, teilte ich die Relative Problemlöseleistung durch die benötigte Zeit. Ein großer Wert steht für schnelles, effektives Arbeiten (mit wenig Zügen weit vorangekommen innerhalb kurzer Zeit). Dieses Maß nenne ich **Effizienz**.

$$\text{Effizienz} = \text{Relative Problemlöseleistung} / \text{benötigte Zeit}$$

Das Subziel, möglichst schnell zu arbeiten um möglichst schnell fertig zu werden (Anhang A2) konkurriert eventuell mit dem Subziel, möglichst wenig Fehler zu machen (ein Aspekt von Genauigkeit). Versucht eine Versuchsperson schnell zu arbeiten, macht sie eventuell viele Fehler; versucht sie keine Fehler zu machen, wird sie möglicherweise langsamer bei der Bearbeitung (zu Strategien des speed-accuracy-trade-off siehe Schöttke, 1988, S.56 ff.; zitiert in Schöttke, 2000). Bei der Berechnung der Korrelation zwischen der Fehleranzahl und der Anzahl Züge ($r = .18$; ns) und der benötigten Zeit ($r = .14$; ns) zeigte sich, dass die Fehler zwar positiv, aber nicht signifikant mit beiden Maßen korrelierten. Man kann also von einer relativen Unabhängigkeit der Fehler ausgehen. Trotzdem integriere ich die Fehler in die Berechnungsformel, weil sie als zusätzliche Züge ansonsten nicht in die Leistungsbewertung eingehen. Arbeitet eine Versuchsperson effizient (kommt in kurzer Zeit mit wenig Zügen sehr nah an die Problemlösung heran) und macht wenig (oder keinen) Fehler, dann hat sie optimal gearbeitet. Ich nenne dieses Maß **Optimale Effizienz**: je größer der Wert, desto genauer und effektiver hat die Versuchsperson gearbeitet.

$$\text{Optimale Effizienz} = \text{Effizienz} / \text{Fehleranzahl}$$

Falls jemand mit 0 Fehlern (Regelverstöße) die Aufgabe bewältigte, wurde nur zur Berechnung des Quotienten Optimale Effizienz der Wert 1 zugewiesen; bei einem Fehler der Wert 2 usw.. Dieser Eingriff war notwendig, weil der Nenner (hier die Anzahl Fehler) nicht Null werden darf. Bei Null Fehlern entspricht die Optimale Effizienz mit dieser Berechnung der Effizienz.

4.5.2 Spezifische und unspezifische persönliche Aufgabenstandards

Der Aufgabenstandard ist der Anspruch an sich selbst bei einer Aufgabe. Gemessen wird er über die Frage: „Welches ist das niedrigste Leistungslevel, mit dem du noch zufrieden bist?“ (Mento, Locke & Klein, 1992). Bei der „Turm von Hanoi“ Aufgabe wurden 3 Aufgabenziele extern vorgegeben (neben dem übergeordneten Ziel, die Aufgabe überhaupt zu bewältigen).

Zunächst wurde untersucht, welche **spezifischen, persönlichen Aufgabenstandards** die Versuchspersonen vor und nach der Bearbeitung des Turms von Hanoi verfolgten. Die drei Fragen dazu lauteten:

- (1) *„Welches ist die längste Zeit bis zur Problemlösung, mit der du noch zufrieden bist?“* („Zeitstandard“)
- (2) *„Welches ist die maximalste Anzahl Schritte (Züge) bis zur Problemlösung, mit der du noch zufrieden bist (63 Züge Minimum)?“* („Zugstandard“)
- (3) *„Welches ist die maximalste Anzahl Regelverstöße bis zur Problemlösung, mit der du noch zufrieden bist?“* („Fehlerstandard“)

Die spezifischen Vorgaben (Zeit, Züge, Fehler) zu den drei Fragen kann den Fragebögen entnommen werden (Anhang A, A4a und A4b). Um zu prüfen, ob die Personen aufgrund des Lobs ihre Aufgabenstandards verändern, wurden diese Fragen vor und nach der Bearbeitung der Aufgabe gestellt.

Anschließend wurden nach der Bearbeitung der Aufgabe die **unspezifischen, persönlichen Aufgabenziele** erhoben. Die persönliche Gewichtung der 3 Aufgabenziele wurde nach der Bearbeitung der „Turm von Hanoi“ Aufgabe mit 3 Fragen erfasst (5 Punkte Likert-Skala, 1 = trifft gar nicht zu; 5 = trifft vollständig zu):

- (1) *„Ich wollte die Aufgabe möglichst schnell lösen“*. („Zeitziel“)
- (2) *„Ich wollte die Aufgabe mit möglichst wenigen Schritten (Zügen) lösen“*.
(Aufgabenziel „Züge“)
- (3) *„Ich wollte die Aufgabe mit möglichst wenigen Regelverstößen bewältigen“*.
(Aufgabenziel „Fehler“)

4.5.3 Ort der Aufmerksamkeit

Einige Studien erfassten den Ort oder die Verteilung der Aufmerksamkeit subjektiv mit Fragebögen (Kanfer & Ackerman, 1989; Kluger, 2001). Der selbst eingeschätzte Zeitanteil an Aufmerksamkeit (Kanfer & Ackerman, 1989) oder die Häufigkeit eines bestimmten Gedankens innerhalb einer bestimmten Zeiteinheit (Kluger, 2001) ist meiner Meinung nach von der Person selbst schwierig einzuschätzen. Die Items der Fragebögen erfassten immer nur einen Teilbereich der Gedanken und somit nicht die gesamte Verteilung der Aufmerksamkeit. Eine objektive Erfassung der Aufmerksamkeit sollte als zusätzliche Informationsquelle herangezogen werden. Verschiedene Facetten und Bereiche der Aufmerksamkeit können über eine sich ergänzende Methodenvielfalt (subjektive und objektive Erfassung) in den Dienst der Konstruktvalidierung gestellt werden. Zur objektiven Erfassung der Art der Gedanken (selbstbezogen, aufgabenbezogen, sonstige) als Indikator für den Ort der Aufmerksamkeit (Selbst, Aufgabe, sonstiger Ort) bot sich die Methode des „Lauten Denkens“ (Ericsson & Simon, 1984) an. Mit dieser Methode können Informationen über kognitive Prozesse gewonnen werden.

In meinem Experiment wurden die Versuchspersonen aufgefordert, jeden Gedanken, der ihnen während der Bearbeitung der Aufgabe durch den Kopf ging, zu äußern. Die Instruktion dazu lautete (nach Sonnentag, 1998): „Ich bin nicht nur an der Lösung des „Turm von Hanoi“ Problems interessiert, sondern auch wie du darauf kommst. Deshalb wäre es sehr wichtig, dass du während der Bearbeitung der Aufgabe „laut denkst“. „Laut denken“ bedeutet, mit lauter Stimme zu überlegen. Bitte sage alles laut, was dir während der Aufgabenbearbeitung durch den Kopf geht, wenn es dir auch unwichtig oder belanglos erscheint und nichts mit der Aufgabe zu tun hat. Es ist

wichtig, dass du deine Gedanken nicht bewertest, bevor du sie laut gesagt hast“ (Anhang A2). Alles, was während der Aufgabenbearbeitung gesprochen wurde, wurde mit einem Tonbandgerät aufgezeichnet, transkribiert und mit Hilfe eines Kategoriensystems kategorisiert (Kategoriendefinitionen Anhang B1a; genaue Beschreibung der Prozedur unter „Auswertung der verbalen Daten“).

Zusätzlich wurde der Ort der Aufmerksamkeit subjektiv mit Skalen von Kanfer und Ackerman (1989) gemessen. Ich übersetzte die Items der Aufmerksamkeitskalen (Skala „Aufgabenbezogene Aufmerksamkeit“, Skala „Leistungsbewertung“, Skala „Positive Selbsteinschätzung“ und Skala „Negative Selbsteinschätzung“; Anhang B2) und verwendete für alle Skalen einheitlich eine 5 Punkte Likert-Skala anstelle der variierenden Abstufungen bei Kanfer und Ackerman (1989). Mittelwerte, Standardabweichungen und Interkorrelation aller erhobenen Maße zur abhängigen Variable „Ort der Aufmerksamkeit“ (verbale Daten und Fragebogendaten) sind am Ende des Methodenkapitels in Tabelle 5 aufgeführt.

Aufgabenbezogene Aufmerksamkeit. Objektiv gemessen habe ich die aufgabenbezogene Aufmerksamkeit über die Anzahl laut geäußerter, aufgabenbezogener Gedanken wie „Soll ich den Klotz hier oder da hinlegen?“. Alle aufgabenbezogenen Kommentare wurden in zwei Kategorien klassifiziert: „Aufgabenbearbeitung“ und „Aufgabenschwierigkeit“. Aufgabenbezogene Gedanken dürften frei von sozialen Erwünschtheitseffekten sein und deswegen immer laut geäußert werden (Kategoriensystem mit Beispielen im Anhang B1a).

Subjektiv gemessen habe ich die aufgabenbezogene Aufmerksamkeit über die Skala „Aufgabenbezogene Aufmerksamkeit“ von Kanfer und Ackerman (1989). Die von mir verwendeten 5 Items lauteten:

„Ich fokussierte meine gesamte Aufmerksamkeit darauf (1 = nie; 5 = die ganze Zeit)

(1) ...wenig Regelverstöße zu begehen. („Fehler“)

(2) ...eine bestimmte Regel zur Problemlösung zu finden. („Regeln“)

(3) ...so schnell wie möglich das Problem zu lösen. („Schnelligkeit“)

(4) ...so wenig Schritte wie möglich zu benötigen, um die Aufgabe zu lösen. („Züge“)

(5) ...das Problem durch Nachdenken zu lösen. („Nachdenken“)

Diese Skala wurde der „Turm von Hanoi“ Aufgabe angepasst (z. Bsp. anstatt „Ich fokussierte meine gesamte Aufmerksamkeit darauf, weniger Fehler zu machen“ formulierte ich „Ich fokussierte meine gesamte Aufmerksamkeit darauf, wenig Regelverstöße zu begehen“). Von den ursprünglich sechs Items konnten zwei Items wegen mangelnder inhaltlicher Übereinstimmung nicht übernommen werden (z. Bsp. „Ich fokussierte meine gesamte Aufmerksamkeit darauf, auf eine Änderung der Windrichtung vorbereitet zu sein“). Ich ergänzte die Skala um das Item „Ich fokussierte meine gesamte Aufmerksamkeit darauf, das Problem durch Nachdenken zu lösen“.

Die Reliabilität der Skala mit diesen 5 Items betrug $\alpha = .33$. Die Skala ist nicht homogen, die Skalenbildung mit diesen 5 Items unzulässig. Nach einer explorativen Faktorenanalyse (Tabelle 2) mit den 5 aufgabenbezogenen Items und den drei unspezifischen Aufgabenzielen (vgl. Abschnitt 4.5.2) ergaben sich drei Faktoren. Auf Faktor 1 luden die Items „Schnelligkeit“, „Zeitziel“ und „AZ Züge“. Auf Faktor 2 luden die Items „Fehler“ und „AZ Fehler“. Auf Faktor 3 luden die Items „Regeln“, „Züge“ und „Nachdenken“. Aus Faktor 1 bildete ich die Skala „**Geschwindigkeit**“ ($\alpha = .76$), aus Faktor 2 die Skala „**Fehlervermeidung**“ ($\alpha = .82$) und aus Faktor 3 die Skala „**Nachdenken**“ ($\alpha = .66$). Aufgrund der mittelmäßigen internen Konsistenz der Skala „Nachdenken“ wurde das Item „Züge“ eliminiert. Cronbachs Alpha nach Löschung dieses Items betrug .80. Die drei Skalen „Geschwindigkeit“, „Fehlervermeidung“ und „Nachdenken“ gingen als Subskalen der aufgabenbezogenen Aufmerksamkeit in die Berechnung ein.

Tabelle 2: Faktorenanalyse mit den Ladungen der aufgabenbezogenen Aufmerksamkeitsitems und den unspezifischen Aufgabenzielen (AZ) auf drei Faktoren

Items	Faktor 1	Faktor 2	Faktor 3
1 Fehler	.04	.89	-.20
2 Regeln	.02	-.27	.81
3 Schnelligkeit	.83	.04	-.05
4 Züge	.28	.48	.61
5 Nachdenken	-.07	-.16	.87
6 Zeitziel	.88	.05	-.08
7 AZ Züge	.71	.36	.28
8 AZ Fehler	.17	.85	-.13

Anmerkungen: Extraktion mittels Hauptkomponentenanalyse
 Rotation: Varimax mit Kaiser-Normalisierung
 Aufgabenbezogene Aufmerksamkeit Items 1-5
 Aufgabenziele (AZ) Items 6-8

Selbstbezogene Aufmerksamkeit. Die objektive Messung der selbstbezogenen Aufmerksamkeit erfolgte über die Anzahl laut geäußerter, selbstbezogener Gedanken wie „Wieso krieg’ ich das nicht hin?“ oder „Das sollte ich doch hinkriegen“. Alle selbstbezogenen Kommentare wurden in zwei Kategorien klassifiziert: „Positive Selbstbewertung“ und „Negative Selbstbewertung“ (Kategoriensystem Anhang B1a, genaue Beschreibung unter „Auswertung der verbalen Daten“).

Subjektiv gemessen habe ich die Aufmerksamkeit auf das Selbst über drei an die Erfordernisse der Aufgabe angepasste Skalen nach Kanfer und Ackerman (1989): „Aufmerksamkeit auf die Leistungsbewertung“, „Positive Selbsteinschätzung“ und „Negative Selbsteinschätzung“ (alle Aufmerksamkeitssskalen mit Codes im Anhang B2). Die interne Konsistenz der drei Items („Positiver Vergleich“, „Negativer Vergleich“, „Vergleich“) der Skala „**Leistungsbewertung**“ betrug .92 (Cronbachs Alpha). Man kann also von einer homogenen Skala ausgehen. Bei der Interpretation dieser Skala als selbstbezogene Aufmerksamkeit ist zu beachten, dass je höher der Wert angegeben wurde, desto mehr versuchten die Versuchspersonen über den Ver-

gleich mit anderen ihre Leistung zu bewerten. Das bedeutet, dass je unsicherer die Versuchsperson in ihrer Selbsteinschätzung ist, desto mehr fragt sie sich, wie die anderen bei dieser Aufgabe wohl abschneiden oder abgeschnitten haben. Hohe Werte bei der Skala „Leistungsbewertung“ bedeutet große Unsicherheit in der Selbsteinschätzung und damit hohe selbstbezogene Aufmerksamkeit.

Nicht homogen ist die Skala „**Positive Selbsteinschätzung**“ ($\alpha = .30$). Nach Eliminierung des Items „Spaß“ ergab sich eine gerade noch akzeptierbare interne Konsistenz von $\alpha = .64$ mit den Items „Verbesserung“, „Positive Selbsteinschätzung“ und „Selbstbeurteilung“. Die interne Konsistenz der zwei Items der Skala „**Negative Selbsteinschätzung**“ betrug $\alpha = .84$ (Item „Selbstkritik“ und Item „Unzufriedenheit“). Die Skala ist somit homogen. Zur Überprüfung der Skalenbildung wurde eine Faktorenanalyse über alle selbstbezogenen Aufmerksamkeitsitems gerechnet. Die Faktorladungen sind in Tabelle 3 zu sehen. Die Faktorenanalyse bestätigte die Skalenbildung.

Tabelle 3: Faktorladungen aller selbstbezogenen Aufmerksamkeitsitems

Items	Faktor 1 (Ladung)	Faktor 2 (Ladung)	Faktor 3 (Ladung)
1 Positiver Vergleich	.91	.07	.19
2 Negativer Vergleich	.95	.11	.15
3 Vergleich	.89	.22	.15
4 Verbesserung	.25	.20	.76
5 Positive Selbsteinschätzung	.16	.09	.79
6 Spaß	-.20	-.55	.31
7 Selbstbeurteilung	.28	.74	.24
8 Selbstkritik	-.02	.88	.15
9 Unzufriedenheit	.08	.83	.22

Anmerkung: Extraktion mittels Hauptkomponentenanalyse
 Rotation: Varimax mit Kaiser-Normalisierung
 Skala Leistungsverbesserung Items 1-3,
 Skala Positive Selbsteinschätzung Items 4-7
 Skala Negative Selbsteinschätzung Item 8 und 9

Die Korrelation der drei Skalen (Tabelle 5) zeigte, dass die Skala „Leistungsbewertung“ unabhängig von der Skala „Negative Selbsteinschätzung“ ist ($r = .22$, n.s.) und abhängig von der Skala „Positive Selbsteinschätzung“ ($r = .47$, $p < .01$). Die Skala „Negative Selbsteinschätzung“ korrelierte hochsignifikant positiv mit der Skala „Positive Selbsteinschätzung“ ($r = .48$, $p < .01$). Der relativ hohe positive Zusammenhang der drei Skalen zeigt, dass alle drei Skalen Subskalen der selbstbezogenen Aufmerksamkeit darstellen.

Sonstige Aufmerksamkeit. Die sonstige Aufmerksamkeit wurde objektiv erfasst über die Anzahl laut geäußerter Gedanken, die sich nicht direkt auf die Aufgabe und nicht auf die Person beziehen (Kategorie „Sonstige Gedanken“). Ein Beispiel dafür wäre der Satz „Was gibt es heute in der Mensa zum Essen?“ (Anhang B1).

4.5.4 Freude und Aktivierung

Alle Freude- und Aktivierungsitems sind in den Fragebögen im Anhang enthalten (Anhang A4a und A4b). Die Stimmung „**Freude**“ wurde über Stimmungitems nach Mano (1991) abgefragt. Diese wurde mit vier Items über eine 5 Punkte Likert-Skala vor und nach der Bearbeitung des „Turm von Hanoi“ gemessen. Die 4 abgefragten Items waren „Zufriedenheit“, „Vergnügtheit“, „glücklich sein“ und „in guter Stimmung sein“. Cronbachs Alpha betrug .75 vorher und .85 nachher, die Stimmung Freude wird also von einer homogenen Skala erfasst. „**Aktivierung**“ wurde unabhängig von der Freude skala mit drei Items über eine 5 Punkte Likert-Skala vor und nach der Bearbeitung der „Turm von Hanoi“ Aufgabe von Mano (1991) abgefragt. Die drei Items waren „Erstauntheit“, „Überraschtheit“ und „Aktiviertheit“. Aktivierung und gleichzeitige „Vergnügtheit“ wurde nach Mano (1991) über die Skala „Freudige Erregung“ abgefragt. Die drei Items lauteten „freudig erregt sein“, „aktiv sein“, und „aufgeregt sein“. Die 3 Items zur „Aktivierung“ und die 3 Items zur „Freudigen Erregung“ gaben zusammen einen ausreichenden Indikator für die „Aktivierung“ der Versuchspersonen. Cronbachs Alpha aller 6 Items betrug vorher .69 und nachher .68.

Tabelle 4 zeigt, dass die Stimmung „Freude“ unabhängig von der „Aktivierung“ erfasst wurde: Interkorrelation der Freudeskala vorher und nachher mit der neugebildeten Aktivierungsskala vorher und nachher ergaben Nullkorrelationen.

Tabelle 4: Korrelationen der Aktivierungs- und Freudeskalen

Messung	1	2	3
1 Freude vorher			
2 Freude nachher	.45**		
3 Aktivierung vorher	.04	.01	
4 Aktivierung nachher	.01	.14	.50**

Anmerkung: ** signifikant auf dem .01 Niveau (zweiseitig)

4.6 Erfassung der aufgabenspezifischen Selbstwirksamkeit

Die aufgabenspezifische Selbstwirksamkeit wurde vor und nach der Durchführung über eine 5 Punkte Likert-Skala (1 bedeutete „trifft gar nicht zu“; 5 bedeutete „trifft vollständig zu“) mit 4 Items erhoben (Thompson & Perlini, 1998). Maurer und Pierce (1998) überprüften sowohl die prädiktive Validität als auch die Reliabilität von Likert-Skalen zur Messung von Selbstwirksamkeit. Ergebnis ist, dass Likert-Skalen eine akzeptable Alternative zu anderen Messmethoden der Selbstwirksamkeit darstellen. Die vier Items – von mir übersetzt – lauteten (siehe auch Fragebögen im Anhang A4a und A4b):

- (1) *„Ich sehe (anticipate) überhaupt keine Probleme, diese Aufgabe zu bewältigen (perform)“.* („Bewältigung“)
- (2) *„Ich glaube (believe), dass meine Leistung bei dieser Aufgabe unter dem Durchschnitt sein wird“.* („Durchschnitt“)
- (3) *„Ich habe vertrauen (confident) in meine Fähigkeiten, diese Aufgabe gut zu machen“.* („Vertrauen“)
- (4) *„Ich glaube (believe), dass ich diese Aufgabe besser bewältige als die meisten anderen Personen“.* („Besser sein“)

Vor der Berechnung der internen Konsistenz wurde Item (2) umgepolt. Die interne Konsistenz der 4 Items betrug bei der ersten Messung .70 (Cronbachs Alpha) und bei der zweiten Messung .75. Man kann also von einer einheitlichen Skala zur Messung von aufgabenspezifischer Selbstwirksamkeit ausgehen.

4.7 Auswertung der Daten

Es wurden zwei Arten von Daten erhoben um die abhängige Variable „Ort der Aufmerksamkeit“ zu messen: Verbale Daten (Denkprotokolle) und Fragebogendaten. Die Mittelwerte, Standardabweichungen, Interkorrelation und Beurteilerübereinstimmungen sind in Tabelle 5 am Ende des Methodenkapitels aufgeführt.

4.7.1 Auswertung der verbalen Daten

64 von 67 Protokollen waren auswertbar, 3 Protokolle bestanden aus weniger als 9 Segmenten und wurden ausgeschlossen. Die verbalen Daten des Denkprotokolls (Beispielprotokoll im Anhang B1b) wurden nach der Transkription in Segmente unterteilt. Jeder Satz (vollständig oder unvollständig) wurde als eine Segmenteinheit betrachtet. Inhaltlich nicht interpretierbare Segmenteinheiten wurden nicht kategorisiert (zum Beispiel „Das ist ...“; siehe Kategorisierung des Beispielprotokolls Anhang B1c). Im Durchschnitt hatte ein Denkprotokoll 48 Segmenteinheiten ($SD = 26$). Die Männer ($n = 34$) hatten im Schnitt 60 Segmente ($SD = 22$), die Frauen ($n = 30$) 34 Segmente ($SD = 23$).

Jede Segmenteinheit wurde mit einem Kategoriensystem (siehe Anhang B1a) von zwei unabhängigen Beurteilern systematisch kategorisiert. Das Kategoriensystem enthält drei große Kategorien: **aufgabenbezogene Gedanken**, **selbstbezogene Gedanken** und **sonstige Gedanken**. Diese Kategorien wurden von Blankstein, Toner und Flett (1989) zur Kategorisierung der Denkprotokolle von Studierende mit Prüfungsangst erfolgreich angewandt. Anders als bei Blankstein et al. (1989) wurden

aufgabenbezogene Gedanken von mir über zwei Unterkategorien operationalisiert: Aussagen, die die Aufgabenbearbeitung potentiell voranbringen („Den blauen Becher stell ich jetzt hierhin“, Kategorie „*Aufgabenbearbeitung*“) und Kommentare zur Aufgabenschwierigkeit („Das ist aber schwierig“, Kategorie „*Aufgabenschwierigkeit*“). Ähnlich wie bei Blankstein et al. (1989) wurden zwei der drei Unterkategorien der Kategorie **Selbstbezogene Gedanken** als Polaritätsdimension (vgl. Cacioppo & Petty, 1981) konstruiert: „*Positive Selbstbewertung*“ („Wie sehr habe ich mich verbessert?“) versus „*Negative Selbstbewertung*“ („Wie schlecht mache ich meine Aufgabe?“). Zusätzlich verwendete ich eine dritte Unterkategorie „Eigene Leistungsbewertung“, um die Daten des verbalen Denkprotokolls analog zu den Aufmerksamkeitsskalen (Anhang B2) auszuwerten. Wegen mangelnder Aussagen, die in diese Kategorie fielen (vom Erstbeurteiler eine Aussage, vom Zweitbeurteiler keine), wurde diese Kategorie aufgehoben und die strittige Aussage unter „*Negative Selbstbewertung*“ kategorisiert. Unter **sonstige Gedanken** fielen alle Kommentare, die nicht in die schon genannten Kategorien passten, zum Beispiel „Schönes Wetter heute!“.

Zur Reliabilitätsbestimmung wurden alle Denkprotokolle von zwei unabhängigen Beurteilern kategorisiert. Auf der Segmentebene wurde jede Kategorisierung des ersten Beurteilers mit der Kategorisierung des zweiten Beurteilers verglichen. Die Übereinstimmung der beiden Beurteiler über alle Segmente hinweg betrug 91 Prozent. Das bedeutet, dass bei 91 Prozent aller Segmente von beiden Beurteilern die gleiche Kategorie gewählt wurde. Aufgrund des Problems der Zufallsübereinstimmung (vor allem bei wenig Kategorien, Bortz & Döring, 2002) berechnete ich Cohen's Kappa als Maß für die wahre Übereinstimmung der beiden Beurteiler. Die erreichte wahre Übereinstimmung von $\kappa = .67$ kann als zufriedenstellend betrachtet werden (Landis & Koch, 1977). Auf der Protokollebene berechnete ich die Korrelationen zwischen den beiden Beurteilern aus den prozentualen Anteilen der Aussagen, wie sie sich pro Person auf die Kategorien verteilten. Dazu mussten in einem ersten Schritt die Häufigkeiten der Segmenteinheiten pro Kategorie und pro Person bestimmt werden. In einem zweiten Schritt wurde der prozentuale Anteil der Segmenteinheiten innerhalb einer Kategorie pro Versuchsperson berechnet. Die Gesamtzahl der Aussagen (Segmenteinheiten) einer Versuchsperson war dabei jeweils 100 Prozent. Mit diesem Vorgehen

wurden die prozentualen Anteile der Segmenteinheiten innerhalb einer Kategorie pro Versuchsperson vor der Manipulation, nach der Manipulation und über die gesamte Zeitspanne von 6 Minuten berechnet.

Das Spektrum der Korrelationen reichte von $r = .57$ (Aufgabenschwierigkeit) bis zu $r = .88$ (Aufgabenbearbeitung). Der Mittelwert aller fünf Korrelationen betrug $r = .77$. Alle Beurteilerübereinstimmungen sind in Tabelle 5 am Ende des Methodenkapitels zu finden. Alle Analysen basieren auf den Kategorisierungen des Beurteilers mit der größeren Erfahrung.

4.7.2 Auswertung der Fragebogendaten

Vor der Skalenbildung wurde jeweils bei allen theoretisch zugehörigen Items die interne Konsistenz geprüft (Cronbachs Alpha). Wenn notwendig (deutlich größeres α) wurden die vom Programm SPSS 10.0 vorgeschlagenen Itemselektionen soweit vorgenommen, das dadurch keine inhaltlichen Nachteile bei der Skalierung entstanden.

Für alle gruppenvergleichenden Analysen wurde die Varianzanalyse gewählt, obwohl die Voraussetzungen (Normalverteilung, Homogenität der Varianzen) nicht immer erfüllt waren. Nach Bortz (1999) verlieren die Voraussetzungen der Varianzanalyse mit zunehmender Stichprobengröße an Bedeutung. Bortz (2002) empfiehlt bei Verletzung der Normalverteilungsannahme mit größeren Stichproben zu rechnen (ab ca. 30). Dann kann auf die Normalverteilungsvoraussetzung verzichtet werden. Bei Verletzung der Homogenität der Varianzen zwischen den Gruppen (heterogene Varianzen) wird das Ergebnis nur unerheblich verändert, wenn die Gruppen gleich groß sind (Glass, Peckham und Sanders, 1972; zitiert in Bortz, 2002). Bei zu kleinen (ca. 10) und verschiedenen großen Gruppen sollte ein verteilungsfreies Verfahren anstelle der Varianzanalyse verwendet werden (Kruskal-Wallis-Test, Bortz, 1999). Um die Ergebnisse der einfaktoriellen Varianzanalysen abzusichern, wurden zusätzlich die Ergebnisse des Kruskal-Wallis-Tests vergleichend in Tabellenform berichtet, obwohl die Stichprobe in meinem Experiment größer als 30 war und die Gruppen meistens gleich groß waren.

Die Überprüfung der Moderatorhypothese 9 erfolgte zuerst über einen Vergleich der Korrelationen zwischen der Bedingung und dem Ort der Aufmerksamkeit (vor und nach der Manipulation) innerhalb niedriger und hoher aufgabenspezifischer Selbstwirksamkeit vorher (vgl. Bortz, 1999, S.212). Ich verwendete die Korrelationen, um mögliche Artefakte zu identifizieren, beispielsweise signifikante Korrelationen zwischen Ort der Aufmerksamkeit und der Bedingung *vor der Manipulation*. Dann wurde eine zweifaktorielle Varianzanalyse mit den Faktoren Bedingung und aufgabenspezifischer Selbstwirksamkeit vorher bei den identifizierten Variablen gerechnet um potentielle Interaktionseffekte auf die abhängige Variable zu überprüfen.

Bei den „Weiterführenden Analysen“ wurden teilweise mit Hilfe der univariaten Varianzanalyse Dreifachinteraktionen gerechnet. Die Gruppengröße war dabei teilweise knapp unter 10 Personen und nicht immer gleich groß. Diese Ergebnisse wurden unter Vorbehalt diskutiert und als Anregung für die weitere Forschung verstanden. Die Vorhersage der Leistung wurde über eine schrittweise Regressionsanalyse berechnet, bei der alle in Frage kommenden Variablen in die Regressionsgleichung aufgenommen wurden und die potentesten Prädiktoren identifiziert wurden. Alle Korrelationen sind Produkt-Moment-Korrelationen nach Pearson. Drei Signifikanzniveaus wurden allen Berechnungen zugrundegelegt: tendenziell signifikant ($\alpha = .10$), signifikant ($\alpha = .05$) und hochsignifikant ($\alpha = .01$).

Messung	M	SD	1	2	3	4	5	6	7	8	9	10
Denkprotokoll												
1 Aufgabenbearbeitung ¹	77.4	18.1	(.88)									
2 Aufgabenschwierigkeit ¹	2.8	4.8	-.43**	(.57)								
3 Sonstige Gedanken ¹	11.4	12.4	-.64**	.05	(.75)							
4 Positive Selbstbewertung ¹	1.6	3.2	-.31*	.43**	.03	(.81)						
5 Negative Selbstbewertung ¹	6.8	7.0	-.56**	.42**	.26*	.21^a	(.84)					
Fragebogen												
6 Nachdenken	3.46	1.09	.16	.06	-.09	-.01	-.24*					
7 Fehlervermeidung	3.11	0.84	.12	-.21^a	-.09	-.01	-.14	-.31*				
8 Geschwindigkeit	3.11	0.84	.12	-.22^a	-.08	-.14	-.11	-.03	.27*			
9 Leistungsbewertung	2.55	1.18	-.22^a	.28*	.15	.15	.15	-.31*	-.03	.08		
10 Positive Selbsteinschätzung	2.82	0.83	-.08	.06	-.08	-.02	.11	-.10	.07	.45**	.47**	
11 Negative Selbsteinschätzung	2.52	1.05	-.09	-.00	-.01	-.09	.03	-.06	.05	.37**	.22^a	.48**
Anmerkungen:	^a signifikant auf dem Niveau von .10 (zweiseitig)						* signifikant auf dem Niveau von .05 (zweiseitig)					
	** signifikant auf dem Niveau von .01 (zweiseitig)						Interrater Übereinstimmung in Klammern					
	¹ in Prozent des verbalen Denkprotokolls (n = 64)											

Tabelle 5: Mittelwerte, Standardabweichung und Interkorrelationen der abhängigen Variablen Ort der Aufmerksamkeit

5 Ergebnisse

5.1 Übersicht

In diesem Kapitel werde ich innerhalb des Abschnitts „Manipulationchecks“ kontrollieren, ob erstens die Ausschlusskriterien eingehalten wurden, zweitens die Manipulation der unabhängigen Variable funktioniert hat und drittens die Manipulation eine angemessene Qualität besitzt. Dann werde ich den Ausschluss bestimmter Datensätze aufgrund der Manipulationchecks erklären. Die so vorbereiteten Datensätze werde ich dann für die anschließenden Hypothesentests. Abschließend überprüfte ich innerhalb der „Weiterführenden Analysen“ erstens, ob sich eventuell die Glaubwürdigkeit des Lobs auf die abhängigen Variablen auswirkte, zweitens versuchte ich Zusammenhänge zu entdecken, die für die Forschung interessant sein könnten, drittens testete ich Geschlechtsunterschiede, die Hypothesen unter Berücksichtigung des Geschlechts und geschlechtsspezifische Zusammenhänge. Zum Schluss identifizierte ich regressionsanalytisch die besten Prädiktoren zur Vorhersage der Leistung.

5.2 Manipulationchecks

Alle Manipulationchecks wurden mit einer 5 Punkte Likert-Skala durchgeführt. 1 bedeutete „trifft überhaupt nicht zu“ und 5 bedeutete „trifft voll zu“.

5.2.1 Bekanntheitsgrad der Aufgabe

Zuerst wurde bei beiden Bedingungen (Fragebogenversionen im Anhang A4a und A4b) das Ausschlusskriterium für die Teilnahme am Experiment erneut abgefragt, nachdem bei der Rekrutierung schon darauf hingewiesen wurde:

- (a) *„Ich kenne die „Turm von Hanoi“ Aufgabe.“*
- (b) *„Ich habe diese Aufgabe schon ziemlich oft gemacht.“*

Die beiden Items korrelierten hoch signifikant mit $r = .78$ ($p < .01$). Versuchspersonen, die die Aufgabe kannten und trotzdem am Experiment teilnahmen, wurden von dem Testen bestimmter Hypothesen ausgeschlossen (siehe 5.2.3).

5.2.2 Manipulation der unabhängigen Variable

Als nächstes wurde mit zwei Items geprüft, ob die Manipulation der unabhängigen Variablen funktioniert hat:

- (1) *„Ich wurde von meinem Versuchsleiter gelobt“.*
- (2) *„Ich fühlte mich von meinem Versuchsleiter gelobt“.*

Beide Items korrelierten mit $r = .86$ ($p < .001$) hoch signifikant. Vergleicht man die Bedingungen hinsichtlich der zwei Items (Kruskal Wallis Test, Items sind nicht normalverteilt) ergaben sich hochsignifikante Unterschiede. Der Mittelwert in der Experimentalgruppe bei Item 1 betrug 3.7 (SD = 0.9), in der Kontrollgruppe 1.1 (SD = 0.4), $\chi^2 = 51.9$, $p < .001$. Der Mittelwert in der Experimentalgruppe bei Item 2 betrug 3.4 (SD = 0.8), in der Kontrollgruppe 1.2 (SD = 0.6), $\chi^2 = 46.7$, $p < .001$. Die Manipulation hat funktioniert: nur die Versuchspersonen in der Experimentalgruppe gaben an, gelobt worden zu sein; und nur in der Experimentalgruppe fühlten sich die Versuchspersonen gelobt.

5.2.3 Qualität der Manipulation

In der Experimentalbedingung wurde darüber hinaus kontrolliert, wie gut, überzeugend und glaubwürdig die Versuchsleiter auf die Versuchspersonen wirkten. Der Manipulationcheck zur Qualität des Lobs erfolgte mit folgenden Items:

- (3) *„Der Versuchsleiter hat mit seinem Lob übertrieben.“*
- (4) *„Der Versuchsleiter hat mit seinem Lob untertrieben.“*
- (5) *„Ich hatte das Gefühl, der Versuchsleiter meinte sein Lob ernst.“*
- (6) *„Ich hatte den Eindruck, der Versuchsleiter lobte mich nicht wegen meiner guten Leistung.“*

- (7) *„Ich fand das Lob des Versuchsleiters gerechtfertigt.“*
- (8) *„Ich empfand das Lob als glaubwürdig.“*
- (9) *„Ich hatte den Eindruck, der Versuchsleiter / die Versuchsleiterin macht mir was vor, als er / sie mich lobte.“*

Die Qualität der unabhängigen Variablen zeigt sich in den Mittelwerten über alle Versuchspersonen der Experimentalbedingung hinweg, und im Vergleich der beiden Versuchsleiter (Tabelle 6). Bevor eine Aussage über die Qualität der Manipulation getroffen werden kann, müssen die Datensätze entfernt werden, bei denen die Manipulation nicht funktioniert hat (siehe 5.2.4).

Um zu entscheiden, wie überzeugend das Lob für die Versuchspersonen war, berechnete ich die Mittelwerte und Standardabweichungen aller sieben Kontrollitems (korrigierter Datensatz mit 32 Personen, Tabelle 6). Die Versuchsleiter haben weder über- noch untertrieben gelobt (3 und 4) und die Versuchspersonen hatten den Eindruck, dass das Lob ernst gemeint war (5). Das der Versuchsleiter nicht wegen guter Leistung gelobt hätte (6), sondern aus anderen Gründen (eventuell experimentelle Situation) wurde zwischen „trifft wenig zu“ und „trifft mittelmäßig zu“ eingestuft. „Gerechtfertigt“ (7) war das Lob nach Einschätzung der Versuchspersonen nicht immer („trifft mittelmäßig zu“). „Glaubwürdig“ (8) wird das Lob zwischen „trifft mittelmäßig zu“ und „trifft größtenteils zu“ eingeschätzt. Dass der Versuchsleiter den Versuchspersonen etwas vormachte, als er lobte (9), trifft für die meisten eher weniger zu.

Die Berechnung eines Kruskal Wallis Test zeigte, dass sich keines der sieben Kontrollitems – weder vor noch nach dem Ausschluss von Datensätzen - sich signifikant zwischen den beiden Versuchsleitern unterscheidet (Tabelle 6). Über alle sieben Kontrollitems hinweg hatten beide Versuchsleiter den gleichen Einfluss auf die Versuchspersonen. Im Originaldatensatz (Experimentalbedingung $n = 37$) zeigten sich tendenzielle Unterschiede zwischen den Versuchsleitern in der Manipulation hinsichtlich der Items (6) und (9). Nach Ausschluss von fünf Datensätzen, bei denen die Manipulation nicht funktioniert hat (siehe 5.2.4) wurden die Items (4) und (9) tendenziell signifikant (Tabelle 6). Das bedeutet, dass der Versuchsleiter mit seinem Lob tenden-

ziell eher untertrieben (4) hat im Vergleich mit der Versuchsleiterin. Bei Item (9) „Ich hatte den Eindruck, der Versuchsleiter / die Versuchsleiterin macht mir was vor, als er / sie mich lobte“ war die Versuchsleiterin tendenziell „echter“: die meisten Versuchspersonen kreuzten bei diesem Item „trifft wenig zu“ an (bei den Männern eher „trifft mittelmäßig zu“).

Tabelle 6: Vergleich der beiden Versuchsleiter innerhalb der Experimentalbedingung, vor und nach Ausschluss von Datensätzen aufgrund der Manipulationchecks

Kontrollitems	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Korrigierter Datensatz ¹	2.5	1.7	3.6	2.3	2.9	3.5	2.3
M (SD), n = 32	(1.2)	(.8)	(.9)	(1.1)	(1.1)	(.9)	(1.2)
Originaldaten ¹ (n = 37)							
Versuchsleiter 1	2.6	1.5	3.6	2.2	2.8	3.4	2.1
M (SD), n = 18	(1.3)	(.8)	(.9)	(1.2)	(1.2)	(1.0)	(1.3)
Versuchsleiter 2	2.9	1.8	3.3	2.9	2.7	3.2	2.8
M (SD), n = 19	(1.2)	(.8)	(1.2)	(1.3)	(1.0)	(1.3)	(1.3)
χ^2	.48	2.25	.47	2.9^a	.00	.43	2.9^a
Nach Daten- cleaning ¹ (n = 32)							
Versuchsleiter 1	2.5	1.5	3.6	2.1	2.9	3.6	1.9
M (SD), n = 17	(1.3)	(.8)	(.9)	(1.2)	(1.2)	(.8)	(1.1)
Versuchsleiter 2	2.5	1.9	3.6	2.5	2.9	3.5	2.7
M (SD), n = 15	(1.1)	(.7)	(1.0)	(1.1)	(1.1)	(1.1)	(1.2)
χ^2	.02	2.78^a	.04	1.34	.03	.08	3.00^a

Anmerkungen: ^a tendenziell signifikant auf 0.10-Niveau

¹ Daten sind nicht normalverteilt: Kruskal Wallis Test

5 Punkte Likert-Skala: 1 = trifft gar nicht zu, 5 = trifft vollständig zu

5.2.4 Ausschluss von Datensätzen

Vor der Überprüfung der Hypothesen wurden alle Datensätze entfernt, bei denen die Manipulation der unabhängigen Variablen nicht funktioniert hat. Kreuzte eine Versuchsperson auf die Frage „Ich fühlte mich von meinem Versuchsleiter gelobt“ (innerhalb der Experimentalbedingung) die „1“ oder „2“ an (1 = „trifft gar nicht zu“, 2 = „trifft wenig zu“), wurde der Datensatz von den Berechnungen ausgeschlossen. Wenn Item „8“ als „trifft gar nicht zu“ oder Kontrollitem „9“ innerhalb der Experimentalbedingung als „vollständig zutreffend“ angegeben wurde, musste der Datensatz auch von den Berechnungen ausgeschlossen werden. Insgesamt trafen diese Ausschlusskriterien auf fünf Datensätze zu. Diese wurden von den gruppenvergleichenden Analysen ausgeschlossen (Korrigiertes $n = 62$). Für die Überprüfung der Hypothesen 5 und 6 wurde mit dem vollständige Datensatz ($n = 67$) gerechnet, weil die mögliche Veränderung der Freude und Aktivierung der Person durch die Manipulation keine Auswirkung haben sollte auf die Überprüfung des in diesen Hypothesen postulierten Zusammenhangs. Für die Überprüfung der Hypothese 8 wurde auch mit dem vollständige Datensatz ($n = 67$) gerechnet, weil die mögliche Veränderung der aufgabenspezifischen Selbstwirksamkeit der Person durch die Manipulation keine Auswirkung haben sollte auf die Überprüfung des postulierten Zusammenhangs in dieser Hypothese. Für die Berechnung der Hypothesen 10 und 11 mussten von dem reduzierten Datensatz ($n = 62$) weitere drei Datensätze von der Analyse ausgeschlossen werden. In diesen drei Datensätzen gaben die Versuchspersonen an, die „Turm von Hanoi“ Aufgabe zu kennen und schon einmal gemacht zu haben (Item 1 und Item 2 des Manipulationchecks).

5.3 Testen der Hypothesen

5.3.1 Der Einfluss von Lob auf Freude (H1) und Aktivierung (H2)

Die Personen, die gelobt werden, sollten mehr Freude empfinden und aktiver sein als die Personen, die nicht gelobt werden (Hypothese 1 und 2). Die einfaktorielle Va-

riananalyse mit Messwiederholung zeigte, dass es keine Unterschiede zwischen den Gruppen gab hinsichtlich der Aktivierungsskala und der Freudeskala (Tabelle 7). Die Hypothesen 1 und 2 bestätigten sich nicht: Die Experimentalgruppe war – nach subjektiver Einschätzung – weder aktiver noch freute sie sich mehr nach der Manipulation als die Gruppe, die nicht gelobt wurde. Sowohl die Interaktion der Bedingung mit der Veränderung der Freude war nicht signifikant $F(1, 60) = 0.86$, ns; als auch die Interaktion der Bedingung mit der Veränderung der Aktivierung, $F(1,60) = 0.26$, ns.

Tabelle 7: Einfaktorielle Varianzanalyse mit Messwiederholung zur Überprüfung des Einflusses von Lob auf die Veränderung der Freude und der Aktivierung (H1 und H2)

Messung	Lob		Neutral		V	B	I
	M (SD)		M (SD)				
	vorher	nachher	vorher	nachher			
Freude	3.48 (.74)	3.08 (.90)	3.33 (.70)	2.98 (.84)	11.52**	0.53	0.86
Aktivierung	2.73 (.64)	3.01 (.64)	2.81 (.70)	2.89 (.65)	4.60*	0.91	0.26

Anmerkungen: * signifikant auf dem Niveau von .05 (zweiseitig)
 ** signifikant auf dem Niveau von .01 (zweiseitig)
 V= Veränderung, B = Bedingung, I = Interaktion, F (1, 60), n = 62

5.3.2 Der Einfluss von Lob auf den Ort der Aufmerksamkeit (H3)

Für die Überprüfung der Hypothese 3 wurde sowohl eine einfaktorielle Varianzanalyse mit Messwiederholung (Tabelle 8a, Denkprotokoll) als auch eine einfaktorielle Varianzanalyse mit einem Kruskal-Wallis-Test (Tabelle 8b, Fragebogendaten) durchgeführt. Die Ergebnisse der Analysen sind in den Tabellen 8a und 8b gegenübergestellt.

Der erste Teil von Hypothese 3 lautete, dass die Personen, die gelobt werden, dem Selbst mehr Aufmerksamkeit schenken als die Personen, die nicht gelobt werden. Dieser erste Teil von Hypothese 3 wurde bei den Fragebogendaten zum Teil wider-

legt. Die Personen, die gelobt wurden, verglichen ihre Leistung signifikant weniger stark mit der Leistung anderer ($M = 2.09$, $SD = 1.03$) als die Personen, die nicht gelobt wurden ($M = 3.04$, $SD = 1.15$, $F = 11.85$, $p < .01$). Keine Unterschiede gab es bei der Skala „Positive Selbsteinschätzung“ ($F = 0.02$, ns) und „Negative Selbsteinschätzung“ ($F = 0.83$, ns). Nach der Manipulation gab es zwar im Denkprotokoll tendenziell mehr positive Selbstbewertungen (Veränderung: $F = 3.43$, $p < .10$) und hochsignifikant mehr negative Selbstbewertungen ($F = 11.7$, $p < .01$), diese Veränderungen sind aber nicht auf die Bedingung zurückzuführen (positive Selbstbewertung: F (Interaktion) = 0.81 , ns, negative Selbstbewertung: F (Interaktion) = $.00$, ns).

Der zweite Teil der Hypothese 3 lautete, dass die Personen, die gelobt werden, mehr sonstige Gedanken und weniger aufgabenbezogene Gedanken haben als die Personen, die nicht gelobt werden. Dieser Teil der Hypothese wurde nicht bestätigt (siehe Tabellen 8a und 8b). Die Personen, die gelobt wurden, hatten genau die gleiche Veränderung der relativen Häufigkeit aufgabenbezogener Gedanken (Aufgabenbearbeitung: F (Interaktion) = 0.39 , ns; Aufgabenschwierigkeit: F (Interaktion) = 1.6 , ns) und sonstiger Gedanken (F (Interaktion) = 0.15 , ns) wie die Personen, die nicht gelobt wurden. Das gleiche gilt für die Fragebogenmaße „Nachdenken“ ($F = 0.73$, ns), „Fehlervermeidung“ ($F = 0.40$, ns) und „Geschwindigkeit“ ($F = 2.71$, ns): die Gruppen unterschieden sich nicht.

Tabelle 8a: Einfaktorielle Varianzanalyse mit Messwiederholung zur Überprüfung des Einflusses von Lob auf den Ort der Aufmerksamkeit (Denkprotokoll; H3)

Messung Denkprotokoll	Lob M (SD)		Neutral M (SD)		V	B	I
	vorher	nachher	vorher	nachher			
1 Aufgabenbearbeitung	75.8 (28.0)	68.3 (26.0)	77.1 (28.6)	74.3 (26.4)	1.80	0.41	0.39
2 Aufgabenschwierigkeit ¹	0.8 (2.4)	4.2 (7.4)	2.9 (8.5)	3.2 (6.8)	2.16	0.24	1.6
3 Sonstige Gedanken ¹	14.2 (22.8)	11.9 (15.8)	10.7 (17.1)	10.4 (10.5)	0.27	0.48	0.15
4 Positive Selbstbewertung ¹	1.0 (2.2)	2.8 (6.6)	0.3 (1.2)	1.7 (6.2)	3.43^a	1.21	0.81
5 Negative Selbstbewertung ¹	5.1 (7.5)	9.8 (10.3)	2.3 (4.2)	7.1 (11.0)	11.7^{**}	2.52	0.00

Anmerkungen: ^asignifikant auf dem Niveau von .10 (zweiseitig)

^{**}signifikant auf dem Niveau von .01 (zweiseitig)

¹ Nicht normalverteilte Variablen

V = Veränderung, B = Bedingung, I = Interaktion (F (1, 60), n = 62)

Tabelle 8b: Einfaktorielle Varianzanalyse und Kruskal-Wallis-Test zur Überprüfung des Einflusses von Lob auf den Ort der Aufmerksamkeit (Fragebogen; H3)

Messung Fragebogen	M (SD) Lob	M (SD) Neutral	F-Wert	χ^2
6 Nachdenken ²	3.6 (1.2)	3.4 (0.8)	0.73	1.55
7 Fehlervermeidung	3.5 (1.2)	3.3 (1.3)	0.40	2.52
8 Geschwindigkeit	2.9 (0.9)	3.3 (1.1)	2.71	0.29
8 Leistungsbewertung	2.1 (1.0)	3.0 (1.1)	11.85^{**}	9.72^{**}
9 Positive Selbsteinschätzung	2.8 (0.9)	2.8 (0.8)	0.02	0.01
10 Negative Selbsteinschätzung	2.4 (1.0)	2.6 (1.0)	0.83	0.82

Anmerkungen: ^{**}signifikant auf dem Niveau von .01 (zweiseitig)

² Heterogene Varianzen

n = 62

5.3.3 Der Einfluss von Lob auf den persönlichen Aufgabenstandard (H4)

Hypothese 4 behauptet, dass die gelobten Personen eher ihren persönlichen, spezifischen und selbstzielrelevanten Aufgabenstandard verändern als die nicht gelobten Personen. Das war nicht der Fall. Die einfaktorielle Varianzanalyse mit Messwiederholung zeigte keine signifikanten Interaktionen der Bedingung mit den Veränderungen der jeweiligen Aufgabenstandards (Aufgabenstandard Zeit: $F(1, 59) = 0.96$, ns; Aufgabenstandard Züge: $F(1, 59) = 1.74$, ns; Aufgabenstandard Fehler: $F(1, 59) = 0.61$, ns). Die Veränderung der drei Aufgabenstandards (Zeit, Züge, Fehler) war in beiden Gruppen gleich (Tabelle 9).

Tabelle 9: Einfaktorielle Varianzanalyse mit Messwiederholung zur Überprüfung des Einflusses von Lob auf die Veränderung des persönlichen, spezifischen Aufgabenstandards (H4)

Messung ¹	M (SD) Lob		M(SD) Neutral		V	B	I
	vorher	nachher	vorher	nachher			
AS Zeit	6.7 (1.3)	7.3 (1.2)	6.7 (1.2)	7.0 (1.1)	9.45**	0.30	0.96
AS Züge	77.7 (4.6)	81.9 (4.9)	77.3 (5.0)	79.7 (5.6)	21.39**	1.53	1.74
AS Fehler	2.4 (1.2)	3.2 (1.6)	2.5 (1.4)	2.9 (1.8)	9.97**	0.10	0.61

Anmerkungen: ¹ Daten sind nicht normalverteilt
 ** signifikant auf dem Niveau von .01 (zweiseitig)
 n = 61 aufgrund eines fehlenden Wertes
 V = Veränderung, B = Bedingung, I = Interaktion, (F (1, 60), n = 62)

5.3.4 Der Zusammenhang von Freude (H5) und Aktivierung (H6) mit der Art der Gedanken

Hypothese 5 behauptete, dass die Versuchspersonen mit zunehmender Freude mehr sonstige Gedanken äußern würden. Die Skala „Freude vorher“ hing signifikant negativ mit der relativen Häufigkeit sonstiger Gedanken zusammen ($r = -.25$, $p = .04$, zweiseitig getestet). Hypothese 5 wurde somit widerlegt: Mit zunehmender Freude äußerten die Versuchspersonen weniger sonstige Gedanken. **Hypothese 6** behauptete, dass die Versuchspersonen mit zunehmender Aktivierung mehr aufgabenbezogene Gedanken äußern würden. Die Skala „Aktivierung vorher“ korrelierte nicht mit den aufgabenbezogenen Gedanken (Aufgabenbearbeitung: $r = .08$, ns; Aufgabenschwierigkeit: $r = .01$, ns, zweiseitig getestet) und nicht mit den Fragebogenmaßen der aufgabenbezogenen Aufmerksamkeit (Nachdenken: $r = .04$, ns; Fehlervermeidung: $r = -.01$, ns; Geschwindigkeit: $r = -.02$, ns). Hypothese 6 wurde nicht bestätigt: Es bestand kein Zusammenhang zwischen der Aktivierung vor der Bearbeitung der Aufgabe und der aufgabenbezogenen Aufmerksamkeit.

5.3.5 Der Einfluss von Lob auf die Veränderung des persönlichen Aufgabenstandards bei Personen mit relativ hoher aufgabenspezifischer Selbstwirksamkeit (H7)

Hypothese 7 lautete, dass Personen, die gelobt werden und über eine hohe aufgabenspezifische Selbstwirksamkeit verfügen, ihren Aufgabenstandard eher nach oben setzen als Personen, die über hohe aufgabenspezifische Selbstwirksamkeit verfügen und nicht gelobt werden. Die Formulierung „nach oben setzen“ bedeutete, sich einen noch schwereren, persönlichen Aufgabenstandard zu setzen (vgl. Abschnitt 3.4.1). Die einfaktorielle Varianzanalyse mit Messwiederholung ergab, dass die Personen, die gelobt wurden und über eine relativ hohe, aufgabenspezifische Selbstwirksamkeit verfügten sich *tendenziell* mit höheren *Zugzahlen* zufrieden gaben als die Personen, die nicht gelobt wurden und über relativ hohe, aufgabenspezifische Selbstwirksamkeit verfügten. Die Veränderung des AS Züge vorher zu nachher betrug in der Experimentalbe-

dingung 4.6 Züge, und in der Kontrollbedingung 1.4 Züge, F (Interaktion) = 3.27, $p < .10$. „Relativ hoch“ bedeutete, dass die aufgabenspezifische Selbstwirksamkeit vorher ≥ 3.0 sein musste. Bei $n = 28$ Versuchspersonen war das der Fall. Die Veränderung der Aufgabenstandards „Zeit“ und „Fehler“ war bei beiden Gruppen gleich (AS Zeit: $F = 2.10$, ns; AS Fehler: $F = 0.19$, ns).

Hypothese 7 wurde hinsichtlich der Aufgabenstandards Zeit und Fehler nicht bestätigt. Hinsichtlich des Aufgabenstandards Züge wurde Hypothese 7 widerlegt. Bei dem Aufgabenstandard Züge förderte das Lob bei Personen mit relativ hoher aufgabenspezifischer Selbstwirksamkeit vorher tendenziell eher das Setzen von einem leichteren Aufgabenstandard nach der Bearbeitung der Aufgabe.

Tabelle 10: Einfaktorielle Varianzanalyse mit Messwiederholung zur Überprüfung des Einflusses von Lob auf die Veränderung des persönlichen, spezifischen Aufgabenstandards bei Personen mit relativ hoher aufgabenspezifischer Selbstwirksamkeit vorher (H7)

Messung ¹	Lob M (SD)		Neutral M (SD)		V	B	I
	vorher	nachher	vorher	nachher			
AS Zeit	6.6 (1.3)	7.2 (1.2)	7.1 (1.2)	7.2 (1.0)	3.24^a	0.37	2.10
AS Züge	77.3 (4.4)	81.9 (4.3)	78.3 (5.9)	79.7 (6.1)	10.74^{**}	0.12	3.27^a
AS Fehler	2.5 (1.3)	3.1 (1.7)	2.1 (1.2)	2.3 (1.6)	1.63	1.80	0.19

Anmerkungen: ¹ Daten sind nicht normalverteilt

^a signifikant auf dem Niveau von .10 (zweiseitig)

^{**} signifikant auf dem Niveau von .01 (zweiseitig)

$n = 28$: Experimentalgruppe 13, Kontrollgruppe 15

V = Veränderung, B = Bedingung, I = Interaktion, $F(1, 26)$

5.3.6 Der Zusammenhang von aufgabenspezifischer Selbstwirksamkeit mit der Art der Gedanken (H8)

Hypothese 8 lautete, dass je niedriger die aufgabenspezifische Selbstwirksamkeit der Versuchspersonen ist, desto mehr selbstbezogene und sonstige Gedanken und desto weniger aufgabenbezogene Gedanken haben sie. Die Ergebnisse der Korrelationsanalyse sind in Tabelle 11 zu finden. Die Skala „Negative Selbsteinschätzung“ korrelierte signifikant negativ mit der aufgabenspezifischen Selbstwirksamkeit danach ($r = -.26$, $p < .05$). Alle anderen Korrelationen waren nicht signifikant. Hypothese 8 bestätigte sich nicht.

Tabelle 11: Korrelationsanalyse zur Überprüfung des Zusammenhangs von aufgabenspezifischer Selbstwirksamkeit (vorher und nachher) mit dem Ort der Aufmerksamkeit (H8)

Messung	r (SW vorher)	r (SW nachher)
<u>Denkprotokoll</u>		
1 Aufgabenbearbeitung	.12	.11
2 Aufgabenschwierigkeit	-.19	-.03
3 Sonstige Gedanken	-.04	-.05
4 Positive Selbstbewertung	-.10	.15
5 Negative Selbstbewertung	-.18	-.17
<u>Fragebogen</u>		
6 Nachdenken	.01	.20
7 Fehlervermeidung	.16	-.03
8 Geschwindigkeit	.14	.14
9 Leistungsbewertung	-.01	.05
10 Positive Selbsteinschätzung	.16	.20
11 Negative Selbsteinschätzung	.12	-.26*

Anmerkungen: *signifikant auf dem Niveau von .05 (zweiseitig)
n = 67

Da die Korrelationsanalyse keine Ergebnisse lieferte, wurden zwei Extremgruppen gebildet. Eine Extremgruppe wurde gebildet mit sehr niedriger aufgabenspezifischen Selbstwirksamkeit nachher (die 20 niedrigsten), und eine Gruppe mit sehr hoher aufgabenspezifischer Selbstwirksamkeit nachher (die 20 höchsten). Die Ergebnisse der Extremgruppenvergleiche in Bezug auf den Ort der Aufmerksamkeit sind in Tabelle 12 zu sehen. Es ergaben sich Unterschiede hinsichtlich der Skala „Negative Selbsteinschätzung“ (Fragebogen) und der Kategorie „Negative Selbstbewertung“ (Denkprotokoll). Personen mit *niedriger* aufgabenspezifischer Selbstwirksamkeit nachher hatten tendenziell mehr negative selbstbezogene Gedanken während der Bearbeitung der Aufgabe und gaben signifikant höhere Werte in der Skala „Negative Selbsteinschätzung“ an als Personen mit *hoher* aufgabenspezifischer Selbstwirksamkeit nachher. Hypothese 8 wurde hinsichtlich der negativen selbstbezogenen Gedanken bestätigt (Denkprotokoll und Fragebogen). Hinsichtlich der aufgabenbezogenen und sonstigen Gedanken bestätigte sich Hypothese 8 nicht.

Tabelle 12: Unterschiede zwischen sehr hoher und sehr niedriger aufgabenspezifischer Selbstwirksamkeit nachher hinsichtlich der Variablen des Orts der Aufmerksamkeit (H8)

Messung	M (SD) SW niedrig	M (SD) SW hoch	F-Wert	χ^2
<u>Denkprotokoll</u>				
1 Aufgabenbearbeitung	80.9 (17.7)	86.1 (10.3)	1.3	0.6
2 Aufgabenschwierigkeit	1.3 (2.8)	1.6 (2.8)	0.11	0.66
3 Sonstige Gedanken	9.7 (10.8)	7.1 (8.7)	0.69	0.41
4 Positive Selbstbewertung	0.9 (1.9)	1.8 (2.2)	1.84	2.49
5 Negative Selbstbewertung	7.1 (7.2)	3.3 (3.7)	4.44*	3.38^a
<u>Fragebogen</u>				
6 Nachdenken ²	3.4 (1.4)	3.8 (0.8)	1.25	0.59
7 Fehlervermeidung	3.6 (1.5)	3.5 (1.1)	0.02	0.22
8 Geschwindigkeit	3.1 (1.0)	3.4 (1.1)	1.04	1.07
9 Leistungsbewertung	2.4 (1.2)	2.3 (1.2)	0.05	0.05
10 Pos. Selbsteinschätzung	2.8 (0.8)	2.9 (1.0)	0.14	0.01
11 Neg. Selbsteinschätzung	2.8 (1.2)	2.1 (1.0)	4.02*	3.58^a

Anmerkungen: *signifikant auf dem Niveau von .05 (zweiseitig)

^a signifikant auf dem Niveau von .10 (zweiseitig)

² Heterogene Varianzen

n = 40

5.3.7 Aufgabenspezifische Selbstwirksamkeit als Moderator zwischen Lob und der Art der Gedanken (H9)

Hypothese 9 postuliert, dass Personen mit *niedriger* aufgabenspezifischer Selbstwirksamkeit nach Lob mehr selbstbezogene und sonstige Gedanken und weniger aufgabenbezogene Gedanken haben als ohne Lob. Personen mit *hoher* aufgabenspezifischer Selbstwirksamkeit dagegen sollten nach Lob weniger selbstbezogene und sonstige Gedanken und mehr aufgabenbezogene Gedanken haben als ohne Lob.

Für die Überprüfung der Moderatorhypothese 9 wurden die Korrelationen zwischen der unabhängigen Variablen „Bedingung“ und der abhängigen Variablen „Ort

der Aufmerksamkeit“ innerhalb der Gruppe mit niedriger aufgabenspezifische Selbstwirksamkeit vorher (SW vorher < 3.0 , $n = 34$) und der Gruppe mit hoher aufgabenspezifische Selbstwirksamkeit vorher (SW vorher ≥ 3.0 , $n = 28$) berechnet (Tabelle 12). Bei der Variable „Positive selbstbezogene Gedanken“ gab es Hinweise in Richtung Moderation des Lobs durch die aufgabenspezifische Selbstwirksamkeit vorher: Es gab zwar weder einen Haupteffekt der Bedingung ($F = .89$, ns) noch der aufgabenspezifischen Selbstwirksamkeit vorher ($F = .10$, ns), dafür einen signifikanten Interaktionseffekt Selbstwirksamkeit mal Bedingung ($F = 5.17$; $p < .05$). Personen mit niedriger aufgabenspezifischer Selbstwirksamkeit vorher haben nach Lob *signifikant weniger* positive selbstbezogene Gedanken (nach der Manipulation: $M = 1.11$ Prozent aller Gedanken, $SD = 2.46$) als Personen, die nicht gelobt wurden ($M = 3.2$ Prozent aller Gedanken, $SD = 8.6$). Personen mit hoher aufgabenspezifischer Selbstwirksamkeit vorher haben nach Lob *signifikant mehr* positive selbstbezogene Gedanken nach der Manipulation ($M = 5.3$, $SD = 9.7$) als Personen die nicht gelobt wurden ($M = 0.1$, $SD = 0.4$). Personen mit hoher aufgabenspezifischer Selbstwirksamkeit vorher haben nach Lob signifikant mehr positive selbstbezogene Gedanken als Personen mit niedriger aufgabenspezifischer Selbstwirksamkeit. Abbildung 3 soll den Zusammenhang verdeutlichen. Die Hypothese 9 wurde nicht bestätigt. Hinsichtlich der positiven selbstbezogenen Gedanken gab es Hinweise auf einen moderierenden Einfluss der aufgabenbezogenen Selbstwirksamkeit. Dieser Einfluss geht aber in die der Hypothese entgegengesetzte Richtung.

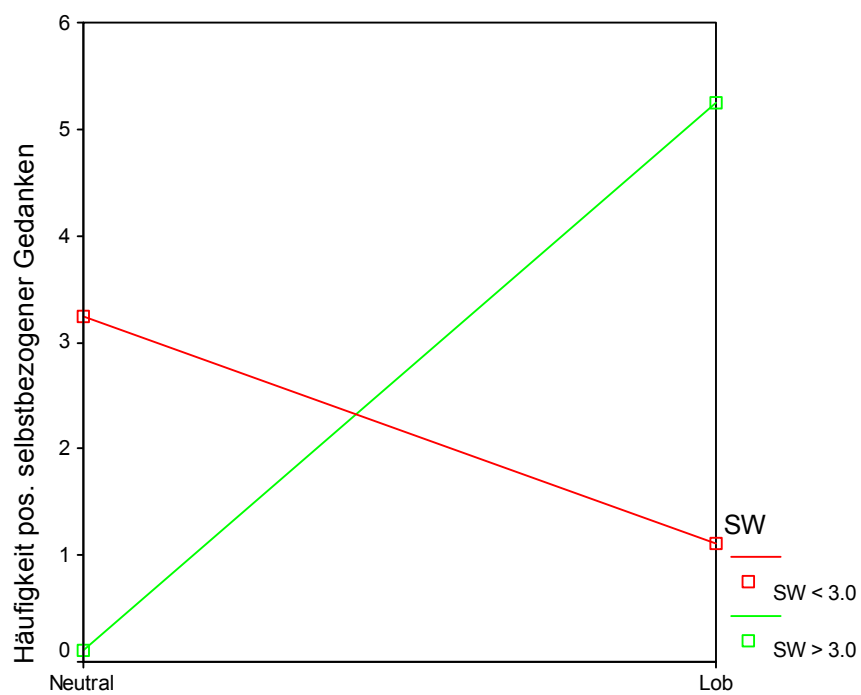


Abbildung 3: Die Auswirkung der Bedingung auf die Häufigkeit von positiven, selbstbezogenen Gedanken (in Prozent) nach der Manipulation in Abhängigkeit von aufgabenspezifischer Selbstwirksamkeit vorher (H9).

Tabelle 13: Korrelationen zwischen der Bedingung und dem Ort der Aufmerksamkeit vor und nach der Manipulation bei Personen mit hoher versus niedriger aufgabenspezifischer Selbstwirksamkeit vorher und Überprüfung der Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher (H9)

Messung	r^2 ³ (SW < 3)		r^2 ³ (SW ≥ 3)		F-Wert Interaktion
	vor Manipulation	nach Manipulation	vor Manipulation	nach Manipulation	
Denkprotokoll					
1 Aufgabenbearbeitung	-.08	-.07	.01	-.19	0.32
2 Aufgabenschwierigkeit	-.15	-.09	-.21	.25	1.57
3 Sonstige Gedanken ¹	.16	-.01	.04	.15	0.31
4 Positive Selbstbewertung ¹	.27	-.18	.14	.38*	5.17*
5 Negative Selbstbewertung ¹	.38*⁴	.35*	-.07	-.10	2.86^a
Fragebogen					
6 Nachdenken	-	.18	-	.00	0.45
7 Fehlervermeidung	-	.27	-	.02	0.38
8 Geschwindigkeit	-	-.15	-	.16	0.14
9 Leistungsbewertung	-	-.55**	-	-.24	2.15
10 Pos. Selbsteinschätzung	-	.03	-	-.07	0.15
11 Neg. Selbsteinschätzung	-	-.15	-	-.08	0.04

Anmerkungen: *signifikant auf dem Niveau von .05 (zweiseitig)

¹ Nicht normalverteilte Variablen

² Niedrige SW: SW < 3, n = 34; hohe SW: SW ≥ 3, n = 28

³ Pearson Korrelation zwischen der Bedingung und den Messungen

⁴ Negative Selbstbewertung wurde nicht ausgewertet (signifikante Korrelation mit der Bedingung *vor* der Manipulation)

5.3.8 Die Auswirkung von Lob auf die Arbeitsleistung (H10)

Hypothese 10 behauptet, dass sich Lob negativ auf die Arbeitsleistung auswirkt. Die Berechnungen der einfaktoriellen Varianzanalyse (und dem Kruskal-Wallis-Test) zeigten, dass sich Lob überhaupt nicht auf die Arbeitsleistung auswirkte. In Tabelle 14 sind die Ergebnisse der Analysen zu sehen. Gerechnet wurde mit $n = 59$ Versuchspersonen. Hypothese 10 findet in diesem Datensatz keine Unterstützung.

Tabelle 14: Einfaktorielle Varianzanalyse und Kruskal-Wallis-Test zur Überprüfung des Einflusses von Lob auf die Arbeitsleistung (Hypothese 10)

Messung	M (SD) Lob	M (SD) Neutral	F-Wert	χ^2
1 Fehler ¹	1.3 (1.5)	1.3 (2.2)	0.00	0.76
2 Züge	62.7 (26.1)	70.1 (33.6)	0.89	0.88
3 Zeit ¹	5.9 (0.7)	5.8 (0.5)	0.08	0.75
4 Absolut ¹	353.3 (100.8)	337.9 (104.9)	0.33	0.42
5 Relativ	6.4 (3.1)	5.7 (2.4)	1.04	1.10
6 Effizienz ¹	1.1 (0.6)	1.0 (0.4)	1.01	0.70
7 Optimal ¹	0.7 (0.6)	0.8 (0.5)	0.05	0.20

Anmerkungen: ¹ Nicht normalverteilte Variablen
n = 59

5.3.9 Der Ort der Aufmerksamkeit als Mediator (H11)

Hypothese 11 behauptet, dass die Auswirkung von Lob auf die Arbeitsleistung durch die Art der Gedanken mediiert wird. Voraussetzung für eine Mediation ist, dass die Bedingung mit den Leistungsmaßen und dem Mediator (Ort der Aufmerksamkeit) signifikant korreliert (Bortz, 1999; S.457 ff). Der Mediator muss darüber hinaus signifikant mit den Leistungsmaßen korrelieren. Diese Voraussetzungen werden in Tabelle 15 (Korrelationen zwischen dem Ort der Aufmerksamkeit und den Leistungsmaßen)

und Tabelle 16 (Korrelationen zwischen dem Ort der Aufmerksamkeit und der Bedingung und den Leistungsmaßen und der Bedingung) geprüft.

Signifikante Korrelationen zwischen dem **Ort der Aufmerksamkeit** und den **Leistungsmaßen** sind in Tabelle 15 zwischen den aufgabenbezogenen Gedanken und der Anzahl Züge zu beobachten ($r = -.25, p < .05$); zwischen den sonstigen Gedanken und der Fehleranzahl ($r = .32, p < .01$); zwischen den negativen, selbstbezogenen Gedanken und der Fehleranzahl ($r = .34, p < .01$), der relativen Problemlöseleistung ($r = -.25, p < .05$), der optimalen Effizienz ($r = -.26, p < .05$); und zwischen der Skala „Nachdenken“ und der Anzahl Züge ($r = -.26, p < .05$). Tendenziell signifikante Korrelationen ergaben sich zwischen den aufgabenbezogenen Gedanken (Aufgabenbearbeitung) und der Fehleranzahl ($r = -.24, p < .10$); zwischen den sonstigen Gedanken und der Anzahl Züge ($r = .22, p < .10$); zwischen den negativen selbstbezogenen Gedanken und der Anzahl Züge ($r = .23, p < .10$) und der Effizienz ($r = -.24, p < .10$).

Tabelle 16 zeigt alle Korrelationen zwischen dem **Ort der Aufmerksamkeit** und der **Bedingung** und den **Leistungsmaßen** und der **Bedingung**. Die Maße für den Ort der Aufmerksamkeit waren unabhängig von der Bedingung. Nur die Skala Leistungsbewertung korrelierte signifikant positiv mit der Bedingung ($r = -.39, p < .01$). Alle Leistungsmaße waren unabhängig von der Bedingung. Es zeigten sich nur bei den *weiblichen* Versuchspersonen signifikante Korrelationen zwischen der Bedingung und der Anzahl Züge ($r = -.42, p < .05$), der relativen Problemlöseleistung ($r = .53, p < .01$) und der Effizienz ($r = .44, p < .05$). Unter Berücksichtigung der Voraussetzungen konnte die Mediatorhypothese 11 nicht getestet werden.

Tabelle 15: Korrelationen zwischen dem Ort der Aufmerksamkeit und den Leistungsmaßen

Messungen	Fehler	Züge	Zeit	Absolut	Relativ	Effiz.	Optimal
<u>Denkprotokoll</u>							
1 Aufgabenbearbeitung	-.23^a	-.26[*]	-.13	-.01	.28[*]	.30[*]	.34^{**}
2 Aufgabenschwierigkeit	-.01	-.01	.10	-.09	.01	-.04	-.05
3 Sonstige Gedanken	.32[*]	.23^a	.09	.08	-.21	-.23^a	-.32[*]
4 Pos. Selbstbewertung	-.06	-.03	-.19	-.02	-.03	.07	-.03
5 Neg. Selbstbewertung	.33^{**}	.22^a	.18	-.18	-.33^{**}	-.36^{**}	-.41^{**}
<u>Fragebogen</u>							
6 Nachdenken	-.06	-.26[*]	-.12	-.03	.33[*]	.33[*]	.32[*]
7 Fehlervermeidung	-.18	.08	-.12	-.08	-.21	-.12	-.13
8 Geschwindigkeit	.06	-.13	-.13	-.07	-.10	-.17	-.11
9 Leistungsbewertung	-.04	-.11	.13	-.04	.09	.01	.10
10 Pos. Selbsteinschätzung	-.06	-.23^a	.13	-.09	.17	.08	.10
11 Neg. Selbsteinschätzung	.13	-.20	.17	-.18	.15	.06	.06

Anmerkungen: ^a signifikant auf dem Niveau von .10 (zweiseitig)

* signifikant auf dem Niveau von .05 (zweiseitig)

** signifikant auf dem Niveau von .01 (zweiseitig)

n = 59

Tabelle 16: Korrelationen zwischen dem Ort der Aufmerksamkeit und der Bedingung und den Leistungsmaßen und der Bedingung

Messungen	r n = 59	r (m) n=27	r (w) n=32
Ort der Aufmerksamkeit			
<u>Denkprotokoll</u>			
1 Aufgabenbearbeitung ¹	-.11	-.19	-.05
2 Aufgabenschwierigkeit ¹	.06	.23	-.09
3 Sonstige Gedanken ¹	.05	.01	.06
4 Pos. Selbstbewertung ¹	.09	.26	-.00
5 Neg. Selbstbewertung ¹	.12	.12	.12
<u>Fragebogen</u>			
6 Nachdenken	.10	.21	.06
7 Fehlervermeidung	-.12	-.19	-.06
8 Geschwindigkeit	.21	.28	.19
9 Leistungsbewertung	-.39**	-.43*	-.35*
10 Pos. Selbsteinschätzung	.01	-.12	.16
11 Neg. Selbsteinschätzung	-.10	-.09	-.11
Leistungsmaße			
12 Fehler	.01	.27	-.14
13 Züge	-.12	.18	-.42*
14 Zeit	.04	.31	-.11
15 Absolut	.08	.07	.07
16 Relativ	.13	-.04	.53**
17 Effizienz	.13	-.11	.44*
18 Optimal	-.03	-.16	.19

Anmerkungen: *signifikant auf dem Niveau von .05 (zweiseitig)
 ** signifikant auf dem Niveau von .01 (zweiseitig)

¹ 1-5: Daten nach der Manipulation

5.4 Weiterführende Analysen

Die folgenden Analysen sollen hauptsächlich Anregungen für die weitere Forschung und für die Diskussion liefern. Es geht hier um die „Entdeckung“ von möglicherweise ungewöhnlichen Zusammenhängen und eventuell interessanten Geschlechtsunterschieden. Einige der Ergebnisse flossen in einen Vorschlag zu einem „Modell zur Auswirkung von gleichgeschlechtlichem Lob auf die Problemlöseleistung, basierend auf Ideen der Feedback Intervention Theory (Kluger & DeNisi, 1996)“. Das Modell wird im Abschnitt 6.5 unter „Vorschläge für die weitere Forschung“ vorgestellt.

5.4.1 Glaubwürdigkeit des Lobs

Wirkt sich die Glaubwürdigkeit des Lobs auf die Berechnungen aus? Um diese Frage zu prüfen wurde eine Variable „Glaubwürdigkeit des Lobs“ aus den Manipulationcheckitems Lob 5 bis Lob 9 gebildet (Abschnitt 5.2.3). Die negativen Items Lob 6 und Lob 9 wurden zuerst umkodiert, dann Cronbachs Alpha aus den 5 Items berechnet ($\alpha = .86$). Der Mittelwert der 5 Items bildet die Variable „Glaubwürdigkeit“. Um zu sehen, ob diese Variable einen Einfluss auf die Ergebnisse hatte, korrelierte ich sämtliche abhängigen Variablen mit der Variable „Glaubwürdigkeit“. Ergebnis ist, dass keine der abhängigen Variablen mit der Variable „Glaubwürdigkeit“ signifikant korrelierte (Anhang B3b, Tabelle 2).

5.4.2 Explorative Datenanalyse

Der Zusammenhang von Freude und Aktivierung mit dem Ort der Aufmerksamkeit. Wie hängen die Freude und die Aktivierung mit dem Ort der Aufmerksamkeit zusammen? Diese Frage wurde in Kapitel 3, Abschnitt 3.3.4 gestellt. Tabelle 17 gibt Antworten über die Stärke der Korrelationen dieser Variablen.

Die *Freude vorher* hing signifikant negativ zusammen mit den sonstigen Gedanken (vgl. Hypothese 5) und der Geschwindigkeit. Die *Freude nachher* hing signifikant positiv zusammen mit der relativen Häufigkeit der positiven, selbstbezogenen

Gedanken und signifikant negativ mit der Negativen Selbsteinschätzung und der Geschwindigkeit. Die *Aktivierung vorher* korrelierte mit keiner der Variablen. Die *Aktivierung nachher* korrelierte signifikant positiv mit der Skala Nachdenken. Geschlechtsspezifische Unterschiede innerhalb dieser Zusammenhänge sind in Abschnitt 5.4.3 zu finden.

Tabelle 17: Korrelation der Freude und der Aktivierung mit dem Ort der Aufmerksamkeit

Korrelationen	Fv	Fn	Av	An
<u>Denkprotokoll</u>				
1 Aufgabenbearbeitung	.08	-.10	-.01	-.13
2 Aufgabenschwierigkeit	.01	.03	-.14	.03
3 Sonstige Gedanken	-.25*	.00	.09	.14
4 Pos. Selbstbewertung	.17	.25*	.06	.08
5 Neg. Selbstbewertung	.03	.09	-.19	.01
<u>Fragebogen</u>				
6 Nachdenken	.07	.15	.04	.27*
7 Fehlervermeidung	.03	-.04	-.01	-.01
8 Geschwindigkeit	-.27*	-.36**	-.02	.10
9 Leistungsbewertung	-.20	-.05	-.08	-.10
10 Pos. Selbsteinschätzung	-.09	-.18	-.07	-.07
11 Neg. Selbsteinschätzung	-.18	-.56**	-.02	-.01

Anmerkungen: ^a signifikant auf dem Niveau von .10 (zweiseitig)
 *signifikant auf dem Niveau von .05 (zweiseitig)
 **signifikant auf dem Niveau von .01 (zweiseitig)
 Fv = Freude vorher, Fn = Freude nachher
 Av = Aktivierung vorher, An = Aktivierung nachher
 n = 67

Der Zusammenhang von Freude, Aktivierung, Selbstwirksamkeit und Aufgabenstandards mit der Leistung. Tabelle 18 zeigt die Korrelationen dieser Variablen. Bei den Variablen, die *vor der Bearbeitung* der Aufgabe gemessen wurden, korrelierten die Aktivierung und der Aufgabenstandard Fehler signifikant mit den Leistungs-

maßen. Aktivierung hing signifikant positiv zusammen mit den Leistungsmaßen absolute Problemlöseleistung, Effizienz und Optimal. Der Aufgabenstandard Fehler hing signifikant positiv zusammen mit der Anzahl Züge, die die Versuchsperson machte. Tendenziell signifikant wurden die positiven Korrelationen zwischen der Selbstwirksamkeit und den Leistungsmaßen Züge und Absolut. Die Aktivierung hing tendenziell positiv zusammen mit dem Leistungsmaß Relativ. Der Aufgabenstandard Zeit hing tendenziell positiv zusammen mit dem Leistungsmaß Absolut.

Bei den Variablen, die *nach der Bearbeitung* der Aufgabe gemessen wurden, korrelierten die aufgabenspezifische Selbstwirksamkeit und die Freude signifikant mit den Leistungsmaßen. Die aufgabenspezifische Selbstwirksamkeit hing signifikant positiv zusammen mit den Leistungsmaßen Effizienz und Optimal und negativ mit der Anzahl Fehler. Die Freude hing signifikant negativ zusammen mit den Leistungsmaßen Zeit und Fehler und positiv mit den Leistungsmaßen Absolut, Relativ, Effizienz und Optimal. Tendenziell signifikant wurde die negative Korrelation zwischen der Selbstwirksamkeit und dem Leistungsmaß Zeit und die positive Korrelation mit dem Leistungsmaß Relativ. Der Aufgabenstandard Züge hing signifikant positiv zusammen mit dem Leistungsmaß Züge.

Tabelle 18: Korrelation der Freude, der Aktivierung, der aufgabenspezifischen Selbstwirksamkeit und den drei persönlichen Aufgabenstandards mit den Leistungsmaßen

Messungen	Zeit	Züge	Fehler	Absolut	Relativ	Effiz.	Optim.
<u>vorher</u>							
SW ¹	-.20	.22^a	-.13	.22^a	.17	.20	.20
Freude	-.12	-.09	-.11	.01	.05	.10	.11
Aktivierung	-.19	.17	-.03	.26*	.24^a	.26*	.26*
AS Zeit	-.06	.20	-.11	.22^a	.18	.11	.12
AS Züge	-.07	.17	-.16	.06	.01	.01	.02
AS Fehler	.01	.37**	-.01	.16	.06	.04	.04
<u>nachher</u>							
SW ¹	-.22^a	-.00	-.29*	.21	.23^a	.27*	.28*
Freude	-.32*	.20	-.30*	.40**	.39**	.42**	.43**
Aktivierung	-.03	.19	-.05	.17	.13	.11	.11
AS Zeit	-.16	.05	.06	.11	.10	.15	.14
AS Züge	-.14	.22^a	-.04	.15	.10	.14	.14
AS Fehler	-.01	.09	.01	.06	.04	.01	.01

Anmerkungen: ^a signifikant auf dem Niveau von .10 (zweiseitig)
^{*}signifikant auf dem Niveau von .05 (zweiseitig)
^{**}signifikant auf dem Niveau von .01 (zweiseitig); n = 59
¹ SW = aufgabenspezifische Selbstwirksamkeit

5.4.3 Geschlecht

Männer und Frauen sind trotz vieler Gemeinsamkeiten unterschiedlich. Worin sie sich innerhalb der Variablen unterscheiden wird im Abschnitt „Haupteffekte“ analysiert. Männer und Frauen reagierten unterschiedlich auf bewertendes Feedback in einer Feldstudie von Johnson und Helgeson (2002). Reagieren Männer und Frauen auch im Labor unterschiedlich auf (künstliches) Feedback? Sollten sich Unterschiede zwischen

den Einschätzung von Frauen und Männern zeigen, werden zwei Effekte wirksam: einerseits die Effekte der Person des Versuchsleiters und andererseits nur die Effekte eines gleichgeschlechtlichen Lobgebers (die weiblichen Versuchspersonen wurden ausschließlich von einer Versuchsleiterin gelobt und die männlichen Versuchspersonen ausschließlich von einem Versuchsleiter). Auswirkungen auf die Hypothesen werden im Abschnitt „Hypothesen“ analysiert.

5.4.3.1 Haupteffekte

Haupteffekte wurden mit einer einfaktoriellen Varianzanalyse anhand des Originaldatensatzes ($n = 67$) untersucht. Im Methodenteil wurde in Tabelle 1 Geschlechtsunterschiede bei der **Leistung** berichtet. Unterschiede zeigten sich hinsichtlich der durchschnittlichen Anzahl Züge (M (Männer) = 50, M (Frauen) = 79 Züge, $F = 21.52$, $p < .01$), der durchschnittlichen Anzahl Fehler (M (Männer) = 0.6, M (Frauen) = 1.9, $F = 10.31$, $p < .01$), der relativen Problemlöseleistung (M (Männer) = 7.36, M (Frauen) = 4.89, $F = 14.97$, $p < .01$), der Effizienz (M (Männer) = 1.26, M (Frauen) = 0.88, $F = 10.06$, $p < .01$), und der optimalen Effizienz (M (Männer) = 1.07, M (Frauen) = 0.49, $F = 19.46$, $p < .01$).

Tabelle 19 überprüft den Einfluss des Geschlechts auf den **Ort der Aufmerksamkeit**. Männer äußerten mehr potentiell aufgabenförderliche Gedanken und mehr Gedanken zur Aufgabenschwierigkeit als Frauen. Frauen äußerten mehr sonstige Gedanken und mehr negative, selbstbezogene Gedanken als Männer. Männer dachten tendenziell mehr nach über die Aufgabe und berichteten von tendenziell mehr positiver Selbsteinschätzung als Frauen. Es zeigte sich tendenziell ein Haupteffekt des Geschlechts auf die *Veränderung der Aufmerksamkeit* auf die Aufgabenbearbeitung (Veränderung vorher zu nachher mal Geschlecht: $F = 3.37$, $p < .10$, Frauen plus 1.5 %, Männer minus 12.4 Prozent aufgabenbezogene Gedanken). Der Einfluss des Geschlechts auf die Veränderung der sonstigen Gedanken ist signifikant ($F = 5.40$, $p < .05$, Frauen minus 7.1 Prozent, Männer plus 4.7 Prozent).

Tabelle 19: Einfaktorielle Varianzanalyse und Kruskal-Wallis-Test zur Überprüfung des Einflusses des Geschlechts auf den Ort der Aufmerksamkeit

Messung	F-Wert	χ^2	M (SD) Frauen	M (SD) Männer
<u>Denkprotokoll</u>				
1 Aufgabenbearbeitung ²	14.10**	8.00**	68.2 (24.8)	85.4 (9.5)
2 Aufgabenschwierigkeit ¹	0.10	4.71*	2.5 (5.8)	2.9 (3.5)
3 Sonstige Gedanken ^{1 2}	11.00**	4.89*	15.6 (15.6)	6.3 (5.0)
4 Positive Selbstbewertung ¹	0.23	1.69	1.7 (4.2)	1.3 (1.6)
5 Negative Selbstbewertung ^{1 2}	9.03**	4.31*	8.9 (8.5)	4.1 (3.9)
<u>Fragebogen</u>				
6 Nachdenken	3.35^a	3.09^a	3.2 (1.2)	3.7 (1.0)
7 Fehlervermeidung	0.02	0.04	3.4 (1.2)	3.5 (1.3)
8 Geschwindigkeit	1.32	1.48	2.9 (1.1)	3.2 (1.0)
9 Leistungsbewertung	1.33	1.48	2.4 (1.2)	2.7 (1.1)
10 Pos. Selbsteinschätzung	3.52^a	2.98^a	2.6 (0.8)	3.0 (0.8)
11 Neg. Selbsteinschätzung	1.11	1.16	2.4 (1.1)	2.7 (1.0)

Anmerkungen: ^a signifikant auf dem Niveau von .10 (zweiseitig)

*signifikant auf dem Niveau von .05 (zweiseitig)

**signifikant auf dem Niveau von .01 (zweiseitig)

¹ Nicht normalverteilte Variablen

² Heterogene Varianzen

n = 66 aufgrund eines fehlenden Werts

Die Männer schätzten ihre **aufgabenspezifische Selbstwirksamkeit** vor der Aufgabe tendenziell höher ein als die Frauen, $F = 3.65$, $p < .10$ (Männer: $M = 2.9$, $SD = 0.65$; Frauen: $M = 2.6$, $SD = 0.66$). Auch nach der Bearbeitung der Aufgabe schätzten sie sich höher ein als die Frauen, $F = 12.25$, $p < .01$ (Männer: $M = 2.9$, $SD = 0.79$; Frauen: $M = 2.30$, $SD = 0.60$). Männer und Frauen unterschieden sich auch in der *Veränderung der aufgabenspezifischen Selbstwirksamkeit vorher zu nachher*, abhängig von der Bedingung: in der neutralen Bedingung senkten die weiblichen Versuchspersonen ihre Einschätzung um 0.5 Punkte (vorher $M = 2.7$, $SD = 0.7$, nachher $M = 2.2$, $SD = 0.4$), in der Lobbedingung nur um 0.1 Punkte (vorher $M = 2.5$, $SD = 0.6$, nachher $M = 2.4$,

SD = 0.7). Die Veränderung der aufgabenspezifischen Selbstwirksamkeit vorher zu nachher war hochsignifikant ($F = 7.59, p < .01$), die Veränderung in Abhängigkeit von der Bedingung tendenziell signifikant ($F = 2.91, p < .10$). Männer verändern ihre Einschätzung nicht (Veränderung vorher – nachher: $F = .33, ns$; Veränderung in Abhängigkeit der Bedingung: $F = 1.96, ns$).

Keine Geschlechtsunterschiede zeigten sich hinsichtlich der **Freude** vor und nach der Bearbeitung der Aufgabe (vorher: $F = 0.10, ns$; nachher: $F = 1.26, ns$). Tendenziell **aktiver** waren die Frauen vor der Aufgabe, $F = 2.83, p < .10$ (Frauen: $M = 2.88, SD = 0.59$; Männer: $M = 2.61, SD = 0.70$). Nach der Aufgabe zeigten sich keine Geschlechtsunterschiede, $F = 0.11, ns$.

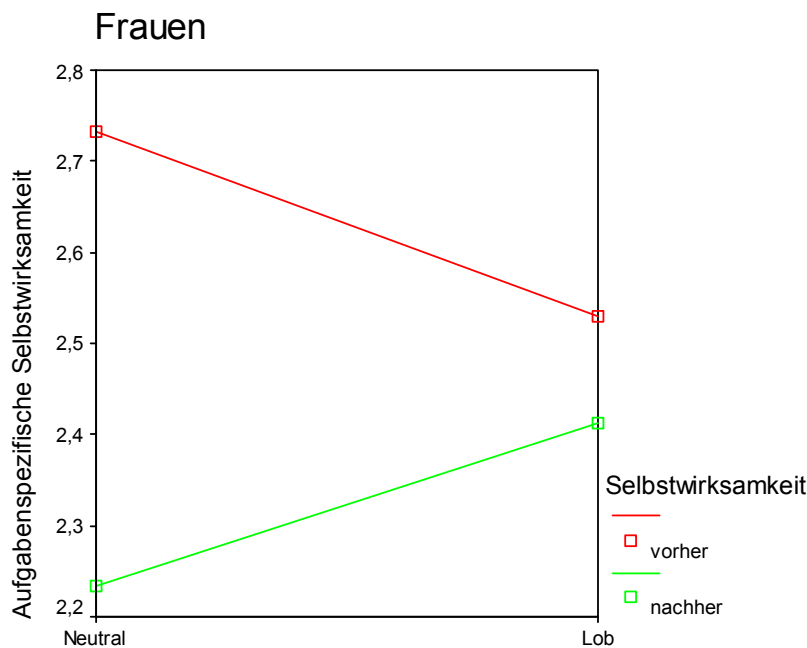


Abbildung 4: Veränderung der aufgabenspezifischen Selbstwirksamkeit bei den weiblichen Versuchspersonen in Abhängigkeit von der Bedingung.

Hinsichtlich der **unspezifischen Aufgabenziele** zeigten sich keine Geschlechtsunterschiede (Schnelligkeit: $F = 0.89, ns$; Züge: $F = 0.92, ns$; Fehler: $F = 0.00, ns$). Hinsichtlich der **spezifischen Aufgabenstandards**, die vor und nach der Bearbeitung

erhoben wurden, zeigten sich signifikante Unterschiede: Vor der Bearbeitung der Aufgaben gaben die Frauen an, im Schnitt mit mehr Fehlern ($M = 2.8$, $SD = 1.4$) zufrieden zu sein als die Männer ($M = 2.1$ Fehler, $SD = 1.1$, $F = 5.71$, $p < .05$). Hinsichtlich des Aufgabenstandards „Schnelligkeit“ ($F = 1.12$, ns) und „Züge“ ($F = 1.47$, ns) gab es keine Unterschiede. Nach der Bearbeitung der Aufgabe gaben die Frauen an, mit einer längeren Zeitdauer bis zur Problemlösung zufrieden zu sein als die Männer, $F = 6.87$, $p < .01$ (Frauen: $M = 7.5$ Minuten, $SD = 0.9$; Männer: $M = 6.7$ Minuten, $SD = 1.3$). Der Aufgabenstandard der Frauen hinsichtlich der Anzahl Züge bis zur Problemlösung war nach der Bearbeitung der Aufgabe signifikant höher als der Aufgabenstandard der Männer, $F = 5.88$, $p < .05$ (Frauen: $M = 82.0$ Züge, $SD = 4.90$; Männer: $M = 78.8$ Züge, $SD = 5.8$). Der Aufgabenstandard „Fehler“ nach der Bearbeitung war bei Männern und Frauen gleich, $F = 2.68$, ns.

5.4.3.2 Hypothesen

Um eine potentielle Interaktion der Bedingung mit dem Geschlecht bei der Interpretation der Ergebnisse zu berücksichtigen, berechnete ich für alle Hypothesen zusätzliche zwei- oder dreifaktorielle Varianzanalysen (mit oder ohne Messwiederholung) beziehungsweise getrennt nach Geschlecht Kruskal-Wallis-Tests oder Korrelationsanalysen.

Bei den **Hypothesen 1 und 2** zeigte sich keine Auswirkung des Geschlechts. Die Veränderung der *Freude* ist unabhängig vom Geschlecht (Veränderung mal Geschlecht: $F = 0.93$, ns). Es zeigte sich auch kein Interaktionseffekt (Veränderung mal Geschlecht mal Bedingung: $F = 0.20$, ns). Die Veränderung der *Aktivierung* ist gleichermaßen unabhängig vom Geschlecht (Veränderung mal Geschlecht: $F = 2.03$, ns). Es zeigte sich auch kein Interaktionseffekt (Veränderung mal Geschlecht mal Bedingung: $F = 1.18$, ns). Keine Auswirkungen des Geschlechts zeigte sich bei **Hypothese 3**. Keine der Interaktionen Bedingung mal Geschlecht wurde signifikant, weder bei den Daten des lauten Denkprotokolls noch bei den Fragebogendaten. Bei **Hypothese 4** zeigten sich ebenfalls keine signifikanten Interaktionen der Bedingung mit dem Geschlecht auf die Veränderung der Aufgabenstandards (AS Zeit: $F = 0.04$, ns; AS Züge:

F = 0.34, ns; AS Fehler: F = 0.19, ns). Bei **Hypothese 5** zeigte das Geschlecht keinen Einfluss (r (Frauen) = $-.34$, $p = .06$; r (Männer) = $-.29$, $p = .09$). Das gleiche galt für **Hypothese 6** (vgl. Tabelle 21) und für **Hypothese 7** (AS Zeit: F = 0.09, ns; AS Züge: F = 0.30, ns; AS Fehler: F = 1.12, ns).

Bei **Hypothese 8** zeigte sich bei den Frauen ein tendenziell signifikanter negativer Zusammenhang der *aufgabenspezifischen Selbstwirksamkeit vorher* mit der relativen Häufigkeit positiver selbstbezogener Gedanken ($r = -.30$, $p < .10$, H8 bestätigt), bei den Männern ein signifikant positiver Zusammenhang ($r = .43$, $p < .01$, H8 widerlegt). Nur bei den Männern zeigte sich ein tendenziell signifikanter positiver Zusammenhang der aufgabenspezifischen Selbstwirksamkeit vorher mit der relativen Häufigkeit sonstiger Gedanken ($r = .32$, $p < .10$, H8 widerlegt). Bei der *aufgabenspezifischen Selbstwirksamkeit nachher* zeigten sich nur bei den Männern signifikante Zusammenhänge. Die aufgabenspezifische Selbstwirksamkeit nachher hing tendenziell negativ zusammen mit der relativen Häufigkeit an Gedanken über die Aufgabenschwierigkeit ($r = .29$, $p < .10$, H8 widerlegt), signifikant positiv mit der relativen Häufigkeit an positiven selbstbezogenen Gedanken ($r = .40$, $p < .05$, H8 widerlegt) und signifikant negativ mit der Skala Negative Selbsteinschätzung ($r = -.42$, $p < .05$, H8 bestätigt).

Rechnete man für **Hypothese 9** für alle Variablen des Orts der Aufmerksamkeit dreifaktorielle Varianzanalysen (Tabelle 20 zeigt nur für die (tendenziell) signifikanten Ergebnisse M und SD), zeigte sich ein signifikanter Effekt der Dreifachinteraktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher mal Geschlecht auf die relative Häufigkeit *negativer selbstbezogener Gedanken* nach der Manipulation (F = 4.54, $p < .05$). Nur bei den Frauen mit relativ niedriger aufgabenspezifischer Selbstwirksamkeit vorher zeigte sich eine signifikante positive Korrelation zwischen der Bedingung und der relativen Häufigkeit negativer selbstbezogener Gedanken *vorher* ($r = .49$, $p < .05$). Ein signifikanter Dreifachinteraktionseffekt zeigte sich auf die Skala *Leistungsbewertung* (F = 4.11, $p < .05$). Tendenziell signifikant wurde die Dreifachinteraktion auf die *sonstigen Gedanken* nach der Manipulation (F = 3.26, $p < .10$). Um die Dreifachinteraktionen besser verstehen zu können, sind sie getrennt nach Geschlecht graphisch in den Abbildungen 5 – 10 zu sehen.

Tabelle 20: Überprüfung von Hypothese 9, getrennt nach Geschlecht

Messung	Männer		Frauen		F-Wert Interaktion ³
	SW > 3 ² N / Lob	SW < 3 N / Lob	SW > 3 N / Lob	SW < 3 N / Lob	
<hr/>					
Sonstige Gedanken ¹					
M	6.0 / 16.4	10.6 / 3.5	12.7 / 9.7	11.7 / 16.2	3.26^a
SD	5.0 / 18.6	13.7 / 4.6	9.7 / 13.9	12.3 / 19.1	
Negative Selbst- bewertung ¹					
M	3.6 / 7.2	6.3 / 7.0	15.4 / 8.1	2.3 / 14.5	4.54*
SD	5.9 / 9.8	8.9 / 4.7	16.1 / 10.3	4.3 / 12.9	
<hr/>					
Fragebogen					
Leistungsbewertung					
M	3.2 / 1.9	3.5 / 2.6	2.3 / 2.5	3.2 / 1.6	4.11*
SD	1.2 / 1.0	1.2 / 0.8	0.8 / 1.4	1.2 / 0.8	

Anmerkungen: ^a signifikant auf dem Niveau von .10 (zweiseitig)

*signifikant auf dem Niveau von .05 (zweiseitig)

¹ Nicht normalverteilte Variable

² Niedrige SW: SW < 3, n = 34; hohe SW: SW ≥ 3, n = 28

³ Varianzen sind nicht homogen

Negative Selbstbewertung wurde ausgewertet (signifikante Korrelation mit der Bedingung vor der Manipulation nur bei Frauen mit niedriger Selbstwirksamkeit). Die Gruppen sind teilweise unterschiedlich groß und knapp unter 10 Personen pro Gruppe.

Abbildung 5 zeigt, dass bei den gelobten *Frauen* mit relativ niedriger aufgabenspezifischer Selbstwirksamkeit die relative Häufigkeit negativer selbstbezogener Gedanken nach der Manipulation *signifikant höher* ist als bei den neutralen Frauen (Lob: M = 2.3, SD = 4.3; Neutral: M = 14.5, SD = 12.9). Bei den gelobten Frauen mit relativ hoher aufgabenspezifischer Selbstwirksamkeit ist die relative Häufigkeit negativer selbstbezogener Gedanken *signifikant niedriger* als bei den neutralen (Lob: M = 8.1, SD = 10.3; Neutral: M = 15.4, SD = 16.1). Frauen mit niedriger aufgabenspezifischer Selbstwirksamkeit haben nach Lob signifikant mehr negative selbstbezogene Gedanken als Frauen mit hoher aufgabenspezifischer Selbstwirksamkeit (niedrige SW: M = 14.5, SD = 12.9; hohe SW: M = 8.1, SD = 10.3, F = 4.89, p < .05; Hypothese 9 zum Teil bestätigt). **Abbildung 6** zeigt, dass bei den gelobten *Männern* mit relativ niedriger aufgabenspezifischer Selbstwirksamkeit die relative Häufigkeit negativer selbstbezogener Gedanken nach der Manipulation gleich ist wie bei den neutralen (Lob: M =

6.3, SD = 8.9; Neutral: M = 7.0, SD = 4.7). Bei den gelobten Männern mit relativ hoher aufgabenspezifischer Selbstwirksamkeit ist die relative Häufigkeit negativer selbstbezogener Gedanken *signifikant höher* als bei den neutralen (Lob: M = 7.2, SD = 9.8; Neutral: M = 3.6, SD = 5.9). Männer mit niedriger aufgabenspezifischer Selbstwirksamkeit haben nach Lob genau so viel negative selbstbezogene Gedanken als Männer mit hoher aufgabenspezifischer Selbstwirksamkeit (niedrige SW: M = 7.0, SD = 4.7; hohe SW: M = 7.2, SD = 9.8, $F = .28$, ns; Hypothese 9 zum Teil nicht bestätigt).

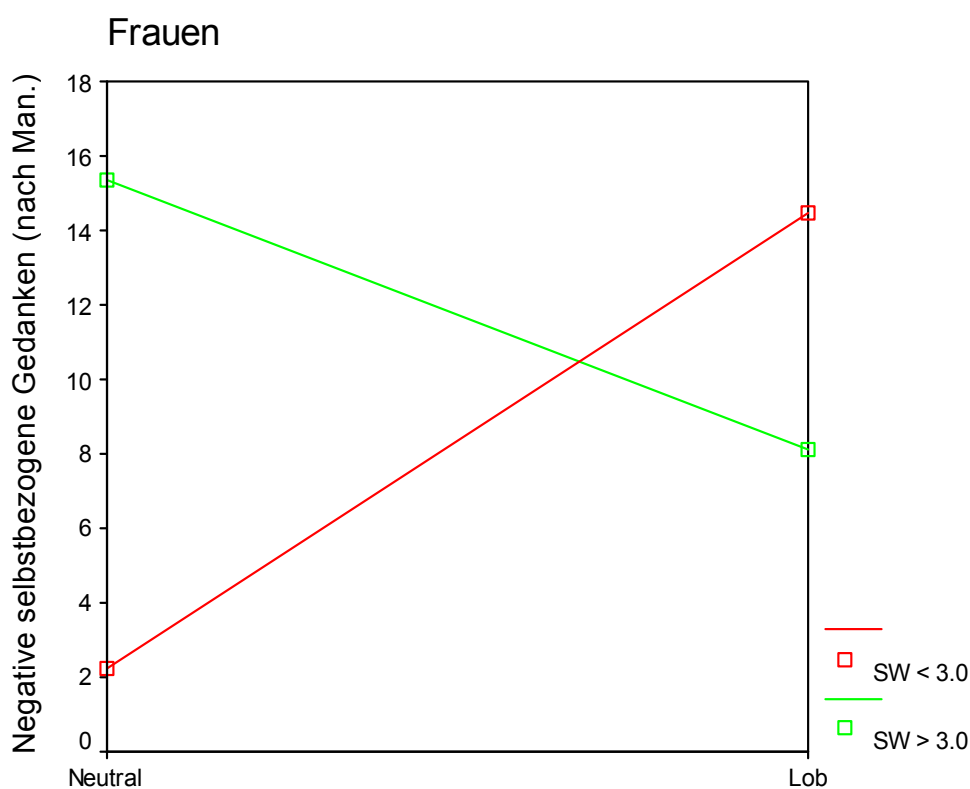


Abbildung 5: Die *signifikante* Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der negativen selbstbezogenen Gedanken nach der Manipulation bei den Frauen (Hypothese 9)

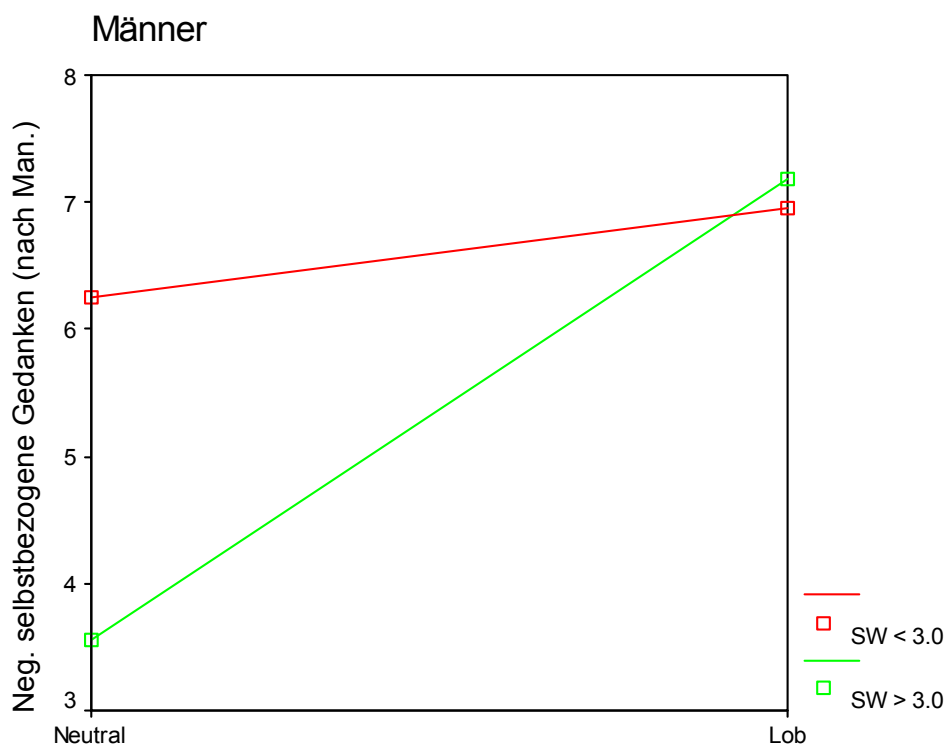


Abbildung 6: Die *nicht signifikante* Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der negativen selbstbezogenen Gedanken nach der Manipulation bei den Männern (Hypothese 9)

Abbildung 7 zeigt, dass die *Frauen* mit relativ niedriger aufgabenspezifischer Selbstwirksamkeit vorher sich nach Lob signifikant seltener mit anderen verglichen als nach dem neutralen Kommentar (Lob: $M = 3.2$, $SD = 1.2$; Neutral: $M = 1.6$, $SD = 0.8$). *Frauen* mit relativ hoher aufgabenspezifischer Selbstwirksamkeit verglichen sich nach Lob gleich oft mit anderen (Lob: $M = 2.5$, $SD = 1.4$; Neutral: $M = 2.3$, $SD = 0.8$). *Frauen* mit niedriger aufgabenspezifischer Selbstwirksamkeit vergleichen sich nach Lob signifikant weniger häufig mit anderen, um ihre Leistung zu bewerten als *Frauen* mit hoher aufgabenspezifischer Selbstwirksamkeit (niedrige SW: $M = 1.6$, $SD = 0.8$; hohe SW: $M = 2.5$, $SD = 1.4$, $F = 5.80$, $p < .05$; Hypothese 9 zum Teil widerlegt)

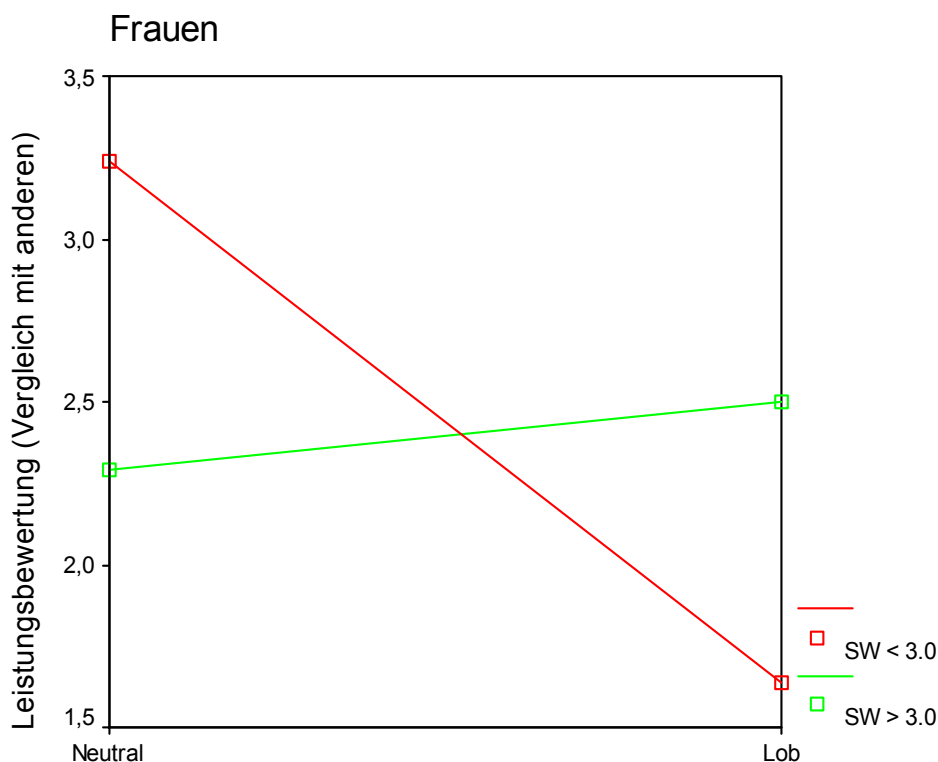


Abbildung 7: Die *signifikante* Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der Leistungsbewertung bei den Frauen (Hypothese 9)

Abbildung 8 zeigt, dass die *Männer* mit relativ niedriger aufgabenspezifischer Selbstwirksamkeit vorher sich nach Lob signifikant seltener mit anderen verglichen als nach dem neutralen Kommentar (Hypothese 9 teilweise widerlegt; Lob: $M = 2.6$, $SD = 0.8$; Neutral: $M = 3.5$, $SD = 1.2$). *Männer* mit relativ *hoher* aufgabenspezifischer Selbstwirksamkeit verglichen sich nach Lob auch signifikant seltener mit anderen (Lob: $M = 1.9$, $SD = 1.0$; Neutral: $M = 3.2$, $SD = 1.2$). Männer mit niedriger aufgabenspezifischer Selbstwirksamkeit vergleichen sich nach Lob genauso häufig mit anderen, um ihre Leistung zu bewerten, als Männer mit hoher aufgabenspezifischer Selbstwirksamkeit (niedrige SW: $M = 2.6$, $SD = 0.8$; hohe SW: $M = 1.9$, $SD = 1.0$, $F = .23$, ns; Hypothese 9 zum Teil nicht bestätigt).

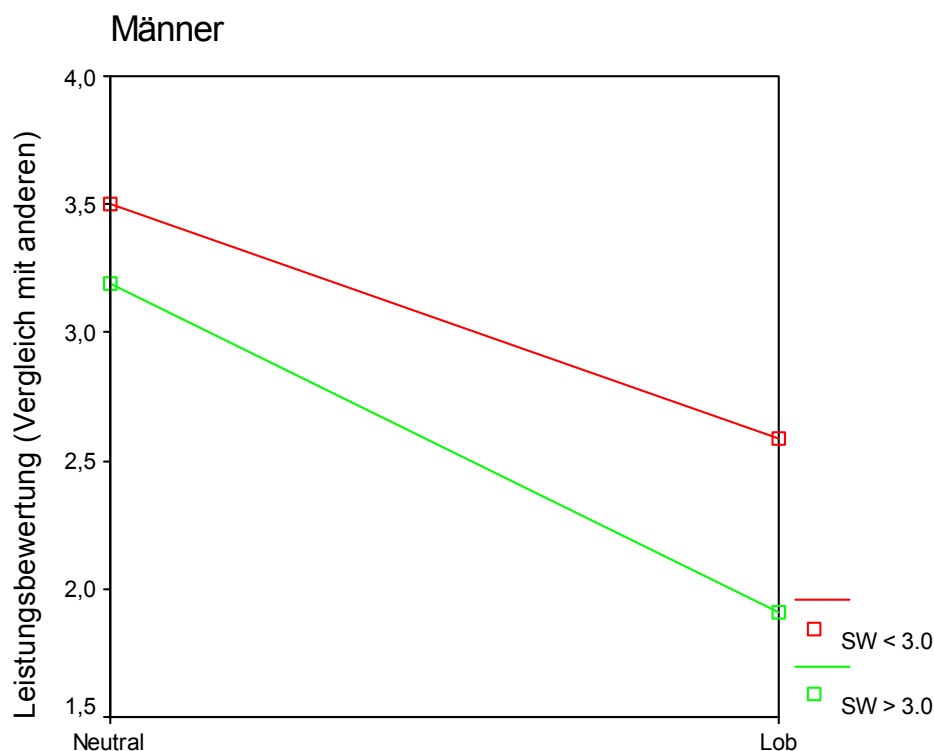


Abbildung 8: Die *nicht signifikante* Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der Leistungsbewertung bei den Männern (Hypothese 9)

Abbildung 9 zeigt, dass bei den gelobten *Frauen* mit relativ niedriger aufgabenspezifischer Selbstwirksamkeit die relative Häufigkeit sonstiger Gedanken nach der Manipulation genau so hoch ist wie bei den neutralen Frauen (Lob: $M = 16.2$, $SD = 19.1$; Neutral: $M = 11.7$, $SD = 12.3$). Bei den gelobten Frauen mit relativ hoher aufgabenspezifischer Selbstwirksamkeit ist die relative Häufigkeit sonstiger Gedanken auch genau so hoch wie bei den neutralen (Lob: $M = 9.7$, $SD = 13.9$; Neutral: $M = 12.7$, $SD = 9.7$). Frauen mit niedriger aufgabenspezifischer Selbstwirksamkeit haben nach Lob nicht mehr sonstige Gedanken als Frauen mit hoher aufgabenspezifischer Selbstwirksamkeit (niedrige SW: $M = 16.2$, $SD = 19.1$; hohe SW: $M = 9.7$, $SD = 13.9$, $F = 4.01$, $p < .10$; Hypothese 9 zum Teil nicht bestätigt). **Abbildung 10** zeigt, dass bei den gelob-

ten *Männern* mit relativ niedriger aufgabenspezifischer Selbstwirksamkeit die relative Häufigkeit sonstiger Gedanken nach der Manipulation *tendenziell geringer* ist als bei den neutralen (Lob: $M = 3.5$, $SD = 4.6$; Neutral: $M = 10.6$, $SD = 13.7$). Bei den gelobten Männern mit relativ hoher aufgabenspezifischer Selbstwirksamkeit ist die relative Häufigkeit sonstiger Gedanken *tendenziell höher* als bei den neutralen (Lob: $M = 16.4$, $SD = 18.6$; Neutral: $M = 6.0$, $SD = 5.0$). Männer mit niedriger aufgabenspezifischer Selbstwirksamkeit haben nach Lob tendenziell weniger sonstige Gedanken als Männer mit hoher aufgabenspezifischer Selbstwirksamkeit (niedrige SW: $M = 3.5$, $SD = 4.6$; hohe SW: $M = 16.4$, $SD = 18.6$, $F = 4.01$, $p < .10$; Hypothese 9 zum Teil widerlegt).

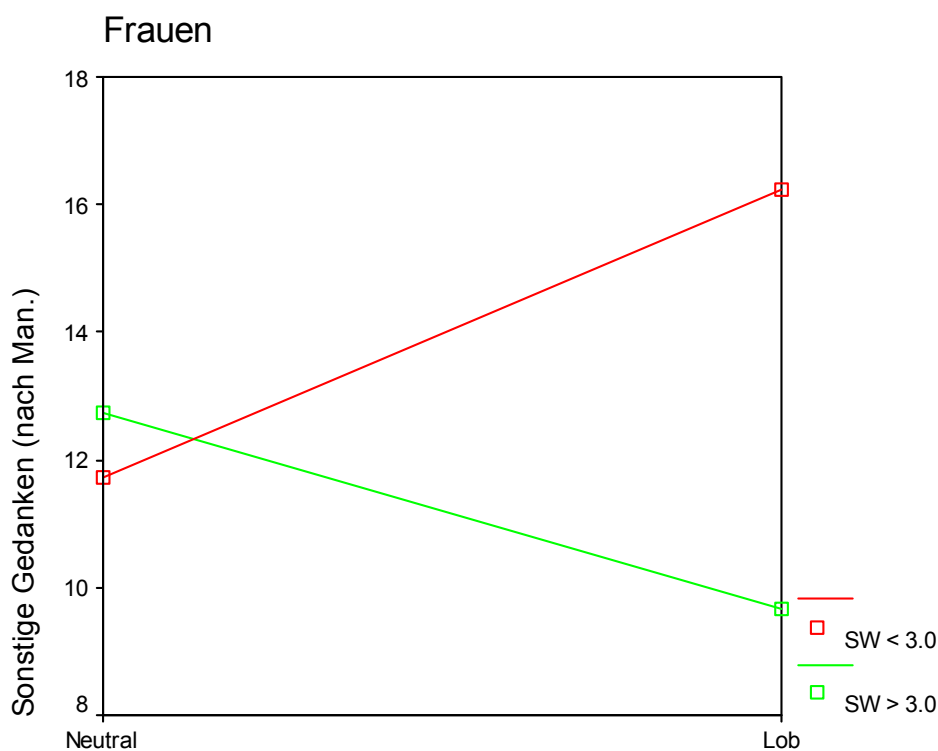


Abbildung 9: Die *nicht signifikante* Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der sonstigen Gedanken nach der Manipulation bei den Frauen (Hypothese 9)



Abbildung 10: Die *tendenziell signifikante* Interaktion Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher innerhalb der sonstigen Gedanken nach der Manipulation bei den Männern (Hypothese 9)

Rechnet man getrennt nach Geschlecht für **Hypothese 10** jeweils für die signifikanten Korrelationen aus Tabelle 15 eine einfaktorielle Varianzanalyse, zeigte sich nur bei den weiblichen Versuchspersonen ein signifikanter Effekt der Bedingung auf die Anzahl Züge ($F = 6.45$, $p < .05$), die Relative Problemlöseleistung ($F = 11.43$, $p < .01$), und die Effizienz ($F = 7.19$, $p < .05$). Die weiblichen Versuchspersonen, die gelobt wurden, machten signifikant weniger Züge ($M = 68.1$, $SD = 24.5$), hatten eine bessere Relative Problemlöseleistung ($M = 5.75$, $SD = 1.76$), und arbeiteten effizienter ($M = 1.06$, $SD = 0.51$), als die, die nicht gelobt wurden (Züge: $M = 91.6$, $SD = 27.8$; Relative Problemlöseleistung: $M = 3.95$, $SD = 5.75$; Effizienz: $M = 0.68$, $SD = 0.23$). Bei den männlichen Versuchspersonen gab es keinen Unterschied zwischen den Bedin-

gungen (Züge: $F = 0.81$, ns; Relative Problemlöseleistung: $F = 0.05$, ns; Effizienz: $F = 0.30$, ns). Die zweifaktorielle Varianzanalyse zeigte einen signifikanten Interaktionseffekt zwischen Geschlecht und Bedingung auf die Anzahl Züge ($F = 5.76$, $p < .05$) und die Effizienz ($F = 4.10$, $p < .05$). Bei den weiblichen Versuchspersonen wurde gegen die Vorhersage von Hypothese 10 die Anzahl Züge durch Lob reduziert (Abbildung 11), die Relative Problemlöseleistung verbessert (Abbildung 12) und effizienter gearbeitet (Abbildung 13). Bei den Männern wirkte sich Lob nicht auf die Leistung aus (vgl. Tabelle 16).

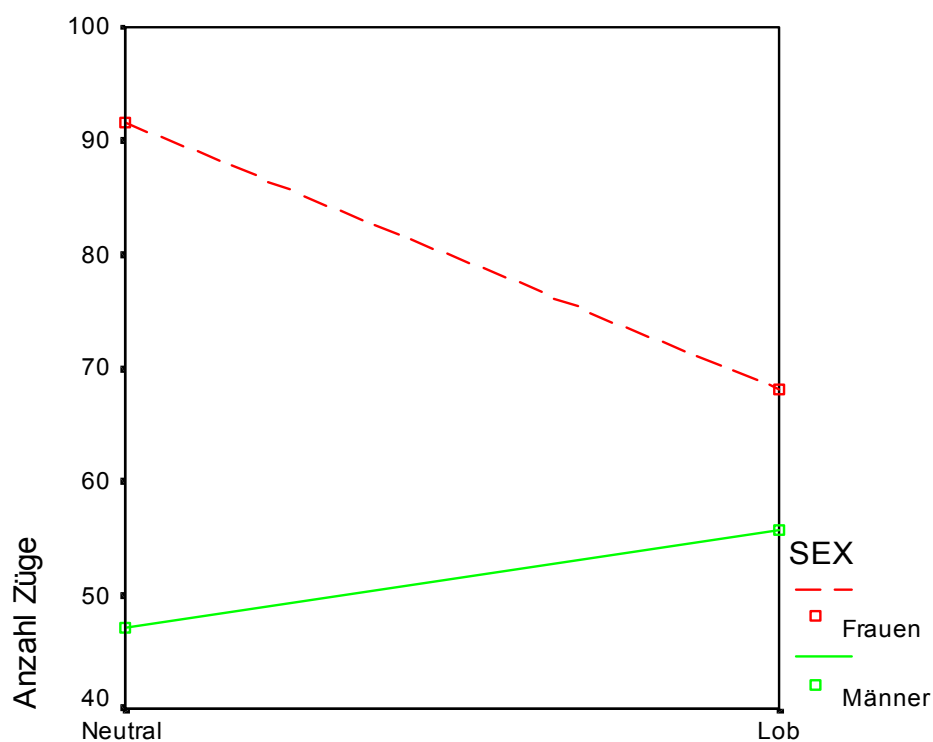


Abbildung 11: Die Auswirkung der Bedingung auf die Anzahl Züge in Abhängigkeit vom Geschlecht (*Signifikante* Interaktion Geschlecht mal Züge; Hypothese 10)

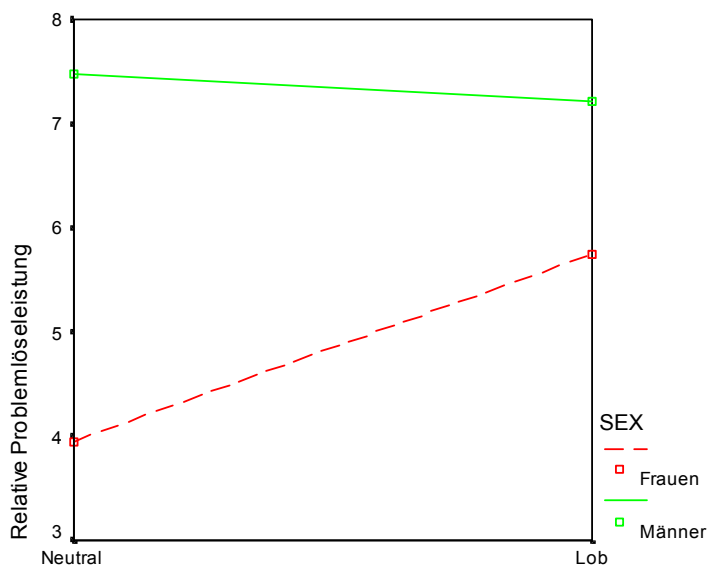


Abbildung 12: Die Auswirkung der Bedingung auf die Relative Problemlöseleistung in Abhängigkeit vom Geschlecht (*Nicht signifikante* Interaktion Geschlecht mal Bedingung, Hypothese 10)

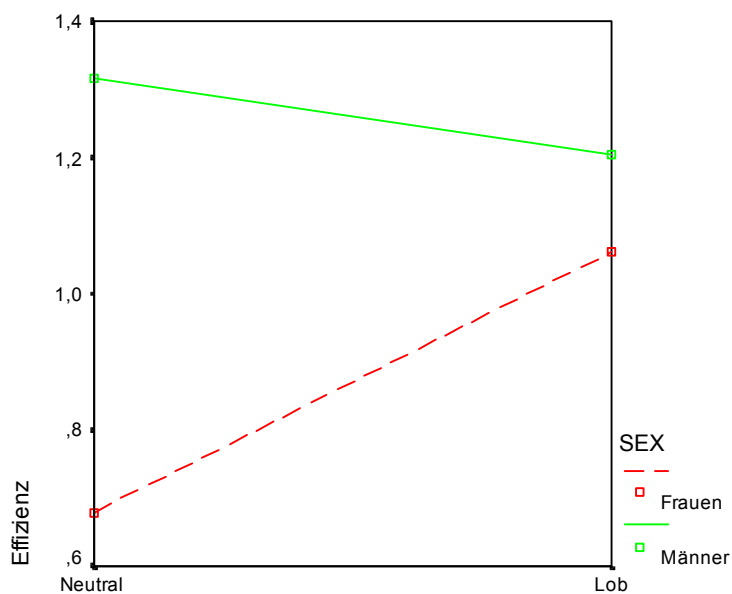


Abbildung 13: Die Auswirkung der Bedingung auf die Effizienz in Abhängigkeit vom Geschlecht (*Signifikante* Interaktion Geschlecht mal Bedingung, Hypothese 10)

Die Mediatorhypothese (**Hypothese 11**) konnte bei den Frauen nur hinsichtlich der Variable Leistungsbewertung getestet werden (signifikante Korrelation mit der Bedingung, Tabelle 16), bei den Männern fehlten die weiteren Voraussetzungen (signifikante Korrelation der Leistungsbewertung mit der Leistung, vgl. Tabelle 16 und Anhang Tabelle 1). Die hierarchische Regression (Tabelle 21) zeigt, dass der Einfluss von Lob auf die Effizienz nicht durch die Leistungsbewertung mediiert wird.

Tabelle 21: Test von Hypothese 11 bei den Frauen

	Effizienz	
	Schritt 1	Schritt 2
Lob (Beta)	.44*	.37*
Leistungsbewertung (Beta)		-.20
R	.44	.48
R-Quadrat (korrigiert)	.17	.17

Anmerkung: * signifikant auf dem Niveau von $p < .05$

5.4.3.3 Geschlechtsspezifische Zusammenhänge

Interessant finde ich zu schauen, ob sich bei Männern und Frauen bei Freude und Aktivierung verschiedene Zusammenhänge zum Ort der Aufmerksamkeit herausstellen. Darauf aufbauend analysierte ich dann, inwieweit es bei Männern und Frauen Unterschiede zwischen den Zusammenhängen zwischen dem Ort der Aufmerksamkeit und der Leistung gab. Tabelle 22 zeigt die **Geschlechtsunterschiede hinsichtlich der Korrelation der Freude und der Aktivierung mit dem Ort der Aufmerksamkeit**. Die *Freude vorher* hing bei Männern und Frauen negativ mit den sonstigen Gedanken und der Aufmerksamkeit auf die Geschwindigkeit zusammen. Nur bei den Männern hing die Freude vorher positiv mit dem Nachdenken über die Aufgabe zusammen. Nur bei den Frauen hing die Freude vorher negativ mit der negativen Selbsteinschätzung zusammen. Die *Freude nachher* hing nur bei den Frauen positiv mit den positiven selbstbezogenen Gedanken zusammen. Nur bei den Männern hing das Nachdenken

über die Aufgabe positiv mit der Freude nachher zusammen. Bei Männern und Frauen hing die Freude nachher negativ mit der Aufmerksamkeit auf die Geschwindigkeit und den negativen Selbsteinschätzung zusammen. Die *Aktivierung vorher* hing nur bei den Männern negativ mit den sonstigen Gedanken zusammen. Nur bei den Frauen hing die Aktivierung vorher negativ mit den negativen, selbstbezogenen Gedanken zusammen. Die *Aktivierung nachher* hing nur bei den Frauen negativ mit der Aufmerksamkeit auf die Aufgabenbearbeitung und positiv mit den sonstigen Gedanken zusammen. Nur bei den Männern hing das Nachdenken über die Aufgabe positiv mit der Aktivierung nachher zusammen.

Tabelle 22: Korrelation der Freude und der Aktivierung mit dem Ort der Aufmerksamkeit, getrennt nach Geschlecht

Korrelationen	Fv	Fn	Av	An
<u>Denkprotokoll</u>				
1 Aufgabenbearbeitung	.10 (.16)	-.08 (.08)	.03 (.23)	-.30^a (.13)
2 Aufgabenschwierigkeit	.01 (.01)	.09 (-.06)	-.15 (-.12)	.14 (-.10)
3 Sonstige Gedanken	-.34^a (-.29^a)	-.05 (-.11)	.13 (-.30^a)	.33^a (-.23)
4 Pos. Selbstbewertung	.25 (.03)	.33^a (.05)	.03 (.09)	.14 (.01)
5 Neg. Selbstbewertung	.05 (-.04)	.07 (-.01)	-.42* (-.11)	-.03 (.05)
<u>Fragebogen</u>				
6 Nachdenken	-.11 (.30^a)	.09 (.31^a)	.02 (.15)	.14 (.43**)
7 Fehlervermeidung	.11 (-.03)	.14 (-.22)	-.17 (.11)	-.13 (.07)
8 Geschwindigkeit	-.29^a (-.25)	-.33^a (-.38*)	-.16 (.15)	.07 (.13)
9 Leistungsbewertung	-.22 (-.16)	-.14 (.11)	.12 (-.20)	.09 (-.25)
10 Pos. Selbsteinsch.	-.19 (.01)	-.25 (-.03)	-.12 (.05)	-.09 (-.05)
11 Neg. Selbsteinsch.	-.41* (.11)	-.60** (-.48**)	-.12 (.12)	-.03 (.02)

Anmerkungen: ^a signifikant auf dem Niveau von .10 (zweiseitig)
 *signifikant auf dem Niveau von .05 (zweiseitig)
 **signifikant auf dem Niveau von .01 (zweiseitig)
 n (Frauen) = 33, n (Männer) = 34; Korrelationen der Männer in Klammern
 Fv = Freude vorher, Fn = Freude nachher
 Av = Aktivierung vorher, An = Aktivierung nachher

Geschlechtsunterschiede gab es auch hinsichtlich der **Zusammenhänge zwischen dem Ort der Aufmerksamkeit und verschiedener Leistungsmaße** (Anhang B3a, Tabelle 1). Nur bei den Frauen gab es tendenziell einen negativen Zusammenhang zwischen der Aufmerksamkeit auf die Fehlervermeidung und der Anzahl Fehler (Frauen: $r = -.30$, $p < .10$; Männer: $r = -.03$, ns). Die Aufmerksamkeit auf die Geschwindigkeit hing bei den Frauen signifikant negativ mit der Anzahl Züge zusammen ($r = -.37$, $p < .05$), bei den Männern signifikant positiv ($r = .46$, $p < .05$). Nur bei den Frauen hing die Aufmerksamkeit auf die Geschwindigkeit tendenziell positiv mit der benötigten Zeit (Frauen: $r = .32$, $p < .10$; Männer: $r = -.21$, ns) und signifikant negativ mit der absoluten Problemlöseleistung (Frauen: $r = -.41$, $p < .05$; Männer: $r = -.25$, ns) zusammen. Nur bei den Männern hing die Aufmerksamkeit auf die Geschwindigkeit hochsignifikant negativ mit der relativen Problemlöseleistung (Frauen: $r = .10$, ns; Männer: $r = -.48$, $p < .01$), der Effizienz (Frauen: $r = .15$, ns; Männer: $r = -.42$, $p < .05$) und der Optimalen Effizienz (Frauen: $r = -.15$, ns; Männer: $r = -.35$, $p < .10$) zusammen. Die Leistungsbewertung hing nur bei den Frauen tendenziell mit der Effizienz zusammen (Frauen: $r = .33$, $p < .10$; Männer: $r = .17$, ns). Die positive Selbsteinschätzung hing bei den Frauen signifikant positiv mit der benötigten Zeit zusammen ($r = .38$, $p < .05$), bei den Männern tendenziell negativ ($r = -.37$, $p < .10$). Die negative Selbsteinschätzung hing nur bei den Frauen tendenziell positiv mit der benötigten Zeit zusammen (Frauen: $r = .33$, $p < .10$; Männer: $r = -.14$, ns).

Keine Geschlechtsunterschiede gab es zwischen den Korrelationen der **Aktivierung vorher, der Freude vorher und der Leistung**. Signifikante Korrelationen gab es nur bei den Männern zwischen der **aufgabenspezifischen Selbstwirksamkeit vorher und der Leistung**: die Anzahl Züge und die Absolute Problemlöseleistung korrelierten (tendenziell) signifikant positiv mit der aufgabenspezifischen Selbstwirksamkeit (Züge: $r = .53$, $p < .01$; Absolute Problemlöseleistung: $r = .34$, $p < .10$); die benötigte Zeit und die Relative Problemlöseleistung korrelierten signifikant negativ mit der aufgabenspezifischen Selbstwirksamkeit (Zeit: $r = -.49$, $p < .01$; Relative Problemlöseleistung: $r = -.42$, $p < .05$).

5.4.4 Vorhersage der Leistung

Um zu schauen, welche Variablen die potentesten Prädiktoren zur Vorhersage verschiedener Leistungsmaße sind, wurden sämtliche Variablen in eine schrittweise Regressionsanalyse aufgenommen. Nur die signifikanten Prädiktoren mit der größten Varianzaufklärung bleiben bei diesem Verfahren übrig. Als *potentielle Prädiktoren* galten: die Bedingung, das Alter, das Geschlecht (nur im gesamten Datensatz), die Semesteranzahl, die aufgabenspezifische Selbstwirksamkeit vorher, die (Vor)-Freude, die Aktivierung vorher, die drei zuvor eingeschätzten Aufgabenstandards Zeit, Züge, Fehler, und alle Variablen zur Erfassung des Orts der Aufmerksamkeit (Denkprotokoll und Fragebogen). Bei allen nicht genannten Leistungsmaßen war eine Vorhersage nicht möglich. Zusätzlich wurden Strukturkoeffizienten (*c*) berechnet (Korrelation zwischen Prädiktor und Kriterium geteilt durch das Beta-Gewicht, Bortz, 1999, S. 438).

Gesamter korrigierter Datensatz (n = 59). Die besten Prädiktoren zur Vorhersage der *Anzahl Züge* waren das Geschlecht ($\beta = -.50$, $p < .01$, $c = -.76$), die aufgabenspezifische Selbstwirksamkeit vorher ($\beta = .30$, $p < .01$, $c = .35$) und das Alter ($\beta = -.28$, $p < .01$, $c = -.53$). Die Varianzaufklärung mit diesen Prädiktoren betrug $R = .62$ (korrigiertes R Quadrat = $.36$). Die *Anzahl Fehler* wurden mit der relativen Häufigkeit negativer selbstbezogener Gedanken am besten vorhergesagt ($\beta = .33$, $p < .01$, $R = .33$, korrigiertes R Quadrat = $.10$, $c = 1.0$). Die *Absolute Problemlöseleistung* wurde von der Aktivierung vor der Aufgabe am besten vorhergesagt ($\beta = .26$, $p < .05$, $R = .26$, korrigiertes R Quadrat = $.05$, $c = 1.0$). Die *Relative Problemlöseleistung* wurde von dem Geschlecht ($\beta = .42$, $p < .01$, $c = .80$), von der aufgabenspezifischen Selbstwirksamkeit vorher ($\beta = -.29$, $p < .05$, $c = -.32$) und von der relativen Häufigkeit negativer selbstbezogener Gedanken ($\beta = -.24$, $p < .05$, $c = -.59$) vorhergesagt ($R = .56$, korrigiertes R Quadrat = $.28$). Wie *effizient* die Versuchspersonen arbeiteten, wurde von dem Geschlecht ($\beta = .29$, $p < .01$, $c = .58$), der Geschwindigkeit ($\beta = -.52$, $p < .01$, $c = -.27$), der relativen Häufigkeit negativer selbstbezogener Gedanken ($\beta = -.42$, $p < .01$, $c = -.56$), der Bedingung ($\beta = .32$, $p < .01$, $c = .20$) und der positiven Selbsteinschätzung ($\beta = .27$, $p < .05$, $c = .11$) vorhergesagt ($R = .64$, korrigiertes R

= .27, $p < .05$, $c = .11$) vorhergesagt ($R = .64$, korrigiertes R Quadrat = .36). Die *Optimale Effizienz* wurde vom Geschlecht ($\beta = .48$, $p < .01$, $c = .71$), der relativen Häufigkeit negativer selbstbezogener Gedanken ($\beta = -.30$, $p < .01$, $c = -.63$), der Geschwindigkeit ($\beta = -.29$, $p < .01$, $c = -.17$) und dem vorher eingeschätzten Aufgabenstandard Anzahl Fehler ($\beta = .26$, $p < .05$, $c = -.17$) vorhergesagt ($R = .65$, korrigiertes R Quadrat = .37).

Frauen (n = 32). Die *Anzahl Fehler* wurde am besten von der Semesteranzahl ($\beta = .41$, $p < .05$, $c = 1.0$) vorhergesagt ($R = .41$, korrigiertes R Quadrat = .14). Die potentesten Prädiktoren zur Vorhersage der *Anzahl Züge* waren das Alter ($\beta = -.54$, $p < .01$, $c = -.77$) und die Bedingung ($\beta = -.44$, $p < .01$, $c = -.61$, $R = 0.69$, korrigiertes R-Quadrat = .43). Die *Absolute Problemlöseleistung* wurde von dem Alter der Versuchsperson ($\beta = -.58$, $p < .01$, $c = -.76$) und der Aufmerksamkeit auf die Fehlervermeidung ($\beta = -.39$, $p < .05$, $c = -.27$) vorhergesagt ($R = .55$, korrigiertes R Quadrat = .25). Die *Relative Problemlöseleistung* wurde von der Bedingung ($\beta = .58$, $p < .01$, $c = .71$), dem Alter ($\beta = .40$, $p < .01$, $c = .48$) und der Häufigkeit negativer selbstbezogener Gedanken ($\beta = -.38$, $p < .01$, $c = -.37$) vorhergesagt (R gesamt = .75, korrigiertes R-Quadrat = .51). Bei der *Effizienz* gab es drei potente Prädiktoren: die Bedingung ($\beta = .55$, $p < .01$, $c = .69$), die Häufigkeit negativer selbstbezogener Gedanken ($\beta = -.41$, $p < .01$, $c = -.47$) und die Geschwindigkeit ($\beta = -.32$, $p < .05$, $c = -.23$; R gesamt = .64, korrigiertes R-Quadrat = .35). Die *Optimale Effizienz* wurde von der Aktivierung vor der Aufgabe ($\beta = .42$, $p < .05$, $c = 1.0$) vorhergesagt ($R = .42$, korrigiertes R Quadrat = .24).

Männer (n = 27). Die *Anzahl Züge* konnte bei den Männern mit 5 Prädiktoren sehr gut vorhergesagt werden ($R = .90$, korrigiertes R Quadrat = .77): mit der aufgabenspezifischen Selbstwirksamkeit vorher ($\beta = .61$, $p < .01$, $c = .59$), mit den sonstigen Gedanken ($\beta = -.53$, $p < .01$, $c = .26$), mit dem Nachdenken über die Aufgabe ($\beta = -.53$, $p < .01$, $c = -.28$), dem vorher gesetzten persönlichen Aufgabenstandard Fehler ($\beta = .46$, $p < .01$, $c = .43$) und der Bedingung ($\beta = .23$, $p < .05$, $c = .20$). Die *Relative Problemlöseleistung* wurde von der Aufmerksamkeit auf die Geschwindigkeit ($\beta = -.53$, $p < .01$,

$c = -.81$) und der Semesteranzahl ($\beta = -.36$, $p < .05$, $c = -.46$) vorhergesagt ($R = .59$, korrigiertes R Quadrat = .29). Die *Effizienz* der Aufgabenbearbeitung wurde von der Aufmerksamkeit auf die Geschwindigkeit ($\beta = -.42$, $p < .05$, $c = -1.0$) vorhergesagt ($R = .42$, korrigiertes R Quadrat = .14).

6 Diskussion

6.1 Übersicht

Nach der Zusammenfassung der Ergebnisse versuche ich, mögliche Ursachen und Erklärungen für die gewonnenen Resultate darzulegen. Danach stelle ich die Stärken und Schwächen der Studie heraus um dann anschließend über die Grenzen der Studie hinaus Vorschläge für die weitere Forschung zu erarbeiten. Zum Schluss meiner Arbeit möchte ich noch die praktische Relevanz der vorliegenden Ergebnisse hervorheben.

6.2 Zusammenfassung der Ergebnisse

6.2.1 Hypothesen und explorative Ergebnisse

Hypothesen. Die Gruppe, die gelobt wurde, freute sich nicht mehr und war nicht aktiver als die nicht gelobte Gruppe (Hypothese 1 und 2 nicht bestätigt). Je größer (kleiner) die (Vor-) Freude der Versuchspersonen war, desto weniger (mehr) sonstige Gedanken hatten sie (Hypothese 5 widerlegt). Die empfundene Aktivierung der Versuchspersonen zeigte keinen Zusammenhang zu der aufgabenbezogenen Aufmerksamkeit (Hypothese 6 nicht bestätigt). Die gelobten Personen hatten genauso viele potentiell aufgabenförderliche Gedanken, Gedanken zur Aufgabenschwierigkeit, sonstige Gedanken und negative, selbstbezogene Gedanken wie die nicht gelobten Personen. Sie verwendeten auch genau so viel Zeit zum Nachdenken über die Aufgabe und fokussierten ihre Aufmerksamkeit gleich lang auf die Aufgabenziele (Hypothese 3 größtenteils nicht bestätigt). Personen, die gelobt wurden, verglichen ihre Leistung weniger stark mit der Leistung anderer (Teilwiderlegung Hypothese 3).

Lob wirkte sich nicht auf die Wahl des persönlichen Aufgabenstandards aus (Hypothese 4 nicht bestätigt). Nur die Personen mit relativ hoher aufgabenspezifischer

Selbstwirksamkeit, die gelobt wurden gaben sich nach der Bearbeitung der Aufgabe tendenziell mit höheren Zugzahlen zufrieden (Aufgabenstandard „Züge“, Hypothese 7 zum Teil widerlegt). Die Auswirkung von Lob auf die Art der Gedanken wurde größtenteils nicht von aufgabenspezifischer Selbstwirksamkeit moderiert (Hypothese 9 größtenteils nicht bestätigt). Personen mit hoher aufgabenspezifischer Selbstwirksamkeit äußerten nach Lob mehr positive, selbstbezogene Gedanken als die Kontrollgruppe. Personen mit niedriger aufgabenspezifischer Selbstwirksamkeit äußerten nach Lob weniger positive, selbstbezogene Gedanken als die Kontrollgruppe (Teilwiderlegung Hypothese 9). Lob wirkte sich nicht (weder positiv noch negativ) auf die Leistung aus (Hypothese 10 nicht bestätigt). Nur die gelobten Frauen benötigten weniger Züge, arbeiteten effizienter und hatten eine bessere Relative Problemlöseleistung als die nicht gelobten Frauen (Hypothese 10 bei den Frauen widerlegt).

Ob sich Lob über den Ort der Aufmerksamkeit der Versuchspersonen auf die Arbeitsleistung auswirkt, konnte aufgrund fehlender Voraussetzungen nicht überprüft werden (Mediatorhypothese 11). Weder wirkte sich Lob auf den Ort der Aufmerksamkeit aus (Ausnahme: die gelobten Personen verglichen sich weniger mit anderen, Skala „Leistungsbewertung“), noch auf die Leistung. Die Variable „Ort der Aufmerksamkeit“ hing mit verschiedenen Leistungsmaßen zusammen und qualifizierte sich als potentieller Mediator: Je mehr (weniger) Gedanken zur *Aufgabenbearbeitung* gemacht wurden, desto weniger (mehr) Fehler und Züge wurden gemacht und desto besser (schlechter) war die Relative Problemlöseleistung, die Effizienz und die Optimale Effizienz. Je mehr (weniger) *sonstige Gedanken* gemacht wurden desto mehr (weniger) Fehler und Züge wurden gemacht und desto ineffizienter (effizienter) arbeiteten die Versuchspersonen und desto schlechter (besser) war die Optimale Effizienz. Je mehr (weniger) *negative selbstbezogene Gedanken* gemacht wurden, desto mehr (weniger) Fehler und Züge wurden gemacht und desto schlechter (besser) war die Leistung (die Relative Problemlöseleistung, die Effizienz und die Optimale Effizienz). Je mehr (weniger) die Versuchspersonen über die Aufgabe *nachdachten*, desto weniger (mehr) Züge machten sie und desto besser (schlechter) war die Leistung (die Relative Problemlöseleistung, die Effizienz und die Optimale Effizienz).

Explorative Ergebnisse. Die *Freude* und die *Aktivierung* hingen mit dem **Ort der Aufmerksamkeit** zusammen. Je größer (kleiner) die (Vor)-Freude der Versuchspersonen, desto weniger (mehr) sonstige Gedanken hatten sie und desto weniger (mehr) fokussierten sie ihre Aufmerksamkeit darauf, die Aufgabe möglichst schnell zu lösen. Je mehr (weniger) positive selbstbezogene Gedanken gemacht wurden, je weniger (mehr) Aufmerksamkeit die Versuchspersonen der Geschwindigkeit der Problemlösung widmeten und je niedriger (höher) die Versuchspersonen ihre Leistung auf der Skala Negative Selbsteinschätzung einstuften, desto größer war die Freude nach der Aufgabe. Je mehr (weniger) die Versuchspersonen über die Aufgabe nachdachten, desto größer (geringer) war ihre Aktivierung nach der Aufgabe.

Die vor der Aufgabe eingeschätzte *aufgabenspezifische Selbstwirksamkeit*, die *Freude*, die *Aktivierung* und die *Aufgabenstandards* hingen mit der **nachfolgenden Leistung** zusammen. Je höher (niedriger) die aufgabenspezifische Selbstwirksamkeit vor der Aufgabe eingeschätzt wurde, desto mehr (weniger) Züge machten die Versuchspersonen und desto besser (schlechter) war die absolute Problemlöseleistung (tendenziell). Je größer (kleiner) die Aktivierung vor der Bearbeitung der Aufgabe war, desto besser (schlechter) waren alle berechneten Leistungsmaße (absolute und Relative Problemlöseleistung, die Effizienz und die Optimale Effizienz). Je höher (niedriger) der Aufgabenstandard Zeit vor der Aufgabe gesetzt wurde, desto besser (schlechter) war die absolute Problemlöseleistung. Je höher (niedriger) der Aufgabenstandard Fehler gesetzt wurde vor der Aufgabe, desto mehr (weniger) Fehler machten die Versuchspersonen.

Andererseits hing die **vollbrachte Leistung** auch mit den nach der Bearbeitung eingeschätzten, oben aufgezählten Variablen zusammen. Je mehr (weniger) Zeit und Anzahl Züge benötigt wurden und je schlechter (besser) die Leistung in den berechneten Leistungsmaßen (außer der absoluten Problemlöseleistung) war, desto geringer (höher) war die *aufgabenspezifische Selbstwirksamkeit* nach der Aufgabe. Das gleiche galt für die Freude nach der Bearbeitung der Aufgabe: Je mehr (weniger) Zeit und Anzahl Züge benötigt wurden und je schlechter (besser) die Leistung in den berechneten Leistungsmaßen war, desto weniger (mehr) *freuten* sich die Versuchspersonen nach der Aufgabe. Bei den Aufgabenstandards wirkte sich nur die Anzahl Züge aus: je

mehr (weniger) Züge benötigt wurden, desto höher (niedriger) wurde der Aufgabenstandard Züge gesetzt.

6.2.2 Geschlecht

Wenn im folgenden Geschlechtseffekte und Geschlechtsunterschiede berichtet werden, muss bei der Interpretation dieser Ergebnisse beachtet werden, dass die weiblichen Versuchspersonen von einer Versuchsleiterin gelobt wurden und die männlichen Versuchspersonen von einem Versuchsleiter. Die Versuchsleiter unterschieden sich aber hinsichtlich der Glaubwürdigkeit ihres Lobes (Tabelle 6) und ihres Vorgehens (Anhang A2) nicht.

Generelle Geschlechtsunterschiede (Haupteffekte). Geschlechtsunterschiede gab es hinsichtlich der *Leistung*. Frauen benötigten mehr Züge, machten mehr Fehler, hatten eine schlechtere Relative Problemlöseleistung, arbeiteten ineffizienter und hatten eine schlechtere Optimale Effizienz. Geschlechtsunterschiede gab es hinsichtlich des *Orts der Aufmerksamkeit*. Männer äußerten mehr potentiell aufgabenförderliche Gedanken und mehr Gedanken zur Aufgabenschwierigkeit. Frauen äußerten mehr sonstige Gedanken und mehr negative selbstbezogene Gedanken. Männer versuchten tendenziell eher, das Problem durch nachdenken zu lösen und hatten tendenziell höhere Werte in ihrer positiven Selbsteinschätzung. Männer hatten in der Arbeitsperiode nach der Manipulation weniger aufgabenbezogene Gedanken als vor der Manipulation, während Frauen eher mehr aufgabenbezogene Gedanken nach der Manipulation äußerten. Dafür hatten die Männer nach der Manipulation mehr sonstige Gedanken als vor der Manipulation, während die Frauen nach der Manipulation weniger sonstige Gedanken hatten als vor der Manipulation.

Männer schätzten ihre *aufgabenbezogene Selbstwirksamkeit* vor der Aufgabe tendenziell und nach der Bearbeitung signifikant höher ein als die Frauen. In der neutralen Bedingung senkten die Frauen ihre aufgabenbezogene Selbstwirksamkeit signifikant ab, während sie in der neutralen Bedingung unverändert blieb. Bei den Männern blieb sie in jeder Bedingung unverändert. Tendenziell *aktivierter* vor der Aufgabe

waren die Frauen. Weiterhin gaben sich Frauen mit einer *höheren Anzahl Fehlern* (Angabe vor der Bearbeitung) zufrieden als Männer. Nach der Bearbeitung gaben sich die Frauen mit einer längeren *Zeitdauer* und mehr *Zügen* bis zur Problemlösung zufrieden.

Hypothesen und Geschlecht. Frauen neigten dazu, bei niedriger (hoher) aufgabenspezifischer Selbstwirksamkeit vorher eher mehr (wenig) positive selbstbezogene Gedanken zu äußern (Teilbestätigung *Hypothese 8*). Männer äußerten bei hoher (niedriger) aufgabenspezifischer Selbstwirksamkeit mehr (weniger) positive selbstbezogene Gedanken (Teilwiderlegung *Hypothese 8*). Je höher (niedriger) die aufgabenspezifische Selbstwirksamkeit von den Männern vor der Aufgabe eingeschätzt wurde, desto mehr (weniger) sonstige Gedanken äußerten sie (Teilwiderlegung *Hypothese 8*). Je mehr (weniger) die Aufmerksamkeit der Männer auf die Aufgabenschwierigkeit gelenkt wurde, je mehr (weniger) die Aufmerksamkeit auf Negative Selbsteinschätzungen und je weniger (mehr) die Aufmerksamkeit auf positive selbstbezogene Gedanken gelenkt wurde, desto schlechter (besser) fiel die Einschätzung der aufgabenspezifischen Selbstwirksamkeit nach der Aufgabe aus.

Hypothese 9 ergab signifikante Dreifachinteraktionen Geschlecht mal aufgabenspezifische Selbstwirksamkeit vorher mal Bedingung auf die abhängigen Variablen negative selbstbezogene Gedanken, die Leistungsbewertung und die sonstigen Gedanken (tendenziell signifikant) nach der Manipulation. Bei den *Frauen* wird die Auswirkung von Lob auf die Verschiebung des Aufmerksamkeitsfokus von der aufgabenspezifische Selbstwirksamkeit vorher teilweise in die gewünschte Richtung moderiert. Wie vorhergesagt haben Frauen mit niedriger aufgabenspezifischer Selbstwirksamkeit vorher nach Lob mehr negative selbstbezogene Gedanken als Frauen mit hoher aufgabenspezifischer Selbstwirksamkeit vorher. Sie verglichen sich jedoch weniger häufig mit anderen (Teilwiderlegung *Hypothese 9*). Bei den *Männern* wird die Auswirkung von Lob auf die Verschiebung des Aufmerksamkeitsfokus von der aufgabenspezifische Selbstwirksamkeit vorher *nicht* in die gewünschte Richtung moderiert. Entgegen der Vorhersage hatten Männer mit niedriger Selbstwirksamkeit vorher nach

Lob tendenziell weniger sonstige Gedanken als Männer mit hoher aufgabenspezifischer Selbstwirksamkeit.

Nur bei den weiblichen Versuchspersonen zeigte sich eine *Verbesserung der Leistung durch Lob*: die gelobten Frauen machten weniger Züge, hatten eine bessere Relative Problemlöseleistung und arbeiteten effizienter an der „Turm von Hanoi“ Aufgabe als die Frauen, die nicht gelobt wurden (Teilwiderlegung *Hypothese 10*). Eine Interaktion zwischen dem Geschlecht und der Bedingung zeigte sich bei der Anzahl Züge und der Effizienz: Frauen hatten bei beiden Maßen schlechtere Werte als Männer, durch Lob verbesserten sich die Werte fast bis auf das Niveau der Männer.

Geschlechtsspezifische Zusammenhänge. Geschlechtsspezifische Zusammenhänge gab es zwischen dem *Ort der Aufmerksamkeit und der Leistung*. Je mehr (weniger) die *Frauen* sich auf die Fehlervermeidung während der Bearbeitung der Aufgabe konzentrierten, desto weniger (mehr) Fehler machten sie. Bei den *Männern* ergab sich kein Zusammenhang. Je mehr (weniger) die *Frauen* sich auf die Geschwindigkeit der Bearbeitung konzentrierten, desto weniger (mehr) Züge machten sie, desto mehr (weniger) Zeit brauchten sie und desto schlechter (besser) war die absolute Problemlöseleistung. Je mehr (weniger) die *Männer* sich auf die Geschwindigkeit der Bearbeitung konzentrierten, desto mehr (weniger) Züge machten sie, desto schlechter (besser) war ihre Relative Problemlöseleistung, ihre Effizienz und ihre Optimale Effizienz. Je mehr (weniger) die *Frauen* ihre Leistung über den Vergleich mit anderen bewerteten, desto ineffizienter (effizienter) arbeiteten sie. Je höher (niedriger) die Positive Selbsteinschätzung bei den *Frauen* ausfiel, desto schneller (langsamer) waren sie. Je höher (niedriger) die Positive Selbsteinschätzung bei den Männern ausfiel, desto langsamer (schneller) waren sie. Je höher (niedriger) die Negative Selbsteinschätzung bei den *Frauen* ausfiel, desto schneller (langsamer) waren sie.

Geschlechtsspezifische Zusammenhänge gab es hinsichtlich der *Zusammenhänge zwischen der Freude und der Aktivierung und dem Ort der Aufmerksamkeit*. Je größer (kleiner) die Freude der *Frauen* vor der Aufgabe war, desto niedriger (höher) fiel die Negative Selbsteinschätzung aus. Je größer (kleiner) die Freude der *Männer* vor der Aufgabe war, desto mehr (weniger) dachten sie über die Aufgabe nach. Je

mehr (weniger) Aufmerksamkeit positiven selbstbezogenen Gedanken gewidmet wurde, desto größer (kleiner) war die Freude der *Frauen* nach der Aufgabenbearbeitung. Je mehr (weniger) Aufmerksamkeit dem Nachdenken über die Aufgabe gewidmet wurde, desto größer (kleiner) war die Freude der *Männer* nach der Aufgabenbearbeitung. Je höher (niedriger) die Aktivierung der *Frauen* vor der Aufgabe war, desto weniger (mehr) Aufmerksamkeit wurde negativen selbstbezogenen Gedanken gewidmet. Je höher (niedriger) die Aktivierung der *Männer* vor der Aufgabe war, desto weniger (mehr) Aufmerksamkeit wurde den sonstigen Gedanken gewidmet. Je mehr (weniger) Aufmerksamkeit die *Frauen* in die Aufgabenbearbeitung steckten, desto kleiner (größer) war die Aktivierung danach. Je mehr (weniger) Aufmerksamkeit die *Frauen* den sonstigen Gedanken widmeten, desto größer (kleiner) war die Aktivierung danach. Je mehr (weniger) die Männer über die Aufgabe nachdachten, desto größer (kleiner) war die Aktivierung danach.

Nur bei den Männern gab es **signifikante Zusammenhänge zwischen der aufgabenspezifischen Selbstwirksamkeit vorher und der Leistung**. Je größer (kleiner) die aufgabenspezifische Selbstwirksamkeit vorher eingeschätzt wurde, desto mehr (weniger) Züge machten die Männer und desto besser (schlechter) war ihre Absolute Problemlöseleistung. Je größer (kleiner) die aufgabenspezifische Selbstwirksamkeit vorher eingeschätzt wurde, desto mehr (weniger) Zeit benötigten sie jedoch und desto schlechter (besser) war ihre Relative Problemlöseleistung.

6.2.3 Vorhersage der Leistung

Die Varianzaufklärung bezieht sich auf den geschätzten wahren Wert in der Population, also auf das korrigierte R Quadrat (Schrumpfungskorrektur berücksichtigt, Bortz, 1999; S. 436ff). Die Interpretation der Einflussrichtung der Prädiktoren auf die Leistung bezieht sich auf das Vorzeichen der Strukturkoeffizienten. Der jeweils beste Prädiktor unabhängig von den anderen Prädiktorvariablen wurde über die Größe des Strukturkoeffizienten identifiziert (Bortz, 1999; S. 438f).

Gesamter Datensatz. Für die Vorhersage der *Anzahl Züge* wurde das Geschlecht, die aufgabenspezifische Selbstwirksamkeit vorher und das Alter zu signifikanten Prädiktoren. Die wenigsten Züge machten junge Männer mit niedriger aufgabenspezifischer Selbstwirksamkeit. Es konnte mit diesen drei Variablen 36 Prozent der Varianz aufgeklärt werden. 10 Prozent der Varianz innerhalb der *Anzahl Fehler* konnte mit den negativen selbstbezogenen Gedanken aufgeklärt werden. Je mehr (weniger) negative selbstbezogene Gedanken die Personen äußerten, desto mehr (weniger) Fehler machten sie. Nur 5 Prozent der Varianz konnte innerhalb der *Absoluten Problemlöseleistung* durch die Aktivierung vor der Aufgabe aufgeklärt werden. Je höher (niedriger) die Aktivierung der Versuchspersonen vor der Aufgabe war, desto besser (schlechter) war die absolute Problemlöseleistung. 28 Prozent der Varianz konnte innerhalb der *Relativen Problemlöseleistung* durch das Geschlecht, der aufgabenspezifischen Selbstwirksamkeit vorher und den negativen selbstbezogenen Gedanken aufgeklärt werden. Die beste Relative Problemlöseleistung hatten Männer mit niedriger aufgabenspezifischer Selbstwirksamkeit vorher und wenig negativen selbstbezogenen Gedanken. 36 Prozent der Varianz konnte innerhalb der Effizienz durch das Geschlecht, die Geschwindigkeit, die relative Häufigkeit negativer selbstbezogener Gedanken, die Bedingung und durch die Positive Selbsteinschätzung aufgeklärt werden. Am *effizientesten* arbeiteten gelobte Männer, die wenig Aufmerksamkeit der Geschwindigkeit widmeten, wenig negative selbstbezogene Gedanken hatten und über eine Positive Selbsteinschätzung verfügten. 37 Prozent der Varianz wurde innerhalb der optimalen Effizienz von dem Geschlecht, den negativen selbstbezogenen Gedanken, der Geschwindigkeit und dem vorher eingeschätzten Aufgabenstandard Anzahl Fehler aufgeklärt. Die optimalste Effizienz hatten Männer mit wenig negativen selbstbezogenen Gedanken, die wenig Aufmerksamkeit in die Geschwindigkeit der Bearbeitung investierten und ihren persönlichen Aufgabenstandard Fehler niedrig ansetzten.

Um die Leistung mit nur einem Prädiktor am besten vorherzusagen, ist bei der *Anzahl Züge*, der *Relativen Problemlöseleistung*, der *Effizienz* und der *Optimalen Effizienz* das Geschlecht als Prädiktor vorzuziehen. Fast genauso potent als Prädiktor wie das Geschlecht war die relative Häufigkeit negativer selbstbezogener Gedanken bei der Vorhersage der *Effizienz* und der *Optimalen Effizienz*.

Frauen. 14 Prozent der Varianz innerhalb der *Fehleranzahl* konnte durch die Semesteranzahl aufgeklärt werden. Am meisten Fehler machten die höheren Semester. 43 Prozent der Varianz innerhalb der *Anzahl Züge* wurde durch das Alter der Versuchspersonen und der Bedingung aufgeklärt. Am wenigsten Züge machten die jüngeren Frauen, die gelobt wurden. 25 Prozent der Varianz innerhalb der *Absoluten Problemlöseleistung* wurde von dem Alter der Versuchspersonen und der Aufmerksamkeit auf die Fehlervermeidung aufgeklärt. Jünger Frauen, die wenig Aufmerksamkeit der Fehlervermeidung widmeten, hatten die beste absolute Problemlöseleistung. 51 Prozent der Varianz wurde innerhalb der *Relativen Problemlöseleistung* von der Bedingung, dem Alter der Versuchspersonen und den negativen selbstbezogenen Gedanken aufgeklärt. Jüngere Frauen, die gelobt wurden und wenig negative selbstbezogene Gedanken hatten während der Bearbeitung hatten die beste Relative Problemlöseleistung. 35 Prozent der Varianz konnte innerhalb der *Effizienz* mit den Prädiktoren Bedingung, negative selbstbezogene Gedanken und Geschwindigkeit aufgeklärt werden. Am effizientesten arbeiteten Frauen, wenn sie gelobt wurden, wenig negative selbstbezogene Gedanken hatten und wenig Aufmerksamkeit der Geschwindigkeit der Aufgabenbearbeitung widmeten. Nur 24 Prozent der Varianz konnte innerhalb der *Optimalen Effizienz* durch die Aktivierung vor der Aufgabe vorhergesagt werden. Die optimalste Effizienz hatten die Frauen mit hoher Aktivierung vor der Aufgabe.

Um die Leistung bei den Frauen mit nur einem Prädiktor am besten vorherzusagen, ist bei der *Anzahl Züge* und der *Absoluten Problemlöseleistung* das Alter als Prädiktor vorzuziehen. Für die Vorhersage der *Relativen Problemlöseleistung* und der *Effizienz* ist die Bedingung am besten geeignet.

Männer. 77 Prozent der Varianz konnte innerhalb der *Anzahl Züge* mit den Prädiktoren aufgabenspezifische Selbstwirksamkeit vorher, sonstige Gedanken, Nachdenken über die Aufgabe, dem vorher gesetzten Aufgabenstandard Fehler und der Bedingung aufgeklärt werden. Männer mit hoher aufgabenspezifischer Selbstwirksamkeit vorher, die viel sonstige Gedanken äußerten und wenig über die Aufgabe nachdachten, sich vor der Aufgabe einen hohen persönliche Aufgabenstandard Fehler setzten und gelobt

wurden, machten am meisten Züge. 29 Prozent der Varianz konnte innerhalb der *Relativen Problemlöseleistung* von der Aufmerksamkeit auf die Geschwindigkeit und der Semesteranzahl aufgeklärt werden. Die schlechteste Relative Problemlöseleistung hatten die Männer, die viel Aufmerksamkeit der Geschwindigkeit widmeten und eine hohe Semesterzahl aufwiesen. 14 Prozent der Varianz konnte innerhalb der *Effizienz* von der Aufmerksamkeit auf die Geschwindigkeit aufgeklärt werden. Am effizientesten arbeiteten die Männer, die wenig Aufmerksamkeit der Geschwindigkeit widmeten.

Um die Leistung bei den Männern mit nur einem Prädiktor am besten vorherzusagen, ist bei der *Anzahl Züge* die aufgabenspezifische Selbstwirksamkeit als Prädiktor vorzuziehen. Für die Vorhersage der *Relativen Problemlöseleistung* und der *Effizienz* ist die Aufmerksamkeit auf die Geschwindigkeit am besten geeignet.

6.3 Mögliche Ursachen und Erklärungen für die Ergebnisse

Vorraussetzung aller Erklärungen für die Ergebnisse ist, dass der Versuchsleiter von den Versuchspersonen als ehrlich und glaubwürdig eingeschätzt wurde (vgl. Tabelle 6).

Hypothese 1 und 2. Auf den ersten Blick ist es schwierig nachzuvollziehen, dass die Personen, die gelobt wurden, sich nicht mehr freuten und auch nicht aktiver waren als die Personen, die nicht gelobt wurden. Eine theoriekonforme Erklärung (Feedback Intervention Theory, Kluger & DeNisi, 1996) für diese Ergebnisse wäre, dass das Lob *nicht positiver war als die persönliche Einschätzung* der Versuchsperson. Wenn das Lob die subjektive Einschätzung der Versuchsperson nicht übertrifft, freut sie sich nicht und ist auch nicht aktiver.

Wenn das Lob die eigene Einschätzung jedoch übertraf, warum freuten sich die gelobten Versuchspersonen dann nicht über das Lob des Versuchsleiters? Die Glaubwürdigkeit des Lobs wurde von den Versuchspersonen größtenteils bestätigt. Was ist, wenn das Lob des Versuchsleiters von den Versuchspersonen zwar als glaubwürdig angesehen wurde (vgl. Tabelle 6), die Versuchspersonen aber trotzdem nicht richtig

vorangekommen sind bei der Aufgabe? Die Versuchspersonen wären doppelt enttäuscht über ihre Leistung: Sie erfüllten die indirekt durch das Lob ausgedrückte, hohe *Erwartung des Versuchsleiters* nicht zur genüge (nur neun Personen bewältigten die Aufgabe; Versuchsleiter sagte: „Du eignest dich gut für diese Sorte Aufgaben“) und ihre eigene Erwartung auch nicht. Tatsächlich zeigte sich in Tabelle 19, dass je weniger (mehr) Zeit benötigt und Fehler gemacht wurden und je besser (schlechter) die Leistung (absolute und Relative Problemlöseleistung, Effizienz und Optimale Effizienz) war, desto mehr (weniger) freuten sich die Versuchspersonen. Dieses Ergebnis stimmt überein mit Resultaten aus der Forschung zur Zielsetzung: Personen freuen sich, wenn sie ihre *Ziele erreichen* und sind enttäuscht, wenn sie sie nicht erreichen (Latham & Locke, 1991). Nur 9 von 67 Versuchspersonen lösten das Problem in der vorgegebenen Zeit. Der Großteil der Versuchspersonen war demnach trotz des Lobs enttäuscht.

Basch und Fischer (2000) konnten zeigen, dass Anerkennung von Vorgesetzten und Kollegen zu positiven Emotionen führt. Die Anerkennung, die der Versuchsleiter in Form eines Lobs aussprach, wird meiner Meinung nach erst durch zwei Aspekte zu einer *wirklichen Anerkennung* für die Versuchsperson: Erstens durch das Verhältnis zwischen Versuchsperson und Versuchsleiter (kein richtiger Vorgesetzter oder richtiger Kollege, sondern nur Student an der gleichen Universität; unbekannte, unwichtige Person...); und zweitens der Konsequenzen (positiv wie negativ), die mit der Erfüllung oder Nicht-Erfüllung der Aufgabe einhergehen (Intelligenztest, hoher Geldbetrag, persönliches Feedback ...).

Ein weiterer Punkt könnte meiner Meinung nach die *Kompetenzeinschätzung des Versuchsleiters* durch die Versuchsperson sein: je kompetenter der Versuchsleiter eingeschätzt wird, desto mehr bedeutet das Lob aufgrund der richtigen Einschätzung der Situation durch den Versuchsleiter. Wie das ehrlich gemeinte Lob bei den Versuchspersonen ankommt, hängt von der jeweiligen *Attribution des Lobs* durch die Versuchsperson ab: Meyer (1992) stellte fest, das Lob vom Lobempfänger auch als Ausdruck niedriger Erwartung in seine Fähigkeiten gewertet werden kann. Der Versuchsleiter „will mich wahrscheinlich nur aufbauen weil er sieht wie schlecht ich bin“.

Eine derartige Interpretation des Lobs wirkt sich natürlich nicht gerade positiv auf die Stimmung oder Aktivierung der Versuchsperson aus.

Entscheidend für die Höhe der Aktivierung nach Lob soll das Ausmaß der Diskrepanz zur eigenen Erwartung bzw. Einschätzung sein (Kluger et al., 1994). Die Versuchspersonen wurden nach zwei Minuten unabhängig von ihrer Leistung während der Bearbeitung der Aufgabe gelobt. Was ist, wenn die Versuchsperson bis zu diesem Zeitpunkt noch *keine Erwartung* hatte, sondern einfach „nur so“ an der Aufgabe arbeitete? Durch die Einschätzung der aufgabenspezifischen Selbstwirksamkeit vor dem Beginn der Aufgabe sollte sich bei den Versuchspersonen eine Erwartungshaltung aufgebaut haben. Inwieweit die Versuchspersonen trotzdem gleichgültig gegenüber ihrer Leistung und Lob des Versuchsleiters waren, ist schwer abzuschätzen. Die fehlende Aktivierung durch Lob lässt sich mit den gleichen Ursachen zusammenbringen wie die fehlende Freude: das Verhältnis zum Versuchsleiter, die Kompetenzeinschätzung des Versuchsleiters, die Attribution des Lobs durch die Versuchsperson.

Hypothese 3. Lob lenkte in diesem Experiment die Aufmerksamkeit nicht auf das Selbst. Das Lob beeinflusste die Verteilung der Aufmerksamkeit in Form der Art der Gedanken nicht. Die einzige Auswirkung des Lobs war, dass die gelobten Personen ihre Leistung weniger stark mit der Leistung anderer verglichen, um zu erfahren, wie gut sie sind. Die Personen, die nicht gelobt wurden, beschäftigten sich mehr mit sich selbst und ihrer Leistung aufgrund der (Selbst) – Unsicherheit, wie gut sie sind. Durch Lob wird diese Selbstunsicherheit durch Information über die Versuchsperson befriedigt, sie kann aufhören danach zu suchen. Lob führt die Personen davon weg, sich zu sehr mit sich selbst und ihrer Leistung zu beschäftigen. In diesem Punkt wurde die Hypothese widerlegt.

Eine theoriekonforme Erklärung für die ansonsten fehlende Auswirkung ist, dass die Versuchspersonen dem Lob zuwenig Aufmerksamkeit schenkten. Basisannahme 5 geht davon aus, dass eine Feedback Intervention meistens ernsthafte Konsequenzen für den Feedbackempfänger hat und er deshalb der Intervention beträchtlich Aufmerksamkeit schenkt. Die Versuchsleiter im Experiment saßen ein Stück hinter den Versuchspersonen und lobten ohne ersichtlichen Grund nach zwei Minuten mitten

während des Problemlöseprozesses. Betrachtet man die Kommentare der Denkprotokolle direkt nach dem Lob oder dem neutralen Statement, fällt auf, dass viele Versuchspersonen die Manipulation mit ein bis zwei Bemerkungen abtaten (zum Beispiel: „Ha, ungeahnte Talente, ähm, dahin, dann dahin...“). Für die Versuchspersonen war das Lob (oder der neutrale Kommentar) eher hinderlich für das weiterarbeiten und hatte für sie wahrscheinlich nicht die Bedeutung, die nötig wäre für eine andere Aufmerksamkeitsverteilung. Wichtiger (relevanter für das Selbst) wäre den Versuchspersonen vermutlich das Lob von einem Dozenten, Kommilitonen oder von einer vertrauten Person; oder wenn das Meistern der Aufgabe mit hohen kognitiven Fähigkeiten einherginge beziehungsweise aus einem anderen Grund relevant für das Selbst der Person wäre. In diesem Experiment hatte das Lob vermutlich keine Selbstrelevanz.

Als andere Ursache für den fehlenden Effekt wäre denkbar, dass auch selbstrelevanteres Lob während der Arbeit, das unabhängig von der Leistung ist, prinzipiell nur kurz Aufmerksamkeit erhält, und die Aufmerksamkeit sofort wieder auf die Aufgabe gelenkt wird. Die Bedeutung einer neuen Verteilung der Aufmerksamkeit aufgrund eines Metaaufgabenprozesses, der durch Lob in Gang gesetzt wird, scheint hier zumindest fraglich.

Eventuell könnte die Höhe des Interesses an der Aufgabe (intrinsische Motivation) eine Rolle spielen. Man könnte annehmen, dass die Beeinflussung von außen *während* der Bearbeitung einer Aufgabe weniger ins Gewicht fällt je höher das Interesse an der Aufgabe selbst ist. Fast alle Versuchspersonen wollten auch nach dem offiziellen Ende des Experiments die Lösung der Aufgabe wissen oder sie noch fertig machen. Das spricht für eine hohe intrinsische Motivation der Versuchspersonen.

Dass die gelobten Personen sich weniger darum kümmern, wie gut andere sind (Leistungsbewertung), kann als indirekter Beleg dafür dienen, dass jeder aktiv nach Feedback sucht aus den verschiedensten Gründen (Ashford & Cummings, 1983). Ein Grund ist der Wunsch, zu Wissen, wie gut man eine Aufgabe macht (Unsicherheitsreduktion). Durch die Formulierung des Lobs „Du eignest dich sicher gut für diese Sorte Aufgabe“ wurde die Unsicherheit über die Leistungsbewertung der Versuchsperson reduziert und die aktive Feedbacksuche verringert. Die dadurch freigewordene Aufmerksamkeitskapazität könnten die Versuchspersonen auf die Aufgabe verwenden. In

dieser Studie war die Leistungsbewertung durch den Vergleich mit anderen unabhängig von der Leistung der Versuchsperson (Tabelle 15).

Hypothese 4. Keiner der Aufgabenstandards (Zeit, Züge, Fehler) wurde nach Lob verändert. Eine Veränderung wurde nur dann vorhergesagt, wenn das Lob die Aufmerksamkeit auf das Selbst lenkt und Selbstziele salient werden. Diese Voraussetzung wurde in dem vorliegenden Experiment nicht erfüllt. Das Lob lenkte die Aufmerksamkeit nicht auf das Selbst, die Versuchspersonen empfanden es nicht selbstzielrelevant. Relevant für das Selbst einer Person könnte Lob durch die bei H3 genannten Gründe werden.

Hypothese 5 und 6. Die Freude und die sonstigen Gedanken hingen nicht positiv zusammen (vgl. Kluger, 2001), sondern negativ. Die Aktivierung hing nicht wie vermutet mit der Häufigkeit aufgabenbezogener Gedanken zusammen. Neben methodischen Problemen, das Konstrukt Aufmerksamkeit von Studie zu Studie vergleichbar zu erfassen (vgl. Methodenkritik, Messung der Aufmerksamkeit), muss die Idee der symmetrischen und asymmetrischen Prozesse hinsichtlich der Zusammenhänge mit den jeweiligen Kognitionen aufgrund der vorliegenden Datenlage in Frage gestellt werden. Je mehr sich die Versuchspersonen vor der Bearbeitung der Aufgabe freuten, desto weniger sonstige Gedanken hatten sie. Wenn die Freude vor dem Beginn einer Aufgabe als hohe intrinsische Motivation, die Aufgabe zu lösen, gewertet wird, kann dieses Ergebnis leicht erklärt werden: die Versuchspersonen mit Freude an der Aufgabe machten sich weniger sonstige Gedanken, weil die Aufgabe sie fesselte. Dass die Aktivierung vorher nicht mehr aufgabenbezogene Gedanken hervorbrachte, kann über die Art und Weise erklärt werden, wie sich Aktivierung äußert: je mehr die Personen in diesem Experiment über die Aufgabenlösung nachdachten, desto aktiver waren sie am Ende der Aufgabe (Tabelle 20). Dieser positive Zusammenhang zwischen Aktivierung und Nachdenken über die Aufgabenlösung ist vor allem bei den Männern zu beobachten (Tabelle 22).

Hypothese 7. Personen mit hoher aufgabenspezifischer Selbstwirksamkeit gaben sich nur nach Lob mit höheren Zugzahlen zufrieden (d. h. mit einem niedrigeren Aufgabenstandard). Die anderen Aufgabenstandards waren bei beiden Bedingungen gleich. Man könnte die Anzahl Züge, mit der die Personen noch zufrieden sind, als selbstzielrelevant bezeichnen: je weniger Züge man braucht, desto gewiefter ist man. Die theoriekonforme Erklärung würde dann lauten, dass die Diskrepanz zwischen Selbstkonzept und Sollselbst durch das Lob verkleinert wurde: Das Sollselbst - Ziel wurde für die Personen mit hoher aufgabenspezifischer Selbstwirksamkeit zum Teil schon erfüllt. Die Folge war eine Anpassung des Aufgabenstandards an die tatsächlich zu erwartende Leistung. Offen bleibt die Frage, warum gerade Personen mit hoher aufgabenspezifischer Selbstwirksamkeit ihren Aufgabenstandard senkten. Eventuell verfügen diese Personen über einen Schutzmechanismus in der Form einer Selbsttäuschung, die den hohen Selbstwert aufrechterhält: kann der Standard nicht erreicht werden, dann wird er so verändert, dass er erreicht werden kann und nicht etwa der Eindruck der Unfähigkeit entsteht. Das Lob bestätigte ihre guten *Fähigkeiten*, ihre Leistung kann ergo gleich bleiben und nur ihr Aufgabenstandard muss verändert werden. Personen mit niedriger Selbstwirksamkeit sehen ihre *Unfähigkeit* durch die Diskrepanz von Leistung zu persönlichem Aufgabenstandard und durch das Lob bestätigt (vgl. das beschriebene Attributionsmuster bei H9: Lob als Aufmunterung). Sie ergeben sich sozusagen in ihr Schicksal, ihren Aufgabenstandard aus Unfähigkeit nicht erreichen zu können.

Hypothese 8. Je niedriger die Personen ihre aufgabenspezifische Selbstwirksamkeit einschätzten, desto mehr gaben sie an, viel Zeit und Aufmerksamkeit in negativer Selbsteinschätzung investiert zu haben. Die 20 Personen mit der niedrigsten aufgabenspezifischen Selbstwirksamkeit hatten mehr als doppelt so viele negative selbstbezogene Gedanken wie die Personen mit hoher aufgabenspezifischer Selbstwirksamkeit. Die Hypothese hat sich in diesem Punkt bestätigt. Dass Personen mit niedriger aufgabenspezifischer Selbstwirksamkeit nicht mehr positive selbstbezogene Gedanken haben als Personen mit hoher, ist einleuchtend. Bei der Formulierung der Hypothese wurde nur an negative selbstbezogene Gedanken und Selbsteinschätzung gedacht. In

dieser Form war die Hypothese hinsichtlich der selbstbezogenen Gedanken zu wenig ausdifferenziert. Das gleiche gilt für die aufgabenbezogenen Gedanken. Für jede Kategorie (Aufgabenbearbeitung und Aufgabenschwierigkeit) und für jede Subskala des Fragebogens (Nachdenken, Fehlervermeidung, Geschwindigkeit) sollte eine spezifische Hypothese formuliert werden. Dasselbe gilt auch für das Geschlecht: je niedriger (höher) beispielsweise die aufgabenspezifische Selbstwirksamkeit vorher bei den Frauen war, desto mehr (weniger) positive selbstbezogene Gedanken äußerten sie; bei den Männern war es genau umgekehrt (je höher (niedriger) die aufgabenspezifische Selbstwirksamkeit, desto mehr (weniger) positive selbstbezogene Gedanken). Es sollten also auch geschlechtsspezifische Hypothesen abgeleitet werden.

An dem Extremgruppenvergleich (Tabelle 12) ist gut zu sehen, dass die Mittelwerte immerhin in die vorhergesagte Richtung gehen (nicht signifikant): die 20 Personen mit der niedrigsten aufgabenspezifischen Selbstwirksamkeit haben ca. 5 Prozentpunkte weniger aufgabenbezogene Gedanken und ca. 2.5 Prozentpunkte mehr sonstige Gedanken.

Hypothese 9. Da Lob keine Verschiebung der Aufmerksamkeit mit sich brachte (Ausnahme: Leistungsbewertung) ist es nicht weiter verwunderlich, dass größtenteils kein Moderatoreffekt der aufgabenspezifischen Selbstwirksamkeit auf den Ort der Aufmerksamkeit zu finden war. Zu erklären wäre umso mehr der Interaktionseffekt Bedingung mal aufgabenspezifische Selbstwirksamkeit vorher auf die positiven selbstbezogenen Gedanken bei gleichzeitigem Fehlen der Haupteffekte: weder aufgabenspezifische Selbstwirksamkeit vorher noch die Bedingung wirkte sich auf die Häufigkeit positiver selbstbezogener Gedanken aus. Eine Erklärung für dieses Ergebnis lieferten Kluger und DeNisi (1996) im Einklang mit einer Interpretation der Kontrolltheorie (Cropanzano, James, & Citera, 1993): Wenn das Feedback kongruent mit dem Selbstwertgefühl ist, dann wird die Aufmerksamkeit vermehrt auf das Selbst gelenkt. Bei negativem Feedback wird die Aufmerksamkeit demnach nur dann auf das Selbst gelenkt, wenn die betreffende Person ein niedriges Selbstwertgefühl besitzt; bei positivem Feedback nur bei Personen mit hohem Selbstwertgefühl. Damit lässt sich das Ergebnis von Abbildung 3 erklären: Die Personen mit hoher aufgabenspezifischer

Selbstwirksamkeit vorher haben ohne Lob fast gar keine positiven selbstbezogenen Gedanken. Werden sie gelobt, schnellst die Häufigkeit auf über 5 Prozent aller Gedanken, die nach der Manipulation geäußert wurden. Bei Personen mit niedriger aufgabenspezifischer Selbstwirksamkeit vorher ist es umgekehrt: ohne Lob sind über 3 Prozent aller Gedanken positiv selbstbezogen, mit Lob nur noch ca. 1 Prozent.

Warum steigt nur bei Kongruenz die Häufigkeit selbstbezogener Gedanken? Warum haben die Versuchspersonen dann nicht mehr negative selbstbezogenen Gedanken? Eine Antwort auf die erste Frage könnte wieder die Attributionstheorie liefern: Personen, die eine geringe aufgabenspezifische Selbstwirksamkeit haben attribuieren das Lob als „Hilfestellung, Ermunterung“; Personen mit hoher aufgabenspezifischer Selbstwirksamkeit als Bestätigung ihrer Fähigkeiten. Eine Antwort auf die zweite Frage könnte sein, dass die negativen selbstbezogenen Gedanken einhergehen und bedingt werden von der selbsteingeschätzten Güte der Leistung, während die positiven selbstbezogenen Gedanken wenig mit der Leistung zu tun haben und eventuell von anderen Faktoren (wie Lob) beeinflusst werden. Tabelle 15 bestätigt diese Vermutung.

Interessant wird die Diskussion der Ergebnisse von Hypothese 9 wenn man das *Geschlecht* als weiteren Faktor hinzuzieht. Bei Frauen wurde die Aufmerksamkeit durch Lob abhängig von der aufgabenspezifischen Selbstwirksamkeit in die vorhergesagte Richtung gelenkt, bei Männern nicht. Das bedeutet, dass die Verschiebung des Aufmerksamkeitsfokus tatsächlich durch Lob stattfindet. Bei der Vorhersage der Richtung der Verschiebung muss jedoch neben der aufgabenspezifischen Selbstwirksamkeit das Geschlecht beachtet werden. Vermutlich verwenden Frauen andere Attributionsmuster wie Männer. Dadurch könnte die geschlechtsspezifisch verschiedene Moderation des Lobs durch die aufgabenspezifische Selbstwirksamkeit erklärt werden.

Hypthese 10. Warum wirkte sich Lob im Gesamtdatensatz nicht auf die Leistung aus? Es wäre zu einfach zu behaupten, die Versuchsleiter wären schlechte Schauspieler gewesen und hätten unglaublich gelobt. Eine getrennte Auswertung der Frauen und Männer zeigte, dass nur die gelobten weiblichen Versuchspersonen weniger Züge benötigten, eine bessere Relative Problemlöseleistung hatten und effizienter arbeiteten. Die veränderte Fragestellung lautet deshalb: Warum verbesserte sich die Leistung

bei den Frauen und warum wirkte sich bei den Männern Lob überhaupt nicht auf die Leistung aus? Die Feedback Intervention Theory macht keine Aussagen über Geschlechtsunterschiede. Roberts und Nolen-Hoeksema (1989, Studie 2; zitiert in Johnson & Helgerson, 2002) fanden in einer experimentellen Studie mit kognitiven Aufgaben heraus, dass nur bei Frauen nach positivem Feedback („Das hast du gut gemacht. Deine Leistung ist über dem Durchschnitt“) das Selbstvertrauen, die Aufgabe zu schaffen, stieg, bei Männern nicht. Frauen waren auch zufriedener mit ihrer Leistung nach positivem Feedback als nach negativem Feedback („Das hast du nicht gut gemacht. Deine Leistung ist unter dem Durchschnitt“). Frauen unterschätzten ihre Leistung nach eigener Einschätzung, während Männer ihre Leistung überschätzten.

Roberts und Nolen-Hoeksema (1994; zitiert in Johnson & Helgerson, 2002) untersuchten die Einschätzung auf Feedback nach einer Rede. Die Einschätzung der Frauen über die Qualität ihrer Rede wurde nach positivem Feedback noch positiver und nach negativem Feedback noch negativer. Die Einschätzung der Männer wurde weniger beeinflusst von dem Feedback, obwohl sie eine niedrigere Leistungserwartung vor der Rede hatten als die Frauen. Eine mögliche Erklärung der Autoren für die höhere Beeinflussbarkeit der Frauen durch Feedback war, dass Frauen Feedback mehr als Information über ihre Fähigkeiten gewichteten als Männer. Tatsächlich sahen die Frauen die Rückmeldung akkurater an als die Männer.

Lenney (1977; zitiert in Johnson & Helgerson, 2002) fasste die Literatur über Geschlechtsunterschiede hinsichtlich des Selbstbewusstseins zusammen. Ergebnis war, dass Frauen vor allem bei typisch maskulinen und leistungsorientierten Aufgaben weniger selbstbewusst waren als Männer. Beyer (1999, zitiert in Johnson & Helgerson, 2002) fand heraus, dass dieser Geschlechtsunterschied durch eine Überschätzung ihres Selbstbewusstseins bei Männern und eine Unterschätzung ihres Selbstbewusstseins bei Frauen zu erklären war.

Jack und Dill (1992, zitiert in Johnson & Helgerson, 2002) erklären die erhöhte Beeinflussbarkeit durch Feedback über das geringere Selbstbewusstsein der Frauen und eine dadurch erhöhte externalisierte Selbstwahrnehmung. Frauen sind auf Informationen von außen angewiesen, wenn sie etwas über ihre Fähigkeiten in Erfahrung bringen wollen. Deshalb nehmen sie Feedback ernster (bzw. persönlicher) als Männer.

Im vorliegenden Datensatz schätzten die Männer ihre aufgabenspezifische Selbstwirksamkeit vorher tendenziell höher ein als die Frauen. Männer und Frauen unterschieden sich auch in der Veränderung ihrer Einschätzung der aufgabenspezifischen Selbstwirksamkeit vorher zu nachher, abhängig von der Bedingung (Abbildung 7): in der neutralen Bedingung senkten die weiblichen Versuchspersonen ihre Einschätzung um 0.5 Punkte, in der Lobbedingung nur um 0.1 Punkte. Männer veränderten ihre Einschätzung nicht (siehe 5.4.3). Diese Ergebnisse bestätigen indirekt die Ergebnisse der genannten Studien: Frauen waren weniger selbstbewusst als Männer (Selbsteinschätzung der aufgabenspezifischen Selbstwirksamkeit vor der Aufgabe) und durch Feedback mehr in ihrer Selbsteinschätzung beeinflussbar als Männer. Mangels Informationen von außen (neutrale Bedingung) korrigierten die Frauen ihre aufgabenspezifische Selbstwirksamkeit nach unten.

Erklärt dieser Unterschied in der aufgabenspezifischen Selbstwirksamkeit auch den Unterschied in der Leistung zwischen gelobten Frauen und nicht gelobten Frauen? Nein, bei der Vorhersage der Leistung (vgl. 5.4.4) wurde bei den Frauen die aufgabenspezifische Selbstwirksamkeit vorher kein Prädiktor der Leistung. Die Bedingung war in fast allen Fällen der potenteste Prädiktor. Eventuell erklärt die besondere Beziehung zwischen der Versuchsleiterin und den weiblichen Versuchspersonen die Unterschiede: die Versuchspersonen bauten in der kurzen Zeit ein Vertrauensverhältnis auf und fühlten sich wohler. Das Arbeitsklima wurde durch das Lob eventuell so verbessert, dass es sich direkt auf die Leistung auswirkte. Johnson und Helgerson (2002) fanden heraus, dass Frauen mehr als Männer mit dem ihnen gegebenen Feedback übereinstimmen und es deshalb annehmen können. Dieses Ergebnis spricht für ein gutes Arbeitsklima bei den gelobten Frauen.

Hypothese 11. Die Frage, ob der Ort der Aufmerksamkeit ein Mediator ist zwischen Lob und der Leistung konnte aufgrund mangelnder Vorraussetzungen nicht beantwortet werden. Es konnte aber gezeigt werden, dass sich der Ort der Aufmerksamkeit auf die Leistung auswirkt (Tabelle 15). Welche Potenz die Aufmerksamkeitsvariablen und die Bedingung bei der Vorhersage der Leistung haben, wurde in 6.2.3 beschrieben.

Hypothesen und Geschlecht. Dass Männer bei hoher aufgabenspezifischer Selbstwirksamkeit vorher mehr positive selbstbezogene Gedanken äußerten, während Frauen bei niedriger aufgabenspezifischer Selbstwirksamkeit vorher mehr positive selbstbezogene Gedanken äußerten (Geschlechtsspezifität von *Hypothese 8*), hat vermutlich zwei verschiedene Gründe. Männer, die sich sicher in ihrer Sache sind, prahlen auch damit und sparen nicht mit „Eigenlob“. Frauen versuchen eventuell durch viele positive selbstbezogene Kommentare einerseits ihre Unsicherheit zu verbergen und andererseits sich selbst Hoffnung zu machen.

Dass nur Frauen mit niedriger aufgabenspezifischer Selbstwirksamkeit nach Lob mehr *negative selbstbezogene Gedanken* äußerten als Frauen mit hoher aufgabenspezifischer Selbstwirksamkeit (Geschlechtsspezifität von *Hypothese 9*), hängt meiner Meinung nach von der subjektiven Attribution des Lobs ab. Welche Ursachen für das Lob herangezogen werden hängt vermutlich vom Geschlecht und der aufgabenspezifischen Selbstwirksamkeit ab. Frauen mit niedriger aufgabenspezifischer Selbstwirksamkeit vorher sehen die Ursache des Lobs in ihrer eigenen Unfähigkeit und interpretieren das Lob deswegen als Aufmunterung und gleichzeitig Bestätigung ihrer Unfähigkeit. Frauen mit hoher aufgabenspezifischer Selbstwirksamkeit sehen die Ursache des Lobs in ihren Fähigkeiten und interpretieren das Lob als Bestätigung ihrer eigenen Einschätzung.

Dass Männer mit hoher aufgabenspezifischer Selbstwirksamkeit nach Lob tendenziell mehr *sonstige Gedanken* äußerten als Männer mit niedriger aufgabenspezifischer Selbstwirksamkeit, könnte wieder an der Attribution des Lobs liegen. Männer mit hoher aufgabenspezifischer Selbstwirksamkeit sehen ihre Fähigkeiten bestätigt, sie brauchen nicht mehr ihre ganze Aufmerksamkeit der Aufgabe widmen. Also schweifen sie ab. Männer mit niedriger aufgabenspezifischer Selbstwirksamkeit trauen sich die Aufgabe jetzt mehr zu, bündeln noch einmal ihre Aufmerksamkeit auf die Aufgabe. Also reduziert sich die Häufigkeit sonstiger Gedanken.

Dass nur die Frauen mit relativ niedriger aufgabenspezifischer Selbstwirksamkeit vorher sich nach Lob *weniger mit anderen verglichen* als die Frauen ohne Lob könnte an der Unsicherheitsreduktion liegen (vgl. Erklärung Hypothese 9). Frauen sind sich vermutlich ihrer Leistung besonders unsicher, wenn sie sich die Aufgabe

nicht zutrauten und tatsächlich nicht richtig vorankommen. Lob reduziert diese Unsicherheit substantiell und damit das Bedürfnis, ihre Leistung durch den Vergleich mit anderen einschätzen zu müssen.

Die ausschließlich bei den Männern gefundenen **Zusammenhänge zwischen der aufgabenspezifischen Selbstwirksamkeit vorher und der Leistung** sind durch eine Selbstüberschätzung der Fähigkeiten der Männer zu erklären. Weil sie denken, sie könnten die Aufgabe, machen sie viele Züge und kommen vermutlich durch Zufall entsprechend weit (Absolute Problemlöseleistung). Entscheidend für die Bewertung der Leistung ist aber auch die Relative Problemlöseleistung. Die Männer mit hoher aufgabenspezifischer Selbstwirksamkeit sind zwar der Problemlösung näher gerückt, haben aber zu viele Züge benötigt. Deshalb war die Relative Problemlöseleistung der Männer mit hoher aufgabenspezifischer Selbstwirksamkeit schlecht.

6.4 Methodenkritik: Stärken und Schwächen der Studie

Experimentelle Situation. Das Problem der Künstlichkeit der experimentellen Situation hat sich nicht übermäßig negativ auf die Glaubwürdigkeit des Lobs ausgewirkt (vgl. 5.1.4, Tabelle 6). Auch in Betrieben, bei denen der Vorgesetzte seine Mitarbeiter nicht so gut kennt und noch keine Vertrauensbasis herrscht zwischen Mitarbeiter und Vorgesetzten, wird sicher nicht jedes Lob des Vorgesetzten als glaubwürdig, gerechtfertigt, ernst gemeint etc. angesehen. Podsakoff und Farr (1989) konnten zeigen, dass Versuchspersonen sich keine höheren Ziele steckten und keine bessere Leistung zeigten wenn sie glaubwürdiges positives Feedback erhalten (Feedback beruhe auf dem Vergleich mit 300 anderen Studierenden) im Vergleich mit weniger glaubwürdigem positivem Feedback (Feedback basiere auf eigener Erfahrung). Bei negativem Feedback wirkte sich die Glaubwürdigkeit sowohl auf die Leistung als auch die Zielsetzung aus. Auch in diesem Experiment wirkte sich die Glaubwürdigkeit des Lobs nicht signifikant auf die abhängigen Variablen aus (Anhang B3b, Tabelle 2). Kritisch anzumerken ist, dass der Versuchsleiter im Vergleich zu der Versuchsleiterin tendenziell

unechter bei den Versuchspersonen ankam. Das kann am Versuchsleiter selbst liegen oder daran, dass Männer es nicht gewohnt sind, von einem anderen Mann gelobt zu werden und deshalb besonders „hellhörig“ wurden. Weiterhin empfanden sie sein Lob tendenziell als „untertrieben“. Das kann an der Tendenz der Männer liegen, sich eher zu überschätzen (Beyer, 1999, zitiert in Johnson & Helgerson, 2002), oder tatsächlich am Versuchsleiter.

Versuchspersoneninstruktion. Durch die Formulierung in der Versuchspersoneninstruktion (Anhang A2) *„Die Auswertung erfolgt auch nicht auf Personenebene, sondern auf Gruppenebene. Die Daten des Einzelnen gehen sozusagen in der Gruppe unter“* könnte den Versuchspersonen den Ansporn zu hoher Anstrengung und Leistung genommen haben im Sinne von „Social Loafing“. Die Leistung des Einzelnen ist nicht mehr nachzuvollziehen, warum sollte sich die Versuchsperson anstrengen? Persönliche, selbstrelevante Folgen waren nicht zu erwarten. Es gab für die Versuchspersonen demnach wenig Grund, ein Lob (vermutlich auch Kritik) ernst bzw. persönlich zu nehmen.

Messung des Konstrukts Aufmerksamkeit. Das Konstrukt Aufmerksamkeit wurde auf subjektive und objektive Weise zu erfassen versucht. Zwischen subjektiven und objektiven Maßen zeigten sich jedoch wenig signifikante Korrelationen (Tabelle 5). Eine positive Korrelation zeigte sich zwischen der Leistungsbewertung (Fragebogen) und der Aufgabenschwierigkeit (Denkprotokoll), eine negative Korrelation zwischen dem Nachdenken über die Aufgabe (Fragebogen) und der negativer Selbstbewertung (Denkprotokoll). Tendenziell signifikante Korrelationen zeigten sich zwischen der Leistungsbewertung (Fragebogen) und der Aufgabenbearbeitung (Denkprotokoll); der Fehlervermeidung (Fragebogen) und der Aufgabenschwierigkeit (Denkprotokoll) und der Geschwindigkeit (Fragebogen) und der Aufgabenschwierigkeit (Denkprotokoll).

Welche Erklärungen gibt es für den geringen Zusammenhang zwischen den objektiven und subjektiven Maßen? Messen die beiden Erhebungsmethoden dasselbe Konstrukt? Ein Hauptgrund für den fehlenden Zusammenhang ist meiner Meinung nach die unterschiedliche Sichtweise und die daraus resultierende Operationalisierung

des Konstrukts „Aufmerksamkeit“. Im *vorliegenden Fragebogen* schätzten die Versuchspersonen auf einer 5 Punkte Likert-Skala für jedes Item ein, wie oft sie über eine bestimmte Sache nachdachten (bei den 9 selbstbezogenen Items). Oder sie schätzten für jedes Item ein, wie oft sie ihre ganze Aufmerksamkeit auf eine bestimmte Sache richteten (bei den 5 aufgabenbezogenen Items). Diese Fragen wurden nach der sechsminütigen Bearbeitungszeit gestellt. Die Versuchspersonen mussten nicht einschätzen, wie viel Zeit sie anteilmäßig dieser oder jener Sache innerhalb der gesamten Bearbeitungszeit widmeten. Es wurde keine Verteilung ihrer Aufmerksamkeit auf die verschiedenen Bereiche erfragt.

Eine subjektive Einschätzung, wie viel Aufmerksamkeit verschiedenen Bereichen innerhalb einer bestimmten Zeitspanne geschenkt wird, halte ich für sehr vage. Genau das kann man mit einer *objektiven Erfassung der Aufmerksamkeit* erreichen. Durch die breite Operationalisierung des Ortes der Aufmerksamkeit über die Art der Gedanken wird über die gesamte Zeitspanne hinweg verfolgt, wo der Fokus der Aufmerksamkeit sich im Moment befindet. Der schnelle Wechsel von visueller Aufmerksamkeit auf die Bauaktivität („Jetzt bau´ ich den gelben Becher auf den blauen“) hin zu akustischer Aufmerksamkeit während des Lobs („Vielen Dank, aber so gut bin ich dann doch nicht“) ist nur über diese Methode durchführbar. Auch nach innen gerichtete Aufmerksamkeit kommt zum Ausdruck („Jetzt muss ich noch mal nachdenken“).

Ein weiterer *Vorteil* der objektiven Methodik liegt in der Analogie zwischen der Auswertung der Denkprotokolle und dem Konstrukt der Aufmerksamkeit als begrenzte kognitive Ressource. Die Gesamtanzahl an Kommentaren wurde mit 100 Prozent angesetzt. Für jede Kategorie (jede Kategorie entspricht einem thematischen Bereich wie Aufgabenbearbeitung, Positive Selbsteinschätzung...) wurde der prozentuale Anteil an Kommentaren berechnet, der in diese Kategorie fiel. Sollte die Theorie stimmen, dass durch das Lenken der Aufmerksamkeit auf einen Bereich gleichzeitig dem anderen Bereich die Aufmerksamkeit entzogen wird (beziehungsweise Ressourcen fehlen), dann entspricht diese Auffassung von Aufmerksamkeit am besten der objektiven Erhebungsmethodik. Ein *Nachteil* ist, dass nicht alle Personen gleich gut ihre Gedanken, Aktivitäten und Gefühle verbalisieren können. Hier könnte der Fragebogen hilfreich sein. Ein weiterer Nachteil ist, dass lautes Denken von den Versuchspersonen

oft als „Kommentierung der Handlung“ (zum Beispiel auf die Bauaktivität, s.o.) verstanden wird. Deshalb ist im Vorfeld immer ratsam, zuerst das laute Denken mit den Versuchspersonen zu üben. Für Experimente, die die Aufmerksamkeit als begrenzte kognitive Ressource ansehen, würde ich die Methode des lauten Denkens (Ericsson & Simon, 1984) empfehlen.

Erfassung der aufgabenbezogenen Aufmerksamkeit. Bei näherer Betrachtung der einzelnen Kategorien der Denkprotokolle und der Skalen des Fragebogens fällt auf, dass innerhalb der aufgabenbezogenen Aufmerksamkeit verschiedene Bereiche durch die objektive und die subjektive Methodik abgedeckt wurden: das Denkprotokoll enthält größtenteils Kommentare über die Bauaktivität (Aufmerksamkeit auf die Handlung, fiel in die Kategorie „Aufgabenbearbeitung“) und wenig bis keine Kommentare über die Aufgabenziele (wie „Geschwindigkeit“ und „Fehlervermeidung“, im Fragebogen). Dasselbe gilt für das „Nachdenken“ über die Aufgabe: nur selten kommentierte die Versuchsperson ihr Schweigen als „Nachdenken“ (zum Bsp. „Erst mal überlegen...“ oder ähnliches). Die subjektive Methodik sollte hier als Ergänzung der objektiven Methodik betrachtet werden: das was die objektive Methodik nicht leisten kann (schweigen darf nicht als „Nachdenken über die Aufgabe“ interpretiert werden) muss mit Fragebögen abgefragt werden, um ein umfassendes Bild des Konstrukts zu bekommen. Für zukünftige Experimente sollte deshalb genau überlegt werden, welche Aspekte der aufgabenbezogenen Aufmerksamkeit abgedeckt werden sollen, um darauf aufbauend die geeignetere Messmethode zu wählen.

Erfassung der selbstbezogenen Aufmerksamkeit. Der Versuch, die Kategorie „Selbstbezogene Aufmerksamkeit“ des Denkprotokolls parallel zu den Skalen von Kanfer und Ackerman (1989) zu konstruieren, ist misslungen (im Sinne einer konvergenten Konstruktvalidierung). Die Kategorie „Eigene Leistungsbewertung“ (Denkprotokoll) musste verworfen werden, weil nur ein einziger Kommentar in diese Kategorie klassifiziert wurde. Die Skalen „Positive Selbsteinschätzung“ und „Negative Selbsteinschätzung“ des Fragebogens korrelierten nicht mit den gleichgeschalteten Kategorien „Positive und Negative Selbstbewertung“ der Denkprotokolle (Tabelle 5). Neben

den schon genannten Gründen könnte hier der „Soziale Erwünschtheitseffekt“ maßgeblich zu Verzerrungen führen: Im Fragebogen trauten sich die Versuchspersonen eher, negative wie positive selbstbezogene Gedanken über ihre Antworten zum Ausdruck zu bringen. Bei den Denkprotokollen wurden eher aufgabenbezogene, sachliche Kommentare bevorzugt geäußert ($M = 77.4$ Prozent, $SD = 18.1$).

Was messen dann die Skalen des Fragebogens im Vergleich zu den Kategorien der Denkprotokolle? Vielleicht messen tatsächlich beide Messmethoden das Konstrukt „Selbstbewertung“ (Denkprotokoll) bzw. „Selbsteinschätzung“ (Fragebogen). Doch gerade bei Einschätzungen über sich selbst liegt man oft (objektiv gesehen) falsch: für das eigene Selbstbild günstige Verzerrungen (zum Beispiel Überschätzung der eigenen Testleistung, Shepperd (1993), zitiert aus *Social Psychology*, 1996) lassen einen in einem besseren Licht erscheinen. Oder man macht sich (unbewusst) viel schlechter als man ist. Oder man hat zu hohe Ansprüche, Erwartungen an sich selbst und schätzt sich deshalb schlechter ein. Die Schere zwischen subjektiver und objektiver Einschätzung ist hier zumindest zu groß.

Erfassung der sonstigen Gedanken. Aus ökonomischen Gründen wurden die „sonstigen Gedanken“ nur objektiv erfasst. Andere Autoren (zum Bsp. Kanfer & Ackerman, 1989; Kluger, 2001) erfassten andere Gedanken über eine Vielzahl von Items, die dann zu einer Skala „sonstige Gedanken“ zusammengefasst wurden (zum Bsp. 9 Items bei Kluger, 2001). Als Maß für die Aufmerksamkeit, die nicht der Aufgabe und nicht dem Selbst gewidmet wurde, eignet sich meiner Meinung nach die objektive Methodik am besten, weil dadurch der prozentuale Anteil „Sonstiger Gedanken“ an allen Gedanken berechnet werden kann. Eine Art Aufmerksamkeitsverteilung über die gesamte Zeitspanne kann mit dieser Methodik quantifiziert werden. Ein Problem dieser Methodik ist, dass sonstige Gedanken oft nicht geäußert werden. Um ein Unterdrücken dieser Gedanken zu vermeiden, ist wahrscheinlich eine angenehme, locker Atmosphäre besonders hilfreich. Diese zu schaffen ist die Aufgabe des Versuchsleiters.

6.5 Grenzen der Studie und Vorschläge für die weitere Forschung

Labor versus Feld. Nur Studierende der Universität Konstanz nahmen an diesem Experiment teil. Es handelte sich um ein Laborexperiment mit instruierten Versuchsleitern, die die Versuchspersonen nicht kannten und etwa das Alter der Versuchspersonen hatten. Die Versuchspersonen wurden bezahlt für ihre Teilnahme. In der Realität definiert entweder der *Status* einer Person oder die *Beziehung* zwischen zwei Personen die Macht, ein Lob auszusprechen. Wird ein Lob von einem peer ausgesprochen – wie hier – fragt sich die gelobte Person automatisch, aus welchen Gründen er lobt. Es scheint an sich durch die experimentelle Situation an sich klar, dass der Versuchsleiter nicht freiwillig lobt. Er folgt einer Anweisung, ein Plan steckt dahinter. Die Vermutung der Versuchsperson über die wahre Absicht des Lobs könnte entweder Misstrauen wecken oder Raum für zwangloses arbeiten schaffen. Voraussetzung dafür ist, dass der Versuchsleiter der Versuchsperson nicht unsympathisch ist und sie keine böse Absicht hinter dem Lob vermutet. Problematisch könnte das Wissen der Versuchsperson sein, dass sie während und nach dem Experiment keinerlei Folgen für ihr Tun tragen muss: es ist im Grunde nur ein Spiel für die Versuchsperson. Das Fehlen von *ernsthaften Konsequenzen* für die Versuchsperson verbunden mit der Frage nach der Absicht des Lobs aufgrund mangelnder Beziehung zum Lobgeber stellt die externe Validität in Frage. In der Realität findet kein Lob im „Vakuum“ statt: die gelobte Person kennt den Lobgeber zumindest flüchtig. Selbst in großen Organisationen kennt man seinen höheren Vorgesetzten. Darüber hinaus ist ein Feedback „von oben“ auch immer mit ernsthaften Konsequenzen verbunden. Um diese praxisrelevanten Aspekte im Labor umzusetzen, müsste beispielsweise vor dem Experiment der Versuchsperson erzählt werden, dass es sich bei der Aufgabe um eine Intelligenzaufgabe handelt. Lob wird dann wichtig durch die damit einhergehende Signalfunktion (zum Beispiel: „Ich bin intelligent“).

Im vorliegenden Experiment wurden die Männer von einem Versuchsleiter und die Frauen von einer Versuchsleiterin gelobt. Das bedeutet, dass einerseits die Person des Versuchsleiters eine Rolle spielte und andererseits nur die Auswirkung von Lob eines gleichgeschlechtlichen Lobgebers ins Gewicht fiel. Die Generalisierung be-

schränkt sich somit auf gleichgeschlechtliche Personenkonstellationen (Feedbackgeber – Feedbacknehmer).

Vorschläge für die Forschung. Die Feedback Intervention Theory (Kluger & DeNisi, 1996) ist ein Rahmenmodell. Detaillierte Vorschläge zur Ausdifferenzierung dieses Modells liefern die vorliegenden Ergebnisse. Abbildung 14 soll als ein „Modellvorschlag zur Auswirkung von *gleichgeschlechtlichem Lob (GG Lob)* auf die *Problemlöseleistung*, basierend auf Ideen der Feedback Intervention Theory“ angesehen werden. Am wichtigsten erscheint mir die Ergänzung des Modells um das *Geschlecht (GE)*. Will man die Auswirkung von Feedback auf die Leistung vorhersagen, muss das Geschlecht als Variable integriert oder für Männer und Frauen zwei getrennte Modelle konstruiert werden: Nur bei Frauen wirkte sich Lob positiv auf die Leistung aus und qualifizierte sich als bester Prädiktor der Relativen Problemlöseleistung und der Effizienz. Um diese positive und direkte Auswirkung von Lob auf die Leistung der Frauen adäquat zu beschreiben, könnte eventuell das *Arbeitsklima* als vermittelnde Variable integriert werden. Der *Ort der Aufmerksamkeit* war entscheidend für die Leistung. Im Gesamtdatensatz war die Effizienz und die Optimale Effizienz am schlechtesten wenn sich die Versuchspersonen viel negative selbstbezogene Gedanken machten, förderlich war die Aufmerksamkeit auf die Aufgabenbearbeitung und das Nachdenken über die Aufgabe, hinderlich waren viel sonstige Gedanken. Für die weitere Forschung am Ort der Aufmerksamkeit sollte die Aufmerksamkeit auf das Selbst immer in positive und negative selbstbezogene Gedanken getrennt werden. Für die geschlechtsspezifische Forschung am Ort der Aufmerksamkeit müssen andere Schwerpunkte gelegt werden. Bei den Männern wirkte sich besonders die Aufmerksamkeit auf die Geschwindigkeit der Aufgabenbearbeitung negativ aus auf ihre Relative Problemlöseleistung und ihre Effizienz. Eventuell überfordern sich die Männer, wenn sie zu schnell die Aufgabe lösen wollen. Bei den Frauen spielte der Ort der Aufmerksamkeit für die Leistung eher eine untergeordnete Rolle, nur die Menge negativer selbstbezogener Gedanken war hinderlich für eine effiziente bzw. optimale Leistung.

Direkte Leistungsauswirkung hatte im gesamten Datensatz die Aufmerksamkeit auf die Aufgabenbearbeitung, die sonstigen Gedanken, die negativen selbstbezogenen

Gedanken und das Nachdenken über die Aufgabe; bei den Männern zusätzlich die Aufmerksamkeit auf die Geschwindigkeit der Aufgabenbearbeitung; bei den Frauen hauptsächlich die negativen selbstbezogenen Gedanken. Die Hauptaussage der Feedback Intervention Theory ist, dass durch Feedback die Aufmerksamkeitsverteilung verändert wird. In diesem Experiment konnte diese direkte Auswirkung von Lob nicht beobachtet werden. Die Aufmerksamkeitsverteilung wird vermutlich erst dann nach Lob verändert wenn zwei weitere, *moderierende Faktoren* eingeschaltet werden: *die aufgabenspezifische Selbstwirksamkeit (ASW)* und das *Geschlecht*. Erst die Interaktion aufgabenspezifische Selbstwirksamkeit mal Geschlecht veränderte die Aufmerksamkeitsverteilung durch Lob innerhalb der sonstigen und der negativen selbstbezogenen Gedanken. Lob wirkt sich abhängig vom Geschlecht und abhängig von der aufgabenspezifischen Selbstwirksamkeit jeweils unterschiedlich auf die sonstigen und negativen selbstbezogenen Gedanken aus. Um diese Auswirkung ursächlich zu beschreiben, könnte in der weiteren Forschung zwischen Lob und dem Ort der Aufmerksamkeit die *Attribution des Lobs* als Mediator eingebaut werden. Männer attribuieren Lob anders wie Frauen, Männer mit niedriger aufgabenspezifischer Selbstwirksamkeit anders als Männer mit hoher aufgabenspezifischer Selbstwirksamkeit.

Welche Variablen wirkten sich auf die leistungsrelevanten Orte der Aufmerksamkeit aus? Je mehr sich die Versuchspersonen auf die Aufgabe *freuten*, desto weniger sonstige Gedanken hatten sie und desto weniger Aufmerksamkeit schenkten sie der Geschwindigkeit der Aufgabenbearbeitung. Auch das *Geschlecht* wirkte sich direkt aus auf die Aufgabenbearbeitung und auf die Häufigkeit sonstiger und negativer selbstbezogener Gedanken (Frauen mehr als doppelt so viele wie Männer). Männer dachten mehr über die Aufgabe nach als Frauen. Die *aufgabenspezifische Selbstwirksamkeit* wirkt sich vermittelt über das Geschlecht auf die Aufmerksamkeitsverteilung aus: Frauen äußerten beispielsweise bei niedriger, Männer bei hoher aufgabenspezifischer Selbstwirksamkeit mehr positive selbstbezogene Gedanken

Direkt positiv auf die Problemlöseleistung wirkt die *Aktivierung*. Vermittelt über das Geschlecht wirkt sich die *aufgabenspezifische Selbstwirksamkeit* auf die Problemlöseleistung aus. Nur bei Männern herrscht ein signifikanter Zusammenhang. Trauen sie sich die Aufgabe zu, kommen sie zwar der Problemlösung an sich näher als wenn

sie sich die Aufgabe nicht zutrauen. Jedoch benötigen sie mehr Züge und hatten insgesamt eine schlechtere relative Problemlöseleistung.

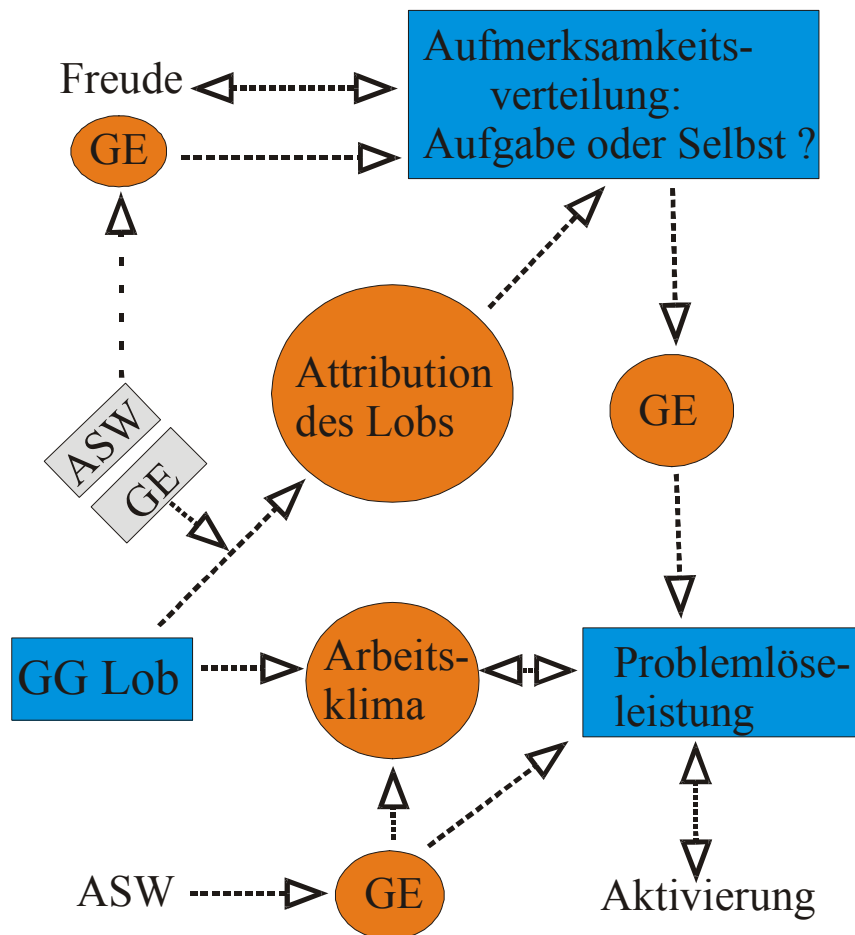


Abbildung 14: Vorschlag eines Modells zur Auswirkung von gleichgeschlechtlichem Lob auf die Problemlöseleistung, basierend auf Ideen der Feedback Intervention Theory (Kluger & DeNisi, 1996).

Nicht im Modell, aber denkbar wäre auch eine Auswirkung auf die Attribuierung des Lobs durch: die Kompetenzeinschätzung des Lobgebers durch den Lobnehmer, den Status des Lobgebers, durch die Beziehung zwischen Lobgeber und Lobnehmer und das daraus resultierende Vertrauensverhältnis und eventuell die Dauer der Beziehung.

6.6 Praktische Relevanz dieser Studie

Dass Frauen nach dem Lob von einer weiblichen Versuchsleiterin eine bessere Relative Problemlöseleistung aufweisen und effizienter an der Lösung des Problems arbeiten spricht für eine bessere Arbeitsatmosphäre nach Lob. Frauen scheinen sensibler auf eine verbesserte Arbeitsatmosphäre zu reagieren wie Männer. Frauen senkten ohne Lob ihre aufgabenspezifische Selbstwirksamkeit ab, das bedeutet sie trauen sich diese Aufgabe beim nächsten Mal weniger zu. Diese Hauptergebnisse lädt zumindest weibliche Vorgesetzte, Lehrer wie Mütter dazu ein, ihre Untergebenen, Schülerinnen oder Töchter gerade während der Bearbeitung einer anspruchsvollen, herausfordernden Tätigkeit zu loben. Bei Männern wirkte sich Lob von einem männlichen Versuchsleiter immerhin nicht negativ aus. Aus dieser Sicht kann die amerikanische Lobkultur, bei der amerikanische Manager versuchen über Lob ihren Mitarbeitern Anerkennung und Wertschätzung zu vermitteln, nur unterstützt werden. Diese Strategie scheint mir auch langfristig gesehen wirkungsvoller zu sein, weil durch die vermittelte Wertschätzung unter anderem die Bindung des Mitarbeiters an sein Unternehmen steigt. Jedoch ist es nicht ratsam, uneingeschränkt und immer während der Bearbeitung einer anspruchsvollen Aufgabe zu loben. Noch nicht abgesichert ist das Ergebnis, das nur Frauen, die sich die Aufgabe zutrauen und nur Männer, die sich die Aufgabe nicht zutrauen von gleichgeschlechtlichem Lob profitieren: Männer machen sich tendenziell weniger sonstige Gedanken und Frauen machen sich weniger negative selbstbezogene Gedanken. Hinderlich für die Bearbeitung einer Aufgabe ist gleichgeschlechtliches Lob für Frauen, die sich die Aufgabe nicht zutrauen und für Männer, die sich die Aufgabe zutrauen: Männer machen sich tendenziell mehr sonstige Gedanken und Frauen machen sich mehr negative selbstbezogene Gedanken. Je weniger negative und sonstige Gedanken sich Personen bei der Bearbeitung kniffliger Problemlöseaufgaben machen, desto effizienter und optimaler ist die Leistung. Traut sich beispielsweise ein *Schüler* eine Matheaufgabe nicht zu, kann Lob von dem *Mathelehrer* hilfreich für ihn sein. Traut sich jedoch eine *Schülerin* eine Matheaufgabe nicht zu, ist Lob von der *Mathelehrerin* vermutlich fehl am Platze. Genau andersrum verhält es sich, wenn sich ein *Schüler* eine Matheaufgabe zutraut: der *Lehrer* sollte eher nicht loben während der

Bearbeitung der Aufgabe. Traut sich jedoch eine *Schülerin* eine Matheaufgabe zu, kann sie ruhig während der Bearbeitung der Aufgabe von der *Lehrerin* gelobt werden.

Lob reduziert bei Männern wie Frauen das Bedürfnis, ihre Leistung durch den Vergleich mit anderen zu bewerten. Lobt der Vorgesetzte seine Mitarbeiter im Großraumbüro, werden sie automatisch weniger auf die Leistung ihrer Kollegen achten. Sie konzentrieren sich vermehrt auf ihre Arbeit. Das schafft wahrscheinlich eine angenehme Arbeitsatmosphäre. Besonders in Großraumbüros scheint diese Auswirkung von Lob bedeutsam zu sein. Fast jeder kennt die unangenehme Nervosität, die einen befällt vor der Lösung schwieriger Aufgaben. Diese ist aber positiv zu bewerten: sie hilft nicht nur bei der Lösung der Aufgabe, sondern hilft auch, effizienter und mit weniger Fehlern zu arbeiten. Die Freude vor der Bearbeitung einer Aufgabe wirkt sich indirekt günstig auf die Lösung eines Problems aus: je mehr man sich freut vor (und während) der Arbeit, desto weniger lässt man seine Gedanken schweifen und desto weniger will man einfach schnell fertig werden.

6.7 Abschluss (Konklusion)

Diese Studie versuchte, die Auswirkung von positivem Feedback während der Aufgabe auf die Aufgabenleistung und die ausgelösten Prozesse zu untersuchen. Neu war die Idee, eine nach der Feedback Intervention Theory postulierte Verschiebung des Orts der Aufmerksamkeit punktgenau objektiv über das laute Denken der Versuchspersonen während der Bearbeitung zu erfassen. Die Denkprotokolle konnten gut als aufgabenbezogene, selbstbezogene und sonstige Gedanken als Indikator für den Ort der Aufmerksamkeit kategorisiert werden. Neu war auch, die weiblichen Versuchspersonen nur von einem weiblichen Versuchsleiter und die männlichen Versuchspersonen nur von einem männlichen Versuchsleiter loben zu lassen. Beide Versuchsleiter wurden als glaubwürdig eingeschätzt. Das Ergebnis, dass nur gelobte Frauen effizienter arbeiten eröffnete neue Sichtweisen bezüglich der internen Strukturierung von Mann und Frau und einem damit einhergehenden anderen Umgangs mit Lob.

7 Literaturverzeichnis

Abele, A. (1995). Stimmung und Leistung. Göttingen: Hogrefe.

Albright, M. D., Levy, P. E., & Williams, J. R. (1992). Reactions to feedback discrepancies from multiple raters. Paper presented at the 7th annual convention of the Society for Industrial and Organizational Psychology, Montreal, Quebec, Canada.

Ammons, R. B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. Journal of General Psychology, *54*, 279-299.

Ashford, S. J., & Cummings, L. L. (1983). Feedback as an individual resource: Personal strategies of creating information. Organisational Behaviour and Human Performance, *32*, 370-398.

Azuma, H., & Cronbach, L. J. (1966). Cue-response correlations in the attainment of a scalar concept. American Journal of Psychology, *79*, 3-49.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, *84*, 191-215.

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. (1991). Social cognitive theory of self regulation. Organizational Behavior and Human Decision Processes, *50*, 248-287.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Basch, J., & Fisher, C. D. (2000). Affective events-emotions matrix: A classification of work events and associated emotions. In Ashkanasy, N. M. (Ed); Haertel, C. E. (Ed). Emotions in the workplace: Research, theory, and practice (pp. 36-48).

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. Journal of Personality and Social Psychology, 46, 610-620.

Baumeister, R. F., Hutton, D. G., & Cairns, K. J. (1990). Negative effects of praise on skilled performance. Basic and Applied Social Psychology, 11, 131-148.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 5, 462-467.

Bell, B. S., & Kozlowski, S. W. J. (2002). Goal orientation and ability: Interactive effects on self-efficacy, performance, and knowledge. Journal of Applied Psychology, 87, 497-505.

Beyer, H.-G. (1986). Zur Wirkung positiver und negativer Leistungsrückmeldung auf die Problemlöseleistung, die emotionale Befindlichkeit und die physiologische Aktivierung. In E. Weiss, Studien zur experimentellen und klinischen Psychologie, Band 4, München.

Blankstein, K. R., Toner, B. B., & Flett, G. L. (1989). Test anxiety and the contents of consciousness: Thought-listing and endorsement measures. Journal of Research in Personality, 23, 269-286.

Bless, H., Mackie, D. M., & Schwarz, N. (1992). Mood effects on encoding and judgmental processes in persuasion. Journal of Personality and Social Psychology, *63*, 585-595.

Bless, H. (2000). The interplay of affect and cognition: The mediating role of general knowledge structures. In J. P. Forgas (Ed.), Feeling and thinking: The role of affect in social cognition (pp. 201-222). New York: Cambridge University Press.

Bodenhauser, G. V., Kramer, G. P., & Süsser, K. (1994). Happiness and stereotypic thinking in social judgment. Journal of Personality and Social Psychology, *66*, 621-632.

Bortz, J. (1999). Statistik für Sozialwissenschaftler. Berlin, Heidelberg, New York: Springer Verlag.

Bortz, J., & Döring, N. (2002). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin, Heidelberg, New York: Springer Verlag.

Brehm, S. S., & Kassin, S. M. (1996). Social Psychology. Boston, Toronto: Houghton Mifflin Company.

Butler, Ruth (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. Journal of Educational Psychology, *79*, 474-482.

Carroll, J. M., & Kay, D. S. (1988). Prompting, feedback and error correction in the design of a scenario machine. International Journal of Man-Machine Studies, *28*, 11-27.

Carver, C. S., & Scheier, M. F. (1981). Attention and self regulation: A control theory to human behavior. New York: Springer Verlag.

Carver, C. S., & Scheier, M. F. (1990). Origins and function of positive and negative affect: A control-process view. Psychological Review, 97, 19-35.

Christianson, S. (1992). Emotional stress and eyewitness memory: A critical review. Psychological Bulletin, 112, 284-309.

Cropanzano, R., James, K., & Citera, M. (1993). A goal hierarchy model of personality, motivation, and leadership. In L. L. Cummings & B. M. Staw (Eds.), Research in organizational behavior (Vol. 15, pp. 267-322). Greenwich, CT: JAI Press.

Delin, Catherine R., & Baumeister, Roy F. (1994). Praise: More than just social reinforcement. Journal for the Theoretical Social Behaviour, 24, 219-241.

Duncker, K. (1945). On problem solving. Psychological Monographs, 58.5. Whole No. 270.

Dwyer, D. J., Hall, J. K., Volpe, C., Cannon-Bowers, J. A., & Salas, E. (1992). A performance assessment task for examining tactical decision making under stress (Spec. Rep. No. 92-002). Orlando, FL: Naval Training Systems Center, Human Factors Division.

Easterbrook, J. A. (1959). The effect of emotion of cue utilization and the organization of behavior. Psychological Review, 66, 183-201.

Erez, M. (1977). Feedback: A necessary condition for the goal setting-performance relationship. Journal of Applied Psychology, 62, 624-627.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis. Verbal reports as data. Cambridge, Massachusetts: The MIT Press.

Eysenck, M. W. (1992). Attention and arousal: Cognition and Performance. Berlin: Springer Verlag.

Feather, N. T. (1966). Effects of prior success and failure on expectations of success and subsequent performance. Journal of Personality and Social Psychology, 3, 287-298.

Feather, N. T. (1968). Change in confidence following success or failure as a predictor of subsequent performance. Journal of Personality and Social Psychology, 13, 129-144.

Fiedler, K. (2000). Towards an integrative account of affect and cognition phenomena using the BIAS computer algorithm. In J. P. Forgas (Ed.), Feeling and thinking: The role of affect in social cognition (pp. 223-252). New York: Cambridge University Press.

Fiedler, K. (2001). Affective influences on social information processing. In J. P. Forgas (Ed.), The handbook of affect and social cognition (pp.163-185). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Fiedler, K., & Bless, H. (2001). The formation of beliefs in the interface of affective and cognitive processes. In N. Frijda, A. Manstead, & S. Bem (Eds.), The influence of emotions on beliefs. New York: Cambridge University Press.

Forgas, J. P., Ciarrochi, J. V., & Moylan, S. J. (2000b). Subjective experience and mood regulation: The role of information processing strategies. In H. Bless & J. P. Forgas (Eds.), The message within: The role of subjective experience in social cognition. Philadelphia: Psychology Press.

Forgas, J. P. (2002). Feeling and doing: Affective influences on interpersonal behavior. Psychological Inquiry, 13, 1-28.

Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunnette, & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (Vol.4, 2nd ed., pp. 271-340). Palo Alto, CA: Consulting Psychologists Press.

Guindon, R. (1990). Designing the design process: Exploiting opportunistic thoughts. Human-Computer Interaction, *5*, 305-344.

Heppner, P. P., & Petersen, C. H. (1982). The development and implications of a personal problem-solving inventory. Journal of Counseling Psychology, *29*, 66-75.

Higgins, E. T. (1989). Self-discrepancy theory: What patterns of self-beliefs cause people to suffer? In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol.22, pp. 93-136). New York: Academic Press.

Higgins, E. T. (1997). Beyond pleasure and pain. American Psychologist, *52*, 1280-1300.

Higgins, E. T. (2001). Promotion and prevention experiences: Relating emotions to non emotional motivational states. In J. P. Forgas (Ed.), The handbook of affect and social cognition (pp.186-211). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ilgen, D. R. (1971). Satisfaction with performance as a function of the initial level of expected performance and the deviation from expectations. Organizational Behavior and Human Performance, *6*, 345-361.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organization. Journal of Applied Psychology, *64*, 349-371.

Isen, A. M., & Means, B. (1983). The influence of positive affect on decision-making strategy. Social Cognition, *2*, 18-31.

Isen, A. M., & Daubman, K. A. (1984). The influence of affect on categorization. Journal of Personality and Social Psychology, *47*, 1206-1217.

Isen, A. M. (1987). Positive affect, cognitive processes and social behavior. Advances in Experimental Social Psychology, *20*, 203-253.

Isen, A. M., Daubmann, K., & Nowicki, G. (1987). Positive affect facilitates creative problem solving. Journal of Personality and Social Psychology, *52*, 1122-1131.

Isen, A. M., Daubman, K. A., & Gorgolione, J. M. (1987). The influence of positive affect on cognitive organization: Implications for education. In R. E. Snow & J. F. Marshall (Eds.), Aptitude, learning and instruction (Vol. 3, pp. 143-164). Hillsdale, NJ: Erlbaum.

Isen, A. M. (1999). Positive affect. In T. Dalgeish, & M. J. Power (Eds), Handbook of Cognition and Emotion (pp. 521-541). Chichester: John Wiley & Sons

Johnson, M. & Helgerson, V. S. (2002). Sex differences in response to evaluative Feedback: A field study. Psychology of Women Quarterly, *26*, 242-251.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integration / aptitude-treatment interaction approach to skill acquisition. Journal of Applied Psychology, *74*, 657-690.

Kluger, A. N., Lewinsohn, S., & Aiello, J. R. (1994). The influence of feedback on mood: Linear effects on pleasantness and curvilinear effects on arousal. Organizational Behavior and Human Decision Processes, *60*, 276-299.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin, 119, 254-284.

Kluger, A. N., & DeNisi, A. (1998). Feedback interventions: Toward a understanding of a double-edged sword. Current Directions in Psychological Science, 7, 67-71.

Kluger, A. N. (2001). Feedback-expectation discrepancy, arousal and locus of cognition. In M. Erez, U. Kleinbeck, & Thierry (Eds.), Work motivation in the context of a globalizing economy. Mahwah, NJ: Erlbaum.

Koestner, R., Zuckerman, M., & Koestner, J. (1987). Praise, involvement, and intrinsic motivation. Journal of Personality and Social Psychology, 53, 383-390.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. Organizational Behavior and Human Decision Processes, 50, 212-247.

Lewin, K., Dembo, T., Festinger, L., & Sears, P.S. (1944). Level of aspiration. In J. M. Hunt (Ed.), Personality and the behavior disorders (pp. 333-377). New York: Ronald Press.

Locke, E. A.; & Latham, G. P. (1990). A theory of goal setting and task performance. Englewood Cliffs, NJ: Prentice Hall.

London, Manuel, Larsen, Henrik Holt, & Thisted, Lars Nellemann (1999). Relationships between feedback and self-development. Group and Organization Management, 24, 5-27.

Lust, J. A., Celuch, K. G., & Showers, L. S. (1993). A note on issues concerning the measurement of self efficacy. Journal of Applied Social Psychology, 23, 1426-1434.

Mackie, D. M., & Worth, L. T. (1989). Cognitive deficits and the mediation of positive affect in persuasion. Journal of Personality and Social Psychology, *57*, 27-40.

Majo, V. D., & Matsumi, J. T. (1996). Think aloud statements and solutions of dysphoric persons on a social problem-solving task. Cognitive Therapy and Research, *20*, 97-113.

Mano, H. (1991). The structure and intensity of emotional experiences: Method and context convergence. Multivariate Behavioral Research, *26*, 389-411.

Martin, L. L., Ward, D. W., Achee, J. W., & Wyer, J. R., Jr. (1993). Mood as input: People have to interpret the motivational implications of their moods. Journal of Personality and Social Psychology, *64*, 317-326.

Matthews, G.; Davies, D. R., & Lees, J. L. (1990). Arousal, extraversion, and individual differences in resource availability. Journal of Personality and Social Psychology, *59*, 150-168.

Mednick, M. T., Mednick, S. A., & Mednick, E. V. (1964). Incubation of creative performance and specific associative priming. Journal of Abnormal and Social Psychology, *69*, 84-88.

Meyer, W. U. (1992). Paradoxical effects of praise and blame on perceived ability. In W. Stroebe & M. Hewstone (Eds.), European review of social psychology, *3*, 259-283. Chichester, England: Wiley.

Mikulincer, M. (1994). Human learned helplessness: A coping perspective. New York: Plenum Press.

Mikulincer, M. (1988). Reactance and helplessness following exposure to unsolvable problems: The effects of attributional style. Journal of Personality and Social Psychology, *54*, 679-686.

Mor, N., & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. Psychological Bulletin, *128*, 638-662.

Mueller, Claudia M., & Dweck, Carol S. (1998). Praise for intelligence can undermine children's motivation and performance. Journal of Personality and Social Psychology, *75*, 33-52.

Murray, N., Sujan, H., Hirt, E. R., & Sujan, M. (1990). The influence of mood on categorization: A cognitive flexibility interpretation. Journal of Personality and Social Psychology, *59*, 411-425.

Nease, AnJanette A., Mudgett, Brad O., & Quiñones (1999). Relationships among feedback sign, self-efficacy, and acceptance of performance feedback. Journal of Applied Psychology, *84*, 806-814.

Otto, J. H., & Schmitz, B. B. (1993). Veränderungen positiver Gefühlszustände durch analytische und kreative Informationsverarbeitung. Zeitschrift für experimentelle und angewandte Psychologie, *41*, 566-583.

Otto, J. H. (1999). Auswirkungen des Denkstils auf den Gefühlszustand: Ein vernachlässigter Leistungsaspekt. In M. Jerusalem & R. Pekrun (Hrsg.), Emotion, Motivation und Leistung. Göttingen; Bern; Toronto; Seattle: Hogrefe Verlag.

Petty, R. E., & Cacioppo, J. J. (1977). Forewarning, cognitive responding, and resistance to persuasion. Journal of Personality and Social Psychology, *35*, 645-655.

Podsakoff, P. M., & Farh, Jiing-Lih (1989). Effects of feedback sign and credibility on goal setting and task performance. Organizational Behavior and Human Decision Processes, 44, 45-67.

Pritchard, R. D., Jones, S. D., Roth, P. L., Stuebing, K. K., & Ekeberg, S. E. (1988). Effects of group feedback, goal setting, and incentives on organizational productivity (Monograph). Journal of Applied Psychology, 73, 337-358.

Sarason, I. G., & Stoops, R. (1978). Test anxiety and the passage of time. Journal of Consulting and Clinical Psychology, 46, 105-109.

Sarason, I. G., Sarason, B. R., Keefe, D. E., Hayes, B. E., & Shearin, E. N. (1986). Cognitive interference: Situational determinants and traitlike characteristics. Journal of Personality and Social Psychology, 51, 215-226.

Sinclair, R. C. (1988). Mood, categorization breadth, and performance appraisal: The effects of order of information acquisition and affective state on halo, accuracy, information retrieval, and evaluations. Organizational Behavior and Human Decision Processes, 42, 22-46.

Sinclair, R. C., & Mark, M. M. (1992). The influence of mood state on judgment and action: Effects on persuasion, categorization, social justice, person perception, and judgmental accuracy. In L. L. Martin, & A. Tesser (Eds.), The construction of social judgment (pp. 165-193). Hillsdale, NJ: Erlbaum.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. Psychological Bulletin, 119, 3-22.

Sonnentag, S. (1998). Expertise in professional software Design: A process study. Journal of Applied Psychology, 83, 703-715.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.

Stajkovic, Alexander D., & Luthans, Fred (1998). Self-efficacy and work-related performance: A meta-analysis. Psychological Bulletin, 124, 240-261.

Taylor, Shelley E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. Psychological Bulletin, 110, 67-85.

Thompson, Roy F., & Perlini, Arthur H. (1998). Feedback and self-efficacy, arousal, and performance of introverts and extraverts. Psychological Reports, 82, 707-716.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. Psychological Review, 94, 3-15.

Waldersee, Robert, & Luthans, Fred (1994). The impact of positive and corrective feedback on customer service performance. Journal of Organizational Behavior, 15, 83-95.

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. Psychological Bulletin, 96, 465-490.

Weary, G., Marsh, K. L., Gleicher, F., & Edwards, J. A. (1993). Social-cognitive consequences of depression. In G. Weary, F. Gleicher & K. L. Marsh (Eds.), Control motivation and social cognition. New York: Springer Verlag.

Weiner, B., Heckhausen, H., & Meyer, W. (1972). Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control. Journal of Personality and Social Psychology, 21, 239-248.

Wicklund, R. A. (1975). Objective self-awareness. In L. Berkowitz (Ed.), Advances in experimental psychology (Vol. 8, pp. 233-237). New York: Academic Press.

8 Anhang

Teil A: Materialien zur Durchführung des Experiments

- 1 Versuchsleiterfahrplan
- 2 Versuchspersoneninstruktion
- 3 Fragebogen vor „Turm von Hanoi“ und Codierung
- 4 Fragebogen nach „Turm von Hanoi“
 - (a) Experimentalbedingung mit Codierung
 - (b) Kontrollbedingung mit Codierung
- 5 Leistungsmessung „Turm von Hanoi“ und Definitionen
- 6 Spielplattform „Turm von Hanoi“

Teil B: Skalen, Auswertungsinstrumente und Tabellen

- 1 Lautes Denken
 - (a) Kategorienschema zur Auswertung der Denkprotokolle
 - (b) Beispiel eines Denkprotokolls
 - (c) Kategorisierungsbeispiel (zwei Beurteiler)
- 2 Verwendete Aufmerksamkeitsskalen mit Code
- 3 Tabellen
 - (a) Tabelle 1: Korrelationen zwischen dem Ort der Aufmerksamkeit und den Leistungsmaßen
 - (b) Tabelle 2: Korrelation der Variablen „Glaubwürdigkeit“ mit sämtlichen abhängigen Variablen

Versuchsleiterfahrplan

1. Bedingungszuweisung per Zufall (per Los)

VP1	VP2	VP3	VP4	VP5	VP6	VP7	VP8	VP9	VP10
VP11	VP12	VP13	VP14	VP15	VP16	VP17	VP18	VP19	VP20
VP21	VP22	VP23	VP24	VP25	VP26	VP27	VP28	VP29	VP30

2. Begrüßung und Einführung

„Hallo, ich bin Ich werde die nächste halbe Stunde diesen Versuch mit dir zusammen durchführen und Ansprechpartner für alles sein. Deshalb, wenn du irgendwelche Fragen hast, wende dich immer an mich. Lies aber bitte zuerst die Handanweisung, die vor dir liegt, durch.

3. Fragen der VP beantworten und Wortfindungsaufgabe vorspielen mit lautem Denken und danach durchführen (zusehen) mit anderen Buchstaben (O,U,E,E,A,I,S,D,R,N)

4. Wortfindungsaufgabe / denken klappt?

5. „Fragen vor TvH“ vorlegen

6. TvH durchführen (Kassettenrekorder an, Pausetaste nicht gedrückt?)

Auf die Regeln und Ziele der TvH Anweisung hinweisen und darauf, das jetzt die Möglichkeit zu Fragen bei Unklarheiten besteht und *keine* Fragen mehr *während*

der Bearbeitung beantwortet werden. Nach zwei Minuten loben oder neutrales Statement, weitere vier Minuten abwarten und dann abbrechen. Insgesamt also sechs Minuten Bearbeitungszeit. Während der Bearbeitung die Tabelle „Leistungsmessung TvH“ ausfüllen. Sagt die VP 15 sek nichts, sagen: „Bitte laut denken“.

TIP: Gerade Uhrzeit abwarten (z.Bsp. 13 Uhr 17) und gleich Manipulationszeit notieren (hier: 13 Uhr 19) und Abbruchzeit (hier: 13 Uhr 23)

7. Abbruch, falls VP nicht fertig wird nach 6 Minuten mit den Worten:

„Ich muß dich hier jetzt leider bei der Bearbeitung unterbrechen. Bevor du dich ein zweitesmal an die „Turm von Hanoi“ Aufgabe machst, bitte ich dich, diesen Fragebogen kurz auszufüllen.“

8. Kassettenrekorder aus und „Fragen nach TvH“ vorlegen

9. Nach Ausfüllen des Fragebogens:

- Versuchsperson danken („Vielen Dank, das du mitgemacht hast“)
- Falls sie Lust hat, kann sie nochmal den TvH machen
- Aufklärung über Zweck der Studie: nebenher einfach sagen, sie sollen zu mir kommen (auch wegen Geld...)

10. Namen eintragen lassen in Liste „Studie60“ und 5,- aushändigen

11. VP-Nummer auf Kassettenrekorder sprechen

Versuchspersoneninstruktion

Liebe Versuchsperson,

du hast dich bereit erklärt, freiwillig an einer arbeitspsychologischen Studie mitzumachen. Deine gewissenhafte Mitarbeit bei allen Aufgaben und deine ehrlichen Antworten auf alle Fragen tragen wesentlich zum Gelingen meiner Diplomarbeit bei und dadurch – hoffentlich – ein klein wenig zu neuen Erkenntnissen in der arbeitspsychologischen Forschung. Danke für deine Mitarbeit. Ich kann dir versichern, dass alle Daten, die von dir erhoben werden, absolut anonym bleiben. Deshalb wird auch kein Namen oder Adresse von dir verlangt. Die Auswertung erfolgt auch nicht auf Personenebene, sondern auf Gruppenebene. Die Daten des Einzelnen gehen sozusagen in der Gruppe unter.

Lies nun als erstes den Text unten und sag deinem Versuchsleiter Bescheid, wenn du fertig bist. Bei allen Fragen, die du hast, wende dich vertrauensvoll an deine(n) VersuchsleiterIn.

Du siehst hier vor dir den sogenannten „Turm von Hanoi“ oder „Turm des Brahmanen“. Weiterhin siehst du zwei weitere, kreisrunde Plattformen. Deine Aufgabe besteht nun darin, diesen Turm genau so, wie er hier steht, auf eine der beiden Plattformen wieder aufzubauen. Du mußt jedoch ein paar **Regeln** beachten:

1. **Du darfst immer nur *einen* Becher bewegen, nie zwei oder drei auf einmal**
2. **Du darfst immer nur einen kleineren über einen größeren Becher bauen**
3. **Du darfst die Becher auf allen Plattformen solange hin und her bewegen, bis der neue Turm fertig gestellt ist.**

Ein **Ziel** ist, **möglichst schnell fertig** zu werden. Jedoch solltest du auch versuchen, mit **möglichst wenig Zügen** und **möglichst wenig Fehlern (Regelverstöße)** den Turm wieder aufzubauen

Ich bin nicht nur an der Lösung des Turm-von-Hanoi-Problems interessiert, sondern auch wie du darauf kommst. Deshalb wäre es sehr wichtig, dass du während der Bearbeitung der Aufgabe „**laut denkst**“. „**Laut denken**“ bedeutet, mit lauter Stimme zu überlegen. Bitte sag´ alles laut, was dir während der Aufgabenbearbeitung durch den Kopf geht, wenn es dir auch unwichtig oder belanglos erscheint und nichts mit der Aufgabe zu tun hat. Es ist wichtig, dass du deine Gedanken nicht bewertest, bevor du sie laut gesagt hast.

Um sich daran zu gewöhnen, laut zu denken beziehungsweise alles laut zu sagen, was einem durch den Kopf geht während der Bearbeitung einer Aufgabe, darfst du zuerst eine **Wortfindungsaufgabe** machen und **dabei „laut denken“**. Wenn du das Gefühl hast, es fällt dir leicht, alle deine Gedanken zu verbalisieren während der Bearbeitung der Wortfindungsaufgabe, dann sag´ deinem Versuchsleiter Bescheid. Du darfst dann einen kurzen Fragebogen ausfüllen und dann geht es weiter mit dem „Turm von Hanoi“. Denk´ bitte immer daran, auch während der „Turm von Hanoi“ Aufgabe laut zu denken. Alles, was gesagt wird, wird mit einem Tonbandgerät aufgezeichnet. Danke für dein Verständnis. Stell´ während der Bearbeitung der Aufgaben bitte keine Fragen mehr an den Versuchsleiter. Wenn irgend etwas unklar ist dann frag´ den Versuchsleiter bitte **vor** der Bearbeitung der Aufgaben.

Aber jetzt zuerst die **Wortfindungsaufgabe**

Die Aufgabe lautet:

Versuche aus den zehn Buchstaben ein sinnvolles Wort zu legen mit mindestens 6 Buchstaben und dabei laut zu denken

Code	Bedeutung	Skala	Quelle	Reliabilität (α)
vas	Vorher Aufgabenstandard	keine	Pilotstudie (n =10)	Einzelitems
vsw	Vorher Selbstwirksamkeit (SW)	„Aufgaben-spezifische SW“	Thompson & Perlini, 1998	.70
vem	Vorher Emotion	„Emotion“	Mano, 1991	.75
va	Vorher Aktivierung	Eine Skala „Aktivierung“	Mano, 1991	.69
ve	Vorher „elation“			

Fragebogen vor der Bearbeitung („Turm von Hanoi“)

„Welches ist die längste Zeit bis zur Problemlösung, mit der du noch zufrieden bist?“ (vas1)	4 min	5 min	6 min	7 min	8 min
„Welches ist die maximalste Anzahl Schritte (Züge) bis zur Problemlösung, mit der du noch zufrieden bist?“ (63 Züge Minimum) (vas2)	65 Züge	70 Züge	75 Züge	80 Züge	85 Züge
„Welches ist die maximalste Anzahl Regelverstöße bis zur Problemlösung, mit der du noch zufrieden bist?“ (vas3)	1 Fehler	2 Fehler	3 Fehler	4 Fehler	5 Fehler
	trifft gar nicht zu	trifft wenig zu	trifft mit-telmäßig zu	trifft größtenteils zu	trifft vollständig zu
„Ich sehe überhaupt keine Probleme“	1	2	3	4	5

me, diese Aufgabe zu bewältigen.“ (vsw1)					
„Ich glaube, dass meine Leistung bei dieser Aufgabe unter dem Durchschnitt sein wird.“(vsw2)	1	2	3	4	5
„Ich habe Vertrauen in meine Fähigkeiten, diese Aufgabe gut zu machen.“(vsw3)	1	2	3	4	5
„Ich glaube, dass ich diese Aufgabe besser bewältige als die meisten anderen Personen.“ (vsw4)	1	2	3	4	5
Ich fühle mich im Moment...	trifft gar nicht zu	trifft wenig zu	trifft mitelmäßig zu	trifft größtenteils zu	trifft vollständig zu
... zufrieden (vem1)	1	2	3	4	5
... vergnügt (vem2)	1	2	3	4	5
... glücklich (vem3)	1	2	3	4	5
... in guter Stimmung (vem4)	1	2	3	4	5
Ich bin im Moment eher...					
... erstaunt (va1)	1	2	3	4	5
... überrascht (va2)	1	2	3	4	5
... aktiviert (von außen) (va3)	1	2	3	4	5
... freudig erregt (ve1)	1	2	3	4	5
... aktiv (ve2)	1	2	3	4	5
... aufgeregt (ve3)	1	2	3	4	5

Alter: (alter)	
Geschlecht (sex)	
Studienfach/fächer (fach)	
Semesteranzahl (sem)	

Code	Bedeutung	Skala	Quelle	Reliabilität (α)
em	Nachher Emotion	„Emotion“	Mano, 1991	.85
a	Nachher Aktivierung	Eine Skala „Akti-	Mano, 1991	.68
e	Nachher „elation“	vierung“		
as	Nachher Aufgaben- standard	keine	Pilotstudie (n =10)	Einzelitems
nsw	Nachher Selbst- wirksamkeit (SW)	„Aufgaben- spezifische SW“	Thompson & Perlini, 1998	.75
on	Aufgabenbezogene Aufmerksamkeit	„On-task attention“ Nachdenken, Ge- schwindigkeit, Feh- lervermeidung	Kanfer & Acker- man, 1989	Nachdenken: .80 Geschw. .76 Fehlerverm. .82
LB	Aufmerksamkeit auf Leistungsbeurteilung	„Leistungs- beurteilung“	Kanfer & Acker- man, 1989	.92
Pos	Aufmerksamkeit auf Positive Selbstein- schätzung	„Positive Selbstein- schätzung“	Kanfer & Acker- man, 1989	.64
Neg	Aufmerksamkeit auf Neg. Selbsteinschät- zung	„Negative Selbst- einschätzung“	Kanfer & Acker- man, 1989	.84
bek	„bekannt“: Kontrolli- tems	keine	Pilotstudie (n =10)	Einzelitems
lob	Kontrollitems	keine	Pilotstudie (n =10)	Einzelitems
az	Aufgabenziele	keine	Pilotstudie (n =10)	Einzelitems

Fragen nach „Turm von Hanoi“ – Experimentalbedingung

Ich fühle mich im Moment...	trifft gar nicht zu	trifft wenig zu	trifft mittelmäßig zu	trifft größtenteils zu	trifft vollständig zu
... zufrieden (em1)	1	2	3	4	5
... vergnügt (em2)	1	2	3	4	5
... glücklich (em3)	1	2	3	4	5
... in guter Stimmung (em4)	1	2	3	4	5
Ich bin im Moment eher...					
... erstaunt (a1)	1	2	3	4	5
... überrascht (a2)	1	2	3	4	5
... aktiviert (von außen) (a3)	1	2	3	4	5
... freudig erregt (e1)	1	2	3	4	5
... aktiv (e2)	1	2	3	4	5
... aufgeregt (e3)	1	2	3	4	5
Welches ist die längste Zeit bis zur Problemlösung, mit der du noch zufrieden bist? (as1)	4 min	5 min	6 min	7 min	8 min
Welches ist die maximalste Anzahl Schritte (Züge) bis zur Problemlösung, mit der du noch zufrieden bist? (63 Züge Minimum) (as2)	65 Züge	70 Züge	75 Züge	80 Züge	85 Züge
Welches ist die maximalste Anzahl Regelverstöße bis zur Problemlösung, mit der du noch zufrieden bist? (as3)	1 Fehler	2 Fehler	3 Fehler	4 Fehler	5 Fehler

	trifft gar nicht zu (nie)	trifft we- nig zu (selten)	trifft mit- telmäßig zu (manch- mal)	trifft größten- teils zu (oft)	trifft voll- ständig zu (die ganze Zeit)
...wenig Regelverstöße zu begehen“.(on1)	1	2	3	4	5
...eine bestimmte Regel zur Problemlösung zu fin- den“(on2)	1	2	3	4	5
...so schnell wie möglich das Problem zu lösen“ (on3)	1	2	3	4	5
...so wenig Schritte wie möglich zu benötigen, um die Aufgabe zu lösen“ (on4)	1	2	3	4	5
...das Problem durch nach- denken zu lösen“(on5)	1	2	3	4	5
„Ich frage mich, wie gut meine Leistung im Ver- gleich mit anderen ist“ (LB1)	1	2	3	4	5
„Ich dachte darüber nach...“					
...wie gut ich meine Sache im Vergleich mit anderen mache“(LB2)	1	2	3	4	5
...wie gut andere bei dieser Aufgabe sind“(LB3)	1	2	3	4	5
...wie sehr ich mich verbes- sere“(Pos1)	1	2	3	4	5
...wie gut ich meine Aufga-	1	2	3	4	5

be mache“(Pos2)					
„Ich dachte darüber nach...	trifft gar nicht zu	trifft wenig zu	trifft mittelmäßig zu	trifft größtenteils zu	trifft vollständig zu
...wieviel Spaß mir die Aufgabe macht“ (Pos3)	1	2	3	4	5
...wie meine bisherige Leistung war“ (Pos4)	1	2	3	4	5
...wie schlecht ich meine Aufgabe gemacht habe (Neg1)	1	2	3	4	5
...wie unzufrieden ich mit meiner Leistung bin“(Neg2)	1	2	3	4	5
„Ich sehe überhaupt keine Probleme, diese Aufgabe zu bewältigen“ (nsw1)	1	2	3	4	5
„Ich glaube, dass meine Leistung bei dieser Aufgabe unter dem Durchschnitt sein wird“. (nsw2)	1	2	3	4	5
„Ich habe Vertrauen in meine Fähigkeiten, diese Aufgabe gut zu machen“ (nsw3)	1	2	3	4	5
„Ich glaube, dass ich diese Aufgabe besser bewältige als die meisten anderen Personen“. (nsw4)	1	2	3	4	5
	trifft gar nicht zu	trifft wenig zu	trifft mittelmäßig zu	trifft größtenteils zu	trifft vollständig zu
„Ich kenne die „Turm von	1	2	3	4	5

Hanoi“ Aufgabe“(Bek1)					
„Ich habe diese Aufgabe schon ziemlich oft gemacht“(Bek2)	1	2	3	4	5
„Ich wurde von meinem Versuchsleiter gelobt“(lob1)	1	2	3	4	5
„Ich fühlte mich von meinem Versuchsleiter gelobt“(lob2)	1	2	3	4	5
„Der Versuchsleiter hat mit seinem Lob übertrieben“(lob3)	1	2	3	4	5
„Der Versuchsleiter hat mit seinem Lob untertrieben“(lob4)	1	2	3	4	5
„Ich hatte das Gefühl, der Versuchsleiter meinte sein Lob ernst“(lob5)	1	2	3	4	5
„Ich hatte den Eindruck, der Versuchsleiter lobte mich nicht wegen meiner guten Leistung.“(lob6)	1	2	3	4	5
„Ich fand das Lob des Versuchsleiters gerechtfertigt“(lob7)	1	2	3	4	5
„Ich empfand das Lob als glaubwürdig“(lob8)	1	2	3	4	5
	trifft gar nicht zu	trifft wenig zu	trifft mittelmäßig zu	trifft größtenteils zu	trifft vollständig zu
„Ich hatte den Eindruck, der / die VersuchsleiterIn macht	1	2	3	4	5

mir was vor, als er/sie mich lobte“ (lob9)					
„Ich wollte die Aufgabe möglichst schnell lösen“ (az1)	1	2	3	4	5
„Ich wollte die Aufgabe mit möglichst wenig Schritten (Zügen) lösen“.(az2)	1	2	3	4	5
„Ich wollte die Aufgabe mit möglichst wenig Regelverstößen bewältigen“(az3)	1	2	3	4	5

Fragen nach „Turm von Hanoi“ – Kontrollbedingung

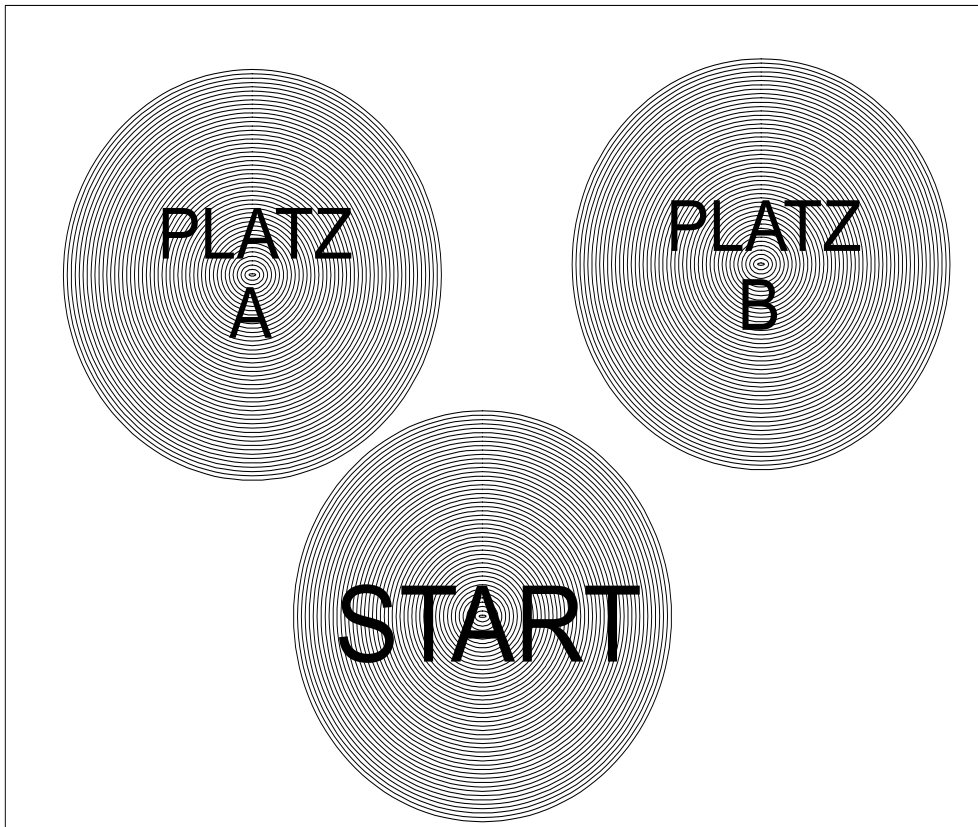
Ich fühle mich im Moment...	trifft gar nicht zu	trifft wenig zu	trifft mittelmäßig zu	trifft größtenteils zu	trifft vollständig zu
... zufrieden (em1)	1	2	3	4	5
... vergnügt (em2)	1	2	3	4	5
... glücklich (em3)	1	2	3	4	5
... in guter Stimmung (em4)	1	2	3	4	5
Ich bin im Moment eher...					
... erstaunt (a1)	1	2	3	4	5
... überrascht (a2)	1	2	3	4	5
... aktiviert (von außen) (a3)	1	2	3	4	5
... freudig erregt (e1)	1	2	3	4	5
... aktiv (e2)	1	2	3	4	5
... aufgeregter (e3)	1	2	3	4	5
Welches ist die längste Zeit bis zur Problemlösung, mit der du noch zufrieden bist? (as1)	4 min	5 min	6 min	7 min	8 min
Welches ist die maximale Anzahl Schritte (Züge) bis zur Problemlösung, mit der du noch zufrieden bist? (63 Züge Minimum) (as2)	65 Züge	70 Züge	75 Züge	80 Züge	85 Züge

Welches ist die maximale Anzahl Regelverstöße bis zur Problemlösung, mit der du noch zufrieden bist? (as3)	1 Fehler	2 Fehler	3 Fehler	4 Fehler	5 Fehler
„Ich richtete meine ganze Aufmerksamkeit darauf...“	trifft gar nicht zu (nie)	trifft wenig zu (selten)	trifft mittelmäßig zu (manchmal)	trifft größtenteils zu (oft)	trifft vollständig zu (die ganze Zeit)
...wenig Regelverstöße zu begehen“ (on1)	1	2	3	4	5
...eine bestimmte Regel zur Problemlösung zu finden“ (on2)	1	2	3	4	5
...so schnell wie möglich das Problem zu lösen“ (on3)	1	2	3	4	5
...so wenig Schritte wie möglich zu benötigen, um die Aufgabe zu lösen“ (on4)	1	2	3	4	5
...das Problem durch nachdenken zu lösen“ (on5)	1	2	3	4	5
„Ich frage mich, wie gut meine Leistung im Vergleich mit anderen ist“ (LB1)	1	2	3	4	5
„Ich dachte darüber nach...“					
...wie gut ich meine	1	2	3	4	5

Sache im Vergleich mit anderen mache“.(LB2)					
„Ich dachte darüber nach...	trifft gar nicht zu	trifft wenig zu	trifft mittelmäßig zu	trifft größtenteils zu	trifft vollständig zu
...wie gut andere bei dieser Aufgabe sind“.(LB3)	1	2	3	4	5
...wie sehr ich mich verbessere“(Pos1)	1	2	3	4	5
...wie gut ich meine Aufgabe mache“(Pos2)	1	2	3	4	5
...wie viel Spaß mir die Aufgabe macht“ (Pos3)	1	2	3	4	5
...wie meine bisherige Leistung war“ (Pos4)	1	2	3	4	5
...wie schlecht ich meine Aufgabe gemacht habe (Neg1)	1	2	3	4	5
...wie unzufrieden ich mit meiner Leistung bin“ (Neg2)	1	2	3	4	5
„Ich sehe überhaupt keine Probleme, diese Aufgabe zu bewältigen“ (nsw1)	1	2	3	4	5
„Ich glaube, dass meine Leistung bei dieser Aufgabe unter dem Durchschnitt sein wird“.(nsw2)	1	2	3	4	5
„Ich habe Vertrauen in meine Fähigkeiten, diese Aufgabe gut zu machen“	1	2	3	4	5

(nsw3)					
„Ich glaube, dass ich diese Aufgabe besser bewältige als die meisten anderen Personen“.	1	2	3	4	5
(nsw4)					
„Ich kenne die „Turm von Hanoi“ Aufgabe“ (Bek1)	1	2	3	4	5
„Ich habe diese Aufgabe schon ziemlich oft gemacht“ (Bek2)	1	2	3	4	5
„Ich wurde von meinem Versuchsleiter gelobt“ (Lob1)	1	2	3	4	5
„Ich fühlte mich von meinem Versuchsleiter gelobt“ (Lob2)	1	2	3	4	5
„Ich wollte die Aufgabe möglichst schnell lösen“ (az1)	1	2	3	4	5
„Ich wollte die Aufgabe mit möglichst wenig Schritten (Zügen) lösen“. (az2)	1	2	3	4	5
„Ich wollte die Aufgabe mit möglichst wenig Regelverstößen bewältigen“ (az3)	1	2	3	4	5

Spielplattform „Turm von Hanoi“



Kategorienschema zur Auswertung der Denkprotokolle) **(Kategoriendefinition mit Beispielen)**

Aufgabenbezogene Gedanken (Kategorie 1)

Alle laut geäußerten Gedanken, die sich direkt auf die Aufgabe und die Aufgabenziele beziehen

Unterkategorien

Bringen die Aufgabenbearbeitung potentiell voran (Kategorie „**Aufgabenbearbeitung**“, 1.1)

- Auf den *Prozess an sich*:
„Den blauen stell` ich jetzt hierhin“
„Ich versuch jetzt den ganzen Turm zu verschieben“
„So geht`s nicht“
- Auf die *Regeln und Aufgabenziele*:
„Also immer nur einen kleineren über einen größeren?!“
„Also möglichst schnell“
- *Verständnisfragen* (auch rhetorische, an sich selbst gerichtete oder an den Versuchsleiter):
„Das grüne ist kleiner als das rote?!“
- *Ideen, Nachdenken, wie das Problem zu lösen ist*
„Bestimmt gibt es ein System das man durchschauen muß“

Kommentare zur Aufgabenschwierigkeit (Kategorie „**Aufgabenschwierigkeit**“, 1.2)

Einschätzung, wie schwierig oder machbar die Aufgabe allgemein ist: „Das ist schwierig“

Selbstbezogene Gedanken (Kategorie 3)

Alle laut geäußerten Gedanken, die sich direkt oder indirekt auf die Person selbst beziehen.

Unterkategorien

Eigene Leistungsbewertung: (Kategorie **Leistungsbewertung**, 3.1)

Wie gut, wie schlecht mache ich meine Sache im Vergleich mit anderen?

(Bsp.: „Stellen sich die anderen auch so blöd an wie ich?“)

Positive Selbsteinschätzung, positive Selbstbewertung

(Kategorie „**Positive Selbstbewertung**“, 3.2)

- Wie sehr hab ich mich verbessert?
- Wie gut mach ich meine Aufgabe?
- Wie gut war meine bisherige Leistung?

Negative Selbsteinschätzung, negative Selbstbewertung

(Kategorie „**Negative Selbstbewertung**“, 3.3)

- Wie schlecht mach` ich meine Aufgabe? („Das war blöd jetzt“)
- Wie unzufrieden bin ich mit meiner Leistung?
- Wie schlecht ist meine Leistung, wie bewerte ich mich? („Jetzt bin ich soweit wie vorher“)

Sonstige Gedanken (Kategorie 2 „**Sonstige Gedanken**“)

Alle laut geäußerten Gedanken, die sich

- nicht direkt auf die Aufgabe beziehen,
- nicht auf die Person (weder direkt noch indirekt) selbst beziehen; zum Beispiel Mißfallensbekundungen, Flüche („Shit“...), Kommentare zum lauten Denken, zur experimentellen Situation an sich, über das Spiel, über das Essen in der Mensa usw.

Beispiel eines DenkprotokollsVPI:

Erstmal überlegen, ähm, nur einen kleineren über einen größeren, das ist wie dieses spiel, dieses "srieschell-spiel" (?), ähm, das ist blöd jetzt, ich darf nur einen kleineren über einen größeren, das ist ...; ich hab mich schon wieder ausmanövriert, ich hab mich schon wieder ausgetrickst, so, so, dann mach wir das da hin, was denkst denn du was ich gemacht hab, das so, dann... ah, ein Regelverstoss, dann so, so, dann kann ich den dahin machen, und dann des dahin, und des dahin, und des dahin, und des hierhin, und dann ..des grüne ist kleiner als des rote, wa? und dann des dahin, und des gelbe ist...und des so, und des so und des so...und das kommt hierdrauf, ach ne, das gelbe ist doch größer, ne? aber,..

LOB

Ha, ungeahnte Talente,...ähm, dahin, und dahin, das hierhin, dann bin ich gleich wieder da wo ich vorhin war,...da vorhin war ich schon besser irgendwie, das war schon geschickter, ich muß kucken das ich das gelbe auf das blaue krieg, ich hab alle freien felder ausgetrickst, das ist immer blöd, so, siehst du, so toll ist es doch nicht weil, so, das stimmt aber, und das machen wir hierdrauf, dann machen wir das da drauf und das hierdrauf, und das da drauf und das da drauf und dann simmer vor dem großen Problem wie so oft, so, nein, das stimmt auch wieder nicht, damn, das ist blöd, das ist eine absolute Zwickmühle, nein, quatsch, den krieg ich nicht mehr raus, da war ich vorhin schon da wo ich jetzt bin, warum krieg ich das gelbe eigentlich nicht frei? ...warum das blaue wieder zugestellt ist, tja, weist du, das ist das wo man dran scheitert letztendlich, das ist..., wenn ich nur eins frei krieg aber dafür das andere wieder nicht, das tun wir dann hierdrauf, oder? ich glaub das war so, so wollt ich das machen, dann ist das gelbe wieder frei, das ist wie mit diesem blöden zauberwürfel wo man irgendwas draus formen muß, ich wette wenn man einmal den trick raus hat dann geht das eigentlich ganz gut, wieviel züge habe ich jetzt eigentlich schon gemacht? ich will das aber gar nicht freikriegen oder, wie mach ich das denn am besten, so, das rote ist größer, ne? das geht aber schlecht,...das blaue ist

Kategorisierungsbeispiel (Zwei Beurteiler: C und K)

	C	K
Erstmal überlegen	2	1.1
ähm, nur einen kleineren über einen größeren,	1.1	1.1
das ist wie dieses spiel, dieses "srieschell-spiel" (?),	2	2
ähm, das ist blöd jetzt,	3.3	3.3
ich darf nur einen kleineren über einen größeren,	1.1	1.1
das ist ...;		
ich hab mich schon wieder ausmanövriert,	3.3	3.3
ich hab mich schon wieder ausgetrickst,	3.3	3.3
so, so, dann mach wir das da hin,	1.1	1.1
was denkst denn du was ich gemacht hab,	2	2
das so,	1.1	1.1
dann... ah, ein Regelverstoss,	1.1	1.1
dann so, so,	1.1	1.1
dann kann ich den dahin machen,	1.1	1.1
und dann des dahin,	1.1	1.1
und des dahin,	1.1	1.1
und des dahin,	1.1	1.1
und des hierhin,	1.1	1.1
und dann ..		
des grüne ist kleiner als des rote, wa?	1.1	1.1
und dann des dahin,	1.1	1.1
und des gelbe ist...	1.1	1.1
und des so,	1.1	1.1
und des so	1.1	1.1
und des so...	1.1	1.1
und das kommt hier drauf,	1.1	1.1
ach ne, das gelbe ist doch größer, ne?	1.1	1.1
aber,..		

LOB

Ha, ungeahnte Talente,	3.2	3.2
...ähm, dahin,	1.1	1.1
und dahin,	1.1	1.1
das hierhin,	1.1	1.1
dann bin ich gleich wieder da wo ich vorhin war,	3.3	3.3
...da vorhin war ich schon besser irgendwie,	3.3	3.3
das war schon geschickter,	3.3	3.3
ich muß kucken das ich das gelbe auf das blaue krieg,	1.1	1.1
ich hab alle freien felder ausgetrickst,	3.3	1.1
das ist immer blöd, so,	3.3	3.3
siehst du, so toll ist es doch nicht weil,	3.3	3.3
so, das stimmt aber,	1.1	3.2
und das machen wir hier drauf,	1.1	1.1
dann machen wir das da drauf und das hier drauf,	1.1	1.1
und das da drauf	1.1	1.1
und das da drauf	1.1	1.1
und dann sind wir vor dem großen Problem wie so oft,	1.2	3.3
so, nein, das stimmt auch wieder nicht,	1.1	3.3
damn, das ist blöd,	3.3	3.3
das ist eine absolute Zwickmühle,	1.2	1.2
nein, quatsch, den krieg ich nicht mehr raus,	1.2	1.1
da war ich vorhin schon da wo ich jetzt bin,	3.3	3.3
warum krieg ich das gelbe eigentlich nicht frei?	3.3	3.3
...warum das blaue wieder zugestellt ist,	3.3	1.1
tja, weist du, das ist das wo man dran scheitert letztendlich,	3.3	3.3
das ist...,		
wenn ich nur eins frei krieg aber dafür das andere wieder nicht,	3.3	1.1
das tun wir dann hierdrauf, oder?	1.1	1.1
ich glaub das war so,	1.1	1.1

so wollt ich das machen,	1.1	1.1
dann ist das gelbe wieder frei,	1.1	1.1
das ist wie mit diesem blöden Zauberwürfel wo man irgendwas draus for- men muß,	2	2
ich wette wenn man einmal den Trick raus hat dann geht das eigentlich ganz gut,	2	1.1
wieviel Züge habe ich jetzt eigentlich schon gemacht?	1.1	1.1
ich will das aber gar nicht freikriegen oder,	1.1	1.1
wie mach ich das denn am besten,	1.1	1.1
so, das rote ist größer, ne?	1.1	1.1
das geht aber schlecht,..	1.2	1.1
das blaue ist	1.1	1.1

Skalen zur Messung der Aufmerksamkeit

(alle Skalen sind Likert–Skalen, die von 1 = nie bis 5 = die ganze Zeit abgestuft sind, nach Kanfer & Ackerman, 1989)

Messung der selbstbezogenen Aufmerksamkeit mit drei Skalen

Aufmerksamkeit auf die Leistungsbewertung (LB)

„Ich dachte darüber nach

- (1) ...wie gut ich meine Sache im Vergleich mit anderen mache“.*
- (2) ...wie gut andere bei dieser Aufgabe sind“.*
- (3) „Ich frage mich, wie gut meine Leistung im Vergleich mit anderen ist“.*

Positive Selbsteinschätzung (POS)

„Ich dachte darüber nach

- (1) ...wie sehr ich mich verbessere“.*
- (2) ...wie gut ich meine Aufgabe mache“.*
- (3) ...wie viel Spaß mir die Aufgabe macht“.*
- (4) ...wie meine bisherige Leistung war“.*

Negative Selbsteinschätzung (NEG)

„Ich dachte darüber nach

- (1) ...wie schlecht ich meine Aufgabe gemacht habe“.*
- (2) ...wie unzufrieden ich mit meiner Leistung bin“.*

Messung der aufgabenbezogenen Aufmerksamkeit (ON)

„Ich fokussierte meine gesamte Aufmerksamkeit darauf

- (1) ...wenig Regelverstöße zu begehen“.*
- (2) ...eine bestimmte Regel zur Problemlösung zu finden“.*
- (3) ...so schnell wie möglich das Problem zu lösen“.*
- (4) ...so wenig Schritte wie möglich zu benötigen, um die Aufgabe zu lösen“.*
- (5) ...das Problem durch nachdenken zu lösen“.*

Tabelle 1: Korrelationen zwischen dem Ort der Aufmerksamkeit und den Leistungsmaßen, getrennt nach Geschlecht

Messungen	Fehler	Züge	Zeit	Absolut	Relativ	Effiz.	Optimal
<u>Denkprotokoll</u>							
1 Aufgabenbearbeitung	-.10 (-.20)	-.13 (.12)	-.16 (-.14)	-.04 (.20)	.22 (.02)	.25 (.05)	.20 (.25)
2 Aufgabenschwierigkeit	-.02 (.10)	.12 (-.17)	.07 (.17)	-.13 (-.05)	-.24 (.20)	-.21 (.16)	-.15 (-.01)
3 Sonstige Gedanken	.23 (.15)	.12 (-.27)	.13 (.04)	.18 (-.24)	-.09 (.01)	-.15 (-.01)	-.22 (-.18)
4 Pos. Selbstbewertung	-.12 (.06)	-.16 (.16)	-.19 (-.22)	-.16 (.21)	-.00 (.03)	.14 (.08)	.09 (-.15)
5 Neg. Selbstbewertung	.27 (.19)	.07 (.15)	.19 (.24)	-.25 (-.24)	-.28 (-.27)	-.30^a (-.32^a)	-.41* (-.31)
<u>Fragebogen</u>							
6 Nachdenken	.05 (-.01)	-.12 (-.25)	-.21 (.04)	.07 (-.10)	.29^a (.24)	.30^a (.22)	.22 (.27)
7 Fehlervermeidung	-.30^a (-.03)	.03 (.14)	-.09 (-.17)	-.15 (-.03)	-.22 (-.24)	-.06 (-.20)	-.05 (-.22)
8 Geschwindigkeit	.16 (.14)	-.37* (.46*)	.32^a (-.21)	-.41* (.25)	.10 (-.48**)	-.15 (-.42*)	-.15 (-.35^a)
9 Leistungsbewertung	.07 (-.02)	.08 (-.12)	.26 (-.10)	-.11 (.05)	-.26 (.15)	-.33^a (.17)	-.20 (.16)
9 Pos. Selbsteinschätzung	-.05 (.22)	-.24 (.04)	.38* (-.32^a)	-.17 (-.00)	.16 (-.01)	-.12 (.06)	-.05 (-.03)
10 Neg. Selbsteinschätzung	.27 (.05)	-.12 (-.18)	.33^a (-.14)	-.21 (-.15)	-.01 (.15)	-.19 (.18)	-.18 (.10)

Anmerkungen:

^a signifikant auf dem Niveau von .10 (zweiseitig)

* signifikant auf dem Niveau von .05 (zweiseitig)

**signifikant auf dem Niveau von .01 (zweiseitig)

In Klammern: Korrelation Männer

Frauen: n = 32; Männer: n = 27

Tabelle 2: Korrelation der Variablen „Glaubwürdigkeit“ mit sämtlichen abhängigen Variablen

Messung	r	r	Messung
1 Aufgabenbearbeitung	-.11	.17	11 Freude nachher
2 Aufgabenschwierigkeit	.04	.27	12 Aktivierung nachher
3 Sonstige Gedanken	.09	-.02	13 Fehler
4 Positive Selbstbewertung	.12	.20	14 Züge
5 Negative Selbstbewertung	.05	.05	15 Zeit
6 Nachdenken	.12	.26	16 Absolut
7 Fehlervermeidung	.09	.24	17 Relativ
8 Geschwindigkeit	.09	.11	18 Effizienz
9 Leistungsbewertung	-.14	.11	19 Optimal
10 Pos. Selbsteinschätzung	.11		
11 Neg. Selbsteinschätzung	-.24		

Anmerkungen: n = 37 (Experimentalgruppe)