

Johannes von Tiling

Neurowissenschaftliche Deutungen von Intelligenz
in den Medien und ihre Folgen für
Lernmotivation, Emotion und Verhalten

Die vorliegende Arbeit wurde vom Fachbereich Wirtschaftswissenschaften der Universität Kassel als Dissertation zur Erlangung des akademischen Grades eines Doktors der Philosophie (Dr. phil.) angenommen.

Erster Gutachter: Prof. Dr. Ernst-Dieter Lantermann, Universität Kassel

Zweiter Gutachter: Prof. Dr. Joachim Funke, Universität Heidelberg

Tag der mündlichen Prüfung

02. Dezember 2008

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar

Zugl.: Kassel, Univ., Diss. 2008

ISBN 978-3-89958-640-4

URN: urn:nbn:de:0002-6419

© 2009, kassel university press GmbH, Kassel

www.upress.uni-kassel.de

Druck und Verarbeitung: Unidruckerei der Universität Kassel
Printed in Germany

Inhaltsverzeichnis

Kurzzusammenfassung / Abstract.....	5
1. Einleitung	6
 Erster Teil: Theoretische Grundlagen	 11
2. Hirnforschung und Öffentlichkeit	11
2.1 Der Einfluss des Hirnforschungsdiskurses auf seine Teilnehmer.....	14
2.1.1 Neuroethik	15
2.1.2 Steuere ich mein Gehirn oder steuert mein Gehirn mich?	16
2.2 Hirnforschung, Lernen und Leistung	17
2.2.1 Die attributionstheoretische Perspektive.....	18
2.2.2 Untersuchungen im Umkreis von Dwecks Leistungsmotivationsmodell.....	19
2.2.3 Mediale Darstellungen neurowissenschaftlicher Intelligenzforschung	24
2.2.4 Die wissenschaftliche Angemessenheit der Medienberichte.....	31
2.3 Fazit	32
 3. Implizite Intelligenztheorien, Emotion und Leistungshandeln.....	 34
3.1 Die Leistungsmotivationstheorie von Dweck	34
3.2 Untersuchungen zum kausalen Einfluss Impliziter Intelligenztheorien	38
3.2.1 Laborexperimentelle Studien	40
3.2.2 Interventionsstudien.....	43
3.2.3 Fazit	46
3.3 Leistungsemotionen und Implizite Intelligenztheorien.....	47
3.3.1 Emotionen in der klassischen Leistungsmotivationsforschung.....	49
3.3.2 Attributionale Ansätze: das Beispiel selbstrelevante Emotionen	50
3.3.3 Emotionen in Abhängigkeit von Zielorientierungen	52
3.3.4 Emotionen in Abhängigkeit von Impliziten Intelligenztheorien	55
3.3.5 Emotionsdarstellung und ihre Rolle in Dwecks Modell.....	59
3.3.6 Fazit	63
 Zweiter Teil: Empirische Untersuchungen	 65
4. Theoretisches Modell, Hypothesen und methodischer Zugang	65
4.1 Fragestellungen und Hypothesen	66
4.1.1 Hauptfrage: Der Einfluss Impliziter Theorien auf Emotion und Verhalten	66
4.1.2 Nebenfragestellungen	69
4.2 Erstellung der Zeitschriftenartikel zur Induktion Impliziter Theorien.....	70
4.2.1 Inhaltliche Gestaltung	71
4.2.2 Formale Gestaltung.....	72

4.3 Auswahl und Erfassung der Emotionskonstrukte.....	76
5. Studie 1: Prüfung des Modells mit Hilfe der Szenario-Methode	81
5.1 Methode	82
5.2 Ergebnisse.....	86
5.2.1 Reaktionen auf negative Leistungsrückmeldungen.....	88
5.2.2 Reaktionen auf positive Leistungsrückmeldungen.....	91
5.3 Diskussion.....	91
6. Studie 2: Prüfung des Modells anhand einer Leistungsaufgabe	97
6.1 Methode	98
6.2 Ergebnisse.....	107
6.2.1 Reaktionen auf negative Leistungsrückmeldungen.....	109
6.2.2 Reaktionen auf positive Leistungsrückmeldungen.....	125
6.3 Diskussion.....	126
7. Studie 3: Prüfung eines erweiterten Modells anhand einer Leistungsaufgabe	134
7.1 Methode	135
7.2 Ergebnisse.....	137
7.2.1 Reaktionen auf negative Leistungsrückmeldungen.....	138
7.2.2 Reaktionen auf positive Leistungsrückmeldungen.....	146
7.3 Diskussion.....	146
Dritter Teil: Schlussfolgerungen und Diskussion.....	149
8. Zusammenfassung und Vergleich der Ergebnisse	149
8.1 Vergleichende Diskussion der drei Untersuchungen	149
8.2 Gegenüberstellung der Befunde zu bestehenden Untersuchungen	151
9. Schlussfolgerungen und praktische Implikationen	156
9.1 Implizite Intelligenztheorien als Ursache von Motivationsproblemen	156
9.2 Implizite Intelligenztheorien und ihre Wirkungsweise im Lernprozess	159
9.3 Von der Neurodidaktik zur Lernpsychologie mit Plastizitätsargumenten ..	161
9.3.1 Die Gefahren eines spekulativ vertretenen Neuro-Determinismus	162
9.3.2 Das Potenzial der neuronalen Plastizität als Metapher des Lernens	166
9.4 Implizite Theorie als Brückenkonzept zur Kulturpsychologie.....	168
Literaturverzeichnis	175

Kurzzusammenfassung / Abstract

Das Leistungsmotivationsmodell von Dweck (1999) geht davon aus, dass die Reaktion einer Person auf Misserfolg davon abhängt, ob sie Intelligenz als stabil oder als veränderbar ansieht. Dieser kausale Einfluss Impliziter Intelligenztheorien auf Motivation und Leistung wurde aber bislang meist nur korrelativ untersucht. Zudem blieb die Rolle spezifischer Emotionen für Dwecks Modell ungeklärt. In drei Experimenten an insgesamt 320 Studierenden wurden die Impliziten Theorien zunächst mit Hilfe von Zeitschriftenartikeln über neurowissenschaftliche Intelligenzforschung induziert. Die Teilnehmer erhielten dann entweder hypothetische Leistungsszenarien zur Einschätzung ihrer möglichen Reaktionen (Studie 1) oder eine Leistungsaufgabe, deren rückgemeldetes Ergebnis Gegenstand der Emotions- und Motivationserhebung war (Studien 2 und 3). Wenngleich sich die erwarteten Zusammenhänge nicht konsistent nachweisen lassen, deutet sich an, dass spezifische Emotionen nach Misserfolg wie Hoffnungslosigkeit, Scham, Schuld und Zuversicht nicht nur von den induzierten Intelligenztheorien abhängen, sondern auch ihrerseits das Bewältigungsverhalten vorhersagen. Teilweise zeigen sich zudem Geschlechtsunterschiede. Auch die Bedeutung des sozialen Kontexts, der Merkmale Attribution, Emotionsdarstellung und Leistung sowie die Reaktionen auf eine Erfolgsrückmeldung werden untersucht. Die Ergebnisse werfen Licht auf die jüngst häufig gestellte Frage, inwieweit neurowissenschaftliche Befunde das Selbstverständnis ihrer Rezipienten verändern. Dabei werden insbesondere die Chancen und Risiken der oft verzerrten medialen Darstellung dieser Befunde diskutiert. Implizite Theorien werden schließlich auch unter kulturpsychologischer Perspektive betrachtet und die daraus entstehenden Implikationen für Dwecks Ansatz diskutiert.

Schlüsselwörter: Implizite Intelligenztheorien, Lernemotionen, Leistungseemotionen, Scham, Schuld, Hirnforschung, Neuroethik, Kulturpsychologie

According to Dweck's (1999) model of achievement motivation, a person's behavior following failure depends on beliefs about the degree to which intelligence can be increased. This causal relationship between implicit theories of intelligence and achievement motivation is mostly investigated by correlative, not experimental designs. Furthermore, the role of specific emotions in this relation has rarely been studied. In three experimental studies, the implicit theories were induced via newspaper articles about neuroscientific intelligence research. After that, the participants either got scenarios to imagine their reactions to hypothetical situations (study 1) or an emotionally activating achievement task (studies 2 and 3). Although the results cannot be interpreted in a clear-cut and consistent way, some indicate that specific emotions following failure (e.g., hopelessness, shame, guilt and hope/confidence) depend on the induced implicit theories and predict coping behavior. The importance of a variety of other variables (e.g., attribution, emotional display and achievement) and conditions (e.g., social context, gender and reactions to success) are discussed as well. The findings shed light on the relation between media coverage of early brain development and achievement beliefs or behavior. The usefulness of a cultural psychological view on implicit theories of intelligence is discussed as well.

Key words: implicit theories of intelligence, achievement emotions, shame, guilt, neurosciences, neuroethics, cultural psychology

1. Einleitung

„Doof bleibt doof, da helfen keine Pillen und keine kalten Umschläge.“ Wenn gleich dieses bekannte deutsche Sprichwort in der Wortwahl etwas altertümlich anmutet, erfreut es sich immer noch großer Beliebtheit. Eltern hört man es sagen, wenn sie versuchen, die gleichbleibend schlechten Schulnoten ihrer Kinder oder die der Mitschüler mit Sinn zu versehen. Neuerdings verwenden es auch Kinder und Jugendliche selbst – kurioserweise in ihren SMS-Kurznachrichten, unter dem Akronym „DBDDHKPUKKU“. Das Überleben des Sprichworts deutet darauf hin, dass es seinen Benutzern bei der Sinngebung und Deutung der Welt gute Dienste leistet. Es legt – freilich halb im Scherz – nahe, dass jemand, der etwas „Doofes“ getan hat, schon immer doof war und auch in Zukunft doof bleiben wird. Von dem „Doofen“ ist nichts mehr zu erwarten, so dass er sich auch selbst vom Vorwurf der Faulheit und Disziplinlosigkeit entlastet sehen kann. Eine inhaltlich verwandte Spruchweisheit lautet: „Was Hänschen nicht lernt, lernt Hans nimmermehr.“ Hier erscheint geistige Leistungsfähigkeit weniger als angeboren denn als Ergebnis frühkindlicher Prägung. Manche Dinge scheint man früh lernen zu müssen, ansonsten lernt man sie gar nicht mehr und bleibt doof. Das Hänschen-Sprichwort läuft insofern auf die Doof-bleibt-doof-Weisheit hinaus, allerdings mit Ausnahme der ersten Lebensjahre. Gibt man es in eine Suchmaschine ein, erhält man eine Fülle von Artikeln über frühkindliche Bildung. Die meisten von ihnen berichten über neue Ergebnisse der Hirnforschung.

Psychologen bezeichnen die Auffassung eines Menschen über die Stabilität oder Veränderbarkeit geistiger Fähigkeiten als seine „Implizite Theorie über Intelligenz“ (im Folgenden als Implizite Intelligenztheorie oder kurz Implizite Theorie bezeichnet). Sie ist deswegen von Interesse, weil sie, wie sich in empirischen Untersuchungen gezeigt hat, mit wichtigen Lern- und Leistungsmerkmalen einhergeht.

Die Frage, ob menschliche Eigenschaften in Stein gemeißelt oder veränderbar sind, ist alt. Doch die Frage, welchen Einfluß es auf unser Leben hat, wenn wir das eine oder das andere glauben, wird erst seit kurzem gestellt: Was ist die Konsequenz, wenn wir glauben, dass wir unsere Intelligenz oder unsere Persönlichkeit weiterentwickeln können, statt zu glauben, es handele sich um unveränderbare und tief verwurzelte Eigenschaften? (Dweck, 2007b, S. 12).

Wer Intelligenz als etwas Gegebenes, Unveränderbares ansieht, läuft Gefahr, auf schulische oder berufliche Misserfolge verzweifelt zu reagieren und schnell aufzugeben. Wer hingegen eher an die Formbarkeit, Veränderbarkeit von Intelligenz glaubt, dem fällt es leichter, Rückschläge zu verkraften und sich für weitere Lernanstrengungen zu motivieren.

Diese Zusammenhänge wurden bereits in vielen Untersuchungen nachgewiesen (vgl. als Überblick Dweck, 1999; Dweck & Molden, 2005; Molden &

Dweck, 2006; populär aufbereitet: Dweck, 2007b). Oft wurde den Untersuchungsteilnehmern ein kurzer Fragebogen über ihr Verständnis von Intelligenz gegeben und danach ihre Schullaufbahn über einen längeren Zeitraum verfolgt. Die gefundenen korrelativen Zusammenhänge zwischen Intelligenztheorie und Lernverhalten können aber nicht kausal interpretiert werden: Vielleicht wurden ja sowohl die Intelligenztheorie als auch das Lernverhaltensmuster von einer dritten Variable verursacht. Wenngleich die tatsächliche Intelligenz der Teilnehmer jeweils miterhoben wurde und nachweislich nicht als eine solche Drittvariable in Frage kommt, gibt es doch eine Vielzahl gut erforschter Motivationskonstrukte, die nicht unwesentlich mit Impliziten Intelligenztheorien korrelieren und somit diese Rolle der verursachenden Drittvariable spielen könnten.

Ob Implizite Intelligenztheorien tatsächlich *kausal* Einfluss auf das Lernverhalten haben, kann nur auf experimentellem Wege untersucht werden, d.h. die Intelligenztheorien müssen den Probanden in der Untersuchungssituation aktiv „eingefloßt“ bzw. induziert werden. Da dies aufwändiger ist als die Erhebung über Fragebögen, sind solche Untersuchungen vergleichsweise selten. In einigen Studien wurde die Induzierung der Impliziten Theorien über das Lesen von Zeitschriftenartikeln vorgenommen, die Intelligenz entweder als stabil oder als veränderbar darstellten. Die Probanden lasen von angeblichen wissenschaftlichen Untersuchungen, die auf den einen oder den anderen Befund schließen ließen. Wurden sie anschließend in eine fordernde Leistungsaufgabe verstrickt, zeigten sich die aus den Korrelationsstudien bekannten Unterschiede (Hong, Chiu, Dweck, Lin, & Wan, 1999). Da diese Untersuchungen jedoch einige methodischen Schwächen aufwiesen und zudem unter inhaltlich eingeschränkter Perspektive durchgeführt wurden (vgl. Abschnitt 3.2), will die vorliegende Arbeit die vorgefundenen Effekte absichern und deren Zustandekommen näher untersuchen.

Im Zuge dessen soll es auch um die inhaltliche Gestaltung der verwendeten Zeitschriftenartikel gehen. Bislang handelten sie von aufwändigen *Zwillingsuntersuchungen*, die genaue Rückschlüsse auf den genetisch bedingten Anteil an der Intelligenz zuließen (der also, je nach experimenteller Bedingung, entweder besonders hoch oder besonders niedrig ausfallen sollte). Gegenwärtig sehr viel einflussreicher und mithin glaubwürdiger ist aber das Vokabular der *Hirnforschung* (vgl. Kuhlmann, 2004). Nahezu jeder Aspekt menschlichen Verhaltens – darunter freilich auch die Intelligenz – wird auf Strukturen und Prozesse im Gehirn kausal zurückgeführt. Auch die psychischen Auswirkungen des Hirnforschungsdiskurses werden bereits öffentlich thematisiert, wenn beispielsweise von einem sich wandelnden Menschenbild oder von „Neuroethik“ die Rede ist. Empirische Studien aber, die diese These von den Auswirkungen medienvermittelter neurowissenschaftlicher Erkenntnisse auf Erleben und Verhalten der Rezipienten untersuchen, liegen meines Wissens noch nicht vor. Daher geht es der

vorliegenden Arbeit auch um diese, hier auf das Thema Intelligenz hin konkretisierte Frage: Beeinflussen medial verbreitete Auffassungen über Intelligenz, die diese beispielsweise als *im Gehirn seit frühester Kindheit festgeschrieben* oder, im Gegenteil, als *repräsentiert in einem sich ständig wandelnden neuronalen Netz* charakterisieren, das Lern- und Leistungsverhalten von Schülern und Studenten?

Damit wird der eher der Grundlagenforschung zuzurechnenden ersten Fragestellung eine eher praktisch bedeutsame zweite gegenübergestellt. Implizite Intelligenztheorien sind einerseits für sich genommen von Interesse, andererseits dienen sie gleichsam als Werkzeug, um die Frage nach den Auswirkungen des Hirnforschungsdiskurses in eine konkrete Forschungshypothese umzusetzen. Da beide Fragestellungen somit in ein und demselben Versuchsdesign untersucht werden können (vgl. Kapitel 4), erscheint diese Kopplung auch unter forschungsökonomischen Aspekten sinnvoll. Allerdings können die in dieser Arbeit vorgestellten empirischen Untersuchungen nur einen bescheidenen, inhaltlich eng umgrenzten Beitrag zu einem besseren Verständnis der alltagspsychologischen Bedeutung der Hirnforschung leisten. Den theoretischen Grundlagen dieser zweiten Fragestellung wird dagegen vergleichsweise viel Raum zugestanden, da diesbezüglich kaum Vorarbeiten vorhanden sind.

Kommen wir noch einmal zurück zu den eingangs erwähnten Sprichwörtern. Stellen wir uns ein zehnjähriges Mädchen vor, das zum wiederholten Male eine schlechte Mathematiknote mit nach Hause bringt. Ihre Mutter quittiert die schlechte Nachricht, halb im Spaß, mit dem Ausruf: „Naja mach’ dir nichts draus, doof bleibt doof, da helfen keine Pillen und keine kalten Umschläge“. Sie will ihre Tochter damit trösten und ihr die Verantwortung für ihr Versagen abnehmen, und doch sollte genau dieses Glaubhaftmachen der stabilen Intelligenzauffassung dem Mädchen in ihrer weiteren schulischen Entwicklung nachhaltig schaden. Dass dies so ist, legen die oben genannten Arbeiten in eindrucklicher Art und Weise nahe. Aber wie kommt es dazu? Welche vermittelnden Prozesse sind für die auf den ersten Blick etwas weit hergeholt erscheinende Beziehung zwischen Impliziten Intelligenztheorien und Leistungsmotivation verantwortlich? Auch zu dieser Frage liegen nur wenige aussagekräftige Untersuchungen vor, wobei insbesondere die Bedeutung von Emotionen kaum erforscht wurde. Dass Emotionen jedoch eine wichtige Rolle spielen könnten, kann anhand des obigen Beispiels leicht plausibilisiert werden: Die Schülerin könnte die Aussage der Mutter im Sinne einer Abwertung ihrer Person auffassen, so dass sie mit Beschämung, Niedergeschlagenheit oder sogar Depression reagiert. Nur insoweit diese gefühlsmäßige Reaktion zustande kommt, sollte die Schülerin dann auch in Zukunft ihre Motivation und Leistung verringern. Ein weiteres Anliegen der vorliegenden Arbeit ist demnach die Bedeutung von Emotionen für die Beziehung zwischen Impliziten Intelligenztheorien, Motivation und Leistung.

Die vorgestellten drei Forschungsfragen – (1) Auswirkungen des Hirnforschungsdiskurses auf das Leistungsverhalten, (2) kausale Wirkungskraft Impliziter Intelligenztheorien, (3) Bedeutung von Emotionen für die Wirkungsweise Impliziter Intelligenztheorien – werden in den folgenden beiden Kapiteln anhand ihrer theoretischen und empirischen Grundlagen näher herausgearbeitet. In Kapitel 2 geht es dabei um die Beziehung zwischen Hirnforschung und Öffentlichkeit. Kapitel 3 beschäftigt sich mit der Leistungsmotivationstheorie von Dweck, deren zentrales Konzept die Implizite Intelligenztheorie ist. Dabei geht es auch um die Frage der kausalen Bedeutung Impliziter Theorien (Abschnitt 3.2) sowie um Ansätze zu Emotionen in Lern- und Leistungssituationen und deren Relevanz für Dwecks Modell (Abschnitt 3.3). In Kapitel 4 werden aus den vorausgegangenen Erörterungen die Hypothesen der drei experimentellen Untersuchungen abgeleitet, die in den Kapiteln 5, 6 und 7 geschildert werden. Das Kapitel 4 fokussiert außerdem auf einige Versuchsmaterialien und methodische Herangehensweisen, die für alle drei Untersuchungen identisch sind. In der ersten Studie (Kapitel 5) werden die Hypothesen anhand von Fallgeschichten untersucht, in die sich die Probanden hineinversetzen sollen, um dann ihr hypothetisches emotionales Erleben und Verhalten einzuschätzen. Die zweite Studie (Kapitel 6) konfrontiert die Teilnehmer hingegen mit einer realen Leistungsaufgabe, deren Resultat als Bezugspunkt der Emotions- und Motivationserhebung verwendet wird. Die dritte Studie (Kapitel 7) geht methodisch ähnlich vor, versucht aber einige in der zweiten Studie offen gebliebene Fragen zu klären. Die Kapitel 8 und 9 fassen die Ergebnisse der Untersuchungen zusammen und diskutieren ihre praktischen Implikationen.

Nachzutragen ist übrigens noch, dass das „Doof-bleibt-doof“-Sprichwort nicht mehr stimmt. Pillen zur vorübergehenden Steigerung der kognitiven Leistungsfähigkeit gibt es mittlerweile in großer Vielfalt (vgl. Kapitel 2.1). Gerade die Neurowissenschaften, die für die Entwicklung dieser Tabletten verantwortlich sind, stützen hiermit also die Veränderbarkeitsthese: Intelligenz ist nicht schicksalhaft in den Genen festgeschrieben, sondern sie lässt sich steigern. Allerdings mit Hilfe von Tabletten, nicht (unbedingt) mit eigenen Lernanstrengungen. Dies jedenfalls könnte der naheliegende Schluss sein, den viele Schüler und Studenten ziehen würden: Was man mit Tabletten verbessern kann, muss naturwüchsig sein; und was naturwüchsig ist, kann nicht mit bloßer geistiger Anstrengung verändert werden. „Wenn es ohnehin nur auf die materiellen Kausalfaktoren ankommt und nicht auf die Einsichten und Erfahrungen eines Heranwachsenden, dann spricht wirklich nichts mehr dagegen, das junge Gehirn entsprechend zu manipulieren“ (Kuhlmann, 2004, S. 152). An diesem Beispiel wird schon deutlich, dass das Vokabular der Hirnforschung nicht durchweg positive oder negative, sondern zutiefst ambivalente Auswirkungen auf Lernen und Leistung hat. Vor einseitigen Vorverurteilungen sei also gewarnt.

An dieser Stelle will ich auf zwei Besonderheiten der in dieser Arbeit verwendeten Sprache aufmerksam machen. Erstens erschien es mir nicht praktikabel, angesichts des ohnehin schon vielschichtigen und zuweilen sprachlich schwer fassbaren Gegenstands meiner Arbeit eine geschlechtergerechte Sprache zu verwenden. Die weibliche Form müssen sich die Leserin und der Leser somit leider immer im Stillen ergänzen. Zweitens wird die Neue Rechtschreibung aus Gründen der Lesbarkeit nur in Teilen umgesetzt. Insbesondere die neuen Regeln der Getrennschreibung werden bewusst ignoriert. Das Kleinkind Adam, von dem meine Probanden in einem fingierten Zeitschriftenartikel lesen, ist beispielsweise weiterhin *hochbegabt* und nicht *hoch begabt*. Letztere, mittlerweile empfohlene Schreibweise könnte unterschiedlich verstanden werden und – wie später deutlich werden wird – im Zuge dessen auch unterschiedliche pädagogische Implikationen haben.

Erster Teil: Theoretische Grundlagen

2. Hirnforschung und Öffentlichkeit

Gary sorgte sich viel um seine geistige Gesundheit (...). Obwohl Gary den aktuellen Trend zur individuellen Gestaltung von Altersgeldanlagen und Ferngesprächstarifen und privaten Bildungsoptionen im Prinzip begrüßte, war er alles andere als begeistert, auch die Verantwortung für seine persönliche Hirnchemie aufgehalst zu bekommen (...). Doch wenn man von Gary eines sagen konnte, dann das: Er war pflichtbewußt. (...) [Er schätzte] daß sein Neurofaktor 3 (also Serotonin, ein sehr, sehr wichtiger Faktor) seit sieben, wenn nicht gar dreißig Tagen einen Höchststand verzeichnete (...). Sein Groll gegen Caroline, seine Frau, hielt sich in wohlkontrollierten Grenzen (...) und seine jahreszeitlich angepaßten Gedanken zur Sinnlosigkeit und Kürze des Lebens korrespondierten mit dem robusten Allgemeinzustand seiner geistigen Ökonomie. Er war nicht das kleinste bißchen klinisch depressiv.

(Jonathan Franzen, *Die Korrekturen*, S. 195-196)

Noch nie sind Menschen in den westlichen Gesellschaften so sehr wie heute mit der Tatsache konfrontiert worden, dass sie ein Gehirn haben und dass dieses hinter ihrem Rücken ihr Tun mehr beeinflusst, als sie glauben. Bereits 1990 erklärte der damalige amerikanische Präsident Bush die bevorstehende Dekade zum Jahrzehnt des Gehirns. Die Förderung neurowissenschaftlicher Forschungsprogramme wurde daraufhin massiv verstärkt, die Auseinandersetzung mit der Öffentlichkeit aktiv gesucht. Spätestens seit Ende 2003 ist die Hirnforschung auch in den deutschen Medien präsent (vgl. Kuhlmann, 2004). Am 1.11.2003 widmet die Zeitschrift „Spiegel special“ dem Thema eine ganze Ausgabe und lässt gleich im Titel des Editorials verlauten: „Die Entschlüsselung des Gehirns – Die Hirnforschung schwingt sich zur Leitwissenschaft des beginnenden Jahrhunderts auf“. Weiter unten wird erläutert, was darunter zu verstehen ist: „Gleichgültig, ob Pädagogik, Religion, Rechtsprechung oder Musikwissenschaft: Überall mischt sie sich inzwischen ein“ (S. 3). Geschickt lassen diese Aussagen offen, weshalb die Hirnforschung zur Leitwissenschaft werden könnte: weil sie in nahezu allen wissenschaftlichen Teildisziplinen etwas *sagen kann*, oder weil sie in ihnen etwas *zu sagen hat*? Ist das „Aufschwingen“ und „Einmischen“ inhaltlich gerechtfertigt oder wissenschaftspolitisch motiviert?

Dass die Hirnforschung in den letzten 20 Jahren große Fortschritte gemacht hat, kann wohl nicht ernsthaft in Zweifel gezogen werden. Die ihr zugewiesenen Fördermittel und die öffentliche Aufmerksamkeit sind somit inhaltlich sicher

nicht ganz unbegründet. Hinzu treten aber einige weitere Gründe für ihren Erfolg, die weniger in harten wissenschaftlichen Ergebnissen als im Umgang mit der Öffentlichkeit zu suchen sind (Könneker, 2007). *Erstens* berühren die Themen der Hirnforschung ähnlich wie die der Psychologie menschliche Alltagsprobleme wie Gedächtnis, Gefühle und Intelligenz. Das dadurch ohnehin schon wachgerufene Interesse wird aber *zweitens* noch verstärkt durch die prägnante und anschauliche Visualisierung neurowissenschaftlicher Erkenntnisse. Anders als die Forschungsergebnisse der Psychologie, die in Form von abstrakten statistischen Kennwerten daherkommen, „vermitteln die bunten Bilder aus dem Tomographen jedem Betrachter den Eindruck, direkt ins Gehirn zu blicken und dort sinnvolle Zusammenhänge zu erkennen. Es scheint, als hätten die Neurowissenschaftler die Black Box endlich geöffnet und brächten jetzt die verborgenen Mechanismen des Geistes ans Tageslicht. Aus diesem Grund finden Hirnscans den Weg von den einschlägigen Fachjournalen bis auf die Psychologieseiten von Wellness- und Frauenmagazinen“ (Könneker, 2007, S. 12).

Visualisierbarkeit allein genügt gleichwohl nicht. Es braucht auch jemanden, der glaubwürdig sagt, was auf den bunten Bildern zu sehen ist und welche Schlussfolgerungen daraus zu ziehen sind. *Drittens* muss also die Kommunikation zwischen Wissenschaft und Medien gut funktionieren. Carsten Könneker (2007) notiert dazu weiter: „Gerade in Deutschland äußerten sich namhafte Hirnforscher auch vielfach außerhalb der Fachzeitschriften und verhalfen ihrer Wissenschaft dadurch zu weiterer Aufmerksamkeit – teils sogar unter Kritik ihrer eigenen Kollegen“ (S. 12). Könneker muss es wissen, ist er doch Chefredakteur des 2002 gegründeten Magazins *Gehirn und Geist*, das gewissermaßen das neurowissenschaftliche Konkurrenzprodukt zu *Psychologie heute* darstellt und in dessen wissenschaftlichem Beirat viele bekannte deutsche Hirnforscher sitzen. In dieser Zeitschrift wurde auch das sogenannte „Manifest“ führender deutscher Hirnforscher publiziert (Ausgabe 6/2004). Dessen Co-Autor Christian Elger räumt freimütig auf die Frage nach den Motiven hinter dem Manifest ein:

FOCUS: Forscher, die Manifeste schreiben, wünschen sich entweder mehr Etat oder mehr Anerkennung. Wie ist es bei Ihnen?

Elger: Wir wollen beides. In Deutschland fehlt einfach ein großes Hirnforschungszentrum im Stil des Deutschen Krebsforschungszentrums – dabei ist die Hirnforschung die Leitwissenschaft des 21. Jahrhunderts. Von den Bürgern wurde das lange nicht so gesehen. Sie assoziierten meine Disziplin eher mit Gedankenkontrolle (*Focus* 43/2004, S. 110).

Schaut man sich einschlägige Veröffentlichungen in den Printmedien zum Thema Hirnforschung an, erhält man tatsächlich den Eindruck, dass sich die Hirnforscher noch mehr für die Medien interessieren als umgekehrt. Sie bedienen sich dabei zum einen der – etwa aus der Genforschung bekannten – Strategie, die baldige Heilbarkeit schwerer Krankheiten zu prognostizieren. Zum anderen fällt

aber des Öfteren auch der diffuse Hinweis, dass allein schon das Rezipieren neurowissenschaftlicher Erkenntnisse unser „Menschenbild“, unser Erleben und Verhalten in irgendeiner Form verändern soll. Zuweilen ist gar von einem „Frontalangriff auf unser Selbstverständnis und unsere Menschenwürde“ (*Gehirn und Geist*, 04/2002, S. 32, Titel) die Rede. Wolf Singer postulierte schon früh, „dass Erkenntnisse der Grundlagenforschung auch ohne Anwendungsbezug unser Menschenbild nachhaltig verändern“ (*Die Zeit*, 50/2000, S. 43). Im gleichen Atemzug fügt er hinzu, dass dies nicht nur eintreten *wird*, sondern auch *soll*, wenn er beklagt, „wie wenig so genannte kultivierte Kreise über naturwissenschaftliche Entwicklungen wissen“, und fordert: „Dieses Wissen muss Auswirkungen haben auf unser Rechtssystem, auf die Art, wie wir Kinder erziehen und wie wir mit Mitmenschen umgehen“ (ebd.). Dieser „sozialreformerische Elan“ (Kuhlmann, 2004, S. 148), der leicht in ein „technokratisches Programm zur rechtzeitigen Auslese“ (ebd., S. 149) umschlagen könne, blieb übrigens nicht unkritisiert. Der Wissenschaftssoziologie Peter Weingart beschreibt das Vorgehen von Hirnforschern wie Singer in einem Interview mit der Zeitschrift *Gehirn und Geist* (1/2005) hingegen eher nüchtern:

Zurzeit befindet sich die Hirnforschung auf dem Vormarsch in die Medien. Das ist ein wunderschönes Beispiel für einen Wissenschaftszweig, der den Weg in die Öffentlichkeit aktiv sucht, um in die öffentliche Diskussion zu geraten und darüber Akzeptanz zu finden. Hirnforscher wie Gerhard Roth und Wolf Singer verkünden ja in den Medien immer wieder, dass sich aus den Ergebnissen der Hirnforschung ein neues Menschenbild und ein neues Strafrecht ableiten ließen. Aber wenn man Vorträge der beiden im wissenschaftlichen und außerwissenschaftlichen Kontext vergleicht, stellt man fest, dass in Letzterem wesentlich weiterreichende Behauptungen aufgestellt werden. Die derzeit um sich greifende Debatte scheint also durchaus gewollt zu sein (S. 75).

Welche Begründung führen die zitierten Hirnforscher für ihre Menschenbildthese – also dass die Ergebnisse ihrer Disziplin eine einschneidende Wirkung auf das Publikum haben sollen – ins Feld? Ist es nicht für den „Mann auf der Straße“ oder auch für die von Singer angeführten kultivierten Kreise von untergeordneter Bedeutung, dass vielleicht Emotionen in einem bestimmten Teil des Gehirns repräsentiert sind, Intelligenz in einem anderem, und dass es zwischen diesen Bereichen vielleicht interessante neuronale Verbindungen gibt? Was unser Menschenbild verändern soll, sind indes nicht die Einzelergebnisse der Hirnforschung, sondern deren übergreifende Interpretation in eine Richtung, die mit dem Begriff *Epiphanomenalismus* gekennzeichnet werden kann. Damit ist die These gemeint, dass unser Erleben und Handeln nicht über die Interaktion mit einer bedeutungshaft strukturierten sozialen Umwelt entsteht (indem wir diese nämlich deuten und Schlüsse ziehen), sondern allein von physikalisch-chemischen Prozessen gesteuert wird, die mit jener sozialen Umwelt nichts zu tun haben. Kurz: Das Mentale ist nur Epiphanomen der Hirnvorgänge (vgl. z.B. Roth, 2004).

Dies kann man als radikale Variante des Freudschen Diktums verstehen, das Ich sei nicht Herr im eigenen Haus: Unsere Handlungen entspringen nicht mehr nur unbewussten, dunklen Motiven, sondern gar keinen Motiven mehr; sie laufen unabhängig vom Ich ab. Das Ich kann nur im Nachhinein die Geschehnisse mit Sinn versehen und sie sich als Handlung zuschreiben. Willensfreiheit, eine der meistzitierten Vokabeln im Epiphänomenalismus-Diskurs (vgl. Geyer, 2004b), ist demnach nichts weiter als eine Illusion oder, noch interessanter, „kulturelle Konstruktion“ (Singer in *Die Zeit*, 50/2000, S. 43).

Die epiphänomenalistische Deutung der Ergebnisse der Hirnforschung ist bekanntlich alles andere als unumstritten (vgl. Bennett & Hacker, 2003; Geyer, 2004b; Krüger, 2007). Wolfgang Prinz (2008), Direktor der Abteilung für Psychologie am Max-Planck-Institut für Kognitions- und Neurowissenschaften in Leipzig, stellte kürzlich auf prägnante Art und Weise dar, wie die epiphänomenalistische Denkweise Gefahr läuft, sowohl reduktionistischen als auch naturalistischen Fehlschlüssen aufzusitzen. Selbst Michael S. Gazzaniga, einer der prominentesten Hirnforscher der USA, stellte kürzlich klar:

„There are important ethical areas to which neuroscientists are being asked to contribute when, in fact, they should not be. For instance, neuroscience has nothing to say about concepts such as free will and personal responsibility, and it probably also has nothing to say about such things as antisocial thoughts“ (Gazzaniga, 2006, S. 141).

Die vorliegende Arbeit setzt sich keineswegs zum Ziel, mehr Licht in diese Frage zu bringen. Sie verfolgt weder einen neurowissenschaftlichen noch einen philosophischen Ansatz. Die Umstrittenheit der in den Medien häufig als unumstritten präsentierten epiphänomenalistischen Thesen lässt aber die von Singer und anderen aufgestellte Menschenbildthese in neuem, ethisch relevantem Licht erscheinen: Sollte sich etwa herausstellen, dass die Konfrontation mit epiphänomenalistischen Thesen abträgliche Auswirkungen für die Medienkonsumenten hat, wäre diese Art der Werbung für die eigene Forschungsrichtung moralisch in Frage zu stellen. Nicht zuletzt deshalb will sich die vorliegende Arbeit weniger mit der epiphänomenalistischen These als mit der Menschenbildthese beschäftigen. Diese soll im Folgenden näher erörtert und spezifiziert werden.

2.1 Der Einfluss des Hirnforschungsdiskurses auf seine Teilnehmer

Die damalige Bundesforschungsministerin Bulmahn verwies bereits vor einiger Zeit bei der Frage nach Zukunftsthemen der Geisteswissenschaften auf „die spannende Frage, wie die Informationstechnologien und neue naturwissenschaftliche Erkenntnisse etwa aus der Hirnforschung unser Menschenbild verändern“ (*Die Zeit*, 24/2005, S. 42). Zwar muss man Bulmahn nicht unbedingt darin zustimmen, dass sich die Geisteswissenschaften nur noch durch die Bezugnahme

auf „härtere“ Wissenschaften profilieren können (vgl. etwa Laucken, 2003). Aber aus sozial-, medien- und kulturpsychologischer Perspektive drängt sich in der Tat die Frage auf, ob die Hirnforschung als deutungsmächtiger Diskurs der Gegenwart (vergleichbar mit einer Religion) tatsächlich Auswirkungen auf individuelles Erleben und Verhalten hat. Um so überraschender ist, dass meines Wissens bislang keine empirischen Untersuchungen zu dieser Frage vorliegen.

Die Skepsis, ob ein solches Forschungsvorhaben erfolgversprechend ist, scheint zu überwiegen. Verschiedene Autoren (Geyer, 2004a; Kuhlmann, 2004; vgl. auch *Gehirn und Geist* 1/2005, S. 75) verweisen darauf, dass die Debatten der Hirnforschung trotz aller Inszenierungsversuche immer noch von zu großer Alltagsferne seien, um wirkungsmächtig zu sein. Ihre alltagspraktischen Implikationen dürften, so wird gemutmaßt, kaum größer sein als die der Einsteinschen Relativitätstheorie.

2.1.1 Neuroethik

Gleichwohl geben Untersuchungen der sogenannten Neuroethik (Farah, 2002; Illes, 2006) erste Hinweise darauf, dass die Themen, die die Hirnforschung berührt, eine deutlich größere praktische Relevanz besitzen als Einsteins Relativierung von Raum und Zeit. Der Neuroethik geht es als noch junge wissenschaftliche Disziplin sowohl um die neurowissenschaftliche Erforschung von ethisch-moralischen Denkvorgängen (sog. „Neurowissenschaften der Ethik“) als auch um die ethische Vertretbarkeit und Regelungsmöglichkeit bestimmter medizinischer Praktiken, die aus der Anwendung neurowissenschaftlicher Erkenntnisse hervorgehen („Ethik der Neurowissenschaften“). Im vorliegenden Rahmen ist nur der letztere Zweig von Interesse. Hier geht es beispielsweise um die Frage, ob eine Steigerung kognitiver und affektiver Fähigkeiten mit Hilfe von Neurotechnologien (z.B. Psychopharmaka) auch bei gesunden Menschen ethisch zu rechtfertigen ist. Während diese Debatte in Deutschland allenfalls innerhalb des engen Kontexts von Aufmerksamkeitsproblemen bei Kindern (Stichwort „Ritalin“) geführt wird, gibt es in den USA bereits eine viel breitere Kultur des Selbstverbesserns („Enhancement“; vgl. Farah et al., 2004). Kritiker dieser Kultur argumentieren, dass das „Gehirndoping“ Einzelner dazu führe, dass im Zuge des schulischen und beruflichen Wettbewerbs langfristig auch alle anderen zur Nachahmung gezwungen seien. Da ein solcher faktischer Zwang zum Enhancement aber nicht ethisch legitimierbar sei, müsse man ihm durch Verbote vorbeugen.

Die Debatte läuft schließlich auf die Frage hinaus, ob das Umsichgreifen der Selbstoptimierung eine Gefahr für die gesellschaftlich unabdingbare Vorstellung des selbstbestimmten und verantwortlichen Individuums darstellt oder ob nicht gerade die Selbstbestimmung über die eigene Hirnbiologie zur Stärkung und

Vervollständigung dieser Vorstellung beiträgt (Parens, 2006). Spätestens hier wird deutlich, dass neurowissenschaftliche Anwendungen – mehr noch als solche anderer Naturwissenschaften – immer auch das Selbstverständnis des Menschen betreffen. Dies wird von den Vertretern der Neuroethik auch explizit so gesehen:

Neurological biotechnologies differ from others in that they ask us to explicitly consider the kind of „self“ we want to have; or, to put it less dualistically, perhaps, the kind of self we want to be. For some, our astounding ability to manipulate our own biology is an integral part of who we are as human animals. For others, it is an affront to our humanity. This is an argument for which there are no right or wrong answers, emerging as it does from two philosophically different visions of human life. Yet herein lies the tension of the enhancement debate (Wolpe, 2002, S. 394).

Fukuyama (2002) sieht infolge neuer Psychopharmaka gar das „Ende des Menschen“ (Titel) kommen, womit er vor allem das Ende eines humanistischen, selbstbestimmten Menschenbildes meint. „Unsere Fähigkeit, menschliches Verhalten zu manipulieren, hängt jedoch nicht von der Weiterentwicklung der Genmanipulation ab. Buchstäblich alles, von dem wir annehmen, daß es durch Genmanipulation erreichbar ist, werden wir wahrscheinlich sehr viel eher mit den Mitteln der Neuroparmakologie bewerkstelligen können“ (S. 242).

Die naheliegende Schlussfolgerung, dass auch die Darstellung der Hirnforschung in den Medien das Selbstverständnis, Denken und Handeln nicht unberührt lässt, ziehen die Vertreter der Neuroethik indes nicht oder nur in allgemeiner und vager Form. Unter der Überschrift „Public Communications“ spricht etwa Kulynych (2002) von der ethischen Verpflichtung der Hirnforscher, „[to] avoid inadvertently fueling misconceptions about the power and the promise of neuroimaging“ (S. 355). Grainger-Monsen und Karetsky (2006) thematisieren den Einfluss der Medien auf das alltagspsychologische Verständnis neuronaler Prozesse immerhin in einem gesonderten Beitrag. Für den vorliegenden Kontext interessant ist etwa eine dort zitierte Umfrage (National Health Council), nach der 58% der Amerikaner angeben, ihr Verhalten infolge eines Gesundheitsbeitrags in den Medien geändert zu haben. Von systematischen empirischen Studien sind ihre Überlegungen aber weit entfernt. Weil sich die Neuroethik primär innerhalb der Neurowissenschaften (nicht aber der Sozialwissenschaften oder Philosophie) entwickelt hat, erscheint diese Fokussierung auf konkrete medizinische Praktiken zulasten „weicher“ öffentlicher Diskurse auch nicht verwunderlich.

2.1.2 Steuere ich mein Gehirn oder steuert mein Gehirn mich?

Auch in der Neuroethik lassen sich also für die Frage, wie Menschen die von außen kommenden Informationen über ihr Gehirn in ihr Selbstbild integrieren, keinerlei empirische Vorarbeiten ausfindig machen. Interessante Hypothesen über

diese Frage werden indes in neuroethischen und anderen Diskussionen immer wieder explizit oder implizit angesprochen: Lässt sich Hirnforschung als letzte große Entzauberung des Menschen verstehen („Liebe und Gottesglaube sind nur chemische Prozesse!“), oder führt sie im Gegenteil zu einer Bestärkung immaterieller Realitäten („Liebe und Gottesglaube gibt es wirklich – man kann sie auf dem Hirnbild sehen!“)? Erfährt der Bereichen des Seelischen bzw. Psychischen im Zuge der Rückführung auf naturwissenschaftliche Prozesse eher eine Abwertung, die sich in Zynismus und größerer Distanz zu den eigenen Gefühlen, Zielen und moralischen Vorstellungen ausdrückt, oder im Gegenteil eine Aufwertung, weil Gefühle, Ziele und moralische Vorstellungen nun nicht mehr bloß schöne Worte bzw. soziale Konstrukte sind, sondern „wirklich existent“?

Dies sind nur einige mögliche Fragestellungen, die allerdings jeweils noch einiger Konkretisierungsarbeit bedürfen. In einem Punkt scheinen sie aber auf dieselbe Frage hinauszulaufen: Erscheinen gemäß dem Hirnforschungsdiskurs bewusste individuelle Entscheidungen und psychische Prozesse eher als *abhängig vom Gehirn und somit der gezielten Kontrolle entzogen*, oder im Gegenteil als ihrerseits *das Gehirn determinierend und von demselben in ihrer Bedeutung bestärkt*? Führt der Bezug auf das Gehirn eher zum Verlust oder zum Gewinn subjektiven Kontrollempfindens? Im folgenden Abschnitt werde ich diese Fragestellung speziell im Lern- und Leistungskontext untersuchen, da hier bereits empirisches Material vorhanden ist. Die obigen Beispiele mögen aber deutlich gemacht haben, dass man die Menschenbildthese auf ganz unterschiedliche Art und Weise konkretisieren kann.

2.2 Hirnforschung, Lernen und Leistung

Sie war Charlotte Simmons. Könnte sie jemals so in sich gehen, wie Momma es von ihr gefordert hatte? Mr. Starling setzte Seele in Anführungszeichen, was nichts anderes bedeutete, als dass dahinter in erster Linie ein Aberglaube steckte, dass „Seele“ nichts war als ein früherer, noch primitiverer Name für den Geist in der Maschine. (...)

Warum also höre ich den Geist immer wieder dieselbe ermüdende Frage stellen: „Ja, aber was heißt das? Wer *ist* diese Charlotte Simmons?“ – „Man kann einen Menschen nicht einfach so definieren“, sagte Charlotte Simmons, „jeder ist einzigartig.“ Und was erwiderte er, der kleine Geist, den es nicht gibt? „Na, dann nenn mir doch ein paar Attribute, die sie von jedem anderen Mädchen in Dumont unterscheiden, ein paar ihrer Träume, ihrer Wünsche? War es nicht Charlotte Simmons, die ein Leben des Geistes führen wollte? Oder war es immer nur ihr Ehrgeiz, als etwas Besonderes dazustehen und allein dafür bewundert zu werden, ganz egal, auf welchem Wege sie dorthin gelangt ist?“ (Tom Wolfe, *Ich bin Charlotte Simmons*)

2.2.1 Die attributionstheoretische Perspektive

Im Lern- und Leistungskontext bietet sich für die Sammlung möglicher Hypothesen über die Auswirkungen des Hirnforschungsdiskurses eine *attributionstheoretische Perspektive* an. Unter ihr geht es um die Konsequenzen, die entstehen, wenn neuronale Prozesse zum Gegenstand von Leistungszuschreibungen werden. Dabei erscheint die Frage zentral, in welche Dimensionen das Gehirn als Gegenstand der Attribuierung der eigenen Leistung einzuordnen ist. Was bedeutet es für Menschen genau, wenn sie ihre Leistung auf Hirnprozesse zurückführen? Wie verstehen sie diese Kategorie „Hirnprozesse“? Unstrittig dürfte sein, dass es sich um eine *internale* Attribution handelt, d.h. wer ein Handlungsergebnis auf neuronale Prozesse zurückführt, rechnet sie damit der eigenen Person zu. Externale Bezugspunkte wie die Mitmenschen, die Gesellschaft oder der Zufall werden damit unattraktiver und unwahrscheinlicher. Hinsichtlich der Dimensionen Stabilität, Kontrollierbarkeit und Verantwortlichkeit (Weiner, 1985, 1995) scheint der Status neuronaler Prozesse hingegen alles andere als klar zu sein. Dies soll an einem konkreten Beispiel veranschaulicht werden.

Der 12-jährige Tim erbringt schon seit Längerem mangelhafte Leistungen in Mathematik. Eines Tages wird er mit einer Fernsehsendung konfrontiert, die über bestimmte Hirnstrukturen und -prozesse berichtet, welche bei der Lösung mathematischer Aufgaben eine Rolle spielen. Welchen Schluss könnte Tim aus dieser Sendung für die Ursache der eigenen Mathematikleistungen ziehen? Assoziiert er mit neuronalen Prozessen eher eine naturwissenschaftliche Sicht auf den Menschen, die ihm nahelegt, dass seine mathematischen Fähigkeiten durch „externe“, unbewusste Kräfte determiniert sind und sie folglich nicht durch ihn selbst beeinflussbar sind (stabil-unkontrollierbare Attribution)? Oder konstruiert er die dynamischen Hirnprozesse eher als Gegensatz zu stabilen genetischen Anlagen und Begabung, so dass ihm seine mathematischen Fähigkeiten plötzlich viel veränderbarer erscheinen als vorher (variabel-kontrollierbare Attribution)? Empfindet er deshalb seine bisherigen schlechten Leistungen als Spiegel seiner Faulheit und fühlt sich schuldig (verantwortliche Attribution)? Oder ist er erleichtert, dass sein schwaches Abschneiden eher an seinem Gehirn liegt als an ihm selbst (nicht-verantwortliche Attribution)?

All diese Möglichkeiten haben eine gewisse Plausibilität, und doch widersprechen sich einige von ihnen. Ihre empirische Gegenüberstellung erscheint somit sinnvoll. Die praktische Relevanz einer solchen empirischen Untersuchung liegt darin begründet, dass Attributionsmuster mit Motivations- und Leistungsvariablen in Verbindung stehen. Internal-stabil-unkontrollierbare Attributionen nach Misserfolg sind mit dysfunktionalem Leistungsverhalten verbunden, internal-variabel-kontrollierbare hingegen mit funktionalem (vgl. z.B. Seligman, Abramson, Semmel, & Baeyer, 1979; Weiner, 1994). Die Verantwortlichkeitsdi-

mension ist in dieses Schema weniger leicht einzuordnen. Zu vermuten ist aber, dass die Attribution auf hohe Verantwortlichkeit eher mit dysfunktionalen Reaktionen verbunden ist: Sie impliziert die Internalisierung des Leistungsversagens von der psychosozialen Umwelt ins „Innere“ des Menschen (vgl. Sonntag, 1991) und verhindert gleichzeitig, dass der Rückbezug auf das Gehirn als Hilfe zur Externalisierung des Problems (vgl. White & Epston, 1990) vom Geist hin zum Gehirn verwendet wird. Eine Attribution auf geringe Verantwortlichkeit sollte in diesem Sinne also eine moralische Entlastung nach sich ziehen, die funktionale Auswirkungen auf die Person mit sich bringt. Assheuer (2005) stellt gar die These auf, dass es eben diese entlastende, „tröstende“ Komponente des Hirnforschungsdiskurses ist, die erst zu seiner enormen Verbreitung und Popularität geführt hat. Der Aufstieg der Hirnforschung sei nämlich einhergegangen mit gesellschaftlichen Veränderungen, die zunehmend mehr nach der Eigenverantwortung des Einzelnen verlangten. Die Deutung menschlichen Verhaltens im Vokabular der Hirnforschung federe diesen Eigenverantwortungsdruck gleichsam ab und werde deshalb aktiv gesucht. Führt man sich nun vor Augen, dass Kontrollierbarkeits- und Verantwortlichkeitsattributionen empirisch oft gemeinsam auftreten (Weiner, 1995), lässt sich schlussfolgern, dass sich der funktionale Effekt der Kontrollierbarkeitsattribution und der dysfunktionale Effekt Verantwortlichkeitsattribution teilweise neutralisieren könnten.

2.2.2 Untersuchungen im Umkreis von Dwecks Leistungsmotivationsmodell

Im obigen Beispiel wurde die Aufmachung und Gestaltung der Fernsehsendung offengelassen. Wurden mathematischen Fähigkeiten eher als etwas dargestellt, das in bestimmten genetisch oder frühkindlich geprägten Hirnarealen gespeichert ist, oder aber als flüchtiges Produkt eines dynamischen neuronalen Netzes, das sich ständig weiterentwickelt? Einige Untersuchungen zeigen, dass diese Frage tatsächlich für die Attribution und Verhaltensreaktion des Betroffenen entscheidend zu sein scheint. Diese Studien hatten die Prüfung des kausalen Einflusses Impliziter Intelligenztheorien (Dweck, 1999) zum Ziel und bedienen sich eher aus methodischen Gründen des Hirnforschungsdiskurses (vgl. Abschnitt 3.2). Sie werden in der folgenden Darstellung jedoch unter dem hier interessierenden Gesichtspunkt betrachtet und in ihren Schlussfolgerungen gewissermaßen umgedeutet.

Gemeinsam ist den Untersuchungen, dass sie ihre Probanden mit Ergebnissen der Hirnforschung zum Thema Intelligenz konfrontieren. Es wird entweder behauptet, dass sich gemäß neuen neurowissenschaftlichen Erkenntnissen Intelligenz als besonders stabil oder aber als besonders veränderbar herausgestellt habe. Damit soll den Teilnehmern eine leistungsthematische Attribution entweder auf Fähigkeit (internal-stabil-unkontrollierbar) oder auf Anstrengung (internal-

veränderbar-kontrollierbar) nahegelegt und entsprechendes Leistungsverhalten (s.o.) angeregt werden.

Blackwell, Trzesniewski und Dweck (2007, Studie 2) unterrichteten 99 Siebtklässler in einem relativ kurzen (8 x 25 Minuten) Workshop über den Aufbau des Gehirns und daraus abzuleitende Lernstrategien. Sie teilten die Schüler in zwei Gruppen auf und lehrten diesen in vier der acht Sitzungen unterschiedliche Lerninhalte. Zentrales Thema in der einen Gruppe war die enorme Plastizität des Gehirns und die damit einhergehende Veränderbarkeit der Intelligenz eines Menschen. In einprägsamen Videoclips sahen die Schüler z.B. die sich ständig verändernden Verknüpfungen zwischen den Nervenzellen. Ihnen wurde mit Hilfe dieser plastischen Illustrationen sowie vertiefender Diskussionen die Botschaft vermittelt, dass ihre eigenen Lernanstrengungen sich direkt in neue synaptische Verbindungen niederschlagen und somit ihre Intelligenz steigern. Das folgende Zitat, in dem die Autoren den Inhalt des Artikels und die anschließende Diskussion beschreiben, macht deutlich, dass das Gehirn nicht nur als veränderbar dargestellt wird, sondern dass diese Veränderbarkeit auch dezidiert *in Abhängigkeit willentlich gesteuerter Einwirkungen* zu verstehen ist (vgl. Abschnitt 2.1).

Der Artikel stellt die Veränderungen im Gehirn dar, die sich als Folge des Lernens vollziehen. Diese bestehen vor allem in der Neubildung und Stärkung von Verbindungen zwischen Nervenzellen. Wissenschaftliche Untersuchungen werden berichtet, die zeigen, wie geistige Aktivität messbare physische Veränderungen im Gehirn zur Folge hat. Der Artikel vergleicht das Gehirn mit einem Muskel, der durch Übung gestärkt werden kann, und zieht daraus die Schlussfolgerung, dass man durch Lernen schlauer wird.

Nach dem Lesen des Artikels wurde mit den Schülern eine Diskussion geführt, in der sie sich an Beispiele aus der Vergangenheit erinnern sollten, in denen sie durch intensive Lernanstrengungen Ziele erreicht und Kompetenzen aufgebaut hatten. Dabei wurde immer auch thematisiert, wie ihre Gehirne sich dabei verändert hatten und wie sie im Zuge dessen schlauer geworden waren (Blackwell, Dweck, & Trzesniewski, 2007, S. 262, Übersetzung durch mich, J.v.T.).

Die andere Gruppe erhielt in großen Teilen denselben Workshop, außer dass sich dieser in jenen vier Sitzungen nicht mit Intelligenz, sondern mit Gedächtnis befasste. Dabei vermittelte er ebenfalls einige nützliche Lernstrategien, etwa hinsichtlich der Frage, wie Lernmaterial am besten vom Kurzzeit- ins Langzeitgedächtnis übertragen werden kann.

Die Auswirkungen dieser beiden Interventionen waren erstaunlich unterschiedlich. Anhand von Fragebogendaten zeigte sich beispielsweise, dass die Schüler aus der Intelligenzgruppe ihre Lernergebnisse stärker auf die eigene Anstrengung zurückführten, während die Schüler der Gedächtnisgruppe eher auf ihre Fähigkeit attribuierten. Einige Zeit nach der Intervention wurden die etatmäßigen Lehrer über Verhaltensveränderungen ihrer Schüler befragt, die sie bereits ein Schuljahr lang kannten, aber über deren Gruppenzugehörigkeit sie nicht Be-

scheid wussten. Sie berichteten bei Schülern aus der Intelligenzgruppe deutlich häufiger über gesteigerte Motivation und Anstrengung als in der Gedächtnisgruppe. Von 17 Schülern, die sich gesteigert hatten, waren 13 in der Intelligenzgruppe. Die Lehrer umschrieben ihre Fortschritte wie folgt:

„L., der vorher nie in irgend etwas besondere Anstrengung investiert hatte und seine Hausaufgabe selten rechtzeitig abgab, stand nun tatsächlich früh auf, um stundenlang an einer Aufgabe zu arbeiten, damit ich sie durchsehen und ihm die Möglichkeit zu Korrekturen geben konnte. So bekam er, der sonst immer nur Vieren und Fünfen bekommen hatte, eine Zwei auf die Aufgabe.“

„M.s Leistungen waren schon seit langem unterdurchschnittlich. In den letzten Wochen bat sie mich plötzlich wiederholt um eine Nachhilfestunde in der Mittagspause. Ihre Noten verbesserten erheblich, in der letzten Klausur schrieb sie eine gute Zwei“ (Blackwell, Dweck, & Trzesniewski, 2007, S. 256, Übersetzung durch mich, J.v.T.).

Am Ende des Schuljahres, etwa ein Vierteljahr nach der Intervention, war die Intelligenzgruppe der Gedächtnisgruppe hinsichtlich der Mathematik-Zeugnisnote signifikant überlegen. Die recht kurze Intervention zog anscheinend bei der Intelligenzgruppe einen so nachhaltigen Motivationsschub nach sich, dass dieser sich deutlich in der Abschlussnote niederschlug.

Dweck (2007a) berichtet von kürzlich abgeschlossenen Untersuchungen, die mit einer computerbasierten Version dieses Workshops durchgeführt wurden. Das Computer-Programm trägt den Namen „Brainology“ und wurde an 20 New Yorker Schulen eingesetzt. Die Schüler setzen sich dabei mit dem Aufbau und der Funktionsweise des Gehirns interaktiv auseinander und führen virtuelle Experimente durch, in denen sie sehen, wie sich das Gehirn beim Lernen verändert und wie man diesen Prozess aktiv beeinflussen kann. Wiewohl die Ergebnisse dieses ambitionierten Projekts noch nicht veröffentlicht sind, scheinen sie laut Dweck (2007a) ähnlich auszufallen wie in der Studie von Blackwell et al..

Auch andere, ähnlich gestaltete Untersuchungen bestätigen diese Befunde. Aronson, Fried und Good (2002) führten eine Laboruntersuchung an 79 Studenten durch, denen der Zweck des Experiments – anders als bei Blackwell et al. – verborgen bleiben sollte. Der Gegenstand der Studie wurde als „long-distance learning“ deklariert, wobei die Teilnehmer die Rolle eines Tutors für jüngere Schüler übernehmen sollten. In kleinen Gruppen wurden den Tutoren nun diejenigen Lerninhalte vermittelt, die sie nachher ihrerseits über schriftlichen Kontakt ihren Schülern überbringen sollten. Dieses Training verlief ähnlich wie das bei Blackwell et al.. Ein Teil der Tutoren wurde angewiesen, ihre Schüler mit Hilfe der oben erläuterten Plastizitätsmetapher zu motivieren, während der andere Teil die Idee verwenden sollte, dass es verschiedene, voneinander unabhängige Formen von Intelligenz gibt und dass demzufolge jeder in bestimmten Bereichen Stärken hat. Nach diesem Training kommunizierten die Tutoren tatsächlich über Briefe mit ihren (fiktiven) Tutanden. Einige Monate später wurde in Form von

Fragebögen (die den Teilnehmern freilich als unabhängige Untersuchung nahegebracht wurden) überprüft, inwieweit die Tutoren von der zu überbringenden Botschaft selbst beeinflusst worden waren. Die Plastizitätsgruppe berichtete über mehr Lernfreude sowie mehr Wertschätzung von Lernen und Leistung als die andere Gruppe. Auch im Vergleich zu einer Kontrollgruppe, die nur die Fragebögen ausgefüllt hatte, zeigten sich diese Effekte. Am Ende des Studienjahres war die Plastizitätsgruppe den anderen beiden Gruppen auch in der Durchschnittsnote überlegen. Interessanterweise profitierten innerhalb der Plastizitätsgruppe afroamerikanische Studenten besonders stark von dem Training (vgl. hierzu Abschnitt 3.2.2).

Good, Aronson und Inzlicht (2003) bestätigten diese Ergebnisse anhand eines ähnlichen Trainings (übrigens hier mit besonders positiven Effekten bei weiblichen Teilnehmern, vgl. ebenfalls Abschnitt 3.2.2). Bei ihnen ging es jedoch um Siebtklässler, die über ein Jahr lang von einem studentischen Tutor in ihrer schulischen Entwicklung begleitet wurden. Alle Schüler erhielten Ratschläge hinsichtlich allgemeiner Themen, etwa der Eingewöhnung in die neue Schule (Junior High School) oder nützlicher Lernstrategien. Die Argumentationshilfen, die die Tutoren dabei verwendeten, unterschieden sich jedoch: In einer Gruppe wurden Leistungsprobleme mit Hilfe der Plastizitätsmetapher erklärt, in einer weiteren auf normale Umgewöhnungsschwierigkeiten durch den Schulwechsel zurückgeführt, während in einer dritten Gruppe beide Erklärungen kombiniert wurden. In einer Kontrollgruppe wurde keine der beiden Argumentationshilfen zugelassen und stattdessen auf das Thema Drogenprävention fokussiert. Zur Festigung der jeweiligen Botschaft sollten die Schüler Internetseiten mit der entsprechenden Botschaft gestalten. Am Ende des Schuljahres zeigte sich die Plastizitätsgruppe der Kontrollgruppe in schulischen Leistungstests überlegen. Allerdings waren die anderen beiden Experimentalgruppen ebenso wirksam im Vergleich zur Kontrollgruppe.

Dennoch lässt sich in der Zusammenschau der geschilderten Untersuchungen die nachhaltige Wirksamkeit der Plastizitätsmetapher konstatieren. Die Vorstellung, man könne auf das Gehirn durch gezieltes Training einwirken wie auf einen Muskel und so seine Intelligenz steigern, scheint eine erstaunliche Attraktivität und Einprägsamkeit zu besitzen. Dass man mit solchen Vorstellungen Schüler motivieren kann, ist dabei weniger verwunderlich als die *Deutlichkeit und Nachhaltigkeit* der Wirkung.

These interventions were relatively modest, compared with many lengthy and costly educational interventions. This fact brings home the idea that when one pinpoints a belief that is at the core of a motivational pattern, simply altering that belief (and demonstrating how this belief can be put into practice) can have a cascade of beneficial effects (Dweck & Grant, 2006, S. 411).

Die Konfrontation mit der Plastizitätsmetapher scheint also den Glauben zu bestärken, dass eine substanzielle Steigerung akademischer Leistung und Leistungsfähigkeit durch Lernen und Anstrengung möglich ist. Müsste dann nicht auch die entgegengesetzte Botschaft, dass also eine solche Steigerung aufgrund fester, unveränderbarer Hirnstrukturen ausgeschlossen ist, motivational abträgliche Konsequenzen haben? Dieser Nachweis wurde bisher nur in Laborstudien erbracht, die auf kurzfristige Effekte hin angelegt sind, da die bewusste Induktion langfristiger Motivationsprobleme ethisch fragwürdig wäre. Dar-Nimrod und Heine (2006) führten ein solches Laborexperiment durch. Sie gaben ihren Probandinnen fingierte Zeitschriftenartikel, die über die Gründe für das (angeblich) schlechtere Abschneiden von Frauen in mathematischen Aufgaben berichten. Entweder wurden diese Gründe in genetisch bedingten Unterschieden im Hirnstoffwechsel oder in der Sozialisation lokalisiert. Probandinnen in der ersteren Bedingung lasen u.a. folgendes:

Die Forscher konnten zeigen, dass zwei Gene auf dem Y-Chromosom (das sich nur bei Männern findet) mit Genen auf Chromosom 5 und 7 interagieren. Im Zuge dessen veranlasst der Hypothalamus hormonelle Veränderungen in einem Seitenteil des Frontallappens, der während mathematischer Denkvorgänge aktiviert ist. F-MRI-Untersuchungen zeigen, dass diese hormonellen Veränderungen eine Steigerung der Ausschüttung von ATP-Molekülen bewirken. Dieses ATP, das als die Währung körpereigener Energie schlechthin gilt, kann nun den Hippocampus versorgen, der bei schwierigen geistigen Tätigkeiten aktiv ist. Es sorgt dafür, dass der Person mehr Speicher im Kurzzeitgedächtnis zur Verfügung steht und fördert so die mathematische Problemlösefähigkeit. Diese Befunde scheinen die Unterschiede in mathematischen Leistungen zwischen Männern und Frauen, die durchschnittlich bei 5 Perzentilpunkten liegen, zu erklären (Online-Zusatzmaterial zu Dar-Nimrod & Heine, 2006, Quellenangabe siehe dort; freie Übersetzung von mir, J.v.T.).

Die Versuchsgruppe, der diese genetische Erklärung angeboten worden war, schnitt in den anschließenden Mathematikaufgaben schlechter ab als die andere, sozialisationsgeprägte Gruppe, der u.a. vermittelt worden war, dass neurowissenschaftliche Untersuchungen keine Unterschiede zwischen dem männlichen und dem weiblichen Gehirn ausmachen konnten. Dar-Nimrod und Heines Studie, die übrigens im renommierten Fachblatt „Science“ publiziert wurde und ihrerseits breite mediale Aufmerksamkeit erfuhr, gibt also erste Hinweise darauf, dass das neurowissenschaftliche Vokabular auch pädagogisch abträgliche Folgen haben kann.

Unter dem Gesichtspunkt seiner Auswirkungen auf das Lern- und Leistungsverhalten scheint der Hirnforschungsdiskurs demnach in zwei Arten von Aussagen unterteilbar zu sein. Das eine Mal erscheint das Gehirn als relative feste Struktur, die Leistungsergebnisse vorhersagbar macht. Dieser *Stabilitätsdiskurs* umfasst alle Thesen, die die genetische, vorgeburtliche oder frühkindliche Prägung des Gehirns betreffen und so die Lernmotivation von Kindern, Jugendlichen und Erwachsenen untergraben kann. Auch die Rede von „sensiblen Pha-

sen“, in denen eine Fähigkeit besonders leicht bzw. gut erlernbar ist, passt zu diesem Diskurs, weil diese Phasen ja in der Regel ebenfalls in frühen Entwicklungsstadien lokalisiert werden. Das andere Mal wird das Gehirn als Sinnbild der Veränderung dargestellt, als ständig im Fluss befindliche Substanz, die sich gerade nicht nach Genen, sondern nach Erfahrungen richtet. Dieser anscheinend der Lernmotivation zuträgliche *Veränderbarkeitsdiskurs* bezieht sich meist auf eher neuronale Prozesse im Mikro-Bereich, während der Unveränderbarkeitsdiskurs meist die übergreifende Hirnarchitektur zum Gegenstand hat.

Während die zitierten empirischen Untersuchungen nur auf die Unterscheidung Stabilität vs. Veränderbarkeit abheben, zeigt sich bei genauerer Analyse, dass bei den dort verwendeten Lehrmaterialien immer auch zwischen *epiphänomenalistischer* und *klassisch-geisteswissenschaftlicher* Sichtweise unterschieden wird. Der Veränderbarkeitsdiskurs geht hier mit einer Sichtweise einher, die das Gehirn als Gegenstand gezielter Einwirkung oder bestenfalls als Korrelat psychischer Prozesse betrachtet, während der Stabilitätsdiskurs die epiphänomenalistische Sichtweise vertritt, wonach psychische Prozesse nichts weiter als Folgescheinungen neuronaler Vorgänge sind. Beide Unterscheidungen sind aber nicht zwingend aneinander gekoppelt (z.B. ist auch ein epiphänomenalistischer Veränderbarkeitsdiskurs denkbar, wonach sich das Gehirn aufgrund von naturwissenschaftlichen Einflüssen verändert) und sollten daher nicht ohne Not vermengt werden.

Diese Gegenüberstellung lässt die Frage aufkommen, wie in deutschen Medien über neurowissenschaftliche Befunde zum Thema Lernen und Intelligenz tatsächlich berichtet wird. Wird das Gehirn eher dargestellt als unkontrollierbare Substanz, mit der wir leben müssen, oder als stoffliches Korrelat unserer selbst, über das wir uns verändern können?

2.2.3 Mediale Darstellungen neurowissenschaftlicher Intelligenzforschung

Zur Beantwortung dieser Frage habe ich die letzten zehn Jahrgänge (1998-2007) der Zeitschriften *Der Spiegel*, *Focus* und *Die Zeit* untersucht. Tabelle 1 zeigt alle dort aufgefundenen Beiträge, die auf der jeweiligen Internet-Homepage unter den Suchbegriffen „Gehirn Intelligenz“ sowie „Gehirn Lernen“ auffindbar waren und deren zentrales Thema erkennbar die neurowissenschaftliche Lern- und Intelligenzforschung betraf. Damit blieben freilich diejenigen Artikel zur Hirnforschung außen vor, in denen Lernen und Intelligenz eher implizit oder nebenbei zum Thema wurde. Somit fanden beispielsweise Beiträge zu populären Themen wie ADHS, gedächtnissteigernde Substanzen oder Willensfreiheit keine Berücksichtigung. Die hier verwendeten Artikel sind also als Stichprobe zu verstehen, die nur die sichtbare Spitze eines in Wirklichkeit wohl viel umfangreicheren Eisbergs darstellt.

Tabelle 1

Übersicht über die untersuchten Beiträge aus Der Spiegel, Focus und Die Zeit (Jahrgänge 1998-2007), in chronologischer Reihenfolge

Quelle	Autor	Titel	Thema	vermittelte Auffassung über	
				„Intelligenz“	Hirnprozesse
Focus 13/2000	Sanides	Neue Zellen für neues Denken	Plastizität	veränder- bar	geistes- wissenschaftlich
Focus 16/2000	Röper, Miket- ta und Stol- lorz	Das Lexikon der Gene	Einfluss der Gene auf (u.a.) Gehirn und Intelligenz	stabil	epiphäno- menalistisch
Zeit 24/2000	Klein	Training fürs Köpfchen	Neurodidaktik	stabil	--
Spiegel 30/2000	Blech (vgl. Abschnitt 4.2)	Gral der Geis- teskraft	Verortung der In- telligenz im Ge- hirn	stabil	epiphäno- menalistisch
Focus 51/2001	Wittlich (In- terview mit d. Neurobiolo- gen Pöppel)	Die ersten zehn Jahre entscheiden	Neurodidaktik	stabil	epiphäno- menalistisch
Spiegel 11/2002	Thimm (Inter- view mit dem Neurologen Hüther)	Funkstille im Frontalhirn	Hirnentwicklung im Kontext von ADHS	--	--
Spiegel 27/2002	Thimm (Titel- story)	Guten Mor- gen, liebe Zahlen	Neurodidaktik	--	--
Spiegel 27/2002	Thimm (Inter- view mit dem Neurobiolo- gen Scheich)	„Begeisterung diszipliniert“	Neurodidaktik	--	--
Focus 43/2002	Siefer	Was Sy- napsen wün- schen	Neurodidaktik	stabil	--
Focus 43/2002	Plewnia und Reinke- Nobbe	Wie man Bil- dung schafft	Lernen im Kon- text gesellschaft- lichen Wandels	stabil	--
Zeit 48/2002	Eberle (Port- rait d. Neuro- biologin Eliot)	Die Hirnerzie- herin	Determiniertheit	stabil	epiphäno- menalistisch

(Fortsetzung siehe nächste Seite)

Fortsetzung Tabelle 1

Quelle	Autor	Titel	Thema	vermittelte Auffassung über	
				„Intelligenz“	Hirnprozesse
Zeit 15/2003	Kahl (Portrait der Psycholo- gin Stern)	Was Hän- schen lernt	Frühkindliches Lernen	veränder- bar	geistes- wissenschaftlich
Focus 27/2003	? (Interview mit dem Hirn- forscher Spit- zer)	Das Gehirn bastelt Visio- nen	Determiniertheit (Größe des Fron- talhirns bestimmt Erfolg im Leben)	stabil	epiphäno- menalistisch
Zeit 38/2003	Paulus	Lernrezepte aus dem Hirnlabor	Neurodidaktik	--	--
Zeit 39/2003	Spitzer (Neu- robiologe)	Medizin für die Pädago- gik	Neurodidaktik	stabil	--
Zeit 39/2003	Scheich (Neurobiolo- ge)	Lernen unter der Dopa- mindusche	Neurodidaktik	--	--
Zeit 40/2003	Stern (Psy- chologin)	Rezepte statt Rezeptoren	Neurodidaktik	--	--
Spiegel 39/2003	Hoppe	Das gierige Gehirn	Besondere Fähig- keiten sogenann- ter Inselbegabter	stabil	epiphäno- menalistisch
Spiegel 43/2003	Thimm (Titel- story)	Jeden Tag ein neues Universum	Hirnforschern und Psychologen über frühkindliches Lernen	stabil	epiphäno- menalistisch
Focus 47/2003	? (Interview mit dem Psy- chologen Neubauer)	Ständiges Training ist nötig	Veränderbarkeit des IQ	veränder- bar	geistes- wissenschaftlich
Focus 47/2003	Kunz, Siefer und Sanides	Damit der Geist blitzt	Beeinflussbarkeit der Intelligenz ü- ber Training, Um- welt und Tabletten	--	--
Focus 2/2004	Siefer (Inter- view mit dem Genetiker Plomin)	Ich suche nur IQ-Gene	Zusammenspiel von Genen und Umwelt bei der In- telligenz	--	--
Spiegel 31/2004	Koch	Feindliche Übernahme	Neurodidaktik	--	--

(Fortsetzung siehe nächste Seite)

Fortsetzung Tabelle 1

Zeit 28/2004	Kerstan und von Thadden (Interview mit Neurobiologen Spitzer und Psychologin Stern)	Wer macht die Schule klug	Neurodidaktik	--	--
Zeit 7/2005	Schnabel	Knetmasse der Kultur	Plastizität	veränderbar	geisteswissenschaftlich
Zeit 6/2006	Degen	Denken hilft	Plastizität	veränderbar	geisteswissenschaftlich
Spiegel 20/2006	Blech (Titelstory)	Hirn, kuriere dich selbst!	Plastizität	veränderbar	geisteswissenschaftlich
Zeit 42/2006	Wüsthof	Bei Schlaun reift die Rinde langsam	Verortung d. Intelligenz im Gehirn	stabil	epiphänomenalistisch
Spiegel 35/2007	Lakotta/Thimm (Interview mit Hirnforscher Roth)	Das Ich ist eine Einbahnstraße	Determiniertheit	stabil	epiphänomenalistisch
Focus 7/2007	Siefer	Wellenreiten ins Gehirn	Wirksamkeit des Neurofeedbacks	veränderbar	geisteswissenschaftlich
Spiegel 12/2007	Grolle und Thimm (Interview mit Hirnforscher Kempermann und Psychologen Lindenberger)	Altern beginnt in der Wiege	Hirnforscher und Psychologen über das Altern	veränderbar	geisteswissenschaftlich
Zeit 33/2007	Schramm	Mein Bauch gehört dir	Vorgeburtliche Einflüsse auf die Entwicklung	stabil	--
Focus 34/2007	?	Pflegetips für Ihr Gehirn	Tips zur Vermeidung von Gedächtnisproblemen	veränderbar	geisteswissenschaftlich

Anmerkungen.

Die Suche wurde anhand der Suchmasken folgender Internet-Seiten durchgeführt (Zugriff: 4.3.2008):

- Der Spiegel: <http://wissen.spiegel.de/wissen/start/home.html?sm=e>
- Focus: <http://www.focus.de/magazin/archiv/suche>
- Die Zeit: <http://www.zeit.de/archiv/index>

Weitere Erläuterungen im Text.

Tabelle 1 fasst demnach alle Artikel zusammen, die sich erstens mit Lernen und Intelligenz beschäftigen und dies zweitens mit Argumenten der Hirnforschung tun. Artikel, die die erste, nicht aber die zweite Bedingung erfüllen, sind allerdings ohnehin sehr selten. Beiträge etwa, die sich ausführlich über frühkindliche Bildung äußern und dabei ohne Hirnforschung auskommen, sind nur zwischen 1999 und 2001 zu finden (vgl. *Focus* 3/1999, 10/2000 und 33/2001). Dazu passt, dass Tabelle 1 keine Einträge für die Jahre 1998 und 1999 aufweist. Der Hirnforschungsdiskurs scheint tatsächlich erst seit Anfang des neuen Jahrtausends in deutschen Medien präsent zu sein (vgl. Kuhlmann, 2004).

Die meisten der gefundenen Artikel lassen sich einem von drei Themen zuordnen (vgl. die dritte Spalte in Tabelle 1). Erstens gibt es Beiträge, die sich mit neurowissenschaftlichen Untersuchungen beschäftigen, die *die genetisch bzw. frühkindlich geprägten Hirnstrukturen als unveränderlich darstellen und somit Lernfähigkeit und Intelligenzentwicklung auf die ersten 3-5 Lebensjahre beschränkt* sehen („Determiniertheit“). Artikel der zweiten Kategorie („Plastizität“) beschreiben Studien, die genau in die Gegenrichtung weisen, indem sie die Bildung neuer Synapsen und Neuronen auch im Jugend- und Erwachsenenalter noch häufig und nachhaltig vorkommen sehen. Drittens geht es vielen Artikeln um die sogenannte „Neurodidaktik“, also um die Implikationen der Ergebnisse der Hirnforschung für Erziehung und Schule.

Da auch die Artikel über Neurodidaktik sowie die übrigen Artikel teilweise über die Frage der Determiniertheit vs. Plastizität Auskunft geben, habe ich in der vierten Spalte alle Artikel daraufhin untersucht, welcher von beiden Auffassungen sie eher zuneigen. Von insgesamt 32 Artikeln stellen 14 geistige Fähigkeiten als stabil dar, neun vertreten die Gegenposition der Veränderbarkeit, während zehn diesbezüglich keine oder eine ausgewogene Position einnehmen. Auch die andere in Abschnitt 2.2.2 getroffene Unterscheidung habe ich auf die Artikel anzuwenden versucht (vgl. die fünfte Spalte in Tabelle 1). Neun Artikel vertreten eher eine epiphänomenalistische Position, acht eine geisteswissenschaftliche, 15 gar keine oder eine ausgewogene. Erwartungsgemäß sind beide Unterscheidungen empirisch eng miteinander verknüpft. Ein Blick auf den zeitlichen Verlauf macht deutlich, dass bis 2003 die stabil-epiphänomenalistische Position vorherrscht, während ab 2005 ein Umschwung zur veränderbar-geisteswissenschaftlichen Auffassung zu beobachten ist.

In Tabelle 2 sind einige Zitate aufgeführt, die illustrieren sollen, wie in den einzelnen Artikeln die erwähnten Positionen dargestellt und vertreten werden. Dabei zeigt sich, wie gegensätzlich und einander widersprechend diese zuweilen geäußert werden. Dies legt die Vermutung nahe, dass jene Artikel eine jeweils einseitige Interpretation der Befundlage vornehmen. Der folgende Abschnitt versucht näher zu untersuchen, inwieweit die Befundlage in der Hirnentwicklungsforschung tatsächlich umstritten ist.

Tabelle 2

Textbeispiele aus den untersuchten Artikeln für die in ihnen geäußerten Auffassungen über Intelligenz und Hirnprozesse

stabil-epiphänomenalistische Auffassung	veränderbar-geisteswissensch. Auffassung
Es tickt eine Uhr für jeden Säugling. So genannte Entwicklungsfenster bestimmen, welche Fähigkeiten er wann erlangt. (...) Nach drei Jahren endet der erste Entwicklungsschub. In einem zweiten, kurz vor der Pubertät, festigen sich in der präfrontalen Hirnrinde Verbindungen, mit deren Hilfe der Mensch Handlungen aufschieben kann. Er lernt moralische Werte und soziale Routine. Doch dann ist Schluß. Allenfalls Kleinigkeiten lassen sich im Gehirn noch verändern (...). Ein gänzlich neues Gerüst synaptischer Verbindungen kann der Mensch ebenso wenig knüpfen wie das bestehende wegradieren. Er muss mit dem leben, was er hat. „Entscheidend für die Zukunft eines Kindes ist also, was es in den ersten Jahren erlebt“, fasst Singer die bittere Wahrheit zusammen. „Man kann ein Menschenleben voll entfalten - oder zerstören. Verpasste Momente für den Spracherwerb oder für feinmotorische Bewegung lassen sich nie wieder nachholen.“ Wer in deutsche Grundschulen geht, trifft eine Menge Jungen und Mädchen, die solche Momente scheinbar versäumt haben (<i>Der Spiegel</i> 43/2003, S.204-205).	In unserem Kopf sitzt ein Verwandlungskünstler. Das Hirn ist keine starre Masse, es ist formbar wie Knete. Seine Zellen sind nicht von Kindheit an abgezählt, sondern in manchen Arealen bis ins Greisenalter erneuerbar. Und die Nervenzellen vernetzen sich nicht nur bis zum Ende der Pubertät, sondern auch noch Jahrzehnte danach. „Erst jetzt können wir genau herausfinden, wie sich unser Verhalten auf das Gehirn auswirkt“, sagt Gerd Kempermann (...) „auf seine Funktionen, auf seine anatomische Struktur, bis hin zu den einzelnen Zellen.“ (...) Mit jedem Lernprozess werden neue Verbindungen zwischen den Nervenzellen geknüpft und alte verstärkt. Nicht nur bei Kindern, sondern eben auch noch bei Erwachsenen (...). Durch die zusätzlichen Nervenbahnen werden Informationen offenbar schneller verarbeitet. Auch die Konzentration der Botenstoffe steigt, mit deren Hilfe die Zellen kommunizieren. Das Großhirn passt sich neuen Ansprüchen an. (...) „Plastizität“ nennen Neurowissenschaftler dieses kleine Wunder, das sich jeden Tag in unserem Kopf abspielt (<i>Die Zeit</i> 6/2006, S. 38).
Darum sei es so wichtig (...), dass das kleinkindliche Nervensystem frühzeitig stimuliert werde: „Synapsen, die nicht angeregt werden – weil eine Sprache nicht gehört, ein Instrument nicht gespielt, ein Sport nicht geübt oder eine Liebe nicht gespürt wird – sterben“, schreibt Eliot. Und was hätte sein können – akzentfreies Russisch, tonreines Singen, gesundes Selbstvertrauen – wird nimmermehr. Die Moral ist klar: Wer nicht stimuliert wird, steht dumm da, buchstäblich (<i>Die Zeit</i> 48/2002, S. 36).	Auch das ist eine Folge des Irrglaubens, wir kämen mit einer fertigen Ausstattung an grauen Zellen auf die Welt - als wären unsere Hirne Computer, deren Speicherchips einmal zusammengebaut wurden und auf die in der Kindheit nur die Software geladen wird. In Wahrheit aber verwandelt sich das Gehirn das ganze Leben hindurch: Hier wachsen Zellen nach, dort sterben andere ab. Deswegen kann das Gehirn bis ins hohe Alter lernen. Nur muss es dazu auch angeregt werden (<i>Die Zeit</i> 24/2000, S. 37, T. Sejnowski im Interview).

(Fortsetzung siehe nächste Seite)

Fortsetzung Tabelle 2

stabil-epiphänomenalistische Auffassung	veränderbar-geisteswissensch. Auffassung
Man kann sich Disziplin von außen ankonditionieren lassen, aber kaum von selbst einhalten. Es ist schwer, gegen unbewusste eigene Kommandos anzugehen. (...) [Das Unbewußte] ist eine Urform unseres Selbst; die psychische Grundausrüstung, mit der wir auf die Welt kommen. Daraus entsteht alles andere. Man kann es auch Temperament nennen. (...) Die Gene bestimmen zwischen 20 und 50 Prozent der Persönlichkeit eines Menschen. Aber auch das, was eine Frau während der Schwangerschaft erlebt, entscheidet mit über das Temperament eines Kindes: Ob es offen oder ängstlich sein wird, ein stabiles oder ein zaghafte Ego entwickelt, ob es pedantisch ist oder lässig. Diese Weichen stellt das limbische System, eine Art Schaltzentrale der Gefühle, das ab der sechsten Schwangerschaftswoche entsteht. In dieser Zeit erlebt das Ungeborene die ersten emotionalen Konditionierungen, die sein Gehirn für sein ganzes Leben prägen. Steht die Mutter etwa durch furchtbare Erlebnisse unter Stress, dann wirken Botenstoffe aus ihrem Körper direkt auf das kindliche Gehirn ein. (...) Der Spielraum, in dem ein Mensch überhaupt empfinden, sich verhalten und sich verändern kann, ist bei der Geburt bereits in beträchtlichem Maße umrissen. (...) Erziehung ist immer der Versuch, dieses entweder verbitterte oder großartige, narzisstische Kleinkind an die Welt anzupassen. (...) Zur Kreativität und Intelligenz kann man kaum erziehen, die sind hochgradig angeboren (<i>Der Spiegel</i> 35/2007, S. 124 Gerhard Roth im Interview).	Ein Jungbrunnen im Gehirn? Neue Denkkraft dank neuer Zellen? Bis vor kurzem noch hätte man Macklis und Goldberg als Phantasten abgetan. Denn [sie] rütteln an einem Dogma, das ein Jahrhundert lang als unumstößlich galt: Schon im Säuglingsalter hören Nervenzellen demnach auf, sich zu teilen. (...) Doch nun tragen Neurologen, Biochemiker und Ärzte immer mehr Hinweise für einen gegenläufigen Trend zusammen: Stund um Stund kommen unter dem Schädeldach junge Nervenzellen auf die Welt. Verdutzt und voller Ehrfurcht erkennen die Forscher: Die Neubildung der Nervenzellen, wissenschaftlich Neurogenese genannt, hält bis ins Greisenalter an und scheint unentbehrlich für das normale Funktionieren des Denkkorgans. „Wir fangen jetzt an, das Gehirn aus einer völlig neuen Perspektive zu sehen“, urteilt Gerd Kempermann (...) „Es gibt da eine positive Tendenz: Die Entwicklung des Gehirns hält ein Leben lang an.“ (...) „Die Leute halten das Gehirn für einen unveränderlichen Computer“, sagt der Neurowissenschaftler Fred Gage (...). „Dabei ist es ein formbares Organ aus Fleisch, Blut und Nervenzellen. Veränderungen in diesem Organ kann man selbst kontrollieren.“ (...) Die wenigen, aber ungemein vielseitigen Neu-Neuronen haben maßgeblich Anteil daran, dass sich das Gehirn das ganze Leben hindurch verformen kann. So wie ein Muskel unter Belastung wächst, so gedeihen die grauen Zellen, wenn man sie fordert (<i>Der Spiegel</i> 20/2006, S. 164). Unser Wissen und unser Verhalten ändern sich ja beständig im Lauf des Lebens, und das Gehirn stellt sich immer wieder darauf ein, (...) und das ist die Grundlage unserer lebenslangen Lern- und Leistungsfähigkeit. Das Gehirn ist kein festverdrahteter Computer. (<i>Der Spiegel</i> 12/2007, S. 152 Gerd Kempermann und Ulman Lindenberger im Interview).

2.2.4 Die wissenschaftliche Angemessenheit der Medienberichte

Thompson und Nelson (2001) untersuchen in der angesehenen Fachzeitschrift „American Psychologist“ die Beziehung zwischen Forschern, die sich mit der Entwicklung des Gehirns beschäftigen, und den darüber berichtenden US-Medien. Sie identifizieren drei Hauptaussagen, die in der Berichterstattung der Medien hervorgehoben werden. Erstens hätten frühe Erfahrungen des Säuglings eine weit größere Bedeutung für seine spätere Entwicklung als bislang angenommen. Zweitens gebe es viele wichtige Kompetenzen, die nur in ganz bestimmten, sogenannten sensiblen oder kritischen Phasen der Entwicklung erlernt werden können. Drittens sei die sozial-emotionale Beziehung zwischen Säugling bzw. Kleinkind und den Bezugspersonen von zentraler Bedeutung für die spätere psychische Gesundheit des Heranwachsenden. Damit sei zwar einerseits die öffentliche Wahrnehmung der Bedürfnisse von Säuglingen und Kleinkindern gestärkt und die Bedeutung verantwortungsvoller Erziehungspraktiken hervorgehoben worden. Andererseits geben die Autoren zu Bedenken, dass „these media accounts have also tended to exaggerate how much is actually known about the developing brain, overinterpreting current scientific knowledge and overpromising its applications while also unduly narrowing public understanding of the range of important influences on brain development“ (S. 7). Worin die Übertreibungen und Verengungen genau bestehen, arbeiten Thompson und Nelson in ihrem Beitrag dann heraus und fassen schließlich wie folgt zusammen:

Critical periods are exceptional rather than pervasive in human brain growth. Windows of opportunity for early stimulation better characterize basic sensory and motor capacities than higher mental and personality processes, and even so, most such windows close very slowly with development. The relative influence of genetic guidance and experiential exposure in shaping the young brain is complex, and it is unclear what specific experiences are important, and when they must occur, for healthy development. The first three years of life are significant, but other periods are also important, and the brain retains its capacity to grow throughout life (Thompson & Nelson, 2001, S. 12).

Für jene Übertreibungen und Verengungen machen Thompson und Nelson nicht nur die Medienvertreter verantwortlich, sondern auch die um Aufmerksamkeit und Fördergelder konkurrierenden Forscher. Auch über die Folgen einer solch verzerrten Berichterstattung bei ihren Rezipienten machen sie sich übrigens Gedanken: „Brain development can be facilitated not only during the first three years but also at other developmental stages. This is important news for adoptive and foster parents, child-care providers, and parents and teachers of children of all ages“ (S. 12). Bei diesen eher oberflächlichen Überlegungen bleiben Thompson und Nelson allerdings auch stehen.

Dennoch ist es interessant, die Situation in den USA, wo die forcierte Berichterstattung über Hirnforschung schon fünf bis zehn Jahre früher einsetzte als in Deutschland, noch kurz näher zu betrachten. Paulus (2003) fasst die Situation lakonisch wie folgt zusammen:

...dass der erste Versuch weitgehend gescheitert ist, die Kinderbetreuung mithilfe der Neurobiologie zu revolutionieren. Vor allem in den USA hämmerten offiziöse Reporte wie populäre Ratgeber den Eltern ein, die Gehirne der Ein- bis Dreijährigen mit möglichst vielen Anregungen zu versorgen. Ansonsten drohe eine „mentale Wüstenei“, wie ein amerikanischer Autor warnte. Deutsche Erziehungsberechtigte dürfen sich mitgruseln. So propagiert das vor zwei Jahren erschienene Buch *Baby Brain* auch hierzulande Parolen wie „Synapsenpflege“, „Jetzt oder nie!“ und „Die Zukunft Ihres Babys liegt in Ihren Händen“. Praktisch heißt das: „Singen Sie Ihrem Baby ein neues Lied vor, entsteht eine neuronale Verknüpfung. Kitzeln Sie seine Zehen, und Sie haben wieder eine neue.“ Vor allem aber: „Bleiben Sie ganz entspannt.“ Letzteres ist der beste Tipp, denn so entscheidend sind die ersten drei Jahre auch wieder nicht. „Wir haben den Mythos ‚Geburt bis drei‘ beerdigt“, fasste US-Kinderpsychologe Mike Posner bereits im Juni 2000 die Ergebnisse einer hochkarätig besetzten OECD-Konferenz zusammen (Paulus, 2003, S. 35).

Dass in den medial aufbereiteten Thesen der Hirnforscher allgemein ein großer Anteil an Spekulation steckt, war des Weiteren bereits in Abschnitt 2.0 thematisiert worden. Daher sei hier auf weitere Belege verzichtet und abschließend lediglich noch auf zwei neurodidaktische Ratschläge angesehener deutscher Professoren hingewiesen, die das Gesagte besonders anschaulich illustrieren mögen:

In den allermeisten Fällen dürfte es aber genügen, darauf zu vertrauen, daß die jungen Gehirne selbst am besten wissen, was sie in verschiedenen Entwicklungsphasen benötigen und dank ihrer eigenen Bewertungssysteme kritisch beurteilen und auswählen können (Singer, 2002).

Kein vernünftiger Mensch kann annehmen, dass Unterrichtsformen, in denen Kinder sich ihre Ziele vollständig selbst setzen, zu etwas führen. Ihre sich entwickelnden Gehirne können das gar nicht leisten; es fehlt noch die Struktur, die dem Kind signalisiert: Dieses und jenes neue Wissenshäppchen ist wichtig (Henning Scheich im Interview, *Der Spiegel* 27/2002, S. 76).

2.3 Fazit

Die vorangegangenen Abschnitte lassen sich zu den folgenden fünf Schlussfolgerungen verdichten:

Erstens: Empirische Untersuchungen legen nahe, dass Hirnforschungsdiskurse im Lern- und Leistungsbereich besonders dann einen Einfluss auf das Verhalten von Schülern und Studenten haben, wenn sie Aussagen über die Formbarkeit der Intelligenz treffen. Eine stabile Intelligenzauffassung hat eher negative, eine veränderbare Intelligenzauffassung eher positive Auswirkungen auf Lernerfolg und Leistung (vgl. Abschnitt 2.2.2).

Zweitens: Diese Aussagen sind in deutschen Medien recht verbreitet. Dabei wird geistige Leistungsfähigkeit in den meisten Fällen als stabil dargestellt, nicht selten aber auch als veränderbar (vgl. Abschnitt 2.2.3).

Drittens: Dass sich die Thesen von der Stabilität bzw. der Veränderbarkeit der Intelligenz widersprechen, legt die Vermutung nahe, dass jene Artikel eine jeweils einseitige Interpretation der Befundlage vornehmen. Die Formbarkeit der Intelligenz ist umstritten (vgl. Abschnitt 2.2.4).

Viertens: Viele der untersuchten Artikel kommen ohne eine Festlegung auf die eine oder die andere These aus. Bei denjenigen, die es doch tun, liegt der Schluss auf „externe“ Motive nahe, etwa auf das Bedürfnis nach Neuigkeit und Aufmerksamkeit bei den Journalisten oder auf forschungspolitische Motive bei den interviewten Wissenschaftlern (vgl. die Abschnitte 2.0 und 2.2.4).

Fünftens: Wenn dies alles stimmt, fügen diejenigen Journalisten und Wissenschaftler, die die stabile Intelligenzauffassung propagieren, den Medienkonsumenten durch eine bewusste Übertreibung oder gar Verzerrung der Befundlage Schaden zu. Vermittelt über diese ethische Komponente kommt der Frage nach den Auswirkungen der Berichte über Hirnentwicklungsforschung eine besondere praktische Relevanz zu. So ist es nicht zuletzt die mit den Auswirkungen des Hirnforschungsdiskurses verbundene ethisch-moralische Fragwürdigkeit, die dazu veranlasst, diese Auswirkungen empirisch zu untersuchen. Dies soll im empirischen Teil dieser Arbeit geschehen.

3. Implizite Intelligenztheorien, Emotion und Leistungshandeln

In einer kürzlich durchgeführten Online-Befragung der Zeitschrift Geo (4.7.2003, vgl. www.geo.de) gaben 1200 Teilnehmer darüber Auskunft, hinsichtlich welcher Eigenschaften anderer Menschen sie Neid empfinden. Nach Reichtum (31%) wurde Intelligenz (23%) am zweithäufigsten genannt, bemerkenswerterweise noch vor Karriere (22%), Schönheit (20%) und Gesundheit (5%). Geistigen Fähigkeiten, so scheint es, wird im Alltag viel Bedeutung für ein erfolgreiches Leben zugeschrieben. Die Bildungsdiskussionen im Zuge der PISA-Studien sowie der in Abschnitt 2.2 bereits angesprochene Anstieg elterlicher Bemühungen zur frühkindlichen Förderung (vgl. auch Abschnitt 9.3.1) sind nur zwei Beispiele für Entwicklungen, die nahelegen, dass der Stellenwert der Intelligenz in jüngster Zeit eher noch zunimmt. Im Zuge der Globalisierung gewinnt – so ist allerorten zu hören – die Bildung eines Menschen erheblich an Bedeutung für seinen beruflichen Erfolg, und als Voraussetzung für Bildung werden zumeist intellektuelle Fähigkeiten gesehen, die es frühzeitig zu fördern gilt.

Diese Überlegungen lassen vermuten, dass Intelligenz ein wichtiges alltagspsychologisches Konzept ist und dass ihre Bedeutung eher zu- als abnimmt. Da erscheint es plausibel, dass die Art und Weise, wie Menschen Intelligenz auffassen, welche Eigenschaften sie ihr zusprechen und welchen Einfluss auf ihr Leben sie ihr beimessen, Auswirkungen hat auf ihr Handeln, insbesondere in Leistungssituationen. Dass sich diese Vermutung in empirischen Studien tatsächlich bestätigt hat, wird in Abschnitt 3.1 deutlich werden, in dem Dwecks Leistungsmotivationsmodell vorgestellt wird. Dessen zentrales Konzept, die Implizite Intelligenztheorie, wird in Abschnitt 3.2 näher untersucht, indem anhand empirischer Studien diskutiert wird, inwieweit ihre Beziehung zum Leistungshandeln kausal interpretiert werden kann. Im Abschnitt 3.3 geht es schließlich um die Beziehung zwischen Impliziten Intelligenztheorien und Emotionen. Da dieses Thema bislang wenig Aufmerksamkeit erfahren hat, behandelt dieser Abschnitt zunächst wichtige Grundlagen aus der pädagogischen Emotionspsychologie, die erste Hypothesen über jene Beziehung zulassen. Diese Hypothesen werden dann mit den wenigen vorliegenden empirischen Befunden verglichen.

3.1 Die Leistungsmotivationstheorie von Dweck

Die Arbeitsgruppe um Carol S. Dweck (Dweck, 1999; Dweck & Leggett, 1988; Dweck & Molden, 2005; Molden & Dweck, 2006) beschäftigt sich seit den 1980er Jahren mit Impliziten Theorien über Intelligenz und ihren Auswirkungen auf Motivation und Leistung. Sie unterscheidet zwischen einer Veränderbarkeitstheorie („incremental theory“) und einer Unveränderbarkeitstheorie („entity the-

ory“). Das Verhalten in Leistungssituationen soll also (unter anderem) davon abhängen, ob die Person Intelligenz als stabil oder als veränderbar ansieht.

Die Implizite Intelligenztheorie einer Person wird als Disposition aufgefasst, die abgesehen von situationsbedingten Schwankungen (vgl. Abschnitt 3.2) recht zeitstabil ist. In Längsschnittstudien, die die Impliziten Theorien Heranwachsender über mehrere Jahre untersuchten, konnte dies auch empirisch bestätigt werden (Robins & Pals, 2002). Für diese Stabilität scheinen bestimmte Erziehungspraktiken in der Kindheit verantwortlich zu sein (vgl. Mueller & Dweck, 1998).

Die Arbeitsgruppe um Dweck hat verschiedene Verfahren zur Messung von Impliziten Intelligenztheorien vorgelegt. Das bekannteste, die „Implicit Theories Scale“ (Dweck, Chiu, & Hong, 1995a; deutsch: Spinath, 2001), umfasst drei Items, die alle nach der Zustimmung zur Unveränderbarkeitstheorie fragen (z.B. „You have a certain amount of intelligence and you really can’t do much to change it.“). Auf Veränderbarkeits-Items wurde verzichtet, da diese sich als zu anfällig für Effekte sozialer Erwünschtheit erwiesen hatten. Das Verfahren trennt Unveränderbarkeitstheoretiker (die von nun an kurz *Statiker* heißen sollen) und Veränderbarkeitstheoretiker (im Folgenden: *Dynamiker*) also anhand der mehr oder weniger ausgeprägten Zustimmung zu den Unveränderbarkeits-Items. In jüngster Zeit wurden weitere, differenziertere Testverfahren entwickelt (vgl. Dweck, 1999), die sich aber inhaltlich und hinsichtlich ihrer Messeigenschaften nicht wesentlich von dem erwähnten Verfahren unterscheiden.

Der überwiegende Teil der Studien über Implizite Intelligenztheorien untersucht deren korrelative Beziehung zu Merkmalen der Motivation und des Verhaltens in Lern- und Leistungssituationen. Für folgende drei Merkmale konnte eine solche Beziehung bereits sowohl in kontrollierten Laborstudien (die die Probanden in eine vorstrukturierte Leistungssituation versetzen) als auch in längsschnittlichen Felduntersuchungen (die die Probanden in ihren natürlichen Leistungsumwelten begleiten) nachgewiesen werden:

- *Zielorientierung*: Statiker neigen eher zu einer Leistungszielorientierung, d.h. sind bemüht, sich vor anderen als kompetent darzustellen, während Dynamiker mehr zu einer Lernzielorientierung tendieren, bei der die Verbesserung der eigenen Kompetenz im Vordergrund steht (z.B. Cury, Elliot, Da Fonseca, & Moller, 2006, Studie 1; Dweck & Leggett, 1988; Robins & Pals, 2002; vgl. auch Abschnitt 3.3.3).
- *Attribution nach Misserfolg*: Statiker führen ein negatives Leistungsergebnis eher auf mangelnde Fähigkeit und Begabung zurück, Dynamiker hingegen auf zu geringe Anstrengung (z.B. Blackwell, Dweck, & Trzesniewski, 2007, Studie 1; Robins & Pals, 2002). In der Laboruntersuchung von Hong et al. (1999, Studie 1) ließ sich ein konsistenter Effekt allerdings nur für Anstrengungsattribution nachweisen. Das Ausbleiben des Effekts bei Fähigkeitsattri-

butionen wird – auch unter Rückgriff auf andere empirische Befunde – darauf zurückgeführt, dass Dynamiker Fähigkeit anders, eben „dynamischer“ verstehen als Statiker, die mit Fähigkeit eher die angeborene Begabung assoziieren. Durch diese unterschiedliche Auffassung des Fähigkeitskonzepts seien beide Gruppen in gleicher Weise dazu bereit, die Ursache des Scheiterns in der eigenen Fähigkeit zu lokalisieren.

- *Bewältigungsverhalten nach Misserfolg*: Statiker reagieren auf ein negatives Leistungsergebnis eher mit aufgebenden Verhaltensweisen als Dynamiker. Sie versäumen es, spezielle Fördermöglichkeiten in Anspruch zu nehmen, um ihren Leistungsrückstand aufzuholen (Hong et al., 1999, Studie 2). Sie geben in Fragebögen an, aus dem Misserfolg den Schluss zu ziehen, von nun an *weniger* für dieses Schul- oder Studienfach zu lernen und es möglichst schnell abzuwählen. Dynamiker hingegen ziehen die entgegengesetzte Schlussfolgerung, dass der Misserfolg sie zu *vermehrter* Anstrengung ansporne (was sich nachweislich auch in konkretem Bewältigungsverhalten niederschlägt). Diese Unterschiede werden begrifflich zumeist so zusammengefasst, dass Dynamiker mit *meisterndem*, Statiker hingegen mit *hilflosem* Bewältigungsverhalten in Verbindung gebracht werden (Blackwell, Dweck, & Trzesniewski, 2007; Dweck & Leggett, 1988; Hong et al., 1999).

In jüngsten Längsschnittstudien (Blackwell, Dweck, & Trzesniewski, 2007, Studie 1; Cury, Elliot, Da Fonseca, & Moller, 2006, Studie 1) konnte sogar ein Effekt auf die *objektive Schul- und Studienleistung* nachgewiesen werden. Blackwell et al. (2007) zufolge meistern Dynamiker den Übergang in die amerikanische „Junior High School“ mit seinen sprunghaft ansteigenden Anforderungen wesentlich besser als Statiker. Während bei ersteren die Schulnoten immerhin konstant bleiben, fallen sie bei letzteren deutlich ab. Eine querschnittliche korrelative Beziehung zwischen Impliziter Theorie und objektiven Leistungsmaßen wie Intelligenz, Schulnoten etc. konnte aber bislang nicht festgestellt werden (Dweck, 1999), was ja gemäß den Längsschnittstudien zu erwarten gewesen wäre. Ihnen zufolge müssten ja Dynamiker mit der Zeit einen immer größeren Lernvorsprung vor Statikern erlangen. Da dies nicht der Fall ist, erscheint die Befundlage widersprüchlich. Grund hierfür könnte sein, dass Dynamiker in kritischen, misserfolgsreichen Lernphasen bessere Leistungen zeigen, Statiker dagegen in leichten, wenig anspruchsvollen Phasen, welche für Dynamiker nicht so motivierend sind. Diese beiden Effekte könnten sich über die Zeit hinweg ausgleichen.

Die oben genannten drei Zusammenhänge, die den Kern von Dwecks Modell ausmachen, sind nicht unplausibel. Wer Intelligenz für unveränderbar hält, dem ist daran gelegen, dass andere einen möglichst positiven Eindruck der eigenen „Menge“ an Intelligenz erhalten. Wer Intelligenz für unveränderbar hält, wird

vor allem bei anspruchsvollen Aufgaben dazu neigen, Misserfolge auf eben diese Intelligenz zurückzuführen, und künftigen Aufgaben resignierend gegenüberzutreten. Wer hingegen Intelligenz für veränderbar hält, sieht sich eher zu deren Steigerung motiviert, attribuiert Misserfolge auf zu geringe Lernbemühungen und wird folglich diese Bemühungen in der Zukunft intensivieren. Die Zusammenhänge werden überdies durch weitere Befunde gestützt, wonach Statiker stärker zu selbstbehinderndem Verhalten („self-handicapping“) und weniger stark zu selbstregulativen Lernstrategien neigen als Dynamiker (vgl. Dweck & Molden, 2005). Das Intelligenzkonzept scheint psychologisch so bedeutsam zu sein, dass es das Lern- und Leistungsverhalten gleich über mehrere Wege beeinflusst. Dweck spricht in neueren Publikationen daher von einem „meaning system“, das von der jeweiligen Impliziten Theorie gleichsam aufgespannt wird. Statiker und Dynamiker betrachten Leistungssituationen demnach aus unterschiedlichen Blickwinkeln, ja mit unterschiedlichen Brillen und interpretieren sie anhand grundverschiedener Begrifflichkeiten und Maßstäbe.

What failures and setbacks *mean* to the person experiencing them and what specific threats they are seen as posing to this person's self-concept should have an important influence on self-regulation in such circumstances. Is failure interpreted in a way that is undermining and debilitating or one that is challenging and energizing? (Molden & Dweck, 2006, S. 194)

People's lay theories appeared to serve as core assumptions that created a larger system of allied beliefs and goals. For example, in the (...) studies, people's entity and incremental theories influenced their academic performance and self-esteem primarily through a network of goals, beliefs, and strategies that arose from these basic theories. (...) Overall, then, people's lay theories may alter their psychological processes through these larger *meaning systems* that come to be organized around these theories (Molden & Dweck, 2006, S. 201).

By illuminating the valuing of different competence goals, self-theories can also give us entrée into the “meaning systems” people use to construct meaning in competence-relevant situations. Often, motivational variables are considered in isolation. Rarely do researchers look at a network of beliefs and goals that work together to produce important behaviors and outcomes; that is, rarely do they look at the meaning systems that give rise to the behaviors and outcomes we care about (Dweck & Molden, 2005, S. 122).

Dieses Konzept der „meaning systems“ stellt einen interessanten Brückenschlag zu neueren kulturpsychologischen und sozialkonstruktionistischen Ansätzen dar (vgl. Grant & Dweck, 2001; Heine et al., 2001; Molden & Dweck, 2006), den Sozialpsychologen schon seit längerem anmahnen (Jost & Kruglanski, 2002). Diskurse über Lernen und Leistung werden in den Blick genommen, ohne dabei gleich in dekontextualisierte Einzelvariablen zerlegt zu werden. Gleichzeitig wird auf den postmodernen Überbau verzichtet, der häufig mit diskursanalytischen Ansätzen verbunden ist. Im Konzept der „meaning systems“ spiegelt sich mithin eine Eigenschaft des Dweckschen Ansatzes wider, der ihm ebenso viel Kritik wie Aufmerksamkeit eingebracht hat: Er überwindet das starre allgemein-

psychologische Variablendenken, nimmt die *Inhalte* kognitiver Strukturen ernst, ohne sich der Beliebigkeit qualitativ-konstruktivistischer Forschungszugänge zu überlassen (von Tiling, 2008b; vgl. genauer: Abschnitt 9.4).

Wie in den Zitaten angedeutet, sieht Dweck die Implizite Theorie als Ursprung des jeweiligen Bedeutungssystems. Diese kausale Deutung korrelativer Daten begründet Dweck weitgehend mit den pfadanalytischen Auswertungsverfahren, die ein entsprechendes Modell mit Impliziten Theorien als Ausgangsgröße bestätigen. Wenngleich man Pfadanalysen durchaus als *Hinweis* auf kausale Zusammenhänge betrachten kann, rechtfertigen sie nicht solche weitgehenden und eindeutigen Schlussfolgerungen wie die oben zitierten (vgl. Backhaus, Erichson, Plinke, & Weiber, 2000; Rietz, Rudinger, & Andres, 1996). Nur durch experimentelle Studien, in denen Implizite Theorien gezielt induziert werden und im Anschluss daran konkretes Leistungsverhalten gemessen wird, kann festgestellt werden, ob Implizite Theorien tatsächlich diesen *ursächlichen* Einfluss besitzen oder einfach mehr oder weniger originelle Korrelate anderer, gut erforschter Leistungsmotivationskonstrukte sind. Einige solcher Studien liegen zwar mittlerweile vor; im Vergleich zu den Korrelationsstudien sind sie aber weiterhin selten. In ihnen wird versucht, Implizite Intelligenztheorien durch die Konfrontation mit bestimmten Leistungsdiskursen gezielt zu induzieren.

Der folgende Abschnitt widmet sich diesen experimentellen und quasi-experimentellen Studien in ausführlicher Weise und versucht im Zuge dessen die Frage zu beantworten, ob sie die weitreichenden kausalen Postulate Dwecks hinreichend rechtfertigen. Dabei wird übrigens auch deutlich werden, dass die beschriebene Anknüpfung an kulturpsychologische Ansätze (vgl. z.B. Heine, 2008; Ratner, 2006) erst durch diese Form der Operationalisierung Impliziter Theorien in überzeugender Weise gelingt (vgl. auch Abschnitt 9.4).

3.2 Untersuchungen zum kausalen Einfluss Impliziter Intelligenztheorien

Zunächst erscheint es wenig plausibel, dass solch grundlegende Auffassungen wie die über das Wesen der Intelligenz kurzfristig verändert werden können – schon gar nicht mit den begrenzten Möglichkeiten eines Laborexperiments. Es wurde oben bereits auf empirische Untersuchungen hingewiesen, die gezeigt haben, dass Implizite Theorien sich primär über Erziehung und Sozialisation entwickeln und fortan wenig veränderbar sind (vgl. Mueller & Dweck, 1998; Robins & Pals, 2002). Anders herum ist die Vorstellung nicht ganz abwegig, dass äußere Einflüsse die Intelligenztheorie einer Person leicht modifizieren können: Sowohl die statische als auch die dynamische Auffassung dürfte jedem Menschen vertraut sein und plausibel erscheinen; die Auffassung einer *vollständigen* Unveränderbarkeit oder auch einer *beliebigen* Veränderbarkeit dürfte empirisch kaum vorkommen (Dweck, 1999). Die Frage, die sich Menschen implizit stellen,

muss also immer lauten: *In welchem Maße* bin ich meiner Intelligenz ausgeliefert, bzw. *in welchem Maße* kann ich sie verändern? Weil nun also die Intelligenzfrage keine einfachen, leicht zu verbalisierenden Antworten zulässt, kann davon ausgegangen werden, dass die Implizite Intelligenztheorie einer Person immer mit einer erheblichen Unsicherheit verbunden und insofern eher Hypothese als Theorie ist. Wird eine Person mit Informationen aus der sozialen Umwelt konfrontiert, die entweder die statische oder die dynamische Position eindeutig zu bestätigen scheinen, erscheint es naheliegend, dass die Person jene Position zwar nicht unbedingt direkt übernimmt, aber immerhin ihre eigene Vorstellung ein wenig in deren Richtung korrigiert. So wird im Übrigen erst in dieser Betrachtungsweise vollends deutlich, worin genau das *Implizite* an Impliziten Intelligenztheorien besteht.

Die Implizite Theorie eines Menschen wird unter diesem Blickwinkel nicht mehr als ein statisches Personmerkmal konzeptualisiert, sondern als aktueller Zustand, in dem sich eine Person befindet. Dieser Zustand resultiert nicht nur aus dispositionalen, sondern auch aus situativen Einflüssen, welche für eine mehr oder weniger kurzfristige Veränderung der dispositionsgemäßen Ausprägung sorgen können (Dweck, Chiu, & Hong, 1995b). Ein so verstandenes Konstrukt der Impliziten Theorie entspricht dem der Motivation oder Emotion, die sich ebenso aus personspezifischen („Motiv“, „emotionale Disposition“) und situationalen Anteilen zusammensetzen. Es dürfte sich von diesen jedoch in der zeitlichen Variabilität unterscheiden, so dass ein „aktueller Zustand Impliziter Intelligenztheorien“ wohl am besten den Eigenschaften einer *Stimmung* entspricht. Er ändert sich zwar, aber nicht so abrupt und situationsbedingt wie etwa eine Emotion. In den Abschnitten 5.3 und 9.4 wird kritisch zu hinterfragen sein, ob diese geistes- und kulturwissenschaftliche Sichtweise auf Implizite Theorien, die bei Dweck bisweilen durchscheint und auf die in dieser Arbeit stärker fokussiert wird, mit der eher positivistischen Erhebungsweise der Disposition Impliziter Theorien vereinbar ist.

Die Vorstellung der Modifizierbarkeit der Intelligenzauffassung einer Person konnte in mehreren Studien bestätigt werden. Dabei zeigten sich größtenteils dieselben Änderungen im Leistungsverhalten wie in den korrelativen Studien. Im Folgenden werden diese Untersuchungen im Detail beschrieben. Unterschieden wird hierbei zwischen laborexperimentellen Studien einerseits und feldexperimentellen bzw. Interventionsstudien andererseits. Erstere untersuchen die kurzfristigen Auswirkungen der Theorie-Induktion im Labor, während letztere ein umfangreiches Training anbieten, in dessen Anschluss die Entwicklung der Probanden im schulischen bzw. studentischen Leistungsverhalten über einen längeren Zeitraum verfolgt wird. Der Vorteil laborexperimenteller Studien besteht u.a. darin, dass der Zweck der Induktion den Probanden vorenthalten werden kann, während die Interventionsstudien besonders geeignet dafür sind, die Induktions-

effekte differenziert und realitätsnah zu untersuchen. Ein wichtiger Nachteil der Interventionsstudien besteht darin, dass aus ethischen Gründen keine Induktion der statischen Intelligenztheorie vorgenommen werden kann (vgl. Abschnitt 2.2.2).

Angemerkt sei noch, dass die Untersuchung von Aronson et al. (2002) quer zu der vorgenommenen Unterscheidung liegt: Sie hat einerseits laborexperimentellen Charakter, indem sie u.a. den Untersuchungszweck vor den Probanden verbergen kann, umfasst andererseits aber eine längerfristige, sich auf die dynamische Bedingung beschränkende „Behandlung“ und untersucht deren Auswirkungen ebenfalls langfristig. Sie wird in der folgenden Darstellung den Interventionsstudien zugeschlagen.

3.2.1 Laborexperimentelle Studien

Laborexperimentelle Untersuchungen nehmen die Induktion Impliziter Theorien zumeist durch die Konfrontation mit fiktiven wissenschaftsjournalistischen Zeitschriftenartikeln vor. Die Probanden erhalten einen Artikel, der entweder die statische oder die dynamische Botschaft enthält und diese anhand (angeblicher) neuer wissenschaftlicher Untersuchungen glaubhaft zu machen versucht. Bergen (1991) leistete hierbei Pionierarbeit, indem er Artikel entwickelte, die, aufgemacht im Stil und Layout der populärpsychologischen US-Zeitschrift „Psychology Today“, eine der beiden Positionen in extremer Zuspitzung vertreten. Die zweiseitigen Artikel berichten von aufwändigen Zwillingsstudien, deren Ergebnisse die alte Kontroverse zwischen Anlage und Umwelt hinsichtlich Intelligenz endgültig in eine Richtung gelöst hätten.

Bergen entwickelte insgesamt vier solcher Artikel. Er variierte nicht nur die Implizite Intelligenztheorie (statisch vs. dynamisch), sondern auch die Bedeutsamkeit, die Intelligenz für Erfolg in sowohl beruflichen als auch privaten Lebensbereichen zugeschrieben wird (bedeutend vs. unbedeutend). Mit diesen führte er eine kleine, unpublizierte Untersuchung durch (Bergen, 1991, Studie 1), die hier aufgrund einiger methodischer Schwächen und der daraus resultierenden begrenzten inhaltlichen Aussagekraft nicht näher erläutert werden soll. Seine Artikel wurden jedoch von einer Reihe anderer Forscher verwendet (s.u.). Dort ist allerdings jeweils nur von zwei Bergen-Artikeln (einem statischen und einem dynamischen) die Rede. Welche beiden in diesen Studien Verwendung finden, bleibt unklar und konnte auch nicht von den betreffenden Forschern in Erfahrung gebracht werden (vgl. auch die Abschnitte 5.1 und 8.2).

Die erwähnten Untersuchungen, die Bergens Texte als Induktionsmaterial einsetzen, weisen alle eine ähnliche Struktur auf. Die Probanden wurden, nachdem sie den Artikel gelesen hatten, zumeist mit einer Leistungsaufgabe, mitunter auch lediglich mit Persönlichkeitsfragebögen konfrontiert. Immer jedoch wurde

versucht, den Zusammenhang zwischen dem Lesen des Artikels und den nachfolgenden Vorgängen zu verschleiern. Infolgedessen ist davon auszugehen, dass die (potenziellen) Auswirkungen des Lesens der Artikel den Probanden nicht bewusst wurden, so dass die Induktion als „Priming“ aufgefasst werden kann. Wie diese Verschleierung in den einzelnen Experimenten jeweils konkret vorgenommen wurde, wird in der folgenden Darstellung dieser Untersuchungen deutlich werden.

Hong et al. (1999, Studie 3) gaben ihren studentischen Versuchspersonen entweder den statischen oder den dynamischen Artikel und forderten sie auf, eine Zusammenfassung zu schreiben. Nach diesem Priming, das als Leseverständnistest deklariert wurde, erhielten die Probanden einige komplexe Aufgaben eines Intelligenztests zur Bearbeitung, bezüglich derer ihnen anschließend entweder eine negative oder eine gemäßigt-positive Leistungsrückmeldung gegeben wurde. In der anschließenden (angeblichen) Versuchspause wurden den Teilnehmern zwei Aufgaben zu deren Überbrückung angeboten: ein Training zur Verbesserung der Leistung in Intelligenztests sowie eine kognitive Leistungsaufgabe aus einem anderen Fähigkeitsbereich. Innerhalb derer, denen Misserfolg rückgemeldet worden war, wählten lediglich zwei (von 15) Statiker, jedoch elf (von 15) Dynamiker das Intelligenztesttraining. (Der Einfachheit halber werde ich auch hier von Statikern und Dynamikern sprechen, obwohl es dabei ja nicht um eine Disposition geht.) Ähnlich deutlich waren die Unterschiede in der Anstrengungsattribution, welche signifikant stärker bei den Dynamikern auftrat. Innerhalb der Erfolgsbedingung gab es jeweils keine Unterschiede zwischen den Gruppen. Für Fähigkeitsattributionen zeigten sich wiederum keine Effekte (vgl. Abschnitt 3.1 für eine Deutung).

Hong et al. konnten also erstmals zeigen, dass Intelligenztheorien das Bewältigungsverhalten nach Misserfolg *kausal* (mit-)beeinflussen. In manchen Aspekten erweckt ihre Untersuchung allerdings den Eindruck der Trivialität. Dem Proband wird zunächst ein Artikel über die Stabilität von Intelligenz gegeben; dann wird er gefragt, ob er seine Leistung in einem *Intelligenztest* mit Hilfe eines Trainings verbessern möchte; im Ergebnis entscheidet er sich seltener für die Teilnahme als sein Kollege, der von der modifizierbaren Intelligenz gelesen hat. Die Tatsache, dass es in der Bewältigungsaufgabe wieder explizit um Intelligenz geht, nimmt Hong et al.s Studie viel an Aussagekraft und praktischer Relevanz. Sie ist dann nur noch Beleg dafür, dass sich durch Bergens Artikel das Verhalten gegenüber Intelligenztests ändern lässt, nicht allgemein gegenüber Leistungsaufgaben, wie sie alltäglich in Schule und Studium vorkommen und bei denen immer eine gewisse *Unsicherheit* darüber besteht, inwieweit für sie eher Intelligenz oder eher Fleiß vonnöten sind. Hätte man im zweiten Versuchsteil auf den Intelligenzbegriff verzichtet und die Leistungsaufgabe anders deklariert, wäre eben diese Generalisierbarkeit größer gewesen.

Ähnliche Probleme liegen bei der Arbeit von Cury et al. (2006, Studie 2) vor, deren Darstellung hier kurz eingeschoben wird, obwohl sie sich nicht der Bergen-Texte bedient. Die 13-15-jährigen Probanden bearbeiteten einen Intelligenztest (der auch explizit als solcher etikettiert wurde), hinsichtlich dessen sie anschließend entweder eine positive oder eine negative Leistungsrückmeldung erhielten. Danach wurde ihnen in wenigen Sätzen schriftlich „mitgeteilt“, dass (die in dem Test gemessene) Intelligenz ein unveränderbares bzw. veränderbares Merkmal ist. Abschließend bearbeiteten sie weitere Intelligenztestaufgaben sowie einige Persönlichkeitsfragebogen. Es zeigte sich unter anderem, dass Statiker im zweiten Intelligenztest schlechter abschnitten sowie weniger intrinsische Motivation und Lernzielorientierung und mehr Leistungszielorientierung zeigten als Dynamiker. Diese Effekte erscheinen jedoch wiederum – und stärker noch als bei Hong et al. – ebenso plausibel wie trivial: Hier bezieht sich sogar die Induktion Impliziter Theorien direkt auf den bearbeiteten IQ-Test („Sie erhalten jetzt einen Intelligenztest, der eine stabile/veränderbare Fähigkeit misst...“). Dass die Probanden dann dieser Aufgabe mit entsprechenden Zielen entgegneten („Wenn da die Intelligenz gemessen wird, muss ich versuchen, gut abzuschneiden...“), überrascht wenig. Für die Leistungseffekte kann Ähnliches in Anschlag gebracht werden: Wer glaubt, dass er seine stabile Intelligenz unter Beweis stellen muss, ist nervöser, ist gedanklich abgelenkter und erbringt folglich schlechtere Leistungen, wie aus der Prüfungsangstforschung (Zeidner, 1998) hinlänglich bekannt ist. Cury et al.s Ergebnisse verlieren damit nicht völlig an Wert; ihre praktische Relevanz jedoch kann, wie die Autoren sogar selbst angeben, lediglich darin bestehen, „that the way in which ability is framed in an IQ assessment affects performance“ (S. 676).

Bei weitem interessanter wäre dagegen zu erfahren, ob die Induktion Impliziter Theorien auch bei alltäglichen Leistungsaufgaben Wirkung zeigt, bei Aufgaben eben, für deren Bewältigung der Intelligenz immer nur eine *mehr oder weniger* große Bedeutung zugeschrieben wird. Eine solche nicht-triviale experimentelle Überprüfung des Dweckschen Modells liegt meines Wissens noch nicht vor. Sie soll im empirischen Teil dieser Arbeit erfolgen.

Die Untersuchung von Niiya, Crocker und Bartmess (2004) kann dafür in methodischer Hinsicht durchaus als vorbildlich gelten, wenngleich sie sich inhaltlich weniger für konkretes Leistungsverhalten interessiert als für Selbstwertgefühl und Emotion. Ihre studentischen Probanden füllten zunächst Fragebögen zur Disposition Impliziter Theorien, zu momentanem Selbstwertgefühl sowie zu der Frage aus, wie eng sie ihr Selbstwertgefühl an ihre akademischen Erfolge zu koppeln gewohnt sind („contingency of self-worth“). Dann erhielten sie einige Aufgaben des „Graduate Record Examination Test“, der in den USA als Auswahlinstrument für Bewerber an Graduiertenschulen verwendet wird. Die von Bergen (1991) erstellten Artikel waren in die erste Aufgabe, einen Leseverständ-

nistest, integriert. Nach einer Erfolgs- bzw. Misserfolgsrückmeldung wurden ihr Selbstwertgefühl und ihre Emotionen mittels Fragebögen erhoben. Dabei zeigte sich, dass Statiker nach einer negativen Leistungsrückmeldung ein geringeres Selbstwertgefühl sowie mehr Depression, Angst und Feindseligkeit aufwiesen als Dynamiker. Dieser Zusammenhang konnte jedoch nur innerhalb einer Teilpopulation nachgewiesen werden, und zwar unter denjenigen Probanden, die ihr Selbstwertgefühl eher eng an akademische Erfolge zu koppeln gewohnt waren.

Wie bereits angedeutet, vermeiden Niiya et al. die methodischen Schwächen der zuvor geschilderten Untersuchungen. Die Induktion (bzw. deren Zweck) wird, anders als bei Cury et al., in angemessener Form verschleiert, und die Leistungsaufgabe kommt, anders als bei Hong et al. und Cury et al., ohne Bezug auf die Intelligenzsemantik aus. Auf die inhaltliche Bedeutung der Studie von Niiya et al. wird in Abschnitt 3.3.4 näher eingegangen, wo es um die Beziehung zwischen Impliziten Intelligenztheorien und Emotionen geht.

3.2.2 Interventionsstudien

In den zuvor geschilderten laborexperimentellen Untersuchungen geht es um die Frage, welche kurzfristigen Auswirkungen die Konfrontation mit bestimmten Intelligenzdiskursen für die Betroffenen hat. Daneben gibt es in jüngster Zeit vermehrt Versuche, die dynamische Intelligenztheorie zum Gegenstand gezielter Intervention zu machen und die – erhofft positiven – langfristigen Auswirkungen eines solchen Trainings zu überprüfen. Dweck (2007b) versucht plausibel zu machen, warum ein solches bloßes Konfrontieren mit der dynamischen Intelligenztheorie pädagogisch erfolgreich sein sollte:

In Buchhandlungen stoßen wir oft auf Bücher mit Titeln wie „Die zehn Geheimnisse erfolgreicher Menschen“, und diese Bücher enthalten in der Tat oft nützliche Hinweise. Doch meist stehen die Ratschläge unverbunden nebeneinander, da heißt es „Gehen Sie Risiken ein!“ oder „Glauben Sie an sich selbst!“. Als Leser können wir zwar die Menschen bewundern, denen dies gelungen ist, aber wir bekommen wenig Hinweise, wie diese Dinge zusammengehören oder wie wir es anstellen sollen, an uns selbst zu glauben. Ein paar Tage lang fühlen wir uns vielleicht inspiriert, doch am Ende behalten die erfolgreichen Menschen ihr Geheimnis für sich.

Wenn wir jedoch verstehen, was es mit dem statischen und dem dynamischen Selbstbild auf sich hat, erkennen wir, wie eins zum anderen führt: wie der Glaubenssatz, dass unsere Eigenschaften in Stein gemeißelt seien, bestimmte Gedanken und Handlungen auslöst, und wie wiederum der Glaubenssatz, dass wir unsere Eigenschaften weiterentwickeln können, andere Gedanken und Handlungen auslöst, die uns einen völlig anderen Weg einschlagen lassen. Das ist das, was Psychologen das Aha-Erlebnis nennen (Dweck, 2007b, S. 19).

Das aktive Lehren der dynamischen Intelligenztheorien vermittelt gemäß dieser Sichtweise einen Kompass bzw. „meaning system“, das einfach genug ist, um vom „Mann auf der Straße“ verinnerlicht zu werden, aber gleichzeitig hinrei-

chend komplex und flexibel, um in dessen wechselhaftem Alltag auch umsetzbar zu sein. Das zitierte populärwissenschaftliche Buch Dwecks (2007b) versteht sich übrigens folgerichtig selbst als einen solchen Interventionsversuch.

Die drei bislang vorliegenden Studien, die von derartigen Interventionen berichten, wurden bereits in Abschnitt 2.2.2 ausführlich dargestellt, da sie sich alle der neurowissenschaftlichen Plastizitätsmetapher als Argumentationshilfe bedienen. Ich werde daher in diesem Rahmen nur kurz ihre Ergebnisse rekapitulieren sowie anschließend auf hautfarben- und geschlechtsspezifische Befunde hinweisen, die in Abschnitt 2.2.2 nicht zur Sprache kamen.

Die Studien von Blackwell et al. (2007), Aronson et al. (2002) sowie Good et al. (2003) belegen in konsistenter Art und Weise, dass die gezielte und wiederholte Vermittlung der dynamischen Intelligenztheorie positive langfristige Auswirkungen auf Lernverhalten und Leistung hat. Die auf diese Art und Weise unterrichteten Schüler und Studenten zeigen bis zu einem Jahr nach der Intervention bessere Leistungen als eine Kontrollgruppe, die ein anderes Training erhielt. Soweit auch Anstrengungsbereitschaft und Lernmotivation erhoben wurden, schneiden die Dynamiker auch hier besser ab.

Die Wirkung der dynamischen Intervention erwies sich als besonders stark bei Probanden, denen aufgrund äußerer Merkmale allgemein eine geringere Leistungsfähigkeit zugeschrieben wird. In der Studie von Aronson et al. (2002) profitierten afroamerikanische Studenten hinsichtlich Lernfreude, Interesse und Leistung mehr von dem Training zur dynamischen Intelligenztheorie als weiße. Aronson (1998, zit.n. Dweck & Molden, 2005) hatte zuvor bereits im Labor gezeigt, dass nicht nur der dynamische Diskurs Afroamerikanern mehr hilft als Weißen, sondern dass ihnen der statische auch mehr schadet. Bei Good et al. (2003) waren es die weiblichen Teilnehmer, die besonders profitierten, so dass nach der Intervention der anfangs bestehende Leistungsvorsprung der Männer in Mathematik verschwunden war (vgl. auch Broome, 2001).

Dieses Phänomen wird in den zitierten Studien mit Hinweis auf Befunde zum sogenannten „Stereotype Threat“ (Steele & Aronson, 1995) erklärt, d.h. dem Phänomen, dass die Konfrontation mit negativen Stereotypen über die Eigengruppe zu verringerter Motivation und Leistung führt. Demnach führt die Verbreitetheit von Vorurteilen über die intellektuelle Leistungsfähigkeit einer bestimmten Gruppe zu deren tendenzieller Bestätigung, indem die Mitglieder der Gruppe einen besonderen Druck verspüren, dem Vorurteil entgegenzuwirken. Da erscheint es einsichtig, dass dieser Druck besonders stark ist, wenn die Vorurteile ein stabiles Merkmal betreffen, weil dann ein einmaliger Misserfolg das Vorurteil ein für allemal bestätigen würde. Dweck und Molden (2005) schließen daher aus den zitierten Untersuchungen, dass eine dynamische Intelligenztheorie vor den negativen Auswirkungen von Vorurteilen schützt, eine statische hingegen für sie anfällig macht.

Eher unklar bleibt jedoch, wie dieser Effekt genau zustandekommt. Die Untersuchung von Dar-Nimrod und Heine (2006), die bereits in Abschnitt 2.2.2 vorgestellt wurde, könnte hier einen Hinweis geben. Sie gaben ihren Probanden Zeitschriftenartikel über die Gründe für das (angeblich) schlechtere Abschneiden von Frauen in mathematischen Aufgaben. Diese Gründe wurden entweder in der genetischen Ausstattung oder in der Sozialisation lokalisiert. Die Versuchsgruppe mit genetischem Erklärungsansatz schnitt in den anschließenden mathematischen Aufgaben schlechter ab als die andere, sozialisationsgeprägte Gruppe.

Dar-Nimrod und Heines Studie lehrt also: Frauen mit ihrer eigenen (angeblichen) angeborenen Inkompetenz zu konfrontieren, „wirkt“. Weshalb aber wirkt ein Text über stabile Intelligenz, wie er etwa von Aronson (1998) verwendet wurde, in ähnlicher Form? Ein Hypothese wäre hier (vgl. Abbildung 1), dass für Frauen dieser Text deswegen „besser“ wirkt als für Männer, weil sie in ihm *impliziert* sehen (vgl. die gestrichelten Linien in Abbildung 1), dass die Leistungsunterschiede im intellektuellen Bereich zwischen Männern und Frauen (die es ja immerhin im Alltagsdiskurs gibt bzw. die aufgrund der ungleichen gesellschaftlichen Machtverteilung zwischen den Geschlechtern scheinbar evident sind) durch eben diese stabile Intelligenz bedingt sind. Wer Intelligenz als stabil annimmt, der glaubt auch, dass „real existierende“ Leistungsunterschiede zwischen Männern und Frauen (oder Weißen und Schwarzen) auf diese stabile Intelligenz zurückzuführen sind. Statikerinnen leiden demnach nicht nur unter dem Kontrollverlust, der von der statischen Intelligenztheorie nahegelegt wird und in gleicher Weise die Männer betrifft, sondern zusätzlich auch unter jener Abwertung weiblicher Kompetenz. Für männliche Statiker hingegen entstehen ambivalente Auswirkungen (Kontrollverlust, aber Aufwertung männlicher Kompetenz). Umgekehrt sind die Auswirkungen der dynamischen Theorie für Frauen doppelt positiv (Kontrollgewinn und Dekonstruktion weiblicher Unterlegenheit), für Männer hingegen wieder ambivalent (Kontrollgewinn, aber Dekonstruktion männlicher Überlegenheit).

Diese Erklärung der Geschlechtsunterschiede passt sich den Datenmustern der berichteten Studien recht gut an. Sie kann problemlos auf die Befunde zu ethnischen Unterschieden übertragen werden. Dennoch bedarf sie zweifellos eingehenderer empirischer Prüfung. Im empirischen Teil dieser Arbeit wird dies hinsichtlich des Merkmals Geschlecht in Angriff genommen, wenngleich diese Fragestellung im Rahmen meiner Untersuchungen eine randständige Stellung mit eher explorativem Charakter einnimmt.

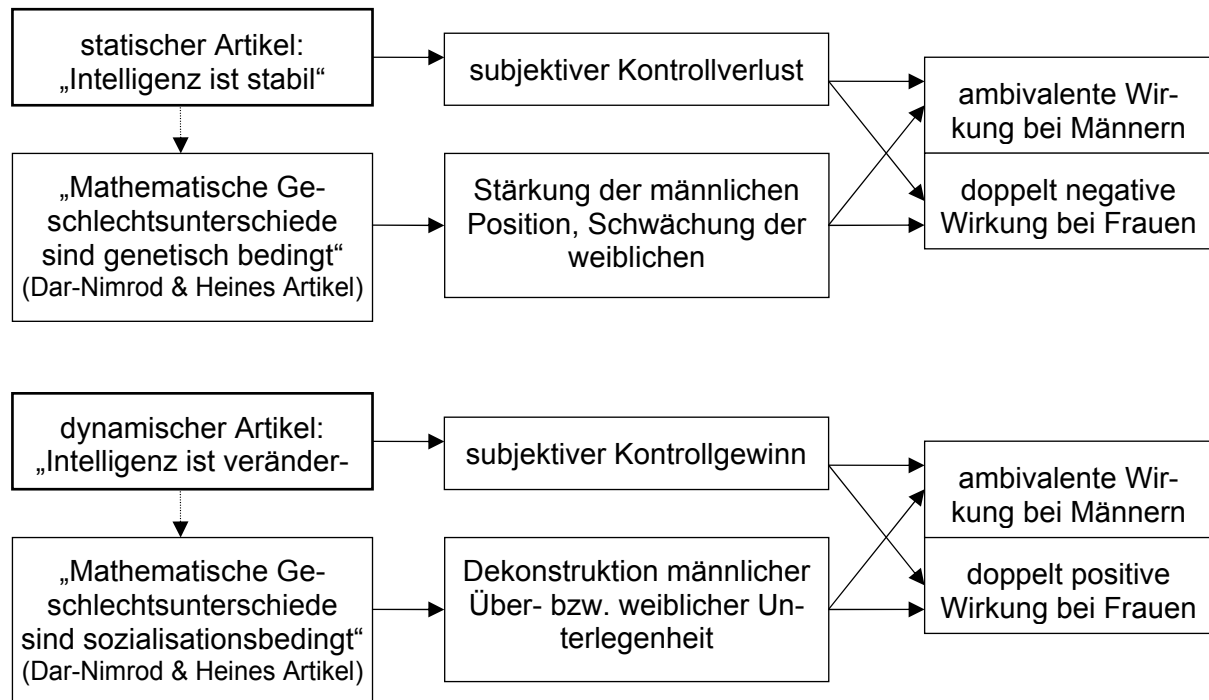


Abbildung 1. Schematische Darstellung der Erklärung der Geschlechtsunterschiede in den Interventionsstudien, unter Rückgriff auf die Befunde von Dar-Nimrod und Heine, 2006.

(Die gestrichelten Pfeile stellen die theoretisch abgeleiteten Hypothesen dar, während die durchgezogenen Pfeile empirisch gefundene Zusammenhänge versinnbildlichen – wenngleich diese, abbildungsbedingt, bereits in theoretische Begriffe überführt sind. Weitere Erläuterungen im Text.)

3.2.3 Fazit

Die kausale Bedeutung Impliziter Theorien für das Lern- und Leistungsverhalten ist bislang eher selten und zumeist auch in inhaltlich eingeschränkter Weise untersucht worden. Die erörterten labor- und feldexperimentellen Befunde liefern aber immerhin einige Hinweise auf den angenommenen ursächlichen Einfluss. Während die Aussagekraft der Laborexperimente durch die diskutierten methodischen Schwächen eher gering ist, sind die Ergebnisse der Interventionsstudien sehr viel besser und in sich kohärenter in diese Richtung interpretierbar. Dies hat allerdings zur Folge, dass nun zwar einiges über die Wirksamkeit der dynamischen Intelligenztheorie bekannt ist, aber kaum etwas über die der statischen. Im empirischen Teil dieser Arbeit werde ich daher einige methodisch sorgfältiger geplante Laborstudien vorstellen, die die Wirksamkeit beider Theorien überprüfen.

Laborstudien sind jedoch auch aus einem anderen Grund angezeigt: Sie eignen sich weit besser als Feldstudien für die Untersuchung der Frage, welche *vermittelnden Prozesse* für jene durchaus erstaunlichen Effekte der experimentellen Induktion Impliziter Theorien verantwortlich sind. Diese Frage ist allgemein bislang kaum untersucht worden. Während Blackwell et al. (2007) zwar

durchaus diverse motivationale Konstrukte wie die Wertschätzung von Lernen und Anstrengung miterheben und als Mediatoren in Erwägung ziehen, geben sie doch zu Bedenken, dass „our findings only begin to answer the question why endorsing an incremental theory is related to better grades“ (S. 259). Ebenso unklar ist auch, welche Vorgänge im Kopf des Teilnehmers der Studie von Hong et al. für die Entscheidung für bzw. gegen die Teilnahme an der Trainingsmaßnahme verantwortlich sind. Vor allem die Untersuchung *emotionaler* Prozesse ist hier, wie ich im Folgenden darlegen will, bislang vernachlässigt worden.

3.3 Leistungsemotionen und Implizite Intelligenztheorien

Diese Vernachlässigung emotionaler Komponenten in Dwecks Modell ist durchaus nicht unbemerkt geblieben. So mahnen Dweck und Molden (2005) am Ende ihres Literaturüberblicks selbst an, die Bedeutung emotionaler Prozesse besser zu untersuchen, und versprechen sich davon eine Stärkung des „much needed link between the study of emotion and the study of motivation“ (S. 137). Bevor ich auf erste Bemühungen in diese Richtung eingehen werde (vgl. Abschnitt 3.3.4), will ich bestehende theoretische Unterscheidungen und empirische Untersuchungen zu Emotionen in Lern- und Leistungssituationen erörtern.

Hier ist zunächst mit Pekrun (1998; vgl. auch Eccles & Wigfield, 2002) zu konstatieren, dass Emotionen nicht nur bei Dweck, sondern in der *gesamten* Pädagogischen Psychologie lange Zeit vernachlässigt worden sind. Während kognitive Prozesse (wie z.B. Kausalattributionen) und motivationale Prozesse (wie z.B. Zielorientierungen) in ihrer Bedeutung für Lernen und Leistung recht gut erforscht sind, steht die Beschäftigung mit Emotionen noch am Anfang. Eine Ausnahme bildet die Prüfungsangst, zu der mittlerweile Arbeiten in kaum überschaubarer Anzahl vorliegen (vgl. Pekrun, Goetz, Titz, & Perry, 2002). Pekrun (1998) zeigt anhand einer Klassifikation von Lern- und Leistungsemotionen, welche unterschiedlichen Arten von Emotionen im Zuge dieser einseitigen Fokussierung auf die Prüfungsangst vernachlässigt wurden.

Danach werden zunächst *aufgabenbezogene* von *sozialen* Lern- und Leistungsemotionen unterschieden. Erstere haben die Lerntätigkeit oder das Lernergebnis selbst zum Gegenstand, während letztere sich auf die in der Leistungssituation beteiligten Menschen beziehen (z.B. Ärger über den Lehrer, Neid auf die Mitschüler). Aufgabenbezogene Emotionen lassen sich anhand ihres Gegenstandes weiter unterteilen. Erstens können sie sich auf die Lerntätigkeit beziehen (*tätigkeitsbezogene* Emotionen, z.B. Lernfreude, Langeweile), zweitens auf eine bevorstehende Leistungssituation (*prospektive* Emotionen, z.B. Zuversicht, Angst, Hoffnungslosigkeit), drittens auf eine zurückliegende Leistungssituation (*retrospektive* Emotionen, z.B. Ergebnisfreude, Stolz, Traurigkeit, Scham, Schuld).

Diese Unterscheidung ist insofern nur eine analytische, als sie nichts darüber aussagt, *wann* bestimmte Emotionen vorkommen (können), sondern nur *unter welcher Gegenstandsfokussierung*. Nach der Rückgabe einer Klausur kann eine Schülerin z.B. nicht nur retrospektive Emotionen empfinden, die sich dann also im Regelfall auf das zurückliegende Leistungsergebnis beziehen, sondern auch prospektive, die sich etwa auf künftige Klausuren richten könnten. Empirisch dürfte hier eine klare Unterscheidung zwischen z.B. retrospektiver Traurigkeit und prospektiver Hoffnungslosigkeit kaum möglich sein.

Im Hinblick auf empirische Untersuchungen halte ich es daher für angebracht, Emotionen eher anhand ihres zeitlichen Vorkommens zu unterscheiden. Der größte Teil aller auftretenden emotionalen Episoden dürfte durch eine Unterscheidung abgedeckt sein, die zwischen Emotionen *kurz vor* einer Leistungssituation, *während* einer Leistungssituation und *kurz nach einer Leistungsrückmeldung* differenziert. Dabei ist der Begriff der Leistungsrückmeldung in einem weiten Sinne zu verstehen. Während zwar der prototypische Fall einer Leistungsrückmeldung im Verkünden eines Leistungsergebnisses durch Lehrer, Prüfer etc. besteht, kann die betroffene Person auch durch Mitschüler oder gar allein durch die Wahrnehmung eigener Lernprobleme und Verständnisschwierigkeiten darauf aufmerksam gemacht werden, dass sie in einer Leistungssituation gut oder schlecht abgeschnitten hat. Emotionen infolge einer Leistungsrückmeldung treten also immer dann auf, wenn die Person ihre eigene Leistung als gut oder schlecht bewertet (vgl. Näheres in Abschnitt 3.3.1).

Im Rahmen von Dwecks Leistungsmotivationsmodell sind besonders diese Emotionen nach Leistungsrückmeldungen von Interesse. Wie in Abschnitt 3.1 dargestellt, betrifft Dwecks Modell nämlich vor allem das Leistungshandeln nach Misserfolg, d.h. nach einer negativen Leistungsrückmeldung. In den folgenden Abschnitten soll es daher vor allem um diese Form der Leistungsemotionen gehen. Zunächst werden klassische Arbeiten zur Beziehung zwischen Emotion und Leistungsmotivation vorgestellt (Abschnitt 3.3.1). Dann geht es um eine spezifische Art von Emotionen nach Leistungsrückmeldungen, nämlich selbstrelevante Emotionen, und um ihre attributionstheoretische Konzeptualisierung (Abschnitt 3.3.2). Abschnitt 3.3.3 schildert dann Befunde zu Leistungsemotionen, die im Rahmen von Studien zur Zielorientierung entstanden sind. Erst in Abschnitt 3.3.4 werden die wenigen bislang vorliegenden empirischen Befunde zum Einfluss von Impliziten Intelligenztheorien auf Leistungsemotionen zur Sprache kommen. Abschließend soll es um ein weithin vernachlässigtes Gebiet der Emotions- und Sozialpsychologie gehen, der sogenannten Emotionsdarstellung (Abschnitt 3.3.5). Das damit gemeinte gezielte Verwenden emotionaler Ausdrucksmittel soll dabei auch hinsichtlich seiner Bedeutung für Dwecks Leistungsmotivationsmodell untersucht werden. Zur Diskussion unterschiedlicher

Erhebungsmöglichkeiten von Emotionen nach Leistungsrückmeldungen sei auf Abschnitt 4.3 verwiesen.

3.3.1 Emotionen in der klassischen Leistungsmotivationsforschung

Bereits im Risikowahlmodell Atkinsons (1957), vor allem aber im darauf aufbauenden Selbstwertungsmodell Heinz Heckhausens (1989) spielen Emotionen eine prominente Rolle. Es wird davon ausgegangen, dass (antizipierte) Gefühle des Stolzes und der Beschämung zu den zentralen Motivatoren des leistungsthematischen Handelns zählen. Stolz über die eigene Leistung hat demnach eine verstärkende Funktion für die Leistungsmotivation, Scham hingegen eine hemmende. Diese retrospektiven Emotionen stehen mit prospektiven Emotionen in einer sich wechselseitig verstärkenden Beziehung. Stolz ist mit Zuversicht bezüglich neuer Leistungsanforderungen verknüpft, Scham mit dem Gefühl der Hoffnungslosigkeit.

Diese theoretischen Vorstellungen werden im Rahmen eines typologischen Modells empirisch umgesetzt: Sogenannte erfolgsmotivierte Personen setzen sich mittelschwere Ziele, erleben deswegen ein ausgewogenes Verhältnis von Erfolgen und Misserfolgen und sind daher geneigt, ihre Ergebnisse eher auf ihre Anstrengung als auf ihre Fähigkeit zurückzuführen, während misserfolgsmotivierte Personen sich besonders leichte oder besonders schwere Ziele setzen, daher entweder nur Erfolge oder nur Misserfolge erleben und, weil sie so den Einfluss ihres eigenen Bemühens nicht wahrnehmen können, ihre Ergebnisse auf stabile Fähigkeiten oder externe Einflüsse attribuieren. Die unterschiedlichen Attributionsmuster bestimmen dabei das Vorkommen der sogenannten Selbstbewertungsemotionen Stolz und Scham: Misserfolgsmotivierte Personen erleben infolge ihres Rekurrerens auf die eigene Fähigkeit mehr Scham und wegen ihrer Tendenz zu externalen Attributionen weniger Stolz als Erfolgsmotivierte (Weiner, 1985; s.u.). Aufgrund dieser eher negativen Affektbilanz sind Misserfolgsmotivierte in einer neuen Aufgabe wiederum weniger zuversichtlich und setzen sich erneut unrealistischere Ziele als Erfolgsmotivierte (vgl. Brunstein & Heckhausen, 2006).

In der Tradition des Risikowahl- und Selbstbewertungsmodells finden Emotionen also einerseits in durchaus differenzierter Weise Berücksichtigung. Zu konstatieren ist andererseits, dass diese Berücksichtigung überwiegend in theoretischer, nur selten in empirischer Form erfolgte. Stolz, Scham, Zuversicht und Hoffnungslosigkeit wurden eher als Erklärungskonstrukte für Befunde im Rahmen des Risikowahlmodells benutzt als selbst einer empirischen Untersuchung unterzogen. Stolz und Scham etwa wurden nur in entwicklungspsychologischer Perspektive empirisch untersucht (vgl. J. Heckhausen & Heckhausen, 2006),

während sie ansonsten als eher nebulöse Konstrukte zur Schließung theoretischer Erklärungslücken Verwendung fanden.

Entsprechend einseitig ist auch die Auswahl der postulierten Emotionen: Weshalb spielt etwa Schuld keine Rolle für das Leistungshandeln? Sollte nicht die Schuld aufgrund zu geringer Anstrengung die als so positiv angenommene Affektbilanz der Erfolgsmotivierten eintrüben? Sollte gemäß dem schlicht verstärkungstheoretischen Modell Heckhausens nicht also Schuld als negative Emotion zu einer Hemmung der Leistungsmotivation führen? Diese Fragen machen deutlich, dass die klassische Leistungsmotivationsforschung zwar wertvolle begriffliche Anregungen lieferte, aber Emotionen – trotz des ihnen zugewiesenen theoretischen Gewichts – empirisch nicht wirklich ernstnahm. Erst die Arbeiten Weiners rückten Emotionen auch in empirischer Hinsicht ins Zentrum der Aufmerksamkeit.

3.3.2 Attributionale Ansätze: das Beispiel selbstrelevante Emotionen

Weiner (1985; 1995) beschäftigte sich mit der Beziehung zwischen Attribution und Emotion und legte damit nicht zuletzt wichtige begriffslogische Grundlagen für die Erforschung von Leistungsemotionen. Insbesondere ist in den Arbeiten Weiners die Grundlage der Erforschung der sogenannten *selbstrelevanten Emotionen* („self-conscious emotions“, „self-relevant emotions“) zu sehen, die sich gegenwärtig wachsender Beliebtheit erfreut (vgl. Lewis, 2000; Tracy, Robins, & Tangney, 2007). Als selbstrelevante Emotionen werden diejenigen Emotionen bezeichnet, die eine Bewertung des eigenen Selbst implizieren (vgl. genauer: Tracy & Robins, 2004). Weiner, der allerdings noch nicht jenes Label benutzte, unterscheidet hierbei vor allem Scham, Schuld und Stolz.

Scham tritt nach Weiner auf, wenn eine Person einen Misserfolg auf interne, unkontrollierbare Faktoren zurückführt. Ein Schüler sollte also z.B. Scham empfinden, wenn er seine schlechte Note auf ein zu geringes Ausmaß an angebotener Intelligenz zurückführt. *Schuld* hingegen sieht Weiner als Folge internalisierbarer Attributionen eines Misserfolgs an, d.h. ein Schüler sollte z.B. dann Schuld erleben, wenn er eine schlechte Note auf seine zu geringen Lernbemühungen zurückführt. Stolz ist gemäß Weiner eine emotionale Reaktion infolge einer internalen Attribution eines Erfolges. Neuere Theorien unterscheiden hier analog zum Misserfallsfall zwischen zwei verschiedenen Emotionen: Nach Tracy und Robins (2004; 2007) lässt sich Stolz von sogenannter Hybris unterscheiden. Demnach wird *Stolz* durch internalisierbare Attributionen eines Erfolges ausgelöst und kann sich somit im obigen Beispiel auf die großen Lernbemühungen des Schülers beziehen, während *Hybris* durch internalisierbare Attributionen eines Erfolges entsteht und sich z.B. auf das hohe Ausmaß eigener angebotener Intelligenz richten könnte.

Tracy und Robins sehen neben der Dimension der Kontrollierbarkeit auch die der Stabilität und Globalität als kennzeichnend für die Differenzierung zwischen den selbstrelevanten Emotionen an. Allgemein besteht noch Uneinigkeit darüber, welche Dimensionen (neben der Lokation) für bestimmte selbstrelevante Emotionen notwendig sind und welche deren Intensität nur verstärken (vgl. Lewis & Sullivan, 2005; Tracy & Robins, 2004, 2007). Einigkeit besteht wohl darüber, dass Scham und Hybris *eher* durch unkontrollierbare, stabile und globale Attributionen ausgelöst werden, Schuld und Stolz *eher* durch kontrollierbare, variable und spezifische. Im vorliegenden Rahmen erscheint eine genauere Klärung dieser Frage aber auch nicht notwendig; ebenso soll hier auf die häufig getroffene Unterscheidung zwischen Scham und Verlegenheit nicht näher eingegangen werden.

Auch für andere Emotionen liegen umfangreiche Untersuchungen über die ihnen zugrundeliegenden Attributions- oder Einschätzungsmuster vor, auch und gerade im Leistungsbereich. Smith und Ellsworth (1987) berechneten Korrelationen zwischen bestimmten Situationseinschätzungen und Emotionen im Anschluss an die Rückgabe einer Klausur. Die Hypothesen über jene Zusammenhänge konnten überwiegend bestätigt werden. Nicht zuletzt wegen solch auffällig guter empirischer Belege wird häufig gegen attributionstheoretische Forschungsprogramme der Einwand vorgebracht, ihre Ergebnisse seien nichts weiter als eine Explikation des Common-Sense (vgl. Reisenzein, Meyer, & Schützwohl, 2003). Dieser Einwand ist einerseits berechtigt, muss aber andererseits nicht als Kritik betrachtet werden. Die Attributionsforschung liefert – gerade für den Bereich der Emotionspsychologie – wichtige Hilfen zur Konzeptualisierung und Definition von Emotionswörtern (vgl. Laucken, 1989; Smedslund, 1988). Kein Emotionspsychologe kann Aussagen darüber treffen, ob eine Probandin Scham empfindet – außer er befragt sie und glaubt ihren Auskünften. Durch die weitgehend fehlenden körperlichen Anzeichen für selbstrelevante Emotionen ist der Forscher noch weit mehr als bei körpernäheren Emotionen wie Angst auf die sprachliche Auskunft seiner Probanden angewiesen (vgl. Tracy & Robins, 2007). Die Attributionsforschung liefert hier einen begrifflichen Rahmen, der deutlich macht, wie Menschen den Schambegriff verwenden und wann man überhaupt „sinnvollerweise“ Scham empfinden kann. So weiß der Forscher, dass es keinen Sinn hat, seinen Probanden z.B. zu fragen, ob er aufgrund seiner geringen angeborenen Intelligenz Schuld empfindet. Würde er diese Frage nämlich bejahen, wäre es ungleich wahrscheinlicher, dass er den Schuldbegriff falsch verwendet, als dass er „unerlaubterweise“ tatsächlich ein Schuldgefühl hinsichtlich seiner Intelligenz verspürt (vgl. Wittgenstein, 1963, 1999).

Diese Einsicht erleichtert beispielsweise das Erstellen von Fragebogen-Items enorm. Anstatt einfach nur die bloße Frage „Schämen sie sich?“ zu stellen und damit zu riskieren, dass die Probanden ganz unterschiedliche Konnotationen mit

dem Begriff der Scham verbinden, lässt sich auf diese Weise eine Emotion begrifflich klar eingrenzen und damit theoretisch fundiert erheben. Man würde den Probanden die Frage „Schämen Sie sich dafür, dass Ihre Intelligenz nicht für das Bestehen ausgereicht hat?“ stellen und damit eine relativ klar definierte Form der Scham erheben. Diese Überlegungen werde ich in Abschnitt 4.3, wo es um die Formulierungen der Items für den empirischen Teil dieser Arbeit geht, wieder aufnehmen.

Untersuchungen über die motivationalen Konsequenzen der selbstrelevanten Emotionen liegen derzeit überraschenderweise nur sehr vereinzelt vor. Hat beispielsweise Scham eher eine positive, gewissermaßen „anstachelnde“ Wirkung auf die Lernmotivation, oder führt sie eher zu defensiven, versteckenden Lernverhaltensweisen, wie es Heckhausens Selbstbewertungsmodell vorhergesagt hatte (vgl. Abschnitt 3.3.1)? Erste Untersuchungen zu dieser Frage (z.B. J. E. Turner & Schallert, 2001; vgl. auch Miller, 2007) geben noch keine befriedigende Antwort. Zu den anderen selbstrelevanten Emotionen liegen sogar noch weniger brauchbare Forschungsansätze vor.

Die attributionstheoretische Emotionsforschung stellt also wenig mehr als ein begriffliches Orientierungsmuster zur Verfügung. Einen vor allem in empirischer Hinsicht fruchtbareren Ansatz zur Entstehung von Leistungsemotionen liefern Untersuchungen zur Zielorientierung, um die es im Folgenden gehen soll.

3.3.3 Emotionen in Abhängigkeit von Zielorientierungen

Die Fülle der pädagogisch-psychologischen Arbeiten, die sich mit dem Konzept der *Zielorientierung* (teilweise auch als *Motivationale Orientierung* bezeichnet) beschäftigen, ist kaum mehr überschaubar. Es kann daher hier kein vollständiger Literaturüberblick angestrebt, sondern lediglich in Grundzügen herausgearbeitet werden, welche Rolle Emotionen innerhalb dieses Forschungsansatzes spielen. Ich orientiere mich dabei zunächst an den frühen Arbeiten von Dweck (Diener & Dweck, 1978; Dweck & Leggett, 1988) und komme anschließend auf neuere empirische Beiträge zu dieser Frage zu sprechen.

Dweck (1986; vgl. ähnlich: Nicholls, 1984) schlug die Unterscheidung zwischen Lern- und Leistungszielen vor, um verschiedene übergeordnete Ziele zu kennzeichnen, die in einer Leistungssituation verfolgt werden können. Empirische Studien hatten gezeigt, dass manche Schüler in Leistungssituationen eher auf Lernzuwachs und die Steigerung der eigenen Fähigkeiten aus sind („Lernziele“), während andere primär bestrebt sind, die eigene Leistungsfähigkeit vor anderen zu demonstrieren und mögliche Defizite zu verbergen („Leistungsziele“). Dabei erwies sich eine Lernzielorientierung für die Lernmotivation und langfristig auch für die Leistung als förderlich, eine Leistungszielorientierung hingegen als abträglich. Trotz der engen begrifflichen Verwandtschaft werden die beiden

Zielorientierungen als unabhängig konzeptualisiert, d.h. eine Person kann gleichzeitig stark ausgeprägte Lernziele *und* stark ausgeprägte Leistungsziele aufweisen. Zumeist werden diese beiden Konstrukte im Sinne von Dispositionen mittels Fragebogen erhoben, teilweise auch über bestimmte Aufgabenstellungen experimentell induziert. Die jüngste Ausdifferenzierung der Konstrukte hinsichtlich Annäherungs- und Vermeidungszielen (z.B. Elliot & Harackiewicz, 1996) spielt im vorliegenden Rahmen keine Rolle.

Die Arbeitsgruppe um Dweck (Diener & Dweck, 1978; Dweck & Leggett, 1988; vgl. auch Dweck, 1999) konnte in experimentellen Untersuchungen zeigen, dass Lern- und Leistungsziele zu ganz bestimmten, distinkten Verhaltensmustern nach Misserfolg führen. Lernziele fördern Anstrengungsattributionen, persistentes Leistungsverhalten und positive Emotionen, während Leistungsziele Fähigkeitsattributionen, aufgebendes Leistungsverhalten und negative Emotionen nach sich ziehen. Erstgenannte Reaktionen wurden in Anlehnung an Befunde zur Erlernten Hilflosigkeit (Seligman, 1979) zusammenfassend als *meisternd* („mastery oriented“) bezeichnet, letztgenannte als *hilflos* („helpless“). Die zusätzlich getroffene Unterscheidung in Abhängigkeit des Fähigkeitsselbstkonzepts wird hier nicht näher diskutiert, da sie mittlerweile fallengelassen wurde.

Leistungsempfindungen erscheinen in dieser Forschungsrichtung also als Teil eines übergreifenden Verhaltensmusters, das durch das relative Überwiegen einer Lern- oder einer Leistungszielorientierung bestimmt ist. Dabei empfinden lernzielorientierte Personen eher positive Emotionen, leistungszielorientierte Personen eher negative. Dweck und Leggett (1988) stellen auch empirisch gestützte Vermutungen darüber an, um welche Emotionen es sich dabei vor allem handeln könnte: Stolz, Zufriedenheit und Entschlossenheit auf der einen Seite, Angst, Depression und Scham auf der anderen.

Erst in jüngster Vergangenheit ist die Frage nach der Beziehung zwischen Zielorientierung und Emotionen eingehender empirisch untersucht worden. Häufig stehen hierbei allerdings tätigkeitsbezogene Emotionen (Lernfreude, Langeweile etc.) im Vordergrund (vgl. Linnenbrink & Pintrich, 2002). Beispielsweise erhob Seiffert (1995) mittels Fragebogen die allgemeine emotionale Gestimmtheit von Schülern in der Schule und setzte sie mit der ebenfalls dispositionell erhobenen Zielorientierung in Beziehung. Dabei versuchte er, über Strukturgleichungsmodelle die Ursache-Wirkungs-Richtung der beiden Konstrukte zu erschließen. Im Zuge dessen kommt er z.B. zu dem Schluss, dass Angst Leistungsziele bedingt und nicht umgekehrt.

Außerdem liegen einige Studien vor, die methodisch ähnlich vorgehen, aber auch prospektive und retrospektive Emotionen miteinbeziehen (Knollmann & Wild, 2004; Pekrun, Elliot, & Maier, 2006; J. C. Turner, Thorpe, & Meyer, 1998). Exemplarisch soll hier die Arbeit von Pekrun et al. (2006) vorgestellt werden. Die Probanden erhielten Fragebögen zu Zielen und Emotionen, wie sie

sie „typischerweise“ im Studium verfolgen bzw. erleben. Knapp zusammengefasst besagen die Ergebnisse, dass Lernzielorientierte eher zu Lernfreude, Hoffnung, Ärger und Stolz/Hybris tendieren, Leistungszielorientierte hingegen zu Angst, Scham und Langeweile. Methodische Schwächen offenbaren sich allerdings an mehreren Stellen. Beispielsweise musste der Zielfragebogen, so die Autoren, zunächst an den Untersuchungszweck angepasst werden, indem das explizite Ansprechen von Emotionen (z.B. „My fear of performing poorly in this class is often what motivates me“, S. 587) vermieden wurde. Einfach Items umzuformulieren ändert indes wenig an dem zugrundeliegenden Problem, dass das Konstrukt der Zielorientierung auf semantischer Ebene in vielerlei Hinsicht mit Emotionen verbunden ist. Ein Lernzielorientierter wird häufig Lernfreude empfinden – warum sonst sollte er Lernzuwachs anstreben?

An vielen Stellen berühren Pekrun et al.s sowie die anderen zitierten Arbeiten somit die Grenzen der Trivialität. Sie werden damit zwar nicht gänzlich überflüssig. Ihre Aussagekraft sollte aber nicht überschätzt werden im Vergleich zu Studien, die Emotionen aktualgenetisch untersuchen oder sogar Zielorientierung experimentell induzieren. Nur wenn mindestens eine dieser beiden Strategien angewandt wird, kann vermieden werden, dass Ziele und Emotionen anhand mehr oder weniger identischer, situationsunspezifischer Fragebogen-Items erhoben werden. Das räumen übrigens auch die zitierten Studien ein: „In subsequent research, more fine-grained analyses addressing goals and emotions in specific situations should be used“ (Pekrun, Elliot, & Maier, 2006, S. 595), „experimental studies need to be conducted to distinguish among equivalent models and to establish causal relationships among these variables“ (J. C. Turner, Thorpe, & Meyer, 1998, S. 770), „es bleibt daher zu hoffen, dass in den kommenden Jahren die Analyse der Aktualgenese von Emotionen zunehmend stärker ins Zentrum pädagogisch-psychologischer Forschung rücken wird“ (Knollmann & Wild, 2004, S. 84). Trotz dieser immer wiederkehrenden Willensbekundungen sind solche Untersuchungen – wohl nicht zuletzt aufgrund eines „leistungszielorientierten“ Schielens auf sichere und somit publikationsträchtige Signifikanzen – immer noch Mangelware.

Jene zwar zahlreichen, aber methodisch unbefriedigenden korrelativen Fragebogen-Untersuchungen lassen nur in eingeschränktem Maße Schlüsse über die Beziehung zwischen Zielorientierung und Emotionen nach Leistungsrückmeldungen zu. In Anschluss an Pintrich und Schunk (2002) ist zu konstatieren, dass sich diese Beziehung überwiegend so gestaltet, wie gemäß einem attributions-theoretischen Ansatz auch zu vermuten gewesen wäre: Lernziele begünstigen aufgrund ihrer Verbindung zu Anstrengungsattributionen vor allem Stolz nach Erfolg und Schuld nach Misserfolg, während Leistungsziele größtenteils negative Emotionen wie Angst, Scham und Depression nach sich ziehen (vgl. auch Jagacinski & Nicholls, 1987). Allgemein ist die Befundlage aber nicht eindeutig,

so dass diese Zusammenhänge eher als Hypothesen denn als Ergebnisse verstanden werden sollten.

3.3.4 Emotionen in Abhängigkeit von Impliziten Intelligenztheorien

Bereits in Dwecks frühen Arbeiten (vgl. Dweck & Leggett, 1988) wird postuliert, dass Implizite Intelligenztheorien die unterschiedlichen Zielorientierungen ursächlich bedingen. Insofern lag von Anfang an nahe, dass sie auch Emotionen zumindest indirekt – nämlich über die Zielorientierung – beeinflussen sollten. Bemerkenswerterweise wurde jedoch darauf verzichtet, Implizite Theorien und Emotionen in ein und derselben Studie zu untersuchen, um so ihre spezifischen Zusammenhangsmuster ausfindig zu machen. Stets geriet nur der Einfluss Impliziter Theorien auf Zielorientierung (z.T. auch auf Verhalten) *oder* der Einfluss von Zielorientierung auf Emotionen ins Blickfeld. Es schien wohl offensichtlich und somit einer näheren Analyse nicht wert, dass zwischen statischen Intelligenztheorien, Leistungszielorientierung, Fähigkeitsattributionen und hilflosem Bewältigungsverhalten immer auch „a variety of negative emotions“ (Dweck, 1999, S. 8) ablaufen mussten. Umgekehrt wurden Dynamiker stets mit ausnahmslos positiven Emotionen in Verbindung gebracht. Erst um die Jahrhundertwende begannen Forscher (auch außerhalb von Dwecks Umfeld), im Rahmen von aufwändigen Längsschnittstudien eine Vielzahl von Merkmalen simultan zu erheben und somit auch die hier interessierende Beziehung zwischen Impliziten Theorien und Emotionen zu untersuchen.

So verfolgten Robins und Pals (2002) die Entwicklung von 363 Studenten über mehrere Jahre hinweg anhand von jährlichen Fragebogenerhebungen. Ihre Hypothese, dass Statiker eher zu negativen Leistungsemotionen neigen, Dynamiker hingegen zu positiven, konnte im Großen und Ganzen bestätigt werden, wenngleich die recht bescheidenen Korrelationen (um .1) wohl vor allem aufgrund der großen Stichprobe statistisch signifikant wurden. Gemäß den Bezeichnungen der verwendeten PANAS-Skalen (Watson, Clark, & Tellegen, 1988) fühlten sich Statiker eher „distressed“, „ashamed“ und „upset“ im Hinblick auf ihren gegenwärtigen Leistungsstand im Studium, während die Dynamiker ihre Stimmungslage eher als „determined“, „enthusiastic“, „excited“, „inspired“ und „strong“ bezeichnen wollten. Die restlichen PANAS-Werte (u.a. für „proud“, „guilty“ und „scared“) waren nicht signifikant. Die Autoren zeigen dann die Beziehungen eines zusammengefassten Emotionsscores (gepolt als „negative Emotionen“) zu den anderen in der Studie erhobenen Variablen in einer Pfadanalyse auf. Negative Emotionen sind danach nicht direkt von der statischen Intelligenztheorie abhängig, sondern nur vermittelt über (Leistungs-)Zielorientierung und fähigkeitszentriertem Attributionsstil. Ihrerseits bedingen sie hilfloses Bewälti-

gungsverhalten und geringes Selbstwertgefühl. Diese und andere von Robins und Pals gefundenen Zusammenhänge sind in der Abbildung 2 wiedergegeben.

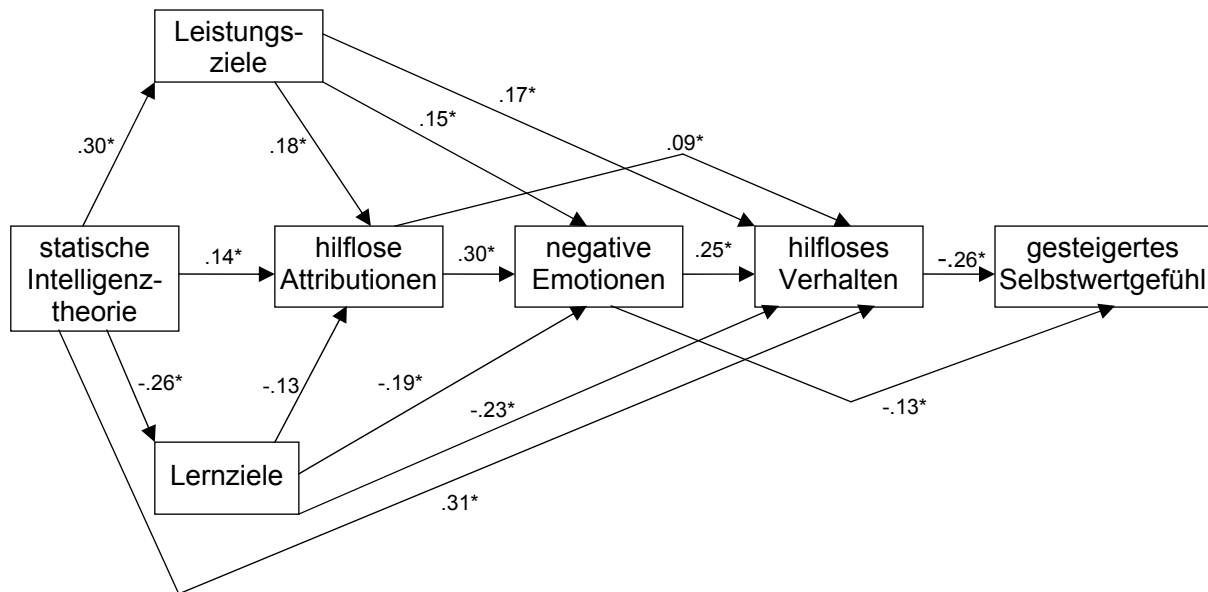


Abbildung 2. Pfadanalytische Zusammenhänge in der Studie von Robins und Pals (2002).

* $p < .05$

Robins und Pals unternehmen einen der wenigen Versuche, Dwecks gesamtes Modell in ein und derselben Studie zu testen. Berücksichtigt werden muss jedoch, dass sie nicht eigentlich Emotionen untersuchen, sondern – ähnlich wie die meisten Studien über den Zusammenhang von Zielorientierung und Emotionen, s.o. – eher die allgemeine Gestimmtheit hinsichtlich des eigenen Studiums. Das macht es um so schwieriger, diesen Stimmungen einen sicheren Platz im kausalen Bedingungsgefüge zuzuweisen – daran kann auch das methodische Instrument der Pfadanalyse nichts ändern: Stimmungen sind immer irgendwie am Leistungsgeschehen beteiligt, aber können eben keinen festen Platz in einem Prozessmodell beanspruchen. Die Ergebnisse der Pfadanalyse sind somit empirisch höchst interessant – und gleichzeitig theoretisch schwer interpretierbar.

Besser deutbar, weil eine konkrete Leistungssituation betreffend, sind da die Ergebnisse der bereits in Abschnitt 3.2.1 beschriebenen Induktionsstudie von Niiya et al. (2004), die Angst, Depression und Feindseligkeit messen und diese jeweils stärker bei den Statikern vorfinden. Wiewohl durchaus ermutigend, sollte die Aussagekraft dieser Ergebnisse jedoch nicht überschätzt werden, vor allem weil die Stichprobengröße (oft weniger als 10 Probanden pro Gruppe) unbefriedigend gering ist. Hinzu kommt, dass Niiya et al. Emotionen mit Hilfe einer allgemeinen Befindensskala erheben, welche nicht auf die Erfassung von Leistungsempfindungen ausgerichtet ist, wie dies bei neueren Instrumenten (Molfenter, 1999; Pekrun et al., 2004; vgl. Abschnitt 4.3) der Fall ist. Schließlich ist zu be-

mängeln, dass Niiya et al. keine Bewältigungs- und Leistungsmaße erheben und diese somit nicht mit den Emotionen in Beziehung gesetzt werden können.

Baer, Grant und Dweck (2005, zit.n. Molden & Dweck, 2006) beschäftigen sich in einer bislang unveröffentlichten Studie mit der Beziehung zwischen Impliziten Theorien und Dysphorie, einer milden Form der depressiven Verstimmtheit. In Tagebuch- und experimentellen Untersuchungen fanden sie heraus, dass Statiker eher zu Dysphorie neigen als Dynamiker. Außerdem zeigte sich, dass nach einer dysphorischen Reaktion auf einen akademischen Misserfolg die Statiker zu aufgebendem, die Dynamiker dagegen zu persistentem Bewältigungsverhalten neigten. Nicht nur das Auftreten von depressivem Affekt war also unterschiedlich, sondern anscheinend auch die motivationale Bedeutung, die dem depressiven Affekt zugewiesen wurde. Inwieweit diese Schlussfolgerungen angemessen sind, kann freilich aufgrund der fehlenden Zugänglichkeit der Originalarbeit nicht eingeschätzt werden.

Bis hierhin lässt sich die Befundlage zur Beziehung zwischen Impliziten Theorien und Emotionen wie folgt zusammenfassen:

- (1) Im Rahmen des klassischen Dweckschen Modells wurden Emotionen als Bestandteil eines übergreifenden Verhaltensmusters aufgefasst, das direkt durch Lern- und Leistungsziele und indirekt durch Implizite Theorien bestimmt wird. Dabei beinhaltet das hilflose Verhaltensmuster ausnahmslos negative, das meisternde Verhaltensmuster ausnahmslos positive Emotionen. Diese Sichtweise scheint sich – soweit man diese Schlüsse aus den wenigen und methodisch fragwürdigen Studien ziehen will – empirisch zu bestätigen.
- (2) Die in Abhängigkeit Impliziter Theorien erhobenen Emotionen wurden zumeist als bloße „Outcomes“ betrachtet. Ihre Auswirkungen auf Leistungsverhalten und Leistung wurden (mit Ausnahme von Robins und Pals, die sich jedoch auf korrelative, situationsunspezifische Daten beziehen, sowie von Baer et al., s.o.) nicht erhoben. Es kann somit nicht eingeschätzt werden, inwieweit Emotionen für die Beziehung zwischen Impliziten Theorien und Bewältigungsverhalten die Rolle von Mediatoren spielen. Diese Frage zu beantworten fällt übrigens um so schwerer angesichts der Tatsache, dass die Auswirkungen spezieller Emotionen auf Lernmotivation und Leistung auch jenseits von Dwecks Theorie bislang kaum untersucht worden sind (vgl. erste Ansätze eines systematischen Forschungsprogramms bei Pekrun, 2000). Zwar sind Leistungsemotionen auch für sich genommen interessant, weil sie – im Sinne eines allgemeinen Wohlbefindens – ein wichtiges Ziel pädagogischen Wirkens darstellen (Goetz, Zirngibl, & Pekrun, 2004). Dies rechtfertigt aber nicht die Vernachlässigung jener anderen, komplexeren Fragestellung.

Während der zweite Punkt zu Beginn des empirischen Teils dieser Arbeit (Abschnitt 4.1) wieder aufgegriffen werden wird, soll im Folgenden die erste Schlussfolgerung anhand neuerem empirischem Material näher diskutiert und

hinterfragt werden. Dabei steht sowohl die Vorstellung zur Debatte, nach der Emotionen direkt an bestimmte hilflose bzw. meisternde Verhaltensweisen gekoppelt sind, als auch die Aussage, dass Statiker ausnahmslos zu negativen Emotionen bzw. Dynamiker ausnahmslos zu positiven Emotionen neigen.

Dass hinsichtlich des letztgenannten Punktes die Dinge etwas komplizierter liegen könnten, legt schon ein Blick auf neuere kognitive Emotionstheorien (Ortony, Clore, & Collins, 1988; Weiner, 1985; vgl. Abschnitt 3.3.2) oder auf die Befunde der Zielorientierungsforschung (vgl. Abschnitt 3.3.3) nahe. Zwar ist die Hypothese theoretisch stimmig, dass Statiker aufgrund ihrer Tendenz zur Fähigkeitsattribution (auch wenn diese sich nicht immer empirisch nachweisen lässt, s.o.) mehr Hoffnungslosigkeit, Depression und Scham nach Misserfolg empfinden. Auch sollten Dynamiker zu mehr Zuversicht neigen, da sie ihren Misserfolg eher auf mangelnde Anstrengung zurückführen. Wie sieht es aber mit Schuld aus? Wiewohl in ihrer Valenz eindeutig negativ, passt sie doch theoretisch eher zum Dynamiker, der seinen Misserfolg mangelndem Bemühen zuschreibt und daher – sei es gegenüber sich selbst oder seinen Gruppenmitgliedern bzw. nahestehenden Menschen – Schuld empfindet.

Grant und Dweck (2001) sind die ersten und wohl bislang einzigen, die sich für diese Verknüpfung zu emotionspsychologischen Befunden interessieren. Sie referieren einige kleine unveröffentlichte Studien, deren Versuchsdesign ähnlich dem von Niiya et al. (2004) ist und hier daher nicht näher beschrieben wird. Von Belang ist hier auch weniger die empirische Evidenz als das theoretische Modell, das die Autorinnen daraus entwickeln. Ihre wichtigste Schlussfolgerung besteht darin, dass Emotionen in Leistungssituationen nicht nur von der Attribution, sondern auch vom sozialen Kontext abhängen (könnten): So erhöhe nicht nur Fähigkeitsattribution, sondern auch die soziale Interdependenz einer kooperativen Leistungssituation die Tendenz, nach einem Misserfolg Scham zu empfinden; ebenso steigere nicht nur Anstrengungsattribution, sondern auch jene Interdependenz (wie sie etwa in Form einer Gruppenarbeit häufig vorkommt) die Tendenz zum Schuldgefühl. Diese Unterschiede zwischen Leistungssettings mit individuellen und solchen mit gemeinsamen Zielen spiegeln sich gemäß Grant und Dweck auch auf interkultureller Ebene wider – hier zwischen Menschen in individualistischen und kollektivistischen Kulturen (bzw. mit independentem und interdependentem Selbst; vgl. Markus & Kitayama, 1991). Daher sollen jene beiden situationalen Arrangements von nun an *individualistische* bzw. *kollektivistische* Leistungssituation heißen.

Die kollektivistische Leistungssituation führt nun aber, so Grant und Dweck (2001) weiter, nicht nur zu einer Verstärkung von Scham und Schuld, sondern motiviert gleichzeitig zu Bewältigungsverhalten (man will ja seine Kollegen nicht im Stich lassen). Menschen in kooperativen Leistungskontexten bzw. kollektivistischen Kulturen zeigen gewissermaßen ein „hybrides“ Verhaltensmuster,

d.h. sie sind einerseits ausdauernd in ihrem Leistungsverhalten, andererseits aber emotional äußerst labil. Grant und Dweck bieten als Erklärungskonzept so etwas wie *Anstrengungs-Leistungsziele* (meine Wortschöpfung) an, d.h. nach dem individuellen Misserfolg in der Gruppensituation entsteht beim Betroffenen plausiblerweise das Bedürfnis, sich vor den Kollegen als bemüht (und nicht als fähig) darzustellen. Diese Anstrengungsorientierung schlägt sich nun nicht nur in erhöhtem Arbeitseinsatz nieder, sondern eben auch in einer stärkeren Neigung zu negativen Emotionen wie Scham und Schuld, die sich aus den Erwartungen der Mitmenschen an die eigene Leistung heraus ergeben.

Allem Anschein nach müssen also die zuvor stets als parallel angenommenen Reaktionsmuster im Verhalten (aufgebend vs. persistent) einerseits und im emotionalen Bereich (negativ vs. positiv) andererseits konzeptuell getrennt werden: Emotionen sind nicht länger Anhängsel eines übergreifenden „hilflosen“ bzw. „meisternden“ Verhaltensmusters, sondern führen gleichsam ein Eigenleben.

Berücksichtigt man nun die empirisch recht gut belegten Beziehungen zwischen Impliziten Theorien und Attribution, liegt es nahe, Leistungsemotionen als von zwei Faktoren abhängig zu sehen: (1) von der Impliziten Theorie, die über ihre Auswirkungen auf die Attribution die jeweilige Emotion (mit-)bestimmt und (2) von der Leistungssituation, die den Misserfolg mehr oder weniger selbstrelevant macht und somit die Tendenz zu Scham und Schuld erhöht bzw. verringert.

Wenngleich Grant und Dwecks Ergebnisse alles andere als eindeutig sind, geben sie doch wichtige Denkanstöße, die in die Hypothesen der vorliegenden Arbeit eingeflossen sind. Bevor ich auf diese näher zu sprechen komme, soll es noch, gewissermaßen in einem kleinen Exkurs, um das Thema der Emotionsdarstellung gehen. Der Anknüpfungspunkt an die eben angestellten Überlegungen liegt allerdings auf der Hand: Wenn sich Menschen aufgrund sozialen Drucks als bemüht (oder als fähig) darstellen sollen bzw. wollen, liegt es ja nahe zu untersuchen, ob sie diese Darstellung auch über das Darbieten bestimmter Emotionen zu Wege bringen: Wer als eifrig gelten will, sollte z.B. nach einem Misserfolg eher Scham darbieten als Schuld, um die Attributionen seiner Mitmenschen in die „richtige“ Richtung zu lenken.

3.3.5 Emotionsdarstellung und ihre Rolle in Dwecks Modell

Die gezielte soziale Darstellung von Emotionen wurde im Lern- und Leistungskontext bislang selten untersucht. Wie Menschen sich gegenüber ihrer sozialen Umwelt emotional positionieren, ob sie beispielsweise gegenüber einer schlechten Note gleichgültig oder traurig erscheinen wollen, sich vor einer Klausur nervös oder zuversichtlich geben möchten, ist dennoch – mindestens implizit – Gegenstand zahlreicher pädagogisch-psychologischer Theorien. Bevor ich auf diese Theorien näher eingehen werde, will ich kurz darlegen, was das Konzept der

Emotionsdarstellung im vorliegenden Kontext beinhalten soll. Am Ende dieses Abschnitts soll es dann um Möglichkeiten gehen, dieses Konzept in Dwecks Leistungsmotivationsmodell zu integrieren und empirisch handhabbar zu machen.

Laux und Weber (1993) verstehen Emotionsdarstellung als spezielle Form der Selbstdarstellung, d.h. als gezielte Lenkung des Eindrucks, den die Mitmenschen von der eigenen Person haben. Im Falle der Emotionsdarstellung vollzieht sich diese Eindruckslenkung mit Hilfe der „Abschwächung, Übertreibung oder qualitativen Modifikation des Emotionsausdrucks“ (ebd., S. 39). Als übergeordnete Ziele der Emotionsdarstellung unterscheiden Laux und Weber Selbst- von Interaktionsregulation. Das Beeinflussen des Emotionsausdrucks wird also mit der Intention vorgenommen, einerseits dem eigenen Selbstkonzept zu entsprechen und das eigene Selbstwertgefühl zu erhöhen und andererseits das Verhalten anderer Menschen zu beeinflussen. Diese Ziele (bzw. deren situationsspezifische Ausformung) müssen dem Handelnden nicht immer vollständig bewusst sein.

Der Ausdruck von Emotionen eignet sich deshalb besonders gut zur Selbstdarstellung, weil Emotionen (in der westlichen Kultur) als eher schwer beeinflussbar, ja als „authentischer“ Ausdruck der eigenen Person gelten (vgl. Parkinson, Fischer, & Manstead, 2005). Wer in einem Streit mit dem Partner anfängt zu weinen, hat gute Chancen, dass der Partner der eigenen Position Aufmerksamkeit schenkt und ihre Glaubwürdigkeit nicht anzweifelt. Der Partner könnte zwar die Echtheit des Weinens anzweifeln, würde aber mit dieser Behauptung so sehr gegen kulturelle Konventionen verstoßen und folglich in die Defensive geraten, dass er – im Regelfall zumindest – keine Chance hat, den Erfolg dieser Selbstdarstellung (wenn es sich denn im gegebenen Fall um eine handelt) zu verhindern.

Das Konzept der Selbstdarstellung – und mithin auch das der Emotionsdarstellung – ist vielen Leistungsmotivationstheorien implizit. Schon in der Tradition des Risikowahlmodells (vgl. Abschnitt 3.3.1) tritt es in Erscheinung, wenn misserfolgsmotivierte Personen nur besonders schwere oder besonders leichte Aufgaben in Angriff nehmen. Damit stellen sie sich nämlich – zumindest in der Logik des Risikowahlmodells – im einen Fall als besonders mutig, zuversichtlich und selbstsicher, im anderen Fall als besonders souverän und erfolgreich dar. Erfolgsmotivierte Personen verzichten hingegen auf solche „Krücken“ zur Selbstdarstellung und stellen sich mittelschweren, potentiell selbstwertgefährdenden Aufgaben.

Ähnlich lassen sich die Befunde der Zielorientierungsforschung (vgl. Abschnitt 3.3.3) deuten. Hier sind es die Leistungszielorientierten, die gern als intelligent, erfolgreich und souverän, keinesfalls aber als bemüht, strebsam und zuweilen durch Misserfolge zurückgeworfen angesehen werden wollen. Um dieser Sichtweise von sich selbst Nachdruck zu verleihen, geben sie etwa vor ihren

Freunden an, nicht für die Klausur gelernt zu haben, oder reduzieren gar ihre tatsächlichen Lernanstrengungen. Letzteres Verhalten ist in der pädagogisch-psychologischen Literatur als *Self-handicapping* bekannt (Rhodewalt & Hill, 1995; vgl. für das verwandte Konzept der Anstrengungsvermeidung im deutschen Sprachraum: Helmke & Rheinberg, 1996). Self-handicapping umfasst alle Maßnahmen, die geeignet sind, den eigenen (potenziellen) Misserfolg vor anderen und vor sich selbst durch andere Faktoren als durch Fähigkeit erklärbar zu machen, also neben geringer Klausurvorbereitung etwa das Konsumieren von Alkohol direkt vor der Klausur. Self-handicapping kann als eine spezielle Form der Emotionsdarstellung gedeutet werden, eine solche nämlich, bei der Emotionen im Hinblick auf eine Leistungssituation (etwa Gleichgültigkeit, Verachtung, Desinteresse, aber auch Zuversicht und Vertrauen in die eigenen Fähigkeiten) durch deutlich sichtbare, selbstschädigende Verhaltensweisen zum Ausdruck gebracht werden (sollen).

Häufiger als Self-handicapping dürfte unter Schülern und Studenten jedoch die Strategie sein, den Eindruck über die eigenen Fähigkeiten und Lernanstrengungen über verbale und mimische Kommunikation (statt über konkrete selbstschädigende Maßnahmen) zu lenken. Juvonen (2001) interessiert sich dabei für die Frage, ob Schüler dabei wirklich unter allen Umständen das Ziel verfolgen, den Eindruck von ausgeprägten Fähigkeiten und geringen Lernbemühungen zu vermitteln, wie es ja aufgrund der Befunde zur Risikowahl, zur Zielorientierung und zum Self-handicapping im Großen und Ganzen zu erwarten wäre. Es zeigte sich, dass die Ziele der Schüler mit dem Alter und der Zielperson variieren. Während Viertklässler sich noch lieber als bemüht darstellen wollen, zeigen Achtklässler bereits den Hang, sich als intelligent zu positionieren. Doch auch Achtklässler zeigen ein deutlich differenzierteres Darstellungsverhalten als erwartet: Einen Misserfolg in einer Klausur erklären sie vor den Mitschülern lieber als Resultat fehlender Anstrengung, während sie vor dem Lehrer lieber als unfähig dastehen wollen. Noch interessanter ist indes, dass die von Juvonen befragten Schüler ihre Misserfolge auch vor ihren Mitschülern nicht selten als Resultat mangelnder Fähigkeiten hinstellen, und zwar besonders dann, wenn sie sich mit ihnen verbunden fühlen bzw. von ihnen gemocht werden wollen.

Juvonens Untersuchung (vgl. für deren attributionstheoretische Einbindung auch Weiner, 2005) macht somit deutlich, dass es auch gute Gründe geben kann, sich „als dumm hinzustellen“ – nämlich besonders dann, wenn es der Person weniger um selbstwertbezogene als vielmehr um affiliative Ziele geht. Diese Befunde weisen eine interessante Übereinstimmung mit denen von Grant und Dweck (2001; vgl. Abschnitt 3.3.4) auf. Was hier die affiliativen Ziele hervorgerufen, bewirkt dort die kollektivistische Leistungssituation bzw. das interdependente Selbst. Ob man sich nun für seine Leistung im Rahmen einer Gruppenarbeit verantworten muss, in einem informellen Gespräch über Prüfungsergebnisse

die Akzeptanz seiner Freunde durch Aufrichtigkeit bzw. Understatement bewahren will oder gar in einer kollektivistischen Leistungskultur lebt, in der der individuelle Misserfolg auf die Familie zurückschlägt: immer scheint ein ähnlicher Mechanismus am Werk, der Menschen dazu veranlasst, sich lieber als dumm und eifrig denn als schlau und faul hinzustellen. Immer scheinen (sub-)kulturelle Normen zu existieren, die das Verhalten der Betroffenen stärker beeinflussen als die „grundlegende“, von der Leistungsmotivations- und Impression-Management-Forschung postulierte Strategie des Selbstwertschutzes.

Wie stark solche Normen auch innerhalb der westlichen Kultur variieren können, zeigen übrigens jüngste Untersuchungen zum Konzept des Strebers (Boehnke, Pelkner, & Kurman, 2004). Während an deutschen Schulen der sozial ausgrenzende Vorwurf, ein Streber (d.h. ein übereifriger, „büffelnder“ Langweiler) zu sein, zum Alltag gehört, existiert in Nordamerika noch nicht einmal ein entsprechender Begriff, geschweige denn nennenswerte soziale Ausgrenzung aufgrund (zu) guter Leistungen.

Allem Anschein nach gibt es keine gleichsam natürliche Tendenz, sich vor anderen Menschen in bestimmter Weise darzustellen. Es ist wohl eher der ganz spezielle soziale und kulturelle Kontext, in dem sich eine Person befindet, der bestimmt, welche Darstellung der Person am angemessensten und am gewinnbringendsten erscheint. Befunde wie die von Juvonen und Boehnke et al. unterstützen damit die jüngsten Bemühungen, sozialkonstruktivistische und kulturpsychologische Denkmuster in die Mainstream-Psychologie zu integrieren (Heine & Norenzayan, 2006; von Tiling, 2008a, 2008b; vgl. Abschnitt 9.4). In diesen Ansätzen wird schon seit geraumer Zeit das Verhältnis von Individuum und Kultur als ein sich wechselseitig konstituierendes gesehen, etwa mit Hilfe der Metapher des „Positioning“ (Davies & Harré, 1990): Kulturelle Ressourcen messen hiernach dem Tun eines Menschen immer schon eine Bedeutung bei, die der einzelne nicht beeinflussen kann; andersherum aber steht es dem Einzelnen immer mehr oder weniger frei, unter den zur Verfügung stehenden kulturellen Ausdrucksmöglichkeiten gezielt auszuwählen und sie mithin für seine Zwecke zu gebrauchen. Positioniert werden und Positionieren sind gleichsam zwei Seiten derselben Medaille. Übertragen auf den vorliegenden Fall heißt das: Die Selbstdarstellungsregeln gibt der soziale und kulturelle Kontext vor; das Individuum ist nicht gezwungen, sich an diese Regeln zu halten – aber es weiß dennoch, dass sein Handeln immer nur relativ zu diesen Regeln bewertet werden kann.

Nicht nur Emotion (wie von Grant und Dweck gezeigt), sondern auch die Darstellung von Emotionen sollte demzufolge mit dem sozialen Kontext variieren. Wer sich im Rahmen einer Gruppenarbeit vor seinen Kollegen als fleißig darstellen will, sollte deren Attributionen eigener Misserfolge auf Fähigkeit, eigener Erfolge hingegen auf Anstrengung zu lenken versuchen. Gemäß der attributionstheoretischen Emotionsforschung (vgl. Abschnitt 3.3.2) müsste ein pro-

bates Mittel hierzu sein, auf eigene Misserfolge mit Scham zu reagieren („Die Aufgabe war einfach zu schwer für mich!“), auf eigene Erfolge mit Stolz („Mein Fleiß hat sich ausgezahlt!“). In einer individualistischen Leistungssituation (vgl. Abschnitt 3.3.4) dagegen, in der man sich als fähig bzw. intelligent positionieren will, sollte man, um die Attributionen der Mitmenschen auf Anstrengung bei Misserfolg bzw. auf Fähigkeit bei Erfolg zu lenken, mit der Darstellung von Schuld („Ich hätte vielleicht doch vorher lernen sollen“) bzw. mit Hybris („Die Aufgaben fand ich viel zu leicht!“) reagieren.

In welcher Beziehung sollte Emotionsdarstellung aber zu Impliziten Intelligenztheorien stehen? Plausiblerweise sollten Statiker infolge ihres Glaubens an eine unveränderliche Intelligenz eher dazu neigen, sich so darzustellen, als hätten sie eine besonders große Menge an Intelligenz „mitbekommen“. Sie sollten sich daher ähnlich wie Personen in individualistischer Lernumwelt präsentieren. Dynamiker sollten sich weniger fähigkeitsorientiert darstellen, wie es ja auch die oben erwähnten Befunde in Abhängigkeit zur Zielorientierung nahelegen. Wenn Dynamiker sich überhaupt gezielt darstellen wollen, dann wohl eher analog zu den Kollektivistinnen als anstrengungsorientiert. Sie sollten außerdem Emotionen wie Lernfreude und Interesse zeigen, während Statiker mit der Darbietung von Langeweile eher wieder von ihren Misserfolgen abzulenken versuchen sollten. Letztere Vermutungen lassen sich bereits aus den Ergebnissen von Dweck und Leggett (1988) ableiten. Einen weiteren Hinweis für die Stichhaltigkeit der hier vorgestellten Hypothesen liefern Rhodewalt und Vohs (2005), die zeigen konnten, dass Statiker eher zum Self-handicapping neigen als Dynamiker.

Der Forschung sind hinsichtlich dieser Fragen allerdings auch Grenzen gesetzt. Da Emotionen in aller Regel über den Selbstbericht erhoben werden, ist es empirisch nur schwer möglich, zwischen tatsächlich erlebten und bewusst dargebotenen Emotionen klar zu unterscheiden. Der Befragte steht ja immer gleichsam im Verdacht, auch auf die Frage nach dem Erleben mit einer Darstellung zu antworten. Daher macht es Sinn, das *berichtete* Erleben und Darstellen mit dem von außen *beobachteten* abzugleichen. Die vorliegende Arbeit wählt diesen anspruchsvollen empirischen Zugang zumindest gelegentlich (vgl. die Abschnitte 6.2.1 und 7.2.1).

3.3.6 Fazit

Emotionen nach Leistungsrückmeldungen waren bislang durchaus Gegenstand mancher Forschungsbemühungen. Einige Arbeiten beschäftigten sich dabei eher mit begrifflich-deskriptiven Fragen, untersuchten etwa die Vorkommenshäufigkeit (Pekrun, 1998) oder die kognitiven Einschätzungsmuster (Smith & Ellsworth, 1987; Weiner, 1985) dieser Emotionen. Relativ unabhängig davon wurde

die Bedeutung von Emotionen nach Leistungsrückmeldungen innerhalb bestimmter pädagogisch-psychologischer Forschungsansätze untersucht.

Insbesondere im Rahmen von Studien zu Dwecks Leistungsmotivationsmodell wurden dabei meist nicht konkrete Emotionen in Leistungssituationen in differenzierter Form erhoben, sondern relativ globale Einschätzungen einer positiven vs. negativen leistungsthematischen Grundstimmung. Diese inhaltlich und methodisch eingeschränkte Zugangsweise zum Phänomen der Emotionen nach Leistungsrückmeldungen versuchen die empirischen Untersuchungen, die im Folgenden geschildert werden, zu überwinden. Emotionen sollen dabei zudem nicht nur als bloßer „Outcome“ kognitiv-motivationaler Größen betrachtet werden, sondern auch als Einflussgröße auf Lernverhalten und Leistung Beachtung finden.

Zweiter Teil: Empirische Untersuchungen

4. Theoretisches Modell, Hypothesen und methodischer Zugang

Die vorliegende Arbeit will die kausale Bedeutung Impliziter Intelligenztheorien im Leistungskontext eingehender als bislang geschehen untersuchen, indem sie neben Attribution und Leistungsverhalten (wie bei Hong et al., 1999) gleichzeitig auch Emotionen (wie bei Niiya et al., 2004) und objektive Leistungsmaße (wie in den Interventionsstudien und bei Cury et al. 2006) miteinbezieht. Sie will damit Robins und Pals' (2002) korrelativ gewonnene Pfadzusammenhänge für den experimentellen Fall validieren. Im Zuge dessen möchte sie den Zusammenhang zwischen Impliziten Intelligenztheorien und Emotionen nach Misserfolg differenziert analysieren, indem sie insbesondere selbstrelevante Emotionen wie Schuld und Scham untersucht. Auch soll auf diesem Wege besser verstanden werden, welche Bedeutung die Leistungssituation (individualistisch vs. kollektivistisch), die neben der Impliziten Theorie der zweite experimentell variierte Faktor ist, für diese Vorgänge hat.

Implizite Intelligenztheorien werden dabei – ähnlich wie in den Studien, die die Texte von Bergen (1991) verwenden (vgl. Abschnitt 3.2.1) – über Zeitschriftenartikel experimentell induziert. Da diese Artikel u.a. neurowissenschaftlich argumentieren, kann auch die Frage nach den Konsequenzen der Ergebnisse der Hirnforschung auf das Leistungsverhalten näher untersucht werden.

Es werden Reaktionen auf Misserfolg und auf Erfolg erhoben. Die Messung beider erfolgt bei den Untersuchungen auf unterschiedliche Weise. Studie 1 verwendet Fallgeschichten und befragt die Probanden nach ihrem hypothetischen Verhalten und Erleben in der geschilderten Leistungssituation. Die Studien 2 und 3 konfrontieren die Teilnehmer mit konkreten Leistungsaufgaben, hinsichtlich derer ihnen Erfolg oder Misserfolg rückgemeldet wird. Die Erhebung der abhängigen Maße erfolgt hier teilweise wieder über den Selbstbericht, teilweise aber auch über verhaltensnähere Merkmale, deren Messung mit einigen Täuschungen über den Untersuchungszweck einhergeht. Um den Faktor Leistungssituation manipulierbar zu machen und die Leistungssituation emotional anregender und realitätsnäher zu gestalten, werden in den Studien 2 und 3 Gruppenversuche (mit jeweils zwei Probanden) durchgeführt.

In diesem Kapitel werde ich auf diejenigen Bestandteile meiner empirischen Studien eingehen, die allen gemeinsam sind. Hierzu gehören die Fragestellungen und Hypothesen (Abschnitt 4.1), die verwendeten Zeitschriftenartikel zur Induktion Impliziter Intelligenztheorien (Abschnitt 4.2) sowie die Erhebung der Leistungsemotionen (Abschnitt 4.3).

4.1 Fragestellungen und Hypothesen

4.1.1 Hauptfrage: Der Einfluss Impliziter Theorien auf Emotion und Verhalten

Im Anschluss an die im theoretischen Teil dieser Arbeit angestellten Überlegungen will ich ein Wirkungsmodell postulieren, das ich nachfolgend empirisch testen werde. Dieses Modell ist in Abbildung 3 veranschaulicht. Danach ist zunächst einmal konzeptuell zwischen Prozessen vor (ggf. auch während) der Durchführung einer Leistungsaufgabe und solchen nach einer negativen Leistungsrückmeldung zu unterscheiden (vgl. die gestrichelte Linie in Abbildung 3).

Vor Inangriffnahme der Leistungsaufgabe sollten es vor allem zwei Faktoren sein, die für die Vorhersage des Verhaltens nach Misserfolg eine Rolle spielen: die Implizite Intelligenztheorie (statisch vs. dynamisch) einerseits und die Leistungssituation (kollektivistisch vs. individualistisch) andererseits (vgl. Abschnitt 3.3.4). Beide Faktoren sind ihrerseits abhängig von kulturellen und sozialen Einflüssen, die in Abbildung 3 als „Leistungskultur“ zusammengefasst werden. Die Leistungskultur umfasst einerseits eher „distale“ Prozesse der Sozialisation und des sozialen Lernens, die sich im Falle der Impliziten Intelligenztheorie in einer relativ stabilen Disposition zur Bevorzugung der einen oder der anderen Auffassung niederschlägt, im Falle der Leistungssituation in ein entweder unabhängiges oder interdependentes Selbstkonzept. Andererseits beinhaltet die Leistungskultur aber auch „proximale“ Prozesse der situationsspezifischen Anregung, welche sich im Falle der Impliziten Intelligenztheorie z.B. über das Lesen eines für die eine oder die andere Auffassung Partei ergreifenden Zeitschriftenartikels vollziehen könnte, im Falle der Leistungssituation z.B. über die Teilnahme an einer Leistungsaufgabe mit oder ohne Gruppencharakter.

Die Auswirkungen jener beiden Hauptfaktoren werden in Abbildung 3 aus Gründen der Übersichtlichkeit immer nur anhand *eines* Extrempols veranschaulicht. Die statische Intelligenztheorie sollte *einerseits* zum Einnehmen eines hilflosen Attributionsmusters nach Misserfolgsrückmeldung führen. In Anschluss an Hong et al. (1999) ist hier vor allem zu erwarten, dass Statiker weniger Anstrengungsattributionen zeigen als Dynamiker. Hinsichtlich Fähigkeitsattributionen sind Unterschiede nicht zwingend zu erwarten, wiewohl tendenziell eher mit einer stärkeren Ausprägung bei den Statikern zu rechnen ist (vgl. Abschnitt 3.1). *Andererseits* bestimmt die Intelligenztheorie die Zielorientierung: Statiker sollten eine ausgeprägtere Leistungszielorientierung und eine weniger ausgeprägte Lernzielorientierung aufweisen als Dynamiker. Die hilflosen Attributionen sowie die stärkere Leistungszielorientierung der Statiker sollten beide zu einem Gefühl der Hoffnungslosigkeit (Depression, Verzweiflung) beitragen und das

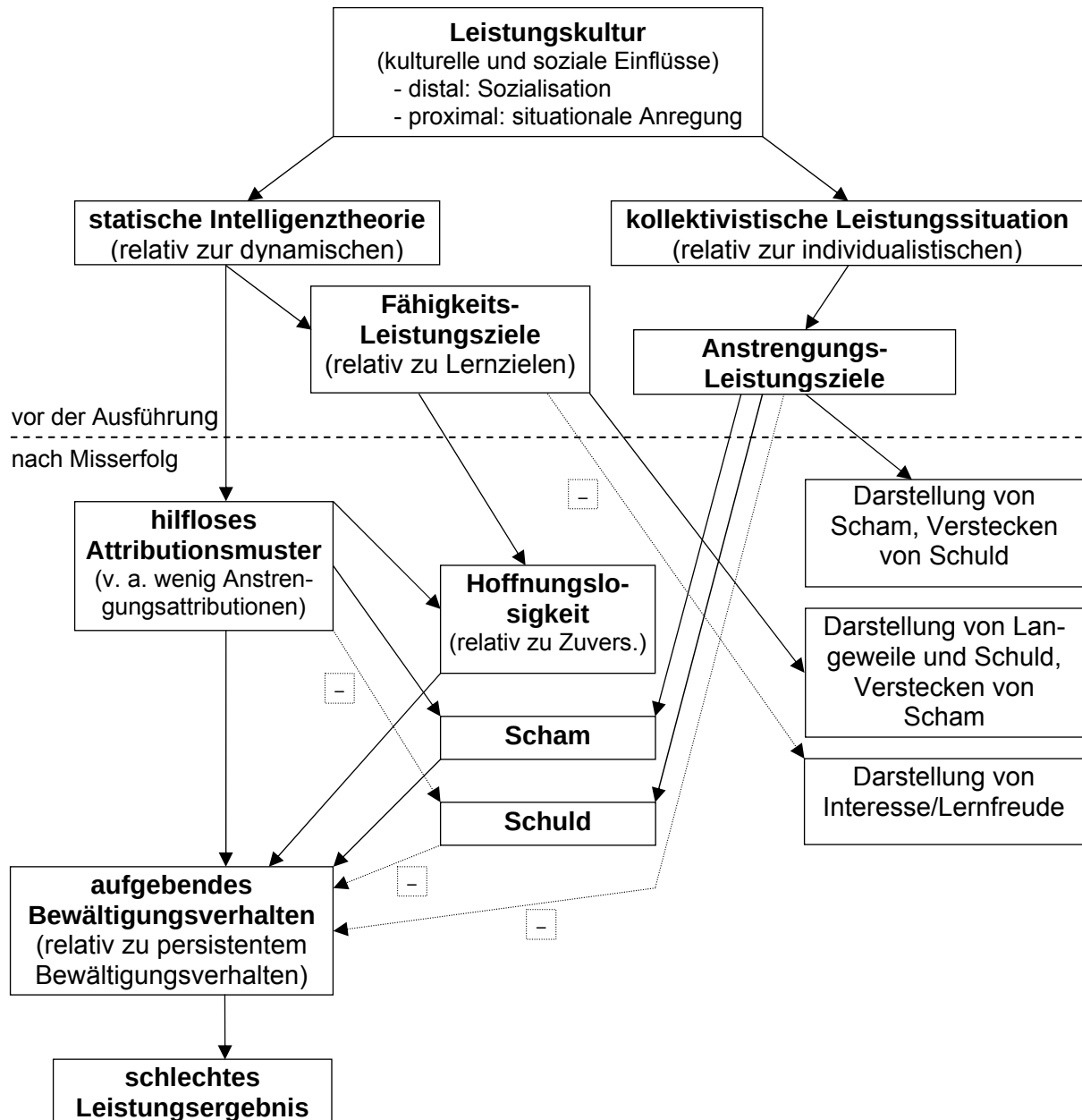


Abbildung 3. Theoretisches Modell der Beziehung zwischen Impliziten Theorien, Zielorientierung, Attribution, Emotionen, Emotionsdarstellung, Bewältigungsverhalten und Leistung.

Gefühl der Zuversicht unwahrscheinlicher machen. Abbildung 3 stellt aus Gründen der Übersichtlichkeit Hoffnungslosigkeit und Zuversicht als zwei Pole desselben Konstrukts dar. Da Hoffnungslosigkeit jedoch anders als Zuversicht nicht nur prospektive, sondern auch retrospektive Anteile aufweist, werden beide als unabhängig konzipiert (vgl. Abschnitt 3.3.0). Nähere Erläuterungen zu den Emotionskonstrukten sind Abschnitt 4.3 zu entnehmen.

Hilflose Attributionen sollten nun sowohl auf direktem Wege als auch vermittelt über das Gefühl der Hoffnungslosigkeit ein aufgebendes („hilfloses“) Bewältigungsverhalten wahrscheinlicher machen. Entsprechend sollte umgekehrt ein anstrengungsorientiertes Attributionsmuster sowohl direkt zu persistentem

(„meisterndem“) Bewältigungsverhalten führen, als auch indirekt über das Gefühl der Zuversicht. In ähnlicher Form sollten hilflose Attributionen zu mehr Scham und weniger Schuld führen als anstrengungsorientierte Attributionen, wobei sich Scham und Schuld – wiederum den Attributionseffekt verstärkend – entsprechend auf das Bewältigungsverhalten auswirken: Scham sollte eher zu hilflosem Verhalten führen, Schuld hingegen eher zu meisterndem (vgl. die gepunktete Linie in Abbildung 3, die einen negativen Zusammenhang verdeutlichen soll). Hoffnungslosigkeit, Zuversicht, Scham und Schuld sollten (mindestens partielle) Mediatoren der Beziehung zwischen Attribution und Bewältigungsverhalten sein, wobei Hoffnungslosigkeit und Scham die abträglichen Auswirkungen hilfloser Attributionen verstärken sollten, Zuversicht und Schuld hingegen die förderlichen Auswirkungen anstrengungsorientierter Attributionen. Damit wird also unter Bezug auf die Ergebnisse von Robins und Pals (2002, vgl. Abschnitt 3.3.4) postuliert, dass Emotionen nicht bloß Resultat einer Konfiguration situationaler Größen sind, sondern auch ihrerseits das Handeln in der Leistungssituation bestimmen. Sie entstehen zwar aus Attributionen heraus, sind aber nicht mit diesen identisch, sondern bringen deren Wirkung durch die ihnen implizite hohe subjektive Bedeutsamkeit (Pekrun, 2000) erst richtig zur Geltung.

Das Vorhandensein einer kollektivistischen Leistungssituation sollte sogenannte Anstrengungs-Leistungsziele nach sich ziehen, d.h. das Bedürfnis, sich vor anderen als fleißig und bemüht darzustellen. Diese durch sozialen Druck gekennzeichnete Leistungssituation sollte außerdem die Tendenz zum Empfinden der selbstrelevanten Emotionen Scham und Schuld sowie die Neigung zu meisterndem Bewältigungsverhalten erhöhen (vgl. Abschnitt 3.3.4). Im Bewältigungsverhalten – d.h. der verhaltensmäßigen Reaktion auf einen Misserfolg, die sich eher in resignativem „Nichtstun“ oder aber in zukunftsorientiertem Lernhandeln ausdrücken kann – laufen somit die Auswirkungen einer ganzen Reihe von Prozessen zusammen. Zusammengenommen sollten diese Prozesse wiederum vorhersagen, wie gut die Person in zukünftigen Leistungsaufgaben abschneidet: Wer beim Lernen schnell aufgibt, sollte langfristig auch schlechtere Leistungen erzielen (vgl. Abschnitt 3.2.2). Diese Hypothese zur objektiven Lernleistung konnte aus methodischen Gründen lediglich in Studie 2 geprüft werden.

Die kollektivistische Leistungssituation sollte außerdem mit der Darstellung von Scham sowie dem Verstecken von Schuld verbunden sein, während die statische Intelligenztheorie, vermittelt über (Fähigkeits-)Leistungsziele, zur Darstellung von Langeweile und Schuld, zum Verstecken von Scham sowie zu geringer Bereitschaft, Lernfreude oder Interesse darzustellen, animieren sollte (vgl. Abschnitt 3.3.5 für eine theoretische Ableitung).

Diese Modellvorstellungen lassen sich nun in konkrete Forschungshypothesen umwandeln. Zu berücksichtigen ist hier, dass in meinen Untersuchungen die Intelligenztheorie sowie die Leistungssituation experimentell manipuliert wer-

den, von der Messung der Zielorientierung aber aus methodischen und inhaltlichen Gründen abgesehen wird. Zielorientierung spielt in Dwecks neueren Arbeiten eine eher untergeordnete oder sogar gar keine Rolle mehr (vgl. Meyer, 2000). Dies dürfte nicht zuletzt daran liegen, dass Zielorientierungen im Vergleich zu Impliziten Theorien die deutlich diffuseren Konstrukte sind und somit erst recht schwieriger experimentell induziert werden können. Allgemein dürfte die Klarheit und einfache Operationalisierbarkeit des Konstrukts der Impliziten Theorie, die auch mit einer vergleichsweise großen Alltagsnähe und praktischen Relevanz verbunden sind (vgl. Abschnitt 2.2.2), wesentlich zu seinem wachsenden Erfolg in der Pädagogischen Psychologie beitragen. Den Zielorientierungen kommt im Rahmen der vorliegenden Untersuchungen nur der Status hypothetischer Erklärungskonstrukte zu. Ähnliches gilt für die Attributionsmerkmale, die zwar empirisch berücksichtigt werden; da diese Berücksichtigung aber mit geringem Erhebungsaufwand erfolgt und umgekehrt die induzierte Implizite Theorie große empirische und theoretische Aufmerksamkeit erhält, werden die wichtigsten Hypothesen ohne eine Mitberücksichtigung der Attribution formuliert, wie sie aus Abbildung 3 abzuleiten gewesen wäre. Insoweit also Zielorientierung und Attribution als vermittelnde Variablen ausgeklammert werden, lassen sich die folgenden Hypothesen aufstellen:

Statiker sollten zu weniger Anstrengungsattributionen, mehr Hoffnungslosigkeit und Scham, weniger Zuversicht und Schuld, mehr aufgebendem Bewältigungsverhalten, schlechteren langfristigen Leistungen, mehr Langeweile- und Schuld-darstellung, mehr Schamverstecken sowie weniger Darstellung von Interesse neigen als Dynamiker. Kollektivisten sollten zu mehr Scham und Schuld, weniger aufgebendem Bewältigungsverhalten sowie mehr Schamdarstellung und Schuldverstecken neigen als Individualisten. Hoffnungslosigkeit, Zuversicht, Scham und Schuld sollten die Beziehung zwischen Impliziter Theorie und Bewältigungsverhalten mindestens teilweise vermitteln.

4.1.2 Nebenfragestellungen

Der Untersuchung von *Reaktionen auf positive Leistungsrückmeldungen* kommt in den vorliegenden empirischen Studien eine eher untergeordnete, explorative Bedeutung zu. Gemäß oben angestellten (vgl. u.a. die Abschnitte 3.3.2 und 3.3.5) bzw. analog zum Misserfolgsfall vorgehenden Überlegungen ist damit zu rechnen, dass Statiker weniger zu Anstrengungsattributionen und Stolz, mehr zu Hybris und Erleichterung, mehr zu meisterndem Bewältigungsverhalten (vgl. Hong et al., 1999), mehr zur Darstellung von Hybris sowie mehr zum Verstecken von Freude neigen als Dynamiker. Des Weiteren sollten Kollektivisten mehr zu selbstrelevanten Emotionen (Hybris, Stolz), zur Darstellung von Stolz und zum Verstecken von Hybris neigen als Individualisten.

Wie in Abschnitt 2 bereits angedeutet, versuchen die vorliegenden Untersuchungen auch Licht auf die medien- und kulturpsychologische Frage zu werfen, welche Konsequenzen die Konfrontation mit neurowissenschaftlichen Erkenntnissen für das Leistungsverhalten haben kann. Diese Frage wird hier dahingehend spezifiziert, ob Implizite Intelligenztheorien besonders gut mit neurowissenschaftlichem Vokabular zu induzieren sind (bzw. ob sich die in Abhängigkeit neurowissenschaftlicher Diskurse zu- oder abträglichen Konsequenzen der Induktion, wie sie in den in Abschnitt 2.2.2 geschilderten Studien gefunden worden sind, replizieren lassen, s.u.). In Studie 1 wird daher neben der Impliziten Theorie und der Leistungssituation noch ein dritter Faktor variiert: der *Leistungsdiskurs*. In einer Bedingung werden die Probanden mit einem Artikel über psychologische Untersuchungen an Zwillingen konfrontiert, in einer anderen mit einem Artikel über neurowissenschaftliche Untersuchungen. Der neurowissenschaftliche Leistungsdiskurs sollte die Impliziten Intelligenztheorien glaubwürdiger und nachhaltiger induzieren und somit die postulierten Zusammenhänge *deutlicher* hervorbringen als der psychologische (vgl. Abschnitt 2.2 sowie zu den Zeitschriftenartikeln Abschnitt 4.2). In den Studien 2 und 3 musste aus ökonomischen Gründen auf den Faktor Leistungsdiskurs verzichtet werden. In diesen Studien kann es daher nur darum gehen, *ob* der neurowissenschaftliche Leistungsdiskurs dazu in der Lage ist, Implizite Intelligenztheorien zu beeinflussen. Von besonderem Interesse ist dabei, ob dies für die statische Variante jenes Diskurses zutrifft, da anders als zur dynamischen hierzu bislang kaum Untersuchungen vorliegen (vgl. Abschnitt 2.2.2).

Schließlich liegt das Interesse auch auf *geschlechtsspezifischen Reaktionsweisen* (auf Misserfolg). Frühere Studien ergaben, dass Frauen sich in ihrem Verhalten stärker von induzierten Intelligenztheorien beeinflussen lassen als Männer (vgl. Abschnitt 3.2.2). Dieser Frage soll hier in zumindest explorativer Art und Weise nachgegangen werden.

Auch das *Zusammenspiel von induzierten und dispositionalen Impliziten Theorien* soll mit den vorliegenden Untersuchungen (Studien 2 und 3) näher analysiert werden, indem sie – was bislang außer bei Niiya et al. (2004) kaum geschah – beides in demselben Design erheben.

4.2 Erstellung der Zeitschriftenartikel zur Induktion Impliziter Theorien

Zur Induktion der Impliziten Intelligenztheorien wurden insgesamt vier Zeitschriftenartikel erstellt. Alle vier Artikel tragen den Titel „Die Wurzeln der Geisteskraft“ und behandeln neue Untersuchungen über die Ursachen der Intelligenz. Zwei Artikel lokalisieren diese Ursachen in den Erbanlagen und/oder in frühen Kindheitserlebnissen (statische Bedingung), während die anderen beiden Intelligenz als erfahrungsabhängig darstellen (dynamische Bedingung). Ferner unter-

scheiden sich die Artikel hinsichtlich der Angaben darüber, welche Art von Untersuchungen durchgeführt wurde. Zwei Artikel berichten von Zwillingstudien (psychologische Bedingung), die anderen beiden von Studien mit Methoden der Hirnforschung (neurowissenschaftliche Bedingung). Die vier Artikel resultieren also aus der Kreuzung der Faktoren Implizite Theorie (statisch vs. dynamisch) sowie Leistungsdiskurs (psychologisch vs. neurowissenschaftlich).

4.2.1 Inhaltliche Gestaltung

Die Artikel sind zwei Din-A-4-Seiten lang und entsprechen damit im Umfang denen von Bergen (1991, vgl. Abschnitt 3.2.1). Auch die narrativ gestaltete Rahmengeschichte wurde von Bergen übernommen. Sie handelt von einem hochbegabten Kleinkind, dessen überragende Intelligenz anhand der daraufhin erörterten Untersuchungsergebnisse erklärt werden soll. Die Wahl dieser Thematik als Rahmenstory ist übrigens durchaus realitätsnah (vgl. etwa *Focus* 47/03 und Tabelle 1).

Die beiden Artikel der psychologischen Bedingung entsprechen im Wesentlichen denen von Bergen (1991). Sie wurden ins Deutsche übersetzt und leicht modifiziert. Vor allem wurde auf die Thematik der Bedeutsamkeit der Intelligenz verzichtet, die Bergen zusätzlich variiert hatte. Während bei Bergen im zweiten Artikelteil Intelligenz einmal als hoch bedeutsam für den Erfolg in allen Lebensbereichen dargestellt wurde, das andere Mal hingegen als völlig unbedeutend, machen die hier verwendeten Artikel zu dieser Frage keine Aussage. Ob andere Studien wie Hong et al. (1999) die Bergenschen Artikel mit oder ohne Bedeutsamkeitsaussage benutzt haben, konnte leider von den betreffenden Forschern nicht in Erfahrung gebracht werden (vgl. Abschnitt 3.2.1). Zwei Möglichkeiten erscheinen theoretisch plausibel: entweder dass die beiden Artikel mit hoher Bedeutsamkeit von Intelligenz übernommen wurden (da diese sich in Bergens Studie als wirksamer herausgestellt hatten), oder aber dass zwei Artikel, die keine zusätzlichen Aussagen über die Bedeutsamkeit von Intelligenz machen, neu erstellt wurden. Da in Bergens Arbeit jedoch die statisch-bedeutsame sowie die dynamisch-unbedeutsame Version beispielhaft abgedruckt sind, erscheint diese „pragmatische“ Textwahl (wohl nicht zuletzt weil die angeschriebenen Forscher keine Auskunft geben wollten) wohl als die wahrscheinlichere (vgl. Abschnitt 8.2 für eine kritische Diskussion).

In dieser Hinsicht muss die Frage der Vergleichbarkeit zwischen den hier verwendeten Artikel und denen anderer Studien also leider offen bleiben. Die vorliegende Untersuchung geht mit Bergens Artikel auf jeden Fall eher „konservativ“ um, d.h. versucht, den Effekt nur über die Manipulation der Impliziten Intelligenztheorie nachzuweisen, ohne dabei auch auf die Bedeutsamkeitsfrage zurückzugreifen.

Die beiden Artikel der neurowissenschaftlichen Bedingung beschreiben eine Studie mit bildgebenden Verfahren (PET), der es um die Lokalisation der Intelligenz im Gehirn geht. Vom Ansatz her ähneln sie damit dem Spiegel-Artikel „Gral der Geisteskraft“ (*Der Spiegel* 30/2000), welcher allerdings die Studie eher kritisch diskutiert und die Lokalisationsfrage offenlässt. Die hier verwendeten Artikel berichten dagegen von eindeutigen Befunden entweder in Richtung einer Lokalisation der Intelligenz im Gehirn oder in Richtung eines Beweises der Unmöglichkeit einer solchen Lokalisation. Der daraufhin vorgenommene Schluss von der Lokalisationsfrage zur Stabilitätsfrage (d.h. lokalisierte Intelligenz = stabile Intelligenz und umgekehrt) wird im Artikel als selbstverständlich hingestellt und nicht weiter hinterfragt. Die neurowissenschaftlichen Fachbegriffe sowie Hirn-Scans werden mehr oder weniger plausibel eingesetzt, ohne den Anspruch zu haben, der wissenschaftlichen Realität möglichst nahe zu kommen. Einige prägnante schlussfolgernde Sätze sind direkt aus den in Abschnitt 2.2.3 vorgestellten realen Zeitschriftenartikeln übernommen worden. Zusätzlich beziehen sich die beiden Texte explizit auf das sogenannte „Manifest“ der Hirnforscher (vgl. Abschnitt 2.0) und zitieren wortgetreu daraus. In den Begrifflichkeiten von Kapitel 2 wird in dem einen Artikel ein stabil-epiphänomenalistischer Diskurs realisiert, in dem anderen ein veränderbar-geisteswissenschaftlicher.

Tabelle 3 illustriert anhand einiger Textausschnitte, hinsichtlich welcher Eigenschaften sich die vier Artikel unterscheiden.

4.2.2 Formale Gestaltung

Die Artikel wurden hinsichtlich Stil und Layout so gut wie möglich der Zeitschrift „Der Spiegel“ nachempfunden und den Probanden auch als „Spiegel-Artikel“ vorgestellt. Von den insgesamt ca. 280 Probanden, die einen der Texte erhielten, äußerte kein einziger von sich aus Zweifel an deren Echtheit.

Die Bilder und Bildunterschriften der Artikel der neurowissenschaftlichen Bedingung wurden teilweise direkt aus dem Artikel „Gral der Geisteskraft“ (*Der Spiegel* 30/2000) übernommen. In der Fußzeile wurde „Spiegel 20/06“ vermerkt, womit die Aktualität des Artikels unterstrichen werden sollte. Die gedruckten Vorlagen der Artikel wurden schließlich in eine Zeitschrift gelegt und so kopiert, dass auch die schwarzen Randbereiche der kopierten Blätter so aussahen, als seien sie tatsächlich beim Vervielfältigen einer Zeitschriftenseite entstanden.

Tabelle 3

Gegenüberstellung der vier Artikel über Intelligenzforschung, die zur Induktion Impliziter Intelligenztheorien verwendet wurden, in Textausschnitten

neurowissenschaftlich		psychologisch	
statisch	dynamisch	statisch	dynamisch
Titel: Die Wurzeln der Geisteskraft			
Untertitel: Neue Erkenntnisse der Neurowissenschaften zeigen die Grundlagen der Intelligenz im Gehirn auf – und entlarven die Idee der Bildungsgerechtigkeit als soziale Konstruktion.	Untertitel: Neue Erkenntnisse der Neurowissenschaften entlarven das Konzept einer allgemeinen und stabilen Intelligenz als soziale Konstruktion.	Untertitel: Neue Erkenntnisse der Zwillingsforschung zeigen die genetischen Grundlagen der Intelligenz auf – und lassen Zweifel aufkommen an der Idee der Bildungsgerechtigkeit.	Untertitel: Neue Erkenntnisse der Zwillingsforschung entlarven das Konzept einer genetisch determinierten Intelligenz als Ideologie.
Adam Belschner ist hochbegabt. Obwohl er erst 18 Monate alt ist, kann er über 2000 Wörter verstehen, 500 Wörter sprechen und fünf verschiedene Vogelarten unterscheiden. (...) In den letzten Jahren wurde in Europa und den USA eine Reihe von bahnbrechenden Studien veröffentlicht, die die bislang klarsten Antworten in der Anlage-Umwelt-Debatte liefern. Die wohl bedeutendste dieser Studien wird diesen Herbst in der amerikanischen Fachzeitschrift „Psychological Review“ publiziert.			
Knowles führte als erster Forscher systematische Längsschnittstudien mit der Positronen-Emissions-Tomografie (PET) durch. Mit PET ist es möglich, ein umfassendes und detailliertes Bild des Gehirns zu erhalten. Dazu injiziert man den Versuchspersonen radioaktive Moleküle; diese senden aus dem Körper heraus Strahlen, aus denen schließlich ein Computer ein Bild errechnet, das die Durchblutung einzelner Gehirnteile zeigt.		Knowles verfolgte über viele Jahre hinweg eineiige Zwillinge, die voneinander getrennt aufwuchsen. In einer unermüdlichen Suche auf der ganzen Welt konnte er 83 solcher Zwillingspaare ausfindig machen. Sie waren zwischen 7 und 51 Jahre alt und stammten aus allen sozialen Schichten.	

(Fortsetzung siehe nächste Seite)

Fortsetzung Tabelle 3

...dass Intelligenz eine sehr starke genetische Komponente hat. (...) „Die funktionale Struktur des Gehirns ist bereits im Alter von drei Jahren relativ festgelegt. Vor allem die Bereiche des linken und rechten Stirnlappens der Großhirnrinde, die für intellektuelle Leistungen verantwortlich sind, sind dann schon am Ende der Entwicklung.“ Diese Hirnregionen seien das physische Substrat dessen, was Psychologen „Intelligenz“ nennen: eine allgemeine, zeitstabile Fähigkeit zu geistigen Leistungen. Knowles ist also gelungen, was Psychologen sich erträumen, seitdem sie den Begriff „Intelligenz“ in die Welt gesetzt haben: Endlich ist Intelligenz nicht mehr nur ein Konstrukt, ein geisteswissenschaftlicher Begriff ohne klare Bedeutung, endlich kann man Intelligenz im wahrsten Sinne des Wortes „sehen“, kann sie mit naturwissenschaftlichen Methoden erforschen.	...dass Intelligenz eine nur sehr geringe genetische Komponente aufweist. (...) „Das Gehirn ändert fortwährend seine funktionale Struktur. Früher dachte man, dass im Bereich des Stirnlappens das Fundament liege für eine allgemeine geistige Leistungsfähigkeit, also das, was Psychologen ‚Intelligenz‘ nennen. Unsere Ergebnisse haben gezeigt, dass es eine solche Hirnstruktur nicht gibt – weder im Stirnlappen noch anderswo.“ Unser Gehirn ist ein Spiegel unserer bisherigen Erfahrung. Die einzelnen Neuronenverbände sind für ganz spezifische Fähigkeiten zuständig. Eine übergeordnete, „allgemeine“ Intelligenz gibt es nicht. Sie ist, so Knowles, „ein Mythos, eine soziale Konstruktion, die uns in den letzten Jahrzehnten in Fleisch und Blut übergegangen ist.“	...dass Intelligenz eine sehr starke genetische Komponente hat. Außerdem scheint die Umwelt eine wichtige Rolle während der ersten drei Lebensjahre zu spielen. Danach jedoch scheinen Umweltfaktoren (außer im Falle einer Gehirnschädigung) fast keinen Einfluß mehr auf die Intelligenz zu haben. (...) Es ergab sich, dass in verschiedenen Umgebungen aufgewachsene Zwillinge sehr ähnliche Intelligenzwerte aufwiesen. Waren die Zwillinge seit ihrer Geburt getrennt, unterschieden sich ihre Werte selten um mehr als zehn Punkte. Wurden sie erst nach dem dritten Lebensjahr getrennt, waren ihre Intelligenzwerte nahezu identisch. Gemäß Knowles' Berechnungen sind bis zu 78 Prozent der Intelligenz einer Person das Resultat genetischer Faktoren. (...)	...dass Intelligenz eine nur sehr geringe genetische Komponente hat. Menschen werden zwar mit einer bestimmten Menge an Intelligenz geboren; diese ist aber eher unbedeutend gegenüber der Höhe der möglichen Steigerung im Laufe des Lebens. So ist unter entsprechend anregenden Umweltbedingungen eine Steigerung von 50 Punkten keine Seltenheit. (...) Es ergab sich, dass in verschiedenen Umgebungen aufgewachsene Zwillinge oft deutlich abweichende Intelligenzwerte aufwiesen. Gemäß Knowles' Berechnungen sind bis zu 78 Prozent der Intelligenz einer Person das Resultat von Umweltfaktoren. Diejenigen Personen in seiner Stichprobe, die in stimulierenden Umgebungen mit engagierten Eltern aufgewachsen waren, hatten eher hohe Intelligenzwerte. (...)
--	--	---	--

(Fortsetzung siehe nächste Seite)

Fortsetzung Tabelle 3

neurowissenschaftlich		psychologisch	
statisch	dynamisch	statisch	dynamisch
Paula Rescorla und ihre Mitarbeiter an der Uni Freiburg sind begeistert über Knowles' Ergebnisse. „Wenn ein Kind mit drei Jahren in seiner intellektuellen Entwicklung weitgehend festgelegt ist, macht es keinen Sinn, alle Kinder so zu behandeln, als hätten sie gleiche Entwicklungschancen.“ Um die Erziehung zu optimieren, müsse man die Kinder früh testen, Begabungen identifizieren und entsprechend der Begabungsspektren früh kanalisieren. Freimachen müsse man sich von der Illusion, dass alle gleich sind und dass aus allen das Gleiche werden kann. „Diese Annahme ist unsinnig und widerspricht elementaren biologischen Gesetzen.“ (...)	Paula Rescorla und ihre Mitarbeiter an der Uni Freiburg sind begeistert über Knowles' Ergebnisse. „Von diesen Studien können wir lernen, dass Intelligenz etwas ist, das motivierte Menschen aneignen können. Diese Einsicht wird vermutlich die Pädagogik des 21. Jahrhunderts prägen.“ Kindern könnten nun, ungeachtet ihrer gegenwärtigen Intelligenz, gezielt Lernumgebungen angeboten werden, die ihnen helfen, ihre Fähigkeiten weiterzuentwickeln. Kinder anhand anfänglicher Erfolge oder Mißerfolge von vornherein auf eine bestimmte Fähigkeitsentwicklung festzulegen, sei hingegen „eine Strategie von gestern“. (...)	Paula Rescorla und ihre Mitarbeiter an der Uni Freiburg sind begeistert über Spoons und Knowles' Ergebnisse. „Es ist jetzt an der Zeit zu akzeptieren, dass Intelligenz eine genetisch determinierte Fähigkeit ist. Durch dieses Wissen kann unsere Gesellschaft auf vielfältige Weise profitieren. Man kann nun schnell diejenigen ausmachen, die von einer spezialisierten intellektuellen Beschäftigung profitieren würden.“ Wenn man den begabtesten Kindern ein entsprechendes Anspruchsvolles Curriculum anböte, könne man ihnen helfen, ihre angeborenen Fähigkeiten voll zur Geltung zu bringen. Dies käme der gesamten Gesellschaft zugute. „Weniger begabten Menschen können wir Umgebungen anbieten, die sie nicht frustrieren oder zu überfordern drohen, sondern sie gemäß ihrer Begabung angemessen zu fördern verstehen.“ (...)	Paula Rescorla und ihre Mitarbeiter an der Uni Freiburg sind begeistert über Spoons und Knowles' Ergebnisse. „Von diesen Studien können wir lernen, dass Intelligenz etwas ist, das motivierte Menschen aneignen können. Diese Einsicht wird vermutlich die Pädagogik des 21. Jahrhunderts prägen.“ Kindern könnten nun, ungeachtet ihrer gegenwärtigen Intelligenz, gezielt Lernumgebungen angeboten werden, die ihnen helfen, ihre Fähigkeiten weiterzuentwickeln. Kinder anhand anfänglicher Erfolge und Mißerfolge von vornherein auf eine bestimmte Fähigkeitsentwicklung festzulegen, sei hingegen „eine Strategie von gestern“. (...)
KATJA THIMM Infos im Internet: www.harvard.edu/jkn.html			

4.3 Auswahl und Erfassung der Emotionskonstrukte

Drei der wichtigsten Probleme, mit denen die Emotionspsychologie zu kämpfen hat, betreffen die Definition, Klassifikation und Messung ihrer Gegenstände (vgl. z.B. Meyer, Schützwohl, & Reisenzein, 2001). Das erstgenannte Problem – also die Frage nach der definitorischen Abgrenzung von Emotionen zu anderen psychischen Zuständen – spielt für die vorliegende Arbeit eine untergeordnete Rolle. Die bisherigen Ausführungen in Abschnitt 3.3 sollten deutlich gemacht haben, dass hier – wie auch im überwiegenden Teil der zitierten Literatur – eine „mentalistische“, d.h. auf das subjektive Erleben fokussierte Definition bevorzugt wird. Wichtiger erscheinen demgegenüber die letzteren beiden Probleme: Verschiedene Emotionen lassen sich erstens oft nicht leicht definitorisch voneinander abgrenzen, und selbst wenn dies gelungen ist, heißt dies zweitens noch lange nicht, dass man an sie anschließend auch in der Art und Weise empirisch herankommt, wie man es gemäß diesen Definitionen vorhatte. In vielen Studien werden diese Probleme kurzerhand dadurch gelöst, dass nur noch „positive“ und „negative“ Emotionalitäten unterschieden werden, weil dies die Missverständnisse zwischen Forscher, Beforschten und Forscherkollegen auf ein Minimum reduziert (vgl. Abschnitt 3.3). Die vorliegende empirische Arbeit will sich dieser Form des Reduktionismus nicht anschließen und ist bemüht, eine breite Palette von Emotionen zu untersuchen. Damit werden die angesprochenen Probleme der Klassifikation und Messung jedoch in besonderem Maße aktuell. Sie sollen in diesem Abschnitt erörtert werden, indem zunächst die theoriegeleitete Auswahl der Emotionen, später dann deren Erhebung thematisiert werden.

In den Abschnitten 3.3 und 4.1 wurde bereits herausgearbeitet, welche Emotionen (bzw. genauer: emotionale Qualitäten) für die vorliegende Fragestellung relevant sind: nämlich alle diejenigen, für die sich erstens gemäß theoretischen Überlegungen und empirischen Befunden Unterschiede in Abhängigkeit von Impliziten Intelligenztheorien erwarten lassen und die sich zweitens häufig nach Leistungsrückmeldungen zeigen. Dies ist im Misserfolgsfall sowohl für die Emotionen Hoffnungslosigkeit und Zuversicht der Fall, die aufgrund ihrer Zukunftsbezogenheit bzw. ihrer theoretischen Verwandtschaft zu Motivationskonstrukten von Impliziten Intelligenztheorien abhängen sollten, als auch für Scham und Schuld, bei denen selbiges aufgrund ihrer Abhängigkeit von bestimmten Attributionen gelten sollte (vgl. Abschnitt 3.3.2). Im Erfolgsfall sprechen wiederum attributionstheoretische Überlegungen für die Bedeutsamkeit von Hybris und Stolz. Unter eher explorativen Gesichtspunkten werden hier außerdem Erleichterung und Freude miterhoben. Erleichterung, die im Zuge eines nicht eingetretenen erwarteten Misserfolgs entsteht (vgl. z.B. Ortony, Clore, & Collins, 1988), könnte besonders unter Statikern verbreitet sein, da diese infolge ihrer

Fähigkeitsorientierung Misserfolg als bedrohlicher empfinden sollten (und ihn zudem vielleicht auch eher erwarten) als Dynamiker. Für Freude lässt sich keine Ergebnistendenz vermuten: Statiker könnten sich einerseits intensiver über Erfolge freuen, weil sie ihre Intelligenz langfristig bestätigt sehen, andererseits aber auch weniger intensiv, weil sie ihren Erfolg aufgrund ihrer „fraglos ausreichenden Intelligenz“ eher als selbstverständlich, als Pflichtaufgabe verstehen (müssen). Theoretisch sinnvoll wäre also, dass Statiker bei leichten Aufgaben weniger, bei schwierigen Aufgaben mehr Freuden empfinden dürften als Dynamiker.

Die resultierenden acht Emotionskonstrukte sind in Tabelle 4 aufgeführt und inhaltlich näher beschrieben. Weiter sind in Tabelle 4 die jeweiligen Attributionen aufgeführt, aus denen die Emotionen resultieren. Wie bereits in Abschnitt 3.3.2 erörtert, sieht sich die vorliegende Arbeit kognitiven Ansätzen der Emotionspsychologie verpflichtet, die ein Emotionskonstrukt nicht einfach anhand eines Emotionswortes oder dessen Synonymen definieren, sondern anhand ganz bestimmter kognitiver Voraussetzungen (vgl. Reisenzein, Meyer, & Schützwohl, 2003). Diese Überlegungen sollen nun hinsichtlich der in meinen Studien verwendeten Items konkretisiert werden.

Wie die allermeisten Studien im Bereich der Emotionspsychologie erfasst auch die vorliegende Arbeit Emotionen über den Selbstbericht (zu den Ausnahmen in den Studien 2 und 3, wo zusätzlich auf Beobachtungsmaße zurückgegriffen wird, vgl. die Abschnitte 6.2.1 und 7.2.1). Anders als diese konfrontiert sie die Probanden aber nicht einfach mit „nackten“ Emotionswörtern, sondern legt Wert auf eine *theoretische und situationale Einbettung der Items*. Die Formulierung der Items berücksichtigt also sowohl die theoretische Konzeption des Emotionskonstrukts (und vertraut nicht darauf, dass die Probanden diese in mehr oder weniger ausreichender Weise teilen) als auch den speziellen Kontext (hier: der Leistungsrückmeldung), in dem nach Emotionen gefragt wird.

Dieses Vorgehen sei an einem Beispiel verdeutlicht. Aus Tabelle 4, der auch die verwendeten Items zu entnehmen sind, geht hervor, dass die Emotion Hybris u.a. anhand des Items „Ich bin stolz auf meine Fähigkeiten und freue mich darüber, dass sie von anderen bemerkt werden“ gemessen wird. Weil nicht zu erwarten ist, dass die Aussage „Ich fühle Hybris“ auch nur annähernd theorieadäquat oder wenigstens konsistent verstanden wird, expliziert das Item gleich auch das Objekt der Emotion und berücksichtigt außerdem den sozialen Kontext der Rückmeldung. Dass das Item nun das Wörtchen „stolz“ verwendet, obwohl es gerade nicht das Konstrukt „Stolz“ messen soll, ist dabei völlig unproblematisch. Ähnlich verhält es sich beim Konstrukt „Schuld“, welches aus theoretischen Gründen sehr weit gefasst wird. Während Schuld in der Alltagssprache eine zutiefst soziale Emotion ist, soll hier auch Schuld gegenüber sich selbst definitiv eingeschlossen sein. Dies führt dazu, dass Schuld zuwei-

Tabelle 4

Überblick über die Emotionskonstrukte, ihre theoretische Verankerung und zugehörige Items

Emotion	Beschreibung	Attribution	Items
Hoffnungslosigkeit	Deprimiertheit, Resignation, Verzweiflung	Fähigkeit	1. Ich bin erstmal ziemlich deprimiert und niedergeschlagen. 2. Ich bin so resigniert, dass ich mich ganz kraftlos fühle. (Mw) 3. Ich fühle mich hoffnungslos. (Mw)
Zuversicht	Hoffnung, Optimismus <i>trotz</i> des Misserfolgs	Anstrengung	1. Ich verspüre ein Gefühl der Herausforderung und freue mich darauf, es das nächste Mal besser zu machen. 2. Ich habe die Hoffnung, dass ich künftig in solchen Aufgaben gut abschneiden werde. (m)
Scham	Beschämung, Verlegenheit, peinliche Berührtheit	Fähigkeit	1. Mein Ergebnis ist mir peinlich. (m) 2. Aus Scham über meine schlechte Leistung würde ich den anderen (Mitstudenten, Versuchsleiterin) am liebsten nie mehr begegnen. (m) 3. Ich schäme mich. (M)
Schuld	Bedauern, Selbstärger, schlechtes Gewissen	Anstrengung	1. Ich ärgere mich über mich selbst, weil ich mir sage, dass meine schlechte Leistung nicht hätte sein müssen. 2. Ich bin mir meiner Schuld an der Sache bewußt.
Hybris	Gefühl der Überlegenheit, Stolz auf die eigenen Fähigkeiten	Fähigkeit	1. Ich bin stolz auf meine Fähigkeiten und freue mich darüber, dass sie von anderen bemerkt werden. 2. Es ist ein gutes Gefühl, wenn die anderen sehen, was man drauf hat. 3. Ich bin stolz darauf, dass ich bei diesen Aufgaben besser als die meisten anderen bin. (m)
Stolz	Stolz auf eigenes Bemühen und den eigenen Fortschritt	Anstrengung	1. Ich bin stolz darauf, wie gut ich die Aufgaben gemeistert habe. (m) 2. Ich bin stolz auf mein Ergebnis und die Arbeit, die dahinter steckt.
Erleichterung	Entspannung aufgrund des nicht eingetretenen Misserfolgs	Fähigkeit (?)	1. Ich bin sehr erleichtert. (M) 2. Ich fühle mich befreit. (M)
Freude	innerliches und äußerliches Jubeln, Glück	--	1. Ich bin richtig glücklich. (M) 2. Ich strahle vor Freude. (M) 3. Vor Freude spüre ich mein Herz höher schlagen. (M)

Anmerkungen.

M: wörtlich übernommen von Molfenter (1999), Bereich „Emotionen nach der Prüfung“

Mw: wörtlich übernommen von Molfenter (1999), Bereich „Emotionen während der Prüfung“

m: übernommen von Molfenter (1999), inhaltlich leicht angepasst, Bereich „Emotionen nach der Prüfung“

Die Items wurden aus gleichnamigen Skalen entnommen, außer bei Zuversicht (heißt bei Molfenter Hoffnung) und Hybris (heißt bei Molfenter Stolz). Items ohne Kennzeichnung wurden selbst formuliert.

len mit weniger moralisch aufgeladenen Emotionswörtern erhoben wird, z.B. mit „Ich ärgere mich über mich selbst, weil ich mir sage, dass meine schlechte Leistung nicht hätte sein müssen“.

Tabelle 4 ist zu entnehmen, dass in einigen Fällen dennoch eine äußerst sparsame Itemformulierung gewählt wurde (z.B. „Ich schäme mich“). Dies hat seinen Grund darin, dass bei der Itemkonstruktion neben der theoretischen und situationalen Einbettung auch die *empirische Bewährung* der Items eine Rolle spielte, d.h. es sollten möglichst viele Items aus bereits bestehenden, empirisch konstruierten und validierten Skalen übernommen werden. Diesbezüglich liegt in deutscher Sprache bislang nur der „Test Emotions Questionnaire“ (TEQ) vor, der sich bei Molfenter (1999) ausführlich beschrieben findet und von Pekrun et al. (2004) in empirisch überarbeiteter, zweisprachig erweiterter Form besprochen wird. Der TEQ beschäftigt sich mit Emotionen vor, während und nach Prüfungen. Die Auswahl der untersuchten Emotionskonstrukte erfolgte u.a. anhand von Studien zu deren Auftretenshäufigkeit (Pekrun, 1998). Die Formulierungen wurden aus aufwändigen Interviews mit Studierenden extrahiert, womit der TEQ ein ähnliches Ziel („use of everyday language“, Pekrun et al., 2004, S. 298) verfolgt wie die oben angesprochene Strategie der situationalen Einbettung. Wie Tabelle 4 ausweist, wurde eine Reihe von Items wörtlich, andere in z.T. angepasster Form übernommen (z.B. wurde aus einer „Klausur“ eine „Aufgabe“, besonders häufig in den Studien 2 und 3). Die Items wurden mit einer Ausnahme (vgl. Tabelle 4) aus dem Item-Pool für Emotionen *nach* der Prüfung übernommen. Die Auswahl erfolgt primär nach inhaltlichen Kriterien, sekundär anhand psychometrischer Eigenschaften. Am Rande sei erwähnt, dass Molfenter bzw. Pekrun et al. die Items zu Emotionen vor, während und nach Prüfungen auf recht atheoretische Weise zu Skalen vereinigen – und dies obwohl gleichzeitig betont wird, dass sie nicht nur Trait-Emotionen, sondern auch State-Emotionen (um die es ja auch hier geht) erfassen sollen. Trotz solcher methodischer Schwächen erschien es für die vorliegende Untersuchung immer noch lohnend, einige TEQ-Items (insbesondere solche mit klarem inhaltlichen Bezug sowie guten psychometrischen Eigenschaften) zu übernehmen, da die empirische Konstruktion eigener Skalen für die vorliegende Arbeit aus ökonomischen Gründen nicht möglich war. Die Konstruktion der übrigen Items erfolgte allein auf der oben bereits erläuterten theoretischen Basis.

Die Items zur Erfassung der Emotionsdarstellung wurden ausnahmslos aufgrund theoretischer Überlegungen erstellt, da hier keinerlei empirische Vorarbeiten existieren. Die Items beginnen zumeist mit der Formulierung „Ich versuche, den anderen durch Worte und Gesten den Eindruck zu vermitteln, dass...“ bzw. „Ich versuche, bei den anderen den Eindruck zu vermeiden, dass...“, je nachdem, ob sie das gezielte Darstellen oder das gezielte Nicht-Darstellen (Verstecken, Verbergen) einer Emotion betreffen. Ansonsten entsprechen die Formulie-

rungen in inhaltlicher Hinsicht weitgehend jenen der Emotions-Items, so dass die Darstellung von Schuld z.B. durch die Zustimmung zu der Aussage „Ich versuche, den anderen durch Worte und Gesten den Eindruck zu vermitteln, dass ich mich selbst sehr über mich ärgere und mein Verhalten bereue“ erfasst wird. Nach negativer Leistungsrückmeldung wird die Darstellung von Zuversicht, Scham, Schuld und Langeweile sowie das Verstecken von Scham und Schuld erhoben, nach positiver Leistungsrückmeldung die Darstellung von Hybris, Stolz, Erleichterung und Freude sowie das Verstecken von Freude. All diese Konstrukte werden mit lediglich einem Item erfasst, mit Ausnahme der Darstellung von Hybris sowie dem Verstecken von Freude, die durch je zwei Items operationalisiert werden. Das Verstecken von Freude wird inhaltlich untergliedert in eine eher selbstwertbezügliche Komponente („...weil das so wirkt, als ob ich's nötig hätte“) und eine empathische Komponente („...weil ich andere, bei denen es vielleicht nicht so gut gelaufen ist, nicht noch mehr runterziehen möchte“). Der vermutete Unterschied in Abhängigkeit Impliziter Theorien sollte sich nur für die selbstwertbezügliche Komponente zeigen.

Die Emotions- und Emotionsdarstellungs-Items werden in Anschluss an Molterer (1999) anhand von fünfstufigen Skalen (stimmt gar nicht vs. stimmt genau) eingeschätzt (vgl. die Abweichung in Studie 3, Abschnitt 7.1). Auf den situationalen Kontext der Emotionserhebung – etwa ob die Emotionen mit Hilfe von Fallgeschichten oder aber anhand konkreter Leistungsaufgaben angeregt werden – wird in den folgenden Kapiteln studienspezifisch eingegangen.

Kritisch ließe sich an dieser Stelle noch einwenden, dass Emotion und Emotionsdarstellung mit vergleichsweise wenig Items erhoben werden. Hierfür lassen sich zwei Begründungen anführen. Zum einen führt die Fülle der in der vorliegenden Arbeit verfolgten Hypothesen zwangsläufig zu Einschränkungen hinsichtlich der Detailliertheit, mit der jede einzelne untersucht wird; die Auf- und Teilnahmebereitschaft der Probanden würde nämlich durch übermäßig viele Items untergraben. Jene geringe Detailliertheit wird bewusst in Kauf genommen, nicht zuletzt weil gerade die wechselseitigen Zusammenhänge der einbezogenen Konstrukte interessieren. Zweitens ist es aber infolge der erwähnten situationalen Einbettung der Items auch schwierig, eine größere Anzahl passender Itemformulierungen zu finden. Die Beschränkung auf wenige Items hat somit nicht nur ökonomische, sondern auch theoretische Gründe.

5. Studie 1: Prüfung des Modells mit Hilfe der Szenario-Methode

Zur empirischen Umsetzung der im vorigen Kapitel explizierten Hypothesen ist es erforderlich, eine Leistungsrückmeldung aktiv (möglichst experimentell kontrolliert) zu induzieren. Nur auf diese Weise kann nämlich weitestmöglich ausgeschlossen werden, dass die Reaktion auf die Leistungsrückmeldung direkt von der tatsächlich erbrachten Leistung abhängt (und mithin den Einfluss der Leistungsrückmeldung selbst, um den es ja primär geht, verwässert würde). Die Versuchsteilnehmer sollten also in eine Situation versetzt werden, die der Anregung von Emotionen nach Leistungsrückmeldungen und anderen Leistungsverhaltensweisen dienlich ist.

Hier bieten sich vor allem zwei Möglichkeiten an. Entweder man konfrontiert die Probanden mit einer konkreten Leistungsaufgabe (z.B. einem Wissenstest), meldet ihnen danach ein besonders gutes oder besonders schlechtes Ergebnis zurück und erfasst die daraufhin auftretenden Emotionen; oder man fordert die Teilnehmer auf, sich in eine hypothetische Leistungssituation gedanklich hineinzuversetzen und dann detailliert zu schildern, wie sie in einer solchen Situation mutmaßlich erleben und handeln würden. Robinson und Clore (2001) bezeichnen letzteres Vorgehen als *Szenario-Methode* und grenzen es von einer verwandten Strategie, der sogenannten Erinnerungs-Methode ab, bei denen der Proband bestimmte emotionale Episoden erinnern und sein damaliges Erleben schildern soll. Hier soll es nur um die Szenario-Methode gehen, wobei für die Erinnerungs-Methode ähnliche Schlussfolgerungen zulässig sein sollten wie die im Folgenden getätigten.

Wenngleich für die Wahl der Szenario-Methode (im Deutschen manchmal auch Vignetten-Methode genannt) häufig in erster Linie ökonomische Gründe genannt werden, gibt es auch inhaltliche Argumente, mit denen sie zuweilen offensiv gerechtfertigt wird. Eine zentrale Rolle spielt dabei das Konzept der externen Validität. Befürworter der Szenario-Methode (vgl. wiederum Robinson & Clore, 2001) würden für die vorliegende Fragestellung etwa geltend machen, dass es weit alltagsnäher sei, die Probanden mit einer vorgestellten Klausurrückgabe zu konfrontieren als mit einer realen Wissenstestrückmeldung im Labor. Diese Argumentation wird besonders häufig von Emotionspsychologen vorgenommen, vermutlich weil gerade Emotionen einen bedeutungsvollen, lebensnahen Kontext voraussetzen, der im Labor schwer herstellbar ist. In diesem Sinne sei es sinnvoller, bloß antizipierte Emotionen zu erheben als reale, aber in einem künstlichen Kontext entstandene Emotionen.

Nicht verwunderlich ist, dass gerade die Vertreter der kognitiven Emotionstheorien (vgl. Abschnitt 3.3.2) diese Argumente in Anschlag bringen und sich häufig der Szenario-Methode bedienen (z.B. Weiner, 1985, 1995; vgl. Laucken,

1989). In ihren Augen spielt es keine Rolle, ob man nun vorgestellte oder erlebte Emotionen untersucht, da in beiden Fällen dieselben Einschätzungsprozesse ablaufen (sollten). Während diese These der Äquivalenz von vorgestellten und erlebten Emotionen anfangs stillschweigend als zutreffend vorausgesetzt wurde (vgl. die Kritik bei Parkinson & Manstead, 1992), existieren mittlerweile einige Untersuchungen, die sie empirisch zu überprüfen versuchen (vgl. den Überblick bei Robinson & Clore, 2001). Zumeist wurde eine recht gute Übereinstimmung in den Ergebnissen von Szenario-Studien und Studien mit direkter Emotionsanregung festgestellt. Smith und Ellsworth (1987) fanden in ihrer Untersuchung von emotionalen Reaktionen nach Rückgabe einer realen Klausur überwiegend dieselben Emotions-Einschätzungs-Beziehungen wie in ihrer zuvor (1985) durchgeführten Szenario-Studie.

Ohne dass hier auf Einzelheiten eingegangen werden soll, lässt sich konstatieren, dass sich die Szenario-Methode für die Erfassung von Emotionen im Großen und Ganzen bewährt hat (vgl. Reisenzein, Meyer, & Schützwohl, 2003; Robinson & Clore, 2001). Sie soll daher auch im vorliegenden Rahmen zum Einsatz kommen. Dies hat nicht zuletzt auch ökonomische Gründe, da die Berücksichtigung aller experimentellen Faktoren ohne die zeitsparende Szenariotechnik kaum möglich wäre (vgl. Abschnitt 6.0).

5.1 Methode

Studie 1 berücksichtigt alle vier in Abschnitt 4.1 herausgearbeiteten experimentellen Faktoren. Die Faktoren Implizite Intelligenztheorie und Leistungsdiskurs werden *zwischen* den Versuchspersonen variiert, d.h. jede Person erhält einen der vier erstellten Zeitschriftenartikel (vgl. Abschnitt 4.2). Eine fünfte Gruppe erhält gar keinen Artikel und dient als Kontrollgruppe. Die Faktoren Leistungssituation und Leistungsrückmeldung werden – aus ökonomischen Gründen – *innerhalb* der Probanden variiert, d.h. jeder Teilnehmer erhält vier Szenarien. Diese unterscheiden sich erstens darin, ob eine kollektivistische oder individualistische Leistungssituation beschrieben wird, und zweitens, ob der Protagonist eine positive oder eine negative Leistungsrückmeldung erhält. Inhaltlich betrachtet stellt die Leistungsrückmeldung allerdings keinen eigenständigen Faktor dar, weil ja kein direkter Vergleich zwischen Misserfolgs- und Erfolgsbedingung angestrebt wird (und dieser aufgrund der größtenteils unterschiedlichen abhängigen Variablen auch empirisch kaum möglich ist). Somit ist es angemessener, von zwei Teilstudien mit je einem 2x2x2-Design (plus Kontrollgruppe) zu sprechen.

Szenarien. Die Erstellung der Szenarien wurde inhaltlich mit Bezug auf die Lebenswelt der studentischen Probanden vorgenommen. Die Geschichten mit individualistischer Leistungssituation berichten von drei Studenten, die ihre Statis-

tik-Klausuren zurückbekommen, für die sie zuvor zusammen gelernt haben. Der Protagonist erzielt dabei die deutlich schlechteste Note (Bedingung negative Leistungsrückmeldung bzw. Misserfolg) bzw. eine sehr gute Note (Bedingung positive Leistungsrückmeldung bzw. Erfolg). Die Geschichten mit kollektivistischer Leistungssituation beschreiben eine Gruppenarbeit im Rahmen eines Seminars, bei der jeder Student einen eigenen, klar definierten Beitrag leisten soll. Der Protagonist erhält nach ein paar Wochen von seinen Gruppenmitgliedern entweder eine positive oder eine negative Rückmeldung zum seinem Beitrag. Tabelle 5 zeigt beispielhaft die beiden Misserfolgs-Szenarien.

Die Szenarien wurden anhand von drei theoretischen Kriterien erstellt. *Ers-*stens sollte die enthaltene Rückmeldung als eindeutiger Misserfolg bzw. Erfolg interpretierbar sein. Eine gesonderte Manipulationskontrolle hinsichtlich der Frage, ob die Probanden die Fallgeschichten in der intendierten Weise verstanden haben, erschien somit nicht notwendig. Außerdem war es auf diese Weise möglich, für ein Szenario nicht alle Items vorzulegen, sondern nur die jeweils passenden (d.h. Misserfolg *oder* Erfolg betreffenden). Insgesamt entspringt das Kriterium der Eindeutigkeit der Rückmeldung also vor allem ökonomischen Motiven. *Zweitens* sollten die Kommunikationspartner des Protagonisten weder ein besonders freundschaftliches noch ein besonders feindschaftliches Verhältnis zu ihm pflegen. Somit bleibt gemäß Juvonen (2001; vgl. Abschnitt 3.3.5) offen, ob der Protagonist sich als bemüht oder als fähig darstellen sollte. Durch diese Ambivalenz sollte die Varianz im Darstellungsverhalten und mithin der Einfluss der Induktion Impliziter Theorien zunehmen. *Drittens* sollte die Situation auch hinsichtlich der Attribution ambivalent erscheinen, d.h. weder Anstrengungs- noch Fähigkeitsattributionen sollten einseitig angeregt werden. Vielmehr sollte das Szenario für beide Attributionsformen Hinweise geben. Wiederum stand die Erwartung dahinter, dass die Probanden je nach Induktion die ambivalenten Informationen unterschiedlich interpretieren.

Die Teilnehmer erhielten alle vier Szenarien nacheinander. Um Reihenfolgeeffekte zu minimieren, wurden sie in zwei unterschiedlichen Abfolgen präsentiert (jeweils beginnend mit den beiden Misserfolgs-Szenarien, weil diese von übergeordnetem theoretischem Interesse waren, und hier dann: indiv.-koll.-indiv.-koll. vs. koll.-indiv.-koll.-indiv.).

Abhängige Maße. Der Teilnehmer wird gebeten, sich in die Lage des jeweiligen Protagonisten des Szenarios zu versetzen und zu überlegen, wie er in dieser Situation reagieren würde. Zunächst wird ihm die Möglichkeit gegeben, in offener Beantwortung darüber Auskunft zu geben, welche Gefühle er an Stelle des Protagonisten empfinden würde. Dieses Item wurde vor allem aus Gründen der

Tabelle 5

Zwei der vier verwendeten Szenarien zur Anregung von Leistungsemotionen und -verhaltensweisen

Szenario Misserfolg – individualistisch	Szenario Misserfolg – kollektivistisch
Sandra Sandra studiert im ersten Semester Psychologie. Nun erwartet sie fieberhaft die Rückgabe der Statistik-Klausuren, denn das Ergebnis der Klausur hat wichtige Auswirkungen auf den Fortgang ihres Studiums. Vor der Prüfung hatte sie manchmal zusammen mit Marc und Manuela gelernt. Sie kennt die beiden aus ihrem Studentenwohnheim. Marc hatte in der Schule Mathe-Leistungskurs und konnte daher den beiden anderen häufig helfen, wenn sie Verständnisprobleme hatten. Sandra und Manuela kompensierten ihre Verständnislücken dadurch, dass sie viel auswendig lernten. Obwohl die drei nicht eng miteinander befreundet sind, sitzen sie heute im Hörsaal nebeneinander. Der Professor teilt die Klausuren aus. Marc hat eine 2, Manuela hat eine 3, und Sandra hat eine 5 und ist damit durchgefallen.	Michael Michael belegt das Seminar „Computerbasiertes Lernen“. Es wird in Projektgruppen gearbeitet: Jeweils vier Studierende arbeiten zusammen und sollen ein computerbasiertes Lernprogramm erstellen. Am Ende sollen sie für ihr Produkt und die gemeinsame Präsentation eine Note erhalten, die in die Magister-Abschlußnote eingeht. Michael und die anderen drei Gruppenmitglieder teilen beim ersten Treffen die anzufallenden Aufgaben unter sich auf und verabreden, dass sich jeder drei Wochen lang in sein Aufgabenfeld einarbeitet. Beim zweiten Treffen stellen sie sich gegenseitig ihre Ergebnisse vor. Mit Erstaunen bemerkt Michael, dass die anderen schon sehr viel detaillierter über ihren Bereich Auskunft geben können als er. Nachdem er über seine Ergebnisse berichtet hat, zeigen sich seine Zuhörer nicht ganz zufrieden. Zwar äußern sie ihre Kritik freundlich und gelassen, aber in der Sache sind sie dennoch direkt: Er habe die Literatur zu oberflächlich gelesen und nicht direkt auf die konkrete inhaltliche Fragestellung des Projektes bezogen. Einer sagt: „Du hast ja eigentlich deine Texte nur zusammengefaßt, anstatt sie hinsichtlich ihrer praktischen Konsequenzen für unser Projektthema auszuwerten! Das bringt uns ja nicht viel weiter.“

Abhängige Maße. Der Teilnehmer wird gebeten, sich in die Lage des jeweiligen Protagonisten des Szenarios zu versetzen und zu überlegen, wie er in dieser Situation reagieren würde. Zunächst wird ihm die Möglichkeit gegeben, in offener Beantwortung darüber Auskunft zu geben, welche Gefühle er an Stelle des Protagonisten empfinden würde. Dieses Item wurde vor allem aus Gründen der Förderung der Teilnahmemotivation in den Fragebogen integriert. Da die nachfolgenden Items nämlich nur ganz bestimmte Emotionen ansprechen, viele eigentlich naheliegende wie Ärger und Mitleid aber aussparen, war zu erwarten, dass die Probanden sich in ihrer Beantwortung zu sehr eingeschränkt fühlen und ihr Erleben nicht angemessen repräsentiert sehen. Auf der nächsten Seite des Frage-

bogens (und somit noch nicht vorher einsehbar) wird der Teilnehmer mit den in Abschnitt 4.3 erläuterten Emotions- und Emotionsdarstellungs-Items konfrontiert. Zwischen Emotion und Emotionsdarstellung erscheinen außerdem selbstkonstruierte Items zur Attribution, die – wiederum auf jeweils fünfstufigen Skalen – Auskunft darüber verlangen, inwieweit der Proband sein Ergebnis auf sein Interesse, seine Anstrengung, Zufall, seine Fähigkeit oder die Schwierigkeit der Aufgabe zurückführen würde. Am Ende der beiden Misserfolgsfragebögen folgen noch je zwei Fragen zum Bewältigungsverhalten, d.h. der Proband wird darüber befragt, ob er an bestimmten Trainingsmaßnahmen oder Tutorien teilnehmen, Hilfsangeboten von Mitstudenten annehmen würde etc., um künftig bessere Leistungen zu erbringen. Es wurden jeweils zwei solcher Fragen gestellt, um den Einfluss des sozialen Kontexts des Hilfsangebots (Dozent, Mitstudent etc.) auf die Entscheidung der Probanden möglichst zu neutralisieren.

Ein Szenario mitsamt Emotionsfragebogen – so nenne ich von nun an den Fragebogen zu Emotionen, Attributionen, Emotionsdarstellung und ggf. Bewältigungsverhalten – umfasst drei Din-A4-Seiten, so dass der Proband insgesamt zwölf Din-A4-Seiten zu bearbeiten hat. Zudem erhält er danach noch einen Bogen, auf dem er Angaben zu seiner Person (Alter, Geschlecht, Herkunftsland, Semesterzahl, Studienfach, Schulabschlussnote) machen sowie die deutsche Version der *Implicit Theories Scale* (Spinath, 2001; Original: Dweck, Chiu, & Hong, 1995a; vgl. Abschnitt 3.1) bearbeiten soll. Letztere wird zur Messung der Disposition Impliziter Theorien eingesetzt und umfasst drei Aussagen („Ich besitze ein bestimmtes Ausmaß an Intelligenz, das ich nicht verändern kann“, „Meine Intelligenz ist etwas, das ich kaum verändern kann“, „Ich kann zwar neue Dinge lernen, aber meine Intelligenz kann ich dadurch nicht wirklich verändern“), zu denen der Proband auf sechsstufigen Skalen seine Zustimmung angeben soll. Die Einbeziehung der Disposition Impliziter Theorien dient u.a. der Manipulationskontrolle der Induktion Impliziter Theorien, wobei die Dispositionswerte der Experimentalgruppe mit denen der Kontrollgruppe verglichen werden.

Teilnehmer. Die 102 Teilnehmer waren Studierende der Universität Kassel, vorwiegend in den Studienfächern Erziehungswissenschaften, Soziologie, Lehramt und Politik. Sie wurden im Rahmen von Lehrveranstaltungen geworben und erhielten sieben Euro oder eine Versuchspersonenbescheinigung (relevant für Studierende mit Nebenfach Psychologie) für ihre Teilnahme. Sie waren durchschnittlich 23 Jahre alt (mit einer Spanne von 19 bis 51), im vierten Semester und hatten eine Abiturnote von 2.8 (mit einer Spanne von 1.3 bis 4.4). Von den 102 Teilnehmern waren 80 weiblich sowie 90 deutscher Herkunft.

Instruktionen. Den Versuchspersonen wurde in konsistenter, schon bei ihrer Werbung einsetzender Weise vorgegeben, sie nähmen an zwei unterschiedlichen

Untersuchungen teil. Eine „medienpsychologische Untersuchung“ wurde in möglichst plausibler Weise (Schriftbild, Darstellungsart etc.) von einer Studie über „Emotionen im Studium“ abgegrenzt. Zunächst erhielten sie einen vierseitigen Fragebogen, der als „Voruntersuchung zu unserem Projekt ‚Verständlichkeit wissenschaftsjournalistischer Artikel in Mainstream-Printmedien‘“ deklariert war. Sie wurden angewiesen, die angeheftete Kopie eines Spiegel-Artikels sorgfältig durchzulesen und anschließend eine Inhaltsangabe (5-10 Sätze) anzufertigen.

Die Teilnehmer erhielten nun den bereits beschriebenen zwölfseitigen Szenario-Fragebogen, der wie gesagt als eigenständige Studie deklariert wurde. Sie wurden zunächst auf die Anonymität ihrer Angaben aufmerksam gemacht und zu ehrlicher Beantwortung ermutigt. Erstsemestern wurde nahegelegt, sich ggf. entsprechende Situationen aus ihrer Schulzeit vorzustellen. Die Bearbeitung dauerte ca. 15-25 Minuten.

Abschließend erhielten die Teilnehmer den („beiden Studien geltenden“) einseitigen Bogen mit Angaben zur Person sowie der Erhebung der Disposition Impliziter Theorien. Sie wurden schließlich über die Gefälschtheit des gelesenen Artikels aufgeklärt.

5.2 Ergebnisse

Bevor auf die Prüfung der Hypothesen für den Misserfolgsfall (Abschnitt 5.2.1) sowie für den Erfolgsfall (Abschnitt 5.2.2) eingegangen wird, soll es zunächst um einige Berechnungen zur Prüfung der Grundannahmen der Untersuchung gehen. Wenn im Folgenden von Statikern und Dynamikern gesprochen wird, so sind damit jeweils die Mitglieder der statischen bzw. dynamischen Induktionsbedingung gemeint; ist von statisch bzw. dynamisch disponierten Probanden die Rede, wird dies explizit gekennzeichnet. Diese Sprachregelung gilt auch für die folgenden Kapitel.

Zum Zweck der *Manipulationskontrolle* wurden die experimentellen Gruppen hinsichtlich der Dispositionswerte Impliziter Theorien verglichen. Es zeigten sich zwar hypothesenkonforme Tendenzen zwischen Statikern ($M = 3.20$, $SD = 0.94$), Dynamikern ($M = 2.91$, $SD = 1.10$) und der Kontrollgruppe ($M = 3.13$, $SD = 0.87$). Die Unterschiede zwischen diesen Werten (bei denen eine hohe Ausprägung eine stärkere Neigung zu einer statischen Intelligenztheorie bedeutet) waren jedoch nicht signifikant. Ebenso konnten keine Unterschiede in Abhängigkeit des Faktors Leistungsdiskurs bzw. Interaktionen zwischen jenen beiden Faktoren nachgewiesen werden.

Zur Einschätzung der Konstruktvalidität der Maße *negativer Emotionen* wurden deren korrelativen Zusammenhänge untersucht. Tabelle 6 gibt die bivariaten Korrelationen zwischen Merkmalen der Emotion und der Attribution für den

Misserfallsfall wieder. (Dabei wurden hier wie auch im Folgenden zunächst die jeweiligen Werte für die individualistische und die kollektivistische Bedingung zu einem gemeinsamen Wert verrechnet und nur bei Bedarf auf Analysen im zweifaktoriellen Messwiederholungsdesign zurückgegriffen, s.u.) Gemäß den in dieser Arbeit angestellten theoretischen Überlegungen sowie den empirischen Erfahrungen von Molfenter (1999) fallen die korrelativen Beziehungen in überwiegend erwarteter Richtung und Stärke aus. Hoffnungslosigkeit, Scham und Fähigkeitsattribution korrelieren untereinander deutlich positiv und zu Zuversicht deutlich negativ. Zudem steht Anstrengungsattribution in positiver Beziehung zu Schuld und Zuversicht. Die selbstrelevanten Emotionen Scham und Schuld korrelieren ebenfalls positiv. Tabelle 6 sind außerdem die internen Konsistenzen (Cronbachs Alpha) der Emotionskonstrukte zu entnehmen. Diese bewegen sich mit Ausnahme von Schuld alle in einem befriedigenden Bereich.

Tabelle 6

Korrelative Beziehungen und interne Konsistenzen der Emotions- und Attributionsmaße (Bedingung Misserfolg)

	Scham	Schuld	Zuversicht	Attr. Anstr.	Attr. Fäh.	Cronbachs Alpha
Hoffnungslosigkeit	.52***	.09	-.44***	-.13	.57***	.76
Scham		.25**	-.22**	.08	.35***	.77
Schuld			.15	.45***	-.05	.59
Zuversicht				.30***	-.58***	.71
Attribution Anstrengung					-.05	

Anmerkungen.

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig).

Die Werte der internen Konsistenz für die *positiven Emotionen* Freude ($\alpha = .78$), Hybris ($\alpha = .84$), Stolz ($\alpha = .70$) und Erleichterung ($\alpha = .73$) fallen ähnlich befriedigend aus. Die Korrelationsmuster lassen hier jedoch auf eine schlechte Trennung zwischen den Konstrukten schließen: Die vier untersuchten Emotionen korrelieren hoch positiv zueinander (jeweils $.3 < r < .6$ und $p < .01$, zweiseitig, mit Ausnahme der Beziehung zwischen Hybris und Erleichterung: $r = .17$, $p < .1$). Hybris steht erwartungsgemäß mit Fähigkeitsattributionen in Beziehung ($r = .37$, $p < .01$), Erleichterung unerwarteterweise mit Anstrengungsattributionen ($r = .23$, $p < .05$). Stolz korreliert zu beiden Attributionsformen hoch positiv (jeweils $.3 < r < .4$, $p < .01$).

5.2.1 Reaktionen auf negative Leistungsrückmeldungen

Tabelle 7 zeigt die Mittelwerte und Standardabweichungen der abhängigen Variablen für die verschiedenen experimentellen Gruppen.

Tabelle 7

Deskriptive Ergebnisse von Studie 1, Bedingung Misserfolg (oben in jeder Zelle: Mittelwert; unten in jeder Zelle: Standardabweichung)

	Implizite Theorie (Gesamt, „Sandra“+„Michael“)			Implizite Theorie (individualistisch, „Sandra“)			Implizite Theorie (kollektivistisch, „Michael“)		
	Kontroll (n=34)	Statiker (n=34)	Dyn. (n=34)	Kontroll (n=34)	Statiker (n=34)	Dyn. (n=34)	Kontr. (n=34)	Statik. (n=34)	Dyn. (n=34)
Hoffnungs- losigkeit	3.13 .59	3.16 .75	3.11 .81	3.58 .70	3.59 .83	3.54 .99	2.68 .79	2.73 .89	2.69 .86
Scham	3.12 .81	3.15 .66	2.92 .85	2.93 .97	2.88 .91	2.71 .98	3.30 .97	3.41 .80	3.13 .99
Schuld	3.80 .66	3.74 .72	3.98 .48	3.60 .75	3.49 1.01	3.74 .70	4.00 .85	4.00 .70	4.22 .68
Zuversicht	3.43 .68	3.37 .74	3.42 .83	3.03 .83	2.96 .97	3.01 1.01	3.84 .86	3.78 .78	3.82 .88
Attribution Anstr.	3.29 .93	3.72 .90	3.60 .78	3.00 1.02	3.53 1.21	3.26 1.08	3.59 1.08	3.91 1.00	3.94 .95
Attribution Fähigkeit	2.63 .75	2.81 1.04	2.78 1.05	3.06 .92	3.15 1.31	3.26 1.31	2.21 .81	2.47 1.11	2.29 1.06
Bew. verhalten	4.22 .73	4.36 .56	4.35 .43	4.51 .69	4.49 .63	4.66 .52	3.93 .91	4.24 .73	4.04 .66
Darstell. Scham	2.66 .71	2.41 .79	2.38 .76	2.76 .82	2.59 1.08	2.47 1.02	2.56 .89	2.24 .85	2.29 .87
Verstecken Scham	2.53 .72	2.46 .73	2.38 .89	2.29 .87	2.24 .89	2.35 1.20	2.76 .92	2.68 1.12	2.41 1.10
Darstell. Schuld	3.26 .85	3.44 .64	3.49 .69	3.15 .93	3.41 .96	3.50 .96	3.38 1.16	3.47 .79	3.47 1.11
Verstecken Schuld	2.47 .76	2.21 .71	2.28 .95	2.53 1.02	2.06 .95	2.21 1.04	2.41 1.02	2.35 .98	2.35 1.20
Darstell. Interesse	2.94 .80	2.94 .69	2.81 .95	2.26 .93	2.35 .98	2.21 1.01	3.62 1.02	3.53 .96	3.41 1.13
Darstell. Langew.	1.66 .66	1.75 .76	1.76 .72	1.88 .88	1.85 .99	1.76 .99	1.44 .61	1.65 .77	1.76 .85

In Varianzanalysen lassen sich keine hypothesenkonformen Einflüsse von *Impliziten Theorien* auf die Emotionen Hoffnungslosigkeit oder Zuversicht nachweisen. Hinsichtlich Schuld wird der erwartete Unterschied nur dann signifikant, wenn man Statiker und Dynamiker direkt (d.h. ohne Kontrollgruppe) miteinander vergleicht, $F(1,66) = 4.50$, $p < .05$, $\eta^2 = .04$. Dieser Unterschied lässt sich in etwa gleicher Stärke auch nur für die kollektivistischen Items nachweisen, innerhalb derer auch Scham schwach signifikant wird, $F(1,66) = 1.37$, $p < .1$, $\eta^2 = .03$. Bezieht man die Kontrollgruppe mit in die Varianzanalysen ein, lassen sich diese Unterschiede jedoch nicht mehr nachweisen. Auch bei multivariater Berücksichtigung aller vier Emotionen zeigen sich keine Effekte. Insgesamt unterscheiden sich also Statiker, Dynamiker und die Kontrollgruppe in ihren emotionalen Reaktionen nicht; allenfalls deutet sich an, dass Dynamiker mehr Schuld und z.T. weniger Scham empfinden als Statiker.

Die Einbeziehung weiterer Faktoren und Kontrollvariablen gibt Hinweise auf weitere Einflussquellen für die Emotionen. Eine multivariate Varianzanalyse der vier Emotionen mit den Faktoren Implizite Theorie und Leistungsdiskurs sowie den Kovariaten Disposition Impliziter Theorien, Geschlecht und Abiturnote kommt zu dem Ergebnis, dass Implizite Theorien ($F(4,89) = 2.04$, $p < .05$, $\eta^2 = .08$) sowie Geschlecht als Kovariate ($F(4,89) = 5.41$, $p < .005$, $\eta^2 = .20$) die größten Erklärungsbeiträge leisten. Die anderen Faktoren und Kovariaten sind nicht signifikant. Insbesondere die Mitberücksichtigung des Geschlechts der Probanden scheint also den Effekt Impliziter Theorien klarer zur Geltung kommen zu lassen (was gleichwohl aufgrund des explorativen Charakters der Berechnung nicht überinterpretiert werden sollte). Dies dürfte zurückzuführen sein auf die stärkere Tendenz der Frauen zu Hoffnungslosigkeit ($M = 3.3$ zu $M = 2.6$) und Scham ($M = 3.1$ zu $M = 2.8$) sowie die stärkere Tendenz der Männer zu Zuversicht ($M = 3.9$ zu $M = 3.3$; jeweils $SD \approx 0.7$). Die vermuteten stärkeren Auswirkungen der Induktion Impliziter Theorien bei Frauen lassen sich anhand zweifaktorieller Varianzanalysen hingegen nicht nachweisen. Ebenso wenig lässt sich die postulierte Interaktion zwischen den Faktoren Implizite Theorie und Leistungsdiskurs in den Daten wiederfinden.

Die Antworten auf das offene Emotions-Item wurden kategorisiert und in Form von Häufigkeitszählungen ausgewertet. Es zeigen sich ähnliche Ergebnistendenzen wie anhand der Items mit vorgegebenen Antwortalternativen, die sich aber nicht inferenzstatistisch absichern lassen.

Für die Merkmale Anstrengungsattribution, Fähigkeitsattribution, Bewältigungsverhalten, Darstellung von Langeweile, Darstellung von Schuld, Darstellung von Lernfreude und Verstecken von Scham liegen entgegen der Hypothesen keine empirischen Hinweise auf einen Einfluss Impliziter Theorien vor, weder in einfachen Varianzanalysen noch bei Berücksichtigung der oben genannten Kontrollvariablen oder dem Faktor Leistungsdiskurs.

Dem Einfluss der Induktion Impliziter Theorien lässt sich derjenige der Disposition Impliziter Theorien direkt gegenüberstellen, indem man letzteren innerhalb der Kontrollgruppe (d.h. unbeeinflusst von der Induktion) analysiert. Hier stehen die Dispositionswerte erwartungsgemäß positiv zu Hoffnungslosigkeit ($r = .43, p < .01$) und Scham ($r = .27, p < .1$) sowie negativ zur Darstellung von Interesse ($r = -.26, p < .1$) in Beziehung. Die übrigen Hypothesen hinsichtlich Impliziter Theorien ließen sich so jedoch nicht bestätigen.

Betrachtet man die Beziehung zwischen der Disposition Impliziter Theorien und den abhängigen Variablen für die gesamte Stichprobe, kommt man zu nahezu identischen Resultaten wie in der Kontrollgruppe. Dies überrascht, da sich in diesen Beziehungen ja jeweils (außer in der Kontrollgruppe) die Einflüsse der Induktion einerseits und der „wirklichen“ Disposition andererseits bündeln sollten (d.h. statisch Induzierte sollten z.B. einen tendenziell hohen Dispositionsscore und einen tendenziell hohen Hoffnungslosigkeitswert haben). Auch die von Dweck et al. (1995a) vorgeschlagene kategoriale Auswertung des Dispositionswerts, bei der Probanden mit Werten größer als vier zu Statikern, Probanden mit Werten kleiner als drei zu Dynamikern erklärt und die übrigen Personen aus der Analyse ausgeschlossen werden, führt zu keinen signifikanten Unterschieden zwischen den beiden so gewonnenen Gruppen. Im Übrigen unterscheidet sich die Größe dieser Gruppen deutlich von denen aus US-amerikanischen Studien. In der hier untersuchten deutschen Stichprobe befinden sich 18 Statiker, 23 Ausgeschlossene und 61 Dynamiker (Kontrollgruppe: 4/10/20), während gemäß Dweck et al. (1995a) das Verhältnis in etwa 42% zu 16% zu 42% betragen sollte. Dieses Ergebnis steht in direkter Übereinstimmung mit der Arbeit Spinaths (2001), die für ihre deutsche Stichprobe ein Verhältnis von 20% zu 20% zu 60% fand.

Der Einfluss des Faktors *Leistungssituation* wurde in Varianzanalysen mit Messwiederholungen untersucht. Die Probanden schildern im Rahmen des kollektivistischen Szenarios erwartungsgemäß mehr Scham, $F(1,100) = 17.79, p < .001, \eta^2 = .15$, und Schuld, $F(1,100) = 24.93, p < .001, \eta^2 = .20$, jedoch entgegen der Hypothesen weniger Bewältigungsverhalten ($M = 4.07$ zu $M = 4.55$) und Schamdarstellung ($M = 2.36$ zu $M = 2.61$) sowie nur geringfügig mehr Schamverstecken ($M = 2.37$ zu $M = 2.26$; jeweils $SD \approx 0.9$) als im individualistischen Szenario. Zweifaktorielle Varianzanalysen mit den Faktoren Implizite Theorie und Leistungssituation erbringen für die Merkmale Scham, Schuld und Bewältigungsverhalten keine signifikanten Interaktionseffekte (weder univariat noch multivariat).

Zwischen den Emotionen und dem Bewältigungsverhalten liegen keine signifikanten korrelativen Beziehungen der erwarteten Richtung vor. Aufgrund der unbefriedigenden empirischen Bestätigung der Hypothesen wurde von der Durchführung von Mediator- und Pfadanalysen Abstand genommen.

5.2.2 Reaktionen auf positive Leistungsrückmeldungen

Die in Abschnitt 4.1.2 vermuteten Unterschiede zwischen Statikern und Dynamikern hinsichtlich Anstrengungsattribution, Stolz, Hybris, Erleichterung, Darstellung von Hybris sowie Verstecken von Freude lassen sich empirisch nicht nachweisen, weder anhand einfacher Varianzanalysen noch unter Berücksichtigung weiterer Faktoren oder Kontrollvariablen. Die Hypothese zum Bewältigungsverhalten konnte in diesem Design aufgrund fehlender Daten nicht überprüft werden (vgl. hierfür die Studien 2 und 3).

Bei Mitberücksichtigung des Faktors Leistungsdiskurs zeigen sich keine Haupt- oder Interaktionseffekte. Bezieht man neben dem Leistungsdiskurs auch die Kontrollvariablen Disposition Impliziter Theorien, Geschlecht und Abiturnote in eine multivariate Varianzanalyse (mit den vier Emotionen als abhängige Variablen) mit ein, wird deutlich, dass – ähnlich wie im Misserfolgsfall – die Kovariaten Geschlecht ($\eta^2 = .15$, $p < .01$) und Abiturnote ($\eta^2 = .08$, ns) weit mehr Varianz erklären als die induzierten Impliziten Theorien ($\eta^2 = .02$, ns). Ein Blick auf die Geschlechtsunterschiede verrät, dass Frauen in *allen* Emotions- und Emotionsdarstellungsmaßen mindestens der Tendenz nach höhere Werte (bzw. niedrigere bei den Items zum Verbergen von Emotionen) aufweisen als Männer. Eine Interaktion zwischen Geschlecht und Implizite Theorie konnte wiederum nicht festgestellt werden.

In der kollektivistischen Situation sollten die Probanden mehr Hybris und Stolz empfinden sowie mehr Stolz und weniger Hybris darstellen als in der individualistischen. Lediglich hinsichtlich Hybris konnten diese Vorhersagen Bestätigung finden, $F(1,100) = 41.69$, $p < .001$, $\eta^2 = .30$.

5.3 Diskussion

Die geschilderten Ergebnisse vermitteln ein recht einheitliches Bild. Die theoretisch abgeleiteten Hypothesen konnten überwiegend nicht bestätigt werden; zwischen Statikern und Dynamikern liegen zumeist keine Unterschiede hinsichtlich ihrer Reaktionen auf Misserfolg und Erfolg vor. Auch zwischen neurowissenschaftlichem und psychologischem Diskurs zeigen sich keine Differenzen bzw. Interaktionen mit dem Faktor Implizite Theorie. Dieser Abschnitt will zunächst diskutieren, wie die wenigen hypothesenkonformen Ergebnisse zu interpretieren sind, um danach zu untersuchen, worin das Ausbleiben der angenommenen Effekte begründet liegen könnte.

Für den Misserfolgsfall besteht das einzige deutlich hypothesenkonforme Resultat darin, dass die Probanden in der kollektivistischen Situation mehr Scham und Schuld empfinden als in der individualistischen. Ob Lernende primär für ihr eigenes oder aber für das Ergebnis einer ganzen Gruppe (mit-)verantwortlich

sind, scheint für die Wahrscheinlichkeit des Auftretens selbstrelevanter Emotionen von Bedeutung zu sein. Dies steht im Einklang mit grundlegenden Befunden zu selbstrelevanten Emotionen, wonach diese vor allem in sozialen Situationen auftreten (z.B. Lewis, 2000), und ist daher nicht überraschend. Auffallend ist aber, dass trotz der stärkeren Tendenz zu Scham und Schuld in der kollektivistischen Bedingung weniger Bewältigungsverhalten gezeigt wird als in der individualistischen. Dies könnte daran liegen, dass das individualistische Setting (Statistik-Klausur) als praktisch bedeutsamer eingeschätzt wird als das kollektivistische (Projektseminar). An diesem Punkt wird ein wichtiger Nachteil der Szenario-Methode deutlich, nämlich dass die Reaktionen auf unterschiedliche Szenarien selten sinnvoll miteinander vergleichbar sind. In den Studien 2 und 3 soll dieses Problem überwunden werden, indem die individualistische und die kollektivistische Bedingung im Rahmen eines weitestmöglich identischen, weil im Labor standardisierten Settings realisiert werden.

Für den Erfolgsfall findet sich lediglich die Bestätigung der Hypothese, dass in der kollektivistischen Bedingung mehr Hybris erlebt wird als in der individualistischen. Dies stellt ein an sich interessantes Ergebnis dar, das allerdings im Lichte der anderen, nicht bestätigten Vorhersagen auch als Zufallsprodukt erscheinen kann. Jedenfalls ist eine Bestätigung dieses Befundes in weiteren Untersuchungen nötig – nicht zuletzt auch wegen der teilweise unerwarteten korrelativen Beziehungen zu den anderen Emotionen.

Der Faktor Implizite Theorie hat sich – mit Ausnahme einiger richtiger Ergebnistendenzen für Scham und Schuld – als nahezu vollkommen bedeutungslos erwiesen. Andere Variablen (wie etwa das Geschlecht der Probanden) haben sich als wirkungsmächtiger herausgestellt. Worin dies begründet liegt, soll nun diskutiert werden.

Die zumindest im Misserfolgsfall theoretisch stimmigen Korrelationsmuster zwischen den Emotionen sowie die recht guten internen Konsistenzen der Skalen sprechen dafür, dass die Versuchsteilnehmer die Fragen mit hinreichender Konzentration bzw. Motivation beantwortet haben. Einfache Erklärungen für das Scheitern der Hypothesen, die in diese Richtung zielen, können also ausgeschlossen werden.

Sodann besteht – im Zuge einer *inhaltlichen Deutung* der Ergebnisse – die Möglichkeit, die kausale Bedeutsamkeit Impliziter Theorien, wie sie in Dwecks Modell postuliert wird, in Zweifel zu ziehen. Hier ist jedoch zu bedenken, dass ebenso für die Disposition Impliziter Theorien kaum Zusammenhänge zu den abhängigen Variablen gefunden werden konnten. Somit müsste man gleichzeitig die korrelative Version des Dweckschen Modells infrage stellen, was im Lichte der diesbezüglich äußerst zahlreichen empirischen Bestätigungen kaum angebracht erscheint. Eine etwas differenziertere Variante der inhaltlichen Deutung der Ergebnisse könnte postulieren, dass Dwecks Modell im europäischen Kul-

turkreis nicht gültig ist, da das Intelligenzkonzept hier weniger verbreitet und gesellschaftlich akzeptiert sein könnte als in Nordamerika, wo die meisten der zitierten Studien durchgeführt wurden. Ein Indiz für diese Erklärung liefert der bei Spinath (2001) erstmals vorgestellte und in der vorliegenden Studie replizierte Befund, dass in Deutschland die dynamische Intelligenztheorie auf Dispositionsebene sehr viel verbreiteter ist als in den USA. Dies könnte ein Hinweis darauf sein, dass Intelligenz bei der Beurteilung von Leistungen in Deutschland als weniger relevant erachtet wird als in Nordamerika und somit auch Implizite Intelligenztheorien keine so zentrale Bedeutung für das Leistungsverhalten haben. Zu bedenken ist aber, dass es durchaus bereits Studien aus dem europäischen Raum gibt, die Dwecks Modell mindestens in seinen Grundzügen bestätigen (vgl. Cury, Elliot, Da Fonseca, & Moller, 2006; Schlangen & Stiensmeier-Pelster, 1997; Ziegler, 2001). Auch Studien aus Dwecks Arbeitsgruppe wurden zuweilen in nicht-amerikanischen, nämlich ostasiatischen Kulturen durchgeführt (z.B. Hong, Chiu, Dweck, Lin, & Wan, 1999) – wenngleich nicht klar ist, inwieweit deren studentische Probanden vom amerikanischen Lebens- und Denkstil beeinflusst waren.

Schließt man sich diesen inhaltlichen Deutungen der Ergebnisse nicht an, bieten sich vor allem zwei Ansatzpunkte an, um die fehlenden Effekte Impliziter Theorien auf *methodische Schwächen der Untersuchung* zurückzuführen: Einerseits könnte die Induktion mit Hilfe der Zeitschriftenartikel nicht erfolgreich gewesen sein, andererseits könnte die Szenario-Methode unter der gewählten Fragestellung ungeeignet für die Erhebung der abhängigen Variablen gewesen sein. Beide Erklärungsansätze schließen sich freilich nicht wechselseitig aus. Dennoch soll die folgende Diskussion helfen zu entscheiden, welcher von beiden theoretisch und empirisch stichhaltiger ist und somit die Konzeption weiterer Untersuchungen bestimmen soll.

Dass die *Induktion Impliziter Intelligenztheorien fehlgeschlagen* sein könnte, legt vor allem die Manipulationskontrolle nahe. Gemäß dem am Ende des Experiments erhobenen Dispositionswert lassen sich keine Unterschiede zwischen Statikern, Dynamikern und Kontrollgruppe ausmachen. Eine Erklärungsmöglichkeit für das Ausbleiben dieses Effekts in der vorliegenden Studie besteht darin, dass die Probanden nach dem zwölfseitigen Emotionsfragebogen die abschließenden Fragen zur Disposition mit geringer Motivation und Konzentration bearbeitet haben könnten. Wichtiger indes erscheinen grundlegendere Überlegungen, die die Notwendigkeit bzw. Aussagekraft einer Manipulationskontrolle hinterfragen. Hiernach lässt gerade der *implizite* Charakter Impliziter Theorien fragwürdig erscheinen, dass diese selbst zuverlässiger messbar sind als ihre kognitiven, emotionalen und behavioralen Auswirkungen. Zur Vertiefung dieser Argumentation sei auf Abschnitt 9.4 verwiesen, wenn es um eine kulturpsychologi-

sche Konzeption Impliziter Theorien geht. Diese wird im Folgenden nur in den Grundzügen skizziert.

In Abgrenzung zur häufig gewählten positivistischen Zugangsweise zu Impliziten Theorien, die sie als feste, klar definierbare und zuverlässig messbare Entitäten in den Blick nimmt (vgl. z.B. Spinath, 2001, aber auch viele Arbeiten Dwecks), verfolgt die vorliegende Arbeit einen eher geistes- und kulturwissenschaftlichen Ansatz (vgl. Kapitel 3.2.0), wonach kulturelle Deutungsangebote wie die verwendeten Zeitschriftenartikel (solange sie verstanden werden und glaubwürdig erscheinen) *immer* die Impliziten Theorien der Rezipienten beeinflussen – die Frage ist nur, inwiefern und mit welchen Konsequenzen. Gerade weil eine Implizite Intelligenztheorie Teil eines „meaning system“ ist, wie es ja Dweck selbst ausdrückt (vgl. Abschnitt 3.1), macht es wenig Sinn, sie so zu behandeln, als wäre sie eine durch Fragebogen problemlos messbare Entität. Wenn beispielsweise eine Manipulationskontrolle nicht die intendierten Effekte ausweist, heißt dies noch lange nicht, dass sich das Lesen des Artikels nicht auf anderen Merkmalen des Erlebens und Verhaltens niederschlägt, die nicht so leicht der bewussten Kontrolle unterliegen wie die Angabe der eigenen Auffassung von Intelligenz. Wenn die Manipulationskontrolle erfolgreich ist, muss das umgekehrt noch nicht unbedingt etwas über die Implizite Theorie der Probanden aussagen, sondern eher darüber, ob sie die Fragen über Intelligenz im Sinne eines Verständnistests verstehen und die Botschaft des Artikels brav wiederholen oder ob sie sich ihre eigene Meinung bewahren wollen bzw. – weil sie den Untersuchungszweck durchschauen – gar Reaktanz zeigen. Da man diesen Faktor – die Bereitwilligkeit, die Botschaft des Artikels um des Versuchsleiters willen zu übernehmen – nicht kontrollieren kann, dürften die nach der Induktion erhobenen Dispositionswerte äußerst fehlerbehaftet sein.

Ähnliche Überlegungen liegen anscheinend auch den in Abschnitt 3.2.1 vorgestellten Studien von Hong et al. (1999) und Niiya et al. (2004) zugrunde, denn sie verzichten ganz auf eine Manipulationskontrolle. In den Studien 2 und 3 der vorliegenden Arbeit wird aus den genannten Gründen ebenfalls auf eine Manipulationskontrolle verzichtet.

Unter der Annahme, dass die Zeitschriftenartikel (etwa weil sie unglaubwürdig oder unverständlich waren) nicht zur Induktion geeignet waren, wäre zu erwarten, dass die Hypothesen immerhin in Abhängigkeit der Disposition Impliziter Theorien Bestätigung finden. Dies ist jedoch nicht – oder nur sehr vereinzelt, s.o. – der Fall, weder innerhalb der Kontrollgruppe, die weniger vom oben erwähnten Problem der Müdigkeit bzw. Unkonzentriertheit betroffen sein sollte, noch in der Gesamtstichprobe. Ein Blick auf die Ergebnisse lässt eher darauf schließen, dass sowohl die Induktion als auch die Disposition Impliziter Theorien nur vereinzelte, zufällig erscheinende Effekte nach sich ziehen.

Gewiss gibt es auch Argumente, die für die These der gescheiterten Induktion sprechen, etwa die Tatsache, dass sich noch nicht einmal an einem so basalen und gut untersuchten Konstrukt wie der Anstrengungsattribution die vorhergesagten Unterschiede finden lassen, wie dies etwa bei Hong et al. (1999) deutlich der Fall war. Auch sollte in künftigen Studien darüber nachzudenken sein, ob das bloße Zusammenfassen des Artikels ausreicht, um die Probanden hinreichend dazu zu motivieren, sich mit der Aussage des Artikels auseinanderzusetzen (vgl. die diesbezüglichen Änderungen in den Studien 2 und 3). Im Ganzen jedoch erscheint jene These als Erklärung für die ausgebliebenen Effekte aus den genannten Gründen wenig befriedigend.

Wie sieht es demgegenüber mit der anderen methodisch ansetzenden Erklärung aus, welche das Scheitern der Hypothesen auf die *Inadäquatheit der Szenario-Methode* zurückführt? Die Szenario-Methode wurde ja in der Emotionspsychologie vor allem mit dem Ziel eingesetzt, das Auftreten qualitativ unterschiedlicher Emotionen auf bestimmte Situationsmerkmale (bzw. Einschätzungsmuster) zurückzuführen (vgl. Weiner, 1994). Für diesen Untersuchungszweck hat sie sich zweifellos als nützlich erwiesen (Reisenzein, Meyer, & Schützwohl, 2003). Die hier verfolgte Fragestellung interessiert sich aber eher für quantitative Unterschiede von Emotionen in Abhängigkeit des Primings mit Zeitschriftenartikeln. Es geht beispielsweise nicht darum, ob Menschen aufgrund des Durchfallens durch eine Prüfung verzweifelt und beschämt sind, sondern ob Statiker diese Emotionen *in stärkerem Ausmaß* zeigen als Dynamiker. Diese quantitativen Unterschiede dürften, wie auch Blackwell et al. (2007) feststellen, im Regelfall eher gering ausfallen, da auch viele andere Einflüsse (Persönlichkeit, Lerngeschichte, Intelligenz) in Leistungssituationen wie der im Szenario beschriebenen wirksam sind. Plausibel erscheint daher, dass die Szenario-Methode aufgrund ihres hypothetischen, nicht am wirklichen Erleben und Verhalten orientierten Charakters nicht in der Lage ist, die zu erwartenden subtilen Induktionseffekte abzubilden.

Ein Indiz für diese Erklärungsvariante sind z.B. die überaus deutlichen Geschlechtsunterschiede. Frauen beschreiben ihre Emotionen als fast durchweg intensiver als Männer. Dies kann man als Hinweis dafür betrachten, dass die Probanden sich bei der Angabe ihrer hypothetischen Reaktionen auf die Leistungsrückmeldungen zu einem nicht geringen Anteil von Rollenerwartungen haben leiten lassen und ihr Verhalten mithin nicht in hinreichender Genauigkeit vorherzusagen in der Lage waren. Dass derartige Tendenzen zu sozialer Erwünschtheit eine Rolle spielen, kann man auch aus den (geschlechtsübergreifend) äußerst hohen Mittelwerten für das Merkmal Bewältigungsverhalten sowie den niedrigen Werten für die Emotionsdarstellungs-Items ableiten (vgl. Tabelle 7): Jeder „will was dagegen tun“, und jeder „verhält sich authentisch“.

Hier ließe sich einwenden, dass derartige Urteilsverzerrungen ja alle Experimentalgruppen gleichermaßen betreffen und daher die angenommenen Unter-

schiede zwischen den Gruppen zwar verwässern, aber nicht gänzlich eliminieren sollten. Dass aber tatsächlich alle Gruppen gleichermaßen betroffen sind, erscheint fraglich, wenn man sich vor Augen führt, dass gemäß Dwecks Modell gerade die Statiker im Zuge ihrer Leistungszielorientierung zu positiver Selbstdarstellung neigen sollten (vgl. Abschnitt 3.3.3). So konnten Mueller und Dweck (1998) zeigen, dass Statiker eher dazu neigen, hinsichtlich ihres Leistungsergebnisses die Unwahrheit zu sagen. Das hieße aber nichts anderes, als dass Statiker ihr (angenommenes) dysfunktionales Leistungsverhalten durch ihre besonders ausgeprägte Tendenz zu positiver Selbstdarstellung verstecken, so dass sie im Endeffekt die Fragen ähnlich beantworten wie die Kontrollgruppe. Diese Argumentation lässt sich umgekehrt auch auf die Dynamiker und ihre weniger ausgeprägte Tendenz zur Selbstdarstellung anwenden.

Freilich hat man mit den beschriebenen Urteilsverzerrungen aufgrund sozialer Erwartungen *immer* zu tun, wenn man Versuchspersonen zur Auskunft über ihr Erleben und Verhalten auffordert. Weil Emotionen anders als über den Selbstbericht kaum zu erheben sind (vgl. z.B. Meyer, Schützwohl, & Reisenzein, 2001), wird man das skizzierte Problem im Rahmen der vorliegenden Fragestellung nicht vollends lösen können. Dennoch ist damit zu rechnen, dass es im Kontext der Szenario-Methode in besonders ausgeprägter Weise auftritt. Diese nämlich fordert von den Probanden nicht nur den Selbstbericht ein, sondern tut dies gleichzeitig in einer Situation, in der das Erleben und Verhalten, über das berichtet werden soll, gar nicht vorliegt. Die Tendenz, sich gemäß bestimmten Rollenerwartungen darzustellen, gerät so gleichsam noch nicht einmal mit dem Bestreben in Konflikt, aufrichtig zu sein und die Wahrheit zu sagen. Weil es keinen deutlich wahrnehmbaren Gegenstand gibt, über den man berichten könnte, kann man auch nicht lügen – denn es lässt sich ja nahezu jedes Szenario so umdeuten, dass es mit den eigenen Selbstdarstellungswünschen kompatibel ist. Verstrickte man hingegen den Probanden in eine konkrete, klar umrissene Leistungssituation, wäre immerhin ein solcher Gegenstand vorhanden, an dem der Proband seinen Selbstbericht ausrichten müsste.

Dieser Weg wird in den Studien 2 und 3 gegangen, welche in den folgenden Kapiteln beschrieben werden. Damit wird der Erklärungsvariante, die die ausbleibenden Effekte auf die Inadäquatheit der Szenario-Methode zurückführt, mehr theoretische und empirische Plausibilität eingeräumt als der die Induktion betreffenden Variante sowie der inhaltlichen Deutung der Ergebnisse, aufgrund derer man andere, weitreichendere Konsequenzen ziehen müsste. Dies soll nicht heißen, dass diese alternativen Erklärungsansätze eindeutig von der Hand zu weisen sind. Inwieweit sie tatsächlich weniger Geltung besitzen als die hier präferierte Erklärung, werden wohl erst die Ergebnisse der im Folgenden beschriebenen Studie 2 zeigen.

6. Studie 2: Prüfung des Modells anhand einer Leistungsaufgabe

Zu Beginn von Kapitel 5 habe ich zwei Strategien zur Anregung von Leistungs-emotionen unterschieden. Aufgrund empirischer Befunde und ökonomischer Erwägungen war ich zu der Auffassung gelangt, dass der Zugang über die Szenario-Methode für die vorliegende Fragestellung lohnend sein könnte. Indes haben die Ergebnisse der ersten Studie gezeigt, dass die aufgestellten Hypothesen anhand von Szenarien überwiegend *nicht* bestätigt werden konnten. Daher soll in einer zweiten Studie untersucht werden, ob die andere in Abschnitt 5.0 beschriebene Forschungsstrategie zu hypothesenadäquateren Ergebnissen führt. Emotionen und andere Reaktionen sollen nun durch die Rückmeldung von Erfolg oder Misserfolg im Rahmen einer konkreten Leistungssituation angeregt werden.

Da dieses Vorgehen deutlich zeitaufwändiger ist und zudem jede Versuchsperson nur noch *einer* experimentellen Bedingung zugewiesen werden kann, ist es in ökonomischer Hinsicht unausweichlich, mindestens einen experimentellen Faktor fallenzulassen. Ich habe mich für den Faktor Leistungsdiskurs entschieden. Im Lichte der Ergebnisse von Studie 1 erscheint es angebracht, eher diejenigen Hypothesen zu prüfen, bei denen aufgrund theoretischer Überlegungen und empirischer Befunde vergleichsweise deutliche Unterschiede zu erwarten sind. Dies scheint in Abhängigkeit des Faktors Leistungsdiskurs nicht der Fall zu sein. Implizite Theorien werden in Studie 2 daher bei *allen* Probanden mit Hilfe der neurowissenschaftlichen Artikel induziert. Demnach kann nicht mehr die Wirkmächtigkeit des neurowissenschaftlichen Leistungsdiskurs *im Vergleich zum psychologischen* eruiert werden; es kann nurmehr darum gehen, ob der neurowissenschaftliche Diskurs *überhaupt* zur Beeinflussung des Leistungsverhaltens tauglich ist.

Ebenfalls aus Ökonomiegründen kann das Forschungsdesign von Studie 2 keine Kontrollgruppe mehr beinhalten. Der Einfluss der statischen Bedingung ist daher nicht mehr separat von dem der dynamischen Bedingung erfassbar. Dies sollte für die große Mehrheit der Hypothesen keinerlei Einschränkungen in der Interpretierbarkeit der Ergebnisse mit sich bringen. Im Übrigen verwenden auch die in Abschnitt 3.2.1 beschriebenen experimentellen Studien überwiegend keine Kontrollgruppe. Einzig für die Fragestellung zu den Auswirkungen neurowissenschaftlicher Leistungsdiskurse ist das Fallenlassen der Kontrollgruppe von Bedeutung: Käme es zu hypothesenkonformen Unterschieden in Abhängigkeit der beiden Artikel, bliebe unklar, ob für diesen Effekt eher die schädliche Wirkung des statischen Artikels oder die förderliche Wirkung des dynamischen Artikels verantwortlich zu machen ist. Schlussfolgerungen in diesem Kontext büßen damit an praktischer Relevanz ein. In Abschnitt 9.3 wird allerdings zu begründen versucht, weshalb die Größe dieser Einbußen auch nicht überschätzt werden sollte.

Studie 2, so kann als Fazit festgehalten werden, prüft das postulierte theoretische Modell mittels einer verhaltensnäheren Operationalisierung und Messung der abhängigen Variablen. Die Verhaltensnähe erkaufte sie sich mit Änderungen im Forschungsdesign, die die Aussagekraft der Studie hinsichtlich der Auswirkungen neurowissenschaftlicher Leistungsdiskurse in mehrfacher Hinsicht einschränken.

6.1 Methode

Studie 2 berücksichtigt die Faktoren Implizite Intelligenztheorie, Leistungssituation und Leistungsrückmeldung. Alle drei Faktoren werden zwischen den Versuchspersonen variiert, d.h. jede Person erhält einen der beiden Zeitschriftenartikel (vgl. Abschnitt 4.2), wird in eine individualistische oder kollektivistische Leistungssituation versetzt und bekommt später – auf die Bearbeitung einer Leistungsaufgabe hin – entweder eine negative oder eine positive Rückmeldung. Da immer zwei Personen simultan untersucht werden und jeweils die objektiv erfolgreichere Person die positive, die andere die negative Rückmeldung erhält, kann der Faktor Leistungsrückmeldung nicht experimentell variiert werden. Es liegen somit zwei Teilstudien mit einem 2x2-Versuchsdesign vor, die Reaktionen auf Misserfolg und Erfolg getrennt voneinander thematisieren. Die Teilstudie zu Erfolgsreaktionen arbeitet mit einer selektiv leistungsstarken, die zu Misserfolgsreaktionen mit einer selektiv leistungsschwachen Stichprobe.

Teilnehmer. Die 88 Versuchsteilnehmer waren Studierende der Universität Kassel, vorwiegend in den Studienfächern Erziehungswissenschaften, Soziologie und Lehramt. Sie wurden im Rahmen von Lehrveranstaltungen sowie über Aushänge in Universitätsgebäuden geworben und erhielten sechs Euro oder eine Versuchspersonenbescheinigung (relevant für Studierende mit Nebenfach Psychologie) für ihre Teilnahme. Diese Vergütung erfolgte unabhängig von der geldlichen Belohnung für das erfolgreiche Abschneiden bei der Leistungsaufgabe (s.u.). Die Probanden waren durchschnittlich 23 Jahre alt (mit einer Spanne von 19 bis 42), im fünften Semester und hatten eine Abiturnote von 2.7 (mit einer Spanne von 1.2 bis 4.4). Von den 88 Teilnehmern waren 65 weiblich sowie 78 deutscher Herkunft.

Leistungsaufgabe. Die beiden Probanden bearbeiten separat voneinander einige Aufgaben des Intelligenztests „Analyse des Schlussfolgernden und Kreativen Denkens“ (ASK) (Schuler & Hell, 2005). Es handelt sich um die Aufgaben 2, 3, 4, 5 und 6 des Untermoduls „Informationen interpretieren“. Die ASK ist speziell für Erwachsene mit Abitur konzipiert und enthält dementsprechend anspruchsvolle Aufgaben. In den ausgewählten Items geht es darum, aus dargebotenen Tabellen und Grafiken abzulesen, welche von mehreren zur Auswahl stehenden

Schlussfolgerungen sie zulassen und welche nicht. Beispielsweise beinhaltet Aufgabe 6 eine Tabelle über die Anzahl und prozentuale Häufigkeit von Ehepaaren in Deutschland in Abhängigkeit ihrer Nationalitäten und der im Haushalt lebenden Kinder. Der Proband soll anhand dieser Tabelle den Wahrheitsgehalt beispielsweise folgender Aussage beurteilen: „Binationale Ehen oder Ehen, bei denen beide Partner ausländischer Staatsangehörigkeit sind, bringen im Verhältnis häufiger mindestens ein Kind hervor als Ehen, bei denen beide Partner deutscher Staatsangehörigkeit sind.“

Der Test besteht aus insgesamt 18 solcher richtig-falsch-Wahlmöglichkeiten, die sich auf die fünf Aufgaben verteilen. Sollte ein Proband bei seiner Bearbeitung der Aufgaben raten, könnte er mit einem Ergebnis von neun richtigen Antworten rechnen. Anhand der Normen der ASK ist zu erwarten, dass etwa jeder Dritte einer studentischen Stichprobe mindestens 12 Punkte, etwa jeder Zehnte mindestens 15 Punkte erreicht. Diese beiden Werte werden bei der Leistungsrückmeldung als Operationalisierung einer erfolgreichen bzw. sehr erfolgreichen Bearbeitung verwendet (s.u.).

Wie anhand der obigen Beschreibung nachzuvollziehen sein sollte, erinnern die Aufgaben wenig an klassische Intelligenztests. Um diese Assoziation gezielt zu vermeiden, wurden sie als Aufgaben „zum Verständnis von Tabellen und Grafiken“ deklariert. Die Versuchsleiterin vermied jeglichen Bezug auf die Intelligenzsemantik, sprach z.B. auch nicht von einem „Test“. Implizit sollten die Aufgaben aber dennoch den Schluss nahelegen, dass ein Misserfolg an zu geringen Kompetenzen des Bearbeitenden liegt (und nicht an Zufall, Fleiß etc.). Die Aufgaben wurden also durchaus so ausgewählt, dass Fähigkeitsattributionen („da bin ich zu blöd zu“) immerhin *möglich* sind. Die hohe Schwierigkeit der Aufgaben sollte es den Teilnehmern außerdem erschweren, eine Erwartungshaltung hinsichtlich des eigenen Abschneidens bei den Aufgaben aufzubauen und somit die gefälschte Rückmeldung (s.u.) zu durchschauen. Schließlich eignen sich die ASK-Aufgaben auch deshalb gut für die vorliegende Untersuchung, weil sie relativ praxisnahe Kompetenzen betreffen, die daher von hoher Relevanz für das Selbstkonzept der Probanden sind. Im Zuge dessen werden den Probanden selbstwertdienliche Rechtfertigungen („bei so abstrakten, künstlichen Aufgaben streng’ ich mich doch nicht an“) erschwert.

Messung der Disposition Impliziter Theorien. In Studie 1 wurde die Disposition Impliziter Theorien *nach* deren experimenteller Induktion erhoben, so dass eine Manipulationskontrolle möglich war. In Studie 2 hingegen wird die Disposition *vor* der Induktion gemessen. Auf die Gründe für den Verzicht auf eine Manipulationskontrolle wurde bereits in Abschnitt 5.3 eingegangen. Studie 2 entscheidet sich dezidiert für eine Dispositionserhebung *vor* der Induktion, weil nur auf diese Weise das Zusammenspiel von Disposition und Induktion angemessen analy-

siert werden kann. Würde man zusätzlich auch nach der Induktion Dispositionswerte erheben, wäre der Untersuchungszweck für die Probanden wahrscheinlich zu leicht durchschaubar.

Die Disposition Impliziter Theorien wird – anders als in Studie 1 – anhand zweier Operationalisierungen bzw. in zwei Dimensionen gemessen. Die Probanden erhalten nicht nur die drei Items der bereits in Studie 1 verwendeten „Implicit Theories Scale“ (vgl. Abschnitt 5.1; hier begrifflich leicht angepasst), sondern auch drei von Spinath (2001; in Anlehnung an Stipek und Gralinski, 1996) konstruierte Items zur Einschätzung der Bedeutsamkeit von Intelligenz für mathematische Leistungen. Die Teilnehmer sollen auf sechsstufigen Skalen ihre Zustimmung quantifizieren hinsichtlich Aussagen wie „Um in mathematischen Aufgaben gut zu sein, muss man intelligent sein“. Untersucht wird damit, ob neben dem Glauben, dass Intelligenz *stabil* (bzw. *veränderbar*) ist, auch der Glaube eine Rolle spielt, dass Intelligenz für die Erbringung von kognitiven Leistungen überhaupt *notwendig* bzw. *förderlich* ist (oder nicht). Spinath (2001) konnte nämlich zeigen, dass beide Faktoren eine Rolle spielen können. Aus ökonomischen Gründen werden allerdings – unter inhaltlichen Auswahlkriterien – lediglich drei Items der im Original längeren Skala verwendet. Nicht zuletzt aus diesem Grund erhält die Berücksichtigung des Bedeutsamkeitsfaktors eher explorativen Charakter. Die neue Dimension soll im Folgenden „Bedeutsamkeit von Intelligenz“ heißen, während die alte Dimension weiterhin als „Disposition Impliziter Theorien“ bezeichnet werden soll.

Abschließend sei noch kurz darauf eingegangen, weshalb jene beiden Dispositionskonstrukte überhaupt als Dimensionen eines übergeordneten Konstrukts betrachtet werden (vgl. auch Spinath, 2001). Beide Auffassungen haben gemeinsam, dass sie die Wichtigkeit von Intelligenz für das eigene Leben betreffen. Diese Wichtigkeit wird das eine Mal durch das Stabilitätsmerkmal unterstrichen, das andere Mal durch die Eigenschaft, Voraussetzung für kognitive Leistungen zu sein. Auch logisch-semantisch hängen beide Auffassungen eng zusammen. Um Intelligenz als bedeutsam für kognitive Leistungen anzusehen, muss man zunächst einmal anerkennen, dass es so etwas wie Intelligenz (die in der Alltagssprache als eher stabil angesehen wird) überhaupt gibt. Aufgrund dieser Verwandtschaft soll bei der Ergebnisdarstellung (vgl. Abschnitt 6.2) aus Gründen der Übersichtlichkeit der Bedeutsamkeitspol „statisch“ genannt werden, der Unbedeutsamkeitspol „dynamisch“.

Ablauf, Instruktionen und abhängige Maße. Da der Ablauf des Experiments nicht unkompliziert ist, wird in Tabelle 8 die Abfolge der im Folgenden näher zu besprechenden Teilschritte veranschaulicht. Wichtig für das Verständnis ist zunächst, dass die Hypothesen anhand zweier Leistungsaufgaben getestet werden – erstens an der Bearbeitung des Zeitschriftenartikels (nachfolgend „erste Leis-

tungsaufgabe“ genannt) und zweitens an der Bearbeitung der ASK-Aufgaben („zweite Leistungsaufgabe“). Demzufolge werden die Reaktionen auf das Abschneiden bei diesen Aufgaben mittels zweier Fragebögen erhoben, die entsprechend „erster Emotionsfragebogen“ bzw. „zweiter Emotionsfragebogen“ heißen sollen. Näheres wird in den folgenden Absätzen erläutert.

Die beiden Versuchspersonen arbeiteten an zwei nebeneinander stehenden Tischen. Ein Sichtschutz in Form einer Milchglasscheibe zwischen den Tischen stellte sicher, dass ein Blick auf den Fragebogen des anderen unmöglich war. Die Möglichkeit eines Blickkontakts wurde dadurch aber nicht beeinträchtigt. Die Versuchsleiterin – eine sorgfältig eingeführte Studentin höheren Semesters – saß den Probanden in einem Abstand von drei Metern an einem Schreibtisch gegenüber.

Tabelle 8

Überblick über den Ablauf und die erfassten Konstrukte des Experiments (Studie 2)

- (1) Einführung (3 Minuten), dabei erhobene Konstrukte:
 - Disposition Impliziter Theorien
 - Bedeutsamkeit von Intelligenz
- (2) Lesen der Spiegel-Artikel (10 Minuten)
- (3) Erste Leistungsaufgabe: Gemeinsame Beantwortung der Fragen zum Spiegel-Artikel (ca. 10-20 Minuten)
 - dabei konstant gehaltene Faktoren: Leistungsrückmeldung und Leistungssituation
- (4) Bearbeitung des Emotionsfragebogens zur ersten Leistungsaufgabe (5 Minuten); erhobene Konstrukte:
 - subjektive Erfolgseinschätzung (4 Items)
 - Emotion (nach Misserfolg *und* nach Erfolg, 9 Items + 1 offenes Item)
 - Emotionsdarstellung (nach Misserfolg *und* nach Erfolg, 6 Items)
- (5) Zweite Leistungsaufgabe: Bearbeitung der Aufgaben zur Interpretation von Tabellen (ASK) (15 Minuten)
- (6) Auswertung und standardisierte Rückmeldung von Misserfolg und Erfolg (3 Minuten)
 - dabei variiertes Faktor: Leistungssituation (individualistisch vs. kollektivistisch)
- (7) Bearbeitung des Emotionsfragebogens zur zweiten Leistungsaufgabe (5 Minuten); erhobene Konstrukte:
 - subjektive Leistungseinschätzung (3 Items)
 - Emotionen (nach Misserfolg *oder* nach Erfolg, jeweils 10 Items + 1 offenes Item)
 - Attribution (nach Misserfolg *oder* nach Erfolg, jeweils 5 Items)
 - Emotionsdarstellung (nach Misserfolg *oder* nach Erfolg, 6 bzw. 7 Items)
- (8) Erhebung des Bewältigungsverhaltens (1 Minute)
- (9) Aufklärung der Probanden über den wahren Untersuchungszweck und die Täuschungen (2 Minuten)

Der Gegenstand des Versuchs wurde den Probanden gegenüber als „Zusammenhang zwischen Leistung und Emotion in Gruppen“ ausgegeben. Er bestehe aus zwei Leistungsaufgaben, die nacheinander zu bearbeiten seien. Bei der ersten Leistungsaufgabe gehe es um das Verständnis eines Zeitschriftenartikels. Vor dem Lesen des Artikels sollten die Teilnehmer angeben, „wie Sie selbst über das Thema des Artikels denken“. Ihnen wurden die sechs Aussagen zur Messung der Disposition Impliziter Theorien sowie der Bedeutsamkeit von Intelligenz zur Einschätzung vorgelegt.

Beide Probanden erhielten jeweils dieselbe Variante des Artikels. Sie hatten nun zwölf Minuten Zeit, um den zweiseitigen Artikel zu lesen. Nach Ablauf dieser Zeit forderte die Versuchsleiterin die Probanden (ohne dass dies vorher angekündigt worden war) auf, sich gemeinsam an einen Tisch zu setzen und einige Fragen zum Artikel auf einem gemeinsamen Antwortblatt zu beantworten (*erste Leistungsaufgabe*). Es handelte sich um vier Fragen über den Inhalt, aber auch die praktische Relevanz des Artikels („Welche Bedeutung haben diese Ergebnisse für die Pädagogik und für die Gesellschaft im allgemeinen?“). Die Probanden sollten, so wurde ihnen weiter mitgeteilt, laut über die Fragen diskutieren und sich nicht daran stören, dass die Diskussion per Videokamera aufgezeichnet wurde (s.u.).

Die kooperative Gestaltung der Bearbeitung der Fragen diente nicht nur der Plausibilisierung des (fingierten) Untersuchungszwecks, sondern auch einer – im Vergleich zur einfachen Inhaltsangabe in Studie 1 – tieferen und nachhaltigeren Beschäftigung mit der Aussage des Artikels. Die häufig lange Dauer der Diskussion über die Fragen – ca. 15-25 Minuten bei vorgeschlagenen 8-10 Minuten – lässt darauf schließen, dass das gewählte Setting in der Tat zur Auseinandersetzung mit dem Artikel motivierte. Die Aufzeichnung der Diskussion auf Videoband sollte explorativ Aufschluss darüber verschaffen, wie sich Personen, die sich von der Botschaft des Artikels in ihren späteren Reaktionen beeinflussen lassen, in ihrem Diskussionsverhalten von den anderen unterscheiden (vgl. Abschnitt 6.2). Die schriftlichen Antworten auf die Fragen zum Artikel wurden keiner Auswertung zugeführt.

Die Probanden sollten – nun wieder getrennt voneinander sitzend – Auskunft darüber geben, wie sie die gemeinsame Beantwortung der Fragen erlebt hätten. Sie erhielten dazu einen dreiseitigen Bogen (*erster Emotionsfragebogen*), der sie über die Einschätzung ihrer Leistung, ihre diesbezüglichen Emotionen und ihre Überlegungen zur Emotionsdarstellung befragte. Die gemeinsame Beantwortung der Fragen wurde somit – vor den Probanden wie auch in der Versuchslogik – als Leistungsaufgabe konstruiert, bei der man gut oder schlecht abschneiden kann. Zunächst wurde der Proband gefragt, ob er oder aber sein Kollege mehr zur Beantwortung beigetragen hat (zunächst im forced-choice-Format, dann in Prozentangaben). Diese Items zielen also auf den individuellen Erfolg oder

Misserfolg bei der Aufgabe ab, hier operationalisiert über den sozialen Vergleich. Dann sollte der Proband die Richtigkeit der gemeinsam gefundenen Antworten einschätzen (zunächst auf einer fünfstufigen Skala, dann unter Angabe der erwarteten Schulnote). Diese Items sollten den kollektiven Erfolg erfassen und bei der Auswertung lediglich als Kontrollvariable fungieren. Anschließend wurden Emotionen anhand eines offenen Items (vgl. Studie 1) sowie einer Auswahl von insgesamt neun Items zu negativer *und* positiver Leistungsrückmeldung erhoben. Aus ökonomischen Gründen wurden Emotionen also nur mit einer Teilmenge der in Abschnitt 4.3 vorgestellten Emotions-Items erhoben; weil kein objektives Erfolgskriterium vorlag, wurden Emotionen zu Misserfolg und zu Erfolg abgefragt. Ebenso wurde nur eine Teilmenge der Emotionsdarstellungs-Items verwendet (insgesamt sechs Items zu: Interesse, Schuld, Stolz, Langeweile, Verstecken von Scham, Verstecken von Freude).

Der erste Emotionsfragebogen dient nicht nur der Plausibilisierung des Untersuchungszwecks (hier: Emotionen bei Gruppenarbeit), sondern auch einer ersten, ökonomisch sparsamen Überprüfung der Hypothesen. Neben den ASK-Aufgaben bietet die Artikel-Aufgabe eine weitere Gelegenheit, die erwarteten Unterschiede in Abhängigkeit des Faktors Implizite Theorie hervorzurufen (der Faktor Leistungssituation wird ja in diesem Kontext auf der kollektivistischen Stufe konstant gehalten). Freilich besteht hier das Problem, dass anders als bei der zweiten Leistungsaufgabe (s.u.) kein objektives Erfolgskriterium vorhanden ist, sondern nur die subjektive Erfolgseinschätzung der Probanden. Immerhin sollen die Resultate des ersten Emotionsfragebogens aber für diejenigen Hypothesen herangezogen werden, für die der zweite Emotionsfragebogen keine Antworten geben kann. Dies ist vor allem bei der Frage nach den Auswirkungen der Impliziten Theorien und der Emotionen auf das objektive Leistungsergebnis der Fall. Letzteres kann in diesem Kontext nämlich als das Abschneiden in der zweiten Leistungsaufgabe operationalisiert werden, auf die nun näher eingegangen wird.

Die Versuchsteilnehmer erhielten nun Erläuterungen zu Ablauf und Spielregeln der *zweiten Leistungsaufgabe*. Sie wurden mit der Logik der ASK-Aufgaben und der Zeitvorgabe vertraut gemacht. Ihnen standen 15 Minuten für die Bearbeitung zur Verfügung; die Versuchsleiterin gab nach 8 sowie nach 13 Minuten die Zwischenzeit an, wodurch eine sinnvolle Zeiteinteilung ermöglicht werden sollte. Die Zeitvorgabe von 15 Minuten entspricht den ASK-Standards und führt dazu, dass die letzten Aufgaben zuweilen nicht mehr oder nur noch flüchtig bearbeitet werden können. Schließlich wurden die Teilnehmer über die Möglichkeit informiert, durch ein gutes Abschneiden Geld zu verdienen („als kleinen Anreiz, dass Sie auch wirklich Ihr Bestes geben“). Bei 12 von 18 richtigen Lösungen sollten sie 5 €, bei 15 von 18 richtigen Lösungen 10 € erhalten. Ersteres entspreche im Vergleich zu anderen Studierenden einer „guten“, letzte-

res einer „sehr guten“ Leistung. In der individualistischen Bedingung sollten die Probanden gemäß ihrem individuellen Ergebnis entlohnt werden, in der kollektivistischen hingegen gemäß ihrem durchschnittlichen Ergebnis. Dies wurde mit Bezug auf den oben genannten Untersuchungszweck plausibilisiert, wie in der folgenden, beispielhaften Wiedergabe der schriftlichen Instruktion nachvollzogen werden kann:

Da die Bearbeitung der Aufgaben recht mühsam ist, bieten wir Ihnen als kleinen Anreiz, dass Sie auch wirklich Ihr Bestes geben, eine geldliche Belohnung. Da wir uns für Motivation in Gruppensituationen interessieren, erhalten Sie die geldliche Entlohnung allerdings *in Abhängigkeit der Gesamtleistung der Gruppe*.

Die Aufgaben beinhalten 18 richtig-falsch-Ankreuzmöglichkeiten. Sie erhalten

- 5 Euro, wenn Ihre Gruppe im Durchschnitt mindestens 12 richtige Lösungen hat (was im Vergleich zu anderen Studenten einer „guten“ Leistung entspricht).
- 10 Euro, wenn Ihre Gruppe im Durchschnitt mindestens 15 richtige Lösungen hat (was im Vergleich zu anderen Studenten einer „sehr guten“ Leistung entspricht).

Die Versuchsleiterin wertete nun die beiden Antwortbögen aus und unterrichtete die Probanden anschließend über ihre jeweilige individuelle Leistung und die daraus resultierenden geldlichen Belohnungen (die allerdings noch nicht ausbezahlt wurden). Die Rückmeldung erfolgte verbal und in Anwesenheit beider Probanden. Sie wurde auf Video aufgezeichnet (was jedoch keine erneute Aufmerksamkeit der Teilnehmer erregt haben dürfte, da das Videoband seit der ersten Leistungsaufgabe durchlief und auch kein rotes Lämpchen an der Kamera zu sehen war). Der Wortlaut war etwa wie folgt:

Sie haben 14 Punkte erreicht, und Sie haben 8 Punkte erreicht. [Die Ansprache erfolgt jeweils mit Blickkontakt bzw. Hindeuten.] Damit erhalten Sie 5 Euro, und Sie gehen leider leer aus [*individualistisch*]. Damit kommen Sie zusammen auf 11 Punkte und kriegen leider kein Geld [*kollektivistisch*]. Damit Sie einschätzen können, wie gut Sie abgeschnitten haben, veranschaulichen wir Ihr Ergebnis anhand eines Vergleichs zu anderen Studenten, die diese Aufgaben bereits bearbeitet haben. Sie befinden sich mit Ihrem Ergebnis auf Prozentrang 72, d.h. 28% der Studierenden sind besser, und 72% schlechter, und Sie sind auf Prozentrang 19, d.h. 81% der Studierenden sind besser, und 19% sind schlechter.

Die Zuteilung der positiven und negativen Rückmeldung erfolgte nicht randomisiert, sondern richtete sich nach dem zuvor ermittelten objektiven ASK-Ergebnis: Der objektiv bessere Proband erhielt die Erfolgsrückmeldung, der objektiv schlechtere die Misserfolgsrückmeldung (bei gleicher Punktzahl entschied der Zufall). Dies geschah aus Gründen der Glaubwürdigkeit und Realitätsnähe der Rückmeldung. Der Nachteil dieses Vorgehens besteht allerdings darin, dass sich auf diese Weise die Erfolgsstichprobe tendenziell aus leistungsfähigen, die Misserfolgsstichprobe hingegen aus leistungsschwachen Individuen zusammensetzen dürfte. Die damit einhergehende eingeschränkte Generalisierbarkeit der Ergebnisse wird jedoch in Kauf genommen, da eine Bestätigung der Hypothesen

auch innerhalb solcher leistungsabhängiger Subpopulationen von theoretischer und praktischer Relevanz wäre.

Der oben angeführte Wortlaut der Rückmeldung macht deutlich, dass den Probanden sowohl über die absolute Punktzahl als auch über den sozialen Vergleich die Bewertung ihres Ergebnisses nahegelegt wurde. Der eine Proband sollte sein Ergebnis als eindeutig unzureichend bewerten, der andere als gut bis sehr gut, aber immer noch steigerungsfähig. In Anschluss an z.B. Hong et al. (1999) wurde kein eindeutig sehr gutes Resultat rückgemeldet, um die Antwort auf die spätere Frage zum Bewältigungsverhalten, bei der es ja um die Steigerung der eigenen Kompetenz ging, nicht zu sehr nahezu legen. In der kollektivistischen Bedingung sollten die rückgemeldeten Ergebnisse außerdem die Attribution eindeutig auf den schlechteren Probanden legen: Seine 8 Punkte sind es, die den Schnitt so weit herunterziehen, dass die Gruppe knapp an der entlohnungsrelevanten 12-Punkte-Marke scheitert.

Der studentische Teilnehmer, so lässt sich zusammenfassen, wurde in eine soziale Leistungssituation versetzt, die ihn mit schwierigen, seine Fähigkeiten vollständig fordernden und gleichzeitig praxisrelevanten Aufgaben konfrontierte. Ihm wurde anschließend – unter Anwesenheit des Mitstudenten sowie einer Videokamera – eine (so er sich in der Misserfolgsgruppe befand) äußerst negative Leistungsrückmeldung gegeben, die entweder ihn selbst oder sogar beide Teilnehmer gleichzeitig um die mögliche geldliche Entlohnung brachte. Kurz: Die Situation wurde so emotional aktivierend gestaltet, wie es in einem laborhaften Setting überhaupt nur möglich erscheint.

Nach der Leistungsrückmeldung bearbeiteten die Teilnehmer den *zweiten Emotionsfragebogen*. Dieser war – anders als der erste Emotionsfragebogen – ergebnisspezifisch, d.h. der Misserfolgsperson wurde ein Bogen speziell zu Reaktionen nach Misserfolg ausgehändigt, der Erfolgsperson ein Bogen zu Reaktionen nach Erfolg. Den Probanden gegenüber wurde dies wie folgt plausibilisiert: „Da Sie eine Punktzahl erreicht haben, die unter [über] dem Durchschnitt liegt, vermuten wir, dass das Ergebnis für Sie ein Misserfolg [Erfolg] ist. Sie erhalten daher einen Fragebogen, der sich vorwiegend mit negativen [positiven] Gefühlen und Gedanken beschäftigt.“ Um die Richtigkeit dieser Einschätzung zu überprüfen, sollten die Teilnehmer nun Fragen zur Einschätzung ihrer Leistung beantworten. Auf fünfstufigen Skalen sollten sie angeben, wie zufrieden und überrascht sie sind sowie ob sie mit einem besseren oder schlechteren Ergebnis gerechnet hätten. Somit stehen neben dem objektiven Leistungsergebnis weitere Variablen zur Kontrolle der Induktion von Erfolg und Misserfolg zur Verfügung (vgl. Abschnitt 6.2). Die übrigen Items des zweiten Emotionsfragebogens entsprechen den in Studie 1 verwendeten bzw. in Abschnitt 4.3 vorgestellten: ein offenes Emotions-Item, zehn Items zu Emotion, fünf zu Attribution (vgl. Ab-

schnitt 5.1) und sechs (im Misserfolgsfall) bzw. sieben (im Erfolgsfall) zu Emotionsdarstellung.

Bewältigungsverhalten wurde hingegen grundlegend anders operationalisiert als in der ersten Studie. Nach Bearbeitung des zweiten Emotionsfragebogens, an dessen Ende auch Angaben zur Person erfragt wurden (Alter, Geschlecht, Herkunftsland, Studienfach, Semesterzahl, Abiturnote), wurde der Versuch für beendet erklärt. Die Probanden erhielten nun einen informell gestalteten Zettel, dessen wichtigste Sätze wie folgt lauteten:

Aufgaben wie diese kommen häufig im Rahmen von Assessment Centern oder anderen Auswahlinstrumenten vor, d.h. es sind Aufgaben, deren Bewältigung für Sie im Rahmen von Bewerbungssituationen eine Rolle spielen kann. Damit Sie als Versuchsperson auch ein wenig von diesem Versuch profitieren, bieten wir Ihnen an, mit der Versuchsleiterin Ihre Fehler kurz durchzusprechen und ähnliche Aufgaben zu trainieren. Die Teilnahme an diesem „Training“, das ca. 30-40 Minuten dauert, ist selbstverständlich freiwillig.

Die Teilnehmer sollten auf dem Zettel ankreuzen, ob sie diesen „Service für die Versuchspersonen“ in Anspruch nehmen und somit am Training teilnehmen wollten, und, falls dies der Fall war, zwecks Vereinbarung eines Termins eine e-Mail-Adresse oder Telefonnummer angeben. Weder die Terminansprache noch das Training wurde später in die Tat umgesetzt.

Bewältigungsverhalten wird hier in verhaltensnaher und eher schwer verfälschbarer Form erhoben. Zwar besteht prinzipiell immer noch die Möglichkeit, dass Probanden aus Gründen sozialer Erwünschtheit zunächst ihre Bereitschaft zur Teilnahme angeben, um dann am Telefon doch unter Ausreden abzusagen. Doch diese Gefahr dürfte wesentlich geringer sein als bei den Bewältigungs-Items in Studie 1. Nachteilig an dieser Operationalisierung ist zum einen, dass das Item bipolar ist und somit eine hinreichende Variabilität der Werte nicht gesichert ist (vgl. aber die diesbezüglichen Änderungen in Studie 3, Abschnitt 7.1). Zum anderen könnte das Item unbeabsichtigt (auch) die Tendenz der Probanden erfassen, sich die eigenen Fehler aufgrund der Unglaubwürdigkeit der Ergebnisse (und nicht aufgrund des Strebens nach Kompetenzzuwachs) erläutern zu lassen. Eine solche Tendenz, die sich etwa in großer Überraschtheit oder dem Verlangen nach sofortigem Aufzeigen der eigenen Fehler ausdrücken sollte, konnte jedoch – um dies vorwegzunehmen – bei keiner Versuchsperson beobachtet werden. Die ASK-Aufgaben waren offensichtlich hinreichend komplex, um dem entgegenzuwirken.

Abschließend erfolgte die vollständige *Aufklärung* der Versuchspersonen über alle drei wichtigen Täuschungen: den gefälschten Zeitschriftenartikel, die falsche Leistungsrückmeldung und die angebliche Trainingsmöglichkeit. Die Geldbeträge wurden schließlich bezüglich der tatsächlich erreichten ASK-Punktzahl ausgezahlt.

6.2 Ergebnisse

Wie bereits im vorangegangenen Abschnitt angedeutet wurde, kann die Überprüfung der Hypothesen zu Misserfolgsreaktionen wie auch zu Erfolgsreaktionen anhand jeweils zweier (sich teilweise überlappender) Stichproben erfolgen. Für den Misserfallsfall lassen sich aus den 88 Teilnehmern erstens diejenigen herausfiltern, die angeben, bei der Beantwortung der Fragen zum Artikel weniger als der Mitspieler beigetragen zu haben. Es resultieren 41 Personen, die in diesem Sinne Misserfolg bei der ersten Leistungsaufgabe erfahren und deren unmittelbare Reaktionen anhand der Daten des ersten Emotionsfragebogens untersucht werden können. Aufgrund dieser eher schlechten Operationalisierung von Misserfolg soll diese Stichprobe allerdings primär für die Frage nach der Beziehung zwischen Impliziten Theorien und Leistung herangezogen werden, da für diese keine bessere vorhanden ist. Zweitens können diejenigen 44 Probanden untersucht werden, denen ein Misserfolg über die Rückmeldung der geringeren ASK-Punktzahl induziert werden sollte. In analoger Weise lassen sich auch zwei Stichproben für den Erfolgsfall definieren.

Unklar bleibt bei den Stichproben in Abhängigkeit der zweiten Leistungsaufgabe aber zunächst, ob die Probanden die Rückmeldung immer in der gewünschten Weise, d.h. als Misserfolg bzw. als Erfolg aufgefasst haben. Zwei Quellen, die eine erfolgreiche Manipulation der Leistungseinschätzung beeinträchtigt haben könnten, erscheinen hier besonders plausibel.

- *Atypische Erwartungshaltung.* In der Misserfolgsbedingung ist es denkbar, dass ein Proband die Aufgaben als unangemessen schwierig empfindet und seine rückgemeldete Punktzahl nicht als Misserfolg auffasst, weil er eher mit einem noch schlechteren Ergebnis gerechnet hat. Zumindest aber wird er emotional relativ neutral reagieren, weil es ihm vorher klar war, dass er nur geraten hat. In Anbetracht der Tatsache, dass die Leistung in den ASK-Aufgaben im Durchschnitt über alle Probanden gerade einmal 8.9 Punkte ($SD = 2.3$) betrug (und damit also sogar knapp unter der Ratewahrscheinlichkeit liegt!), dürfte dieses Problem nicht selten aufgetreten sein. Die Aufgaben waren offensichtlich etwas zu schwierig für die gewählte Stichprobe. Spiegelbildlich ergibt sich für die Erfolgsbedingung das Problem, dass die rückgemeldeten 14 Punkte eher als Misserfolg denn als Erfolg verstanden werden könnten, nicht zuletzt weil ja im Zuge dessen die höhere geldliche Belohnung (in der individualistischen Bedingung) knapp verfehlt wurde.
- *Unglaublichkeit.* Ein Proband in der Misserfolgsbedingung könnte die Rückmeldung als unglaublich empfinden, da sein Eindruck von seiner Leistung deutlich positiver war. Zwar sind die Aufgaben so gestaltet, dass man sich oftmals nicht hundertprozentig sicher sein kann, dass man sie richtig gelöst hat, aber kompetente Probanden sollten doch die mitunter einge-

bauten Aha-Effekte für den Aufbau einer bestimmten Ergebniserwartung nutzen können. In der Erfolgsbedingung könnte ein Proband die Rückmeldung als unglaublich empfinden, weil er mit einem deutlich schlechteren Abschneiden gerechnet hat. In diesem Fall müsste er sein Ergebnis entweder auf Zufall zurückführen (was plausibel ist, weil es sich um multiple-choice-Aufgaben handelt) oder gar mutmaßen, dass die Rückmeldung vom Versuchsleiter gezielt manipuliert wurde. Die Konsequenz wäre in der Misserfolgs- wie der Erfolgsbedingung dieselbe: Der Proband könnte sich sein Ergebnis nicht als Misserfolg bzw. als Erfolg zurechnen, wodurch die Manipulation misslungen und die Verhaltensreaktion des Probanden im Kontext der vorliegenden Fragestellung unbrauchbar wäre.

Wie lässt sich angesichts dieser Schwierigkeiten der Manipulationserfolg der Leistungsrückmeldung untersuchen? Zunächst bieten sich hier die eigens zu diesem Zweck erhobenen Kontrollvariablen an, die die Probanden nach ihrer ergebnisbezogenen Unzufriedenheit ($M = 2.8$, $SD = 1.3$), Überraschtheit ($M = 2.8$, $SD = 1.3$) sowie der Tendenz, ein schlechteres (vs. besseres) Ergebnis erwartet zu haben ($M = 3.2$, $SD = 1.1$), fragen. Wie an den Standardabweichungen bereits abzulesen ist, weisen diese Items jedoch eine zu geringe Variabilität auf, als dass sie für die angestrebte Stichprobenselektion sinnvoll einzusetzen wären. Die geringe Streuung hat nicht zuletzt damit zu tun, dass es sich jeweils nur um fünfstufige Skalen handelt. Als Alternative bietet sich daher das wesentlich differenzierter gemessene objektive Leistungsergebnis in der zweiten Leistungsaufgabe an, das im Folgenden kurz „Leistung“ genannt werden soll. Dieses Merkmal erscheint auch deshalb geeignet, weil es beide erwähnten Probleme gleichzeitig berücksichtigen kann: Probanden, die die Misserfolgsmeldung nicht als Misserfolg empfinden oder jedenfalls mit einem schlechten Ergebnis gerechnet haben, sollten tendenziell eine besonders schlechte Leistung erbracht haben. Umgekehrt sollten Probanden, die die Richtigkeit der Rückmeldung aufgrund von Aha-Effekten anzweifeln, eine besonders gute Leistung erbracht haben. Entsprechendes gilt für die Erfolgsbedingung: Probanden, die die Erfolgsmeldung als Misserfolg verstehen, sollten eine besonders hohe Leistung erbracht haben, während diejenigen, die die Glaubwürdigkeit der Rückmeldung anzweifeln, eine besonders niedrige Leistung aufweisen sollten. Schließlich kann die Leistungsvariable – besser als die Items zur subjektiven Leistungseinschätzung – Zufälligkeiten des sozialen Vergleichs auffangen, d.h. dass manche Probanden einen besonders guten, andere einen besonders schlechten Partner hatten.

Um die aus den genannten Quellen resultierende Störvarianz zu reduzieren, erschien es daher angebracht, diejenigen Versuchspersonen aus der Analyse auszuschließen, deren Leistung deutlich von dem rückgemeldeten Wert (d.h. 8 oder 14 Punkte) abwich. Diese Form der Stichprobenselektion wurde auch in anderen

Studien über Reaktionen auf Leistungsrückmeldungen angewandt (vgl. z.B. Kessels, Warner, Holle, & Hannover, 2008). Die Angemessenheit dieser Vorgehensweise wurde außerdem durch explorative Datenanalysen nahegelegt, wonach in den Extrembereichen des Leistungsmerkmals extrem inkohärente und zufällig erscheinende Ergebnisse bezüglich der Hauptfragestellung anzutreffen sind, während Probanden im mittleren Leistungsbereich ein recht kohärentes Antwortmuster aufweisen.

Aus pragmatischen Erwägungen heraus musste nun noch ein Kriterium festgelegt werden, das bestimmt, wie deutlich ein Proband in seiner Leistung vom rückgemeldeten Wert abweichen muss, damit er ausgeschlossen wird. Dabei war insbesondere darauf zu achten, dass die Anzahl der auszuschließenden Probanden nicht zu groß wurde. Für den Misserfolgsfall wurde dieses Kriterium daher auf eine Abweichung von mehr als zwei Punkten (entspricht einer Standardabweichung) festgesetzt, für den Erfolgsfall (wo eine „schiefe“ Werteverteilung vorliegt, vgl. Tabelle 9) auf eine Abweichung von mehr als fünf Punkten. Tabelle 9 veranschaulicht, was dies für die jeweiligen Stichprobengrößen bedeutet: In der Misserfolgsbedingung wurden sieben besonders „schlechte“ und vier besonders „gute“ Fälle ausgeschlossen, in der Erfolgsbedingung zehn besonders „schlechte“ Fälle und kein einziger besonders „guter“ Fall (dieser war erstens aufgrund des gewählten Kriteriums theoretisch nicht möglich und zweitens auch empirisch nicht denkbar, weil es keinen einzigen Probanden gab, der mehr als die rückgemeldeten 14 Punkte erreichte). Die Stichprobengrößen verringerten sich daher von 44 auf 33 bzw. 34 Probanden. Die Gruppen in Abhängigkeit der experimentell variierten Faktoren blieben annähernd gleich groß.

Tabelle 9

Ergebnisse der Stichprobenselektion anhand der objektiv erzielten Leistungsergebnisse

	Misserfolgsbedingung			Erfolgsbedingung		
Rohpunktwert Leistung	2-5	6-10	11-18	5-8	9-14	15-18
Anzahl Probanden	7	33	4	10	34	0
Berücksichtigung	nein	ja	nein	nein	ja	--

6.2.1 Reaktionen auf negative Leistungsrückmeldungen

Zur Überprüfung der Hypothesen stehen, wie oben erläutert, zwei Stichproben zur Verfügung ($N = 33$ und $N = 41$), die sich teilweise (nämlich mit $n = 14$) überlappen. Die folgenden Ausführungen beziehen sich, soweit nicht anders vermerkt, auf die Stichprobe, die in Abhängigkeit der zweiten Leistungsaufgabe ge-

bildet wurde ($N = 33$). Angesichts einer solch vergleichsweise kleinen Stichprobe stellt sich zunächst die Frage, ob sich die Gruppen vielleicht von vornherein hinsichtlich wichtiger Merkmale unterscheiden. Als relevante Merkmale kommen hier die Disposition Impliziter Theorien, Bedeutsamkeit von Intelligenz, Leistung, Abiturnote und Geschlecht in Betracht. Tabelle 10 zeigt die deskriptiven Ergebnisse solcher Gruppenvergleiche. Dabei wird deutlich, dass Statiker in den beiden Dispositionsmaßen von vornherein eher zur statischen Intelligenzauffassung neigen als Dynamiker. Dieser Unterschied zeigt sich allerdings nur tendenziell, nicht in signifikanter Weise (jeweils $p > .4$ für alle in Tabelle 10 aufgeführten Vergleiche). Auch Leistung, Abiturnote und Geschlecht stehen nicht systematisch mit dem Faktor Implizite Theorie in Beziehung.

Frauen und Männer unterscheiden sich weder hinsichtlich der Disposition Impliziter Theorien ($p > .25$, $\eta^2 = .04$) noch hinsichtlich der Bedeutsamkeit von Intelligenz ($p > .5$, $\eta^2 = .01$). Untersucht man die Verteilung der Geschlechter auf die Induktionsgruppen (Statiker vs. Dynamiker) in Bezug auf die beiden Dispositionsmaße, zeigen sich keine Effekte (zweifaktorielle bivariate Varianzanalyse: für alle Haupt- und Interaktionseffekte gilt $p > .35$, $\eta^2 < .07$). Für eine Deutung dieser Ergebnisse sei auf Abschnitt 6.3 verwiesen.

Tabelle 10

Mittelwerte und Standardabweichungen einiger Kontrollvariablen in Abhängigkeit der experimentellen Faktoren

		Gesamt		individualistisch		kollektivistisch	
		Statiker	Dyn.	Statiker	Dyn.	Statiker	Dyn.
Disposition Impliziter Theorien		3.19 1.15	2.96 1.17	3.33 1.28	3.33 1.30	3.00 1.04	2.63 1.01
Bedeutsamkeit von Intelligenz		3.23 .71	2.98 .93	3.33 .76	3.29 .88	3.10 .66	2.70 .93
Leistung		8.31 1.20	8.18 1.33	8.44 1.01	8.38 1.51	8.14 1.46	8.00 1.22
Abiturnote		2.49 .54	2.67 .70	2.41 .50	2.70 .76	2.60 .62	2.64 .69
Geschlecht	weiblich (Anzahl)	12	13	5	6	7	7
	männlich (Anzahl)	4	4	4	2	0	2

Zur Einschätzung der *Konstruktvalidität* der Emotionen wurden deren korrelative Zusammenhänge untersucht. Tabelle 11 gibt die bivariaten Korrelationen zwischen Merkmalen der Emotion und der Attribution für den Misserfolgsfall

wieder. Die Ergebnisse weisen eine ausgeprägte Ähnlichkeit mit den in Studie 1 vorgefundenen auf: Hoffnungslosigkeit, Scham und Fähigkeitsattribution korrelieren untereinander positiv und zu Zuversicht negativ. Zudem steht Anstrengungsattribution in positiver Beziehung zu Schuld und Zuversicht. Die selbstrelevanten Emotionen Scham und Schuld korrelieren ebenfalls positiv. Diese Ergebnisse entsprechen wiederum weitgehend den theoretisch und empirisch begründeten Erwartungen, wenngleich die Zusammenhänge teilweise nicht so eindeutig und statistisch signifikant ausfallen wie in Studie 1. Einzig Schuld weist überraschende Beziehungen auf, indem sie stärker zu Hoffnungslosigkeit korreliert als zu Zuversicht sowie nicht nur mit Anstrengungsattributionen, sondern auch mit Fähigkeitsattributionen einhergeht.

Aus Tabelle 11 lassen sich außerdem die *internen Konsistenzen* (Cronbachs Alpha) der Emotionskonstrukte entnehmen. Diese bewegen sich alle in einem befriedigenden bis guten Bereich, mit Ausnahme von Schuld. Auch diese Ergebnisse decken sich nahezu vollständig mit den Resultaten aus Studie 1.

Tabelle 11

Korrelative Beziehungen und interne Konsistenzen der Emotions- und Attributionsmaße (Bedingung Misserfolg)

	Scham	Schuld	Zuvers.	Attr.Anstr.	Attr.Fäh.	Cronbachs Alpha
Hoffnungslosigkeit	.72***	.30*	-.46***	-.01	.42**	.83
Scham		.39**	-.27	-.01	.35**	.78
Schuld			.17	.31*	.27	.45
Zuversicht				.28	-.30*	.89
Attr. Anstr.					.08	

Anmerkungen.

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig), * $p < .1$ (zweiseitig).

In Tabelle 12 sind die wichtigsten deskriptiven Ergebnisse von Studie 2 zusammengefasst. Angegeben sind die Mittelwerte und Standardabweichungen der abhängigen Variablen in den verschiedenen experimentellen Gruppen. Diese Ergebnisse werden in den folgenden Absätzen Schritt für Schritt erläutert.

Emotion. In einer multivariaten Varianzanalyse wird zunächst untersucht, ob die vier Emotionen vom Faktor Implizite Theorie abhängen. Es lässt sich kein signifikanter Einfluss nachgewiesen, $F(4,28) = 1.43$, $p = .13$, $\eta^2 = .17$. In univariaten Varianzanalysen zeigen sich hingegen bei zwei Emotionen hypothesen-

Tabelle 12

Deskriptive Ergebnisse (Mittelwerte und Standardabweichungen) in Abhängigkeit Impliziter Theorien unter Berücksichtigung der Leistungssituation, Bedingung Misserfolg

	Gesamt		individualistisch		kollektivistisch	
	Statiker (n=16)	Dyn. (n=17)	Statiker (n=9)	Dyn. (n=8)	Statiker (n=7)	Dyn. (n=9)
Hoffnungslosigkeit	2.40 .87	2.08 .89	2.30 .96	1.96 .68	2.52 .79	2.19 1.07
Scham	2.19 1.05	2.20 .99	1.89 1.13	2.38 .70	2.57 .88	2.04 1.21
Schuld	2.44 1.06	2.97 1.05	2.33 1.27	2.75 1.20	2.57 .79	3.17 .94
Zuversicht	2.38 .87	2.97 1.07	2.28 .87	2.88 1.09	2.50 .91	3.06 1.10
Attribution Anstrengung	1.81 .75	2.06 .90	1.56 .88	2.00 1.07	2.14 .38	2.11 .78
Attribution Fähigkeit	2.75 1.53	3.12 1.11	3.11 1.83	3.38 .92	2.29 .95	2.89 1.27
Darstellung Scham	2.13 1.26	2.65 1.06	2.11 1.54	3.00 1.07	2.14 .90	2.33 1.00
Verstecken Scham	2.31 1.35	2.29 .99	2.00 1.41	2.63 1.19	2.71 1.25	2.00 .71
Darstellung Schuld	2.25 1.24	2.24 1.03	2.22 1.30	2.00 .76	2.29 1.25	2.44 1.24
Verstecken Schuld	2.31 1.40	2.53 1.07	2.22 1.64	2.75 1.16	2.43 1.13	2.33 1.00
Darstellung Interesse	1.88 .81	2.41 1.33	1.78 .97	2.25 1.16	2.00 .58	2.56 1.51
Darstellung Langeweile	1.94 1.06	1.53 .72	1.44 .73	1.75 .71	2.57 1.13	1.33 .71
Bewältigungsverhalten (Anteil der Bewältiger)	38%	71%	22%	75%	57%	67%

getreue Unterschiede: Dynamiker berichten tendenziell mehr über Schuld, $F(1,31) = 2.09$, $p < .1$, $\eta^2 = .06$, und Zuversicht, $F(1,31) = 3.07$, $p < .05$, $\eta^2 = .09$, als Statiker. (In Anbetracht der kleinen Stichprobe und der zu erwartenden geringen Unterschiede zwischen den Gruppen schildere ich hier und im Folgenden auch Signifikanzen auf 10%-Niveau.) Hoffnungslosigkeit wird nur der Ergebnisrichtung nach eher von Statikern berichtet, $F(1,31) = 1.08$, $p = .15$, $\eta^2 = .03$, während die Schamwerte für beide Gruppen nahezu identisch ausfallen. Berücksichtigt man Leistungssituation als zweiten Faktor in jenen univariaten Varianz-

analysen, sollten sich auf diesen Haupteffekte für die Merkmale Scham und Schuld zeigen. Dies ist nicht der Fall. Verwendet man stattdessen das Geschlecht der Probanden als zweiten Faktor, ergeben sich erwartungsgemäß Interaktionseffekte, z.B. für die multivariate Varianzanalyse mit allen vier Emotionen: $F(4,26) = 5.52$, $p < .005$, $\eta^2 = .46$. Tabelle 13 veranschaulicht, dass in der Gruppe der Frauen ($n = 25$) die Hypothesen für Hoffnungslosigkeit, Scham und Zuversicht deutlich klarer bestätigt werden können als in der Gesamtstichprobe ($N = 33$). Interessanterweise zeigt sich bei jener zweifaktoriellen multivariaten Varianzanalyse auch ein schwach signifikanter Haupteffekt für Implizite Theorien ($p < .1$, $\eta^2 = .23$). Analog zu Studie 1 (vgl. Abschnitt 5.2) erbrachte die kategorisierende Auswertung des offenen Emotions-Items keine interpretierbaren Ergebnisse.

Wie aus Tabelle 13 weiter hervorgeht, zeigen sich die erwarteten Unterschiede zwischen Statikern und Dynamikern teilweise auch, wenn man nicht die induzierten, sondern die dispositionalen Impliziten Theorien (in Tabelle 13 medianengesplittet *dargestellt*) betrachtet. Für Hoffnungslosigkeit und Scham lassen sich die Unterschiede inferenzstatistisch absichern. Interaktionen zwischen einem Dispositionsmaß und Geschlecht konnten für keine Emotion nachgewiesen werden (jeweils $p > .3$). Zu berücksichtigen ist bei diesen Ergebnissen, dass anders als bei vielen Untersuchungen der Arbeitsgruppe um Dweck die Disposition Impliziter Theorien nicht durch Gegenüberstellung zweier Extremgruppen (extreme Statiker vs. extreme Dynamiker, bei Ausklammern der Unentschlossenen), sondern in ihrer ursprünglichen, metrisch skalierten Form analysiert wurde. Dies geschah nicht nur aufgrund der Tatsache, dass sich die Gruppengrößen aus kulturellen Gründen deutlich unterscheiden würden (vgl. Abschnitt 5.2.1), sondern auch aus methodischen Überlegungen heraus, die das Umwandeln metrischer in kategoriale Größen kritisch hinterfragen.

Der Einsatz der Dispositionsmaße als Kontrollvariablen kann, anders als erwartet, die Effekte in Abhängigkeit Impliziter Theorien nicht vergrößern. Auch können – unter Verwendung eines regressionsanalytischen Ansatzes (Richter, 2007; als methodisch adäquatere Alternative zur Medianhalbierung) – keine Interaktionen zwischen induzierter und dispositionaler Impliziter Theorie festgestellt werden ($p > .7$), d.h. die anfängliche Disposition scheint für die Frage, ob jemand sich in seinen emotionalen Reaktionen von der Induktion beeinflussen lässt, keine Rolle zu spielen. Vom statischen Induktionstext lassen sich demnach z.B. extrem statisch disponierte Probanden ebenso beeinflussen wie extrem dynamisch disponierte Probanden.

Tabelle 13

Emotionen in Abhängigkeit Impliziter Intelligenztheorien (unter Berücksichtigung von Geschlecht) sowie der beiden Dispositionsmaße

		Hoffnungs- losigkeit	Scham	Schuld	Zuversicht
Induzierte Implizite Theorien (N=33)	Statiker	2.40	2.19	2.44	2.38
	Dynamiker	2.08	2.20	2.97	2.97
	p			*	**
-- Frauen (n=25)	Statiker	2.42	2.47	2.58	2.29
	Dynamiker	1.82	1.87	3.04	3.35
	p	**	**		***
-- Männer (n=8)	Statiker	2.33	1.33	2.00	2.63
	Dynamiker	2.92	3.25	2.75	1.75
	p				
Disposition Impliziter Theorien (N=33)	Statiker	2.54	2.49	2.58	2.46
	Dynamiker	2.03	2.00	2.80	2.83
	p	*	**		
Bedeutsamkeit von Intelligenz (N=33)	Statiker	2.52	2.57	2.78	2.50
	Dynamiker	1.89	1.73	2.63	2.90
	p	**	***		

Anmerkungen.

Die Standardabweichung liegt jeweils bei etwa 1 (in einem Bereich zwischen 0.7 und 1.3). Die Signifikanzangaben betreffen t-Tests, für die Dispositionsmaße lineare Regressionen.

*** $p < .01$, ** $p < .05$, * $p < .1$.

Nun soll auf die Ergebnisse in Abhängigkeit der anderen, durch Beobachtung vorgenommenen Operationalisierung der Emotionen eingegangen werden. Wie oben bereits erwähnt, wurde die Rückmeldungssituation auf Video aufgezeichnet. Die in dieser kurzen Sequenz (ca. 45-60 Sekunden) gezeigten Emotionen des Misserfolgsprobanden wurden später von zwei Personen eingeschätzt. Die Beurteiler waren Psychologie-Dozenten der Universität Kassel mit langjähriger Erfahrung in emotionspsychologischer Forschung. Sie sollten ohne wechselseitige Kommunikation das Verhalten des Probanden hinsichtlich folgender Merkmale auf fünfstufigen Skalen einschätzen: Hoffnungslosigkeit (Interrater-Reliabilität: $ICC_{just,MW} = .69$), Scham ($ICC_{just,MW} = .57$), Zuversicht/Gelassenheit ($ICC_{just,MW} = .20$). Die Werte beider Beurteiler wurden zu gemeinsamen Scores zusammengefasst. Während für Hoffnungslosigkeit und Zuversicht/Gelassenheit keine Unterschiede hinsichtlich Impliziter Theorien festgestellt werden können, zeigen sich für Scham immerhin schwache Tendenzen in hypothesengetreuer Richtung, $p = .13$, $\eta^2 = .04$. Bezieht man nun diese Operationalisierung von Scham statt der Fragebogen-Scham in die oben erläuterte einfaktorielle multivariate Varianzanalyse mit ein, wird diese schwach signifikant, $p < .1$, $\eta^2 = .28$. Die

erwarteten Unterschiede auf dem Faktor Leistungssituation zeigen sich für beobachtete Scham nicht. Beide Schamwerte korrelieren überraschenderweise kaum miteinander ($r = .05$, ns).

Schließlich soll es noch kurz um die emotionalen Reaktionen auf die erste Leistungsaufgabe ($N = 41$) gehen. Hier zeigt sich, dass Statiker ($M = 2.2$) erwartungsgemäß stärker als Dynamiker ($M = 1.5$) zu Hoffnungslosigkeit tendieren, $F(1,39) = 4.92$, $p < .05$, $\eta^2 = .11$. Bei den anderen drei Emotionen zeigen sich keine hypothesenkonformen Ergebnisse, wenngleich sich für Zuversicht eine richtige Ergebnistendenz andeutet, $M = 3.4$ zu $M = 3.7$, $p = .12$.

Attribution. Vorliegende Studien ließen erwarten, dass Anstrengungsattributionen stärker von Dynamikern vorgenommen werden als von Statikern. Überraschenderweise findet sich dazu in den Daten jedoch nur eine bescheidene Tendenz (vgl. Tabelle 12), die nicht signifikant wird. Auch die korrelativen Beziehungen zu den Emotionen sind nicht durchweg hypothesenkonform: Anstrengungsattribution hängt zwar erwartungsgemäß positiv mit Schuld und Zuversicht zusammen, ist zu Hoffnungslosigkeit und Scham aber unkorreliert (vgl. oben, Tabelle 11). Für Fähigkeitsattributionen zeigten sich ebenfalls keine Unterschiede zwischen Statikern und Dynamikern. Hier war allerdings aufgrund empirischer Vorerfahrungen (vgl. Abschnitt 3.1) auch keine Hypothese formuliert worden.

Bewältigungsverhalten. Tabelle 14 zeigt, dass die induzierten Impliziten Theorien erwartungsgemäß das Bewältigungsverhalten bei Misserfolg beeinflussen: Dynamiker tendieren dazu, das Trainingsangebot anzunehmen, während Statiker es eher ablehnen, $\chi^2(1, N = 33) = 3.64$, $p < .05$ (bei Verwenden von Yates' Kontinuitätskorrektur ergibt sich χ^2 zu 2.42, $p < .1$).

Aus Tabelle 14 ist auch der Vergleich zum Verhalten der Probanden nach positiver Rückmeldung abzulesen. Diese zeigen ein tendenziell entgegengesetztes Bewältigungsverhalten (d.h. Statiker neigen zu Bewältigung, Dynamiker zu Nicht-Bewältigung), was Hong et al.s (1999) Ergebnisse bestätigt. Eine 2x2x2-loglineare Analyse erbrachte dementsprechend auch eine deutliche Dreifach-Interaktion, $\chi^2(1, N = 33) = 4.9$, $p < .05$. (Hierbei wurde die Misserfolgsstichprobe ($N = 33$) zugrunde gelegt, da nur auf diese Weise eine loglineare Analyse durchführbar war; in Abschnitt 6.2.2 werden die Ergebnisse hinsichtlich der Bewältigungsreaktionen nach Erfolg für die hier eigentlich angemessene Erfolgsstichprobe ($N = 34$) vorgestellt, die aber fast identisch ausfallen.)

Interaktionen zwischen Intelligenztheorie und Geschlecht sind – schon allein aufgrund der geringen Gruppengröße ($n = 8$) der Männer – nicht auszumachen.

Während die Disposition Impliziter Theorien (mediangesplittet) keinen erkennbaren Einfluss auf das Bewältigungsverhalten hat, ist dieser für Bedeutsamkeit (mediangesplittet) erstaunlich ausgeprägt, $\chi^2(1, N = 33) = 3.92$, $p < .05$ (Ya-

tes: $\chi^2 = 2.64$, $p < .1$). Dieser Effekt zeigt sich auch, wenn man Bedeutsamkeit als kontinuierliche Variable beibehält und die dortigen Gruppenunterschiede in Abhängigkeit vom Bewältigungsverhalten analysiert, $p < .05$, $\eta^2 = .11$. Kollektivisten sollten entsprechend meinen Hypothesen mehr zum Bewältigen neigen als Individualisten. Dies trifft für die Daten nur der Ergebnisrichtung nach zu.

Tabelle 14

Bewältigungsverhalten in Abhängigkeit von Impliziten Intelligenztheorien, Leistungssituation, Geschlecht sowie den beiden Dispositionsmaßen (Bedingung Misserfolg, soweit nicht anders vermerkt)

Quelle der Varianz		Bewältigungsverhalten		Bewältigungsverhalten, prozentualer Anteil „Ja“
		Nein	Ja	
Implizite Theorien	Statiker	10	6	38
	Dynamiker	5	12	71
Implizite Theorien, Bedingung Erfolg	Statiker	6	11	65
	Dynamiker	9	7	44
Implizite Theorien (Gruppe der Frauen)	Statiker	7	5	42
	Dynamiker	3	10	77
Implizite Theorien (Gruppe der Männer)	Statiker	3	1	25
	Dynamiker	2	2	50
Leistungssituation	Individual.	9	8	47
	Kollektiv.	6	10	63
Disposition Impliziter Theorien	Statiker	5	8	62
	Dynamiker	10	10	50
Bedeutsamkeit von Intelligenz	Statiker	11	7	39
	Dynamiker	4	11	73

Emotionsdarstellung. Zwischen den verschiedenen Merkmalen der Emotionsdarstellung bestehen überwiegend positive korrelative Beziehungen (vgl. Tabelle 15). Dabei verwundert insbesondere der deutlich positive Zusammenhang zwischen dem Darstellen und dem Verstecken von Scham. Auch bei den Schuld-Items deutet sich solch eine paradox erscheinende Korrelation an. Einige theoriekonforme Beziehungen sind jedoch auch auszumachen, etwa die zwischen der Darstellung von Langeweile und dem Verstecken von Scham sowie die zwischen der Darstellung von Scham und dem Verstecken von Schuld.

Die Hypothesen zur Darstellung von Interesse und von Langeweile ließen sich vergleichsweise gut bestätigen. Dynamiker neigen etwas stärker als Statiker zur Darstellung von Interesse, $F(1,31) = 1.94$, $p < .1$, $\eta^2 = .06$. Dieser Effekt ist in der Gruppe der Frauen besonders stark ausgeprägt, $F(1,23) = 6.99$, $p < .01$, $\eta^2 = .23$. Statiker dagegen stellen eher Langeweile dar als Dynamiker, $F(1,31) = 1.69$, $p < .1$, $\eta^2 = .05$ (Frauen: $p < .1$, $\eta^2 = .09$).

Das Darstellen bzw. Verstecken von Scham und Schuld konnte hingegen weniger gut durch den Faktor Implizite Theorie vorhergesagt werden. Bezieht man

diese vier Merkmale in eine multivariate Varianzanalyse ein, zeigen sich keine Unterschiede, $p > .7$. Auch in univariaten Analysen ergeben sich keine Effekte.

In der kollektivistischen Bedingung sollten die Probanden zu mehr Schamdarstellung und Schuldverstecken tendieren (vgl. Tabelle 12). Auch diese Hypothesen ließen sich durch Varianzanalysen nicht bestätigen. Für die Darstellungsreaktionen auf die erste Leistungsaufgabe ($N = 41$) zeigten sich ebenfalls keine theoriekonformen Ergebnisse.

Tabelle 15

Korrelative Beziehungen der Emotionsdarstellungsmaße (Bedingung Misserfolg)

	D. Interesse	D. Scham	V. Scham	D. Schuld	V. Schuld
Darstellung Interesse	.011	.167	.228	-.005	.202
Darstellung Langeweile		.250	.466***	.036	.415**
Darstellung Scham			.393**	.283	.599***
Verstecken Scham				-.034	.743***
Darstellung Schuld					.105
Verstecken Schuld					

Anmerkungen.

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig).

Prüfung des Gesamtmodells. In diesem Unterkapitel soll das in Abschnitt 4.1 entwickelte Gesamtmodell schrittweise an den Daten getestet werden. Dabei wird zunächst die Beziehung zwischen Emotion und Bewältigungsverhalten untersucht, dann überprüft, inwieweit die vier erhobenen Emotionen als Mediatoren der Beziehung zwischen Impliziten Theorien und Bewältigungsverhalten aufzufassen sind, und schließlich ein solches Modell – unter zusätzlicher Einbeziehung von Leistungssituation und Attribution – in Pfadanalysen geprüft. Dabei wird anhand der Stichprobe der ersten Leistungsaufgabe ($N = 41$) auch untersucht, inwieweit sich aus Impliziter Theorie und Emotionen auf die objektiv erbrachte Leistung schließen lässt.

Zunächst wurden punktbiseriale Korrelationen zwischen den Emotionen und dem Bewältigungsverhalten berechnet. Demnach geht ein hohes Maß an Bewältigungsverhalten erwartungsgemäß mit viel Schuld ($r = .33$, $p < .05$) und Zuvorsicht ($r = .38$, $p < .05$) sowie – der Einflussrichtung nach – mit wenig Hoffnungslosigkeit ($r = -.18$, ns) und Scham ($r = -.13$, ns) einher. Dann wurde in einer *logistischen Regression* untersucht, inwieweit diese vier Merkmale die Auftretenswahrscheinlichkeit des Merkmals Bewältigungsverhalten (ohne Berück-

sichtigung von Interaktionen) *gemeinsam* erklären. Die folgenden Ergebnisse sollten jedoch mit Vorsicht interpretiert werden, weil erstens die in der Literatur empfohlene Mindestanzahl von 50 Beobachtungen hier nicht erreicht wird und zweitens das Problem der Multikollinearität besteht (vgl. Rese, 2005). Der Likelihood-Quotienten-Test ergibt eine befriedigende Modellanpassung, $\chi^2(4, N = 33) = 9.17$, $p = .06$ (zweiseitig). McFaddens R^2 liegt bei .202, was allgemein noch als gute Anpassung gilt (Rese, 2005). Gemäß Nagelkerkes R^2 wird 32% Varianz von den vier Prädiktoren erklärt. Von den einzelnen Prädiktoren sind Schuld, $Exp(B) = 0.40$, $p < .05$, und Zuversicht, $Exp(B) = 0.52$, $p = .11$, bedeutender als Hoffnungslosigkeit und Scham.

Ersetzt man in diesem Modell Scham durch die mittels Beobachtung gewonnene Schamvariable, wird die Modellanpassung deutlich besser, $\chi^2(4, N = 33) = 11.31$, $p = .02$, McFaddens $R^2 = .248$, 38% Varianzaufklärung. Dabei erweisen sich nun neben Schuld, $Exp(B) = 0.51$, $p < .1$, und Zuversicht, $Exp(B) = 0.41$, $p < .1$, auch die neue Schamvariable, $Exp(B) = 0.24$, $p < .1$, als einflussreiche Prädiktoren. Ins Korrelationsmuster hingegen will sich die neue Variable nicht „richtig“ einordnen, d.h. nicht so wie die Fragebogen-Scham. Mit allen vier Fragebogen-Emotionen korreliert sie etwa zu Null, mit dem Bewältigungsverhalten besteht eine positive Korrelation, $r = .27$, $p = .13$.

In einer *Mediatoranalyse* nach Baron und Kenny (1986, hier angewandt auf logistische Regressionen, vgl. MacKinnon & Dwyer, 1993) wurde nun untersucht, ob sich die vier Emotionen als Mediatoren der Beziehung zwischen Impliziten Theorien und Bewältigungsverhalten auffassen lassen. Baron und Kenny stellten vier Kriterien für das Vorhandensein einer Mediation auf: Der Prädiktor sollte erstens mit dem Kriterium und zweitens mit dem (potenziellen) Mediator zusammenhängen. Bei simultaner Berücksichtigung von Prädiktor und Mediator sollte drittens der Mediator mit dem Kriterium in Beziehung stehen und viertens der Zusammenhang zwischen Prädiktor und Kriterium verschwinden. Ist die vierte Bedingung nicht oder nur teilweise erfüllt, spricht man von partieller Mediation. Gemäß diesen Kriterien lassen sich Zuversicht und Schuld als partielle Mediatoren auffassen: Die (statisch gepolte) Intelligenztheorie beeinflusst das Bewältigungsverhalten ($\beta = -.36$, $p < .05$) sowie Schuld ($\beta = -.25$, $p < .1$) bzw. Zuversicht ($\beta = -.30$, $p < .05$); bei Kontrolle des Einflusses von Schuld ($\beta = .31$, $p < .1$) bzw. Zuversicht ($\beta = .37$, $p < .05$) wird der Einfluss der Intelligenztheorie (für Schuld: $\beta = -.29$, $p < .1$; für Zuversicht: $\beta = -.26$, $p < .1$) auf das Bewältigungsverhalten jedoch geringer. Etwas deutlicher werden diese Beziehungen, wenn man Schuld und Zuversicht (aufgrund ihrer identischen erwarteten Beziehungen zu Impliziten Theorien und Bewältigungsverhalten) in ein übergeordnetes, „anstrengungsorientiertes“ Emotionskonstrukt zusammenfasst: Hier beeinflusst nun die Intelligenztheorie das Bewältigungsverhalten (wieder: $\beta = -.36$, $p < .05$) sowie Schuld/Zuversicht ($\beta = -.36$, $p < .05$); der Einfluss der Intelligenz-

theorie ($\beta = -.19$, *ns*) auf das Bewältigungsverhalten verschwindet jedoch bei gleichzeitiger Berücksichtigung von Schuld/Zuversicht ($\beta = .46$, $p < .05$). Hoffnungslosigkeit und Scham erfüllen die Mediatorkriterien nicht. Auch die beobachtete Scham kann – aufgrund ihrer zu geringen Abhängigkeit von Impliziten Theorien – nicht als Mediator gelten, wenngleich sich interessante Beziehungen zu den anderen Variablen andeuten: Die statische Intelligenztheorie führt zwar zu *weniger* Bewältigungsverhalten, bedingt aber gleichzeitig tendenziell auch beobachtete Scham, die wiederum mit *mehr* Bewältigungsverhalten verbunden ist (s.o.).

Wie hängen nun Implizite Theorien, Leistungssituation, Attribution, Emotion und Bewältigungsverhalten zusammen, wenn man sie in einem gemeinsamen Modell untersucht? Die diesbezüglichen Hypothesen wurden anhand einer *Pfadanalyse* überprüft (LISREL, Jöreskog & Sörbom, 2001), wobei hier und im Folgenden zu berücksichtigen ist, dass deren Ergebnisse aufgrund der geringen Stichprobengröße nur eingeschränkte Gültigkeit beanspruchen können. Sie liefern aber immerhin eine gute Möglichkeit, die Beziehungen zwischen den Modellvariablen graphisch zu veranschaulichen. Abbildung 4 illustriert die Ergebnisse eines theoretisch abgeleiteten Modells, bei dem die beiden experimentell variierten Faktoren die vier Emotionen bestimmen sollten, die ihrerseits das Bewältigungsverhalten beeinflussen sollten. Die Emotionen wurden wiederum anhand ihrer theoretischen Beziehungen zu den anderen Merkmalen zu zwei übergeordneten Konstrukten zusammengefasst, um die Modellkomplexität zu begrenzen. Zwischen den Impliziten Theorien und Bewältigungsverhalten wurde ein zusätzlicher, direkter Pfad eingeräumt. Wenngleich nicht alle Pfadbeziehungen inferenzstatistisch abgesichert werden können (vgl. Abbildung 4), sind sie zumindest dem Vorzeichen nach hypothesenkonform. Die Anpassung des Gesamtmodells an die Daten kann, wie wiederum Abbildung 4 zu entnehmen ist, als gut bis sehr gut bezeichnet werden (d.h. $RMSEA < 0.06$ und $CFI > 0.9$; vgl. Hu & Bentler, 1999). Die Graphik lässt insbesondere deutlich werden, dass der Pfad von Intelligenztheorien über Zuversicht/Schuld weiter zum Bewältigungsverhalten besonders ausgeprägt ist.

Bezieht man Anstrengungsattribution als Vermittlerin zwischen Impliziten Theorien und Emotionen mit ein, ist die Modellanpassung zwar schlechter, aber immer noch befriedigend (vgl. Abbildung 5). Anstrengungsattribution wird von Dynamikern sowie (unerwarteterweise) von Kollektivist*innen bevorzugt und führt ihrerseits zu Zuversicht/Schuld, die wiederum Bewältigungsverhalten wahrscheinlicher machen. Die Beziehungen zwischen Anstrengungsattribution und Hoffnungslosigkeit/Scham und weiter zum Bewältigungsverhalten sind jedoch nicht signifikant und im ersten Fall sogar unerwartet in der Einflussrichtung.

120 STUDIE 2: PRÜFUNG DES MODELLS ANHAND EINER LEISTUNGSAUFGABE

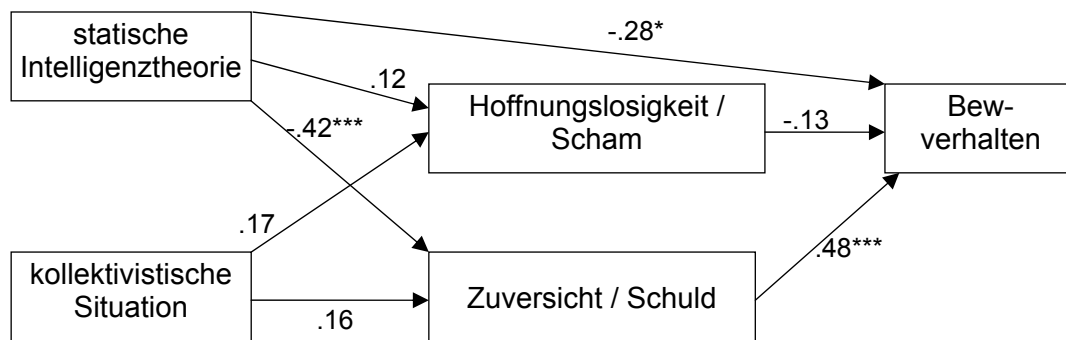


Abbildung 4. Pfadanalytisches Modell zur Prüfung des Gesamtmodells ($N = 33$).

Die unabhängigen Variablen sind gepolt dargestellt (z.B. auf „statische Intelligenztheorie“), beinhalten aber selbstverständlich jeweils beide Ausprägungen. Dies gilt auch für die Abbildungen 5 und 6.

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig), * $p < .1$ (zweiseitig).

Modellanpassung: χ^2 ($df = 2$, $N = 33$) = 1.28, $p = .53$; $RMSEA = 0.00$; $NFI = 0.96$; $CFI = 1.00$.

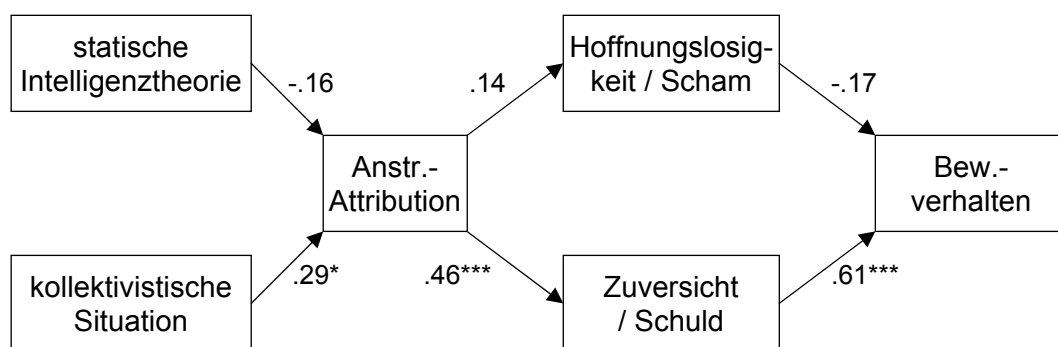


Abbildung 5. Pfadanalytisches Modell zur Prüfung des Gesamtmodells ($N = 33$), inklusive Attribution.

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig), * $p < .1$ (zweiseitig).

Modellanpassung: χ^2 ($df = 8$, $N = 33$) = 9.29, $p = .32$; $RMSEA = 0.07$; $NFI = 0.72$; $CFI = 0.87$.

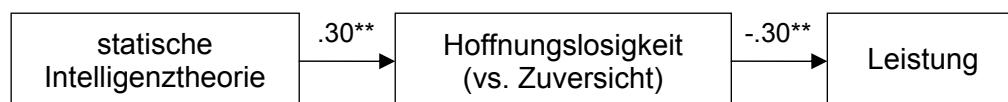


Abbildung 6. Pfadanalytisches Modell zum Einfluss Impliziter Theorien auf Emotion und Leistung ($N = 41$).

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig), * $p < .1$ (zweiseitig).

Modellanpassung: χ^2 ($df = 1$, $N = 41$) = 0.09, $p = .76$; $RMSEA = 0.00$; $NFI = 0.99$; $CFI = 1.00$.

Abschließend kann mit Hilfe der Stichprobe von $N = 41$ (d.h. diejenigen, die nach der ersten Leistungsaufgabe Misserfolg erfahren, s.o.) untersucht werden, ob Implizite Theorien vermittelt über Emotionen auch die Leistung beeinflussen. Gefragt wird hierbei also, ob Statiker nach Misserfolg infolge ihres speziellen Emotionsmusters in einer nachfolgenden Aufgabe (bei der diese Emotionen gewissermaßen „nachwirken“ sollten) weniger Leistung erbringen als Dynamiker. Weil zu vermuten war, dass die selbstrelevanten Emotionen Scham und Schuld eher ambivalente Implikationen für die Leistung aufweisen (vgl. Miller, 2007), wurden hier lediglich die zukunftsbezogenen Emotionen Hoffnungslosigkeit und Zuversicht berücksichtigt. Da diese jeweils nur mit einem einzigen Item gemessen wurden, wurde zur Verringerung der Fehlervarianz ein gemeinsamer Score berechnet (Hoffnungslosigkeit minus Zuversicht, zuvor jeweils z-standardisiert), der im Folgenden Hoffnungslosigkeit heißen soll. Statiker tendieren hier mehr zu Hoffnungslosigkeit als Dynamiker, $F(1,39) = 3.99$, $p < .05$, $\eta^2 = .09$; Hoffnungslosigkeit korreliert ihrerseits mit Leistung negativ ($r = -.25$, $p < .1$), d.h. Hoffnungslosigkeit im Anschluss an die Artikel-Aufgabe geht mit schlechterer Leistung in der ASK-Aufgabe einher. Diese Ergebnisse werden durch eine Pfadanalyse (LISREL) gestützt, deren Ergebnisse und – wiederum ordentlichen – Anpassungswerte in Abbildung 6 zusammengefasst sind. Zu beachten ist jedoch, dass zwischen Impliziten Theorien und Leistung *kein* direkter Zusammenhang besteht ($M = 9.14$ für die Dynamiker zu $M = 9.00$ für die Statiker, $SD = 2.2$). Die statische Intelligenztheorie wirkt sich also *nur in dem Maße* negativ auf die Leistung aus, in dem die Person mit hoffnungslosen bzw. wenig zuversichtlichen Emotionen auf den vorhergehenden Misserfolg reagiert. Folglich muss die statische Intelligenztheorie über einen anderen (unbekannten) Pfad gleichzeitig positiv auf die Leistung einwirken. Bemerkenswert ist zudem, dass die zusätzliche Einbeziehung der Abiturnote der Teilnehmer in das Pfadmodell keinen zusätzlichen Vorhersagebeitrag für die Leistung erbringt, $\beta = -.09$, *ns* (obwohl sie erwartungsgemäß zur Leistung korreliert ist, $r = -.28$, $p < .05$).

Explorative Analysen zu motivationalen Tendenzen. Beobachtete Scham lässt sich aufgrund der beschriebenen Berechnungen als interessante „fünfte Emotion“ verstehen. Einerseits wird sie von Impliziten Theorien in tendenziell hypothesenkonformer Weise beeinflusst, andererseits ist sie nicht mit weniger, sondern mit weniger Bewältigungsverhalten verbunden. Allem Anschein nach wird mit diesem Merkmal eine „produktive“ Form der Scham gemessen, eine Schamkomponente, die nicht zu Rückzug und Verstecken führt, sondern eher zu vermehrter Anstrengung „anstachelt“. Zur näheren Erkundung dieser Deutungsmöglichkeit wurde eine Hauptkomponentenanalyse durchgeführt, die neben beobachteter Scham die Merkmale Hoffnungslosigkeit, Scham, Anstrengungsattribution und Bewältigungsmotivation beinhaltete. Schuld und Zuversicht wurden

nicht berücksichtigt, da diese, wie sich explorativ gezeigt hatte, die Faktorstruktur zu eindeutig bestimmt und somit die Zusammenhangsmuster der übrigen Variablen verdeckt hätten.

Tabelle 16 zeigt die Faktorladungen der ersten vier extrahierten Faktoren (Faktor 4 wurde, trotz seines vergleichsweise geringen Eigenwerts, aus Interpretationsgründen in die Darstellung miteinbezogen). Faktor 1 lädt hoch positiv auf Hoffnungslosigkeit und Scham, hoch negativ auf Bewältigungsverhalten. Faktor 2 hingegen lädt positiv auf beobachteter Scham und Bewältigungsverhalten, negativ auf Anstrengungsattribution. Diese beiden Faktoren deuten also in der Tat auf zwei unterschiedliche motivationale Auswirkungen von Scham hin: eine motivationshemmende in Faktor 1, die hier mit dem Label „Resignation“ versehen werden soll, und eine positive in Faktor 2, die im Folgenden als „Anstachelung“ bezeichnet wird. Die Interpretation der anderen beiden Faktoren ist etwas schwieriger, doch auch sie lassen sich als motivationale Tendenzen deuten: Faktor 3 wird vor allem durch Anstrengungsattribution und Bewältigungsverhalten beschrieben und könnte als „Kompetenzbedürfnis“ betitelt werden, während Faktor 4 durch hohe Anstrengungsattribution sowie geringe Bewältigungsmotivation gekennzeichnet ist und durch so etwas wie „Gleichgültigkeit“ bestimmt sein könnte: Im einen Fall erscheint der Person eine zukünftige Leistungssteigerung sowohl möglich als auch erstrebenswert, im anderen Fall hingegen trifft nur ersteres zu. Beobachtete Scham allerdings lädt auf Faktor 4 etwa gleich hoch positiv wie Anstrengungsattribution, was die Interpretation dieses Faktors erschwert.

Tabelle 16

Explorative Hauptkomponentenanalyse der Schamvariablen und verwandter Konstrukte

Einbezogene Merkmale	Faktor			
	1	2	3	4
Hoffnungslosigkeit	.90	.08	.19	-.10
Scham	.88	.21	.22	-.01
beobachtete Scham	-.13	.90	-.01	.40
Attribution Anstrengung	-.11	-.42	.81	.39
Bewältigungsverhalten	-.42	.44	.66	-.44
Eigenwert	1.8	1.2	1.2	0.5
erklärte Varianz	36%	25%	24%	10%
Interpretation	Resignation	Anstachelung	Kompetenz- bedürfnis	Gleichgültig- keit

Für die Deutung der Faktoren als motivationale Tendenzen (die übrigens in Abschnitt 6.3 näher ausgeführt wird) spricht auch die Tatsache, dass alle vier Faktoren auf Bewältigungsverhalten relativ hoch laden, d.h. alle vier Faktoren lassen sich durch dieses motivationsnahe Konstrukt recht gut beschreiben. Jeder der vier Faktoren lässt sich zudem durch das Merkmal Bewältigungsverhalten in signifikanter Weise trennen (t-Tests: jeweils $p < .02$).

Merkmale der Induktionssituation. Die Bearbeitung der ersten Leistungsaufgabe wurde, wie oben beschrieben, auf Video aufgezeichnet. Damit bietet sich die Gelegenheit, in explorativer Form zu untersuchen, inwieweit eher qualitative Merkmale der Induktion die Wirkungsweise des Faktors Intelligenztheorie beeinflussen: In welcher Form machen sich die Probanden den Spiegel-Artikel verständlich, inwieweit eignen sie sich dessen Aussage an? Derartige Fragestellungen verlangen eigentlich nach einer sorgfältigen qualitativen Analyse der Video-Sequenzen. Aus ökonomischen Gründen konnte eine solch aufwändige Auswertung jedoch nicht realisiert werden. Stattdessen wurde eine deutlich rudimentärere Analysemethode gewählt, die immerhin erste Einblicke sowie Anregungen für künftige Untersuchungen geben sollte. Die Video-Sequenzen der 33 Misserfolgs-Probanden gemäß der ASK-Aufgabe wurden dabei vom Verfasser anhand folgender Kriterien auf siebenstufigen Rating-Skalen eingeschätzt:

- *Glaubwürdigkeit.* Als wie glaubwürdig bzw. seriös wird der Artikel vom Misserfolgsprobanden betrachtet? Wird die Hauptaussage eher „für bare Münze genommen“ und überrascht und erstaunt darauf reagiert, oder herrscht eher Skepsis vor („ich hab’ da auch schon mal was anderes zu gelesen“, „der Spiegel übertreibt immer mal gern“).
- *verbale Bewertung.* Schätzt der Misserfolgsproband die Implikationen des Artikels für sich selbst eher positiv oder negativ ein? Scheint ihm die Botschaft zu gefallen oder zu missfallen? In dieses Merkmal geht nur die verbale Kommunikationsform ein, während der nonverbale Aspekt wie folgt erhoben wird:
- *nonverbale Bewertung.* In welcher Stimmung beantwortet der Misserfolgsproband die Fragen? Wie scheint sich die Botschaft des Artikels auf seine Stimmung (Gestik, Mimik) niederzuschlagen? Theoretisch soll hiermit die nonverbale Bewertung des Artikels erhoben werden; weil es aber kein Vergleichsmaß zur Stimmung vor Versuchsbeginn gibt, ist mit der Einschätzung dieses Merkmals ein beträchtliches Maß an Spekulation verbunden.
- *Kritik.* Sieht der Misserfolgsproband die Aussage des Artikels als moralisch fragwürdig an? Hier geht es also nicht um Implikationen für die eigene Person (wie bei den beiden Bewertungsmerkmalen), sondern um allgemeinere, die ganze Gesellschaft betreffende Konsequenzen (z.B. „das spaltet die Gesellschaft“, „das demotiviert“, „dann ist ja jeder an seinem Misserfolg selbst schuld“).

Anders als bei der Einschätzung der Rückmeldungssequenzen war es hier freilich nicht möglich, die Gruppenzugehörigkeit der Probanden für den Beurteiler unsichtbar zu machen, da diese ja direkt über die Wortbeiträge explizit wurde. Erwartungseffekte sind daher nicht auszuschließen.

In Übereinstimmung mit theoretischen und alltagsweltlichen Erwartungen wird der dynamische Text deutlich positiver bewertet als der statische Text, und

dies sowohl in verbaler Form ($M = 5.2$ zu 3.3 , $SDs \approx 1.0$), als auch in nonverbaler Form ($M = 5.0$ zu 3.1 , $SDs \approx 1.3$). Er erfährt auch deutlich weniger Kritik ($M = 1.5$ zu 3.1 , $SDs \approx 1.7$). Beide Texte werden fast durchgängig als glaubwürdig empfunden ($M = 5.5$), wobei dies beim dynamischen Text etwas weniger der Fall ist als beim statischen ($M = 5.2$ zu 5.8 , $SDs \approx 1.3$). Als fünftes Merkmal wurde außerdem die objektive Bearbeitungszeit erhoben. Diese schwankte zwischen 8 und 25 Minuten und war für den dynamischen Text etwas länger ($M = 15.1$ zu 13.5 , $SDs \approx 4.1$).

Inwieweit beeinflussen diese Merkmale nun die Wirksamkeit der Induktion? Theoretisch plausibel wäre z.B., dass die Merkmale Glaubwürdigkeit und Bearbeitungszeit als Moderatoren auf die Beziehung zwischen Intelligenztheorie und Emotionen einwirken, d.h. dass die Wirkung der Artikel bei höherer Glaubwürdigkeit und bei längerer Beschäftigung deutlicher zutage tritt als bei der jeweils geringeren Ausprägung. Hierzu wurden diese beiden Merkmale zunächst mediandichotomisiert (aufgrund der geringen Datenqualität sowie des explorativen Charakters der Berechnungen wurde auf aufwändigere Analysen verzichtet). In zweifaktoriellen multivariaten Varianzanalysen der vier untersuchten Emotionen ließen sich (für $N = 33$) diese Hypothesen jedoch nicht bestätigen, da sich zwar hypothesengemäße, aber numerisch schwache Interaktionen zwischen Intelligenztheorie und dem jeweiligen Induktionsmerkmal zeigten (für Glaubwürdigkeit: $p = .13$, $\eta^2 = .18$; für Bearbeitungszeit: $p = .23$, $\eta^2 = .12$).

Für die Merkmale Bewertung und Kritik könnte man hingegen erwarten, dass sie sich unabhängig vom Faktor Implizite Theorie auf die gezeigten Emotionen ($N = 33$) auswirken sollten. So sollte etwa eine positive Bewertung des Artikels eher mit Zuversicht, eine negative eher mit Hoffnungslosigkeit einhergehen, während für das Merkmal Kritik beide Hypothesenrichtungen plausibel erscheinen (z.B. weniger Hoffnungslosigkeit durch kritische Distanzierung oder mehr Hoffnungslosigkeit durch Zynismus). Ein Blick auf die Korrelationsmuster zeigt, dass für die beiden Bewertungskonstrukte die erwarteten Ergebnisrichtungen (Hoffnungslosigkeit: $r = -.11$ und $-.11$; Zuversicht: $r = .15$ und $.19$; jeweils *ns*) vorhanden sind. Kritik an der Botschaft des Artikels geht außerdem der Ergebnisrichtung nach mit mehr Hoffnungslosigkeit ($r = .17$, *ns*) und weniger Zuversicht ($r = -.17$, *ns*) einher. Aus Darstellungsgründen werden erst weiter unten (vgl. Abschnitt 9.1) einige Diskussionssequenzen wörtlich wiedergegeben, die veranschaulichen sollen, welche Art von Kritik die Probanden an den Texten übten. Diese Zitate könnten verstehen helfen, weshalb den obigen Ergebnissen zufolge also tendenziell die „Zynismus-Hypothese“ zuzutreffen scheint.

6.2.2 Reaktionen auf positive Leistungsrückmeldungen

Für die Erfolgsbedingung liegen, wie oben erläutert, ebenfalls zwei sich überlappende Stichproben vor. Im Folgenden werden wiederum primär die Ergebnisse der Stichprobe in Abhängigkeit der zweiten Leistungsaufgabe ($N = 34$) vorgestellt. Dort, wo auch die andere Stichprobe ($N = 47$) in die Analyse einbezogen wurde, wird dies kenntlich gemacht.

Tabelle 17 weist interne Konsistenzen und korrelative Beziehungen der Emotions- und Attributionsmaße aus. Ähnlich wie in Studie 1 sind die internen Konsistenzen als ausreichend bis gut zu bezeichnen, während die Korrelationen wiederum wenig befriedigen können: Die vier Emotionen sowie die beiden Attributionsformen korrelieren überwiegend äußerst positiv zueinander. Konstrukte wie Hybris und Stolz oder auch Anstrengungs- und Fähigkeitsattribution lassen sich demnach kaum sinnvoll voneinander abgrenzen.

Gemäß den oben entwickelten Hypothesen sollten Statiker weniger zu Anstrengungsattributionen und Stolz, mehr zu Hybris und Erleichterung, mehr zu meisterndem Bewältigungsverhalten, mehr zur Darstellung von Hybris sowie mehr zum Verstecken von Freude neigen als Dynamiker. Des Weiteren sollten Kollektivistinnen mehr zu selbstrelevanten Emotionen (Hybris, Stolz), zur Darstellung von Stolz und zum Verbergen von Hybris neigen als Individualisten. Diese Hypothesen konnten überwiegend nicht bestätigt werden; bei nahezu allen Gruppenvergleichen (uni- und multivariate Varianzanalysen, auch mit den Dispositionsmaßen oder Geschlecht als Kovariaten oder als Prädiktoren) konnten keine Unterschiede festgestellt werden. Ich beschränke mich im Folgenden darauf, die wenigen hypothesenkonformen Resultate zu schildern. Die deskriptiven Ergebnisse sind Tabelle 18 zu entnehmen.

Das Bewältigungsverhalten ist zwar in hypothesengemäßer Richtung mit dem Faktor Intelligenztheorie verbunden, indem 11 von 17 Statiker, aber nur 7 von 16 Dynamiker das Trainingsangebot annehmen wollen. Diese Beziehung lässt sich jedoch nicht inferenzstatistisch absichern, $\chi^2(1, N = 33) = 1.50$, *ns* (Yates: $\chi^2 = 0.74$, *ns*).

Die Disposition Impliziter Theorien sagt hypothesengemäß Anstrengungsattribution voraus ($p < .1$), die Bedeutsamkeit von Intelligenz das Merkmal Stolz ($p < .05$). Angesichts der Fülle der nicht bestätigten Hypothesen erscheinen diese Unterschiede aber eher zufällig.

Die Darstellung der Abwesenheit von Freude, um den Mitstudenten nicht „herunterzuziehen“, ist in der kollektivistischen Situation stärker ausgeprägt ($p < .05$, $\eta^2 = .11$). Dieses Ergebnis ist zwar wenig überraschend, weil die kollektivistische Bedingung per definitionem mehr soziale Sensibilität verlangt, aber es deutet immerhin darauf hin, dass die Studierenden sich mit den Items ernsthaft auseinandergesetzt bzw. deren sozial-emotionale „Logik“ verstanden haben.

Tabelle 17

Korrelative Beziehungen und interne Konsistenzen der Emotions- und Attributionsmaße (Bedingung Erfolg)

	Hybris	Stolz	Erleicht.	Attr. Anstr.	Attr. Fäh.	Cronbachs Al- pha
Freude	.67***	.75***	.44**	.31*	.40**	.79
Hybris		.80***	.31*	.44***	.66***	.82
Stolz			.29*	.42**	.53***	.91
Erleichterung				-.04	.15	.74
Attribution Anstrengung					.54***	

Anmerkungen.

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig), * $p < .1$ (zweiseitig).

Hinsichtlich Freude gab es keine festen Hypothesen. Vermutet wurde aber (vgl. Abschnitt 4.3), dass Statiker bei leichten Aufgaben weniger, bei schwierigen Aufgaben mehr Freude empfinden als Dynamiker. Dies wurde anhand einer zweifaktoriellen Varianzanalyse mit den Faktoren Intelligenztheorie und Überraschtheit untersucht, wobei hohe Überraschtheit als Indikator für eine hohe (subjektiv empfundene) Aufgabenschwierigkeit verwendet wurde. Es zeigten sich jedoch weder der erwartete Interaktions- noch ein Haupteffekt.

Für die Stichprobe von $N = 47$ konnte keine der Hypothesen bestätigt werden.

6.3 Diskussion

In diesem Abschnitt sollen die Ergebnisse in ihrer Bedeutung für die Fragestellungen und Hypothesen untersucht werden, die in Kapitel 4 vorgestellt worden sind. Für eine eher theoretisch fundierte Einordnung der Resultate, die beispielsweise auch auf die Implikationen für die Frage nach den Auswirkungen neurowissenschaftlicher Diskurse eingeht, sei auf Kapitel 9 verwiesen.

Die Hypothesen zum *Zusammenhang zwischen Intelligenztheorie, Emotion und Bewältigungsverhalten nach Misserfolg*, die gleichsam den Kern meiner theoretisch abgeleiteten Erwartungen ausmachen, konnten überwiegend bestätigt werden. Die dynamische Intelligenztheorie führt zu Zuversicht und Schuld, welche wiederum meisterndes Bewältigungsverhalten nach sich ziehen und mindestens teilweise als Mediatoren zwischen Intelligenztheorie und Bewältigungsverhalten fungieren. Umgekehrt deutet sich immerhin an, dass die statische Intelligenztheorie Hoffnungslosigkeit und Scham begünstigt und dass diese ihrerseits

Tabelle 18

Deskriptive Ergebnisse (Mittelwerte und Standardabweichungen) in Abhängigkeit Impliziter Theorien unter Berücksichtigung der Leistungssituation, Bedingung Erfolg

	Gesamt		individualistisch		kollektivistisch	
	Statiker (n=17)	Dyn. (n=17)	Statiker (n=9)	Dyn. (n=8)	Statiker (n=8)	Dyn. (n=9)
Freude	3.27 .78	2.84 1.03	3.41 .80	2.58 .79	3.13 .80	3.07 1.21
Hybris	3.80 .88	3.49 1.16	3.85 .67	3.00 1.11	3.75 1.12	3.93 1.08
Stolz	3.76 .97	3.12 1.28	3.83 .61	2.63 1.22	3.69 1.31	3.56 1.24
Erleichterung	3.38 1.05	3.68 .88	3.56 1.13	3.75 .76	3.19 1.00	3.61 1.02
Attribution Anstrengung	3.94 .83	3.18 1.38	3.67 .87	3.00 1.51	4.25 .71	3.33 1.32
Attribution Fähigkeit	3.59 .62	3.41 1.18	3.33 .50	3.00 1.41	3.88 .64	3.78 .83
Darstellung Hybris	2.97 .51	2.79 .99	2.89 .49	2.56 1.02	3.06 .56	3.00 .97
Darstellung Stolz	3.12 .70	3.00 1.27	3.22 .67	2.88 1.36	3.00 .76	3.11 1.27
Darstellung Erleicht.	2.88 1.05	3.35 1.11	2.89 .93	3.13 1.25	2.88 1.25	3.56 1.01
Darstellung Freude	2.82 1.19	2.82 1.24	2.89 .78	3.00 1.20	2.75 1.58	3.33 1.00
Verstecken Freude (hab's nicht nötig)	2.88 1.17	3.59 1.06	2.67 1.00	2.50 1.20	3.13 1.36	3.11 1.27
Verstecken Freude (andere nicht runterziehn)	3.59 1.12	1.53 .72	3.33 1.32	3.13 1.13	3.88 .83	4.00 .87
Bewältigungsverhalten (Anteil der Bewältiger)	65%	44%	67%	57%	63%	34%

hilfloses Verhalten hervorbringen. Dabei ist zu berücksichtigen, dass für die Beziehung zwischen Emotion und Bewältigungsverhalten kein strenger kausaler Schluss möglich ist. Ein solcher lässt sich aber aus der Erhebungsreihenfolge sowie pfadanalytischen Berechnungen heraus vermuten. Zudem kann zumindest die direkte Beziehung zwischen Impliziten Theorien und *Bewältigungsverhalten* kausal gedeutet werden. Dass ein derart verhaltensnahes Maß in so eindeutiger

Weise von Impliziten Theorien beeinflusst wird, stellt einen wesentlichen Befund der Untersuchung dar. Die quantitativ ähnlichen Ergebnisse von Hong et al. (1999) wurden nicht bloß repliziert, sondern – über die Vermeidung der Intelligenzsemantik bei der Operationalisierung des Bewältigungsverhaltens (s.o.) – erst auf allgemeine kognitive Leistungsaufgaben hin generalisierbar gemacht.

Unklar bleibt indes, weshalb *Anstrengungsattributionen* in diesem Modell nicht (oder nur in Ansätzen) die erwartete vermittelnde Rolle zwischen Impliziten Theorien und Emotionen spielen. Die geringen korrelativen Beziehungen zu den Emotionen legen die Vermutung nahe, dass eine unzureichende Operationalisierung hierfür verantwortlich sein könnte (vgl. die Änderungen in Studie 3).

Da die Einbeziehung von *Emotionen* in das Dwecksche Modell die wohl wichtigste Innovation der vorliegenden Arbeit ist, sollen die diesbezüglichen Ergebnisse etwas ausführlicher diskutiert werden. Zwar lässt sich die Befundlage keineswegs als eindeutig bezeichnen, aber in der Zusammenschau der unterschiedlichen Ergebnisquellen erscheint ein zufälliges Zustandekommen der beobachteten Effekte unwahrscheinlich: Zuversicht und Schuld werden wie erwartet von Dynamikern stärker empfunden; Effekte für Hoffnungslosigkeit und Scham lassen sich – vielleicht aufgrund von Tendenzen zu sozialer Erwünschtheit und Selbstdarstellung – in der ursprünglichen Operationalisierung hingegen nicht ausmachen. Die für Studie 1 (Abschnitt 5.3) erörterten Erklärungsmöglichkeiten lassen sich auch hier anwenden, da ja (zwar nicht mehr das Problem des hypothetischen Erlebens, aber) immer noch das Problem des Selbstberichts besteht. Um so aufschlussreicher ist daher das Resultat, dass gerade die sozial so unerwünschte Emotion der Scham sich beim Heranziehen des Beobachtungsmaßes als von Intelligenztheorien tendenziell abhängig erweist. In ähnlicher Form zeigen sich bei Hoffnungslosigkeit die erwarteten Unterschiede zwar nicht nach der Leistungsrückmeldung, wohl aber im Anschluss an die Gruppenarbeit, wo das Bedürfnis, sich positiv darzustellen, weniger naheliegend sein dürfte (vgl. Abschnitt 5.3). Zudem lassen sich die erwarteten Effekte für Hoffnungslosigkeit und Scham auch in Abhängigkeit beider Dispositionsmaße nachweisen.

Die induzierten Leistungsemotionen scheinen außerdem die Lernleistung in nachfolgenden Aufgaben zu bestimmen. Wie im Lichte ähnlicher Befunde aus der Prüfungsangstforschung (vgl. Zeidner, 1998) zu vermuten war, bewirken dysfunktionale Gefühle der Hoffnungslosigkeit und Depression nicht nur einen Motivationsverlust, sondern stören auch bei der Bearbeitung fordernder Leistungsaufgaben. Obwohl Hoffnungslosigkeit durch eine statische Intelligenztheorie angeregt wird ($N = 41$), besteht allerdings kein direkter Zusammenhang zwischen dieser und der erbrachten Leistung. Eine weitere, noch unbekannte Einflussgröße scheint den negativen Effekt von Hoffnungslosigkeit zu kompensieren. Wie unten deutlich werden wird, könnten spezifische, „anstachelnde“ Formen der Scham diese Erklärungslücke in künftigen Studien schließen. Diese Er-

gebnisse sind dennoch auch für sich genommen bemerkenswert – nicht nur angesichts des vagen Misserfolgskriteriums, sondern auch infolge der Tatsache, dass die Einbeziehung der Abiturnote keinen zusätzlichen Vorhersagebeitrag für die Leistung erbringt. Innerhalb dieses Modells (vgl. Abbildung 6) verschwindet also der Einfluss intellektueller Grundvoraussetzungen (soweit sie aus der Abiturnote abzulesen sind) auf Kosten der situational bedingten Emotionen.

In der Frage nach den vermittelnden Prozessen zwischen Impliziten Intelligenztheorien und Bewältigungsverhalten haben sich Emotionen als *eine* Antwort erwiesen. Die „qualitativen“ Merkmale der Diskussion über die Induktionstexte konnten wenig zu dieser Frage beitragen. Zwar gibt es Anzeichen, dass die Texte insbesondere dann ihre Wirkung nicht verfehlen, wenn die Person sie als negativ für ihre eigene Position bewertet sowie als gesellschaftlich schädlich kritisiert. Im Ganzen können diese Analysen jedoch, nicht zuletzt aufgrund des eher grobkörnigen methodischen Vorgehens, nicht überzeugen. Eine weitere Möglichkeit, vermittelnde Prozesse zu identifizieren, besteht in der Ableitung motivationaler Tendenzen aus den erhobenen Emotionen. Diese wird weiter unten näher beschrieben.

Emotionsdarstellung hat sich nur in eingeschränktem Umfang als abhängig von Intelligenztheorien erwiesen. Statiker bevorzugen die Darstellung von Langeweile, Dynamiker die von Interesse. Erstere versuchen damit, von ihrem Misserfolg abzulenken und den Fokus auf die unzureichende Aufgabe zu lenken, während Letztere ihr Scheitern zugeben und mit dem dargestellten Ehrgeiz den Schluss auf ihre (mangelnde) Intelligenz riskieren (wenn nicht jetzt, dann spätestens beim nächsten Scheitern). Dass die übrigen Hypothesen nicht bestätigt werden konnten, könnte an einer ungenügenden Operationalisierung liegen, auf welche die unerwarteten Korrelationsmuster schließen lassen. Naheliegend erscheint etwa, eine unzureichende Operationalisierung an der beträchtlichen Länge und Komplexität der Emotionsdarstellungs-Items festzumachen.

Das *Zusammenspiel von induzierten und dispositionalen Impliziten Theorien* scheint additiven Gesetzen zu gehorchen, d.h. unabhängig von ihrer Disposition lassen sich Menschen von den Induktionstexten in ihrer Auffassung beeinflussen. Beide Dispositionsmaße haben zudem etwa gleich viel Einfluss auf die Emotionen. Dass die Haupteffekte der induzierten Intelligenztheorien bei Kontrolle der Dispositionsmaße nicht anwachsen, könnte daran liegen, dass im Zuge der Zufallszuteilung geringfügig mehr statisch Disponierte in die statische Bedingung gelangten als dynamisch Disponierte und umgekehrt. Als Kontrollvariablen würden die Dispositionsmaße somit diesen Zufallseinfluss wieder herausziehen und so ihre andere, eigentlich intendierte Wirkungsweise – die Bindung von Fehlervarianz – gleichsam kompensieren.

In der kollektivistischen *Leistungssituation* werden nach individuellem Misserfolg zwar tendenziell mehr selbstrelevante Emotionen wie Scham und Schuld

erlebt als im individualistischen Fall (vgl. die vergleichsweise deutlichen deskriptiven Unterschiede in Tabelle 12). Die Unterschiede sind aber nicht statistisch bedeutsam. Die kollektivistische Situation scheint zwar emotional anregender zu sein; die emotionalen Kosten kollektivistischer Einstellungen bzw. kooperativer Lern- und Arbeitsformen, wie sie bei Grant und Dweck (2001) postuliert werden, können so aber bestenfalls erahnbar gemacht, nicht nachgewiesen werden. Das Bewältigungsverhalten hängt ebenfalls zwar in hypothesenkonformer Richtung, aber nicht in statistisch signifikanter Weise von der Leistungssituation ab. Die Hypothesen zur Emotionsdarstellung ließen sich deutlich nicht bestätigen. Die tatsächliche Bedeutung des Faktors Leistungssituation ist angesichts dieser unklaren Ergebnismuster schwer einzuschätzen. Zur besseren Sichtbarmachung relativ kleiner Effekte könnte es in zukünftigen Studien Sinn machen, die Versuchspersonenanzahl zu erhöhen und das Merkmal Bewältigungsverhalten auf einem differenzierterem Skalenniveau zu erheben (vgl. Studie 3).

Das Merkmal *Geschlecht* interagiert mit dem Faktor Intelligenztheorie in überraschend deutlicher Art und Weise. Frauen lassen sich von den Spiegel-Artikeln weitaus stärker in ihrem emotionalen Erleben und z.T. auch Darstellen beeinflussen als Männer. (Für das Bewältigungsverhalten lässt sich aufgrund des geringen Datenniveaus kein sinnvoller quantitativer Gruppenvergleich durchführen.) Oft zeigen Männer gar ein spiegelbildliches, also hypothesenentgegengesetztes Reaktionsverhalten. Wiewohl nicht außer Acht gelassen werden sollte, dass lediglich acht Männer (bei $N = 33$) untersucht wurden, sind jene Unterschiede wohl zu deutlich und konsistent, um einfach als zufällig abgetan werden zu können. Auch die sicher nicht unplausible Einschätzung, dass Frauen die aufrichtigeren, engagierteren, gegenüber dem Emotionsthema aufgeschlosseneren Versuchspersonen sind und somit für weniger Störvarianz sorgen als Männer, reicht hier wohl nicht aus.

Weiterhin konnten auch einige naheliegende methodische und theoretische Erklärungen für die Geschlechtsdifferenzen ausgeschlossen werden: Etwa zeigten sich keine Zufälligkeiten in der Gruppenzuteilung, d.h. dass zufallsbedingt beispielsweise mehr statisch disponierte Frauen in die statische Bedingung geraten sein könnten als dynamisch disponierte Frauen. Frauen sind zudem keineswegs von Anfang an stärker zur statischen Intelligenztheorie disponiert. Schließlich lassen sich auch keine Interaktionen zwischen Geschlecht und den Dispositionsmaßen nachweisen. Dies wiederum legt nahe, die Ursache für die angesprochenen Geschlechtsunterschiede in der konkreten Situation der Induktion zu vermuten, d.h. vor allem in inhaltlichen Aspekten der Zeitschriftenartikel.

Weshalb tatsächlich die Induktionstexte für die Geschlechtsdifferenzen verantwortlich sein könnten, wurde bereits in Abschnitt 3.2.2 empirisch und theoretisch hergeleitet: Probandinnen assoziieren demnach mit der Botschaft einer stabilen Intelligenz die intellektuelle Unterlegenheit von Frauen, während Männer

jener Botschaft immerhin die positive Konnotation abgewinnen, dass die vorteilhafte gesellschaftliche Stellung ihres Geschlechts eben doch nicht soziokulturelle, sondern genetische Gründe hat und sie deshalb bessere Erfolgschancen haben als Frauen. Die dynamische Botschaft wirkt in spiegelbildlicher Weise (vgl. Abschnitt 3.2.2). Die vorliegenden Daten stützen demnach nicht nur die Hypothese, dass Frauen von den Induktionstexten stärker beeinflusst werden, sondern stehen auch im Einklang mit der vorgeschlagenen theoretischen Erklärung.

Einschränkend muss allerdings auch betont werden, dass vor der Untersuchung keine eindeutigen Hypothesen zum Faktor Geschlecht formuliert worden waren, die gefundenen Effekte mithin eher posthoc-Charakter haben. Zwar hatten frühere Studien bereits gezeigt, dass geschlechtsabhängige Unterschiede vorkommen; sie standen jedoch nicht im Zentrum des Interesses der vorliegenden Arbeit, weshalb auch keine Anstrengungen unternommen wurden, Frauen und Männer in gleicher Anzahl für den Versuch zu rekrutieren. Auf jeden Fall sollten die gefundenen Unterschiede an einer größeren Stichprobe validiert werden (vgl. Studie 3).

Die Reaktionsmuster nach *positiver Leistungsrückmeldung* sind in deutlich geringerem Ausmaß abhängig von den experimentellen Faktoren. Zwar waren die betreffenden Hypothesen aufgrund fehlender empirischer Vorerfahrungen ohnehin in eher explorativer Form formuliert worden, so dass das Ausbleiben der Gruppenunterschiede nicht vollends überraschen kann. Bemerkenswert ist jedoch, dass sich noch nicht einmal die verschiedenen Emotionskonstrukte sinnvoll empirisch trennen ließen. Die Probanden können bzw. wollen wenig differenzieren zwischen verschiedenen positiven Emotionen; allem Anschein nach fällt es ihnen schwer, ihre Gefühle differenzierter als anhand der positiv-negativ-Unterscheidung zu beschreiben. Sie geben lediglich an, in welchem Ausmaß sie sich nach der Rückmeldung „gut fühlen“ – weswegen und in welcher Qualität wird kaum unterschieden.

Dass das Bewältigungsverhalten immerhin eine vergleichsweise deutliche hypothesenkonforme Tendenz aufweist, spricht für die Qualität und Trennschärfe der vorgenommenen Operationalisierung und gegen ein zufälliges Zustandekommen der Unterschiede im Misserfolgsfall. Einziger Nachteil dieser Operationalisierung ist wie gesagt das geringe Skalenniveau (vgl. Studie 3). Für die übrigen Items muss offen bleiben, ob die Hypothesen aufgrund einer unzureichenden Operationalisierung oder aber aus inhaltlichen Gründen gescheitert sind.

Die bis hierhin referierten Befunde sind von begrenzter *Generalisierbarkeit*. Sie wurden – wenn man nun wieder nur den Misserfolgsfall betrachtet – anhand einer selektiv leistungsschwachen Studierendenpopulation gewonnen. Die Leistungsschwäche resultiert einerseits aus der Versuchslogik, wonach jeweils der in der ASK objektiv schwächeren Person Misserfolg rückgemeldet wurde. Andererseits ergab sie sich aus praktischen Gründen, etwa der Akquirierung der Pro-

banden in bestimmten Studienfächern. Die Versuchspersonen waren daher auch tendenziell überfordert mit der für eine akademisch gebildete Stichprobe genormten ASK: Bekanntlich war die Durchschnittsleistung kaum so hoch wie die Ratewahrscheinlichkeit. Doch selbst wenn die Befunde nur für eine leistungsschwache Subpopulation Gültigkeit hätten – was theoretisch durchaus plausibel wäre –, verlören sie wohl kaum an praktischer Relevanz, eher im Gegenteil.

Bei den meisten der geschilderten Befunde fallen die zugehörigen *Effektstärken* eher klein aus. Etwas anderes konnte allerdings nicht erwartet werden angesichts der bekanntermaßen vielfältigen Einflussvariablen auf Lernen und Leistung, angefangen von Persönlichkeit und Selbstkonzept bis hin zu situationalen Kontextbedingungen. Beispielsweise begründeten einige Lehramtsstudierende ihre Ablehnung des fingierten Trainingsangebots explizit damit, dass sie aufgrund ihrer Studienwahl nie in die dort angesprochenen Bewerbungssituationen kommen würden. Dass sich gegen derartige Störeinflüsse immer noch nennenswerte Effekte durchsetzen konnten, erscheint bemerkenswert. Blackwell et al. (2007) argumentieren hinsichtlich ihrer Ergebnisse in ähnlicher Weise und geben zu Recht zu bedenken, dass, gerade bei der vorliegenden Fragestellung, „small effect sizes can have a major impact on outcomes over time“ (S. 259).

Nicht zuletzt wegen dieser kleinen Effektstärken hat die Diskussion der Ergebnisse von Studie 2 bislang an vielen Stellen zu der Schlussfolgerung geführt, dass eine Replikation der Untersuchung an einer größeren Stichprobe sinnvoll wäre. Auf diese Weise nämlich ließen sich die zwar interessanten, aber eben numerisch meist geringen Unterschiede zwischen den experimentellen Gruppen besser von Zufallseinflüssen abgrenzen. Abschließend soll aber noch auf ein Ergebnismuster eingegangen werden, das – jenseits einer einfachen Replikation der Untersuchung – eine Erweiterung des bislang zugrundegelegten theoretischen Modells zu rechtfertigen scheint.

Die überraschenden empirischen Zusammenhänge der über Beobachtung operationalisierten Schamvariablen hatten den Anstoß dazu gegeben, im Rahmen einer Faktorenanalyse verschiedene Konstrukte hinsichtlich der ihnen innewohnenden *motivationalen Tendenzen* zu untersuchen. Dabei wurde deutlich, dass beispielsweise Scham neben einer – ursprünglich erwarteten – leistungshinderlichen Komponente („Resignation“, im Sinne eines Versteckens bzw. Rückzugs, stark assoziiert mit Hoffnungslosigkeit) auch einen leistungsförderlichen Aspekt („Anstachelung“) beinhalten könnte. Letzterer war in Form von jener beobachteten Scham zutage getreten. Scham sollte also nicht nur dazu motivieren, gleichsam aus dem Feld zu gehen, ohne dass die schamauslösenden Probleme gelöst sind, sondern auch – unter bestimmten Umständen – dazu veranlassen, die Quelle der mit Scham einhergehenden Selbstwertverletzung aktiv anzugehen. Die uneindeutige empirische Beziehung zwischen Scham und Bewältigungsverhalten wird unter Bezug auf diese zwei gegenläufigen Tendenzen erklärlich.

Auch weitere der gemessenen Konstrukte scheinen nicht nur eine einzige motivationale Wirkungsrichtung aufzuweisen. Anstrengungsattribution steht einerseits mit einem leistungsförderlichen Faktor, der als „Kompetenzbedürfnis“ gelabelt wurde, in Beziehung, andererseits aber auch mit einem leistungshinderlichen („Gleichgültigkeit“). Wer auf Anstrengung attribuiert, ist demnach nicht nur geneigt, Lernfortschritt als möglich zu erachten, sondern tendiert auch dazu, die Notwendigkeit eines Lernfortschritts infrage zu stellen, weil er ihn weniger mit seiner Fähigkeit bzw. Intelligenz in Verbindung bringt und sich somit weniger mit ihm identifiziert („Ich hätte mich ja auch mehr anstrengen können!“). In ähnlicher Form sollte Schuld einerseits eine aversive, anstachelnde Komponente aufweisen, andererseits aber auch mit Kompetenzbedürfnis einhergehen. Schuld hätte somit zwei leistungsförderliche Komponenten, was durch den klar positiven Zusammenhang zum Bewältigungsverhalten empirisch untermauert wird. Für Zuversicht scheint eine solche Analyse wenig gewinnbringend – handelt es sich doch bei ihr bereits um eine motivational eindeutig bestimmte Emotion.

Es resultieren vier Motivationskonstrukte, die die vielschichtigen empirischen Beziehungen zwischen Emotion und Bewältigungsverhalten erklären sollen. Diese werden in Abbildung 7 graphisch veranschaulicht. Dabei erscheint das beobachtete Bewältigungsverhalten gleichsam als Summe aller vier motivationalen Einflüsse sowie Zuversicht.

Diese Überlegungen entspringen explorativen Datenanalysen. Es erscheint angebracht, die postulierten motivationalen Tendenzen nicht mehr nur als faktorenanalytisch extrahierte Konstrukte zu analysieren, sondern sie unabhängig von den Emotionen direkt zu erfassen. Dieses Vorhaben wird in Studie 3 umgesetzt.

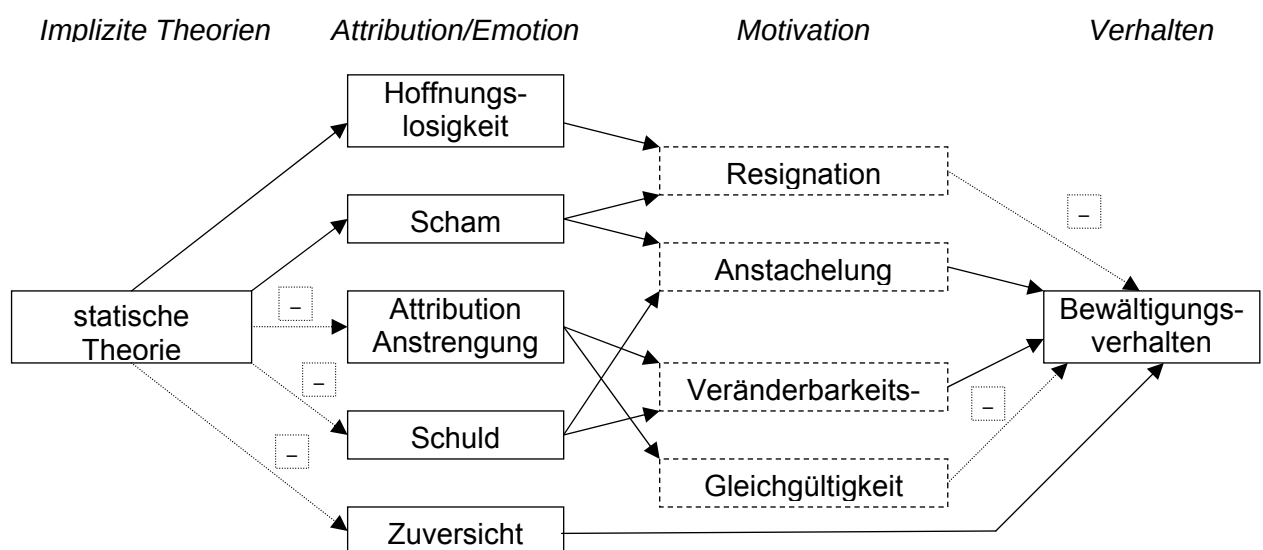


Abbildung 7. Vier hypothetische Motivationskonstrukte zur Erklärung der empirisch gefundenen Beziehungen zwischen Emotion und Bewältigungsverhalten.

7. Studie 3: Prüfung eines erweiterten Modells anhand einer Leistungsaufgabe

Die Ergebnisse von Studie 2 konnten die Hypothesen im Vergleich zu denjenigen der ersten Studie deutlich besser bestätigen. Es macht daher Sinn, in einer weiteren Studie die grundlegenden Merkmale der Untersuchung beizubehalten und eine größere Stichprobe ins Auge zu fassen. Aus ökonomischen Gründen erschien es jedoch naheliegend, das recht aufwändige Forschungsdesign von Studie 2 in einigen Aspekten zu vereinfachen. Neben diesen ökonomisch bedingten Veränderungen, die in Abschnitt 7.1 näher beschrieben werden, sollen jedoch auch neue inhaltliche Hypothesen geprüft werden.

Dabei handelt es sich um die Einbeziehung motivationaler Tendenzen und deren Beziehung zu Emotion und Bewältigungsverhalten. Entsprechend der in Abschnitt 6.3 (Abbildung 7) entwickelten Hypothesen sollen vier motivationale Zustände erfasst werden: Resignation, Anstachelung, Kompetenzbedürfnis und Gleichgültigkeit. Erfasst werden soll also, *weshalb* die Probanden Bewältigungsverhalten zeigen bzw. nicht zeigen (wollen): Erfolgt das Ausschlagen von Bewältigungsangeboten eher deswegen, weil die Person nicht glaubt, dass sie sich verbessern kann (Resignation) – oder weil sie es nicht für nötig hält, sich zu verbessern (Gleichgültigkeit)? Erfolgt das Annehmen der Angebote eher aus Interesse und dem Ehrgeiz, sich zu verbessern (Kompetenzbedürfnis) – oder aus dem Verlangen, in künftigen sozialen Vergleichen weniger stark von Schamgefühlen geplagt zu werden (Anstachelung). An dieser Stelle wird deutlich, dass Resignation und Anstachelung Ähnlichkeiten zum Konstrukt der Leistungszielorientierung aufweisen, Kompetenzbedürfnis und – teilweise – Gleichgültigkeit hingegen an das Konstrukt der Lernzielorientierung erinnern (vgl. Abschnitt 3.3.3). Ein wichtiger Unterschied ist allerdings, dass die hier eingeführten Konstrukte nicht auf Trait-, sondern auf State-Ebene anzusiedeln und entsprechend sinnvoll in ein Prozessmodell einzubinden und situationsbezogen zu erfassen sind. Die hierbei zu erwartenden Beziehungen zu Emotion und Bewältigungsverhalten wurden bereits in Abschnitt 6.3 konkretisiert. Danach sollte Resignation von Hoffnungslosigkeit und Scham abhängen, Anstachelung von Scham und Schuld, Kompetenzbedürfnis von Schuld und Anstrengungsattribution, Gleichgültigkeit von Anstrengungsattribution. Resignation und Gleichgültigkeit sollten zu hilflosem, Anstachelung und Kompetenzbedürfnis zu meisterndem Bewältigungsverhalten führen. Wie in der Beschreibung der methodischen Umsetzung dieser Hypothesen (vgl. Abschnitt 7.1) deutlich werden wird, lassen sich diese kausal formulierten Hypothesen im gewählten Forschungsdesign leider nur korrelativ überprüfen. Immerhin lässt sich sicher aber testen, ob die vier Motivationskonstrukte die Beziehung zwischen Emotion und Bewältigungsverhalten vermit-

teln. Auch sollte untersucht werden, inwieweit die vier Konstrukte vom Faktor Intelligenztheorie abhängen. Aus den in Abschnitt 6.3 angestellten Überlegungen lässt sich ableiten, dass Statiker eher zu Resignation und – tendenziell – zu Anstachelung neigen sollten, Dynamiker hingegen zu Kompetenzbedürfnis und Gleichgültigkeit.

Einige aus der Diskussion in Abschnitt 6.3 abgeleitete Desiderata konnten aus pragmatischen Gründen nicht umgesetzt werden. Erstens wurden die Items für die Emotionen nach Erfolg trotz der geringen Konstruktvalidität nicht verändert, weil die betreffende Fragestellung für meine Arbeit von untergeordneter Bedeutung ist. Zweitens konnte wiederum kein ausgeglichenes Geschlechterverhältnis realisiert werden, weil dies ökonomisch nicht möglich war bzw. eine entsprechende Verringerung der Stichprobenanzahl nicht sinnvoll erschien (wiederum standen deutlich mehr Frauen als Männer für den Versuch zur Verfügung). Schließlich wurde in Studie 3 auf einige Teilfragestellungen zugunsten anderer, erfolgversprechenderer Hypothesen verzichtet. Namentlich geht es dabei vor allem um die Emotionsdarstellung und den Zusammenhang zwischen Emotion und Leistung. Näheres ist dem folgenden Abschnitt zu entnehmen.

7.1 Methode

Teilnehmer. Es nahmen 130 Studierende der Universität Kassel an der Untersuchung teil, vorwiegend aus den Studienfächern Lehramt, Erziehungswissenschaften und Soziologie. Sie wurden im Rahmen von Lehrveranstaltungen geworben und erhielten zehn Euro oder eine Versuchspersonenbescheinigung (relevant für Studierende mit Nebenfach Psychologie) für ihre Teilnahme. Diese Vergütung erfolgte – analog zum Vorgehen bei Studie 2 – unabhängig von der geldlichen Belohnung für das erfolgreiche Abschneiden bei der Leistungsaufgabe. Die Probanden waren durchschnittlich 23 Jahre alt (mit einer Spanne von 18 bis 33) und hatten eine Abiturnote von 2.7 (mit einer Spanne von 1.5 bis 4.1). Von den 130 Teilnehmern waren 105 weiblich.

Änderungen im Vergleich zu Studie 2. In vielen Punkten wurde die Vorgehensweise und die Materialien direkt von Studie 2 übernommen. Im Folgenden werden daher nur die Unterschiede zwischen beiden Studien aufgeführt. Zur Verringerung der Versuchsdauer von 75 auf 60 Minuten wurden folgende Veränderungen vorgenommen:

- Die Anzahl der Fragen zum Spiegel-Artikel wurde von vier auf drei verringert.
- Nach der Diskussion über den Spiegel-Artikel erhielten die Teilnehmer keinen („ersten“) Emotionsfragebogen mehr. Entsprechend wurde den Teilnehmern der Untersuchungszweck anders plausibilisiert. Ähnlich wie in Studie 1 war nun wieder von zwei separaten Untersuchungen die Rede, wobei sich die erste um das Verständnis wissenschaftlicher Zusammenhänge drehen sollte.

- Beim („zweiten“) Emotionsfragebogen wurde auf die Erhebung der subjektiven Leistungseinschätzung und der Emotionsdarstellung sowie auf das offene Emotions-Item verzichtet.

Folgende Veränderungen wurden zur Verbesserung der Versuchsökonomie vorgenommen, v.a. zur Verringerung der Auswertungszeit bzw. leichteren Anwerbung der Teilnehmer.

- Es wurden keine Video-Aufnahmen mehr angefertigt. Einzig die Variable Scham wurde auch in Studie 3 mittels Beobachtung erhoben: Die Versuchsleiterin notierte sich direkt nach der Rückmeldung einen Wert, der das Schamerleben des Misserfolgsprobanden in sechsstufiger Form einschätzen sollte. Dafür wurde sie kurz darin eingewiesen, anhand welcher Kriterien dieses Merkmal beim Video-Rating in Studie 2 bewertet wurde. Diese ökonomische Erfassung von (beobachteter) Scham hatte jedoch den Nachteil, dass die Beurteilerin gegenüber der jeweiligen experimentellen Bedingung nicht blind war.
- Die Rückmeldung wurde besser an die Fähigkeit der Versuchspersonengruppe angepasst, indem die Gesamtschwierigkeit der ASK-Aufgaben etwas verringert wurde. Eine gemäß den Normwerten schwierige ASK-Aufgabe (Nr. 2 des Moduls „Informationen interpretieren“) wurde gegen eine einfache (Nr. 1) ausgetauscht. Nominal blieb die Rückmeldung gleich (14 zu 8 Punkte). Auf diese Weise sollten (so die Intention) nicht mehr so viele „zu schlechte“ Probanden aus der Analyse ausgeschlossen werden müssen wie in Studie 2.
- Die Probanden erhielten für ihre Teilnahme einen größeren geldlichen Anreiz (zehn statt sechs Euro „Grundgehalt“).

Aus theoretischen Erwägungen heraus kam es zu folgenden Änderungen bzw. Erweiterungen:

- Die Erfassung von Emotion in der Misserfolgsbedingung sowie von Attribution in beiden Bedingungen wurde in differenzierterer Form durchgeführt. Emotionen nach Misserfolg werden nun mit 15 statt mit 10 Items erhoben. Neben Hoffnungslosigkeit (vier statt vorher drei Items), Scham (weiterhin drei Items), Schuld (drei statt vorher zwei Items) und Zuversicht (drei statt vorher zwei Items) wurde nun auch aus explorativen Gesichtspunkten heraus Interesse und Langeweile mit je einem Item erhoben. Zu erwarten ist, dass Dynamiker eher zu Interesse (an den eigenen Fehlern etc.), Statiker eher zu Langeweile (d.h. Abwertung der Aufgaben) tendieren. Die Hypothese, dass Statiker zur Darstellung von Langeweile neigen, wird hier also nicht mehr über Darstellungs-Items operationalisiert, sondern über die konkrete Gelegenheit für die Probanden, sich über den Fragebogen zum eigenen Erleben selbst darzustellen. Anstrengungsattribution wird nun nicht mehr mit einem, sondern mit zwei Items gemessen („Konzentration und Anstrengungsbereitschaft“ und „Motivation“). Alle neuen Items wurden anhand theoretischer Überlegungen erstellt, mit Ausnahme des neuen Scham-Items, das von Molfenter (1999) übernommen wurde.
- Zwischen dem Bogen zur Erfassung des Bewältigungsverhaltens und der Aufklärung über den wahren Untersuchungszweck wurde ein weiterer kurzer Fragebogen eingeschoben. Zur Erfassung der motivationalen Tendenzen (s.o.) werden die Probanden dabei mit der Frage konfrontiert, aus welchen Gründen sie sich für bzw. gegen das Trai-

ningsangebot entschieden haben. Jede motivationale Tendenz wird mit zwei Items abgefragt. Wer dem Trainingsangebot zugestimmt hat, erhält zwei Items zur Anstachelung und zwei Items zum Kompetenzbedürfnis. Wer abgelehnt hat, bekommt zwei Items zur Resignation und zwei Items zur Gleichgültigkeit. Der Wortlaut dieser selbstkonstruierten Items ist in Tabelle 19 wiedergegeben.

- Auf dem Bogen zur Erfassung der motivationalen Tendenzen wurde außerdem die sechsstufig zu beantwortende Frage platziert, wie sicher sich der Proband bei Annahme bzw. Ablehnung des Trainingsangebots war. Damit sollte eine differenziertere Messung des Bewältigungsverhaltens ermöglicht werden.
- Es wurden nun durchgehend sechsstufige Rating-Skalen verwendet, während in den Studien 1 und 2 größtenteils fünfstufige Skalen benutzt worden waren. Angestrebt wurde damit eine genauere Abbildung der oftmals geringen Unterschiede zwischen den experimentellen Gruppen.

Tabelle 19

Items zur Erfassung der motivationalen Tendenzen nach Misserfolg und nach Erfolg

	Item 1	Item 2
Resignation	Ich bin in diesen Aufgaben so oder so nicht besonders gut. Da braucht es kein Training, das dies extra unter Beweis stellen muss.	Ich habe keine Lust, mich in einem solchen Training vielleicht vor anderen zu blamieren.
Anstachelung	Es war so ein schlechtes Gefühl, als ich mein Ergebnis gesagt bekommen habe, dass ich mir einfach gesagt habe: „Du musst was dagegen tun!“	Ich hasse es, wenn andere Leute sehen, dass ich etwas nicht kann. Daher werde ich dieses blöde Training machen, es dauert ja nicht lang.
Kompetenzbedürfnis	Ich denke, wenn diese Aufgaben mir wirklich mal in Zukunft gestellt werden könnten, dann sollte man sich dransetzen und sie trainieren.	Wenn ich etwas nicht hinkriege, dann weckt das meinen Ehrgeiz. Ich bin dann neugierig darauf, wie schnell ich mich verbessern kann.
Gleichgültigkeit	Ob ich in solchen abstrakten Aufgaben gut oder schlecht abschneide, ist mir eigentlich egal. Daher gibt es für mich auch keinen Grund, zu dem Training zu gehen.	Natürlich könnte man sich in diesen Aufgaben durch Training verbessern, aber das ist mir die Sache nicht wert. Da gibt es wichtigeres für mich zu lernen.

7.2 Ergebnisse

Anders als in Studie 2 stehen in Studie 3 nicht zwei, sondern nur eine Stichprobe zur Hypothesenprüfung zur Verfügung. In Analogie zu Studie 2 werden auch hier diejenigen Probanden aus der Analyse ausgeschlossen, deren erzielter ASK-Wert deutlich von dem rückgemeldeten Wert abweicht (vgl. die Begründung in Abschnitt 6.2). Legte man jedoch das in Studie 2 verwendete Kriterium für die Misserfolgsbedingung an, gemäß dem alle Probanden mit weniger als sechs und

mehr als zehn Punkten ausgeschlossen werden, blieben lediglich 45 von 65 Probanden übrig. Um die Stichprobe nicht in solch großem Umfang zu reduzieren, wird ein etwas liberaleres Kriterium gewählt, das Probanden nicht mehr bei mindestens drei, sondern erst bei vier Punkten Abweichung vom rückgemeldeten Wert ausschließt. In der Erfolgsbedingung wird das in Studie 2 verwendete Kriterium (mehr als fünf Punkte Abweichung) beibehalten. Tabelle 20 verdeutlicht, dass in der Misserfolgsbedingung die Stichprobengröße nun 57, in der Erfolgsbedingung 61 beträgt. Zuvor waren in der Misserfolgsbedingung bereits zwei weitere Probanden ausgeschlossen worden, die den Versuch (durch monotones Ankreuzen des kleinsten Skalenwerts etc.) erkennbar boykottiert hatten. Die durchschnittliche ASK-Leistung fällt erwartungsgemäß etwas besser aus als in Studie 2 ($M = 9.6$ zu $M = 8.9$).

Tabelle 20

Ergebnisse der Stichprobenselektion anhand der objektiv erzielten Leistungsergebnisse

	Misserfolgsbedingung			Erfolgsbedingung		
Rohpunktwert Leistung	2-4	5-11	12-18	5-8	9-15	16-18
Anzahl Probanden	3	57	3	4	61	0
Berücksichtigung	nein	ja	nein	nein	ja	--

7.2.1 Reaktionen auf negative Leistungsrückmeldungen

Zunächst soll wiederum die Frage untersucht werden, ob sich die experimentellen Gruppen von vornherein hinsichtlich wichtiger Merkmale unterscheiden. Als relevante Merkmale kommen hier die Disposition Impliziter Theorien, Bedeutsamkeit von Intelligenz, Leistung, Abiturnote und Geschlecht in Betracht. Die – in Tabelle 21 deskriptiv aufgeführten – Ergebnisse solcher Gruppenvergleiche kommen zu dem Ergebnis, dass sich keine Unterschiede zwischen den Gruppen nachweisen lassen. Zwischen Implizite Theorie und Geschlecht besteht lediglich eine leichte Abhängigkeit ($p = .13$).

Die Korrelationsmuster zwischen den Emotions- und Attributionskonstrukten fallen in vielen Fällen ähnlich wie in Studie 2 aus (vgl. Tabelle 22). Teilweise ergeben sich aber auch neue, hypothesenabweichende Beziehungen, so etwa zwischen Zuversicht und Anstrengungsattribution. Die neuen Konstrukte Interesse und Langeweile sind zu den anderen Merkmalen überwiegend unkorreliert. Lediglich Interesse und Zuversicht stehen in deutlicher (positiver) Beziehung. Überraschenderweise sind auch Interesse und Langeweile untereinander unkorreliert. Aus Tabelle 22 lassen sich außerdem die internen Konsistenzen (Cron-

bachs Alpha) entnehmen. Diese fallen in den meisten Fällen etwas geringer aus als in Studie 2, was angesichts der Skalenerweiterungen überrascht. Die Skala zu Anstrengungsattribution, die nun mit zwei Items gemessen wurde, erscheint befriedigend reliabel.

Tabelle 21

Mittelwerte und Standardabweichungen einiger Kontrollvariablen in Abhängigkeit der experimentellen Faktoren

		Gesamt		individualistisch		kollektivistisch	
		Statiker	Dyn.	Statiker	Dyn.	Statiker	Dyn.
Disposition Impliziter Theorien		2.80 .96	2.75 1.00	3.05 .99	2.98 .85	2.55 .88	2.53 1.10
Bedeutsamkeit von Intelligenz		2.93 .78	2.93 .82	3.02 .81	2.62 .77	2.83 .76	3.22 .77
Leistung		8.39 1.66	8.30 1.92	8.21 1.53	8.54 2.18	8.57 1.83	8.07 1.69
Abiturnote		2.72 .57	2.62 .47	2.74 .68	2.64 .43	2.70 .47	2.61 .53
Geschlecht	weiblich (Anzahl)	22	26	11	13	11	13
	männlich (Anzahl)	6	3	3	1	3	2

Tabelle 22

Korrelative Beziehungen und interne Konsistenzen der Emotions- und Attributionsmaße (Bedingung Misserfolg)

	Scham	Schuld	Zuvers.	Inter.	Lang.	A. An.	A.Fäh	Cronb. α
Hoffnungslosigkeit	.66***	.29**	-.23*	.10	-.08	.12	.34**	.71
Scham		.40**	.05	.19	-.03	.08	.25*	.75
Schuld			.26*	-.07	.17	.35***	.01	.54
Zuversicht				.32**	-.05	-.04	-.34***	.69
Interesse					.01	-.10	.07	
Langeweile						.10	.01	
Attr. Anstrengung							-.33**	.76

Anmerkungen.

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig), * $p < .1$ (zweiseitig).

Emotion. Aus Tabelle 23 wird ersichtlich, dass alle vier („alten“) Emotionen in der erwarteten Richtung von Impliziten Theorien abhängen. Allerdings sind die Unterschiede z.T. recht gering, so dass eine multivariaten Varianzanalyse keine signifikanten Ergebnisse erbringt, $F(4,52) = 1.09$, $p = .19$, $\eta^2 = .08$. In univariaten Varianzanalysen werden die Differenzen nur für Scham (schwach) signifikant, $F(1,55) = 2.67$, $p < .1$, $\eta^2 = .05$, d.h. Statiker empfinden erwartungsgemäß mehr Scham als Dynamiker.

Berücksichtigt man Leistungssituation als zweiten Faktor in den univariaten Varianzanalysen, sollten sich auf diesem Haupteffekte für Scham und Schuld zeigen. Dies ist nicht der Fall. Auch das Geschlecht der Probanden hat keinen Einfluss auf ihr Antwortverhalten, weder direkt noch in Interaktion mit dem Faktor Intelligenztheorie. Tabelle 24 veranschaulicht dieses Fehlen von Geschlechtsdifferenzen anhand deskriptiver Werte und separater t-Tests.

Auf den dispositionalen Impliziten Theorien (in Tabelle 24 mediangesplittet dargestellt) ergeben sich teilweise die erwarteten Unterschiede zwischen Statikern und Dynamikern. Dabei sagt die Disposition Impliziter Theorien Schuld vorher, Bedeutsamkeit hingegen Hoffnungslosigkeit und Scham. Wiederum wurden die Dispositionsvariablen nicht wie bei Dweck als kategoriale, sondern als metrische Größen betrachtet (vgl. detaillierte Erläuterungen in Abschnitt 6.2.1).

Analog zu Studie 2 kann der Einsatz der Dispositionsmaße als Kontrollvariablen auch hier die Effekte in Abhängigkeit Impliziter Theorien nicht vergrößern. Ebenso liegt wiederum keine Interaktion zwischen induzierter und dispositionaler Impliziter Theorie vor (vgl. die Erläuterungen in Abschnitt 6.2.1).

Abschließend soll noch kurz auf die explorativ erhobenen Emotionsvariablen eingegangen werden. Scham war wiederum nicht nur durch Fragebogen, sondern auch durch Beobachtung erhoben worden. Dieser Schamwert steht mit der Fragebogen-Scham in positiver Beziehung, $r = .32$, $p < .05$, zweiseitig). Zudem steht er mit Hoffnungslosigkeit und Schuld in Beziehung, jeweils $r = .20$, *ns*. In Varianzanalysen lässt sich aber keine Abhängigkeit von den experimentellen Faktoren feststellen. Interesse wird, entgegen der Erwartung, geringfügig stärker von Statikern als von Dynamikern gezeigt (vgl. Tabelle 25). Hinsichtlich Langeweile gibt es keine Unterschiede zwischen Statikern und Dynamikern. Beide Dispositionsmaße Impliziter Theorien sagen Interesse in erwarteter Richtung vorher (jeweils $p < .1$), d.h. dynamisch Disponierte zeigen tendenziell mehr Interesse als statisch Disponierte.

Attribution. Anstrengungsattributionen werden erwartungsgemäß tendenziell stärker von Dynamikern vorgenommen als von Statikern, $F(1,55) = 2.04$, $p < .1$, $\eta^2 = .04$. Fähigkeitsattributionen werden hingegen von Statikern präferiert, $F(1,55) = 4.60$, $p < .05$, zweiseitig, $\eta^2 = .08$. Bei den Dispositionsmaßen finden sich keine hypothesenkonformen Unterschiede.

Tabelle 23

Deskriptive Ergebnisse (Mittelwerte und Standardabweichungen) in Abhängigkeit Impliziter Theorien unter Berücksichtigung der Leistungssituation, Bedingung Misserfolg

	Gesamt		individualistisch		kollektivistisch	
	Statiker (n=28)	Dyn. (n=29)	Statiker (n=14)	Dyn. (n=14)	Statiker (n=14)	Dyn. (n=15)
Hoffnungslosigkeit	2.43 .88	2.21 .89	2.21 .81	2.20 .96	2.64 .92	2.22 .87
Scham	3.20 1.12	2.72 1.09	3.33 1.27	2.57 1.09	3.07 .98	2.87 1.12
Schuld	3.12 1.17	3.28 1.12	3.40 1.21	3.21 1.19	2.83 1.09	3.33 1.08
Zuversicht	3.75 1.08	3.91 1.08	4.12 .97	3.93 1.05	3.38 1.09	3.89 1.14
Interesse	4.64 1.39	4.03 1.76	4.71 1.14	3.57 1.70	4.57 1.65	4.47 1.77
Langeweile	2.68 1.28	2.59 1.45	2.93 1.27	2.79 1.48	2.43 1.28	2.40 1.45
Attribution Anstrengung	2.68 1.29	3.17 1.37	2.68 1.14	3.75 1.27	2.68 1.48	2.63 1.16
Attribution Fähigkeit	3.39 1.35	2.62 1.37	3.29 1.59	2.50 1.29	3.50 1.16	2.73 1.44
Bewältigungsverhalten (metrischer Wert)	-.68 4.61	1.07 4.56	-1.29 4.27	.36 4.60	-.07 5.01	1.79 4.58
Bewältigungsverhalten (Anteil der Bewältiger)	43%	61%	36%	57%	50%	64%

Bewältigungsverhalten. Anders als in Studie 2 lässt sich Bewältigungsverhalten nun auf metrischem Skalenniveau messen: Die Antwort auf das Trainingsangebot (ja vs. nein) kann gewichtet werden mit der Angabe darüber, wie sicher sich der Proband bei dieser Antwort gewesen ist (sechsstufige Skala). Es ergibt sich so eine Skala von -6 (nein, ganz sicher) bis +6 (ja, ganz sicher), wobei allerdings der Nullpunkt empirisch nicht vorkommen kann. Weiterhin steht freilich auch das kategoriale Bewältigungsmaß zur Verfügung, wie es in Studie 2 verwendet wurde. Auf beiden Maßen lässt sich eine schwach signifikante Beziehung zwischen Intelligenztheorie und Bewältigungsverhalten nachweisen, $F(1,54) = 2.04$, $p < .1$, $\eta^2 = .04$, bzw. $\chi^2(1, N = 56) = 1.79$, $p < .1$ (Yates: $\chi^2 = 1.14$, ns), d.h. Dynamiker tendieren erwartungsgemäß dazu, das Trainingsangebot anzunehmen, während Statiker es eher ablehnen.

Eine Interaktion zwischen Intelligenztheorie und Leistungsrückmeldung liegt, anders als in Studie 2, nicht vor. Auch die erwartete Interaktion zwischen Intelligenztheorie und Geschlecht lässt sich statistisch nicht nachweisen, wenngleich eine entsprechende Tendenz in den Daten auszumachen ist: Dynamikerinnen ($M = 1.38$) zeigen mehr Bewältigungsverhalten als Statikerinnen ($M = -0.59$), Dynamiker ($M = -3.00$) aber weniger als Statiker ($M = -1.00$; $SDs \approx 4$). Die beiden Dispositionsmaße haben keinen Einfluss auf das Bewältigungsverhalten. Kollektivisten sollten gemäß meinen Hypothesen mehr zum Bewältigen neigen als Individualisten. Auch dies trifft empirisch nicht zu, wenngleich sich hier eine (bescheidene) Ergebnistendenz zeigt, $M = 0.86$ zu $M = -.46$ bei $SD \approx 4.5$.

Tabelle 24

Emotionen in Abhängigkeit Impliziter Intelligenztheorien (unter Berücksichtigung von Geschlecht) sowie der beiden Dispositionsmaße

		Hoffnungs- losigkeit	Scham	Schuld	Zuversicht
Induzierte Implizite Theorien (N=57)	Statiker	2.43	3.20	3.12	3.75
	Dynamiker p	2.21	2.72 *	3.28	3.91
-- Frauen (n=48)	Statiker	2.44	3.30	3.15	3.76
	Dynamiker p	2.23	2.78 *	3.27	3.91
-- Männer (n=9)	Statiker	2.38	2.83	3.00	3.72
	Dynamiker p	2.00	2.22	3.33	3.89
Disposition Impliziter Theorien (N=57)	Statiker	2.38	3.06	2.58	2.46
	Dynamiker p	2.27	2.87	2.80 **	2.83
Bedeutsamkeit von Intelligenz (N=57)	Statiker	2.63	3.38	3.46	3.78
	Dynamiker p	2.13 **	2.71 **	3.05	3.86

Anmerkungen.

Die Standardabweichung liegt jeweils bei etwa 1 (in einem Bereich zwischen 0.8 und 1.3). Die Signifikanzangaben betreffen t-Tests, für die Dispositionsmaße lineare Regressionen.

** $p < .05$, * $p < .1$.

Motivationale Tendenzen. Hinsichtlich der Antwort auf die Bewältigungsfrage lassen sich zwei Gruppen voneinander unterscheiden: diejenigen, die das Trainingsangebot annehmen, und diejenigen, die es ablehnen. *Innerhalb* dieser beiden Gruppen lassen sich nun separat voneinander die Hypothesen zu den motivationalen Tendenzen überprüfen: Bei Annehmern sollte das (metrisch erfasste) Bewältigungsverhalten besonders stark ausfallen (d.h.: die angegebene Sicher-

heit der Entscheidung sollte besonders hoch sein), wenn die betreffende Person über Anstachelung und/oder Kompetenzbedürfnis berichtet. Umgekehrt sollte bei den Ablehnern das Bewältigungsverhalten besonders gering ausgeprägt sein, wenn sie sich resigniert und/oder gleichgültig fühlen. Tabelle 25 zeigt, dass sich diese Beziehungen mit Ausnahme von Resignation empirisch nachweisen lassen. Zumindest drei der vier Motivationskonstrukte scheinen also tatsächlich mit dem Bewältigungsverhalten in Verbindung zu stehen. Wenngleich somit eine Grundvoraussetzung für eine sinnvolle Interpretation der Konstrukte erfüllt ist, ist zu beachten, dass deren interne Konsistenzen unbefriedigend klein ausfallen (vgl. Tabelle 25). Zudem korrelieren sie untereinander positiv (Anstachelung zu Kompetenzbedürfnis: $r = .54$, $p < .01$; Resignation und Gleichgültigkeit: $r = .28$, ns , jeweils zweiseitig), was theoretisch freilich nicht zu erklären ist. Auch die Korrelation von Interesse zu Kompetenzbedürfnis, die sehr viel niedriger ausfällt als theoretisch erwartet ($r = .20$, ns ; zum inhaltlich interesseneren Item aber immerhin $r = .39$, $p < .05$, zweiseitig), deutet auf eine unbefriedigende Konstruktvalidität hin.

Wie sieht es nun mit der Beziehung der motivationalen Tendenzen zu den Emotionen aus? Auch hier gibt Tabelle 25 Aufschluss. Anstachelung ist erwartungsgemäß mit Scham und Schuld, nicht aber mit Zuversicht verbunden. Resignation geht hypothesengetreu mit Hoffnungslosigkeit und Scham einher und ist außerdem mit Zuversicht negativ verknüpft. Kompetenzbedürfnis geht zwar mit Schuld einher, entgegen der Erwartung aber nicht mit Anstrengungsattribution. Überhaupt ist Anstrengungsattribution mit allen vier Konstrukten unkorreliert, so dass sich auch die erwartete positive Beziehung zu Gleichgültigkeit empirisch nicht bestätigen lässt. Die Beziehung der motivationalen Tendenzen zu Emotion und Attribution erscheint auch bei Berücksichtigung der anderen Zusammenhänge unklar: So korrelieren etwa *alle* vier Konstrukte hoch positiv mit Hoffnungslosigkeit, Scham und Schuld.

Aufgrund dieser diffusen Ergebnismuster macht es allenfalls für Anstachelung und Kompetenzbedürfnis Sinn, Mediatoranalysen (Baron & Kenny, 1986; vgl. Abschnitt 6.2.1) durchzuführen. Innerhalb der Gruppe der Annehmer ($n = 29$) ist Schuld mit Bewältigungsverhalten ($\beta = .35$, $p < .05$) sowie mit Kompetenzbedürfnis ($\beta = .36$, $p < .05$) verknüpft. Bei simultaner Berücksichtigung von Schuld und Kompetenzbedürfnis verringert sich die Beziehung zwischen Schuld und Bewältigungsverhalten ($\beta = .20$, ns), während Kompetenzbedürfnis in der erwarteten Weise mit Bewältigungsverhalten ($\beta = .45$, $p < .01$) verbunden ist. Kompetenzbedürfnis kann demnach als partieller Mediator der Beziehung zwischen Schuld und Bewältigungsverhalten aufgefasst werden. In ähnlicher Form scheint Anstachelung den Einfluss von Scham und Schuld auf Bewältigungsverhalten zu vermitteln, wenngleich hier geringere empirische Evidenz vorliegt: Zwar gehen Scham ($\beta = .28$, $p < .1$), beobachtete Scham ($\beta = .41$, $p < .05$) und

Schuld ($\beta = .35, p < .05$) in erwarteter Form mit Bewältigungsverhalten und außerdem mit Anstachelung (vgl. Tabelle 25) einher. Bei gleichzeitiger Berücksichtigung von Anstachelung und der jeweiligen Emotion zeigt sich einerseits zwar auch ein Absinken der jeweiligen Korrelationen zu Bewältigungsverhalten (Scham: $\beta = .14, ns$; beobachtete Scham: $\beta = .34, p < .1$; Schuld: $\beta = .29, p < .1$), andererseits aber sind die Werte zwischen Anstachelung und Bewältigungsverhalten nicht (mehr) signifikant (direkter Zusammenhang: $\beta = .31, p < .1$; bei Kontrolle von Scham: $\beta = .23, ns$; bei Kontrolle von beobachteter Scham: $\beta = .19, ns$; bei Kontrolle von Schuld: $\beta = .21, ns$).

Tabelle 25

Korrelative Beziehungsmuster und interne Konsistenzen der motivationalen Tendenzen

	Hoffn.	Scham	Scham (beob.)	Schuld	Zuversicht	Attr. Anstr.	Bew. verh.	Cronbachs Alpha
Anstachelung	.55***	.59***	.36**	.35**	.06	.04	.31*	.45
Kompetenzbed.	.46***	.41**	.24	.36**	.03	.03	.52***	.57
Resignation	.50***	.40**	.33**	.19	-.32**	-.03	.19	.71
Gleichgültigkeit	.42**	.38**	.26*	.29*	-.06	.01	-.32*	.48

Anmerkungen.

Anstachelung und Kompetenzbedürfnis wurden nur für die Gruppe der „Annehmer“ ($n = 29$) erhoben, Resignation und Gleichgültigkeit nur für die Gruppe der Ablehner ($n = 36$).

*** $p < .01$ (einseitig), ** $p < .05$ (einseitig), * $p < .1$ (einseitig).

Statiker und Dynamiker unterscheiden sich nicht signifikant hinsichtlich der vier motivationalen Tendenzen, wenngleich die Ergebnisrichtungen hypothesenkonform sind: Statiker neigen mehr zu Resignation ($M = 2.19$ zu 1.96) und Anstachelung ($M = 2.72$ zu 3.38) und weniger zu Kompetenzbedürfnis ($M = 4.38$ zu 4.53) und Gleichgültigkeit ($M = 4.04$ zu 4.13 ; SD jeweils ≈ 1.2) als Dynamiker.

Prüfung des Gesamtmodells. Analog zu Studie 2 soll abschließend versucht werden, das in Abschnitt 4.1 entwickelte Gesamtmodell an den empirischen Daten zu testen. Zunächst stellt sich wieder die Frage, ob Emotion und Bewältigungsverhalten in der erwarteten Weise miteinander in Beziehung stehen (im vorigen Abschnitt war dieser Frage lediglich in Teilstichproben nachgegangen worden). Überraschenderweise steht das Bewältigungsverhalten mit allen vier Emotionen in positiver Beziehung (Hoffnungslosigkeit: $r = .08$; Scham: $r = .19$; beobachtete Scham: $r = .25$; Schuld: $r = .15$; Zuversicht: $r = .12$). Nur für Schuld und Zuversicht verlaufen die Zusammenhänge also in hypothesengemäßer Richtung, lassen

sich dabei jedoch nicht statistisch absichern. Auf eine Mediator- und Pfadanalyse des Zusammenhangs zwischen Intelligenztheorie, Emotion und Bewältigung, wie sie in Studie 2 zum Einsatz kam, kann demnach verzichtet werden.

Eine Pfadanalyse mit Berücksichtigung der Motivationskonstrukte erscheint aus zwei Gründen ebenfalls nicht erfolgversprechend. Erstens könnte man jeweils nur zwei dieser Konstrukte in der Analyse berücksichtigen und müsste die Stichprobe dabei jeweils reduzieren. Zweitens finden sich zwar einerseits, wie oben erläutert, viele vorhergesagte Zusammenhänge in den Daten wieder, andererseits aber lassen sich an fast ebenso vielen Stellen signifikante Beziehungen ausmachen, die *nicht* erwartet worden sind. Entsprechend unbefriedigend sind die Anpassungswerte des in Abbildung 8 dargestellten Modells (LISREL). Es umfasst nur die erwarteten Teilprozesse um die Motivationskonstrukte Anstachelung und Kompetenzbedürfnis herum, d.h. betrifft diejenigen Probanden, die das Trainingsangebot angenommen haben ($n = 27$). Immerhin lassen sich aber die meisten der vorhergesagten Pfade inferenzstatistisch absichern. Alle Pfade außer demjenigen zwischen Anstrengungsattribution und Kompetenzbedürfnis weisen eine hypothesengemäße Ergebnistendenz auf und nehmen einen größeren Wert als .15 an. Von Leistungssituation zu Anstrengungsattribution wurde ein Pfad eingeräumt, obwohl hier keine Hypothese formuliert worden war. Trotz dieser sowie der oben angesprochenen Probleme liefern die pfadanalytischen Ergebnisse einen interessanten Überblick über die mit Scham und Schuld zusammenhängenden Prozesse. Im Gegensatz zur anderen Modellhälfte passen sich die Daten hier – so kann mit aller Vorsicht zusammengefasst werden – doch ver-

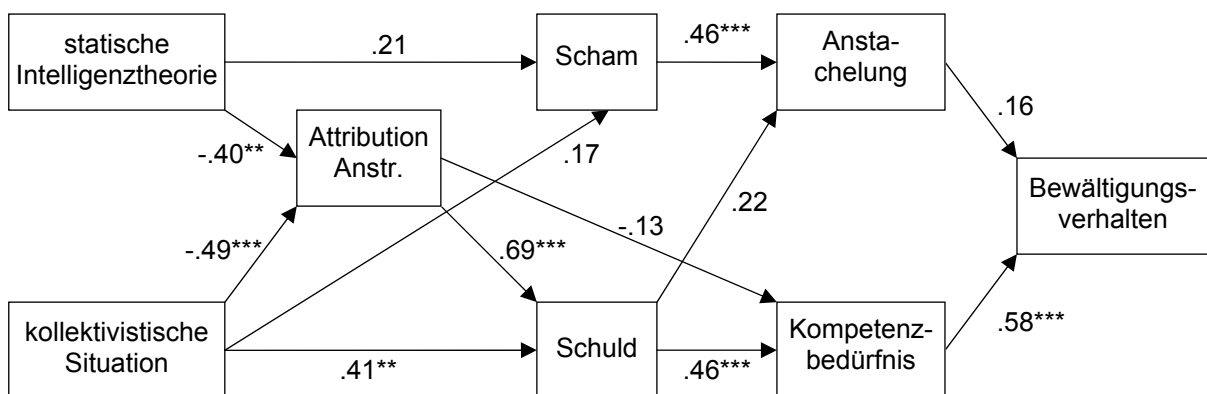


Abbildung 8. Pfadanalytisches Modell zur Prüfung der Hypothesen zu Scham, Schuld, Anstachelung und Kompetenzbedürfnis ($N = 27$).

*** $p < .01$ (zweiseitig), ** $p < .05$ (zweiseitig), * $p < .1$ (zweiseitig).

Modellanpassung: χ^2 ($df = 15$, $N = 27$) = 20.36, $p = .16$; RMSEA = 0.12; NFI = 0.69; CFI = 0.80.

gleichsweise gut den Hypothesen an.

7.2.2 Reaktionen auf positive Leistungsrückmeldungen

In den Studien 1 und 2 konnten die Hypothesen zu den Reaktionen auf positive Leistungsrückmeldungen deutlich nicht bestätigt werden. Daher wäre eine erneute Berücksichtigung dieser Bedingung inhaltlich eigentlich nicht vonnöten gewesen. Da die Erfolgsbedingung aber zur Herstellung der Misserfolgsbedingung zwingend erforderlich ist (jedenfalls insoweit man die Rückmeldung über einen direkten sozialen Vergleich mit einer anderen „natürlichen“ Versuchsperson vornehmen will), liegen auch in Studie 3 wieder Daten für die Reaktionen auf positive Leistungsrückmeldungen vor. Aus Platzgründen soll hier jedoch auf ihre Darstellung verzichtet werden.

7.3 Diskussion

Mit Studie 3 waren vor allem zwei Ziele verfolgt worden. Erstens sollten die interessanten, aber nicht ganz eindeutigen Ergebnisse von Studie 2 an einer größeren Stichprobe repliziert werden. Im Zuge dessen wurden das Versuchsdesign sowie die methodische Vorgehensweise an einigen Stellen in Richtung einer besseren Ökonomie modifiziert. Zweitens wurde mit den sogenannten motivationalen Tendenzen eine neue Gruppe von abhängigen Variablen in die Analyse miteinbezogen, um deren Bedeutung im Zusammenhangsgefüge der anderen Merkmale besser zu verstehen. Im Folgenden soll vor allem die Einlösung dieser beiden Vorhaben diskutiert werden.

Der Faktor Intelligenztheorie übt nur auf einige der Merkmale Einfluss aus, für die Hypothesen gebildet worden waren. Namentlich sind dies Anstrengungsattribution, Scham und Bewältigungsverhalten. In Studie 2 waren hingegen Effekte für Schuld, Zuversicht und Bewältigungsverhalten zu beobachten gewesen. Allein für Bewältigungsverhalten kommen also beide Studien zu einem hypothesenbestätigenden Ergebnis. In Studie 3 liegen immerhin bei allen wichtigen abhängigen Variablen erwartete Ergebnisrichtungen vor. Dem Faktor Leistungssituation kommt ähnlich wie in Studie 2 kaum Bedeutung zu. Allerdings ist zu beachten, dass sich dieser in pfadanalytischen Berechnungen eher stärker bemerkbar macht als in einfachen Gruppenvergleichen ohne Kontrolle andere Variablen. In deutlichem Unterschied zu Studie 2 konnte kein Zusammenhang zwischen Emotion und Bewältigungsverhalten festgestellt werden. Ebenfalls blieben die erwarteten Interaktionen von Geschlecht und Intelligenztheorie aus. Letzteres könnte schlicht als Folge fehlender bzw. geringerer Haupteffekte erklärt werden. Die Deutung dieses Ergebnisses bleibt aber vor allem deswegen Spekulation,

weil die Anzahl von Männern, die für die Studie rekrutiert werden konnten, wiederum äußerst gering war.

Nimmt man all diese Ergebnisse zusammen, ist evident, dass die Resultate von Studie 3 in deutlich geringerem Umfang hypothesenkonform ausgefallen sind als diejenigen von Studie 2. Die Bestätigung der vergleichsweise ermutigenden Befunde der zweiten Studie ist mindestens teilweise fehlgeschlagen. Hierfür lassen sich verschiedene Erklärungsansätze in Anschlag bringen:

- Die *Teilnahmemotivation der Probanden* könnte geringer gewesen sein als in Studie 2, weil sie keine Möglichkeit mehr hatten, ihre „wahren“ Emotionen mit Hilfe der offenen Items zum Ausdruck zu bringen. Auch der höhere geldliche Anreiz zur Teilnahme könnte dazu geführt haben, dass mehr in diesem Sinne extrinsisch motivierte Probanden teilgenommen haben und im Zuge dessen die Motivation, konstruktiv mitzuarbeiten, geringer ausgeprägt war als in Studie 2. Für diese Erklärung spricht erstens, dass die Versuchsleiterin in der Tat beobachtete, dass viele Versuchspersonen den Emotionsfragebogen relativ lustlos bearbeiteten; die beiden Extremfälle in dieser Hinsicht wurden wie gesagt aus der Analyse ausgeschlossen. In diese Richtung lassen sich auch die geringeren Reliabilitäten und Konstruktvaliditäten der Emotionsskalen deuten. Zweitens scheint für diese Erklärungsvariante auch die Tatsache zu sprechen, dass der Effekt für das Bewältigungsverhalten als sehr verhaltensnahe Maß repliziert werden konnte, die Effekte für die meisten auf Selbstbericht beruhenden Merkmale hingegen nicht.
- Die *zeitliche Dauer der Induktion Impliziter Intelligenztheorien* fiel aus versuchsökonomischen Gründen deutlich geringer aus als in Studie 2. Die Probanden mussten statt vier nur noch drei Fragen zu dem Spiegel-Artikel beantworten. Außerdem mahnte die Versuchsleiterin eindringlicher dazu, die vorgeschlagene Bearbeitungszeit nicht zu überschreiten. Die Beschäftigung mit der Botschaft des Artikels verkürzte sich auf diese Weise nach Schätzungen der Versuchsleiterin von 15-25 Minuten (Studie 2) auf 8-15 Minuten (Studie 3). Wenngleich diese Modifikation im Sinne einer „unvollständigen Induktion“ durchaus gewissen Einfluss auf die Ergebnisse ausgeübt haben könnte, so erscheint doch unwahrscheinlich, dass sie allein für das Scheitern der Hypothesen verantwortlich ist.
- Schließlich lassen die Diskrepanzen zwischen den Befunden der Studien 2 und 3 auch den *inhaltlichen Schluss* zu, dass die Konfrontation mit den Induktionstexten hier wie dort keinen Einfluss auf die nachfolgenden Reaktionen auf Misserfolg hat. Waren die Resultate von Studie 2 bereits unklar und inkonsistent, so hat sich in Studie 3 anhand einer größeren Stichprobe ein Ergebnismuster durchgesetzt, das bis auf wenige Ausnahmen für die Annahme der Nullhypothese spricht. Diese Ausnahmen wären in dieser Sichtweise – dann ebenso wie die Effekte in Studie 2 – als dem Zufall geschuldet zu interpretieren.

Welcher Erklärung ist mehr zu trauen? Zeigen die Ergebnisse von Studie 3 lediglich, dass die Induktion Impliziter Theorien nicht gelingt, wenn nicht genug Sorgfalt auf die Induktion der Impliziten Theorien sowie auf die Motivierung der Probanden gelegt wird? Oder legen sie vielmehr den Schluss nahe, dass eine Beeinflussung des Verhaltens über die verwendeten Induktionstexte nicht möglich

ist und somit die bislang vorliegenden Untersuchungen – die sich ja in oft trivialer Manier bei der Charakterisierung der Leistungsaufgabe auf die Intelligenzsemantik rückbezogen haben – in ihrer Aussagekraft infragegestellt werden müssen? Sprechen andererseits nicht die fehlenden Beziehungen zwischen der Disposition Impliziter Theorien und den abhängigen Variablen gerade für einen übergeordneten, d.h. nicht speziell induktionsbezogenen Grund für das Scheitern der Hypothesen, der etwa in der Kultur oder Subkultur der Probanden zu suchen ist? Die Daten lassen hier sicherlich keinen letztgültigen Schluss zu. Mit Bezug auf theoretische und empirische Quellen sollen diese Fragen jedoch in den Abschnitten 8.2 und 9 näher untersucht werden.

Aufgrund interessanter explorativer Befunde in Studie 2 wurden in Studie 3 vier motivationale Tendenzen erhoben, die die Beziehung zwischen Emotion und Bewältigungsverhalten aufzuklären helfen sollten. Dieser Anspruch konnte nur bedingt eingelöst werden. Probleme zeigten sich bereits bei der Reliabilität der Items sowie den Beziehungen zwischen den vier Skalen. Allem Anschein nach ist eine solch direkte Befragung über die Gründe, aus denen heraus mehr oder weniger Bewältigungsverhalten gezeigt wird, wenig erfolgversprechend. Dennoch ließen sich einige Zusammenhangshypothesen bestätigen, wobei insbesondere Anstachelung und Kompetenzbedürfnis die Beziehung zwischen Emotion und Bewältigungsverhalten teilweise zu vermitteln scheinen. Einiges deutet also darauf hin, dass Emotionen wie Scham und Schuld nicht einseitig positive oder negative, sondern oftmals ambivalente motivationale Auswirkungen haben (vgl. Abschnitt 6.3 für eine eingehendere inhaltliche Plausibilisierung). Auf diesem Wege ließe sich auch die Uneindeutigkeit in den Beziehungen zwischen Intelligenztheorie und Motivation bzw. Verhalten erklären, indem nämlich geltend gemacht werden könnte, dass beispielsweise die statische Intelligenztheorie nicht nur motivationsabträgliche, sondern auch motivationsförderliche Implikationen in sich trägt. Die Ergebnismuster sind allerdings alles andere als eindeutig, so dass aus den hier vorgestellten Befunden wenig mehr abzuleiten ist als die Vermutung, dass zukünftige, methodisch adäquatere Studien zu dieser Fragestellung lohnenswert sein dürften.

Die Reaktionsmuster nach positiver Leistungsrückmeldung entsprechen weitgehend den in Studie 2 gezeigten. Ähnlich wie dort konnten die Hypothesen nicht bestätigt werden. Auch die Probleme hinsichtlich Reliabilität und Konstruktvalidität sind dieselben wie in Studie 2 und werden daher hier nicht noch einmal diskutiert. Auch für die allgemeinen Bemerkungen über die Generalisierbarkeit sowie die Effektstärke der referierten Befunde sei auf Abschnitt 6.3 verwiesen.

Dritter Teil: Schlussfolgerungen und Diskussion

8. Zusammenfassung und Vergleich der Ergebnisse

In diesem Kapitel soll eine zusammenfassende Bewertung und Einordnung der Ergebnisse der drei durchgeführten Untersuchungen vorgenommen werden. Dafür werden zunächst deren Befunde einander gegenübergestellt und unter Berücksichtigung der jeweils gewählten Operationalisierungen in ihrer Aussagekraft eingeschätzt. Daran anschließend geht es um eine Bewertung der oft uneindeutigen Ergebnisse im Vergleich zu früheren Untersuchungen anderer Forscher. Für eine Diskussion der inhaltlichen Implikationen sowie praktischen Konsequenzen der Befunde sei auf Kapitel 9 verwiesen.

8.1 Vergleichende Diskussion der drei Untersuchungen

Die Hypothesen zu Reaktionen auf negative Leistungsrückmeldungen konnten anhand von in-vivo-Leistungsaufgaben (Studien 2 und 3) besser bestätigt werden als anhand von hypothetischen Szenarien (Studie 1). Die Erwartungen hinsichtlich positiver Leistungsrückmeldungen hingegen fanden hier wie dort keine Entsprechung in den Daten. Leistungsverhalten nach Erfolg scheint nicht davon abzuhängen, wie die Person über Intelligenz denkt oder in welchen sozialen Abhängigkeiten sie sich sieht. Die unbefriedigenden Messeigenschaften der Emotionsskalen deuten ferner daraufhin, dass die Probanden Schwierigkeiten hatten, ihre affektiven Reaktionen so differenziert zu beschreiben, wie es die Items verlangten und wie es mithin auch Voraussetzung für das Bestätigen der diesbezüglichen Hypothesen war. Die folgende Diskussion konzentriert sich daher auf die Misserfolgsbedingung.

In Abschnitt 5.3 waren bereits Erklärungsmöglichkeiten für das Ausbleiben der Effekte in der Szenario-Studie erörtert worden. Ein wichtiger Nachteil im Vergleich zu in-vivo-Studien besteht u.a. darin, dass die Probanden über Erleben und Verhalten berichten sollen, von dem sie sich lediglich vorstellen, dass sie es in einer bestimmten Situation zeigen könnten. Dieses methodische Vorgehen mag bei emotionspsychologischen Studien, die sich mit allgemeinen Einschätzungsgrundlagen von Emotionskonzepten beschäftigen, angebracht sein (vgl. Robinson & Clore, 2001); für die Abbildung subtiler Gruppenunterschiede in Abhängigkeit Impliziter Intelligenztheorien jedoch scheint es zu grobkörnig zu sein. Darüber hinaus lässt sich aus einem Vergleich der Ergebnisse auch ableiten, dass die Szenario-Methode zuweilen sogar irreführende Resultate liefert. So zeigen die Probanden in Studie 1 deutlich mehr Scham und Schuld in der kollektivistischen als in der individualistischen Situation, während sich in den in-vivo-

Studien keine bzw. keine durchgehenden Unterschiede finden. Möglicherweise richten sich die Probanden in Studie 1 also lediglich nach einer sozialen Erwartung, gemäß der man in Gruppenkontexten mehr selbstrelevante Emotionen empfindet bzw. empfinden sollte, und verkennen dabei, dass sie sich bei einer konkreten Leistungsaufgabe gar nicht in entsprechender Form verhalten (wollen) würden. In diesem Sinne gibt die vorliegende Arbeit durchaus dazu Anlass, die Angemessenheit der Szenario-Methode auch für klassische emotionspsychologische Fragestellungen in Zweifel zu ziehen (vgl. Parkinson & Manstead, 1992).

Die in-vivo-Studien können die Hypothesen einerseits allgemein vergleichsweise gut bestätigen. Andererseits zeigen sich hier aber deutliche Unterschiede zwischen den untersuchten Merkmalen. Am besten lässt sich der Einfluss von Impliziten Theorien auf das *Bewältigungsverhalten* nachweisen. Dies könnte mit der besonders großen Verhaltensnähe dieses Merkmals zusammenhängen, d.h. es ist am schwierigsten durch Tendenzen zu sozialer Erwünschtheit zu beeinflussen: Während ein Proband ohne Mühe ankreuzen kann, er sei durch die negative Rückmeldung überhaupt nicht emotional berührt, geht er eine soziale Verpflichtung ein, wenn er sich zur Teilnahme an dem Training bereiterklärt. Eine positive Selbstdarstellung wird somit erschwert. Die Tendenz etwa, sich in solch einem Training nicht vor anderen blamieren zu wollen, lässt sich so kaum verbergen. Dieser Aspekt scheint nicht zuletzt deshalb von so großer Bedeutung, weil Statiker ja gemäß einschlägigen Befunden stärker zu positiver Selbstdarstellung tendieren sollten als Dynamiker (vgl. Kapitel 5.3). In diesem Sinne könnte das Selbstdarstellungsbedürfnis der Statiker ihr dysfunktionales Leistungsverhalten (z.B. das Erleben von Hoffnungslosigkeit) empirisch verwischen bzw. „kompensieren“.

Umgekehrt lässt sich für die Beziehung zwischen Intelligenztheorien und *Emotionen* geltend machen, dass hier, wo es um den Selbstbericht über innere Zustände geht, jener Kompensationseffekt eine bedeutende Rolle spielen könnte. Die eher unklaren Befunde hinsichtlich Emotion könnten also teilweise so zu erklären sein. Insbesondere die Ergebnisse von Studie 2 unterstützen diese Deutungsweise, indem gerade für die sozial unerwünschten Emotionen Hoffnungslosigkeit und Scham keine Unterschiede auffindbar sind, für Schuld, Zuversicht und tendenziell sogar für die durch Beobachtung erhobene Schamvariable dagegen schon. Studie 3 allerdings liefert geradezu spiegelbildliche Resultate, indem nun einzig für Scham Unterschiede feststellbar sind. Auch für *Anstrengungsattribution* ergeben sich ähnliche Inkonsistenzen zwischen den Studien, so dass der Kompensationseffekt als Erklärung für die unklaren Befunde sicherlich nicht ausreicht (vgl. Abschnitt 9.2 für stärker theoretisch orientierte Erklärungen).

Was für das Erleben von Emotionen gilt, gilt für das *Darstellen von Emotionen* erst recht. Jeder Fragebogen ist auch ein Mittel zur Selbstdarstellung – und

sei es nur vor sich selbst (vgl. Laux, 1993). Wenn Fragen zum Erleben die Darstellungsziele der Probanden anregen, dann tun es ebenso – oder vielleicht sogar noch stärker – Fragen zum Darstellen. Die Probanden müssen sich dann gewissermaßen in ihrer Darstellung darstellen, beispielsweise indem sie vorgeben, „authentisch“ zu sein bzw. sich „nicht“ darzustellen. Als weiteres Problem kommt bei der Erfassung der Emotionsdarstellung hinzu, dass die verwendeten Items notwendigerweise recht kompliziert ausfallen müssen. Deren geringe Verständlichkeit lässt sich etwa aus teilweise absurden korrelativen Beziehungen zwischen den Konstrukten ableiten. So korrelieren in Studie 2 Schamdarstellen und Schamverstecken hoch positiv zueinander. Dazu passt, dass es die eher schlicht formulierten Items zu Interesse und Langeweile sind, bei denen Gruppenunterschiede festgestellt werden konnten. Im Ganzen erscheint die Befundlage zur Emotionsdarstellung aber unklar: Teils sind die erwarteten Unterschiede vorhanden, meist aber bleiben sie – sei es aus methodischen oder inhaltlichen Gründen – aus.

Ähnlich ambivalent fallen die Ergebnisse auch bei anderen Hypothesen aus. Nur in Studie 2 gibt es die erwarteten Zusammenhangsmuster zwischen Emotion und Bewältigungsverhalten sowie die vorhergesagten Geschlechtsdifferenzen (vgl. hierzu Abschnitt 9.2). Die weiteren Auswirkungen Impliziter Theorien auf die Lernleistung ließen sich nur anhand problematischer Operationalisierungen überprüfen; es fanden sich zwar einige interessante Mediationsprozesse, aber nicht der vorhergesagte Gruppenunterschied zwischen Statikern und Dynamikern. Die Hypothesen zum Faktor Leistungssituation konnten nur in Einzelfällen bestätigt werden, die durchaus auch dem Zufall zugeschrieben werden können (vgl. auch Abschnitt 9.2). Die Wirkungen der induzierten und dispositionalen Intelligenztheorie erfolgen unabhängig voneinander. Unterschiede zwischen neurowissenschaftlichem und psychologischem Leistungsdiskurs waren in Studie 1 nicht vorhanden, während in den Studien 2 und 3 dieser Faktor nicht mehr variiert wurde. Inwieweit die dortigen Ergebnisse dennoch praktische Implikationen besitzen, wird in Abschnitt 9.3 diskutiert.

8.2 Gegenüberstellung der Befunde zu bestehenden Untersuchungen

Bis hierhin sollte deutlich geworden sein, dass die wichtigsten Forschungsfragen der vorliegenden Arbeit nicht eindeutig beantwortet werden können, da die Befundlage widersprüchlich ist. Bei aller Unklarheit ist nichtsdestotrotz evident, dass die hier durchgeführten empirischen Studien dem Dweckschen Modell *weniger gut* entsprechen als die bislang vorliegenden Untersuchungen. Prototypisch lässt sich etwa die Untersuchung von Hong et al. (1999, Studie 3) als Vergleich heranziehen, deren methodisches Vorgehen dem der hier durchgeführten Studien 2 und 3 sehr ähnlich ist und bei der sich hoch signifikante Unterschiede für Be-

wältigungsverhalten und Attribution zeigen (vgl. auch Abschnitt 3.2.1). Zeigen sich bei Hong et al. etwa für Anstrengungsattribution Mittelwertsunterschiede von 4 zu 2, betragen sie in Studie 2 – bei ähnlichen Skaleneigenschaften – lediglich 2 zu 1.8. Solch geringe Unterschiede werden oft bestenfalls auf dem 10%-Niveau signifikant. Auch die Interventionsstudien konnten deutlichere und konsistentere Verhaltensänderungen durch die Induktion Impliziter Theorien nachweisen, wobei hier die Induktionsphase zwar teilweise etwas länger und intensiver, aber qualitativ prinzipiell vergleichbar war. Im Folgenden sollen daher mögliche Ursachen dieser Diskrepanz in ihrer Stichhaltigkeit und ihren Implikationen für künftige Studien erörtert werden. Dabei sollen lediglich die Studien 2 und 3 einem solchen Vergleich unterzogen werden, da das Ausbleiben der Effekte in Studie 1 bis hierhin bereits recht gut aufgeklärt erscheint.

Zunächst lassen sich einige Sachverhalte identifizieren, die darauf hindeuten, dass die Studien 2 und 3 einige in früheren Untersuchungen begangene methodische Fehler vermeiden und den kausalen Einfluss Impliziter Intelligenztheorien mithin auf eine inhaltlich angemessenere Art und Weise prüfen. Zu nennen sind hier vor allem zwei Punkte:

- Im Gegensatz zu den Texten Bergens (1991) treffen die hier verwendeten Artikel keine Aussage über die Bedeutsamkeit von Intelligenz. Es geht explizit allein um die Frage der Stabilität von Intelligenz. Zwar wird implizit gelegentlich die Auffassung vertreten, dass Intelligenz wichtig ist für den Lebensweg eines Menschen (vgl. die Rahmengeschichte über das Kleinkind Adam, Abschnitt 4.2). Diese Auffassung wird aber sowohl im statischen als auch im dynamischen Artikel gleichermaßen geltend gemacht. Bergen hingegen variiert beide Faktoren unabhängig voneinander, so dass vier Texte resultieren, von denen also zwei behaupten, Intelligenz korreliere hoch mit Erfolg in allen Lebensbereichen. Wenngleich dies von den Autoren leider nicht in Erfahrung gebracht werden konnte, spricht einiges dafür, dass Studien wie Hong et al. (1999) die beiden „zugespitzten“ Versionen verwenden, d.h. die statisch-bedeutsame und die dynamisch-unbedeutsame Variante (vgl. Abschnitt 4.2.1). Dass durch diese – recht künstlich anmutende – Konfundierung der Faktoren Stabilität und Bedeutsamkeit größere Effekte erzielt werden könnten als durch die gemäß Dwecks Modell vorgesehene Beschränkung auf den Faktor Stabilität, erschien wenig erstaunlich.
- Studien wie die von Hong et al. verwenden bei der inhaltlichen Gestaltung der Leistungsaufgabe die Intelligenzsemantik, d.h. sie lassen ihre Probanden z.B. explizit einen „Intelligenztest“ bearbeiten (vgl. Abschnitt 3.2.1). Wenn der Proband vorher etwas über die äußerst hohe Stabilität von Intelligenz gelesen hat, erscheint es nur noch trivial, dass der Proband in der Rückmeldesituation an einem Training seiner Intelligenz nicht besonders interessiert ist.

Es handelt sich dann gerade nicht um eine hilflose, sondern um eine rationale Verhaltensreaktion (zu vergleichen mit der eines Querschnittsgelähmten, der seine Zeit nicht mit dem Versuch verschwendet, wieder laufen zu lernen). Während die externe Validität früherer Untersuchungen also in Zweifel gezogen werden kann, dürften die in den hier durchgeführten Studien 2 und 3 erzielten, vergleichsweise moderaten Effekte auf das Bewältigungsverhalten ein realistischeres Bild über den motivationalen Einfluss induzierter Intelligenztheorien liefern.

Vor dem Hintergrund dieser beiden methodischen Probleme erscheint es unverständlich, dass kaum Versuche zur Replikation der experimentellen Untersuchungen á la Hong et al. unternommen wurden, während korrelative Untersuchungen über Implizite Theorien mittlerweile in kaum mehr überblickbarer Fülle vorliegen. Überblicksarbeiten wie die von Dweck und Molden (2005) suggerieren eine Vielzahl von Untersuchungen zu diesem Thema und verweisen doch immer nur auf dieselben zwei Studien. Zwar liegen außerdem eine Reihe von Arbeiten vor, die die Induktion der Intelligenztheorien durch schlichte Kennzeichnung der zu bearbeitenden Leistungsaufgabe als „eine, bei der es vor allem auf geistige Fähigkeiten bzw. Konzentration und Anstrengung ankommt“ vornehmen wollen (z.B. Martocchio, 1994). Anzuzweifeln ist bei dieser – verführerisch ökonomischen – Operationalisierung jedoch, ob überhaupt noch irgendeine Beziehung zu dem Konzept der Impliziten Intelligenztheorie sensu Dweck vorliegt. Jene Untersuchungen spielten daher in der vorliegenden Arbeit keine Rolle.

Während die obigen Argumente dafür sprechen, dass der Einfluss induzierter Intelligenztheorien bislang eher *überschätzt* wurde, lässt sich in manchen Punkten jedoch auch das hier gewählte methodische Vorgehen kritisieren. Folgende Sachverhalte könnten demnach dafür sprechen, dass die Induktionseffekte in der vorliegenden Arbeit eher *unterschätzt* werden:

- Die etwas geringeren Effekte für das Merkmal Bewältigungsverhalten könnten darauf zurückzuführen sein, dass Studien wie Hong et al. eine noch verhaltensnähere Operationalisierung wählen, die den Probanden nicht nur eine Absichtserklärung zur Teilnahme an einer Trainingsmaßnahme abverlangt, sondern sie diese auch gleich im Rahmen des Versuchs durchführen lässt (vgl. Abschnitt 3.2.1). Auf diese Weise können Hong et al. erfassen, wer an der Maßnahme tatsächlich teilnimmt, während die Studien 2 und 3 der vorliegenden Arbeit lediglich erheben, wer sich dazu bereit erklärt, demnächst einen Termin für die Maßnahme zu verabreden. Insofern könnte man statt „Bewältigungsverhalten“ den Begriff „Bewältigungsmotivation“ als eine angemessenere Kennzeichnung des Merkmals betrachten. Denkbar ist etwa, dass manche Statiker zwar aus Gründen der Selbstdarstellung zunächst ihre

Teilnahme am Training signalisiert haben, diese später jedoch unter Ausreden wieder abgesagt hätten, so dass der wahre Unterschied im Bewältigungsverhalten unterschätzt worden ist. Mit der Operationalisierung von Hong et al. hätte zudem vielleicht das bereits erwähnte Problem vermieden werden können, dass viele Lehramtsstudierende ihr Ablehnen des Trainingsangebots nach dem Versuch informell damit begründeten, dass sie in ihrem weiteren Karriereweg nicht mit derartigen (in der Instruktion angesprochenen) Einstellungstests konfrontiert würden. Führt man sich jedoch die Vorteile der hier gewählten Operationalisierung des Bewältigungsverhaltens vor Augen, die etwa in ihrer Realitätsnähe, Glaubwürdigkeit und Plausibilität zu sehen sind, ist damit zu rechnen, dass diese kaum weniger Störvarianz verhindert haben, als die Nachteile produziert haben. Eine nachhaltige Beeinträchtigung der Reliabilität oder Validität des Bewältigungsmaßes erscheint somit zwar nicht ausgeschlossen, aber doch unwahrscheinlich.

- Auch der in den Studien 2 und 3 gewählte soziale Charakter der Rückmeldungssituation könnte problematisiert werden. Die Tatsache, dass immer zwei Probanden gleichzeitig mit ihrer Leistung konfrontiert wurden, könnte nämlich zu einer übermäßigen Aktivierung sozialer Einstellungen und Verhaltensstile geführt haben, die die eigentlich interessierenden Effekte Impliziter Theorien überlagert haben könnten. Anders herum lassen sich aber auch Vorteile eines sozialen Versuchsaufbaus ausmachen, etwa die größere Realitätsnähe und stärkere emotionale Aktivierung. Gemäß dieser Sichtweise, die bei der Planung der Studien 2 und 3 eine zentrale Rolle spielte, wäre also eher mit stärkeren als mit schwächeren Gruppenunterschieden zu rechnen gewesen. Klar ist in diesem Kontext wohl nur, dass das simultane Testen zulasten der Versuchsökonomie ging, da dies die Terminabsprachen erschwerte und die Erfolgsprobanden von geringem inhaltlichem Interesse waren.

Schließlich kommen als Erklärungsmöglichkeiten für die erwähnte Diskrepanz zwischen den hier vorgestellten und anderen Untersuchungen freilich auch inhaltliche Ansätze infrage. Etwa lässt sich – auch angesichts der wenigen vorliegenden Studien – eine experimentelle Wirkung induzierter Intelligenztheorien gänzlich in Zweifel ziehen. Gegen diese Deutung spricht allerdings die Tatsache, dass in der vorliegenden Arbeit nicht nur die induzierten, sondern auch die dispositionalen Impliziten Theorien deutlich geringer und inkonsistenter mit den abhängigen Variablen in Verbindung stehen, als dies bei den einschlägigen korrelativen Studien der Fall ist (vgl. Abschnitt 3.1). Man müsste also auch die korrelative Version von Dwecks Modell infragestellen, was im Lichte zahlreicher empirischen Bestätigungen nicht haltbar erscheint. Angemessener ist da wohl eine Erklärungsvariante, die jene Diskrepanz auf die Kultur bzw. Subkultur der Probanden zurückführt. Hier ist allerdings zu beachten, dass die internationale

Befundlage keineswegs eindeutige Aussagen in diese Richtung zulässt (vgl. Abschnitt 5.3); eine genauere Betrachtung dieser Frage in metaanalytischer Form wäre vonnöten. Die Stichhaltigkeit der kulturell ausgerichteten Erklärungsvariante kann daher hier nicht näher eingeschätzt werden. Für eine eher theoretisch gestützte Analyse der Beziehung zwischen Impliziten Theorien und Kultur sei aber auf Abschnitt 9.4 verwiesen.

Zusammenfassend erscheinen zwei Erklärungen für die uneinheitlichen und oft schwachen Auswirkungen induzierter Impliziter Theorien auf das Leistungsverhalten nach Misserfolg stichhaltig. Erstens könnte die Auseinandersetzung mit Zeitschriftenartikeln wie den verwendeten keinen so großen Einfluss ausüben, wie frühere Studien dies nahelegen. Die in diesen Studien gefundenen Effekte könnten auf die angesprochenen methodischen Schwächen zurückzuführen sein, so dass die vorliegende Arbeit jene Effekte in neuem, kritischerem Licht erscheinen ließe. Zweitens könnten sich die Probanden der vorliegenden Untersuchungen so sehr in ihrer (Sub-)Kultur von denjenigen früherer Studien unterscheiden haben, dass eine Replikation ihrer Befunde aus diesen populationsspezifischen Gründen gescheitert sein könnte. Diese Erklärungsansätze schließen sich im Übrigen nicht gegenseitig aus, so dass beide ihren Teil zu den uneindeutigen Ergebnissen beigetragen haben könnten. Den allgemeinen methodischen Begrenzungen der vorliegenden empirischen Studien, die etwa in der Beschränkung auf künstliche Leistungsumwelten sowie dem Messen der abhängigen Variablen anhand weniger, kaum vorgetesteter Items bestehen, sollte dagegen nicht zu viel Bedeutung zugemessen werden, da diese Probleme in den angesprochenen früheren Untersuchungen in ähnlicher Form vorliegen.

9. Schlussfolgerungen und praktische Implikationen

Während es in Kapitel 8 um die Güte und Aussagekraft der empirischen Befunde und im Zuge dessen um eher methodische Aspekte ging, sollen nun ihre inhaltlichen Implikationen für Forschung und Anwendung thematisiert werden. Dabei wird die Uneinheitlichkeit der Befunde zumeist nicht mehr ausführlich zur Sprache gebracht, sondern mitunter eine Einschätzung ihrer Konsequenzen angeboten, die ihre hypothesengemäße Deutung voraussetzt.

Abschnitt 9.1 thematisiert dabei die Rolle von Impliziten Intelligenztheorien für Lernmotivation und -verhalten, während es Abschnitt 9.2 stärker um das prozessuale Zustandekommen dieses Zusammenhangs geht. Abschnitt 9.3 unternimmt den Versuch, die Bedeutung der empirischen Befunde für die Frage nach dem Einfluss neurowissenschaftlicher Leistungsdiskurse auf das Lernverhalten auszuloten. Der abschließende Abschnitt 9.4 stellt es sich zur Aufgabe, das Konzept der Impliziten Intelligenztheorie von einem kulturpsychologischen Standpunkt aus theoretisch anzureichern, um es auf diese Weise neuen empirischen Forschungsbemühungen zugänglich zu machen.

9.1 Implizite Intelligenztheorien als Ursache von Motivationsproblemen

Die empirischen Untersuchungen haben deutlich gemacht, dass die Konfrontation mit wissenschaftlichen Argumenten über die Natur der Intelligenz das Leistungsverhalten von Studierenden beeinflussen kann. Die Vorstellung einer stabilen Intelligenz hat dabei überwiegend dysfunktionale, die Vorstellung einer veränderbaren Intelligenz hingegen meist funktionale Auswirkungen. Diese Resultate stehen im Einklang mit früheren Untersuchungen (vgl. Abschnitt 3.2), wobei die vorliegende Arbeit erstmals auch Hinweise darauf gibt, dass Intelligenz nicht als angeboren angesehen werden muss, um die leistungsabträglichen Effekte hervorzurufen. Es reicht bereits aus, wenn die Veränderbarkeit von Intelligenz *nach einem bestimmten Zeitpunkt der Entwicklung* (frühe Kindheit, Pubertät etc.) in Abrede gestellt wird.

Dieser Befund passt sich nicht nur gut in die Literatur zu Impliziten Intelligenztheorien ein, sondern auch in verwandte Forschungsansätze wie die attributionstheoretische Leistungsmotivationsforschung oder die Zielorientierungsforschung (vgl. die Literaturangaben in Kapitel 3), die anhand von Konzepten wie Fähigkeitsattribution oder Leistungszielorientierung zu ähnlichen Befunden kommen wie der Dwecksche Ansatz anhand des Konzepts der statischen Intelligenztheorie. Das Innovative an diesem Konzept liegt jedoch darin begründet, dass es nicht nur korrelative, sondern auch experimentelle Studien zulässt, ohne dabei den Bezug zur praktischen Anwendung zu verlieren. Zwar konnte auch im

Rahmen von Studien zu Reattributionstrainings (z.B. Dresel & Ziegler, 2006; Ziegler & Heller, 1998) gezeigt werden, dass Attributionsprozesse nicht nur Begleitvorgänge anderer, grundlegenderer Abläufe sind, sondern eine eigenständige Wirkungskraft auf Motivation und Verhalten ausüben. Im Vergleich zu diesen Programmen, die meist eine künstliche, stark individualisierte Lernumwelt voraussetzen und somit aufwändig und gleichzeitig in ihren Auswirkungen inhaltlich begrenzt sind, liefert das Konzept der Impliziten Intelligenztheorie jedoch eine einfachere, direktere Möglichkeit zur Einflussnahme auf die Lernmotivation. Wie diese praktisch aussehen könnte, wird am Beispiel des neurowissenschaftlichen Diskurses in Abschnitt 9.3 näher ausgeführt.

Dass subtile sozial-kognitive Prozesse bedeutende und langfristige Auswirkungen auf die Lernmotivation haben können, weiß man wie gesagt nicht erst seit den Untersuchungen Dwecks. Stichworte wie Erlernte Hilflosigkeit, Bezugsnorm, Anstrengungsvermeidung oder Pygmalion-Effekt stehen stellvertretend für eine Fülle von Ansätzen, die die leistungsbezügliche Bedeutung von Merkmalen und Prozessen jenseits der intellektuellen Grundvoraussetzungen nachgewiesen haben. Dennoch mag erstaunen, dass allein die Lektüre und Diskussion eines Zeitschriftenartikels eine nicht unerhebliche Wirkung auf das Lernverhalten hat. Allem Anschein nach sprechen die Artikel fundamentale Einstellungen über Lernen und Leistung an und verändern – in Dwecks Worten – das „meaning system“ des Betroffenen in nachhaltiger Art und Weise. Dass diese Prozesse in einigen Diskussionen über die Artikel auch explizit auszumachen sind, sollen die folgenden vier Auszüge aus den in Studie 2 angefertigten Videomitschnitten beispielhaft illustrieren.

Ausschnitt 1 (die Spiegelstriche trennen zwischen den Beiträgen der beiden Diskutanten)

- Ja ist doch scheiße einzusehen dass man nicht weiterkommt im Leben. Man denkt ja immer man macht viel und erreicht damit was... aber wenn alles schon vorbestimmt ist, ist ja irgendwie...
- Perspektivlosigkeit...
- Ja schon irgendwie, ne? Es geht eh nicht weiter, also du kannst dir zwar sicher Wissen aneignen, aber an deiner Intelligenz selber kannst du eh nichts ändern.
- Zukunftspessimismus...
- Ja, dann kannst dir halt nicht sagen, ich reiße aus, also... mein Umfeld war zwar so, dass ich 'n totaler Trottel sein müsste, aber ich werd jetzt Bundespräsident, aber das wirste halt nicht, weil dir fehlt die Intelligenz...

Ausschnitt 2

- Wenn seine Begabung optimal gefördert wird..., dass er mal eine bedeutende Persönlichkeit wird..., dass er mal gute Chancen hat, eine gute gesellschaftliche Position zu erlangen.
- Bescheuert, echt. Das erinnert mich voll an diese Selektionsprozesse im Dritten Reich. Keine Ahnung, ich find das irgendwie gefährlich. [...]
- Da siehste dann ja, die Pädagogik ist ja auch ab dem dritten Lebensjahr so ungefähr... nutzlos, quasi, weil dann ist ja sowieso schon alles gelaufen ungefähr, dann müssen wir nur noch lernen, wie man zählt, schreibt, fertig.

Ausschnitt 3

- ...dass man früh auch schon quasi selektieren sollte. Aber das fand ich auch 'n bisschen krass, so, im Alter von drei Jahren ist eh schon alles abgeschlossen, deshalb können wir dann ja auch schon anfangen auszusortieren...
- Ach ja genau, man soll früh die Talente erkennen und dann kanalisieren.
- Kanalisieren, genau, das war das Wort, genau. (schmunzelt)

Ausschnitt 4

- Die Lehrer, wenn die schon erkennen, dass das Kind in Mathe nichts taugt, dann...
- ...dann werden die schneller bevurteilt.
- Ich fand das nämlich auch sehr eigenartig, dass die geschrieben haben..., dass die Intelligenz...
- ...dann abgeschlossen ist und dann nichts mehr großartig möglich ist, das fand ich auch etwas eigenartig.
- ...nee, also das fand ich... überraschend. Aber dass sie das halt so trennen wollen, dass man es einsehen soll, es gibt dumme Kinder und es gibt schlaue Kinder, irgendwie hab ichs Gefühl, und das fand ich irgendwie... so ne Klassifizierung von Menschen, nur weil man schlau ist, ist man was Besseres.
- Ich würde schreiben, es ist ne Gefahr für die Pädagogik, weil man halt so zu Vorurteilen neigt. Wenn man gesagt bekommt, er ist dumm, dann behandelt man ihn auch so.
- Man sollte trotzdem gleichbehandelt werden, auch wenn es da Unterschiede gibt, in der Intelligenz...

Die Ausschnitte dokumentieren verschiedene Facetten der negativen Bewertung, die die Probanden dem Artikel über den stabilen Charakter der Intelligenz gewöhnlich entgegenbrachten. Unterscheiden könnte man hier etwa zwischen individueller Hoffnungs- und Perspektivlosigkeit (Ausschnitt 1), der besonderen Förderung Hochbegabter bei gleichzeitiger früher Aussortierung Minderbegabter (Ausschnitte 2 und 3) und der Bewertung eines Menschen allein aufgrund seiner Intelligenz (Ausschnitt 4). Während derartig explizite Kritik am Stabilitätstext in sieben von 16 Fällen geäußert wurde, kam es nur in einem einzigen von 17 Fällen zu einer negativen Bewertung des Veränderbarkeitstextes (Zitat: „Dann müsste es ja prinzipiell ideale Umgebungen geben, die man schaffen könnte für Kinder, damit die Intelligenz sich möglichst gut entwickeln kann. Was 'ne Manipulation, eh!“).

Dieser kleine Exkurs zurück in die Empirie sollte die vorgefundenen Effekte von Impliziten Intelligenztheorien plausibilisieren. Einerseits ist die Erkenntnis sicher nicht neu, dass die Vorstellung einer unveränderbaren Intelligenz bei Heranwachsenden Unbehagen auslöst. Andererseits deutet empirisch eben vieles darauf hin, dass man dieses Unbehagen nicht als abstrakte „altlinke“ Denkgewohnheit abtun, sondern seine realen Auswirkungen im Schul- und Studienalltag ernstnehmen sollte.

9.2 Implizite Intelligenztheorien und ihre Wirkungsweise im Lernprozess

Welche Prozesse laufen zwischen der Partizipation an dem statischen bzw. dynamischen Leistungsdiskurs und dem beobachteten dysfunktionalen Bewältigungsverhalten ab? Der vorige Abschnitt hat lediglich anhand beispielhafter Zitate deutlich zu machen versucht, weshalb die Botschaft des statischen Artikels Unbehagen bei seinen Rezipienten auslöst. Wie sieht es aber mit den konkreten psychischen Prozessen aus, die in den empirischen Untersuchungen erfasst worden sind und von denen erwartet wurde, dass sie für die Wirkungsweise der Intelligenztheorien eine Rolle spielen?

Der wichtigste dieser Prozesse ist zweifelsohne die emotionale Reaktion auf eine Leistungsrückmeldung. *Emotionen* wurden in Dwecks Modell weitgehend vernachlässigt und gehören daher zu den wichtigsten Innovationen der vorliegenden Arbeit (vgl. Abschnitt 3.3). In den empirischen Daten zeichnet sich die erwartete Tendenz ab, dass die Induktion der statischen Intelligenztheorie zu verstärktem Erleben von Hoffnungslosigkeit und Scham führt, die Induktion der dynamischen Intelligenztheorie dagegen mit Zuversicht und Schuld verbunden ist. Für Schuld muss allerdings hinzugefügt werden, dass deren geringe interne Konsistenz die Interpretation erschwert; die unüblich weite Definition von Schuld, die beispielsweise auch Selbstärger miteinschließt, hat sich in dieser Hinsicht empirisch nicht bewährt.

Zudem konnte etwas Licht in die Auswirkungen jener Emotionen auf das Bewältigungsverhalten gebracht werden. Teilweise scheinen die erwarteten Zusammenhänge zwischen Hoffnungslosigkeit, Scham und hilflosem Bewältigungsverhalten sowie zwischen Zuversicht, Schuld und meisterndem Bewältigungsverhalten vorzuliegen. Eine Reihe von Ergebnissen spricht aber auch dafür, dass die Dinge etwas komplizierter liegen, d.h. Emotionen durchaus ambivalente motivationale Auswirkungen haben können. Auch die spiegelbildliche Deutung, wonach die motivationalen Konsequenzen Impliziter Theorien in ambivalenter Form mit den Emotionen zusammenhängen, erscheint – insbesondere im Hinblick auf die Ergebnisse von Studie 3 – nicht völlig abwegig. Die zeitliche Abfolge der ablaufenden Prozesse ist aber wohl auch gar nicht von größter praktischer Bedeutung, ja z.T. vielleicht prinzipiell nicht empirisch zu entscheiden. Dwecks Rede von „meaning systems“, die durch Implizite Theorien aufgespannt werden, legt ja gerade nahe, dass emotionale, motivationale und behaviorale Prozesse sich wechselseitig bedingen und verstärken.

Die Untersuchung von Baer, Grant und Dweck (2005; vgl. Abschnitt 3.3.4) liefert empirische Hinweise, die der Vorstellung eines solchen komplexen Zusammenhangsgefüges gut entsprechen. Depressiver Affekt hat hier für Statiker und Dynamiker eine vollkommen unterschiedliche Bedeutung: „in the context of one lay theory, depressed affect appeared to signal defeat, whereas in the context

of another it appeared to serve as a call for renewed action” (Molden & Dweck, 2006, S. 196). In diesem Kontext erscheint auch die Frage interessant, inwieweit Implizite Intelligenztheorien mit impliziten *Emotionstheorien* einhergehen. So ließe sich beispielsweise mutmaßen, dass Statiker negative Emotionen eher als unhinterfragbare und unabwendbare Signale des eigenen Körpers verstehen, während Dynamiker sie mehr als Motivation verstehen, Maßnahmen zu ihrer Überwindung zu ergreifen. Die bisherigen Untersuchungen legen allerdings eher nahe, dass Implizite Theorien über verschiedene alltagspsychologische Gegenstände unabhängig voneinander sind, d.h. dass jemand, der Intelligenz als stabil ansieht, gleichzeitig durchaus Schüchternheit als veränderbar betrachten kann und umgekehrt (vgl. Dweck, 1999). Nichtsdestotrotz bietet die Frage, wie Emotionen im Leistungskontext verstanden werden, ob sie etwa eher als kontrollierbar oder als unkontrollierbar angesehen werden (vgl. Weber, 2004; Weber & Titzmann, 2003), ein interessantes Feld für zukünftige Untersuchungen.

Attributionen spielen – in ähnlicher Form wie Emotionen – in der Beziehung zwischen Intelligenztheorie und Bewältigung eine vermittelnde Rolle. Dieses Parallelgehen von Emotion und Attribution ist einerseits nicht erstaunlich, da die Emotions-Items oft unter direktem Bezug zu ihrer attributionstheoretischen Definition formuliert wurden (vgl. Abschnitt 4.3). Angesichts dessen erscheint es – andererseits – eher erstaunlich, dass die korrelativen Beziehungen zwischen Attribution und Emotion oft gering ausfielen. Dies zeigt aber, dass Emotionen trotz ihrer engen attributionstheoretischen Definition nicht identisch sind mit Attributionen, sondern immer noch eine zusätzliche Komponente aufweisen, die zuweilen als „Wert“ oder „Bedeutsamkeit“ des Emotionsgegenstands bezeichnet wird (vgl. Pekrun, 2000).

Welche Bedeutung die soziale Struktur der *Leistungssituation* für die von Impliziten Theorien angeregten differenziellen Lernverhaltensweisen hat, bleibt weitgehend unklar. Empirisch deutet sich an, dass in Lernformen mit kooperativen Zielstrukturen bei gleichzeitiger individueller Verantwortlichkeit (vgl. z.B. Slavin, 1993) einerseits selbstrelevante Emotionen wie Scham und Schuld vermehrt auftreten, andererseits aber die Motivation zur Bewältigung von Misserfolgen ausgeprägter ist. Vor diesem Hintergrund wäre bei der Implementierung dieser Lernformen also künftig abzuwägen, inwieweit die Vorteile solcher Maßnahmen jene emotionalen Kosten rechtfertigen. Die hier gefundenen Ergebnisse lassen aufgrund ihrer fehlenden Klarheit und Konsistenz aber wie gesagt keineswegs so weitreichende Schlussfolgerungen zu.

Auch die Bedeutung des Faktors *Geschlecht* konnte nicht endgültig geklärt werden. Studie 2 gibt starke Hinweise darauf, dass das Geschlecht der Probanden als Moderator wirkt, d.h. dass Frauen sich von den Texten über Intelligenz deutlich mehr beeinflussen lassen als Männer. Studie 3 konnte dies indes nicht bestätigen. Da in beiden Untersuchungen nur äußerst wenige Männer untersucht

werden konnten, sind klare Schlussfolgerungen nicht möglich. Die bislang vorliegenden Untersuchungen zu dieser Frage (vgl. Abschnitt 3.2.2) bieten aber Anlass genug, die Frage nach geschlechtsspezifischen Auswirkungen Impliziter Theorien in künftigen Forschungsbemühungen stärker in den Mittelpunkt zu rücken.

Selbstverständlich sind außer den genannten auch noch andere Prozesse vorstellbar, die bei der Wirkung Impliziter Intelligenztheorien von Bedeutung sein könnten. Verwiesen sei hier etwa auf die Analysen inhaltlich-qualitativer Aspekte der Induktionssituation in Studie 2 (Abschnitt 6.2.2), die aus ökonomischen Gründen leider nur in rudimentärer Form vorgenommen werden konnten, jedoch prinzipiell erfolgversprechend erscheinen (vgl. auch Abschnitt 9.4). Seit neuestem wird zur Erklärung der Wirkungsweise Impliziter Theorien sogar der neurowissenschaftliche Zugang gewählt. So unternehmen Mangels, Butterfield, Lamb, Good und Dweck (2006) den Versuch, die neurophysiologischen Prozesse zu erkunden, hinsichtlich derer sich statisch und dynamisch Disponierte bei der Leistungsrückmeldung unterscheiden. Da dieser Ansatz jedoch noch nicht hinreichend ausgereift ist, habe ich nicht nur auf eine Erhebung neurophysiologischer Merkmale, sondern auch auf die theoretische Berücksichtigung der Studie von Mangels et al. verzichtet. Der neurowissenschaftliche Denkansatz wurde bekanntlich lediglich als Forschungsgegenstand, nicht als Forschungsmethodologie verwendet, wie im folgenden Abschnitt noch einmal deutlich werden wird.

9.3 Von der Neurodidaktik zur Lernpsychologie mit Plastizitätsargumenten

Eine wichtige Implikation meiner empirischen Untersuchungen betrifft die in Kapitel 2 aufgeworfene Frage nach dem Einfluss des Hirnforschungsdiskurses auf das Lern- und Leistungsverhalten seiner Teilnehmer. Wenngleich meine Studien auf diese Frage sicherlich keine abschließende Antwort geben können, sollen im Folgenden dennoch einige Überlegungen darüber angestellt werden, welche Schlussfolgerungen meine Ergebnisse für Hirnforscher, Journalisten und Pädagogen zulassen.

Ein Vergleich der Tabellen 2 (Abschnitt 2.2.3) und 3 (Abschnitt 4.2.1) macht deutlich, dass die Inhalte der Texte zur Induktion Impliziter Intelligenztheorien durchaus als repräsentativ gelten können für die mediale Berichterstattung zu diesem Thema. Lediglich hinsichtlich der Zuspitzung und Pointierung seiner Botschaft könnte der erstellte Text als untypisch angesehen werden. Die Schlussfolgerung vom hier verwendeten Induktionstext auf zeitgenössische Beiträge in deutschen Printmedien erscheint demzufolge zulässig.

Die Studien 2 und 3 haben gezeigt, dass Berichte über neurowissenschaftliche Befunde, wonach Intelligenz stabil und nur ein Epiphänomen des Gehirns ist, zu abträglichen Folgen im Lernverhalten führen können, während Berichte,

die Intelligenz als beeinflussbar und im Gehirn repräsentiert ansehen, positive Verhaltensreaktionen nach sich ziehen können. Eine solche Interpretation der Ergebnisse ist jedoch nur unter zwei Einschränkungen möglich. Erstens muss berücksichtigt werden, dass ein direkter Vergleich mit dem „psychologischen“ Leistungsdiskurs aus versuchsökonomischen Gründen nicht vorgenommen wurde. Klar ist also nur, dass der neurowissenschaftliche Leistungsdiskurs *auch* wirksam ist, nicht aber, ob er etwa noch besser wirksam ist als etwa der klassisch-psychologische, wie er z.B. bei Hong et al. (1999) verwendet wurde. Diese Einschränkung spielt aber für die praktischen Konsequenzen der Befunde, um die es im Folgenden gehen soll, eine untergeordnete Rolle.

Zweitens bleibt in Ermangelung einer Kontrollgruppe empirisch gesehen unklar, ob für den Unterschied zwischen beiden Gruppen eher die eine oder die andere Bedingung verantwortlich ist. Theoretische Überlegungen lassen dieses Problem jedoch als weniger schwerwiegend erscheinen. Gemäß Abschnitt 3.2.0 ist die Implizite Intelligenztheorie einer Person als Kontinuum zu verstehen und kann durch äußere Einflüsse in eine Richtung verändert werden. Die Gruppenunterschiede ergeben sich also nicht daraus, dass die Probanden in der Statiker-Gruppe „statisch“ geworden sind und in der Dynamiker-Gruppe „dynamisch“, sondern dass sie jeweils etwas statischer bzw. dynamischer geworden sind, als sie vorher waren. Aus theoretischer Perspektive erscheint somit plausibel (wenn auch sicherlich nicht logisch zwingend), dass *beide* Bedingungen zu etwa gleichen Teilen zu dem Effekt beigetragen haben. Diese Deutung erfährt im Übrigen durch die Erfahrung aus Dar-Nimrod und Heines (2006) Kontrollgruppenstudie empirische Unterstützung.

Demzufolge können zwei Implikationen der Befunde unterschieden werden: zum einen solche in Abhängigkeit der pädagogisch negativen Auswirkungen des stabil-epiphänomenalistischen Diskurses, zum anderen solche aufgrund der positiven Auswirkungen des veränderbar-geisteswissenschaftlichen Diskurses. Um diese beiden Implikationen soll es in den folgenden zwei Abschnitten gehen.

9.3.1 Die Gefahren eines spekulativ vertretenen Neuro-Determinismus

Die Studien 2 und 3 haben gezeigt, dass sich die – hier über Zeitschriftenartikel induzierte – Vorstellung einer stabilen Intelligenz negativ im Leistungsverhalten der Betroffenen manifestieren kann. Obwohl der Zeitschriftenartikel eine eher plumpe und leicht zu hinterfragende Spielart neurowissenschaftlichen Denkens repräsentiert (frei nach dem Motto: „die Größe des Intelligenzbereichs im Gehirn entspricht der Intelligenz eines Menschen“), scheinen die allermeisten studentischen Probanden dem Artikel Glauben geschenkt zu haben. Dass die massenhafte Verbreitung populärwissenschaftlicher Deutungen der Hirnforschung solcherlei Gefahren in sich trägt, wird erst in jüngster Zeit erkannt.

So genau weiß tatsächlich niemand, ob die Tiere bei dem Schockversuch dasselbe intellektuelle Vergnügen empfinden wie die Forscher. Aber in Zeitungsinterviews verkündet Scheich, dass erfolgreiches Lernen genau über solche Belohnungen funktioniert – natürlich auch bei Schülern. Deshalb sei der Unterricht bitte entsprechend umzukrempeln. Würde eine Pharmafirma aufgrund eines Tierversuchs und einer vagen Theorie ein Medikament auf den Markt bringen, käme sofort der Staatsanwalt. In der Bildungsdiskussion, wo jeder ungeprüfte Rezepte propagieren darf, bringt man es so schnell zum gefragten Experten. Scheich weiß aber noch mehr. „Auf keinen Fall ist es so, dass man Menschen nur mit Lob und Spaßlebnissen dazu bringt, dauerhaft etwas zu lernen“, verlautbarte er in der *Apotheken Umschau*. (Wenn sie ihre großen Ideen verkünden, dann meiden Neurodidaktiker gern Fachzeitschriften samt deren kritischen Gutachtern.) Denn ohne gelegentliche Herausforderungen in Form von Schocks lässt bei den Mäusen die Dopaminproduktion nach. Leichter Stress fördert das Lernen. So spricht die Neurobiologie (Paulus, 2003, S. 35).

In der Diskussion über Neurodidaktik mehren sich kritische Stimmen, die den medialen Geltungsdrang einiger Hirnforscher infragestellen. Gerade mit Bezug auf die Situation in den USA wird auch bereits erkannt, dass das leichtfertige Spiel der Hirnforscher mit den Medien ethisch höchst fragwürdige Konsequenzen nach sich zieht. Eberle (2002) porträtiert die US-Bestsellerautorin Lise Eliot, die mit ihrem Buch den amerikanischen Neuro-Erziehungsboom maßgeblich angefacht haben dürfte. Nach dem Interview mit Eliot resümiert Eberle:

Diese Erkenntnis – vorgetragen auch von anderen Neurologen – katapultierte Amerikas Kleinkinder in einen Strudel elterlicher Ängste und Eitelkeiten. Die Regierung von Florida befahl Kindergärten, klassische Musik zu spielen, nachdem Forscher erkannt haben wollen, dass dies die Intelligenz steigere. Übereifrige Schwangere bedudeln schon ihren Bauch mit Mozartsonaten – die Weisen des Wunderkinds sollen besonders förderlich wirken. (...) Mal heißt es, durch die richtige Strategie ließe sich der Intelligenzquotient um 20 Punkte verbessern, mal nur um 12. Doch selbst 12 IQ-Punkte, rechnet Eliot vor, machten den Unterschied aus zwischen schnödem Mittelmaß und dem oberen Fünftel der Schlauköpfchen. (...)

Kritik daran tut weh. Hämisches warf ihr eine angesehene britische Zeitung vor, auf dem Intelligenzfimmel mitzureiten, bei dem Gehirnwindungen gezüchtet werden sollen wie Hollandtomaten im Treibhaus. Eliot knetet sich gequält die Brauen bei der Erinnerung. Heute würde sie lauter davor warnen, dass zu viel Stimulation auch schaden kann, würde sagen, dass Kinder „ausbrennen“, wenn sie schon als Fünfjährige mit den Spielzeugen für Ältere spielen müssen, „nur weil die Eltern denken, dann sind sie gleich so schlau wie Achtjährige“. Aber wer konnte denn 1999, als sie das Buch schrieb, ahnen, wie hysterisch manche Eltern reagieren würden? (Eberle, 2002, S. 36).

Doch auch in Deutschland wird bereits von dieser übertriebenen „Angst, die wichtigste Zeit in der Gehirnentwicklung ihres Kindes ungenutzt verstreichen zu lassen“ (Otto, 2007, S. 73), berichtet. Da wirkt es dann grotesk, wenn der prominente Neurodidaktiker Henning Scheich die im Frühförderungsbereich zu beobachtenden Auswüchse von „Scharlatanerie“ und „Pseudowissenschaft“ (zit.n. Otto, 2007, S. 73) kritisiert, die er mit seinem aggressiven Drängen in die Öffentlichkeit selbst mit herbeigeführt hat.

Sogar Neurowissenschaftler von internationalem Ansehen reagieren, angesprochen auf die ethischen Konsequenzen ihres öffentlichen Auftretens, nicht selten mit erstaunlicher Naivität. Eric Kandel, Nobelpreisträger und Entwickler gedächtnissteigernder Psychopharmaka, wirkt bei Fragen auf die praktischen Konsequenzen solcher „Gedächtnispillen“ ahnungs- bis hilflos:

FOCUS: Nehmen wir als Beispiel die Schule. Wie soll man Testergebnisse vergleichen, wenn ein Teil der Schüler mit Pillen gebüffelt hat, der andere aber nur Fleiß (...)?

Kandel: Sie springen da an einen Punkt, an den ich nicht will. Ich habe dazu eine klare Haltung: Man darf keine Prüfung bestehen, wenn man Pillen nimmt. Das sollten Gesetze ganz klar regeln. Genauso wenig darf an den Olympischen Spielen teilnehmen, wer Steroide schluckt.

FOCUS: Wie lässt sich überprüfen, ob jemand Lern-Doping betrieben hat?

Kandel: Wieso sollte man Menschen nicht einfach trauen?

FOCUS: So, wie man den Athleten bei Wettkämpfen traut?

Kandel: Wir sollten Vorschriften aufstellen, an die sich alle zu halten haben. Herrgott, ich weiß es doch auch nicht! Es ist ein Problem, das streite ich gar nicht ab.

FOCUS: Hinzu kommt, dass die Pillen mit Sicherheit teuer sein werden.

Kandel: Richtig. Das bedeutet, nur die Wohlhabenden werden sich die chemische Nachhilfe leisten können. Das macht die Situation schlimmer, und ich weiß nicht, was man dagegen tun könnte. (...)

[Später im Gespräch kommt Kandel von sich aus auf das Thema zurück:] Ich muss gestehen, mir will die Sache mit den Schülern nicht aus dem Kopf gehen. Wenn die das Gefühl haben, Pillen könnten ihnen bei Prüfungen helfen, dann werden sie diese garantiert nehmen. Denn die Intelligenz und damit auch das Gedächtnis beeinflussen sehr stark, welche Karriere jemand einschlagen kann. (...)

FOCUS: Ich bin überrascht, dass Sie sich mit den ethischen Aspekten Ihrer Forschung bisher nicht mehr beschäftigt haben.

Kandel: Vielleicht ist das eine Schwäche von mir. Ich dachte immer, wir wüssten, wie jungen Menschen am besten zu helfen sei: Man muss ihnen die Gelegenheit geben, gut lernen zu können. Eine Gedächtnis-Pille ist dafür sicher nicht die Lösung. Aber mit dieser Haltung habe ich es mir wohl etwas zu leicht gemacht. Ich bin zwar eine ethisch denkende Person, aber kein Spezialist für Medizinethik. Ich bin Naturwissenschaftler (*Focus* 1/2005, S. 74).

Allem Anschein nach haben wir es bei den Hirnforschern, die sich in der Öffentlichkeit äußern, nicht selten mit Menschen zu tun, die über erhebliche Deutungsmacht verfügen, sich aber dieser Macht und ihrer Konsequenzen kaum bewusst sind. Prinz (2008) hat sicher recht, wenn er jene Macht auch auf die entsprechende Nachfrage eines „wissenschaftsgläubigen Publikum(s)“ zurückführt, „das sich schon lange danach sehnt, die verwickelten Angelegenheiten des Geistes und der Kultur endlich mit der gleichen Klarheit und Verbindlichkeit aufgeklärt und entzaubert zu sehen, die für die Angelegenheiten der belebten und unbelebten Natur schon lange gelten“ (jeweils S. 48). Wie in Kapitel 2 ausführlich dargelegt wurde, sind es aber auch und vor allem die sich öffentlich äußernden Hirnforscher selbst, die für Auswüchse wie die oben zitierte Scharlatanerie und Hysterie verantwortlich zu machen sind.

Verständlich ist zwar, dass sich die Hirnforschung als noch junge Disziplin mit provokanten Thesen Gehör zu verschaffen versucht. Fragwürdig aber wird diese Praxis, sobald sich herausstellt, dass diese Thesen ethisch bedenkliche Konsequenzen nach sich ziehen. Dass dies im vorliegenden Fall beim stabil-epiphänomenalistischen Diskurs gegeben sein könnte, habe ich bereits in Abschnitt 2.3 gezeigt, indem ich auf die Umstrittenheit und Spekulativität der vorgebrachten Thesen hingewiesen habe. Die schwierige Position der Hirnforschung als Mittlerin zwischen Naturwissenschaft, Psychologie und Philosophie bringt es mit sich, dass gesicherte und gleichzeitig alltagsrelevante Aussagen (noch?) äußerst selten sind.

Welche Konsequenzen sind aber zu ziehen? Der Wissenschaftssoziologe Weingart mahnt im Interview mit der Zeitschrift „Gehirn und Geist“ (1/2005) zu einem „verantwortlichen Handeln“ der Wissenschaftler und Journalisten, das sich „über die möglichen Folgen ihrer Arbeit im Klaren“ sein müsse, seinerseits allerdings auch „ein gewisses Verständnis der Gesellschaft, der Politik und der Medien“ (S. 75) voraussetze. Dem ist eigentlich wenig hinzuzufügen (vgl. auch Thompson & Nelson, 2001, sowie Abschnitt 2.2.4). Um nicht missverstanden zu werden: Plädiert wird hier keineswegs für ein Redeverbot für Forscher, deren Studien „politisch unkorrekte“ Schlussfolgerungen zulassen. Unsere Auffassung von Wissenschaft als Mittel der Aufklärung hat sich gut bewährt, d.h. auch unbequeme Wahrheiten müssen ausgesprochen werden dürfen. In dem Maße aber, in dem die Datenlage unklar und die Deutung von Untersuchungen höchst umstritten und spekulativ ist, kann es manchmal tatsächlich besser sein zu schweigen. In ihren Artikeln in Fachzeitschriften hüten sich Wissenschaftler in der Regel äußerst gewissenhaft vor voreiligen Schlussfolgerungen aus ihren Ergebnissen. Die vorliegende Untersuchung kann ihnen Anregungen geben darüber nachzudenken, dass es gute Gründe gibt, sich vor diesen voreiligen Schlüssen auch – oder besser: ganz besonders – im Umgang mit der Öffentlichkeit zu hüten.

Auch an Journalisten und Politiker muss appelliert werden. Sie sind es, die die Wissenschaftler dazu motivieren (Politiker) und ggf. dabei unterstützen (Journalisten), die Öffentlichkeit mit überinterpretierten Forschungsberichten zu täuschen. Die Medienwissenschaftlerin Senja Post (2008) zeigte dies kürzlich am Beispiel der Klimaforschung. Diejenigen Klimaforscher, die besonders dramatische Zukunftsszenarien entwerfen, erhalten nicht nur besonders viel Aufmerksamkeit durch die Medien, sondern auch besonders viele Forschungsgelder. Die zahlenmäßig nahezu gleichgroße Fraktion der sogenannten „vorsichtigen Beobachter“ hingegen geht hier wie dort leer aus. In der Öffentlichkeit entsteht damit ein stark verzerrtes Bild der Klimaforschung – das sich in der Zukunft durch die selektive finanzielle Förderung immer weiter verzerren wird. Es besteht wohl kein Grund zu der Annahme, dass die Dinge bei der Hirnforschung wesentlich anders liegen.

9.3.2 *Das Potenzial der neuronalen Plastizität als Metapher des Lernens*

Selbst wenn sich Hirnforscher und Medienvertreter eines Besseren besinnen würden und neurowissenschaftliche Ergebnisse künftig mit der gebotenen Vorsicht und kritischen Distanz behandeln würden, stünde dennoch zu befürchten, dass die Vorstellung einer stabilen Intelligenz nicht aus der Welt wäre, sondern weiterhin – sei es in Familie, Peer-Gruppe oder Schule bzw. Universität – das Selbstbild von Schülern und Studierenden nachhaltig prägen dürfte. Daher stellt sich die Frage, wie sich der derzeit so populäre Hirnforschungsdiskurs in offensiver Form pädagogisch einsetzen lässt.

Die Studien 2 und 3 lassen den Schluss zu, dass die neurowissenschaftliche Deutung des Gehirns als plastisch und mithin der Intelligenz als veränderbar einen positiven Einfluss auf Lernen und Leistung ausüben kann. Demnach scheint es zunächst einmal geboten, in Berichten über Hirnforschung immer wenigstens diese Plastizitätskomponente mitzubetonen, um die dysfunktionalen Konsequenzen der Vorstellung eines frühkindlich determinierten Gehirns zu vermeiden. Beispiele für eine solch ausgewogene Darstellungsweise sind übrigens keineswegs selten; verwiesen sei etwa auf die in Tabelle 1 (Abschnitt 2.2.3) aufgeführten Zeitschriftenartikel, die weder einer statischen noch einer dynamischen Auffassung zugeordnet werden konnten. Dass dennoch häufig eine einseitige Betrachtungsweise vorgezogen wird, dürfte auf wissenschaftsexterne Gründe zurückzuführen sein (vgl. Kapitel 2). (Übrigens sei an dieser Stelle noch einmal ausdrücklich betont, dass die Probanden der hier durchgeführten Untersuchungen ausführlich darüber aufgeklärt wurden, dass die von ihnen gelesenen Artikel keine realitätsangemessene Sichtweise von Intelligenz vermitteln. Ethische Probleme konnten so weitestmöglich vermieden werden.)

Darüber hinaus ist sicher auch bedenkenswert, in Interventionen an Schulen die neuronale Plastizität als Metapher für den Lernfortschritt gezielt einzusetzen. Die in Abschnitt 2.2.2 vorgestellten Untersuchungen (z.B. Blackwell, Dweck, & Trzesniewski, 2007 und das Projekt „Brainology“) haben gezeigt, dass derartige Programme bei geringem ökonomischem Aufwand beeindruckende Auswirkungen haben. Die Plastizitätsmetapher mit all ihren Visualisierungsmöglichkeiten scheint den Schülern auf höchst einprägsame Weise zu verdeutlichen, dass Lernen „was bringt“ und Anstrengung sich lohnt, dass alles, mit dem man sich beschäftigt, „das Gehirn verändert“ und somit nicht vertane Zeit ist, dass Intelligenz nicht nur die Voraussetzung für Lernen ist, sondern immer auch dessen Ergebnis.

Zwar haben Untersuchungen wie die von Hong et al. (1999) gezeigt, dass sich auch mit nicht-neurowissenschaftlichen Deutungsmustern ähnliche Effekte erzielen lassen. Dass jedoch die Metapher der neuronalen Plastizität aufgrund ihrer Anschaulichkeit, intuitiven Plausibilität und – nicht zuletzt – auch wegen der

momentanen Popularität der Hirnforschung in der Öffentlichkeit diesbezüglich besonders geeignet ist, kann wohl kaum in Zweifel gezogen werden.

Die Botschaft der neuronalen Plastizität könnte den Schülern dabei nicht nur in Form von eigens abgehaltenen Workshops vermittelt werden, wie dies bei Blackwell et al. sowie bei den computergestützten Trainings des Brainology-Programms (Dweck, 2007a) der Fall war, sondern auch durch Kommentare des Lehrers im regulären Unterricht. Wenn es einzelne Schüler oder sogar die ganze Klasse an Motivation und Anstrengungsbereitschaft fehlen lassen, könnte der Lehrer etwa die Plastizitätsmetapher durch folgende Kommentare umsetzen:

- „Das Gehirn speichert alles, was wir lernen. Es ist daher auch wichtig, sich bei langweiligen und scheinbar nicht so wichtigen Themen anzustrengen.“
- „Das hat doch jetzt sogar die Hirnforschung gezeigt: Anstrengung lohnt sich! Das Gehirn wächst immer weiter. Da frag ich mich wirklich, warum ihr hier so unmotiviert rumsitzt.“
- „Das könnt ihr doch jeden Tag in der Zeitung lesen: Das Gehirn ist keine Maschine, mit der man zur Welt kommt. Es entwickelt sich ständig weiter, je nachdem, was man lernt. Entscheidend ist, dass ihr diese Entwicklung selbst steuern könnt.“
- „Kennt ihr nicht diese Darstellungen des Gehirns im Fernsehen? Die Nervenzellen gehen ständig neue Verbindungen ein und entwickeln sich ständig weiter. Ihr habt diese Veränderungen in der Hand. Intelligenz ist nicht die Voraussetzung des Lernens, sondern das Ergebnis des Lernens.“
- „Ihr lest doch jeden Tag in der Zeitung: Die Hirnforschung hat gezeigt, dass man nicht mit irgendeiner Intelligenz auf die Welt kommt, mit der man dann leben muss. Man bestimmt seine Intelligenz selbst, durch das was man lernt. Das ist doch wunderbar. Allerdings kann man sich dann auch nicht mehr hinstellen und sagen: ‚Ich kann das nicht, ich bin halt zu blöd dazu.‘“

Sicherlich sollten solche Kommentare mit Bedacht eingesetzt werden. Insbesondere das letzte Beispiel macht deutlich, dass der Plastizitätsdiskurs immer auch eine individualisierende Komponente hat (vgl. Abschnitt 2.2.1), die mit negativen Gefühlen und Selbstzweifeln verbunden sein kann. Der Lehrer muss im Einzelfall entscheiden, ob einer der oben vorgeschlagenen Kommentare sinnvoll und situational angebracht ist oder nicht. Auch mag es Lehrer geben, die Schwierigkeiten haben, Kommentare wie diese glaubwürdig vorzubringen. Ihr pädagogischer Nutzen sollte aber durch die vorgestellten Ergebnisse deutlich geworden sein.

9.4 Implizite Theorie als Brückenkonzept zur Kulturpsychologie

Ein weiterer Einwand gegen die im vorigen Abschnitt vorgeschlagenen Lehrerkommentare könnte sein, dass den Schülerinnen und Schülern so eine unrealistische Sicht auf die Beziehung zwischen Geist und Gehirn nahegebracht wird oder dass das Vertreten einer Geist-Gehirn-Dichotomie (hier konkretisiert als „Geist steuert Gehirn“) wissenschaftlich fragwürdig ist. Wie mittlerweile klar geworden sein sollte, positioniert sich die vorliegende Arbeit in dieser „neurophilosophischen“ Frage nicht. Sie gibt lediglich Hinweise darauf, wie sich ohnehin bestehende Diskurse über Lernen und Leistung für pädagogisch sinnvolle Zwecke einsetzen lassen. Dem liegt die Einschätzung zugrunde, dass Intelligenz und Lernanstrengung, Gehirn und Geist wichtige alltagspsychologische Konzepte sind, die sich nicht von heute auf morgen abtrainieren lassen. Eltern, Lehrer und Dozenten sollten sie lieber aufgreifen als ignorieren, da sie ein – derzeit zumindest – unverzichtbarer Teil unserer Kultur sind. Nicht zuletzt um dieses Argument näher zu erläutern, soll in diesem Abschnitt eingeschätzt werden, wie die Beziehung zwischen Impliziter Intelligenztheorie und Kultur konzeptualisiert werden kann und inwiefern künftige Untersuchungen von einem kulturpsychologisch erweiterten Konzept Impliziter Theorien profitieren könnten. Wenn gleich auf diese Thematik bereits an mehreren Stellen dieser Arbeit Bezug genommen wurde, bietet es sich an, diese verstreuten Hinweise nun im Zusammenhang zu betrachten und ihre theoretischen Grundlagen zu explizieren.

In Abschnitt 4.1.1 wurde der Einfluss kultureller Faktoren auf Lernen und Leistung mit dem Begriff der Leistungskultur zu fassen versucht. Mit diesem Fokus auf „weiche“, soziale Komponenten des Lernprozesses lässt sich die vorliegende Arbeit als Beitrag zum konstruktivistischen bzw. kulturalistischen Ansatz in der Pädagogischen Psychologie verstehen, wie er etwa von Lave und Wenger (1991) vertreten und von Reinmann-Rothmeier und Mandl (2001) wie folgt charakterisiert wird:

[1.] Das Wissen in einer Gesellschaft stellt immer „geteiltes Wissen“ dar, d.h., Wissen wird von Individuen im Rahmen sozialer Transaktionen gemeinsam entwickelt (...). [2.] Das konkrete Denken und Handeln eines Individuums lässt sich jeweils nur auf dem Hintergrund eines konkreten (sozialen) Kontexts verstehen. [3.] Lernen ist stets situiert, d.h., es ist an die inhaltlichen und sozialen Erfahrungen der Lernsituation gebunden. [4.] Wissen wird nicht passiv erworben, sondern aktiv konstruiert (Reinmann-Rothmeier & Mandl, 2001, S. 615).

Lernen erscheint unter dieser Perspektive zu allererst als sozialer, als interaktiver Prozess, als praktischer Austausch mit der Umwelt. Inneren Vorgängen, insbesondere abstrakten, deduktiv sich vollziehenden Denkvorgängen wird weniger Aufmerksamkeit geschenkt. Während sich die Arbeiten von Lave und anderen primär mit der Aneignung von Wissen in ganz bestimmten Lernumgebungen be-

schäftigt, untersucht die vorliegende Arbeit den Einfluss allgemeiner gesellschaftlicher Diskurse über Lernen und Leistung auf die Motivation und Anstrengungsbereitschaft der Diskursteilnehmer. Sie verfolgt damit einen kulturpsychologischen Ansatz, der sich zwar am Leistungsmotivationsmodell von Dweck inhaltlich orientiert, dessen zentrales Konzept der Impliziten Intelligenztheorie aber kulturalistisch umdeutet, indem er sie nicht mehr als stabiles Personmerkmal, sondern als situational beeinflussbaren Zustand versteht. Die Implizite Intelligenztheorie eines Menschen ist hiernach empirisch kaum zu trennen von seiner sozialen und kulturellen Umwelt. Die Begrifflichkeiten, die ihm seine Sprache zur Verfügung stellt, die Erwartungen, die seine soziale Umwelt an ihn heranträgt, und schließlich auch wissenschaftliche Deutungsangebote, die ihm nahelegen, wie seine Zukunft aussehen wird und ob sich seine Lernanstrengungen lohnen, – all dies steht in direkter Beziehung mit dem Konzept der Impliziten Intelligenztheorie. Dass aus all diesen Erfahrungen schließlich eine mehr oder weniger stabile Disposition zu einer bestimmten Intelligenztheorie erwächst, ist dabei nicht unplausibel. Dies jedoch kann nicht das Vorgehen der allermeisten Forscher rechtfertigen, sich mit der Erfassung dieser Disposition über einen schlichten Fragebogen zufriedenzugeben (vgl. Abschnitt 3.2). Man kann dann zwar Korrelationen zu allerlei Lern- und Leistungsmerkmalen berechnen, aber weiß am Ende doch nicht, wie es zu dem gefundenen Dispositionswert einer Person kommt, geschweige denn wie dieser zu verändern ist.

Auf welche Art von „Kulturpsychologie“ bezieht sich diese Reformulierung des Konzepts Impliziter Intelligenztheorien genau? Zunächst ist gemäß Straub (2003; Straub & Layes, 2002) die Kulturpsychologie von der Kulturvergleichenden Psychologie zu unterscheiden. Während letztere die Begrifflichkeiten der Mainstream-Psychologie weitgehend übernimmt und lediglich die Gültigkeit ihrer Theorien in verschiedenen Ländern empirisch zu überprüfen versucht, versteht sich erstere als alternativer psychologischer Denkansatz, der die individualistische, positivistische, empirizistische und inhaltsferne Ausrichtung des Mainstreams kritisiert und die Beziehung zwischen Individuum und Umwelt neu zu konzeptualisieren versucht. Beispielfhaft soll hier auf den Ansatz Carl Ratners eingegangen werden (vgl. aber auch Laucken, 2000; Straub, 1999). Wiewohl Ratner (2006) mit Bezug auf Vygotsky und Foucault den prägenden Einfluss gesellschaftlicher und sozialer Kräfte auf die individuelle Psyche betont, gesteht er ihr – anders als postmodern-sozialkonstruktivistische Ansätze, z.B. Gergen, 1994 – gleichzeitig ein Eigenleben zu. Dass innerpsychische Vorgänge „nur“ sozial konstruiert sind, impliziert eben keineswegs, dass sie nicht auch eine Wirkung entfalten und – unabhängig von den kulturellen Einflüssen, die für sie verantwortlich sind – Verhalten bestimmen. Ein paar programmatische Zitate, in denen Ratner sein zentrales Konzept des „macro cultural factor“ erläutert, wel-

ches er in Institutionen, Artefakte und kulturelle Konzepte untergliedert, sollen dies illustrieren.

Cultural factors are envisioned and maintained by subjective processes such as thinking, perceiving, emotions, and motives. However, the products that subjectivity produces reciprocally structure it. (...) Because macro cultural factors are the means by which humans survive and fulfill themselves, psychological phenomena must be devoted to constructing, maintaining and refining them. (...) These psychological phenomena must have a form and content that is congruent with the form and content of the macro factors they construct, maintain, and refine“ (Ratner, 2006, S. 13-14).

Psychological phenomena have the properties of a macro cultural factor. They are socially constructed and shared; they are artifacts rather than natural phenomena; they are emergent formations that transcend individual processes, although they certainly depend on individual processes (ebd., S. 116).

Psychological phenomena are distinctive and need to be investigated in their own right. They cannot be presumed, or read off from, macro cultural factors. Macro cultural psychology is not sociology. It does not diminish or abandon psychology by reducing it to macro factors. An emotion is not a social institution. Although emotions have a form and content that derives from social institutions, emotions are distinct (ebd., S. 23).

So ist das Gefühl der Angst zunächst einmal ein privates, individuelles Phänomen und kann auch nur als solches empirisch angegangen werden. Untrennbar mit dem Angstgefühl verbunden ist aber der kulturelle Stellenwert von Angst, der „Angstdiskurs“, der angibt, wann und wovon man Angst haben kann, darf und soll. Weder macht es Sinn, sich in der Art der Mainstream-Psychologie auf das Angstgefühl zu beschränken und alles andere als „äußere Einflussfaktoren der Angst“ abzutun, noch ist es ratsam, sich nach Art qualitativ-konstruktivistischer Forschungsprogramme damit zufriedenzugeben, Ratgeberliteratur, Alltagsgespräche oder Fernsehwerbung zum Thema Angst zu untersuchen. Beides muss zusammengedacht werden: Wie reagieren Menschen in der konkreten Situation auf die Lektüre des Angststratgebers? Wie hängen konkrete kulturelle Darstellungsformen der Angst mit individuellem Angsterleben zusammen? Dass Ratners Ansatz diesen Fragen gewachsen ist, dürfte überdies dadurch wahrscheinlicher werden, dass er „does not necessarily reject mainstream research“, sondern „reinterprets data and conclusions produced by mainstream psychologists“ (jeweils Ratner, 2006, S. 20). Was hier am Beispiel der Angst verdeutlicht wurde, lässt sich übrigens problemlos auf die Auffassung über die eigene Intelligenz (als kulturelles Konzept, s.o.) übertragen.

Neuere Ansätze wie der von Ratner betonen die Bedeutung von Sprache und Diskurs und sind so von gemäßigten Spielarten des Sozialkonstruktivismus bzw. Sozialen Konstruktionismus (Burr, 2003; Davies & Harré, 1990, vgl. hierzu Abschnitt 3.3.5; Nightingale & Cromby, 1999; Parker, 2002) kaum mehr unterscheidbar (vgl. von Tiling, 2008a, 2008b; Zielke, 2004). Wenn hier von kulturpsychologischen Ansätzen die Rede ist, sind damit also auch jene Strömungen

des Sozialkonstruktivismus mitgemeint, die sich vom radikalen Relativismus und Anti-Empirismus Gergens (1985; 2002) distanzieren.

Obwohl bei Ratner erste Bemühungen festzustellen sind, sich dem Dialog mit dem Mainstream-Ansatz in der Psychologie zu stellen, gibt es weiterhin äußerst selten Versuche, aus dessen überaus reichhaltigem Reservoir an empirischen Befunden kulturpsychologisch Interessantes aufzugreifen und entsprechend umzudeuten. Kulturpsychologische Ansätze werten gewöhnlich den Mainstream-Ansatz einseitig ab oder geben sich damit zufrieden, dessen zentrale Begriffe auf einer abstrakten Ebene umzudeuten, so dass beispielsweise (innere) Kognition als (äußeres) Sprachspiel „entlarvt“ wird (vgl. Edwards, 1997). Unabhängig von ihrer Qualität dürften solche theoretischen Abhandlungen für den Dialog zwischen Kulturpsychologie und Mainstream kaum produktiver sein als die bevorzugte Strategie des Mainstreams, die kulturpsychologischen Ansätze schlichtweg zu ignorieren. Immerhin werden diese oftmals unnötig tiefen Gräben zwischen den beiden Forschungsrichtungen in den letzten Jahren vermehrt kritisiert.

[S]ocial constructionists have generally emphasized the significance of different contents of lay beliefs or narratives, whereas experimental social psychologists have typically focused on the processes whereby people perceive and think about others, engage in social interaction, form and dissolve groups, and carry out intergroup relations. (...) Within cultural psychology, we have seen that constructionist preoccupations with the contents of shared representations have influenced mainstream social psychology (e.g., Markus & Kitayama [...]), where, appropriately enough, they have been approached via empirical (and often experimental) research. Along these lines, there has been intriguing work on the specific structures of various lay concepts, such as child development, anger, intelligence, understanding, ideals, and oughts. Social developmentalists (...), social representationalists (...), and, more recently, social cognitivists (e.g., Dweck [...]) have focused increasingly on the contents of mental representation (Jost & Kruglanski, 2002, S. 180-181).

Die Ursache für das vermehrte Interesse des Mainstreams an kulturalistischen Deutungen psychischer Prozesse liegt vermutlich in dem durchschlagenden Erfolg der Kulturvergleichenden Psychologie. Angeblich universelle psychologische Theorien haben sich in den letzten Jahren als kulturabhängiger herausgestellt, als es sich die allermeisten wissenschaftlich arbeitenden Psychologen jemals hätten vorstellen können (Heine & Norenzayan, 2006). So ist es mit Steven Heine (2008) auch ein Kulturvergleichender Psychologe, der bislang am nachdrücklichsten darauf hingewiesen hat, dass das Vorfinden interkultureller Unterschiede wichtige Implikationen auch für intrakulturelle Fragestellungen hat. Heine und Norenzayan (2006) schwebt eine experimentell ausgerichtete Kulturpsychologie vor, deren zentrale Methode das Priming kultureller Konstrukte ist (welches bekanntlich auch in der vorliegenden Arbeit Verwendung fand). Prototypisch für diese aus der Kulturvergleichenden Psychologie hervorgegangene Form der Kulturpsychologie ist die Studie von Heine et al. (2001, Studie 3), in der durch „kultur-entgegengesetztes“ Priming amerikanische Probanden stärker

das anstrengungsorientierte Leistungsverhalten von Asiaten annehmen und asiatische Probanden stärker das fähigkeitsorientierte Verhalten von Amerikanern. Studien dieser Art machen deutlich, dass Leistungsverhaltensweisen weder in universell gültigen noch in intrakulturell gültigen Mustern auftreten, sondern entscheidend vom unmittelbaren situationalen Kontext abhängig sind. Das Kulturelle in der Kulturpsychologie erschöpft sich eben nicht in der Gegenüberstellung von „westlichen“ und „östlichen“ Kulturen, sondern impliziert auch Subkulturen, lokale Gewohnheiten und bisweilen sogar familiäre Deutungsmuster (vgl. Straub & Layes, 2002). Heines Ansatz macht deutlich, dass Implizite Theorien sich hervorragend als Kernkonzepte in eine so verstandene Kulturpsychologie einpassen lassen. Wie sie kulturpsychologisch reformuliert werden können, wurde bereits dargestellt. Welche Schlussfolgerungen lassen sich nun aber aus einem solchermaßen modifizierten Konzept Impliziter Intelligenztheorien für dessen Erforschung ziehen? Folgende Punkte erscheinen hier von besonderer Wichtigkeit:

- Die Implizite Theorie einer Person ist, wie der Name eigentlich schon nahelegen sollte, nicht explizit repräsentiert oder gar sprachlich genau fassbar. Sie kann eher als Punkt auf einem Kontinuum zwischen statischer und dynamischer Auffassung verstanden werden. Vor allem aber ist sie vom situationalen Kontext abhängig. Daher ist es wenig fruchtbar, sie im Sinne einer Disposition zu betrachten und entsprechend durch kontextlose Fragebogen-Items zu erheben. Unter diesem Blickwinkel erscheint beispielsweise die Durchführung einer Manipulationskontrolle nach der Induktion Impliziter Theorien nicht mehr so fruchtbar bzw. notwendig, wie dies zuweilen geäußert wird. Eben weil eine Implizite Theorie Teil eines „meaning system“ ist, erscheint eine Trennung zwischen ihrer Ausprägung und ihren Auswirkungen künstlich. Folgerichtig wurde in den hier durchgeführten Studien 2 und 3 auf eine Manipulationskontrolle verzichtet, da diese notwendigerweise in dekontextualisierter Form erfolgen muss und somit weniger valide sein dürfte als die Erfassung ihrer konkreten Auswirkungen (vgl. Abschnitt 5.3). Dass die Ergebnisse der Studien 2 und 3 eher für eine erfolgreiche Induktion sprechen, die in Studie 1 durchgeführte Manipulationskontrolle (derselben Induktionstexte) dagegen fehlschlug, lässt sich als Bestätigung dieser Überlegungen verstehen.
- Wenn Implizite Intelligenztheorien ein wichtiger Bestandteil der Kultur sind, können sie auch als Ideologie missbraucht werden. Während dies für die statische Auffassung schon eingehend erörtert wurde, sollte auch das Ideologiepotenzial der dynamischen Theorie nicht unterschätzt werden. Dweck (2007b) scheint dies durchaus bewusst zu sein, wenn sie in ihrem populärwissenschaftlichem Buch warnt: „[H]eißt das, dass Menschen, die an einer bestimmten Situation scheitern, selbst daran schuld sind, weil sie sich nicht

genug angestrengt haben? Nein, auf keinen Fall. Natürlich ist die Anstrengung entscheidend für den Erfolg, doch sie ist bei weitem nicht der einzige Faktor“ (S. 60-61). Die Gefahr, dass die dynamische Auffassung als Mittel sozialer Kontrolle missbraucht werden könnte, wurde auch in einer der im Rahmen von Studie 2 aufgezeichneten Diskussionen geäußert: „Was gibt es sonst noch für Implikationen für die Gesellschaft? Schaffung von 'ner – angeblichen – Gleichheit, Chancengleichheit. Ich würd' eher *angeblichen* Chancengleichheit schreiben, weil, das hat man ja trotzdem nicht...“. Kritische Aufmerksamkeit in Richtung solchen Missbrauchs der dynamischen Intelligenzauffassung scheint also angebracht.

- Gemäß der hier vorgeschlagenen kulturpsychologischen Perspektive lassen sich Implizite Intelligenztheorien als Merkmal einer bestimmten Kultur verstehen. Dies legt die Vermutung nahe, dass sich verschiedene Kulturen oder Subkulturen hinsichtlich der jeweils vorherrschenden Intelligenztheorie und der einhergehenden Leistungsverhaltensweisen unterscheiden. Zu dieser Frage liegen noch keine systematischen Untersuchungen vor. Die hier gewählte kulturpsychologische Ausrichtung lässt indes erwarten, dass kultur- und milieuvergleichende Studien nicht nur für sich genommen interessant sind, sondern auch das Verständnis der allgemeinen Wirkmechanismen Impliziter Theorien erweitern könnten.
- Schließlich macht die kulturpsychologische Perspektive auch den Einbezug qualitativer Methoden unvermeidlich. Dabei könnten beispielsweise Gruppendiskussionen über die Induktionstexte, wie sie auch in Studie 2 aufgezeichnet wurden, diskursanalytisch ausgewertet werden (vgl. die Vorarbeiten bei Spinath, 2001). So ließe sich ergründen, anhand welcher Konzepte, Metaphern etc. sich die Probanden die Aussagen des Textes zu eigen machen. Darüber hinaus jedoch erscheint auch der Einsatz gezielter Interviews notwendig, die darüber aufklären könnten, welche Gedanken, Gefühle, Hoffnungen und Befürchtungen nach Lektüre des Induktionstextes auftreten bzw. ganz allgemein mit der jeweiligen Vorstellung von Intelligenz verbunden sind. Nur ein enges Zusammenspiel von experimentellen und qualitativen Verfahren lässt es aussichtsreich erscheinen, die in dieser Arbeit vorgefundenen uneinheitlichen Ergebnismuster aufzuklären und schließlich verstehbar zu machen, welcher Proband sich aus welchen Gründen und unter welchen Umständen von bestimmten Leistungsdiskursen in seinem Verhalten beeinflussen lässt. Typische Probleme des hier gewählten experimentellen Zugangs – etwa die Verwendung künstlicher Leistungsumwelten, das Messen von Emotionen anhand weniger, vorgegebener Aussagen oder die Beschränkung auf extrem kurzfristige Auswirkungen der Induktion Impliziter Theorien – könnten so kompensiert werden.

Die aufgeführten Implikationen legen nahe, dass eine theoretische Klärung des Konzepts Impliziter Intelligenztheorien maßgeblich zu einer Aufklärung der empirischen Inkonsistenzen beitragen könnte. Für eine solche Klärung sind die hier angestellten Überlegungen sicherlich nur ein erster Schritt. Dass sich hierbei aber der kulturpsychologische Blickwinkel als besonders fruchtbar erweist, sollte deutlich geworden sein. Das bereits implizit kulturalistische Modell Dwecks kann auf diese Weise zu Ende gedacht werden, indem eine Implizite Theorie weniger als „leere“ Disposition denn als kulturell aufgeladener Zustand verstanden wird. Damit rückt stärker der *Inhalt* Impliziter Theorien in den Vordergrund, welcher nicht zuletzt auch bestimmte Leistungsdiskurse wie den der Hirnforschung berühren kann. Mit einem Augenzwinkern könnte man daher die hier vorgeschlagene Neuausrichtung als Dynamisierung der bislang allzu statischen Impliziten Theorie Impliziter Intelligenztheorien bezeichnen.

Literaturverzeichnis

- Aronson, J. (1998). *The effects of conceiving ability as fixed or improvable on responses to stereotype threat*. Unpublished manuscript, University of Texas, Austin.
- Aronson, J., Fried, C., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, 38, 113-125.
- Assheuer, T. (2005). Hartz IV in der Synapse. *Die Zeit*, 14/2005, S.43, online abrufbar unter: <http://www.zeit.de/2005/14/hirn-politik>.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64, 359-372.
- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2000). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (9. Aufl.). Heidelberg: Springer.
- Baer, A. R., Grant, H., & Dweck, C. S. (2005). *Personal goals, dysphoria, & coping strategies*: Unpublished manuscript, Columbia University.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience*. Malden, MA: Blackwell.
- Bergen, R. (1991). *Beliefs about intelligence and achievement-related behaviors*. Unpublished doctoral dissertation, University of Illinois, Champaign-Urbana.
- Blackwell, L. S., Dweck, C. S., & Trzesniewski, K. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78(1), 246-263.
- Boehnke, K., Pelkner, A.-K., & Kurman, J. (2004). On the interrelation of peer climate and school performance in mathematics: A German-Canadian-Israeli comparison of 14-year-old school students. In B. N. Setiadi, A. Supratiknya, W. Lonner & Y. P. Poortinga (Hrsg.), *Ongoing themes in psychology and culture* (S. 415-432). Yogyakarta: IACCP.
- Broome, P. (2001). The gender-related influence of implicit theories of one's intelligence with regard to academic performance in introductory physics classes. *Psychologische Beiträge*, 43, 100-128.
- Brunstein, J., & Heckhausen, H. (2006). Motivation und Handeln. In H. Heckhausen & J. Heckhausen (Hrsg.), *Motivation und Handeln* (3. Aufl., S. 143-192). Heidelberg: Springer.
- Burr, V. (2003). *Social constructionism* (2. Aufl.). London: Routledge.
- Cury, F., Elliot, A. J., Da Fonseca, D., & Moller, A. C. (2006). The social-cognitive model of achievement motivation and the 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, 90(4), 666-679.
- Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science*, 314, 435.
- Davies, B., & Harré, R. (1990). Positioning: the discursive production of selves. *Journal for the Theory of Social Behavior*, 20(1), 43-63.
- Diener, C. I., & Dweck, C. S. (1978). An analysis of learned helplessness: Continuous changes in performances, strategy and achievement cognitions following failure. *Journal of Personality and Social Psychology*, 36, 451-462.
- Dresel, M., & Ziegler, A. (2006). Langfristige Förderung von Fähigkeitsselbstkonzept und impliziter Fähigkeitstheorie durch computerbasiertes attributionales Feedback. *Zeitschrift für Pädagogische Psychologie*, 20(1/2), 49-63.

- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.
- Dweck, C. S. (1999). *Self-Theories: Their role in motivation, personality and development*. Philadelphia: Taylor and Francis/Psychology Press.
- Dweck, C. S. (2007a). The perils and promises of praise. *Educational Leadership*, 65(2), 34-39.
- Dweck, C. S. (2007b). *Selbstbild. Wie unser Denken Erfolge oder Niederlagen bewirkt*. Frankfurt a.M.: Campus.
- Dweck, C. S., Chiu, C., & Hong, Y. Y. (1995a). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry*, 6, 267-285.
- Dweck, C. S., Chiu, C., & Hong, Y. Y. (1995b). Implicit theories: Elaboration and extension of the model. *Psychological Inquiry*, 6, 322-333.
- Dweck, C. S., & Grant, H. (2006). Self-theories, goals, and meaning. In J. Shah & W. Gardner (Hrsg.), *The handbook of motivational science* (S. 405-416). New York: Guilford.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.
- Dweck, C. S., & Molden, D. C. (2005). Self-theories: Their impact on competence motivation and acquisition. In A. J. Elliot & C. S. Dweck (Hrsg.), *The handbook of competence and motivation* (S. 122-140). New York: Guilford.
- Eberle, U. (2002). Die Hirnerzieherin. *Die Zeit*, 48/2002, S. 36.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109-132.
- Edwards, D. (1997). *Discourse and cognition*. London/Beverly Hills, CA: Sage.
- Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, 70, 461-475.
- Farah, M. J. (2002). Emerging ethical issues in neuroscience. *Nature Neuroscience*, 5(11), 1123-1129.
- Farah, M. J., Illes, J., Cook-Deegan, R., Gardner, H., Kandel, E., King, P., et al. (2004). Neurocognitive enhancement: What can we do and what should we do? *Nature Reviews Neuroscience*, 5, 421-425.
- Fukuyama. (2002). *Das Ende des Menschen*. Stuttgart: Deutsche Verlags Anstalt.
- Gazzaniga, M. S. (2006). Facts, fictions, and the future of neuroethics. In J. Illes (Hrsg.), *Neuroethics: Defining the issues in theory, practice, and policy* (S. 141-148). New York: Oxford University Press.
- Gergen, K. J. (1985). The social constructionist movement in modern psychology. *American Psychologist*, 40, 266-275.
- Gergen, K. J. (1994). *Realities and relationships. Soundings in social construction*. Cambridge, MA: Harvard University Press.
- Gergen, K. J. (2002). *Konstruierte Wirklichkeiten. Eine Hinführung zum sozialen Konstruktivismus*. Stuttgart: Kohlhammer.
- Geyer, C. (2004a). Vorwort. In C. Geyer (Hrsg.), *Hirnforschung und Willensfreiheit. Zur Deutung der neuesten Experimente* (S. 9-19). Frankfurt a.M.: Suhrkamp.
- Geyer, C. (Hrsg.). (2004b). *Hirnforschung und Willensfreiheit. Zur Deutung der neuesten Experimente*. Frankfurt/Main: Suhrkamp.
- Goetz, T., Zirngibl, A., & Pekrun, R. (2004). Lern- und Leistungsempfinden von Schülerinnen und Schülern. In T. Hascher (Hrsg.), *Schule positiv erleben* (S. 49-66). Bern: Haupt Verlag.
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, 24, 645-662.

- Grainger-Monson, M., & Karetzky, K. (2006). The mind in the movies: a neuroethical analysis of the portrayal of the mind in popular media. In J. Illes (Hrsg.), *Neuroethics: Defining the issues in theory, practice, and policy* (S. 297-311). New York: Oxford University Press.
- Grant, H., & Dweck, C. S. (2001). Cross-cultural response to failure: Considering outcome attributions within different goals. In F. Salili, C. Chiu & Y. Y. Hong (Hrsg.), *Student motivation: The culture and context of learning* (S. 203-219). New York: Plenum.
- Heckhausen, H. (1989). *Motivation und Handeln* (2. Aufl.). Berlin: Springer.
- Heckhausen, J., & Heckhausen, H. (2006). Motivation und Entwicklung. In H. Heckhausen & J. Heckhausen (Hrsg.), *Motivation und Handeln* (3. Aufl., S. 393-454). Heidelberg: Springer.
- Heine, S. J. (2008). *Cultural psychology*. New York: Norton.
- Heine, S. J., Kitayama, S., Lehman, D. R., Takata, T., Ide, E., Leung, C., et al. (2001). Divergent consequences of success and failure in Japan and North America. An investigation of self-improving motivations and malleable selves. *Journal of Personality and Social Psychology*, 81, 599-615.
- Heine, S. J., & Norenzayan, A. (2006). Towards a psychological science for a cultural species. *Perspectives on Psychological Science*, 1, 251-269.
- Helmke, A., & Rheinberg, F. (1996). Anstrengungsvermeidung - Morphologie eines Konstruktes. In C. Spiel, U. Kastner-Koller & P. Deimann (Hrsg.), *Motivation und Lernen aus der Perspektive lebenslanger Entwicklung* (S. 207-224). Münster: Waxmann.
- Hong, Y.-y., Chiu, C.-y., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, 77, 588-599.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Illes, J. (Hrsg.). (2006). *Neuroethics: Defining the issues in theory, practice, and policy*. New York: Oxford University Press.
- Jagacinski, C. M., & Nicholls, J. G. (1987). Competence and affect in task involvement and ego involvement: The impact of social comparison information. *Journal of Educational Psychology*, 79(2), 107-114.
- Jöreskog, K., & Sörbom, D. (2001). LISREL (Version 8.50): Scientific Software International.
- Jost, J. T., & Kruglanski, A. W. (2002). The estrangement of social constructionism and experimental social psychology: History of the rift and prospects for reconciliation. *Personality and Social Psychology Review*, 6, 168-187.
- Juvonen, J. (2001). The social functions of attributional face saving tactics. In F. Salili, C. Chiu & Y. Y. Hong (Hrsg.), *Student Motivation: The Culture and Context of Learning* (S. 61-77). New York: Kluwer Academic.
- Kessels, U., Warner, L. M., Holle, J., Hannover, B. (2008). Identitätsbedrohung durch positives schulisches Leistungsfeedback. Die Erledigung von Entwicklungsaufgaben im Konflikt mit schulischem Engagement. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40(1), 22-31
- Knollmann, M., & Wild, E. (2004). Schüleremotionen in schulischen und häuslichen Kontexten. In T. Hascher (Hrsg.), *Wohlbefinden in der Schule* (S. 67-87). Bern: Haupt.
- Könneker, C. (2007). Wer erklärt den Menschen? Warum die Neurowissenschaften heute en vogue sind - und wir den Dialog über sie dringend brauchen. In C. Könneker (Hrsg.), *Wer erklärt den Menschen? Hirnforscher, Psychologen und Philosophen im Dialog* (S. 11-18). Frankfurt a.M.: Fischer.
- Krüger, H.-P. (Hrsg.). (2007). *Hirn als Subjekt? Philosophische Grenzfragen der Neurobiologie*. Berlin: Akademie Verlag.
- Kuhlmann, A. (2004). Menschen im Begabungstest. Mutmaßungen über Hirnforschung als soziale Praxis. *WestEnd. Neue Zeitschrift für Sozialforschung*, 1(1), 143-153.

- Kulynnych, J. (2002). Legal and ethical issues in neuroimaging research: human subjects protection, medical privacy, and the public communication of research results. *Brain and Cognition*, 50, 345-357.
- Laucken, U. (1989). *Denkformen der Psychologie. Dargestellt am Entwurf einer Logographie der Gefühle*. Bern: Huber.
- Laucken, U. (2000). Sozialkonstruktivismus und die semantische Wissenschaft vom Menschen. *Handlung Kultur Interpretation*, 9, 37-65.
- Laucken, U. (2003). Über die semantische Blindheit einer neurowissenschaftlich gewendeten Psychologie. Oder: Was hätte uns eine so gewendete Psychologie zum "Dialog der Kulturen" zu sagen? *Journal für Psychologie*, 11, 149-175.
- Laux, L. (1993). Angst, Selbstwert und Selbstdarstellung. In L. Laux & H. Weber (Hrsg.), *Emotionsbewältigung und Selbstdarstellung* (S. 121-141). Stuttgart: Kohlhammer.
- Laux, L., & Weber, H. (1993). *Emotionsbewältigung und Selbstdarstellung*. Stuttgart: Kohlhammer.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.
- Lewis, M. (2000). The self-conscious emotions. In M. Lewis & J. Haviland-Jones (Hrsg.), *Handbook of emotions* (2. Aufl., S. 623-636). New York: Guilford.
- Lewis, M., & Sullivan, M. W. (2005). The development of self-conscious emotions. In A. J. Elliot & C. S. Dweck (Hrsg.), *The handbook of competence and motivation* (S. 185-201). New York: Guilford.
- Linnenbrink, E. A., & Pintrich, P. R. (2002). Achievement goal theory and affect: An asymmetrical bidirectional model. *Educational Psychologist*, 37(2), 69-78.
- MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, 17, 144-158.
- Mangels, J. A., Butterfield, B., Lamb, J., Good, C., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective Neuroscience*, 1, 75-86.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224-253.
- Martocchio, J. J. (1994). Effects of conceptions of ability on anxiety, self-efficacy, and learning in training. *Journal of Applied Psychology*, 79, 819-825.
- Meyer, W.-U. (2000). *Gelernte Hilflosigkeit: Grundlagen und Anwendungen in Schule und Unterricht*. Göttingen: Huber.
- Meyer, W.-U., Schützwohl, A., & Reisenzein, R. (2001). *Einführung in die Emotionspsychologie. Band I: Die Emotionstheorien von Watson, James und Schachter* (2. Aufl.). Bern: Huber.
- Miller, R. S. (2007). Is embarrassment a blessing or a curse? In J. L. Tracy, R. W. Robins & J. Tangney (Hrsg.), *The self-conscious emotions: theory and research* (S. 245-262). New York: Guilford.
- Molden, D. C., & Dweck, C. S. (2006). Finding "meaning" in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist*, 61(3), 192-203.
- Molfenter, S. (1999). *Prüfungsemotionen bei Studierenden: Explorative Analysen und Entwicklung eines diagnostischen Instrumenatriums*. Unveröff. Diss., Universität Regensburg.
- Mueller, C. M., & Dweck, C. S. (1998). Intelligence praise can undermine motivation and performance. *Journal of Personality and Social Psychology*, 75, 33-52.
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91, 328-346.
- Nightingale, D. J., & Cromby, J. (Hrsg.). (1999). *Social constructionist psychology. A critical analysis of theory and practice*. Buckingham: Open University Press.

- Niiya, Y., Crocker, J., & Bartmess, E. N. (2004). From vulnerability to resilience. Learning orientations buffer contingent self-esteem from failure. *Psychological Science*, 15(12), 801-805.
- Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.
- Otto, J. (2007). Meines kann schon mehr! *Die Zeit*, 37/2007, S. 73.
- Parens, E. (2006). Creativity, gratitude, and the enhancement debate. In J. Illes (Hrsg.), *Neuroethics: Defining the issues in theory, practice, and policy* (S. 75-86). New York: Oxford University Press.
- Parker, I. (2002). *Critical Discursive Psychology*. London: Palgrave.
- Parkinson, B., Fischer, A. H., & Manstead, A. S. R. (2005). *Emotions in social relations: Cultural, group and interpersonal processes*. New York: Psychology Press.
- Parkinson, B., & Manstead, A. S. R. (1992). Appraisal as a cause of emotion. In M. S. Clark (Hrsg.), *Review of personality and social psychology* (Vol. 13, S. 122-149). Newbury Park, CA: Sage.
- Paulus, J. (2003). Lernrezepte aus dem Hirnlabor. *Die Zeit*, 38/2003, S. 35.
- Pekrun, R. (1998). Schüleremotionen und ihre Förderung: Ein blinder Fleck in der Unterrichtsforschung. *Psychologie in Erziehung und Unterricht*, 44, 230-248.
- Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Hrsg.), *Motivational psychology of human development* (S. 143-163). Amsterdam: Elsevier.
- Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology*, 98(3), 583-597.
- Pekrun, R., Goetz, T., Perry, R. P., Kramer, K., Hochstadt, M., & Molfenter, S. (2004). Beyond Test Anxiety: Development and Validation of the Test Emotions Questionnaire (TEQ). *Anxiety, Stress, and Coping*, 17(3), 287-316.
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91-105.
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2. Aufl.). New Jersey: Pearson.
- Post, S. (2008). *Klimakatastrophe oder Katastrophenklima? Die Berichterstattung über den Klimawandel aus Sicht der Klimaforscher*. München: Fischer (Reinhard).
- Prinz, W. (2008). Im Nebelgebiet des Geistes. *Die Zeit*, 24/2005, S. 48.
- Ratner, C. (2006). *Cultural psychology: A perspective on psychological functioning and social reform*. Upper Saddle River, NJ: Erlbaum.
- Reinmann-Rothmeier, G., & Mandl, H. (2001). Unterrichten und Lernumgebungen gestalten. In A. Krapp & B. Weidemann (Hrsg.), *Pädagogische Psychologie* (S. 601- 646). Weinheim: Beltz.
- Reisenzein, R., Meyer, W.-U., & Schützwohl, A. (2003). *Einführung in die Emotionspsychologie. Band III: Kognitive Emotionstheorien*. Bern: Huber.
- Rese, M. (2005). Logistische Regression. In K. Backhaus, B. Erichson, W. Plinke & R. Weiber (Hrsg.), *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (9. Aufl., S. 105-143). Berlin: Springer.
- Rhodewalt, F., & Hill, S. K. (1995). Self-handicapping in the classroom: The effects of claimed self-handicaps in responses to academic failure. *Journal of Personality and Social Psychology*, 16, 397-416.
- Rhodewalt, F., & Vohs, K. D. (2005). Defensive strategies, motivation, and the self: A self-regulatory process view. In A. J. Elliot & C. S. Dweck (Hrsg.), *The handbook of competence and motivation* (S. 548-565). New York: Guilford.

- Richter, T. (2007). Wie analysiert man Interaktionen von metrischen und kategorialen Prädiktoren? Nicht mit Median-Splits! *Zeitschrift für Medienpsychologie*, 19, 116-125.
- Rietz, C., Rudinger, G., & Andres, J. (1996). Lineare Strukturgleichungsmodelle. In E. Erdfelder, R. Mausfeld, T. Meiser & G. Rudinger (Hrsg.), *Handbuch Quantitative Methoden* (S. 253-268). Weinheim: PVU.
- Robins, R. W., & Pals, J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity*, 1, 313-336.
- Robinson, M. D., & Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin*, 27, 1520-1532.
- Roth, G. (2004). Worüber dürfen Hirnforscher reden - und in welcher Weise? In C. Geyer (Hrsg.), *Hirnforschung und Willensfreiheit. Zur Deutung der neuesten Experimente* (S. 66-85). Frankfurt a.M.: Suhrkamp.
- Schlangen, B., & Stiensmeier-Pelster, J. (1997). Implizite Theorien der Intelligenz als Determinanten der Leistungsmotivation. *Zeitschrift für Pädagogische Psychologie*, 11, 167-176.
- Schuler, H., & Hell, B. (2005). *Analyse des schlussfolgernden und kreativen Denkens (ASK)*. Bern: Verlag Hans Huber.
- Seiffert, T. L. (1995). Academic goals and emotions: A test of two models. *The Journal of Psychology*, 129(5), 543-552.
- Seligman, M. (1979). *Erlernte Hilflosigkeit*. München: Urban & Schwarzenberg.
- Seligman, M., Abramson, L. Y., Semmel, A., & Baeyer, C. (1979). Depressive attributional style. *Journal of Abnormal Psychology*, 88, 242-247.
- Singer, W. (2002). Was kann ein Mensch wann lernen? In W. Singer (Hrsg.), *Der Beobachter im Gehirn* (S. 43-59). Frankfurt a.M.: Suhrkamp.
- Slavin, R. (1993). Kooperatives Lernen und Leistung: Eine empirisch fundierte Theorie. In G. Huber (Hrsg.), *Neue Perspektiven der Kooperation* (S. 151-170). Baltmannsweiler: Schneider-Verlag.
- Smedslund, U. (1988). *Psycho-Logic*. Berlin: Springer.
- Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, 48, 813-838.
- Smith, C. A., & Ellsworth, P. C. (1987). Patterns of appraisal and emotion related to taking an exam. *Journal of Personality and Social Psychology*, 52(3), 475-488.
- Sonntag, M. (1991). Klinische Psychologie zwischen Theorie und Praxis. In G. Hörmann & W. Körner (Hrsg.), *Klinische Psychologie. Ein kritisches Handbuch* (S. 15-38). Reinbek: Rowohlt.
- Spinath, B. (2001). *Implizite Theorien über die Veränderbarkeit von Intelligenz und Begabung als Bedingungen von Motivation und Leistung*. Lengerich: Pabst.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology*, 68, 797-811.
- Stipek, D. J., & Gralinski, H. J. (1996). Children's beliefs about intelligence and school performance. *Journal of Educational Psychology*, 3, 384-391.
- Straub, J. (1999). *Handlung, Interpretation, Kritik. Grundzüge einer textwissenschaftlichen Handlungs- und Kulturpsychologie*. Berlin/New York: de Gruyter.
- Straub, J. (2003). Kulturwissenschaftliche Psychologie. In F. Jäger & J. Straub (Hrsg.), *Kulturwissenschaften. Ein Handbuch. Band 2: Paradigmen und Disziplinen* (S. 568-591). Stuttgart: Metzler.
- Straub, J., & Layes, G. (2002). Kulturpsychologie, Kulturvergleichende Psychologie, Interkulturelle Psychologie - Forschung im Kontext der 'Glokalisierung'. *Handlung Kultur Interpretation*, 11, 334-381.
- Thompson, R. A., & Nelson, C. A. (2001). Developmental science and the media: Early brain development. *American Psychologist*, 56, 5-15.

- Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, 15, 103-125.
- Tracy, J. L., & Robins, R. W. (2007). The self in self-conscious emotions: A cognitive appraisal approach. In J. L. Tracy, R. W. Robins & J. Tangney (Hrsg.), *The self-conscious emotions: theory and research* (S. 3-20). New York: Guilford.
- Tracy, J. L., Robins, R. W., & Tangney, J. (Hrsg.). (2007). *The self-conscious emotions: theory and research*. New York: Guilford.
- Turner, J. C., Thorpe, P. K., & Meyer, D. K. (1998). Students' reports of motivation and negative affect: A theoretical and empirical analysis. *Journal of Educational Psychology*, 90(4), 758-771.
- Turner, J. E., & Schallert, D. L. (2001). Expectancy-value relationships of shame reactions and shame resiliency. *Journal of Educational Psychology*, 93(2), 320-329.
- von Tiling, J. (2008a). *Anwendungen des Sozialkonstruktivismus: Zwischen kritischer Psychologie und Psychotherapie*. Saarbrücken: VDM.
- von Tiling, J. (2008b). Sozialkonstruktionistische Psychologie und ihre praktische Anwendung. Möglichkeiten einer Neuausrichtung [39 Absätze]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [Online Journal]*, 9(1), Art. 44, Online verfügbar über <http://www.qualitative-research.net/fqs-texte/41-08/08-41-44-d.htm>.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Weber, H. (2004). Explorations in the Social Construction of Anger. *Motivation & Emotion*, 28(2), 197-219.
- Weber, H., & Titzmann, P. (2003). Ärgerbezogene Reaktionen und Ziele: Entwicklung eines neuen Fragebogens. *Diagnostica*, 49(3), 97-109.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573.
- Weiner, B. (1994). *Motivationspsychologie* (3. Aufl.). Weinheim: Beltz.
- Weiner, B. (1995). *Judgments of responsibility. A foundation for a theory of social conduct*. New York: Guilford.
- Weiner, B. (2005). Motivation from an attribution perspective and the social psychology of perceived competence. In A. J. Elliot & C. S. Dweck (Hrsg.), *The handbook of competence and motivation* (S. 73-84). New York: Guilford.
- White, M., & Epston, D. (1990). *Die Zähmung der Monster. Literarische Mittel zu therapeutischen Zwecken*. Heidelberg: Auer (Orig. 1990: Narrative means to therapeutic ends. New York: Norton).
- Wittgenstein, L. (1963). Philosophische Untersuchungen. In *Schriften*. Frankfurt/Main: Suhrkamp.
- Wittgenstein, L. (1999). Bemerkungen über die Philosophie der Psychologie, II. In L. Wittgenstein (Hrsg.), *Werkausgabe Band I: Bemerkungen über die Philosophie der Psychologie u.a.* Frankfurt a.M.: Suhrkamp.
- Wolpe, P. R. (2002). Treatment, enhancement, and the ethics of neurotherapeutics. *Brain and Cognition*, 50, 387-395.
- Zeidner, M. (1998). *Test anxiety: The state of the art*. New York: Plenum Press.
- Ziegler, A. (2001). Implicit theories as antecedents of motivation and behavior [Special Issue]. *Psychologische Beiträge*, 43, 1-199.
- Ziegler, A., & Heller, K. A. (1998). Motivationsförderung mit Hilfe eines Reattributionstrainings. *Psychologie in Erziehung und Unterricht*, 45, 216-229.
- Zielke, B. (2004). *Kognition und soziale Praxis. Der Soziale Konstruktivismus und die Perspektiven einer postkognitivistischen Psychologie*. Bielefeld: Transcript Verlag.