

BETTINA SEIPP & CHRISTINE SCHWARZER

Angst und Leistung – Eine Meta-Analyse empirischer Befunde

Anxiety and academic achievement – A meta-analysis of empirical findings

Summary: 126 studies regarding the relationship between anxiety and achievement were meta-analytically integrated. Studies date from 1975 to 1988 and have a total sample size of 36 626 and 156 independent effects. The overall-analysis yielded a population effect size of $r_g = -0.21$. Further analyses tried to find moderater variables which could explain the remaining variance. Unexpectedly, no moderating effect was found for gender, culture and stability of anxiety. Analyses for the anxiety components worry and emotionality, for general and test anxiety as well as for time of the anxiety measurement, however, showed systematic differences: The more the anxiety scale focussed on the cognitive aspect and the more situation-specific it was, the closer was the relationship towards academic achievement. A closer relationship was also found when anxiety was measured *after* the achievement situation and not before.

Zusammenfassung: 126 Studien zum Zusammenhang zwischen Angst und Leistung aus den Jahren 1975 bis 1988 mit einer Gesamtstichprobe von 36 626 Personen und 156 unabhängigen Effekten wurden meta-analytisch integriert. Eine alle Effekte umfassende Analyse führte zu einer Populations-Effektstärke von $r_g = -0.21$. Weitere Analysen versuchten, Moderator-Variablen aufzufinden, die die verbleibende Varianz aufklären könnten. Dabei zeigte sich erwartungswidrigerweise keine Moderatorwirkung für «Geschlecht», «Kultur» und «zeitliche Stabilität von Angst». Demgegenüber erbrachten die Analysen, die für die Angstkomponenten «Besorgtheit» und «Aufgeregtheit», «allgemeine Ängstlichkeit» und «Leistungsängstlichkeit» sowie «Zeitpunkt der Angstmessung» durchgeführt wurden, systematische Unterschiede: Je stärker das Angstinstrument auf die kognitive Komponente ausgerichtet und je situationspezifischer es ist, desto enger wird die Beziehung zu schulischer und akademischer Leistung. Ein engerer Zusammenhang zeigt sich ebenfalls dann, wenn Angst *nach* der Leistungssituation erhoben wird statt davor.

Im Rahmen der heute vorherrschenden kognitiven Gefühlstheorien wird das Gefühl Angst als Resultat der Einschätzung einer Situation als bedrohlich angesehen. Das Entsprechende gilt für Leistungsangst in schulischen beziehungsweise universitären Kontexten, wobei sich die Bedrohung gegen das Selbstkonzept richtet. Während hinsichtlich dieses weitgefaßten deskriptiven Konstrukts Einigkeit besteht und im Bereich der Beziehung zwischen Angst und Leistung eine negative Beziehung dokumentiert ist, zeichnen sich die Einschätzungen der Größe, der Generalisierbarkeit sowie die theoretischen Erklärungen dieses Zusammenhangs durch wenig Konsistenz aus. Dies beruht vor allem auf häufigen Widersprüchlichkeiten empirischer Ergebnisse.

Vereinzelt finden sich positive Beziehungen zwischen Angst und Leistung, die als förderlicher Effekt von Angst interpretiert werden, wohingegen negative Zusammenhänge auf deren hinderlichen Effekt zurückgeführt werden. Wann Angst tatsächlich förderlich und wann sie hinderlich ist, bringt am deutlichsten die «Drive-Theory» um Spence & Taylor (1951) zum Ausdruck: Unter Berücksichtigung der Überle-

gungen von Yerkes & Dodson (1908) wird Angst dann als förderlich angesehen, wenn die zu erbringende Leistung «einfach» ist, hinderlich dagegen dann, wenn sie «schwierig» ist. Vergleichbar damit werden negative Effekte bei «schwierigen» Aufgaben im klassischen Prüfungsangstansatz durch vermehrte aufgabenirrelevante Kognitionen erklärt (Mandler & Sarason 1952; Sarason 1960; Wine 1971). Wenn auch die Auffassung von «leicht» und «schwierig» bei verschiedenen Autoren variiert, sind intellektuelle Leistungsanforderungen durchweg als «schwierig» zu bezeichnen.

Damit gilt für den Zusammenhang zwischen Angst und schulischer oder universitärer Leistung zwar der negative Zusammenhang, und die Mehrzahl empirischer Ergebnisse bestätigt diese Annahme, jedoch variieren die ermittelten Zusammenhänge so beträchtlich, daß es unmöglich wird, von *dem* Zusammenhang zu sprechen oder ihn gar auf eine bestimmte Größe

festlegen zu wollen. Zahlreiche Studien machen es sich deshalb zum Ziel, die Uneinheitlichkeit durch systematisch wirkende Faktoren, vorzugsweise innerhalb des Angstkonstrukts, einseitig zu machen. In diesem Bemühen sind mehrere unterscheidbare Facetten des Angstkonstrukts zutage getreten:

- (a) Eine wesentliche Trennung unterscheidet *Allgemeine Angst* von *Leistungsangst* wobei letztere als situationspezifische Angst von Personen, die Leistungsanforderungen gegenüberstehen, definiert ist. *Allgemeine Angst* umfaßt demgegenüber unspezifische Gefühle von Bedrohung sowohl der körperlichen Unversehrtheit als auch des Selbstkonzepts.
- (b) Auf der zeitlichen Ebene bzw. am Schnittpunkt von Eigenschaftspsychologie und Situationismus wird zwischen *Ängstlichkeit als Eigenschaft* von Personen, als latente Bereitschaft, mit Angst zu reagieren, und einer *akuten Zustandsangst* unterschieden (Spielberger 1966). Dabei wird davon ausgegangen, daß die Bereitschaft zu Angst allein nicht genügen kann, um Leistung zu beeinträchtigen, wohingegen der akute Zustand die betroffene Person kognitiv in Anspruch nimmt und somit von der Aufgabe ablenkt.
- (c) In Einklang damit wird die Trennung in die beiden Komponenten Besorgtheit und Aufgeregtheit (Liebert & Morris 1967) zunächst auf State-Ebene angesiedelt und erbringt das Ergebnis, daß Besorgtheit eine deutlichere negative Beziehung zu Leistung aufweist als Aufgeregtheit. Bei der Übertragung auf Trait-Niveau wird dieser Nachweis schwieriger.

Selbst unter Berücksichtigung dieser Sub-Konstrukte bleibt jedoch die Ergebnis-Inkonsistenz bestehen. Offenbar müssen weitere Faktoren, die beispielsweise mit Merkmalen der untersuchten Stichprobe zusammenhängen oder sogar noch weiter konstrukt-fremd im Design oder in anderen Merkmalen der jeweiligen Studie begründet liegen, die Zusammenhänge spürbar mitbestimmen, so daß Aussagen kaum generell gemacht werden können, sondern nur differenzielle Gültigkeit haben.

1. Fragestellungen

Treffen die oben genannten Überlegungen zu, kann es nicht länger sinnvoll erscheinen, dem bereits bekannten Bestand weitere Studien hinzuzufügen. Stattdessen erscheint es angezeigt, die vorhandenen Ergebnisse zunächst einmal zusammenzufassen, um sich einen Überblick über das zu verschaffen, was bereits als hinreichend gesichert angesehen werden kann, und andererseits über das, was einer gezielten Weitererforschung bedarf. Von zentralem Interesse ist es dabei zu erfahren, wie groß, d.h. wie praktisch bedeutsam, die Beziehung zwischen Angst und Leistung überhaupt ist bzw. unter welchen spezifischen Bedingungen die darauf basierende Leistungsprognose optimierbar wird.

Im einzelnen ergeben sich folgende Fragen für die Meta-Analyse(n):

- (a) Stellt die Stichprobe von Untersuchungen eine homogene Gruppe dar, d.h. ist die mittlere Effekt-Stärke generalisierbar und somit sinnvoll interpretierbar? Wenn dies nicht der Fall ist:
 - (b) Welche Faktoren moderieren den Zusammenhang zwischen Angst und Leistung?
 - (c) Welche Effekte (Größe und Richtung) lassen sich für spezielle Populationscharakteristika und Konstrukte finden?
- Angesichts der bereits unübersehbar gewordenen Anzahl widersprüchlicher Resultate ist eine traditionelle Form der Zusammenfassung, wie ein Übersichtsartikel sie leistet, nicht ausreichend, um die oben genannten Anforderungen einlösen zu können. Stattdessen bedarf es einer quantitativen Integration.

2. Methode

2.1. Meta-Analyse

Zur Beantwortung der oben genannten Fragestellungen wird die Methode der *quantitativen Forschungsbefundintegration* oder *Meta-Analyse* herangezogen. Mit ihrer Hilfe ist es möglich, auch große Datenmengen systematisch und objektiv zu integrieren, Unterschiede in den Ergebnissen auf unterschiedliche Studienmerkmale zurückzuführen und die gefundenen Effekte bzw. die Beziehungen zu quantifizieren (d.h. Effekt-Stärken anzugeben). Das Resultat dieser Analyse ist unabhängig von den verwendeten Stichprobengrößen und der Art der statistischen Hypothesentestung, da sowohl signifi-

kante als auch nicht signifikante Ergebnisse integriert werden (vgl. Hornke 1983; Hornke & Tries 1979).

Der Begriff «Meta-Analyse» geht auf Glass zurück, der dieses Verfahren unter diesem Namen 1976 einführt. Er bezeichnet Meta-Analysen als «Analysen von Analysen», d.h. als statistische Analysen einer großen Zahl von Analyseergebnissen mit der Absicht, deren Resultate zu integrieren. Vergleichbar definieren Hunter, Schmidt & Jackson (1982, 137) Meta-Analysen als quantitative Sammlung und Analyse deskriptiver Statistiken über Studien hinweg.

Während Glass mit seiner Methode der Frage nach Treatment-Effekten einer Variablen nachgeht, versuchen Hunter, Schmidt & Jackson (1982) sowie Hedges & Olkins (1985) aufgrund einer vergrößerten Stichprobe die «wahre» Populations-Effektstärke zu schätzen. Da dieses zweite Verfahren hier angewendet wird, soll es kurz erläutert werden: In einem ersten Schritt werden (a) der gewogene Mittelwert und die beobachtete gewogene Varianz der Studieneffekte sowie (b) die auf Stichprobenfehler zurückgehende Varianz berechnet. Zur Ermittlung der korrigierten bzw. Residualvarianz wird die Stichprobenfehlervarianz von der beobachteten Varianz abgezogen (vgl. Hunter, Schmidt & Jackson, 1982, 35–44).

Im zweiten Schritt wird zunächst die Homogenität bzw. Heterogenität des Datensatzes bestimmt. Im Laufe ihrer Überlegungen kommen die Autoren zu drei Homogenitätskriterien für Datensätze: Hunter, Schmidt & Jackson (1982) entwickeln zunächst einen χ^2 -Test mit besonders hoher statistischer Stärke. Wird er nicht signifikant, ist bedeutsame Restvarianz auszuschließen. Weiterhin stellen die Autoren (Hunter, Schmidt & Jackson 1982) die Faustregel vor, daß wenigstens 75% der beobachteten Varianz durch Artefakte erklärt sein sollten, um von homogenen Gruppen sprechen zu können. Letztendlich wird von Stoffelmayr, Dillavou & Hunter (1983) und McDaniel, Hirsh, Schmidt, Raju & Hunter (1986) betont, daß die tatsächlich nach Abzug von Fehlervarianzen verbleibende Varianz (bzw. Standardabweichung) entscheidender sei als der Prozentsatz nicht erklärter Varianz. Hier gilt als Faustregel, daß eine Reststreuung, die kleiner als 25% der Populationeffektstärke ist, als vernachlässigbar angesehen werden kann. Da diese Reststreuung zur Berechnung des sogenannten Variationsbereichs herangezogen wird, ist dieses dritte Kriterium das wichtigste. Die Bestimmung homogener Gruppen erfolgt in den hier vorliegenden Analysen daher nach diesem Kriterium.

In Fälle homogener Effekte läßt sich um den Mittelwert «über alles», ähnlich dem Konfidenz-Intervall, der Variationsbereich dieser mittleren Korrelation bestimmen, wobei, wie oben erwähnt, nicht der Standardfehler des Mittelwertes, sondern die korrigierte Standardabweichung verwendet wird:

$$P(r - 1.96 \times S_c \leq \rho \leq r + 1.96 \times S_c) = 0.95 \quad (1)$$

Es wird deutlich, daß im Falle 100%iger Varianzaufklärung, wenn also keine Reststandardabweichung übrig bleibt, dieses Intervall verschwindet, so daß der geschätzte mittlere Koeffizient dem Populationskoeffizienten entspricht. Schließt im Falle vorhandener Restvarianz das Intervall einen Null-Effekt aus, so ist der Zusammenhang eindeutig negativ/positiv.

Im Falle heterogener Effekte dagegen müssen Moderatorvariablen zur Erklärung dieser Varianz herangezogen werden, anhand derer die Studienstichprobe aufgebrochen wird. Dabei wird für jede Stufe des potentiellen Moderators eine eigene Meta-Analyse durchgeführt, bis keine Restvarianz mehr vorhanden ist. Von tatsächlichem Vorhandensein eines Moderators kann nur dann gesprochen werden, wenn zwei Bedingungen erfüllt sind:

- (1) «Die mittlere Korrelation variiert von Subgruppe zu Subgruppe»; und
- (2) «die korrigierte Varianz ist in den Subgruppen im Mittel niedriger als für den gesamten Datensatz» (Hunter, Schmidt & Jackson 1982, 48).

2.2. Datenbasis

Aufgrund einer On-line-Computer-Literaturrecherche und einer unsystematischen Suche in einschlägigen Fachzeitschriften wurden 359 in- und ausländische Studien aus den Jahren 1975 bis 1988 lokalisiert. Studien wurden jedoch dann nicht in die Analyse aufgenommen, wenn

- (a) Angst allein aufgrund physiologischer Maße operationalisiert wurde,
- (b) Leistung als Intelligenzquotient, Anagramm-Lösungen, Gedächtnisleistungen etc. operationalisiert wurde,
- (c) die Richtung und/oder Größe der Effekte nicht als Produkt-Moment-Korrelation berechenbar war (z.B. im Falle von multiplen Regressionsgewichten, kanonischen Korrelationen).

In den verbleibenden Untersuchungen wurde (trotz der Charakterisierung der Mehrzahl der Studien als Korrelationsstudien) Angst als die unabhängige und Leistung als die abhängige Variable angesehen. Beide wurden durch die zu ihrer Messung in den Original-Untersuchungen eingesetzten Instrumente operationalisiert. Im Falle von Angst sind dies größtenteils Instrumente, die Eigenschaftsangst messen und die ferner die beiden Komponenten Aufgeregtheit und Besorgtheit zusammen erfassen. Leistung wurde verstanden als Punktwert in Meßinstrumenten wie z.B. standardisierte Tests (beispielsweise Subskalen des *Achievement College*

Tests, des Stanford Achievement Tests) oder als informelle Maße (wie Zeugnisnoten, Grade Point Average GPA) oder mündliche Prüfungen. Nach Ausschluß nicht integrierbarer Studien verblieben 126 Untersuchungen (35.1% der lokalisierten) mit 156 unabhängigen Effekten (zwei pro Studie bei getrennten Ergebnissen für männliche und weibliche Personen, sonst eine). Ihre Verteilung auf verschiedene Quellen zeigt Tabelle 1.

Tabelle 1: Verteilung der Studien auf verschiedene Quellen

QUELLE	HÄUFIGKEIT
USA	78
BRD	29
andere	19
Σ	126
<hr/>	
Zeitschriftenartikel	75
Buchkapitel	6
Monographien	5
Dissertationen	19
Tagungsberichte	21
Σ	126

In den 126 Studien wurden insgesamt 36 626 Personen ($\bar{N}=235$, $S=432.2$) untersucht. 126 Effektstärken wurden als Korrelationen berichtet; 30 waren Mittelwerte mit Streuungen oder t- bzw. F-Werte, die ebenfalls in Korrelationen transformiert wurden.

Wenn es darum geht, Subanalysen durchzuführen, stützt sich die Auswahl der jeweiligen Studien auf ein Kodiersystem, in dem theoretisch (und methodisch) relevante Merkmale der Einzelstudien a priori festgehalten sind. Im vorliegenden Fall wurde ein System mit 41 Kategorien erstellt (vgl. Seipp 1990).

3. Ergebnisse

3.1. Über alle Merkmale hinweg generalisiertes Ergebnis

Das erste Ergebnis befaßt sich mit der Frage nach der Homogenität der Effektstärken ES

und damit mit der Frage nach der Generalisierbarkeit der mittleren Effektstärke. Die Stamm- und-Blatt-Darstellung² der Effektstärken in Abbildung 1 zeigt eine symmetrische unimodale Verteilung.

Nach dem zentralen Grenzwerttheorem legte dieser Umstand der annähernden Normalverteilung die Vermutung nahe, daß die Effekte sämtlich derselben Grundgesamtheit entstammen, d.h. alle Effekte denselben zugrundeliegenden Populationsparameter schätzen und somit alle Abweichungen auf Stichprobenfehler zurückzuführen sind.

Derartige graphische Darstellungen sind jedoch keine Homogenitätstests, sondern lediglich erste Hinweise auf Homo- oder Heterogenität, die durch die von Hunter, Schmidt & Jackson (1982) begründeten statistischen Verfahren nachzuweisen waren.

Der von den Autoren genannte χ^2 -Test zur Überprüfung der Homogenität der Effektstärken wurden beim vorliegenden Datensatz hochsignifikant: $\chi^2=381$; $p<0.001$, was allerdings bei der statistischen Stärke dieses Tests nicht verwundert. Nach dem impliziten Homogenitätstest jedoch zeigt sich dasselbe Ergebnis: Nach Abzug des geschätzten Stichprobenfehlers blieb eine Residualvarianz von $S_c^2=0.006$, was

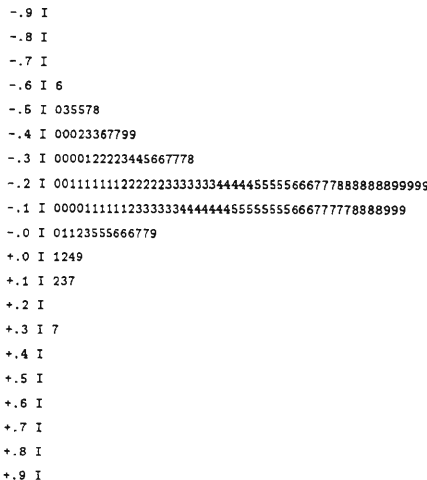


Abbildung 1: Stem-and-Leaf-Darstellung für 156 Effektstärken

bei einer beobachteten Varianz von $S_r^2=0.009$ einer prozentualen Varianzaufklärung durch Stichprobenfehler von nur 41% entspricht. Die Residualstreuung betrug $S_c=0.075$. Sie war größer als ein Viertel der mittleren ES ($=0.053$) und damit als nicht-trivial zu bezeichnen. Demnach war der Datensatz eindeutig als nicht homogen ausgewiesen, und eine Suche nach moderierenden Einflüssen und homogenen Untergruppen innerhalb der Gesamtstichprobe war notwendig, um die verbleibende Variation der ES von -0.36 bis -0.07 zu erklären. Im Überblick stellt sich das Ergebnis der Gesamtanalyse also wie in Tabelle 2 gezeigt dar.

Tabelle 2: Meta-Analyse für den Gesamtdatensatz zu Angst und Leistung

Variable	k	N	\bar{r}_g	S_r^2	S_c^2	S_c^2	S_c	% erklärte Varianz	95% Variationsbereich
alle	156	36626	-.212	.009*	.004*	.006*	.075	41.0	-.36 bis -.07

$$|\bar{r}_g| > 2S_c = .15, S_c > |\bar{r}_g| \times .25 = .053, \text{heterogen}$$

* Ungenauigkeiten durch Rundungsfehler

Dabei (und in allen weiteren Tabellen dieser Art) bedeuten: k: Anzahl der ES, N: Anzahl aller Personen, \bar{r}_g : gewogenes arithmetisches Mittel der Korrelationskoeffizienten, S_r^2 : beobachtete Varianz, S_c^2 : Stichprobenfehlervarianz, S_c^2 : korrigierte (Residual-)Varianz, S_c : korrigierte (Residual-)Streuung, % erklärte Varianz: prozentuale Varianzaufklärung durch Stichprobenfehler, $|\bar{r}_g| > / < 2S_c$: Signifikanzkriterium, $S_c > / < |\bar{r}_g| \times .25$: Homogenitätskriterium

Tabelle 3: Meta-Analysen für weibliche und männliche Personen

Variable	k	N	\bar{r}_g	S_r^2	S_c^2	S_c^2	S_c	% erkl. Varianz	95% Var.-Bereich
weibl.	38	5835	-.215	.009	.006	.004	.059	62.6	-.33 bis -.10

$$|\bar{r}_g| > .12, S_c > .054, \text{heterogen}$$

männl.	40	10921	-.182	.009	.003	.005	.073	38.8	-.33 bis -.04
--------	----	-------	-------	------	------	------	------	------	---------------

$$|\bar{r}_g| > .15, S_c > .046, \text{heterogen}$$

Das Ergebnis ist als eins der wichtigsten dieser Meta-Analysen anzusehen: Auf oberster Gliederungsebene ist das Geschlecht *kein* Moderator für die Beziehung zwischen Angst und Leistung. Die Differenz zwischen den jeweiligen mittleren Effektstärken beträgt nur $d=0.033$. Beide Gruppen sind zudem stark heterogen, so daß innerhalb jeder Geschlechtergruppe weitere Merkmale Einfluß nehmen.

3.2.2. Üben Kulturunterschiede eine Moderatorwirkung aus?

Die nationale Herkunft der untersuchten Personen ist aufgrund unterschiedlicher Studien- und Prüfungsbedingungen ein weiterer potentieller Moderator für die Enge und die Richtung des Zusammenhangs. Die folgende Analyse ging dieser Idee nach, indem sie die Effekte für Studien aus den USA, der Bundesrepublik Deutschland und einer Gruppe von weiteren Ländern, die Australien, Kanada, Griechenland, Holland, Italien, Indien, Israel, die Türkei und Ungarn umfaßt, verglich (s. Tabelle 4). Wie im vorhergehenden Fall sind auch hier die mittleren Effektstärken vergleichbar und von so weiten Variabilitätsbereichen umgeben, daß weitere Faktoren auf nächsttieferer Ebene wirksam sein müssen.

Tabelle 4: Meta-Analysen für die nationale Herkunft

Variable	k	N	\bar{r}_g	S_r^2	S_c^2	S_e^2	S_c	% erkl. Varianz	95% Var.-Bereich
USA	96	15003	-.207	.012	.006	.006	.076	50.6	-.35 bis -.06
$ \bar{r}_g > .15, S_c > .052, \text{heterogen}$									
BRD	37	14849	-.217	.006	.002	.004	.064	35.8	-.34 bis -.09
$ \bar{r}_g > .13, S_c > .054, \text{heterogen}$									
Rest	23	6774	-.211	.012	.003	.009	.093	26.6	-.39 bis -.03
$ \bar{r}_g > .19, S_c > .053, \text{heterogen}$									

3.2.3. Gibt es Geschlechts-Kultur-Interaktionen?

Die folgenden Analysen untersuchten eine eventuell differenzielle Beziehung zwischen Angst und Leistung in gemeinsamer Abhängigkeit vom Geschlecht und der nationalen Herkunft der untersuchten Personen, um so nochmals der vielzitierten Unterschiedlichkeit zwischen den Geschlechtern und der Heterogenität aller bisher besprochenen Gruppen Rechnung zu tragen (s. Tabelle 5).

Die Ergebnisse zeigen, daß sowohl für die Bundesrepublik als auch für dritte Länder kein Geschlechtsunterschied besteht. Für die Bundesrepublik sind zudem beide Gruppen absolut homogen, d.h. die jeweiligen Werte sind verläßlich geschätzt bzw. von weiteren Faktoren unbeeinflußt. Ähnliches gilt für weibliche Probanden in Drittländern. Hier besteht zwar eine gewisse Restvarianz, die aber im Vergleich zum Mittelwert so gering ist, daß sie nicht durch systematisch wirkende Faktoren, sondern nur durch (andere als Stichproben-) Fehler zu erklären ist.

Das Entscheidende dieser Analyse jedoch liegt im Moderatoreffekt des Geschlechts für die Gruppe der Studien aus den USA: Es zeigt sich hier eine Differenz von $d=0.074$, wobei der Effekt für weibliche Personen stärker negativ ist. Dies impliziert, daß die generelle Vorstellung, es bestehe eine engere negative Beziehung zwischen Angst und Leistung bei weiblichen Probanden, auf US-amerikanische Befunde zu-

Tabelle 5: Meta-Analysen für Geschlecht und Nation

Variable	k	N	\bar{r}_g	S_r^2	S_c^2	S_e^2	S_c	% erkl. Varianz	95% Var.-Bereich
BRD									
weiblich	8	2092	-.175	.002	.004	-.002	.000	100	-
<i>homogen</i>									
männlich	9	2154	-.183	.003	.004	-.001	.000	100	-
<i>homogen</i>									
USA									
weiblich	22	2926	-.241	.012	.007	.006	.076	53.9	-.39 to .09
$ \bar{r}_g > .15, S_c > .060, \text{heterogen}$									
männlich	18	3937	-.167	.008	.004	.004	.062	52.6	-.29 to .04
$ \bar{r}_g > .12, S_c > .042, \text{heterogen}$									
Andere									
weiblich	7	693	-.213	.009	.009	.000	.015	97.6	-.24 to -.18
$ \bar{r}_g > .03, S_c < .053, \text{homogen}$									
männlich	9	4310	-.206	.009	.002	.008	.088	19.9	-.38 to -.03
$ \bar{r}_g > .18, S_c > .054, \text{heterogen}$									

rückgeht: Die Mehrzahl an theoretischen und empirischen Arbeiten über das Thema stammt aus den USA, und die Postulierung des Geschlechtseffekts ist als eine Übergeneralisierung US-amerikanischer Ergebnisse zu verstehen.

3.2.4. Gibt es Unterschiede zwischen den Angst-Komponenten?

Ein dritter Befund, der in zahlreichen Studien nachgewiesen wird, postuliert eine deutlicher negative Beziehung zwischen den in Frage stehenden Variablen für die Angstkomponente Besorgtheit als für Aufgeregtheit. Die folgende Analyse (Tabelle 6) überprüfte, ob dies tatsäch-

lich für die Gesamtheit der integrierten Studien zutrifft.

Die Ergebnisse bestätigen den moderierenden Effekt sowie weitere Annahmen bezüglich der beiden Komponenten:

- Die Beziehung zwischen Angst und Leistung fällt für Aufgeregtheit niedriger aus als für Besorgtheit
- Die Differenz zwischen den beiden mittleren Effektstärken beträgt $d=0.07$ und ist aufgrund der gegenüber dem gemeinsamen Datensatz gesunkenen mittleren Restvarianz ($S_c^2=0.004 < S_e^2=0.006$) als systematisch zu betrachten.

Tabelle 6: Meta-Analysen für Worry und Emotionality

Variable	k	N	\bar{r}_g	S_r^2	S_c^2	S_c^2	S_c	% erkl. Varianz	95% Var.-Bereich
Worry	38	6885	-.219	.013	.005	.008	.090	38.0	-.40 bis -.04
$ \bar{r}_g > .18, S_c > .055, \text{heterogen}$									
Emo.	37	5182	-.147	.006	.007	-.001	.000	100	-
<i>homogen</i>									

- (c) Die Daten zur Komponente Aufgeregtheit stellen eine homogene Menge dar, wohingegen der Mittelwert für Besorgtheit von einem weiten Variabilitätsbereich umgeben ist. Die Beziehung Aufgeregtheit-Leistung wird somit von keinem weiteren Faktor beeinflusst, während weitere Merkmale innerhalb des Zusammenhangs Besorgtheit-Leistung imstande sind, diesen bis zu $\bar{r}_g = -0.40$ ansteigen oder aber ihn auf fast 0 sinken zu lassen.
- (d) Die Beziehung Besorgtheit-Leistung ist dem Gesamtergebnis vergleichbar, sowohl was die Populationseffektstärke als auch was den Variabilitätsbereich angeht. Dies legt nahe, daß «Angst» und «Besorgtheit» als äquivalent bezeichnet werden können, wenn es um die Vorhersage von Leistung aus Angst bzw. umgekehrt geht.

3.2.5. Gibt es Unterschiede zwischen Eigenschafts- und Zustandsangst?

Ausgehend von Spielbergers Konzeption von State- und Trait-Angst ist Eigenschaftsangst (bzw. Ängstlichkeit) eine dem Motiv vergleichbare latente Dimension, die einer ich-involvierenden Stimulation bedarf, um zur handlungswirksamen Zustandsangst aktiviert zu werden. Daher sollte die Beziehung zur Leistung für Zustandsangst stärker negativ sein als für Eigenschaftsangst. Die Ergebnisse zeigt Tabelle 7. Den Erwartungen entgegen fallen die Beziehungen identisch aus; das gleiche gilt für die Variabilitätsbereiche. An diesem Ergebnis änderte sich auch dann nichts, wenn nur diejenigen Studien integriert wurden, in denen State- und Trait-Angst an denselben Personen erfaßt wurden, oder in Analysen, für die die State- und Trait-Werte per Zufall zu gleichen Teilen aus dem Gesamtdatensatz ausgewählt wurden.

Tabelle 7.: Meta-Analysen für State- und Trait-Angst

Variable	k	N	\bar{r}_g	S_r^2	S_c^2	S_c^2	S_c	% erkl. Varianz	95% Var.-Bereich
State	29	3867	-.210	.014	.007	.007	.084	49.0	-.38 bis -.04
$ \bar{r}_g > .17, S_c > .053, \text{heterogen}$									
Trait	137	30670	-.210	.010	.004	.006	.079	39.3	-.37 bis -.05
$ \bar{r}_g > .16, S_c > .053, \text{heterogen}$									

3.2.6. Gibt es Unterschiede zwischen Allgemeiner Angst und Leistungsangst?

Eine letzte Frage aus dem Bereich des Angstkonstrukts befaßte sich mit Unterschieden zwischen dem Zusammenhang *Allgemeine Angst und Leistung* und *Leistungsangst und Leistung*. Es erscheint logisch, daß die Beziehung um so enger sein sollte, je situationsspezifischer das Angstinstrument ausgerichtet ist. Die folgenden Analysen (Tabelle 8) bestätigen diese Annahme: Die mittleren Effektstärken unterscheiden sich um $d=0.07$, und die mittlere Residualvarianz ist gegenüber der des Gesamtdatensatzes gesunken ($S_c^2=0.004 < S_c^2=0.006$). Selbst wenn auch in diesem Fall beide Subgruppen auf dieser Analyseebene nicht hinreichend charakterisiert sind, gilt die Situationsspezifität von

Angst - über alle weiteren denkbaren Faktoren hinweg generalisiert - als Moderator erster Ordnung (auf der obersten Analyseebene).

3.2.7. Gibt es Unterschiede für die Meßzeitpunkte?

Eine im vorliegenden Kontext interessante Frage befaßte sich mit dem Zusammenhang der interessierenden Variablen in Abhängigkeit vom Zeitpunkt, zu dem Angst relativ zur Leistungssituation erhoben wurde. Theoretisch wäre eine engere negative Beziehung zu erwarten, wenn zuerst die Angst gemessen und dann die Leistung zu erbringen ist, weil durch die Beschäftigung mit dem Instrument die darin zum Ausdruck kommenden negativen Kognitionen der Probanden verstärkt auf sie selbst gerichtet werden und so Kapazität von der Aufgabenlö-

Tabelle 8: Meta-Analysen für Allgemeine und Leistungsangst

Variable	k	N	\bar{r}_g	S_r^2	S_c^2	S_c^2	S_c	% erkl. Varianz	95% Var.-Bereich
Allg. Angst	53	11680	-.163	.007	.004	.003	.050	63.6	-.26 bis -.07
$ \bar{r}_g > .10, S_c > .041, \text{heterogen}$									
Leistungsangst	114	28424	-.233	.009	.004	.005	.070	42.0	-.37 bis -.09
$ \bar{r}_g > .14, S_c > .058, \text{heterogen}$									

Tabelle 9: Meta-Analysen für verschiedene Zeitpunkte der Angstmessung

Variable	k	N	\bar{r}_g	S_r^2	S_c^2	S_c^2	S_c	% erkl. Varianz	95% Var.-Bereich
vorher	31	4498	-.211	.009	.006	.003	.051	70.3	-.32 bis -.12
$ \bar{r}_g > .10, S_c < .055, \text{homogen}$									
nachher	9	2056	-.283	.001	.004	-.003	.000	100	-
<i>homogen</i>									
unabh.	35	10906	-.212	.009	.003	.006	.079	31.8	-.37 bis -.06
$ \bar{r}_g > .16, S_c > .053, \text{heterogen}$									

sung abgezogen wird. Um diese Annahme zu überprüfen, wurden aus den Untersuchungen drei unterschiedliche Datensätze zusammengestellt: In der ersten Gruppe erfolgten die Messungen in der Reihenfolge 1) Angst, 2) Leistung; in der zweiten Gruppe war diese Reihenfolge umgekehrt, und in der dritten war für die Probanden keine Beziehung zwischen beiden Messungen ersichtlich. Die Ergebnisse zeigt Tabelle 9.

Die Tabelle zeigt, daß eine deutlich stärker negative Beziehung besteht, wenn die Angst nach der Leistungssituation erhoben wird. Somit existiert zwar ein Unterschied für die Meßzeitpunkte, jedoch in einer der Erwartung widersprechenden Richtung.

Das Differenzmuster, das sich aus diesen Analysen ergibt (Tabelle 10), zeigt eine mittlere Differenz von $Diff=0.07$.

Tabelle 10: Differenzen und Residualvarianzen für die Zeitpunkte

	vorher	nachher	unabhängig
vorher	.003	.07	(.00)
nachher		-.003	.07
unabhängig			.006

Die mittlere Residualvarianz ist gegenüber der zusammengesetzten Gruppe auf 0.002 gesunken. Somit ist der Zeitpunkt der Angstmessung als Moderator erster Ordnung im Zusammenhang zwischen Angst und Leistung qualifiziert. Dieser Befund legt die Möglichkeit nahe, daß Leistung nicht unbedingt in so hohem Maße durch Angst beeinträchtigt wird, sondern daß auch das tatsächliche oder erwartete Leistungsergebnis für das Ausmaß an Angst entscheidend ist, das im Anschluß an eine Leistung erlebt wird. Eine alternative Erklärung läßt sich darin sehen, daß nachträglich berichtete Angst von den Probanden absichtlich als Entschuldigung für schlechte Leistung herangezogen wird (vgl. Laux & Glanzmann 1987).

4. Zusammenfassung und Diskussion

Die hier durchgeführten Meta-Analysen zielten darauf ab,

(a) die Generalisierbarkeit der Beziehung zwi-

schen Angst und Leistung zu erfassen sowie deren Größe und Richtung zu bestimmen, (b) unterschiedliche Effekte für unterscheidbare Populationen und Angstfacetten zu ermitteln.

Die 156 integrierten Effekt-Stärken zeigten eine Spannweite von $r=-0.66$ bis $r=+0.37$. Das macht die Inkonsistenz der empirisch ermittelten Zusammenhänge und die bisherige Unmöglichkeit, Aussagen über die tatsächliche Beziehung zu machen, deutlich. Da unterschiedliche Operationalisierungen der beiden Variablen offensichtlich einen Einfluß auf die einzelnen Ergebnisse haben, kann der berechneten mittleren ES von $r_g=-0.212$ nur geringe Aussagekraft beigemessen werden. Sie stellt lediglich eine grobe Schätzung des über alle denkbaren Randbedingungen hinweg generalisierten Zusammenhangs dar, ohne zu berücksichtigen, daß wesentlich niedrigere beziehungsweise höhere Zusammenhänge unter spezifischen Bedingungen auftreten. Diese hier ermittelten Ergebnisse sowie die von Hembree (1988) unter ähnlichen Fragestellungen ermittelten Effektstärken stützen sich gegenseitig.

Dem auf der höchsten Generalisierungsebene angesiedelten und als niedrig zu bezeichnenden negativen Zusammenhang kommt, gemessen am herkömmlichen Kriterium für praktische Relevanz, r^2 , eine nur geringe Bedeutung zu: Nur 4.5% der Leistungsvariation werden durch Unterschiede in der Angstausrprägung erklärt. Der Interpretationsrahmen gerade kleiner Effektstärken oder mittlerer Korrelationen jedoch ist belangvoller, als es der numerische Wert suggeriert. Die bloße Orientierung am Korrelations- oder Determinationskoeffizienten ist unergiebig. Dies zeigt sich unter anderem daran, daß sich das identische Ergebnis anders darstellt, wenn man das Absinken der Erfolgsrate zwischen Ängstlichen und Nicht-Ängstlichen anhand des «Binomial Effect Size Display» (BESD; Rosenthal & Rubin 1982) betrachtet. Der Zusammenhang zwischen Angst und Leistung von $r_g=-0.212$ besagt, daß 60% der Ängstlichen Leistungseinbußen hinnehmen müssen gegenüber nur 40% der Nicht-Ängstlichen (vorausgesetzt, die Personen sind zu annähernd gleichen Teilen auf Ängstliche und Nicht-Ängstliche aufgeteilt (s. Tabelle 11).

Tabelle 11: BESD für den Gesamtdatensatz

	Erfolg	Mißerfolg	
Ängstliche	7142 (39%)	11171 (61%)	18313
Nicht-Ängstliche	11171 (61%)	7142 (39%)	18313
	18313	18313	36626

Aus einer Gruppe von 18 313 Ängstlichen sind demnach 7142 erfolgreich und 11 171 nicht erfolgreich, während in einer Gruppe von 18 313 Nicht-Ängstlichen 11 171 Erfolg haben gegenüber 7142, die Mißerfolg haben.

Eingeschränkt wird die Bedeutung solcher, über das BESD ermittelter, Unterschiede in derartigen Fällen gegenüber solchen mit dichotomen Ereignissen allerdings dadurch, daß Erfolg und Mißerfolg auf einer «uneindeutigen» Skala liegen. Neben einem objektiven Gütemaßstab, anhand dessen Erfolg und Mißerfolg zu weitreichenden Konsequenzen führen, existiert ein subjektiver Maßstab, der «lediglich» Konsequenzen für das Wohlbefinden hat und bei unterschiedlichen Personen unterschiedliche Wendepunkte annimmt. Die Produkt-Moment-Korrelation gibt keinerlei Hinweis darauf, wie Erfolg in Mißerfolg übergeht, – das selbe gilt für Hoch- und Niedrigängstlichkeit. Dennoch ist festzuhalten, daß die Beziehung für diejenigen Personen, die zu den 11 171 der «Mißerfolgs-Gruppe» gehören, subjektiv von Bedeutung ist.

Im übrigen läßt sich die Korrelation von $r_g=-0.212$ im Sinne von Glass auch als standardisierte Mittelwertdifferenz von $d=-0.43$ darstellen. Unter der Voraussetzung der Normalverteilung bedeutet dies, daß Ängstliche und Nicht-Ängstliche um fast eine halbe Standardabweichung auf einer fiktiven Leistungsskala auseinander liegen.

Diese drei Darstellungsarten charakterisieren den groben Zusammenhang zwischen Angst und Leistung. Dieser trägt jedoch der angetroffenen Inkonsistenz der Ergebnisse keine Rechnung. Nach Abzug von Stichprobenfehler als Varianzquelle blieb noch immer ein auf systematische Quellen zurückführbarer Variationsbereich von $r_g=-0.36$ bis $r_g=-0.07$. Damit wurden zwar auf höchster Generalisierungsebene alle positiven Effekte durch Stichprobenfehler erklärlich, so daß der tatsächliche Zusammenhang als eindeutig negativ zu charakterisieren war. Dennoch war der Datensatz bei einer derartig großen Restvariation als heterogen zu be-

trachten. Deshalb wurde in weiteren Analysephasen versucht, einige derjenigen Quellen ausfindig zu machen, die für die verbleibende Variation verantwortlich sind. Dazu wurden die im Kodiersystem enthaltenen Kategorien und Items herangezogen¹.

Verschiedene Faktoren konnten als Moderatoren erster Ordnung des Zusammenhangs zwischen Angst und Leistung aufgezogen werden:

- Besorgtheit anstelle von Aufgeregtheit,
- Leistungsangst anstelle von Allgemeiner Angst,
- nachträglich anstelle von vorher gemessener Angst haben eine stärker negative Beziehung zu Leistung.

Die oft besprochene enger negative Beziehung zwischen den beiden Größen für weibliche als für männliche Befragte stellte sich als eine Übergeneralisierung von Befunden aus den USA heraus (Moderator zweiter Ordnung). Erwartungswidrigerweise machte es keinen Unterschied, ob Angst als Eigenschafts- oder als Zustandsangst erfaßt wurde. Möglicherweise ist das Ausfüllen eines Angstfragebogens für ängstliche Personen bereits hinreichend, um ein erhöhtes Maß an Zustandsangst auszulösen oder die Eigenschaft-Instruktion wird von den Probanden ignoriert.

Die trotz vielfacher Aufspaltung des Konstrukts noch zu verzeichnende Heterogenität legt es nahe, zukünftig verstärkt auf solche Einflußfaktoren einzugehen, die außerhalb des Angstkonstruktes selbst zu suchen sind. Ergebnisse, die in diese Richtung weisen, liegen beispielsweise von Helmke (1988), Krohne (1980), Nickel & Schlüter (1970) oder Nickel, Schlüter & Fenner (1973) vor, in denen Variablen wie Klassen- bzw. Kommunikationsstrukturen, Coping-Stile, Unterrichtsstile, Lehrerpersönlichkeit und Schularten mitherangezogen werden. Auf der einen Seite ist Angst somit als ein vielschichtiges Konstrukt bestätigt worden, dem jede weitere Forschung auf diesem Gebiet Rechnung tragen muß. Dies bedeutet aber, gerade im Zusammenhang mit Leistungsprognosen aus Angst, daß es aufschlußreich und gewinnbringend sein kann, sich auf einige Konstruktfacetten zu beschränken. Beispielsweise ließe sich die Vorhersage erhöhen, würde Angst

zu diesem Zweck nur als Leistungsangst und diese wieder nur als Besorgtheit erfaßt. Zum anderen zeigte sich, daß gerade die kognitive Komponente der Angst weiterführender Forschung bedarf: Die Heterogenität dieses Konstrukts spiegelt sich ja bereits in Untersuchungen von Hodapp (1989), Rost & Schermer (1987a, b; 1989a, b), Schwarzer & Kim (1984), Schwarzer (1984), Schwarzer & Quast (1985) wider. Mit der vorliegenden Meta-Analyse ist ein Schritt zur Differenzierung des Angstkonstrukts erfolgt, und zwar durch die Validierung am Außenkriterium Leistung. Dabei ist dieses Kriterium selbst noch nicht hinreichend differenziert worden. In späteren Arbeiten sollte geklärt werden, welche Arten von Leistung mit welchen Komponenten der Angst in Beziehung stehen.

Anmerkungen

- 1 Diese hohe Ausschlußrate steht im Einklang mit derjenigen anderer Meta-Analytiker (vgl. Seipp 1990).
- 2 Im *stem-and-leaf-display* beinhaltet der Stamm die Dezimalstellen von +9 bis -9, die Blätter zeigen in aufsteigender Richtung die Anzahl und Größe der Effektstärken der jeweiligen Dezimalordnung.
- 3 Zahlreiche weitere Relationen sowie homogene Subdatensätze finden sich in Seipp & Schwarzer (1987), Schwarzer, Seipp & Schwarzer (1989), Seipp (1990).

Literatur

- Baron-Boldt, J., Schuler, H. & Funke, U. (1988). Prädiktive Validität von Schulabschlußnoten: Eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 2, 79-90.
- Fricke, R. & Treinies, G. (1985). Einführung in die Meta-Analyse. Bern: Huber.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3-8.
- Hedges, L. V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Helmke, A. (1988). The role of classroom context for the achievement-impairing effect of test anxiety. *Anxiety Research*, 1, 37-52.
- Hembree, R. (1988). Correlates, causes, effects and treatment of test anxiety. *Review of Educational Research*, 58, 47-77.
- Hodapp, V. (1989). An expanded version of the German Test Anxiety Inventory. Paper presented at the 20. Conference of the Society for Test Anxiety Research. Amsterdam, June/July 1989.
- Hornke, L. F. (1983). Integration empirischer Forschungsergebnisse? *Psychologie in Erziehung und Unterricht*, 30, 54-63.
- Hornke, L. F. & Tries, J. (1979). Zusammenfassung empirisch-pädagogischer Ergebnisse im Sinne der Meta-Analyse nach Glass. Dokumentation der Beiträge der 27. Tagung der Arbeitsgemeinschaft für Empirisch Päd-

- agogische Forschung. Bremen: Universität Bremen, 158-171.
- Hunter, J. E., Schmidt, F. L. & Jackson, G. B. (1982). *Meta-analysis. Cumulating research findings across studies*. Beverly Hills: Sage.
- Krohne, H. W. (1980). Prüfungsangst: Defensive Motivation in selbstwertrelevanten Situationen. *Unterrichtswissenschaft*, 3, 226-242.
- Laux, L. & Glanzmann, P. (1987). A self-presentational view of test anxiety. In Schwarzer, R., van der Ploeg, H. M. & Spielberger, C. D. (Eds.). *Advances in test anxiety research*, Vol. 5. Lisse: Swets and Zeitlinger, 31-37.
- Levine, R. L. & Hunter, J. W. (1983). Regression methodology: Correlation, meta-analysis, confidence intervals, and reliability. *Journal of Leisure Research*, 15, 323-343.
- Liebert, R. M. & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 20, 975-978.
- Mandler, G. & Sarason, S. B. (1952). A study of anxiety and learning. *Journal of Abnormal and Social Psychology*, 47, 166-173.
- McDaniel, M. A., Hirsh, H. R., Schmidt, F. L., Raju, N. S. & Hunter, J. E. (1986). Interpreting the results of meta-analytic research: A comment on Schmitt, Gooding, Noe, and Kirsch (1984). *Personnel Psychology*, 39, 141-148.
- Nickel, H. & Schlüter, P. (1970). Angstwerte bei Hauptschülern und ihr Zusammenhang mit Leistungs- sowie Verhaltensmerkmalen, Lehrerurteil und Unterrichtsstil. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 2, 125-136.
- Nickel, H., Schlüter, P. & Fenner, H.-J. (1973). Angstwerte, Intelligenztest- und Schulleistungen sowie der Einfluß der Lehrerpersönlichkeit bei Schülern verschiedener Schularten. *Psychologie in Erziehung und Unterricht*, 20, 1-13.
- Rosenthal, R. & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, 74, 166-169.
- Rost, D. H. & Schermer, F. J. (1987a). Auf dem Weg zu einer differentiellen Diagnostik der Leistungsangst. *Psychologische Rundschau*, 38, 14-36.
- Rost, D. H. & Schermer, F. J. (1987b). Emotion and cognition in coping with test anxiety. *Communication and Cognition*, 20, 225-244.
- Rost, D. H. & Schermer, F. J. (1989a). «Reaktionsweisen gegenüber Tests» (RTT) und «Manifestationen von Leistungsangst» (DAIMAN): una eademque res? *Zeitschrift für Differentielle und Diagnostische Psychologie*, 10, 169-179.
- Rost, D. H. & Schermer, F. J. (1989b). Diagnostik des Leistungsangsterlebens. *Diagnostica*, 35, 284-314.
- Sarason, I. G. (1960). Empirical findings and theoretical problems in the use of anxiety scales. *Psychological Bulletin*, 57, 403-415.
- Schwarzer, C. & Kim, M.-J. (1984). Adaptation of the Korean form of the Test Anxiety Inventory: A research note. In van der Ploeg, H. M., Schwarzer, R. & Spielberger, C. D. (Eds.). *Advances in test anxiety research*, Vol. 3. Lisse: Swets and Zeitlinger, 277-285.
- Schwarzer, R. (1981). *Streß, Angst und Hilflosigkeit*. Stuttgart: Kohlhammer.
- Schwarzer, R. (1984). Worry and emotionality as separate components in test anxiety. *International Review of Applied Psychology*, 33, 205-220.

- Schwarzer, R. & Quast, H.-H. (1985). Multidimensionality of the anxiety experience: Evidence for additional components. In van der Ploeg, H. M., Schwarzer, R. & Spielberger, C. D. (Eds.). *Advances in test anxiety research*, Vol. 4. Lisse: Swets and Zeitlinger, 3-14.
- Schwarzer, R., Seipp, B. & Schwarzer, C. (1989). Mathematics performance and anxiety: A meta-analysis. In Schwarzer, R., van der Ploeg, H. M. & Spielberger, C. D. (Eds.). *Advances in test anxiety research*, Vol. 6. Lisse: Swets and Zeitlinger, 105-119.
- Seipp, B. (1990). Angst und Leistung in Schule und Hochschule: Eine Meta-Analyse. Frankfurt: Peter Lang.
- Seipp, B. & Schwarzer, C. (1987). Anxiety and mathematics achievement: A meta-analysis. In Schwarzer, C. & Seipp, B. (Eds.). *Trends in European educational research*. Braunschweig: Braunschweiger Studien zur Erziehungs- und Sozialarbeitswissenschaft, 89-99.

- Spence, K. W. & Taylor, J. A. (1951). Anxiety and strength of the US as determiners of the amount of eyelid conditioning. *Journal of Experimental Psychology*, 42, 183-188.
- Spielberger, C. D. (1966). *Theory and research on anxiety*. In Spielberger, C. D. (Ed.). *Anxiety and behavior*. New York: Academic Press, 3-20.
- Stoffelmayr, B. E., Dillavou, D. & Hunter, J. E. (1983). Premorbid functioning and outcome in schizophrenia: A cumulative analysis. *Journal of Consulting and Clinical Psychology*, 51, 338-352.
- Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.
- Wine, J. D. (1971). Test anxiety and direction of attention. *Psychological Bulletin*, 70, 92-104.

Prof. Dr. Christine Schwarzer, Dr. Bettina Seipp, Heinrich-Heine-Universität Düsseldorf, Abteilung für Bildungsforschung und Pädagogische Beratung, Universitätsstraße 1, D-W-4000 Düsseldorf, Telefon (0211) 311 2039 bzw. 311 3474