



Universität Potsdam

Andreas Krapp, Ulrich Schiefele, Inge Schreyer

Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung

first published in:
Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 10
(1993) 2, S. 120-148

Postprint published at the Institutional Repository of Potsdam University:
In: Postprints der Universität Potsdam
Humanwissenschaftliche Reihe ; 56
<http://opus.kobv.de/ubp/volltexte/2009/3356/>
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-33561>

Postprints der Universität Potsdam
Humanwissenschaftliche Reihe ; 56

Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung*

Ulrich Schiefele, Andreas Krapp und Inge Schreyer
Universität der Bundeswehr München

Der Artikel berichtet von den Ergebnissen einer Metaanalyse des Zusammenhangs zwischen der Ausprägung fachspezifischer Interessen und schulischer Leistung. Es konnten für den Zeitraum 1965–1990 21 relevante Untersuchungsberichte identifiziert werden, die insgesamt 127 voneinander unabhängige (d. h. auf verschiedenen Stichproben beruhende) Korrelationskoeffizienten berichten. Ungeachtet der Einflüsse von Moderatorvariablen ergab sich für den Zusammenhang von Interesse und Leistung eine Durchschnittskorrelation von .30. Darüber hinaus zeigte sich, daß für Jungen signifikant höhere Interesse-Leistungs-Korrelationen bestehen als für Mädchen. Zwischen den verschiedenen Fächern ergaben sich teilweise ebenfalls signifikante Unterschiede. Dagegen war kein signifikanter Moderatoreffekt der Klassenstufe zu beobachten. Abschließend erfolgt eine Diskussion der Befunde auf dem Hintergrund interessentheoretischer und methodischer Überlegungen.

Die Erklärung und Vorhersage der Schulleistung ist ein wichtiges Forschungsgebiet der Pädagogischen Psychologie. Die besondere Bedeutung dieser Forschungsrichtung resultiert aus dem Sachverhalt, daß in einem modernen Bildungssystem sehr viele Entscheidungen getroffen werden müssen, die sich auf prognostische Schätzungen des Schulerfolgs stützen. Dazu gehören z. B. die Wahl des optimalen Zeitpunkts für den Schulbeginn (Schulreifeprüfung), die Auswahl eines geeigneten Schulzweigs (z. B. neusprachlich vs. naturwissenschaftlich), die Aufnahme an eine Universität oder die Entscheidung für eine bestimmte Studienrichtung. Diese Entscheidungen sind für die Betroffenen meist von großer Tragweite. Häufig lassen sich Eltern und Schüler von Experten beraten. Manche Bildungsinstitutio-

* Die Autoren möchten sich herzlich bei Herrn Dr. Adi Winteler bedanken, der bei der Suche und Aufarbeitung der Literatur wesentliche Hilfe geleistet hat.

Alle Korrespondenz sollte an Ulrich Schiefele, Universität der Bundeswehr, Institut für Erziehungswissenschaft und Pädagogische Psychologie, Werner-Heisenberg-Weg 39, 8014 Neubiberg, Tel. (089) 6004 2719, gerichtet werden.

nen haben eigene Aufnahmeprüfungen entwickelt und stützen ihre Entscheidungen auf systematisch gewonnene diagnostische Daten.

Ein zentraler Gesichtspunkt bei der Vorbereitung bildungsrelevanter Entscheidungen sind Einschätzungen des künftigen Schul- oder Studienerfolgs. Prognosen sind ein elementarer Bestandteil rationaler Entscheidungen, gleichgültig ob sie im Rahmen individueller Wertbezüge von einzelnen Personen oder im Kontext institutioneller Maßnahmen getroffen werden (Cronbach & Gleser, 1965; Jungermann, 1976; Krapp, 1979; Wiczerkowski & zur Oeveste, 1982).

Jede bildungsrelevante Prognose stützt sich im Prinzip auf zwei Gruppen von Daten: (1) Daten über die Anforderungen und Bedingungen des gewünschten Ausbildungswegs (z. B. die curricularen Ziele, das Anforderungsniveau oder die Qualifikation der Lehrenden) und (2) Daten über individuelle Lernvoraussetzungen (z. B. kognitive Fähigkeiten oder motivationale Orientierungen). Je weitreichender die Entscheidung ist, d. h. je ferner die endgültige Bewertung des Erfolgs einer Schullaufbahn in der Zukunft liegt, desto stärker muß sich die Entscheidung auf dauerhafte Merkmale des Schülerverhaltens stützen. Weil für solche Entscheidungen Theorien, Befunde und Methoden zur Erfassung prognostisch relevanter Merkmale notwendig sind, befaßt sich die Pädagogische Psychologie nicht nur mit prozessualen und situationsspezifischen Faktoren des Lernens (z. B. Mechanismus der Informationsspeicherung, Lehrer-Schüler-Interaktion), sondern ebenso mit längerfristig wirksamen, relativ stabilen Determinanten des Schulerfolgs (Heller, 1991).

Die jahrelangen Bemühungen um eine möglichst exakte Prognose des Schul- und Studienerfolgs haben ihren Niederschlag in einer schier unübersehbaren Fülle von Untersuchungen gefunden. Übersichten finden sich u. a. bei Lavin (1965), Krapp (1973), Bloom (1976), Tiedemann (1977), Fleming und Malone (1983), Kühn (1983), Steinkamp und Maehr (1983) und Sauer und Gattringer (1985). Gelegentlich wurde auch der Versuch unternommen, die Befunde zu integrieren und die große Zahl von Einzelprädiktoren in theoretisch plausible Klassen einzuordnen (vgl. Kühn, 1983; Krapp, 1984; Haertel, Walberg & Weinstein, 1983).

Die meisten Autoren gelangen zur Unterscheidung von drei Gruppen leistungsbeeinflussender Faktoren: Schülermerkmale (z. B. Intelligenz), häusliche Umwelt (z. B. Sozialstatus der Eltern) und schulische Umwelt (z. B. Qualität des Unterrichts). Bildungsrelevante Entscheidungen stützen sich in erster Linie auf die Diagnose von Schülermerkmalen, die wiederum in drei Gruppen klassifiziert werden können: allgemeine kognitive Faktoren (z. B. Vorwissen, Intelligenz), allgemeine motivationale Faktoren (z. B. generelles Leistungsmotiv) und spezifische Präferenzen für bestimmte Lerninhalte. Die letztgenannte Gruppe bezeichnet man üblicherweise als *Interessen*.

Als gesichert gilt, daß die kognitiven Faktoren, insbesondere Intelligenz, die höchste prognostische Valenz besitzen. In empirischen Untersuchungen erklären

sie den größten Anteil an beobachteter Varianz der Leistung (z. B. Bloom, 1976; Kuusinen & Leskinen, 1988; Schneider & Bös, 1985). Man ist allgemein der Auffassung, daß motivationale und emotionale Faktoren weniger wichtig sind (z. B. Kühn, 1983; Steinkamp & Maehr, 1983). Auch neuere, kausalanalytische Studien (Parkerson, Lomax, Schiller & Walberg, 1979; Quack, 1979; Schneider & Bös, 1985) bestätigen die überlegene Bedeutung kognitiver Faktoren. Als Prädiktoren erklären sie, berechnet mit Korrelations- und Regressionsanalysen, bis zu 50% der Leistungsvarianz.

Genauere Analysen der Varianzaufklärung kommen jedoch zu einem differenzierteren Bild. Quack (1979) hat in einer empirischen Studie kognitive und nicht-kognitive Faktoren gleichzeitig berücksichtigt und für beide Gruppen sowohl die spezifischen als auch die konfundierten, d. h. in ihrem Einfluß nicht trennbaren Anteile berechnet. Die Ergebnisse ergaben, daß ungefähr 25% bis 30% der beobachtbaren Schulleistungsvarianz ausschließlich mit kognitiven Faktoren erklärt werden kann. Weitere 25% sind mit nichtkognitiven Faktoren konfundiert. Es ergeben sich somit zwei Grenzwerte für den durch Intelligenz aufklärbaren Anteil der Schulleistungsvarianz: „Ein *unterer Grenzwert* für den ‚reinen‘ Anteil von 25 bis 30% — d. h. Korrelationen um .50 — und ein *oberer Grenzwert* von 50% — d. h. Korrelationen um .70 — für den mit nichtkognitiven Variablen unkontrolliert konfundierten Anteil“ (Quack, 1979, S. 112).

Schneider und Bös (1985) kommen mit anderen Daten und einer anderen Technik der Varianzaufklärung in der Tendenz zu einem ähnlichen Ergebnis. Mittels kausalanalytischer Methoden konnten sie zeigen, daß der Einfluß motivationaler Faktoren in der Vergangenheit oft unterschätzt wurde, da diese Faktoren sich zu einem erheblichen Teil nur indirekt, vermittelt über andere Variablen, auf die Schulleistung auswirken (s. a. Meece, Wigfield & Eccles, 1990; Reynolds & Walberg, 1991).

Die letztgenannten Untersuchungen machen deutlich, daß die Suche nach motivationalen Prädiktoren der Schulleistung nicht von vorneherein, wegen des übermäßigen Effektes kognitiver Variablen, vergeblich ist. Dies hat uns ermutigt, in der vorliegenden Arbeit alle diejenigen Studien zusammenzutragen, die sich empirisch mit dem Zusammenhang von *Interesse* und Schulleistung beschäftigt haben. Interesse besitzt aus unserer Sicht eine herausragende Bedeutung innerhalb der Gruppe motivationaler Schülermerkmale. Dafür lassen sich verschiedene Gründe anführen. Erstens wurde in der Vergangenheit wiederholt auf theoretischer Ebene die Behauptung geäußert, daß sich die Stärke fachlicher Interessen signifikant auf die Schulleistung auswirkt (vgl. Todt, 1978; Krapp, 1992b) und gerade auch im Bereich der Hochbegabung von Bedeutung ist (z. B. Gagné, 1985; Klauer, 1992; Renzulli, 1988). Zweitens spielt Interesse in den subjektiven Theorien von Lehrern eine wichtige Rolle, wenn es darum geht, Schülerverhalten und -leistungen zu erklären (Krapp, 1989). Drittens ist Interesse von namhaften Pädagogen der

Vergangenheit und Gegenwart (Herbart, 1806/1965; Kerschensteiner, 1922; Dewey, 1913; Rubinstein, 1958; H. Schiefele, 1978; H. Schiefele, Haußer & Schneider, 1979; Travers, 1978) wiederholt als ein wesentliches Ziel schulischen Unterrichts genannt und begründet worden. Viertens ist Interesse u. E. ein zentrales Konzept wenn es um die Erklärung des Auftretens intrinsischer Motivation geht (Deci, 1992). Dieser Zusammenhang gewinnt seine Bedeutsamkeit wiederum aus Befunden, die den Einfluß intrinsischer Motivation auf die Qualität des Lernens aufzeigen (z. B. Deci & Ryan, 1985; s. Überblick bei Schiefele & Schreyer, in Druck).

Im folgenden geben wir zunächst anhand älterer, deskriptiver Übersichtsarbeiten einen Überblick über Ergebnisse der Forschung zur Interesse-Leistungs-Relation vor 1965. Daran anschließend erfolgt die Darstellung einer Metaanalyse der Untersuchungen, die zwischen 1965 und 1990 verfaßt worden sind. Dieser Zeitraum wurde deshalb so gewählt, weil die Forschung bis etwa 1965 recht gut in verschiedenen Reviews dokumentiert wurde. Ziel der Metaanalyse ist die Schätzung der Stärke des korrelativen Zusammenhangs von Interesse und Schulleistung. Da wir davon ausgehen, daß die Höhe der Interesse-Leistungs-Korrelation nicht für alle Schülergruppen und Fächer gleich hoch ausfällt, besteht ein weiteres Ziel in der Identifizierung bedeutsamer Moderatorvariablen.

Während ältere Untersuchungen der nichtkognitiven Bedingungen des Schulerfolgs häufig auch die Wirkung von Interessen einbezogen haben, ist diese Variablengruppe später von der Forschung zunehmend vernachlässigt worden. Dies schlägt sich auch in zusammenfassenden Darstellungen einschlägiger Forschungsarbeiten nieder: In früheren Übersichtsarbeiten wird regelmäßig auch über den Einfluß von Interessen referiert (z. B. Lavin, 1965; Super, 1960). Neuere Darstellungen (z. B. Fleming & Malone, 1983; Steinkamp & Maehr, 1983; Uguroglu & Walberg, 1979; Willson, 1983) nehmen dagegen keinen Bezug auf das Interessenkonzept. Im folgenden referieren wir kurz die wichtigsten Ergebnisse der älteren Reviews, die den Forschungsstand bis ca. 1965 zusammengefaßt haben. Vorauszuschicken ist, daß sich die Autoren dieser Reviews, von geringfügigen Überschneidungen abgesehen, auf jeweils andere Studien beziehen.

Fishman und Pasanella (1960) sichtetten insgesamt 580 Studien aus dem Publikationszeitraum 1949 bis 1959, die sich mit dem Zusammenhang zwischen kognitiven bzw. nichtkognitiven Prädiktoren und der Schulleistung (Durchschnittsnoten) befaßten. Davon untersuchten lediglich sieben Studien die Interesse-Leistungs-Beziehung. Die Korrelationen liegen zwischen .05 und .26.

Der Beitrag von Super (1960) stützt sich auf Studien, die in den USA vor 1957 publiziert wurden. Die Korrelationen zwischen Interessenwerten (in der Regel mit Berufsinteressentests gemessen) und Schulleistungsindikatoren erreichen selten höhere Werte als .30. Höhere Korrelationen wurden lediglich bei Stichproben gefunden, die sich entweder durch ein homogenes Fähigkeitsniveau oder hohe

Varianz der Interessenwerte auszeichneten. Auch Merkmale des Schulkontexts scheinen die Höhe der Interesse-Leistungs-Korrelation zu beeinflussen. In einigen Studien wurde festgestellt, daß das Vorherrschen einer wettbewerbsorientierten Atmosphäre den Einfluß von Interesse reduziert bzw. behindert.

Lavin (1965) kam zu einer ähnlichen Gesamteinschätzung wie Super (1960). Die von ihm referierten Korrelationen zwischen Interesse und Schulnoten gingen nicht über den Wert .30 hinaus. Für den Bereich des Studiums, der in seinem Review enthalten war, begründete Lavin dies damit, daß bei den Studenten einer bestimmten Fachrichtung das Interessenniveau relativ homogen ist, da in den meisten Fällen eine freiwillig getroffene Entscheidung für das Fach vorausgegangen ist. Interesse wies jedoch hohe Korrelationen mit Leistungsdaten auf (bis zu .70), wenn beide Variablen in bezug auf eine bestimmte Lehrveranstaltung innerhalb eines Studienfachs erfaßt worden waren. Lavin betont, daß nur wenige der von ihm einbezogenen Studien geschlechtsspezifische Unterschiede analysierten oder Fähigkeitsunterschiede berücksichtigten. Diejenigen Studien, die das Fähigkeitsniveau der untersuchten Personen einbezogen, konnten jedoch zeigen, daß Interesse auch dann mit Leistung signifikant korreliert, wenn Fähigkeitsunterschiede konstant gehalten werden.

Die berichteten Ergebnisse bestätigen die Vermutung, daß fachliche Interessen für die Vorhersage und Erklärung von Leistungsunterschieden in Schule und Studium einen signifikanten Beitrag liefern. Allerdings wird der Einfluß von Interesse auf die Leistung gelegentlich verdeckt oder deshalb nicht richtig eingeschätzt, weil in den untersuchten Stichproben die Varianz der Interessenwerte stark reduziert ist, oder die Verwendung unspezifischer (kumulativer) Leistungskriterien den eigentlichen Effekt des Interesses nicht zum Vorschein kommen läßt. Darüber hinaus ist mit Moderatoreffekten zu rechnen, die allerdings in den älteren Erhebungen nicht systematisch untersucht wurden.

Methode

Hinweise zur Metaanalyse

Wir wenden uns nun der Metaanalyse der Studien zur Interesse-Leistungs-Relation zu, die im Zeitraum 1965—1990 erschienen sind. Die folgenden Fragen sollten mit unserer Analyse beantwortet werden: Wie groß ist die Korrelation zwischen Interesse und Leistung im allgemeinen? Gibt es Unterschiede zwischen den verschiedenen Fächern? Wird der Einfluß des Interesses auf die Leistung über die Schuljahre hinweg stärker oder schwächer? Lassen sich geschlechtsspezifische Unterschiede beobachten?

Um eine möglichst objektive Zusammenfassung der einschlägigen Studien zu erhalten, verwenden wir die Methode der Metaanalyse. Von den mittlerweile relativ zahlreich vorhandenen Monographien zur Metaanalyse (z. B. Fricke & Treinies, 1985; Glass, 1976; Glass, McGaw & Smith, 1981; Hunter & Schmidt, 1990; Hunter, Schmidt & Jackson, 1982; Kulik & Kulik, 1989; Schwarzer, 1989)

stützten wir uns vor allem auf die Arbeiten von Hunter und Mitarbeitern (Hunter & Schmidt, 1990; Hunter et al., 1982). Diese Autoren beschäftigten sich insbesondere mit der Integration von Korrelationsergebnissen.

Auswahl der Studien

Eines der größten Probleme bei der Zusammenfassung interessenbezogener Studien besteht in dem äußerst eklektizistischen Gebrauch des Interessenkonzeptes. Der Begriff Interesse wird oft synonym gebraucht mit Begriffen wie z. B. „intrinsische Motivation“, „fachbezogener Affekt“, „Einstellung“ oder „kognitive Motivation“. So wird in einigen Studien, die sich angeblich mit Interesse beschäftigen, eigentlich etwas anderes gemessen. Andererseits sprechen etliche Studien, die tatsächlich Interesse untersucht haben, stattdessen von Einstellung oder Neugier. Durch diese Begriffsverwirrung ergeben sich erhebliche Probleme bei der Auswahl relevanter Studien.

Unser Verständnis von Interesse folgt der Konzeptualisierung von H. Schiefele und Kollegen (vgl. Prenzel, Krapp & H. Schiefele, 1986; H. Schiefele, 1978), die Interesse als *inhaltsspezifisches* Motivationsmerkmal beschreiben, das sich durch eine emotionale und eine wertbezogene Komponente auszeichnet (s. a. Krapp, 1992a; Schiefele, 1990, 1991b). Ein weiterer zentraler Gesichtspunkt ist die Selbstintentionalität. Sie besagt, daß die emotionale und wertbezogene Beziehung zu einem Gegenstand auf Gründe zurückzuführen ist, die vornehmlich im Gegenstand selbst bzw. in den auf ihn bezogenen Handlungen liegen (z. B. Erleben positiver Gefühle). Von Interesse kann keine Rede sein, wenn die Beschäftigung mit einem Gegenstand aus Gründen erfolgt, die mit dem Gegenstand selbst i. e. S. nichts zu tun haben (z. B. materielle Belohnung). Zum Zwecke des vorliegenden Überblicks wurden zwei Definitionskriterien von Interesse festgelegt. Es handelt sich um Minimalkriterien für die Bestimmung von Interesse als eigenständiges theoretisches Konstrukt: (1) Angabe einer Präferenzeinschätzung und (2) Inhalts- bzw. Gegenstandsspezifität dieser Präferenzeinschätzung. Studien, die Interesse nicht fachspezifisch verwendeten, wurden von vornherein ausgeschlossen. Typische Beispiele sind die Studien von Khan (1969) und Lloyd und Barenblatt (1984). Khan untersuchte den Prädiktor „schulisches Interesse“, das er jedoch als generelle Einstellung gegenüber schulbezogenen Arbeiten und Unterrichtsmethoden definierte. In ähnlicher Weise bestimmten Lloyd und Barenblatt (1984) das Konstrukt „intrinsische intellektuelle Motivation“. Es umfaßt die emotionalen Reaktionen einer Person auf die Inhalte und Prozesse schulischen Lernens, wobei nicht nach Fächern differenziert wird. Beide Studien erfüllen somit nicht das Kriterium der Bereichs- bzw. Inhaltsspezifität.

Die abhängige Variable, nämlich schulische Leistung, stellte dagegen ein deutlich geringeres Problem bei der Auswahl der Studien dar. So gut wie alle einschlägigen Forschungsarbeiten operationalisierten Schulleistung mittels Noten oder standardisierten Wissenstests. Studien, die extrem spezifische Leistungsindikatoren verwendeten, z. B. Lösung konkreter Aufgaben oder Verstehen eines Lehrbuchtextes, wurden nicht berücksichtigt. Die Befunde aus Untersuchungen über die Beziehung zwischen spezifischen thematischen Interessen und Indikatoren des Textverstehens sind an anderer Stelle zusammengefaßt (vgl. Hidi, 1990; Schiefele, 1988; Wade, 1992).

Die Suche nach relevanten Studien wurde unter Zuhilfenahme der Datenbasen PSYCINFO (anglo-amerikanische Literatur ab 1967), PSYINDEX (deutschsprachige Literatur ab 1977) und DISSERTATION ABSTRACTS (amerikanische Dissertationen ab 1970) durchgeführt. Für die Suche wurde eine Vielzahl von Stichworten (u. a. Interesse, Schulleistung, Motivation, Einstellung, Affekt) benutzt, da sich gelegentlich auch hinter weniger spezifischen Titeln wichtige Studien verbergen können. Außerdem wurden zahlreiche einschlägige Zeitschriften (z. B. Journal of Educational Psychology, Review of Educational Research) durchgesehen.

Insgesamt konnten 15 Publikationen und 6 Dissertationen identifiziert werden, die sich mit Interesse und Leistung beschäftigen und unseren Kriterien entsprechen. Diese 21 Studien berichten Befunde von insgesamt 127 unabhängigen Stichproben aus 19 verschiedenen Ländern. Mit Ausnahme

von vier Stichproben setzen sich alle sowohl aus Mädchen als auch Jungen zusammen. Die Größe der Stichproben variiert von 49 bis 18,215 und die Klassenstufen reichen von der 5. bis zur 12. Klasse. Neun verschiedene Fachgebiete wurden in den Studien miteinbezogen. Insgesamt finden sich in den berücksichtigten Arbeiten 211 verwertbare Korrelationen zwischen Interesse und Leistung. In dieser Zahl sind die für Mädchen und Jungen getrennt berechneten Einzelkorrelationen nicht enthalten. Die einbezogenen Studien sind in Tabelle 1 in der Reihenfolge ihres Erscheinens aufgelistet. Für jede Studie wurden die einbezogenen Fächer, die Klassenstufe, die Größe der Stichprobe und die berichteten Korrelationen angeführt. Einige Studien berichten auch getrennte Korrelationen für Jungen und Mädchen. In diesen Fällen haben wir nur die geschlechtsspezifischen Korrelationen, nicht jedoch die Gesamtkorrelationen (über alle Versuchspersonen hinweg) in die Tabelle 1 aufgenommen.

Es ist bemerkenswert, daß 108 der insgesamt 211 Korrelationen aus Studien eines internationalen Projektes stammen, das von der „International Association for the Evaluation of Educational Achievement“ (IEA, z. B. Husén, 1967) initiiert wurde. Das IEA-Projekt wurde in 21 Ländern durchgeführt und umfaßte insgesamt sechs Schulfächer (Mathematik, Naturwissenschaft, Englisch oder Französisch als Fremdsprache, Literatur, Lesen, und "civic education"). In den meisten Ländern wurden zwei Gruppen untersucht: 13- und 14-jährige Schüler, die normalerweise die 8. Klasse besuchen, und 18-jährige Schüler, die im Regelfall die 12. Klasse besuchten. Da sich die 13- von den 14-jährigen Schülern kaum unterschieden, haben wir nur die letzteren in die Metaanalyse einbezogen.

Statistische Verfahrensweisen

Kodierung. Als erster Schritt der Metaanalyse wurden alle relevanten Informationen aus den Studien kodiert: Höhe der Korrelationskoeffizienten, Stichprobengröße, Zusammensetzung der Stichprobe nach Geschlecht, Fachgebiet, Klassenstufe, Erscheinungsjahr der Studie, Nationalität der Stichprobe, Art des Leistungsmaßes (Noten vs. Leistungstests), Art des Interessenmaßes (Fragebogen vs. Test), Reliabilität von Interesse- und Leistungsmaßen, Zeitspanne zwischen Interessen- und Leistungsmaß, Herkunft der Studie (IEA vs. Nicht-IEA), Art der Veröffentlichung (Dissertation vs. reguläre Publikation).

Bestimmung möglicher Moderatoreffekte. Da es ein wichtiges Ziel dieser Metaanalyse ist, die Effekte von Moderatorvariablen (Geschlecht, Fachgebiet, Klassenstufe) zu untersuchen, mußte zuerst festgestellt werden, ob die Varianz zwischen den Korrelationen auf den Stichprobenfehler zurückzuführen ist oder aber eine tatsächliche Variation zwischen den Populationswerten widerspiegelt. Gesetzt den Fall, daß ein großer Teil der beobachteten Varianz nicht dem Stichprobenfehler zugeschrieben werden kann, dann sind die Korrelationen als heterogen anzusehen und die Existenz von relevanten Moderatorvariablen ist sehr wahrscheinlich („Modell der heterogenen Effekte“). Wird die beobachtete Varianz dagegen in erster Linie durch den Stichprobenfehler verursacht, dann ist von dem „Modell der homogenen Effekte“ auszugehen. In diesem Fall wäre eine Suche nach Moderatorvariablen nicht gerechtfertigt.

In der Literatur werden drei verschiedene Indikatoren der Heterogenität diskutiert (Schwarzer, 1989): (1) Hunter et al. (1982) haben vorgeschlagen, daß der Teil der beobachteten Varianz, der durch den Stichprobenfehler erklärt wird, kleiner als 75% sein sollte. (2) Zusätzlich schlugen die selben Autoren vor (s. a. Hunter & Schmidt, 1990), einen χ^2 -Test durchzuführen, in dem die beobachtete Varianz für eine Gruppe von Korrelationen zum Mittelwert dieser Korrelationen in Beziehung gesetzt wird. (3) Schließlich sollte die Populations- bzw. Residualstandardabweichung (die Wurzel aus der Differenz zwischen beobachteter Varianz und der Varianz des Stichprobenfehlers) größer sein als ein Viertel der Populationskorrelation (McDaniel, Hirsh, Schmidt, Raju & Hunter, 1986).

Aggregation und Gewichtung der Korrelationen. Nach Hunter et al. (1982) ergibt sich die beste Schätzung für eine Populationskorrelation, wenn man den Mittelwert der individuellen Korrelationen mit den Stichprobengrößen der entsprechenden Studien gewichtet. Umstritten ist bei diesem

Tab. 1. Die in die Metaanalyse einbezogenen Studien

Autor	Fach	Stichprobe Klasse Größe	Korrelationen
Skager et al. (1965)	Mathematik	10/11 524	m: .35, w: .31
	Physik	10/11 524	m: .32, w: .25
	Biologie	10/11 524	m: .20, w: .21
	Sozialkunde	10/11 524	m: .35, w: .22
	Fremdsprachen	10/11 524	m: .31, w: .16
	Kunst/Musik	10/11 524	m: .22, w: .38
Husén (1967/IEA)	Mathematik	(Stichproben aus 11 Ländern: 8 $n > 841$ $n < 6544$)	.23, .24, .26, .27, .28, .29, .32, .35, .38, .39, .42
		(Stichproben aus 9 ^a Ländern: 12 $n > 369$ $n < 4372$)	.16, .29, .29, .29, .30, .32, .33, .37, .39, .40, .43, .47, .51, .52
Todt (1967)	Mathematik	12 208 (m)	.37
	Physik	12 208 (m)	.28
	Sozialkunde	12 208 (m)	.26
Wendeler (1968)	Mathematik	10 227	.38
	Physik	10 227	.45
	Biologie	10 227	.27
	Chemie	10 227	.43
	Sozialkunde	10 227	.36
	Literatur	10 227	.32
Carlson (1970/Diss)	Mathematik	9 70	.26
Lowe (1970/Diss)	Sozialkunde	12 187	.22
Comber und Keeves (1973/IEA)	Naturwissen- schaft	(Stichproben aus 15 Ländern: 8 $n > 697$ $n < 7363$)	.09, .18, .23, .26, .26, .27, .32, .35, .35, .36, .37, .38, .39, .42, .49
		(Stichproben aus 14 Ländern: 12 $n > 491$ $n < 15719$)	.21, .22, .27, .33, .40, .43, .48, .50, .51, .53, .58, .60, .64, .67

Fortsetzung Tabelle 1

Autor	Fach	Stichprobe Klasse Größe	Korrelationen
Purves (1973/IEA)	Literatur	(Stichproben aus 9 Ländern:)	.10, .12, .13, .16, .17, .17,
		8 $n > 548$ $n < 7228$.18, .22, .22
		12 $n > 464$ $n < 14204$.12, .14, .15, .16, .17, .17, .21, .24, .34
Carroll (1975/IEA)	Fremdsprachen	(Stichproben aus 6 Ländern:)	.22, .24, .26, .38, .42, .47
		8 $n > 839$ $n < 4420$	
		(Stichproben aus 7 Ländern:)	.26, .28, .33, .35, .39, .41, .43
		12 $n > 378$ $n < 3230$	
Hall (1975)	Sozialkunde	12 159 (m)	.29
Harty und Beall (1975)	Naturwissen- schaft	5 95	.30
Lewis und Massad (1975/IEA)	Fremdsprachen	(Stichproben aus 7 Ländern:)	.22, .28, .28, .38, .48, .49, .49
		8 $n > 687$ $n < 2331$	
		(Stichproben aus 9 Ländern:)	.18, .26, .27, .27, .27, .27, .29, .31, .33
		12 $n > 323$ $n < 2310$	
Eisenhardt (1976/Diss)	Mathematik	6—9 1: 17808	m: .19, w: .18
		2: 18092	m: .17, w: .16
		9—11 1: 17580	m: .24, w: .19
		2: 18215	m: .25, w: .20
	Naturwissen- schaft	6—9 1: 17808	m: .15, w: .08
		2: 18092	m: .17, w: .16
		9—11 1: 17580	m: .27, w: .17
		2: 18215	m: .28, w: .18
	Sozialkunde	6—9 1: 17808	m: .16, w: .12
		2: 18092	m: .17, w: .11
		9—11 1: 17580	m: .25, w: .20
		2: 18215	m: .26, w: .19

Fortsetzung Tabelle 1

Autor	Fach	Stichprobe Klasse	Größe	Korrelationen
Eisenhardt	Literatur	6—9	1: 17808	m: .01, w: —.01
			2: 18092	m: .01, w: .00
		9—11	1: 17580	m: .10, w: .09
			2: 18215	m: .11, w: .09
Oehlert (1977/Diss)	Physik	5/6	100	.52
Kelly (1978/IEA)	Naturwissen- schaft	(Stichprobe aus 7 Ländern:)		m: .22, .31, .37, .38, .41, .46, .52,
		8	$n > 932$ $n < 3823$	w: .18, .23, .24, .25, .30, .34, .36
Todt (1978)	Mathematik	10/11	220	m: .50, w: .47
		10—12	158	m: .53, w: .40
		11/12	158 (w)	.45
	Physik	8/9	234	m: .24, w: .39
		8—10	526	m: .22, w: .11
		10/11	220	m: .17, w: .33
		10—12	146	m: .43, w: .10
		11/12	158 (w)	.39
	Biologie	8/9	234	m: .08, w: .07
		8—10	526	m: .13, w: .12
		10/11	220	m: .16, w: .11
		10—12	125	m: .22, w: —.08
		11/12	158 (w)	.30
	Sozialkunde	10/11	220	m: .41, w: .51
		10—12	154	m: .31, w: .20
		11/12	158 (w)	.32
Sjöberg (1983)	Mathemaik	12	174	.53
	Physik	12	174	.48
	Sozialkunde	12	174	.41
	Fremdsprachen	12	174	.41
Sjöberg (1984)	Mathematik	12	100	.44
	Physik	12	100	.49
	Sozialkunde	12	100	.26
	Fremdsprachen	12	100	.30
Napier und Riley (1985)	Naturwissen- schaft	11	3135	.22

Fortsetzung Tabelle 1

Autor	Fach	Stichprobe Klasse Größe		Korrelationen
Goldstein (1986/Diss)	Naturwissen- schaft	5/6	217	m: .21, w: .03
Huetinck (1990/Diss)	Naturwissen- schaft	10	120	m: .31, w: .10

Anmerkungen. m = männlich, w = weiblich.

^a Die Zahl der Korrelationen ist hier größer als die Zahl der Länder, da in einigen Ländern zwei Stichproben aus derselben Population untersucht wurden.

Verfahren lediglich, ob die Korrelationen vor ihrer Aggregation in Fishers z-Werte umgewandelt werden sollen (s. z. B. Fricke & Treinies, 1985, S. 98). Wir haben uns der Argumentation von Hedges und Olkin (1985) angeschlossen und eine z-Transformation durchgeführt.

Ein Nachteil der Gewichtung von Korrelationskoeffizienten besteht darin, daß diejenigen Korrelationen, die auf extrem großen Stichproben beruhen, viel Einfluß erhalten. Im vorliegenden Fall betrifft dies die meisten Stichproben der IEA-Studien (vgl. Tabelle 1). Um den übermäßigen Einfluß dieser Studien auf die Ergebnisse der Metaanalyse zu kontrollieren, führten wir sowohl Berechnungen mit gewichteten als auch ungewichteten Korrelationen durch. Da sich dabei nur sehr geringe Unterschiede ergaben, werden im folgenden nur Ergebnisse berichtet, die auf gewichteten Durchschnittswerten beruhen.

Die meisten Studien berechneten auf der Grundlage einer einzigen Stichprobe mehrere, und somit voneinander abhängige, Korrelationskoeffizienten. In diesen Fällen bildeten wir gewichtete Mittelwerte. Als Gewichte wurden wieder die jeweiligen Stichprobengrößen verwendet. Für alle im Rahmen der Metaanalyse durchgeführten Berechnungen wurden voneinander abhängige Korrelationen aggregiert. Würde man diese Vorgehen unterlassen, käme es zu einer ungebührlichen Vergrößerung des Einflusses des Stichprobenfehlers.

Ergebnisse

Der generelle Zusammenhang von Interesse und Leistung

Die Verteilung der 127 unabhängigen Korrelationskoeffizienten ist in Tabelle 2 als sogenanntes "stem and leaf"-Diagramm dargestellt (Tukey, 1977). Jede Ziffer auf der linken Seite der vertikalen Linie ist Teil des „Stammes“ und entspricht jeweils der ersten Ziffer einer Korrelation. Die Zahlen auf der rechten Seite der vertikalen Linie („Blätter“) entsprechen jeweils der zweiten Ziffer einer Korrelation. Aus Tabelle 2 ist z. B. abzulesen, daß sich vier Korrelationen größer oder gleich .60 ergaben (nämlich .60, .64, .66, .67) und insgesamt acht Korrelationen den Wert .26 erreichten.

Tab. 2. Unabhängige Korrelationen
zwischen Interesse und Leistung

.6	7
.6	04
.5	8
.5	011223
.4	67788999
.4	0012223334
.3	55556677778888899999
.3	000012222333334
.2	666666667777777888899999
.2	00011122222233444
.1	55666677778888
.1	0223334
.0	9

Anmerkung. k (Zahl der unabhängigen Korrelationen) = 127.

Die Werte in Tabelle 2 ergeben annähernd eine Normalverteilung. Es ist bemerkenswert, daß ausnahmslos positive Korrelationen berichtet wurden. Die Werte reichen von .09 bis .67. Der Mittelwert der Korrelationen liegt bei .30 mit einer Standardabweichung von .134 (SD_r). Der Stichprobenfehler erklärt nur 1.9% der beobachteten Varianz. Die Populationsstandardabweichung (SD_p) beträgt .132. Benutzt man diesen Wert, um das 95%-Konfidenzintervall für die Populationskorrelation zu berechnen, ergeben sich als Grenzwerte .04 und .56.¹

Um eine präzisere Schätzung der durchschnittlichen Populationskorrelation zu bekommen, schlagen Hunter et al. (1982) vor, eine Minderungskorrektur vorzunehmen. Diese sieht vor, daß jeder Korrelationskoeffizient mit den Reliabilitäten des jeweiligen Interessen- und Leistungsmaßes gewichtet wird. Die aufgrund der korrigierten Einzelwerte berechnete Durchschnittskorrelation für den Zusammen-

¹ Die Signifikanz gemittelter Korrelationen kann nicht auf dieselbe Weise bestimmt werden wie die einzelner Korrelationen. Das Problem besteht darin, die korrekte Stichprobengröße festzulegen. Basiert der Signifikanztest auf einem n , das die Summe aller einbezogenen Stichproben darstellt, so wird der Fehler 1. Art vergrößert. Legt man dagegen die Zahl der Korrelationen als Stichprobengröße zugrunde, riskiert man einen Fehler 2. Art. Es wird deshalb von verschiedenen Autoren (z. B. Schwarzer, 1989) vorgeschlagen, 95%-Konfidenzintervalle zu berechnen. Darüber hinaus kann die Faustregel angewendet werden, daß die Populationskorrelation mindestens doppelt so hoch wie die Populationsstandardabweichung sein sollte. In der vorliegenden Studie wurden Signifikanzen allein auf der Basis von Konfidenzintervallen bestimmt.

hang zwischen Interesse und Leistung in der Gesamtstichprobe erreichte den Wert .41 ($SD_r = .172$). In den Fällen, wo keine Reliabilitäten angegeben waren, wurde der Mittelwert der verfügbaren Reliabilitätskoeffizienten als Wert eingesetzt. Die Minderungskorrektur wurde nur im Fall der Gesamtkorrelation angewendet. Da eine fehlerfreie Messung von Variablen wie Interesse und Leistung nicht möglich ist, führt die Minderungskorrektur zu unrealistischen Ergebnissen. Außerdem ist die Vergleichbarkeit mit anderen Forschungsergebnissen größer, wenn unkorrigierte Koeffizienten verwendet werden (Kulik & Kulik, 1989).

Die Analyse der Daten in der Gesamtstichprobe läßt vermuten, daß die Höhe der Korrelationen in Subgruppen systematisch variiert und mit Moderatoreffekten zu rechnen ist. Die oben erwähnten Indikatoren einer auf Moderatoreffekte hinweisenden Heterogenität sind alle erfüllt: die Populationsstandardabweichung ist größer als ein Viertel der Populationskorrelation, nur 1.9 % der beobachteten Varianz werden durch den Stichprobenfehler erklärt und der Chi^2 -Test ist hochsignifikant ($p < .001$).

Der Einfluß untersuchungstechnischer Merkmale

Neben der Überprüfung einiger theoretisch bedeutsamer Moderatoreffekte (s. u.), wurde der Einfluß folgender untersuchungstechnischer Merkmale überprüft: Erscheinungsjahr der Studie, Art der Interessen- und Leistungsmaße, Größe der Stichprobe, Herkunft der Studie (IEA vs. Nicht-IEA) und Art der Veröffentlichung (Dissertation vs. reguläre Publikation). Die Beurteilung des Einflusses dieser Merkmale ist nicht unproblematisch, da es sich in der Mehrzahl um IEA-Studien handelt und die genannten Merkmale daher mehr oder weniger miteinander konfundiert sind. Vergleichen wir deshalb zuerst IEA- und Nicht-IEA-Studien miteinander. Die mittlere Interesse-Leistungs-Korrelation für IEA-Studien beträgt .32 ($SD_p = .132$, $k = 108$), die für Nicht-IEA-Studien .184 ($SD_p = .051$, $k = 19$). Dieser Unterschied ist signifikant.² Läßt man jedoch die Dissertationen bei den Nicht-IEA-Studien weg, so erhöht sich die mittlere Korrelation auf .25 ($SD_p = .058$, $k = 10$) und der signifikante Unterschied zwischen IEA- und Nicht-IEA-Studien verschwindet wieder. Vergleicht man die letztgenannte Gruppe von Studien ($r_M = .25$) mit den Dissertationen ($r_M = .176$, $SD_p = .046$, $k = 9$), so ergibt sich wiederum ein signifikanter Unterschied.

² Das Problem des Vergleichs zweier Durchschnittskorrelationen ist ähnlich dem der Testung einzelner Werte (s. Fußnote 1). Es gibt jedoch für diesen Fall keine allgemein akzeptierten Faustregeln. Wir folgten jedoch einem Vorschlag von R. Schwarzer (Persönliche Mitteilung, 23. 6. 1990), wonach der Unterschied zwischen zwei gemittelten Korrelationen größer als die mittlere Populationsstandardabweichung sein sollte, um als signifikant zu gelten. Alle folgenden Signifikanzangaben zu Korrelationsunterschieden beruhen auf dieser Faustregel.

Es konnten keine signifikanten Unterschiede zwischen älteren (vor 1974 erschienen; $r_M = .31$, $SD_p = .144$, $k = 76$) und neueren Studien (nach 1974 erschienen; $r_M = .28$, $SD_p = .108$, $k = 51$) und zwischen Studien mit großen ($r_M = .28$, $SD_p = .121$, $k = 59$) oder kleinen Stichproben ($r_M = .36$, $SD_p = .153$, $k = 68$) festgestellt werden. Es wurde allerdings eine signifikante Korrelation zwischen Stichprobengröße und Höhe der Korrelationen gefunden ($r = -.22$, $k = 127$, $p < .01$). Erwartungsgemäß ergaben sich also bei großen Stichproben kleinere Korrelationen.

Ohne signifikanten Einfluß auf die Höhe der Interesse-Leistungs-Korrelationen blieben die Art des Interessenmaßes (standardisierter Test: $r_M = .22$, $SD_p = .047$, $k = 5$; Fragebogen: $r_M = .30$, $SD_p = .132$, $k = 122$) und des Leistungsmaßes (Noten: $r_M = .28$, $SD_p = .092$, $k = 9$; Tests: $r_M = .30$, $SD_p = .133$, $k = 117$).

Schließlich untersuchten wir den Einfluß der Zeitspanne zwischen Interessen- und Leistungsmessung. In den meisten Fällen fand sich entweder keine Angabe zur Zeitspanne oder die Prädiktor- und Kriteriumsmaße wurden zum gleichen Zeitpunkt erhoben. Bei einigen der Studien, die Noten als Leistungskriterium verwendeten, ist den Angaben zu entnehmen, daß Interesse erst *nach* der Notenvergabe erfaßt wurde. Nur in 2 der insgesamt 21 Studien konnte eindeutig festgestellt werden, daß Interesse in größerem zeitlichen Abstand *vor* der Leistungsmessung erhoben worden war. Bei diesen beiden Studien handelt es sich um die Dissertationen von Eisenhardt (1976) und Goldstein (1986). In der Studie von Eisenhardt wurde Interesse 24 (Stichprobe 2) bzw. 30 Monate (Stichprobe 1) vor der Erfassung des Leistungskriteriums erhoben. Bei Goldstein beträgt der zeitliche Abstand drei Monate. Die Durchschnittskorrelation dieser beiden Studien beträgt .17 und liegt somit niedriger als die Gesamtkorrelation. Aufgrund der geringen Zahl echter prognostischer Studien kann natürlich keine aussagekräftige Schlußfolgerung gezogen werden. Als indirekter Beleg dafür, daß die Größe der Zeitspanne, sofern sie nicht eine bestimmte Grenze überschreitet, keinen besonderen Einfluß auf die Stärke der Beziehung zwischen Interesse und Leistung ausübt, kann ein Befund von Uguroglu und Walberg (1979) dienen. Diese Autoren führten eine Metaanalyse zur Beziehung zwischen motivationalen Prädiktoren und Leistungskriterien durch und fanden keinen Unterschied zwischen „gleichzeitigen“ und prognostischen Korrelationen.

Zusammenfassend zeigt die Analyse der untersuchungstechnischen Merkmale, daß in den Dissertationen niedrigere Interesse-Leistungs-Korrelationen berichtet wurden als in den anderen Studien. Im vorliegenden Fall ist dies jedoch hauptsächlich auf die Studie von Eisenhardt (1976) zurückzuführen. Die in dieser Studie gefundenen niedrigen Korrelationen sind vermutlich vor allem durch den großen zeitlichen Abstand zwischen Interesse- und Leistungsmessung bedingt.

Moderatoreffekt des Fachs

Hängt die Stärke der Interesse-Leistungs-Relation von der Art der Schulfächer ab? Die Prüfung dieses Moderatoreffekts führte zu dem in Tabelle 3 dargestellten Ergebnis.

Tab. 3. Durchschnittliche Interesse-Leistungs-Korrelationen
in den einzelnen Fächern

Fach	r_M	k	SD_p
Mathematik	.28	36	.090
Naturwissenschaft	.31	45	.150
Physik	.31	9	.117
Biologie	.16	6	.027
Sozialkunde	.20	11	.052
Fremdsprachen	.33	32	.099
Literatur	.14	23	.052

Anmerkungen. r_M = durchschnittliche, gewichtete Korrelation; k = Zahl der unabhängigen Korrelationen; SD_p = Populationsstandardabweichung.

Für Musik und Kunst konnten keine Durchschnittskorrelationen berechnet werden, da in diesen Fächern jeweils nur eine Korrelation berichtet wurde.

Die Durchschnittskorrelationen für Mathematik, Naturwissenschaft³, Physik und Fremdsprachen liegen alle um den Wert .30. Deutlich niedriger sind die Korrelationen in den Fächern Biologie, Sozialkunde und Literatur⁴ ausgefallen. Alle Werte unterscheiden sich auf dem 5%-Niveau signifikant von Null. Die Werte für Biologie, Sozialkunde und Literatur sind signifikant niedriger als die aller anderen Schulfächer. Ein signifikanter Unterschied ergab sich zudem zwischen den Korrelationsmittelwerten für Literatur und Sozialkunde.

Um die oben genannten Ergebnisse zu vereinfachen und allgemeinere Trends herauszuarbeiten, wurden zwei Gruppen von Schulfächern gebildet: Naturwissenschaften und Sozial- bzw. Humanwissenschaften. Die erste Gruppe besteht aus Mathematik, Naturwissenschaft, Physik und Biologie. Die zweite Gruppe enthält Sozialkunde, Fremdsprachen, Literatur und Kunst/Musik. Die Ergebnisse dieser

³ Naturwissenschaft ("science") als Schulfach ist vor allem in den USA verbreitet. Es umfaßt üblicherweise die Teilfächer Biologie, Physik, Geologie und Chemie.

⁴ Dem Fach Literatur entspricht in unserem Schulsystem das Fach Deutsch.

Analyse sind in Tabelle 4 dargestellt. Obwohl die Interesse-Leistungs-Korrelation für Naturwissenschaften höher ist als die für Sozialwissenschaften, überschreitet die Differenz zwischen den beiden Werten nicht die durchschnittliche Populationsstandardabweichung ($SD_p = .118$), und kann daher nicht als signifikant angesehen werden.

Moderatoreffekt der Klassenstufe und des Geschlechts

Barak (1981), Gottfredson (1981) und andere Autoren behaupten, daß die Beziehung zwischen Interesse und Leistung mit zunehmendem Alter stabiler wird. Es wäre daher zu erwarten, daß die Korrelationen zwischen Interesse und Leistung für ältere Schüler höher ausfallen als für jüngere Schüler.

Leider sind die in unserer Metaanalyse einbezogenen Korrelationen relativ unregelmäßig auf die Klassenstufen 5 bis 12 verteilt. In manchen Studien wurden zudem Schüler verschiedener Klassenstufen in einer Stichprobe zusammengefaßt. Deshalb schien es zweckmäßig zu sein, zwei Altersgruppen miteinander zu vergleichen, nämlich einerseits die Klassen 5 bis 10 und andererseits die Klassen 9 bis 12.

Erwartungsgemäß zeigt sich in der zweiten Gruppe eine höhere Durchschnittskorrelation (vgl. Tabelle 4). Der zu beobachtende Unterschied erwies sich jedoch nicht als signifikant.

Tab. 4. Durchschnittliche Interesse-Leistungs-Korrelation unterteilt nach Fachgebiet, Klassenstufe und Geschlecht

Moderatorvariable	r_M	k	SD_p
Fachgebiet:			
Naturwissenschaft	.32	79	.132
Sozialwissenschaft	.23	58	.105
Klassenstufe:			
5–10	.28	63	.115
9–12	.32	64	.149
Geschlecht:			
Jungen	.24	20	.111
Mädchen	.16	19	.078

Anmerkungen. r_M = durchschnittliche, gewichtete Korrelation; k = Zahl der unabhängigen Korrelationen; SD_p = Populationsstandardabweichung.

Der letzte Teil von Tabelle 4 zeigt die Höhe der Interesse-Leistungs-Korrelationen für männliche gegenüber weiblichen Schülern. Mit Ausnahme der Studie von Kelly (1978) ist hier keine weitere IEA-Studie miteinbezogen. Die Korrelationen fallen zwar bei Jungen höher aus als bei Mädchen, der Unterschied erreicht jedoch kein akzeptables Signifikanzniveau.

Aufschlußreich ist hier die Tatsache, daß die Einbeziehung der Dissertationen zu einem anderen Ergebnis führt, als eine Analyse von Publikationen in Zeitschriften und Sammelbänden. In den regulär publizierten Studien findet man für Jungen eine mittlere Korrelation von .35 ($SD_p = .102$, $k = 14$) und für Mädchen eine von .25 ($SD_p = .063$, $k = 13$). Dieser Unterschied ist signifikant.

Zusammenfassung der Befunde und Analyse von Wechselwirkungen

Zusammenfassend kann man feststellen, daß der Vorhersagewert von Interesse für die meisten Schulfächer, mit Ausnahme von Biologie, Sozialkunde und Literatur, gleichermaßen stark ist. Ein Einfluß des Alters konnte statistisch nicht abgesichert werden und der des Geschlechts erwies sich nur bei Ausschluß der Dissertationen als signifikant.

Es war nicht möglich, Wechselwirkungen der drei Moderatorvariablen zu untersuchen, da nicht genügend unabhängige Stichproben dafür vorhanden waren. Für jeweils zwei Moderatorvariablen konnten jedoch Wechselwirkungsanalysen durchgeführt werden. Die Ergebnisse dieser Analysen zeigten nur additive Wirkungen in der jeweils erwarteten Richtung. Die höchste Korrelation ergab sich für Schüler der Klassenstufen 9 bis 12 im Fach Naturwissenschaft ($r_M = .37$, $k = 19$, $p < .05$). Die niedrigste Korrelation ergab sich für Schüler der Klassenstufen 5 bis 10 im Fach Literatur ($r_M = .11$, $k = 11$, $p < .05$).

Diskussion

Die Stärke des Zusammenhangs von Interesse und Leistung in der Gesamtstichprobe

Die vorliegende Metaanalyse zeigt, daß die Interessenausprägung im Durchschnitt (ohne Einbeziehung von Moderatorvariablen) ungefähr 10% der beobachteten Varianz der Leistung erklärt. Dies steht in Übereinstimmung mit den Ergebnissen früherer Studien (z. B. Super, 1960). Auf den ersten Blick scheint dies ein geringer Wert zu sein. Doch wenn man bedenkt, daß Untersuchungen zur Vorhersage und Erklärung von Schulleistung selten höhere Aufklärungsquoten als 50% erreichen (Krapp, 1979), so ist dies ein durchaus beachtlicher Betrag — umso mehr, als in Subgruppen und speziellen Fachgebieten durchaus höhere Werte auftreten. Allerdings ist hier auch zu bedenken, daß fachliches Interesse mit anderen

hypothetischen Bedingungsvariablen konfundiert ist und deshalb die Schätzung von 10% aufgeklärter Varianz nicht als isolierter „Interesseneffekt“ interpretiert werden darf.

Es liegt nahe, die vorliegenden Befunde mit den empirischen Befunden über den Einfluß anderer motivationaler und emotionaler Faktoren zu vergleichen. Dazu gibt es bisher nur wenige zusammenfassende Darstellungen. Die Reviews von Bloom (1976), Steinkamp und Maehr (1983), Uguroglu und Walberg (1979), Rennie und Punch (1991) und Willson (1983) sind nur von eingeschränktem Wert, da sie die verschiedenen Kategorien emotionaler oder motivationaler Prädiktoren nicht deutlich genug unterschieden haben. So werden z. B. bei Uguroglu und Walberg (1979) so verschiedene Variablen wie schulisches Selbstkonzept und Leistungsmotivation summarisch der Kategorie „Motivation“ zugeordnet. Die ermittelte Durchschnittskorrelation beträgt in diesem Fall .34. Bloom (1976) fand einen ähnlichen Wert für die Beziehung zwischen fachbezogenen „affektiven Merkmalen“ und Schulleistung. Steinkamp und Maehr (1983) und Willson (1983) sichteteten die Forschung im Bereich Naturwissenschaften und fanden deutlich niedrigere Korrelationen zwischen affektiven Prädiktoren und Schulleistung ($r_M = .19$ bzw. $.16$). Willson (1983) führte getrennte Analysen für drei verschiedene Arten von Einstellungsvariablen durch (Interesse, Einstellung per se und beobachtetes Verhalten) und fand, daß Interesse mit Leistung höher korreliert als die beiden anderen Merkmale ($r_M = .27$ vs. $.14$ und $.05$). Insgesamt kann der Schluß gezogen werden, daß die Höhe der von uns ermittelten Interesse-Leistungs-Korrelation in etwa der Höhe der Durchschnittskorrelationen entspricht oder darüber liegt, die für andere Kategorien emotional-motivationaler Prädiktoren berichtet wurde.

Moderatoreffekt des Fachs

Der Einfluß von Interesse auf die Leistung hängt teilweise vom jeweiligen Gegenstand (Fachgebiet) des Interesses ab. Für die „harten“ naturwissenschaftlichen Fächer wurden in der Regel höhere Korrelationen berichtet als für die Fächer Biologie, Sozialkunde oder Literatur (s. a. Lehrke, 1988; Löwe, 1992). Dieser Sachverhalt wird z. T. dadurch erklärt, daß der höhere Schwierigkeitsgrad der erstgenannten Fächer bei Fachwahlentscheidungen nur solche Schüler anzieht, die sich ihrer fachbezogenen Fähigkeit relativ sicher sind (vgl. Lehrke, 1988). Es wäre jedoch auch denkbar, daß es in schwierigen Fächern generell stärker auf motivationale Bedingungen ankommt. Ein ausgeprägteres fachliches Interesse veranlaßt den Schüler, während der frühen Phasen des Lernprozesses nicht gleich aufzugeben. Leichtere Schulfächer erlauben dagegen auch dem weniger interessierten Schüler, gute Noten zu erzielen. Dieses Argument paßt recht gut zu den Schwierigkeitseinschätzungen der Fächer. Biologie-, Sozialkunde- und Literaturunterricht gelten als leicht. Dagegen werden Mathematik, Physik, Chemie und Fremdsprachen als vergleichsweise schwierig eingestuft (Löwe, 1992; Stodolsky, 1988).

Moderatoreffekt der Klassenstufe

Die Vermutung, daß die Bedeutung von Interessen für die Leistung mit zunehmendem Alter steigt, wird durch unsere Befunde nur in Form eines nicht signifikanten Trends bekräftigt. Die Resultate anderer Untersuchungen scheinen die Altersabhängigkeit des Interesseneffekts besser bestätigen zu können. Uguroglu und Walberg (1979) berichten z. B. über einen signifikanten Anstieg der Korrelationen zwischen Motivation und Leistung von der ersten ($r_M = .24$) bis zur 12. Klasse ($r_M = .34$).

In der Literatur findet man mindestens vier mögliche Erklärungen für diesen Trend (Evans, 1971; Gottfredson, 1981; Lehrke, 1988; Todt, 1978, 1985): (1) Über die Schuljahre hinweg findet eine Selektion derjenigen Schüler statt, die in der Lage sind, ihre Interessen weiterverfolgen zu können. (2) Interesse und Leistung beeinflussen sich im Laufe der Schulzeit gegenseitig und konvergieren deshalb zunehmend. (3) Die Stabilität von Interessen nimmt mit dem Alter zu. (4) Die Struktur der Interessen erfährt während der Pubertät eine grundlegende Veränderung in Richtung einer stärkeren Realitätsorientierung: Die Ziele und Interessen des Schülers werden an die eigenen Fähigkeiten und künftigen Berufsmöglichkeiten angepaßt und bleiben dann stabil.

Moderatoreffekt des Geschlechts

In der Vergangenheit wurden Geschlechtsdifferenzen besonders im Bereich der Naturwissenschaften untersucht. Die Ergebnisse dieser Forschungsrichtung stimmen mit unserem Befund überein, wonach die Interessen bei Jungen einen stärkeren Einfluß auf die Leistung haben als bei Mädchen. In vielen Studien wurde zudem festgestellt, daß sich Jungen sehr viel stärker für Mathematik und Naturwissenschaften (mit Ausnahme der Biologie) interessieren als Mädchen und daß dieser Unterschied mit zunehmendem Alter größer wird (Comber & Keeves, 1973; Gardner, 1985; Kelly, 1978, 1985, 1987; Kelly & Smail, 1986; Schibeci, 1985). Doch trotz ihres geringeren Interesses für Mathematik und Naturwissenschaften haben Mädchen (auch) in diesen Fächern oft bessere Schulleistungen als Jungen (Kimball, 1989).

Einige Autoren (z. B. Banreti-Fuchs & Meadows, 1976) sind der Meinung, daß sich Mädchen konformer verhalten als Jungen und daß sie eher bereit sind, sich in allen Fächern anzustrengen, ungeachtet ihrer Interessen. Sie haben deshalb auch in solchen Fächern gute Noten, in denen ihr Interesse gering ist. Neben dieser *Konformitätshypothese* gibt es jedoch auch andere Erklärungsmöglichkeiten. Es ist z. B. davon auszugehen, daß das Geschlechtsstereotyp zu einem insgesamt geringeren Interesse an naturwissenschaftlichen Themen führt (z.B. Kelly, 1978) und mit der Reduktion der Varianz der Interessenausprägung auch die Korrelation mit

Leistungsindikatoren sinkt. Darüber hinaus gibt es einige Befunde, die nahelegen, daß auch Faktoren des schulischen Unterrichts zu einem geringen Interesse der Mädchen an Naturwissenschaften beitragen (s. Kelly, 1987; Kubli, 1987; Wilkinson & Marrett, 1985).

Interesse, Fähigkeit und Leistung

Die Bedeutung von Interesse für die Erklärung von Leistungsunterschieden kann ohne die Berücksichtigung anderer Faktoren nicht hinreichend genau geschätzt werden. Die meisten Studien, die in die Metaanalyse Eingang fanden, kontrollierten z. B. nicht das Fähigkeitsniveau der Probanden. Es gibt jedoch eine Reihe von empirischen Belegen, daß Interesse neben kognitiven Faktoren einen eigenständigen Beitrag zur Erklärung von Leistungsunterschieden liefert (Barilleaux, 1961; Frandsen & Sessions, 1953; Frandsen & Sorenson, 1969; Geffert, 1985; Hall, 1975; Skager, Bussis & Schultz, 1965; s. a. die Übersichtsarbeiten von Evans, 1971; Super, 1960). Zudem wurden wiederholt relativ geringe Korrelationen zwischen fachspezifischen Interessen und allgemeinen Fähigkeitsfaktoren festgestellt (Gardner, 1985; Hungerman, 1967; Schibeci, 1985).

Steinkamp und Maehr (1983) diskutieren die Möglichkeit, daß das relative Gewicht von Interesse und Fähigkeit zur Vorhersage von Leistung mit dem individuellen Fähigkeitsniveau variiert. Sie postulieren eine Art Schwellenwert. Unterhalb dieses Schwellenwerts hängt die Leistung primär von der Fähigkeit ab, oberhalb des Schwellenwerts gewinnen die motivationalen Faktoren zunehmend an Bedeutung. Diese Vermutung stimmt mit Beobachtungen bei Hochbegabten oder besonders erfolgreichen Studenten oder Berufstätigen gut überein. Interindividuelle Unterschiede zeigen sich hier meist besonders deutlich im Bereich der Interessen und anderer motivationaler Merkmale und weniger im Bereich allgemeiner kognitiver Fähigkeiten (vgl. Hany & Nickel, 1992).

Die Tatsache, daß Interesse und Fähigkeit als voneinander unabhängige Prädiktoren der Leistung interpretiert werden können, schließt nicht aus, daß sich beide Faktoren im Entwicklungsverlauf gegenseitig beeinflussen. Mehrere Autoren (Bloom, 1976; Evans, 1971; Frandsen & Sorenson, 1969; Johnson, 1969) postulieren eine reziproke Beziehung zwischen Interesse und Fähigkeit: Hohes Interesse führt mit größerer Wahrscheinlichkeit zum Erfolg, der seinerseits positive Gefühle und Bewertungen zur Folge hat (z. B. Erhöhung des Selbstvertrauens, s. Helmke, 1992). Erfolgskontingente Gefühle und Bewertungen steigern wiederum das Interesse und die künftige Anstrengungsbereitschaft, die langfristig die Entwicklung kognitiver Fähigkeiten fördern. Leider gibt es so gut wie keine empirischen Belege für die Wechselbeziehung zwischen Interesse, Fähigkeit und Leistung. Möglicherweise variiert das Gewicht der Faktoren auch über die Zeit. Die Befunde der Metaanalyse von Willson (1983) belegen, daß sich die Einfluß-

richtung im Laufe der Schulzeit ändert: In niedrigeren Klassen bzw. in jüngerem Alter bestimmt die Leistung das Interesse. Bei älteren Schülern ist es eher umgekehrt, hier bestimmt das Interesse das erreichte Leistungsniveau.

Kritische Anmerkungen zur Methodik

Die meisten bisherigen Studien haben die prognostische Beziehung zwischen Interesse und Schulleistung mit Hilfe einfacher Korrelations- und Regressionsverfahren analysiert. Dabei wurde in der Regel hauptsächlich festgestellt, ob eine beobachtbare Leistungsvarianz mit hinreichender Sicherheit auf die Ausprägung eines Interessensfaktors zurückgeführt, d. h. statistisch „erklärt“ werden kann. Die wenigsten der von uns gesichteten Untersuchungen können als Prognosestudien i. e. S. bezeichnet werden, da sie Prädiktor und Kriterium gleichzeitig oder nur mit geringer zeitlicher Verzögerung erhoben haben. Nur in den Studien von Eisenhardt (1976) und Goldstein (1986) wurde ein größerer Zeitabstand zwischen Prädiktor- und Kriteriumsmessung realisiert. Eisenhardt konnte mit Hilfe der Methode der "cross-lagged panel"-Korrelation zeigen, daß Leistung zum Zeitpunkt t_1 höher mit Interesse zum Zeitpunkt t_2 korreliert als Interesse zum Zeitpunkt t_1 mit Leistung zum Zeitpunkt t_2 . Das deutet darauf hin, daß Interesse zwar tatsächlich Leistung prognostizieren kann, aber daß der umgekehrte Wirkzusammenhang ebenfalls besteht und möglicherweise stärker ist.

Ein weiteres Defizit der bisherigen Forschung betrifft die Tatsache, daß die meisten Vorhersagestudien keine systematische Kontrolle alternativer Einflußgrößen vorgenommen haben. In den wenigen Untersuchungen, die eine solche Kontrolle vorsahen, wurde entweder nach der Methode des Ausbalancierens darauf geachtet, daß die zu kontrollierenden Einflußgrößen in den Vergleichsgruppen annähernd gleich verteilt waren oder der Einfluß dieser Größen wurde auf statistischem Wege eliminiert (z. B. mittels Kovarianzanalysen).

Schließlich ist kritisch hervorzuheben, daß bisher — von einzelnen Versuchen abgesehen (z. B. Sjöberg, 1984) — keine Kausalanalysen durchgeführt wurden. Aussagefähige Kausalanalysen sind nur auf der Grundlage längsschnittlicher Daten und unter Einbeziehung zusätzlicher Einflußgrößen möglich.

Auch in inhaltlicher Hinsicht hat die Metaanalyse einige Forschungsdefizite offengelegt. Ein erstes Problem betrifft Geschlechtsunterschiede bezüglich der Höhe der Interesse-Leistungs-Korrelation. Obwohl immer wieder bestätigt wurde, daß Mädchen und Jungen sich im Hinblick auf ihre Fähigkeiten und Interessen in Natur- und Sozialwissenschaften unterscheiden, stellen diese Befunde jedoch noch keine hinreichende Erklärung für die generell niedrigen Interesse-Leistungs-Korrelationen der Mädchen dar. Studien, die sich speziell mit diesem Problem befassen, wären daher außerordentlich wünschenswert.

Ein weiteres Problem ist die genauere Untersuchung altersabhängiger Trends. Ein differenziertes Wissen auf diesem Gebiet hätte erhebliche praktische Konsequenzen. Wenn z. B. zutrifft, daß Schüler während der Pubertät nicht nur ihre Freizeitinteressen verändern, sondern auch die Struktur ihrer schulbezogenen Interessen, dann sollten sich Lehrer und Eltern gemeinsam rechtzeitig darauf einstellen und nach Möglichkeiten suchen, wie dieser Entwicklung produktiv begegnet werden kann.

Was auf der Ebene der interessenthematischen Forschung fehlt, sind Studien zur funktionalen oder kausalen Analyse der Zusammenhänge zwischen Interesse, Lernen und Leistung. Das betrifft in erster Linie die Faktoren und Prozesse im Lerngeschehen, die die Auswirkungen von Interesse erklären können (vgl. Krapp, 1992b; Schiefele, 1990, 1992). Neben Prozeßvariablen wie Aufmerksamkeit oder Aktivierung (Eysenck, 1982) sind dabei besonders Variablen zu berücksichtigen, die der Selbststeuerung des Lernprozesses dienen (z. B. Lernstrategien; vgl. Mandl & Friedrich, 1992; Schiefele, Krapp, Wild & Winteler, 1992; Schiefele & Pekrun, in Druck; Wild, Krapp & Winteler, 1992; Wild, Schiefele & Winteler, 1992). Eine Klärung dieser Fragen ist nur mit längsschnittlich angelegten Studien oder experimentellen Analysen möglich (z. B. Kuusinen & Leskinen, 1988). Auch quasi-experimentelle Studien können wichtige Befunde liefern (z. B. Schiefele, 1990, 1991a).

Insgesamt bestätigen die vorliegenden Untersuchungsbefunde, daß Interessen für die Erklärung von Lernen und Leistung in der Schule eine wichtige Rolle spielen. Die empirisch-psychologische Forschung hat die Untersuchung des Zusammenhangs von Interesse, Lernen und Leistung lange Zeit vernachlässigt. Andere motivationale Konzepte, insbesondere die Konzepte der Leistungsmotivationstheorie fanden eine viel stärkere Beachtung (Krapp, in Druck). Doch inzwischen hat man erkannt, daß die motivationalen Grundlagen des Lernens und der Leistung nur dann angemessen erfaßt werden können, wenn Faktoren wie Inhaltsspezifität, Gegenstandsbezug und intrinsische Motivation in die theoretischen und empirischen Analysen einbezogen werden. Dies genau ist das Anliegen der interessenthematischen Forschung, die gegenwärtig innerhalb und außerhalb der Pädagogischen Psychologie zunehmend Beachtung findet (vgl. Krapp & Prenzel, 1992; Renninger, Hidi & Krapp, 1992). Wir sind der Auffassung, daß die meisten Schüler keine „Allesfresser“ sind, d. h., daß sie nicht motiviert sind, alles mögliche zu lernen, ungeachtet der Bedeutung, die es für sie hat. Ebenso wird ihr „Wissenshunger“ nicht von ihren (erlernten oder angeborenen) Fähigkeiten determiniert. "It is not that students cannot learn, it is that they do not wish to" (Csikszentmihalyi, 1990, S. 115). Oder in John Nicholls Worten: "It is simply that more students might gain a sense of accomplishment and experiencing their work as meaningful if more researchers had been asking about the meaning of students' work rather than about their ability" (1990, S. 40).

Summary

This article presents results from a meta-analysis of studies on the relation between subject-matter-related interest and school achievement. For the time period between 1965 and 1990 a total of 21 studies reporting 127 independent correlations (i. e., correlations based on independent samples) were identified. For the overall relation between interest and achievement a mean correlation of .30 was found. Male students exhibited significantly higher interest-achievement correlations than female students. In addition, significant differences among school subjects were observed. Grade level, however, did not produce a significant moderator effect. Finally, the results are discussed on the basis of theories of interest and methodological considerations.

Literatur

- Banreti-Fuchs, K. M. & Meadows, W. M. (1976). Interest, mental health, and attitudinal correlates of academic achievement among university students. *British Journal of Educational Psychology*, 46, 212–219.
- Barak, A. (1981). Vocational interests: A cognitive view. *Journal of Vocational Behavior*, 19, 1–14.
- Barrilleaux, L. E. (1961). High school science achievement as related to interest and IQ. *Educational and Psychological Measurement*, 21, 929–936.
- Bloom, B. S. (1976) *Human characteristics and school learning*. New York: McGraw-Hill.
- Carlson, S. L. (1970). *Differences in aptitude, previous achievement, and nonintellectual traits (personality, values, interest, and attitude toward mathematics) of freshmen mathematics majors and transfers from the mathematics major at the University of Northern Colorado*. Unpublished doctoral dissertation, University of Northern Colorado, Greeley, CO.
- Carroll, J. B. (1975). *The teaching of French as a foreign language in eight countries*. Stockholm/New York: Almqvist & Wiksell/Wiley.
- Comber, L. C. & Keeves, J. P. (1973). *Science education in nineteen countries. An empirical study*. New York: Wiley.
- Cronbach, L. J. & Gleser, G. C. (1965). *Psychological tests and personnel decisions*. Urbana: University of Illinois Press.
- Csikszentmihalyi, M. (1990). Literacy and intrinsic motivation. *Daedalus*, 199, 115–140.
- Deci, E. L. (1992). The relation of interest to the motivation of behavior: A self-determination theory perspective. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 43–47). Hillsdale, NJ: Erlbaum.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Dewey, J. (1913). *Interest and effort in education*. Boston: Riverside Press.
- Eisenhardt, W. B. (1976). *A search for the predominant causal sequence in the interrelationship of interest in academic subjects and academic achievement. A cross-lagged panel correlation study*. Unpublished doctoral dissertation, Duke University, Durham, NC.
- Evans, K. M. (1971). *Attitudes and interests in education*. London: Routledge & Kegan Paul.

- Eysenck, M. W. (1982). *Attention and arousal*. Berlin: Springer.
- Fishman, J. A. & Pasanella, A. K. (1960). College admission-selection studies. *Review of Educational Research*, 30, 298–310.
- Fleming, M. L. & Malone, M. R. (1983). The relationship of student characteristics and student performance in science as viewed by meta-analysis research. *Journal of Research in Science Teaching*, 20, 481–495.
- Frandsen, A. N. & Sessions, A. D. (1953). Interests and school achievement. *Educational and Psychological Measurement*, 13, 94–101.
- Frandsen, A. N. & Sorenson, M. (1969). Interests as motives in academic achievement. *Journal of School Psychology*, 7, 52–57.
- Fricke, R. & Treinies, G. (1985). *Einführung in die Metaanalyse*. Bern: Huber.
- Gagné, F. (1985). Giftedness and talent: Reexamining a reexamination of the definitions. *Gifted Child Quarterly*, 29, 103–112.
- Gardner, P. L. (1985). Students' interest in science and technology: An international overview. In M. Lehrke, L. Hoffmann & P. L. Gardner (Eds). *Interests in science and technology education* (pp. 15–34). Kiel: Institut für die Pädagogik der Naturwissenschaften.
- Geffert, E. (1985). Motivationale Grundlagen der mathematischen Begabung. *Zeitschrift für Psychologie*, 193, 431–441.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3–8.
- Glass, G. V., McGaw, B. & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Goldstein, S. P. (1986). *The gifted child: An investigation of science achievement with respect to spatial ability, science interest, curiosity, level of giftedness, and sex of child*. Unpublished doctoral dissertation, George Washington University.
- Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology Monograph*, 28, 545–579.
- Haertel, G. D., Walberg, H. J. & Weinstein, T. (1983). Psychological performance models of educational performance: A theoretical synthesis. *Review of Educational Research*, 53, 75–91.
- Hall, E. R. (1975). Motivation and achievement in black and white junior college students. *Journal of Social Psychology*, 97, 107–113.
- Hany, E. A. & Nickel, H. (Hrsg.). (1992). *Begabung und Hochbegabung*. Bern/Göttingen: Huber.
- Harty, H. & Beall, D. (1984). Toward the development of a children's science curiosity measure. *Journal of Research in Science Teaching*, 21, 425–436.
- Hedges, L. V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Heller, K. (Hrsg.). (1991). *Begabungsdagnostik in der Schul- und Erziehungsberatung*. Bern/Göttingen: Huber.
- Helmke, A. (1992). *Selbstvertrauen und schulische Leistung*. Göttingen: Hogrefe.
- Herbart, J. F. (1806/1965). Allgemeine Pädagogik, aus dem Zweck der Erziehung abgeleitet. In J. F. Herbart, *Pädagogische Schriften* (Bd. 2, S. 9–155). Düsseldorf: Küpper.

- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, 60, 549–571.
- Huetinck, L. (1990). *Gender differences on science exams with respect to item type, format, and student interests and experience*. Unpublished doctoral dissertation, University of California, Los Angeles, CA.
- Hungermann, A. D. (1967). Achievement and attitude of sixthgrade pupils in conventional and contemporary mathematics programs. *Arithmetic Teacher*, 14, 30–39.
- Hunter, J. E. & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.
- Hunter, J. E. & Schmidt, F. L. & Jackson, G. B. (1982) *Meta-analysis: Cumulating research findings across studies*. Beverly Hills, CA: Sage.
- Husén, T. (Ed.) (1967). *International study of achievement in mathematics (Vols. 1 and 2)*. Stockholm/New York: Almqvist & Wiksell/Wiley.
- Johnson, R. W. (1969). Effectiveness of SVIB academic interest scales in predicting college achievement. *Journal of Applied Psychology*, 53, 309–316.
- Jungermann, H. (1976). *Rationale Entscheidungen*. Bern: Huber.
- Kelly, A. (1978). *Girls and science: An international study of sex differences in school science achievement*. Stockholm: Almqvist & Wiksell.
- Kelly, A. (1985). The development of girls' and boys' attitudes to science: A longitudinal study. In M. Lehrke, L. Hoffmann & P. L. Gardner (Eds.), *Interests in science and technology education* (pp. 269–280). Kiel: Institut für die Pädagogik der Naturwissenschaften.
- Kelly, A. (1987). Die Entwicklung naturwissenschaftlicher Interessen und Einstellungen bei Mädchen und Jungen. In M. Lehrke & L. Hoffmann (Hrsg.), *Schülerinteressen am naturwissenschaftlichen Unterricht*. Köln: Aulis-Verlag Deubner.
- Kelly, A. & Smail, B. (1986). Sex stereotypes and attitudes to science among 11-year-old children. *British Journal of Educational Psychology*, 56, 158–168.
- Kerschensteiner, G. (1922). Der Interessenbegriff in der Pädagogik. *Pädagogische Blätter*, 5, 349–354.
- Khan, S. B. (1969) Affective correlates of academic achievement. *Journal of Educational Psychology*, 60, 216–221.
- Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin*, 105, 198–214.
- Klauer, K. J. (1992). Zur Diagnostik von Hochbegabung. In E. A. Hany & H. Nickel (Hrsg.), *Begabung und Hochbegabung* (S. 205–214). Bern/Göttingen: Huber.
- Krapp, A. (1973). *Bedingungen des Schulerfolges*. München: Oldenbourg.
- Krapp, A. (1979). *Prognose und Entscheidung*. Weinheim: Beltz.
- Krapp, A. (1984). Forschungsergebnisse zur Bedingungsstruktur der Schulleistung. In K. Heller (Hrsg.), *Leistungsdiagnostik in der Schule* (pp. 46–62). Bern: Huber.
- Krapp, A. (1989). Der Stellenwert des Interessenkonzepts in der pädagogisch orientierten Forschung. *Empirische Pädagogik*, 3, 233–255.

- Krapp, A. (1992 a). Das Interessenkonstrukt. Bestimmungsmerkmale der Interessenhandlung und des individuellen Interesses aus der Sicht einer Person-Gegenstands-Konzeption. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung* (S. 297–329). Münster: Aschendorff.
- Krapp, A. (1992 b). Konzepte und Forschungsansätze zur Analyse des Zusammenhangs von Interesse, Lernen und Leistung. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung* (S. 9–52). Münster: Aschendorff.
- Krapp, A. (in Druck). Psychologie der Lernmotivation – Perspektiven der Forschung und Probleme ihrer pädagogischen Rezeption. *Zeitschrift für Pädagogik*, 39.
- Krapp, A. & Prenzel, M. (Hrsg.). (1992). *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung*. Münster: Aschendorff.
- Kubli, F. (1987). *Interesse und Verstehen in Physik und Chemie*. Köln: Aulis-Verlag Deubner.
- Kühn, R. (1983). *Bedingungen für Schulerfolg. Zusammenhänge zwischen Schülermerkmalen, häuslicher Umwelt und Schulnoten*. Göttingen: Hogrefe.
- Kulik, J. A. & Kulik, C. C. (1989). Meta-analysis in education. *International Journal of Educational Research*, 13, Issue 3.
- Kuusinen, J. & Leskinen, E. (1988). Latent structure analysis of longitudinal data on relations between intellectual abilities and school achievement. *Multivariate Behavioral Research*, 8, 103–118.
- Lavin, D. E. (1965). *The prediction of academic performance*. New York: Russell Sage Foundation.
- Lehrke, M. (1988). *Interesse und Desinteresse am naturwissenschaftlich-technischen Unterricht*. Kiel: Institut für die Pädagogik der Naturwissenschaften (IPN).
- Lewis, E. G. & Massad, C. E. (1975). *The teaching of English as a foreign language in ten countries*. Stockholm/New York: Almqvist & Wiksell/Wiley.
- Lloyd, J. & Barenblatt, L. (1984). Intrinsic intellectuality: Its relations to social class, intelligence, and achievement. *Journal of Personality and Social Psychology*, 46, 655–668.
- Löwe, B. (1992). *Schülerinteressen zur Biologie und Biologieunterricht*. Weinheim: Deutscher Studienverlag.
- Lowe, A. (1970). *Cognitive teaching effectiveness: The effects of student motivation, interest, and teacher influence upon junior college students' achievement*. Unpublished doctoral dissertation, Florida State University, Tallahassee, FL.
- Mandl, H. & Friedrich, H. F. (Hrsg.). (1992). *Lern- und Denkstrategien*. Göttingen: Hogrefe.
- 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.
- Meece, J. L., Wigfield, A. & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, 82, 60–70.
- Napier, J. D. & Riley, J. P. (1985). Relationship between affective determinants and achievement in science for seventeen-year-olds. *Journal of Research in Science Teaching*, 22, 365–383.

- Nicholls, J. G. (1990). What is ability and why are we mindful of it? A developmental perspective. In R. J. Sternberg & J. Kolligian, Jr. (Eds.), *Competence considered* (pp. 11–40). New Haven/London: Yale University Press.
- Oehlert, P. (1977). *Aufstellung und erste empirische Überprüfung eines Konstruktes zur Methode der Steuerung von kognitiven Lernprozessen durch Unterrichtsimpulse*. Unveröffentlichte Dissertation, Pädagogische Hochschule Westfalen-Lippe.
- Parkerson, J. A., Lomax, R. G., Schiller, D. P. & Walberg, H. J. (1984). Exploring causal models of educational achievement. *Journal of Educational Psychology*, 76, 638–646.
- Prenzel, M., Krapp, A. & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik*, 32, 163–173.
- Purves, A. C. (1973). *Literature education in ten countries. An empirical study*. New York: Halsted Press.
- Quack, L. (1979). Zur Bedingungsanalyse der Schulleistung: Der Beitrag kognitiver und nicht-kognitiver Merkmale der Schülerpersönlichkeit. In K. J. Klauer & H. J. Kornadt (Hrsg.), *Jahrbuch für empirische Erziehungswissenschaft* (S. 93–116). Düsseldorf: Schwann.
- Rennie, L. J. & Punch, K. F. (1991). The relationship between affect and achievement in science. *Journal of Research in Science Teaching*, 28, 193–209.
- Renninger, K. A., Hidi, S. & Krapp, A. (Eds.). (1992). *The role of interest in learning and development*. Hillsdale, NJ: Erlbaum.
- Renzulli, J. S. (1988). A decade of dialogue on the three-ring conception of giftedness. *Roeper Review*, 11, 19–25.
- Reynolds, A. J. & Walberg, H. J. (1991). A structural model of science achievement. *Journal of Educational Psychology*, 83, 97–107.
- Rubinstein, S. L. (1985). *Grundlagen der allgemeinen Psychologie*. Berlin: Volk & Wissen.
- Sauer, J. & Gattringer, H. (1985). Soziale, familiale, kognitive und motivationale Determinanten der Schulleistung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 37, 288–309.
- Schibeci, R. (1985). Students' attitudes to science: What influences them, and how these influences are investigated. In M. Lehrke, L. Hoffmann & P. L. Gardner (Eds.), *Interests in science and technology education* (pp. 35–48). Kiel: Institut für die Pädagogik der Naturwissenschaften.
- Schiefele, H. (1978). *Lernmotivation und Motivlernen* (2. Aufl.). München: Ehrenwirth.
- Schiefele, H., Hausser, K. & Schneider, G. (1979). „Interesse“ als Ziel und Weg der Erziehung. Überlegungen zu einem vernachlässigten pädagogischen Konzept. *Zeitschrift für Pädagogik*, 25, 1–20.
- Schiefele, U. (1988). Motivationale Bedingungen des Textverstehens. *Zeitschrift für Pädagogik*, 34, 687–708.
- Schiefele, U. (1990). Thematisches Interesse, Variablen des Leseprozesses und Textverstehens. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 37, 304–332.
- Schiefele, U. (1991a). Interesse und Textrepräsentation – Zur Auswirkung des thematischen Interesses auf unterschiedliche Komponenten der Textrepräsentation unter Berücksichtigung kognitiver und motivationaler Kontrollvariablen. *Zeitschrift für Pädagogische Psychologie*, 5, 245–259.
- Schiefele, U. (1991b). Interest, learning, and motivation. *Educational Psychologist*, 26, 299–323.

- Schiefele, U. (1992). Topic interest and levels of text comprehension. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 151–182). Hillsdale, NJ: Erlbaum.
- Schiefele, U., Krapp, A., Wild, K.-P. & Winteler, A. (1992). *Eine neue Version des „Fragebogen zum Studieninteresse“* (Gelbe Reihe, Arbeiten zur Empirischen Pädagogik und Pädagogischen Psychologie, Nr. 21). München: Universität der Bundeswehr, Institut für Erziehungswissenschaft und Pädagogische Psychologie.
- Schiefele, U. & Pekrun, R. (in Druck). Psychologische Modelle des selbstgesteuerten und fremdgesteuerten Lernens. In F. E. Weinert (Hrsg.), *Psychologie des Lernens und der Instruktion* (Enzyklopädie der Psychologie). Göttingen: Hogrefe.
- Schiefele, U. & Schreyer, I. (in Druck). Intrinsische Lernmotivation und Lernen. Ein Überblick zu Ergebnissen der Forschung. *Zeitschrift für Pädagogische Psychologie*, 7.
- Schneider, W. & Bös, K. (1985). Exploratorische Analysen zu Komponenten des Schulerfolgs. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 17, 325–340.
- Schwarzer, R. (1989). *Meta-analysis programs (unpublished manuscript)*. Berlin: Free University of Berlin, Department of Psychology.
- Sjöberg, L. (1983). Interest, achievement and vocational choice. *European Journal of Science Education*, 5, 299–307.
- Sjöberg, L. (1984). Interests, effort, achievement and vocational preference. *British Journal of Educational Psychology*, 54, 189–205.
- Skager, R. W., Bussis, A. M. & Schultz, C. B. (1965). Comparison of information scales and like-indifferent-dislike scales as measures of interest. *Psychological Reports*, 16, 251–261.
- Steinkamp, M. W. & Maehr, M. L. (1983). Affect, ability and science achievement: A quantitative synthesis of correlational research. *Review of Educational Research*, 53, 369–396.
- Stodolsky, S. S. (1988). *The subject matters*. Chicago: University of Chicago Press.
- Super, D. E. (1960). Interests. In C. W. Harris (Ed.), *Encyclopedia of educational research* (pp. 728–733). New York: Macmillan.
- Tiedemann, J. (1977). *Leistungsversagen in der Schule*. München: Goldmann.
- Todt, E. (1967). *Differentieller Interessentest (DIT)*. Bern: Huber.
- Todt, E. (1978). *Das Interesse*. Bern: Huber.
- Todt, E. (1985). Elements of a theory of science interests. In M. Lehrke, L. Hoffmann & P. L. Gardner (Eds.), *Interests in science and technology education* (pp. 59–69). Kiel: Institut für die Pädagogik der Naturwissenschaften.
- Travers, R. M. W. (1978). *Children's interests*. Kalamazoo, Michigan: Michigan University, College of Education.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- Uguroglu, M. E. & Walberg, H. J. (1979). Motivation and achievement: A quantitative synthesis. *American Educational Research Journal*, 16, 375–389.
- Wade, S. E. (1992). How interest affects learning from text. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 255–277). Hillsdale, NJ: Erlbaum.

- Wendeler, J. (1968). Schülerinteressen bei Gymnasiasten der 10. Klasse. *Schule und Psychologie*, 15, 114–119.
- Wieczerkowski, W. & zur Oeveste, H. (1982). Zuordnungs- und Entscheidungsstrategien. In K. J. Klauer (Hrsg.), *Handbuch der pädagogischen Diagnostik* (Bd. 2, S. 919–951). Düsseldorf: Schwann.
- Wild, K.-P., Krapp, A. & Winteler, A. (1992). Die Bedeutung von Lernstrategien zur Erklärung des Einflusses von Studieninteresse auf Lernleistungen. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung* (S. 279–295). Münster: Aschendorff.
- Wild, K.-P., Schiefele, U. & Winteler, A. (1992). *LIST: Ein Verfahren zur Erfassung von Lernstrategien im Studium* (Gelbe Reihe, Arbeiten zur Empirischen Pädagogik und Pädagogischen Psychologie, Nr. 20). München: Universität der Bundeswehr, Institut für Erziehungswissenschaft und Pädagogische Psychologie.
- Wilkinson, L. C. & Marrett, C. B. (1985). *Gender influences in classroom interactions*. Orlando, FL: Academic Press.
- Willson, V. L. (1983). A meta-analysis of the relationship between science achievement and science attitude: Kindergarten through college. *Journal of Research in Science Teaching*, 20, 839–850.

Anschrift der Verfasser:

Dr. Ulrich Schiefele
Prof. Dr. Andreas Krapp
Inge Schreyer
Universität der Bundeswehr, Fakultät für Sozialwissenschaften
Werner-Heisenberg-Weg 39, 8014 Neubiberg bei München