

**Leistungsentwicklungen und berufliche Interessen
in der gymnasialen Oberstufe**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von

Jasmin Warwas

aus Northeim

Göttingen 2008

D 7

Referent: Prof. Dr. Marcus Hasselhorn

Korreferent: Prof. Dr. Rainer Watermann

Tag der mündlichen Prüfung: 24. Oktober 2008

Danksagung

Diese Arbeit wurde durch ein Stipendium im Rahmen des Graduiertenkollegs GRK 1195 „Passungsverhältnisse schulischen Lernens – Verstehen und Optimieren“ von der Deutschen Forschungsgemeinschaft gefördert. Vielen Dank für die finanzielle und materielle Unterstützung. Die empirischen Studien dieser Arbeit fanden in Kooperation mit der Arnoldi-Schule in Göttingen statt. Insbesondere möchte ich hier dem Koordinator Herrn Rölleke, dem Team der beteiligten Fachlehrerinnen und -lehrer sowie der Direktorin Frau Wucherpfennig für die gelungene Zusammenarbeit danken.

Mein besonderer Dank gilt meinen Betreuern Herrn Prof. Dr. Rainer Watermann und Herrn Prof. Dr. Marcus Hasselhorn. Herrn Watermann danke ich für den wertvollen methodischen und inhaltlichen Austausch während der gesamten Projektlaufzeit. Herrn Hasselhorn danke ich für die hilfreichen Anmerkungen zur Entwicklung des Dissertationsprojekts und zu den Manuskripten. Eng mit der Entstehung dieser Arbeit verbunden möchte ich auch Gabriel Nagy für ausschlaggebende methodische Ideen zu Analysemethoden danken. Jennie Auffenberg gilt mein Dank für die hilfreiche Unterstützung bei Literaturverwaltung und Datenerhebung.

Nicht weniger bedeutend für das Entstehen dieser Dissertation waren meine Kolleginnen und Kollegen bzw. Freundinnen und Freunde im GRK. Danke an euch. Ein besonderer Dank gilt Adrienne Schmeling für anregende Gespräche und für die stets herzliche Atmosphäre in unserem Büro. Bastian Funken danke ich für abwechslungsreiche Freizeitgestaltung, kulinarische Experimente und intensives Korrekturlesen. Ich danke Felix Brümmer für die Bereitschaft sich jederzeit mit meinen Fragen zu verschiedenen Inhalten und Methoden auseinanderzusetzen und wertvolle Anregungen zu geben. Ein herzlicher Dank geht an Klaudia Schulte für 1a-Copingstrategien, für emotionalen Beistand und für all die (wahn-)witzigen Momente. Matthias Martens gilt mein Dank für die unvergessen bleibenden Bewirtungen und bei Sven Lindberg möchte ich für die Reanimation meiner sportlichen Fertigkeiten bedanken.

Einen großen Dank möchte ich auch meinen Freundinnen aus Schul- und Studienzeit bedanken. An erster Stelle gilt hier ein Dank an Marina Pätzold für uneingeschränkte Loyalität und Empathie. Ich danke Petra Fischer für die emotionale Unter-

stützung insbesondere in schwierigen Phasen der Dissertation sowie den nötigen Ausgleich im sportlichen Bereich. Schließlich möchte ich noch Simone Salzer danken für einfühlsame Gespräche und die ausgestrahlte Ruhe.

Tobias Koch möchte ich für emotionales Beiseitestehen auf einem Grossteil des Weges danken und für die Unkompliziertheit und Verständnis auch in Phasen hoher Arbeitsbelastung.

Schließlich möchte ich einen herzlichen Dank meiner Familie aussprechen, meinen Eltern Rita und Klaus-Peter Warwas und meinen Geschwistern Patricia Thienel, Constanze Hügel und Vanessa Pätzold. Insbesondere danke ich meiner Mutter für das stete Vertrauen und die gewährten Freiheiten sowie meiner Schwester Patricia für fortwährendes Zuhören.

Inhaltsverzeichnis

Tabellenverzeichnis	iii
Abbildungsverzeichnis	iv
Einleitung	1
1. Leistungsentwicklungen in der gymnasialen Oberstufe.....	5
1.1 Strukturelle Ausgestaltung und Zielsetzungen der gymnasialen Oberstufe	5
1.2 Erfassung schulischer Leistungen und Leistungsentwicklungen	8
1.3 Leistungsentwicklungen in Mathematik und Englisch in der gymnasialen Oberstufe.....	12
2. Die Bedeutung beruflicher Interessen für schulische Leistungen	15
2.1 Die Person-Gegenstands-Konzeption des Interesses	15
2.2 Die Theorie der Berufswahl nach Holland	17
2.2.1 Kernannahmen in Hollands Berufswahltheorie	18
2.2.2 Charakteristika beruflicher Interessenprofile sensu Holland.....	20
2.2.3 Berufliche Interessenprofile im Zusammenhang mit schulischen Leistungen	23
2.3 Die Bedeutung beruflicher Interessen für den Zusammenhang von Selbstkonzept und Kursniveau mit schulischen Leistungsentwicklungen	25
3. Fragestellungen	29
4. Leistungsentwicklungen und berufliche Interessen in der gymnasialen Oberstufe	33
4.1 Leistungsentwicklungen in Mathematik und Englisch im ersten Jahr an einem Fachgymnasium.....	33
<i>Zusammenfassung und Abstract</i>	33
4.1.1 <i>Theoretischer Hintergrund</i>	34
4.1.2 <i>Fragestellungen</i>	38
4.1.3 <i>Methode</i>	40
4.1.4 <i>Ergebnisse</i>	42
4.1.5 <i>Diskussion</i>	45

4.2 The Relations of Vocational Interests and Mathematical Literacy: On the Predictive Power of Interest Profiles	49
<i>Abstract</i>	49
4.2.1 <i>Theoretical Background</i>	49
4.2.2 <i>Research Questions</i>	54
4.2.3 <i>Method</i>	55
4.2.4 <i>Results</i>	60
4.2.5 <i>Discussion</i>	70
4.3 Zum Effekt von beruflichen Interessenprofilen, Selbstkonzept und Kursniveau auf Leistungsentwicklungen in Mathematik und Englisch	74
<i>Zusammenfassung und Abstract</i>	74
4.3.1 <i>Theoretischer Hintergrund</i>	75
4.3.2 <i>Fragestellungen</i>	81
4.3.3 <i>Methode</i>	83
4.3.4 <i>Ergebnisse</i>	86
4.3.5 <i>Diskussion</i>	91
5. Zusammenfassende Schlussdiskussion und Ausblick.....	95
5.1 Empirische Erträge und Einordnung in den aktuellen Forschungsstand	95
5.2 Methodische Entwicklungen im Verlauf der empirischen Studien	98
5.2.1 <i>Missing Data</i>	98
5.2.2 <i>Clustereffekte</i>	100
5.3 Integrative Diskussion zur Bedeutung von beruflichen Interessen für schulische Leistungsentwicklungen.....	101
5.3.1 <i>Fachspezifische Unterschiede bei der Bedeutung beruflicher Interessen</i>	101
5.3.2 <i>Implikationen für den Unterricht</i>	102
5.4 Grenzen der Studie	105
5.4.1 <i>Effektrichtungen</i>	105
5.4.2 <i>Selektivität der Stichprobe</i>	106
Zusammenfassung	111
Literaturverzeichnis	113
Anhang.....	128

Tabellenverzeichnis

Tabelle 1	Charakterisierung der sechs Orientierungen nach Holland.....	20
Tabelle 2	Ausgangsleistungen pro Kohorte in Mathematik und Englisch.....	43
Tabelle 3	Mittlere Leistungsentwicklungen in mathematischer Grundbildung und Englisch nach dem Übergang und am Ende der Einführungsphase....	43
Tabelle 4	Kursniveauspezifische mittlere Leistungsentwicklungen in mathematischer Grundbildung und Englisch nach dem Übergang und am Ende der Einführungsphase.....	44
Table 5	Means, Standard Deviations, and Correlations Between Interest Scale Scores, Mathematical Literacy at t1 and t2, and Gender	61
Table 6	Correlations Between Interest Profile Parameters (Level, Unweighted Interest Orientation, and Interest Orientation Weighted by Differentiation) and Mathematical Literacy at t1 and t2, and with Gender	62
Table 7	Multiple Regression of Mathematical Literacy at t1 and t2 on Interest Scale Scores and Covariates.....	64
Table 8	Multiple Regression of Mathematical Literacy at t1 and t2 on Interest Profile Parameters (Level and Unweighted Interest Orientation) and Covariates.....	65
Table 9	Multiple Regression of Mathematical Literacy at t1 and t2 on Interest Profile Parameters (Level and Interest Orientation Weighted by Differentiation) and Covariates.....	68
Tabelle 10	Korrelationen zwischen Mathematik- und Englisch-Leistungen sowie beruflichen Interessen, Selbstkonzept und Kursniveau.....	87
Tabelle 11	Befunde zur Vorhersage von Mathematikleistungen zu T2; standardisierte Regressionskoeffizienten und in Klammern Standardfehler	88
Tabelle 12	Befunde zur Vorhersage von Mathematikleistungen zu T2; standardisierte Regressionskoeffizienten und in Klammern Standardfehler.....	89

Abbildungsverzeichnis

Figure 1	Interest profiles based on Holland's RIASEC model.....	52
Figure 2	Representation of model parameters in interest profiles	59
Figure 3	Predicted mathematical literacy at t1 and t2 by unweighted profile orientation and profile orientation weighted by amplitudes.....	67
Figure 4	Amount of variance explained, R^2 , in the regression analyses with vocational interest as predictors of mathematical literacy	69
Abbildung 5	RIASEC-Modell nach Holland (1997).....	77

Einleitung

Eine sichere Anwendung mathematischer Konzepte und Regeln sowie gute Englischkenntnisse in Wort und Schrift gelten als Basisqualifikationen zur Bewältigung schulischer, beruflicher und gesellschaftlicher Anforderungen (vgl. Kultusministerkonferenz [KMK], 1995; Tenorth, 2001). Mathematische Grundkenntnisse können als elementare Kulturwerkzeuge angesehen werden, mit denen eine erfolgreiche Teilhabe an der Gesellschaft möglich ist. Englisch wird als Weltsprache (Lingua Franca) angesehen. Im Zuge einer zunehmenden Internationalisierung erlangen kommunikative Fähigkeiten in Englisch als unentbehrliche Kulturtechniken zentrale Bedeutung (Tenorth, 2001).

Eine diesen Beurteilungen entsprechende Betonung erfahren die beiden Fächer auch in der gymnasialen Oberstufe. So ist es ein zentrales Ziel des Unterrichts in der gymnasialen Oberstufe, den Schülerinnen und Schülern eine vertiefte Allgemeinbildung zu vermitteln. In diesem Kontext wird nicht zuletzt durch aktuelle Reformen (vgl. „Husumer Beschluss“ der KMK vom 22. Oktober 1999) die Wichtigkeit von Mathematik und Englisch noch einmal betont. Die zentrale Bedeutung spiegelt sich auch in der aktuellen Fassung der Vereinbarung zur Gestaltung der gymnasialen Oberstufe wider, in der festgelegt ist, dass Mathematik und Englisch durchgängig bis zum Abitur belegt werden müssen (vgl. KMK, 2006).

Da ein wesentliches Ziel von Schule und Unterricht die Förderung schulischer Leistungen ist, stellt sich im Rahmen pädagogisch-psychologischer Forschung die Frage nach den Bedingungen für positive Leistungsentwicklungen in Mathematik und Englisch in der gymnasialen Oberstufe (vgl. Helmke & Weinert, 1997). So besteht weitgehend Konsens, dass Interessen einflussreiche Bedingungsfaktoren des Lernens und somit aussagekräftige Prädiktoren schulischer Leistungen sind (vgl. Krapp, 1998). Köller, Baumert und Schnabel (2000) verweisen auf die besondere Bedeutung der strukturellen Ausgestaltung der gymnasialen Oberstufe hinsichtlich der Rolle von Interessen für schulische Leistungen. So sei es (erst) in der gymnasialen Oberstufe aufgrund von Freiräumen, wie Kurswahlen, möglich, dass Interessen einen Effekt auf Leistungen ausüben. Eine zusätzliche Bedeutung kommt Interessen in Hinblick auf das Anforderungsniveau der besuchten Kurse in der gymnasialen Oberstufe zu. Schülerinnen und Schüler wählen aufgrund von selbst wahrgenommenen Fähigkeiten und Interessen (vgl.

z.B. Hodapp & Mißler, 1996; Köller, Daniels, Schnabel & Baumert, 2000; Köller, Trautwein, Lüdtke & Baumert, 2006) Kurse auf unterschiedlichem Anforderungsniveau. Mit einem erhöhten Anforderungsniveau geht in der Regel ein kognitiv anspruchsvollerer Unterricht in dem Fach einher. Im Rahmen einer Person-Gegenstands-Konzeption sollte sich aufgrund der intensiveren Auseinandersetzung mit dem Thema wiederum das Interesse an diesem Gegenstand erhöhen (Krapp, 1992). Während erste längsschnittliche Studien in der gymnasialen Oberstufe entsprechende Befunde für das Fach Mathematik bestätigen (vgl. Köller, Baumert et al., 2000; Köller et al., 2006), liegen Untersuchungen für das Fach Englisch bislang nicht vor.

Im deutschsprachigen Raum wurden in den letzten Jahren theoretische Modelle und empirische Zugänge zu *beruflichen* Interessen im Rahmen der Person-Gegenstands-Konzeption ausgearbeitet (Bergmann & Eder, 2005). Berufliche Interessen bilden sich im Jugendalter heraus (z.B. Tracey, 2001); ihnen liegt eine kognitive Repräsentation der gesamten Berufslandschaft zugrunde. Diese lässt sich wiederum in bestimmte Berufsfelder bzw. Tätigkeitsklassen einteilen (Rolfes, 2001). Auch berufliche Interessen nehmen eine handlungsleitende Funktion ein. Sie sind wichtige Determinanten von Studienfachwahlen (Nagy, 2006) und stehen in Zusammenhang mit mathematischen und sprachlichen Leistungen (vgl. Ackerman & Heggestad, 1997). In einer aktuellen Längsschnittstudie ließen sich keine zeitversetzten Effekte zwischen beruflichen Interessen und Leistungen in Mathematik und Englisch nachweisen (Tracey, Robbins & Hofsess, 2005). Bisherige Befunde beziehen sich jedoch auf einzelne Skalenwerte beruflicher Interessen. Ein umfassendere und gleichzeitig sparsame Variante sind Profile beruflicher Interessen. Hierzu belegte Nagy (2006) Zusammenhänge mit Testleistungen in Mathematik und Englisch; eine längsschnittliche Prüfung fehlt bislang.

Für ein Verständnis der Leistungsentwicklungen in der gymnasialen Oberstufe (Kapitel 1) erscheint es zunächst einmal notwendig, die strukturelle Ausgestaltung der gymnasialen Oberstufe und die damit verbundenen Zielsetzungen in Hinblick auf die Kernfächer Mathematik, Deutsch und erste Fremdsprache zu beschreiben. Vor dem Hintergrund inhaltlicher und methodischer Aspekte zur Erfassung und Beschreibung schulischer Leistungen werden bisherige Befunde zu Leistungsentwicklungen in Mathematik und Englisch berichtet. In Kapitel 2 wird auf die Bedeutung von Interessen für schulische Leistungen eingegangen. Anhand einer Person-Gegenstands-Konzeption des

Interesses werden Gemeinsamkeiten und Abgrenzungen zwischen fachlichen und beruflichen Interessen ausgearbeitet und daran anschließend die empirisch breit untermauerte Theorie der Berufswahl von Holland (1997) vorgestellt. Im Anschluss werden verschiedene Methoden zur Modellierung beruflicher Interessen nach Holland angeführt und kritisch reflektiert. Auf dieser Grundlage werden dann berufliche Interessen in Zusammenhang mit weiteren Einflussgrößen gebracht, die für die Leistungsentwicklung in der gymnasialen Oberstufe als zentral angesehen werden. Aus den angeführten theoretischen Aspekten und empirischen Befunden werden in Kapitel 3 die übergeordneten Fragestellungen dieser Arbeit abgeleitet und expliziert. Diesen wird in den nachfolgenden Abschnitten des vierten Kapitels in empirischen Studien nachgegangen. In Studie 1 (Abschnitt 4.1) werden Leistungsentwicklungen in zwei Kernfächern der gymnasialen Oberstufe untersucht. In Studie 2 (Abschnitt 4.2) wird ein methodischer Zwischenschritt vorgenommen, indem exemplarisch am Fach Mathematik die Validität von Interessenprofilen im Vergleich zu Skalenwerten beruflicher Interessen geprüft wird. Auf dieser Basis werden dann im dritten empirischen Teil dieser Arbeit Determinanten der Leistungsentwicklungen in Mathematik und Englisch untersucht (Abschnitt 4.3). Dazu wird die Rolle beruflicher Interessenprofile für Leistungsentwicklungen analysiert. Zum anderen werden berufliche Interessen in Theorien und Modelle zur Erklärung schulischer Leistungen integriert und deren Wechselspiel mit fachspezifischem Selbstkonzept und Kursniveau in der gymnasialen Oberstufe untersucht. Die zentralen Ergebnisse dieser Studien werden im Kapitel 5 zusammengefasst, in den gegenwärtigen Forschungsstand eingeordnet und integrativ diskutiert sowie die Grenzen dieser Arbeit aufgezeigt.

1. Leistungsentwicklungen in der gymnasialen Oberstufe

1.1 Strukturelle Ausgestaltung und Zielsetzungen der gymnasialen Oberstufe

In der gymnasialen Oberstufe soll den Schülerinnen und Schülern eine vertiefte Allgemeinbildung, eine allgemeine Studierfähigkeit sowie eine wissenschaftspropädeutische Bildung vermittelt werden (vgl. KMK, 2006). Die gymnasiale Oberstufe umfasst traditionell die drei Jahrgangsstufen 11 bis 13,¹ die sich in eine einjährige Einführungs- und eine zweijährige Qualifikationsphase einteilen lassen. Die Jahrgangsstufe 11 ist als Einführungsphase anzusehen; ihr kommt beim Übergang in die Qualifikationsphase eine Brückenfunktion zu. Unterschiedliche Voraussetzungen bei den Schülerinnen und Schülern sollen während dieses Schuljahres ausgeglichen und ihr Wissensstand angeglichen werden, um auf die Qualifikationsphase der Jahrgangsstufen 12 und 13 vorzubereiten. Diese Zielsetzung spiegelt sich auch im Curriculum der Einführungsphase wider: So ist es im Mathematikunterricht der Jahrgangsstufe 11 ein übergeordnetes Ziel, den Schülerinnen und Schülern vergleichbare Lernvoraussetzungen für die Mitarbeit in der Qualifikationsphase zu ermöglichen (vgl. Niedersächsisches Kultusministerium, 1990). Auch in den Rahmenrichtlinien für das Fach Englisch wird der Ausgleich von Defiziten bei unterschiedlichen Eingangsvoraussetzungen als ein wesentliches Ziel des Fachunterrichts betrachtet (vgl. Niedersächsisches Kultusministerium, 2003).

Die Struktur und Organisation der heutigen gymnasialen Oberstufe geht im Wesentlichen auf die von den Kultusministern der Länder in Bonn beschlossene „Vereinbarung zur Neugestaltung der gymnasialen Oberstufe vom 7. Juli 1972“ zurück. Seit dieser Reform besteht die gymnasiale Oberstufe aus einer Kombination von Pflicht-

¹ Diese Struktur entspricht der längeren Form der gymnasialen Oberstufe. Aufgrund von Kritik an der international vergleichsweise langen Ausbildungszeit existiert aktuell in allen Bundesländern parallel dazu eine verkürzte Form, die nach der Jahrgangsstufe 12 endet. Hierbei kann die Einführungsphase auf die Jahrgangsstufe 10 vorgezogen werden oder auch dem zweiten Halbjahr der Jahrgangsstufe 11 eine Doppelfunktion aus Einführungs- und Qualifikationsphase zukommen. Ergebnisse aus bisherigen Schulleistungsstudien und auch die vorliegende Arbeit beziehen sich auf die traditionelle Form der gymnasialen Oberstufe.

kursen und Wahlmöglichkeiten, mit der den Schülerinnen und Schülern zum einen ein Grundkanon an Allgemeinbildung und Studierfähigkeit vermittelt und zum anderen eine individuelle Spezialisierung ermöglicht werden soll (vgl. KMK, 2006). Aktuelle Reformen fanden vor allem auf Grundlage des „Husumer Beschlusses“ statt und beziehen sich insbesondere auf die stärkere Betonung einer vertieften Allgemeinbildung. Die strukturelle Umsetzung dieser Forderung fand durch die Pflichtbelegung der Kernfächer Deutsch, Fremdsprache und Mathematik bis zum Abitur sowie der Vorgabe, dass zwei davon Abiturprüfungsfächer sein müssen, statt. Ein erfolgreiches Absolvieren der gesamten gymnasialen Oberstufe inkl. Abschlussprüfung führt zum Erwerb der Allgemeinen Hochschulreife als höchste schulische Abschlussqualifikation und somit zur Berechtigung für ein Hochschulstudium (vgl. KMK, 2006). Bezogen auf die gleichaltrige Bevölkerung erreichten in Deutschland im Jahr 2005 knapp 29% der Schülerinnen und Schüler diese Qualifikation. Ein vergleichsweise geringerer Anteil war in Bayern zu verzeichnen (21%), während der Prozentsatz von 36 in Berlin deutlich über dem Bundesdurchschnitt lag (vgl. Niedersächsisches Kultusministerium, 2006).

Die gymnasiale Oberstufe besteht aus einem Kanon aus Pflicht- und Wahlmöglichkeiten. Neben den bereits angesprochenen Pflichtbelegungen der Kernfächer, können Schülerinnen und Schüler im Rahmen der Vereinbarung zur Gestaltung der gymnasialen Oberstufe wählen, ob und auf welchem Kursniveau sie an bestimmten Fächern teilnehmen möchten. So gibt es Kurse mit grundlegenden (Grundkurs) und erhöhten Anforderungen (Leistungskurs). Unterricht im Kurs mit grundlegenden Anforderungen repräsentiert das Lernniveau der gymnasialen Oberstufe unter dem Aspekt einer wissenschaftspropädeutischen Bildung. Im Kurs mit erhöhten Anforderungen werden zusätzlich einzelne Themenbereiche exemplarisch vertieft. Mit dieser Struktur soll den Schülerinnen und Schülern eine individuelle Profilbildung ermöglicht und zugleich eine Sicherung von Standards einer vertieften Allgemeinbildung erzielt werden. An allgemein bildenden Gymnasien ist diese vergleichsweise freier gestaltet als an beruflichen Gymnasien bzw. Fachgymnasien.² Fachgymnasien beinhalten jeweils eine berufsbezogene Schwerpunktsetzung, die in weiteren Pflicht- bzw. Profulfächern entsprechend der

² Die Bezeichnung variiert zwischen den Bundesländern; im Folgenden wird der für Niedersachsen gültige Terminus „Fachgymnasium“ verwendet.

beruflichen Ausrichtung strukturell verankert ist. Somit legen sich Schülerinnen und Schüler durch die Wahl eines Fachgymnasiums frühzeitig auf ein bestimmtes Profil fest. Zurzeit existieren in Deutschland Fachgymnasien mit ganz verschiedenen Schwerpunkten. Diese sind auf die anschließende Berufswelt ausgelegt, haben aber gleichzeitig ihre Entsprechung in universitären Studiengängen. Prominente Vertreter solcher Profile sind beispielsweise das der Wirtschaft oder das der Technik. An Fachgymnasien findet ausschließlich Unterricht der gymnasialen Oberstufe statt. Fachgymnasien sind in der Regel Gymnasien in Aufbauform, an denen Schülerinnen und Schüler im Anschluss an einen Realschulabschluss oder einen als gleichwertig anerkannten Abschluss die allgemeine Hochschulreife erwerben können.

In ihrer aktuellen Form sind Fachgymnasien das Resultat zahl- und umfangreicher Reformprozesse (einen historischen Überblick geben bspw. Brauckmann & Neumann, 2004). Seit der Vereinbarung zur Neugestaltung der gymnasialen Oberstufe 1972 sind Fachgymnasien mit den allgemein bildenden Gymnasien gleichgestellt – an ihnen erworbene Abiturzeugnisse haben den gleichen Stellenwert wie die allgemein bildender Gymnasien, Qualifikationen werden gegenseitig anerkannt.³ Der Unterricht an beiden Schulformen beinhaltet, entsprechend der Vereinbarung zur Gestaltung der gymnasialen Oberstufe, einen festgelegten Pflichtkanon an Grundbildung, mit dem eine breite und vertiefte Allgemeinbildung der Schülerinnen und Schüler und eine allgemeine Studierfähigkeit angestrebt werden (vgl. KMK, 2006). Von allen Schülerinnen und Schülern, die die Allgemeine Hochschulreife im Jahr 2005 erworben haben, erreichten diese im Bundesdurchschnitt 15% an Fachgymnasien. Auch hier zeigen sich deutliche länderspezifische Unterschiede: Den größten Anteil verzeichnete Baden-Württemberg (32%), besonders niedrige Quoten gab es im Saarland (4%). In Niedersachsen erreichten 17% der Abiturientinnen und Abiturienten ihre allgemeine Hochschulreife an Fachgymnasien (Niedersächsisches Kultusministerium, 2006, eigene Berechnungen).

³ Eine notwendige Voraussetzung hierfür ist nach dem am 28. Oktober 1964 verabschiedeten „Hamburger Abkommen“ von den Ministerpräsidenten der Länder, dass Schülerinnen und Schüler, die in den Jahrgangsstufen 7 bis 10 nicht am Unterricht in einer zweiten Fremdsprache teilgenommen haben, beginnend in der Jahrgangsstufe 11 in der gymnasialen Oberstufe am Unterricht in einer zweiten Fremdsprache teilnehmen.

1.2 Erfassung schulischer Leistungen und Leistungsentwicklungen

Der Erwerb der allgemeinen Hochschulreife gilt als die höchste schulische Abschlussqualifikation. Das Abiturzeugnis spiegelt die Leistungen der Schülerinnen und Schüler in der Qualifikationsphase und den Abiturprüfungen in Form von Noten- bzw. Notenpunktprofilen wider. Diesen Profilen kommt insbesondere an Übergängen im Bildungssystem eine entscheidende Funktion zu: Schülerinnen und Schüler an Haupt- oder Realschulen können mit einem gewissen Notenprofil im Abschlusszeugnis den erweiterten Sekundarschulabschluss I erwerben und haben somit die Möglichkeit, direkt in die gymnasiale Oberstufe zu wechseln. Und auch am Ende der gymnasialen Oberstufe ist die Abiturnote mitbestimmend für weiterführende Bildungswege: Je besser, desto mehr Studiengänge stehen den Schülerinnen und Schülern offen. Ebenso haben Notenprofile eine gewisse Relevanz für die Vergabe von Ausbildungsstellen.

Noten sind keine ausschließlichen Leistungsindikatoren, sondern stellen ein breiteres Maß dar, das zahlreiche weitere Merkmale beinhaltet (Archer & McCarty, 1998). So fließen in die Wahrnehmung und Benotung der Lehrkräfte neben kognitiven Aspekten auch motivationale Faktoren, wie Anstrengungsbereitschaft, mit ein. Zudem zeigte Ingenkamp (1995), dass Noten eher sozialnorm- als sachnormorientiert vergeben werden, dass sie also in hohem Maß vom Leistungsspektrum einer Bezugsgruppe (z. B. Schulklasse oder Schule) abhängig sind (vgl. auch Baumert, Trautwein & Artelt, 2003). Objektivität und Vergleichbarkeit von Leistungen in Form von Noten sind somit eingeschränkt. Dieses zeigt sich besonders bei der Erfassung von Leistungsentwicklungen. Aufgrund einer meist sozialnormorientierten Bewertung spiegeln Veränderungen in den Noten in hohem Maße eine geänderte Rangposition innerhalb der Klasse wider.

Eine möglichst objektive Erfassung von Leistungen wird im Kontext der insbesondere im letzten Jahrzehnt durchgeführten Schulleistungstudien angestrebt (z. B. *Programme for International Student Assessment, PISA*; OECD, 2001). Eine damit verbundene wesentliche Zielsetzung ist die bestmögliche Vergleichbarkeit der Ergebnisse zur Evaluation von Bildungssystemen. Die große Mehrheit der diesbezüglich entwickelten und verwendeten Tests basiert auf der *Item Response Theory* (IRT) (z. B. Hambleton & Swaminathan, 1985). Im Rahmen der IRT bildet oftmals das eindimensionale Rasch-Modell die Grundlage der Leistungstests. Im dichotomen Raschmodell wird nur zwi-

schen zwei Antwortformaten (z. B. richtig/falsch) unterschieden. Anhand dieser so eingeteilten Antworten wird auf eine latente kontinuierliche Personenfähigkeit (Personenparameter) sowie die Itemschwierigkeit (Itemparameter) geschlossen. Je höher ein Parameter, desto größer die Fähigkeit einer Person bzw. die Schwierigkeit einer Aufgabe. Je weiter der jeweilige Fähigkeitsparameter den Itemparameter übersteigt, desto größer ist die Wahrscheinlichkeit, dass die Person die Aufgabe löst. Die Beziehung zwischen Personenparameter und Antwortwahrscheinlichkeit pro Item lässt sich als eine Funktion beschreiben und graphisch mit Hilfe einer *Item Characteristic Curve* (ICC) veranschaulichen. Im Wendepunkt einer ICC beträgt die Lösungswahrscheinlichkeit .50, hier wird üblicherweise die Itemschwierigkeit definiert. Um mit hinreichender Sicherheit anzugeben, dass eine Person eine Aufgabe lösen kann, wurde beispielsweise in TIMSS (*Third International Mathematics and Science Study*, Mullis et al., 1998; Baumert et al., 1997; Baumert, Bos & Watermann, 1999; Baumert et al., 2000) hingegen die Itemschwierigkeit bei einer Wahrscheinlichkeit von .65 festgelegt.

Gegenüber der Klassischen Testtheorie (KTT, z. B. Lord & Novick, 1968) hat das Rasch-Modell insbesondere zwei ökonomische Vorteile: Zum einen wird bei letzterem angenommen, dass Itemparameter unabhängig von der Personenstichprobe sind. Das bedeutet, dass Itemschwierigkeiten zwischen Stichproben gleich bleiben und Personenparameter für Stichproben anderer Studien anhand bereits vorliegender, aus hinreichend großen Stichproben ermittelter, Itemparameter geschätzt werden können. Zum anderen lassen sich Personenparameter unabhängig von der Itemstichprobe bestimmen. Somit können Personen, auch wenn sie unterschiedliche Aufgaben bearbeitet haben, auf einer gemeinsamen Fähigkeitsskala abgebildet werden. Dies hat insbesondere dann praktische Relevanz, wenn ein Test sehr umfangreich ist und nicht alle Aufgaben eingesetzt werden können. Für längsschnittliche Erhebungen erlauben *Test Equating*-Prozeduren (vgl. Hambleton & Swaminathan, 1985) im Rahmen der IRT, dass Schätzungen mit unterschiedlichen Items auf einer *Common Metric* abgebildet werden können, sofern die verschiedenen Tests eine Gruppe gemeinsamer Items (Ankeritems) enthalten (*Common Item Design*) (s. a. Lord, 1980). Neben diesen ökonomischen Aspekten, bieten IRT-Modelle entscheidende Vorteile hinsichtlich curricularer, didaktischer oder psychologischer Fragestellungen: Sie erlauben die Verbindung von norm- und kriterienbezogener Interpretation von Testergebnissen und tragen somit zum Verständnis bei, was ein Test-

wert inhaltlich bedeutet. Die kontinuierliche Fähigkeitskala wird dazu in Abschnitte unterteilt. Über Aufgabenanalysen lassen sich anschließend diese Bereiche und somit bestimmte Fähigkeitsniveaus bzw. Testwerte inhaltlich beschreiben (vgl. Watermann & Klieme, 2002).

Im Kontext der Erfassung schulischer Leistungen und Leistungsveränderungen wurden weitere methodische Aspekte diskutiert. Hierzu zählen insbesondere der Umgang mit *Missing Data* und die Berücksichtigung der hierarchischen Struktur der Daten und damit einhergehenden Clustereffekten im Vordergrund. Das Problem fehlender Werte tritt insbesondere in Längsschnittmessungen auf. So stehen Personen aus motivationalen oder organisatorischen Gründen teilweise nicht für alle Erhebungen zur Verfügung. Für Erhebungen in der gymnasialen Oberstufe kommt hinzu, dass einzelne Personen Klassen wiederholen oder in eine berufliche Ausbildung wechseln und somit vorzeitig aus der Schule ausscheiden. Mit fehlenden Werten gehen jedoch bestimmte Probleme einher (vgl. Lüdtke, Robitzsch, Trautwein & Köller, 2007). So liegt aufgrund der eingeschränkten Stichprobengröße ein Verlust an Effizienz und statistischer Power vor. Zum anderen setzen statistische Standardverfahren vollständige Datenmatrizen voraus, sodass der Umgang mit den Daten erschwert ist. Und drittens besteht, sofern sich beobachtete und fehlende Daten systematisch voneinander unterscheiden, die Gefahr verzerrter Parameterschätzungen. Zum Umgang mit fehlenden Werten werden oftmals klassische Verfahren, wie *Pairwise* oder *Listwise Deletion*, eingesetzt. Diese können jedoch die mit fehlenden Werten einhergehenden Probleme meist nicht zufrieden stellend angehen, sodass alternativ verschiedene moderne Verfahren zum Ersetzen fehlender Werte vorgeschlagen wurden (vgl. Collins, Schafer & Kam, 2001). So zeigte sich in Simulationsstudien, dass der *Expectation-Maximization*-Algorithmus und *Multiple Imputation* im Vergleich zu klassischen Verfahren insbesondere bei einer Ausfallquote über 5% zu weniger verzerrten Schätzungen führen können (z. B. Schafer & Graham, 2002).

Studien im Rahmen der empirischen Bildungsforschung beziehen sich in der Regel nicht auf Zufallsstichproben, sondern auf bereits bestehende Einheiten (Cluster), nämlich Schülerinnen und Schülern in bestimmten Klassen, die wiederum bestimmten Schulen zugeordnet werden können. Mit dieser hierarchischen Datenstruktur geht oftmals einher, dass sich Schülerinnen und Schüler innerhalb von Klassen und Schulen

einander ähnlicher sind als eine zufällig zusammengesetzte Stichprobe. Inferenzstatistische Analysen setzen zur Berechnung des Standardfehlers (als Maß für die Ungenauigkeit der geschätzten Kennwerte) unabhängige Beobachtungen voraus. Bei hierarchischer Datenstruktur wird somit der Standardfehler unterschätzt, sodass Signifikanzprüfungen tendenziell zu liberal angesetzt werden. Zum Umgang mit Clustereffekten existieren verschiedene Ansätze. Einige setzen an einer Korrektur der Standardfehler an (z. B. die sogenannten *Jackknife*-Methoden; Wolter, 1985). Bei einem anderen Verfahren, dem *Hierarchical Linear Modeling* (HLM, z. B. Bryk & Raudenbush, 1989), wird die hierarchische Struktur der Daten direkt abgebildet, wodurch nicht nur Standardfehler korrigiert werden, sondern sich dabei auch simultan Effekte auf Individual- und Clusterebene beschreiben lassen.

Neben der Bearbeitung methodischer Fragestellungen fanden ebenso intensive inhaltliche Diskussionen zwischen Erziehungswissenschaften, Fachdidaktiken und Psychologie über die Erfassung von Leistungen statt. Exemplarisch sei hier die Debatte zur Testkonzeption auf Basis lehrplanbasierter Modelle der Wissensüberprüfung versus eines *Literacy*-Konzepts angeführt. Unter Literacy werden Fähigkeiten verstanden, die notwendig sind, um auf die Herausforderung des Berufslebens und der mündigen Teilhabe am gesellschaftlichen Leben vorbereitet zu sein (vgl. Baumert, Artelt et al., 2003). Bei den auf diesem Konzept basierenden Tests wurde jedoch in Frage gestellt, ob sich mit diesen Testergebnissen ein realistisches Bild des Leistungsstands von Schülerinnen und Schülern abbilden lässt. Zur Beantwortung fanden beispielsweise differenzierte Analysen der PISA-Testaufgaben statt. Hierbei wurden nur Aufgaben in die Auswertung mit einbezogen, die von Lehrplanexpertinnen und -experten des jeweiligen Fachs in einem jeweiligen Bundesland als curricular valide eingestuft wurden. Es zeigte sich, dass Ergebnisse dieser lehrplanoptimierten Ländertests vergleichbar waren mit den ursprünglichen PISA-Testergebnissen. Lehrplanvalide bzw. nicht lehrplanvalide Literacy-Aufgaben wurden somit von den Schülerinnen und Schülern ähnlich gut gelöst (ebd.).

1.3 Leistungsentwicklungen in Mathematik und Englisch in der gymnasialen Oberstufe

Der im Rahmen von TIMSS entwickelte Test zur mathematischen Grundbildung basiert auf dem Literacy-Konzept und lässt sich dabei insgesamt als ein „Kompromiss zwischen Anwendungsorientierung und curricularer Anbindung“ (vgl. Baumert, Stanat & Demmrich, 2001, S. 19) beschreiben. Die Testaufgaben beziehen sich auf zentrale Inhalte des mathematischen Unterrichts der Mittelstufe, sollten also von Schülerinnen und Schülern ab dem Ende der Mittelstufe bearbeitet werden können und werden als Basisqualifikation für die Oberstufenmathematik im Allgemeinen angesehen (vgl. Watermann, Nagy & Köller, 2004). Die empirische Befundlage zur Entwicklung der mathematischen Grundbildung in der gymnasialen Oberstufe ist uneinheitlich. So lassen sich in einer Querschnittstudie aus der Schweiz einerseits im Mittel für Berufsbildende Schülerinnen und Schüler keine Veränderungen der Grundbildung im Verlauf der gymnasialen Oberstufe feststellen, für diejenigen an allgemein bildenden Gymnasien hingegen ein Zuwachs verzeichnen (vgl. Ramseier, Keller & Moser, 1999). In den Hamburger Untersuchungen *Aspekte der Lernausgangslage und der Lernentwicklung* (LAU 11, Lehmann, Hunger, Ivanov & Gänsfuß, 2004; LAU 13, Lehmann, Vieluf, Nikolova & Ivanov, 2006) ließ sich hingegen ein Leistungszuwachs über alle betrachteten gymnasialen Schulformen hinweg aufzeigen. In Englisch belegten empirische Studien mit standardisierten Tests überwiegend Leistungsverbesserungen (vgl. LAU 11 und 13, Lehmann et al., 2004, 2006; *Bildungsprozesse und psychosoziale Entwicklung im Jugendalter und jungen Erwachsenenalter*, BIJU, Baumert et al., 1996). Eine Ausnahme in der BIJU-Studie bildeten Schülerinnen und Schüler an Integrierten Gesamtschulen. Hier konnte bei Schülerinnen und Schülern von der 12. zur 13. Jahrgangsstufe keine Leistungsverbesserung nachgewiesen werden (vgl. Köller, Baumert, Cortina, Trautwein & Watermann, 2004).

In Hinblick auf die strukturelle Ausgestaltung der gymnasialen Oberstufe in Form von unterschiedlichen Kursniveaus und damit einhergehenden spezifischen Zielsetzungen, liegen neben diesen mittleren Leistungsverläufen auch differentielle Entwicklungen in Abhängigkeit vom besuchten Kursniveau nahe. So sollen laut der Vereinbarung zur Gestaltung der gymnasialen Oberstufe (KMK, 2006) in Kursen mit erhöhten Anfor-

derungen einzelne Bereiche exemplarisch vertieft werden. Mit einem erhöhten Kursniveau geht in der Regel ein kognitiv anspruchsvollerer Unterricht in dem Fach einher, der zu einem tieferen Verständnis der Inhalte beitragen und somit positive Effekte auf den Wissenserwerb haben sollte. Dabei ist anzunehmen, dass sich Leistungsunterschiede zwischen Schülerinnen und Schülern unterschiedlicher Kursniveaus zum einen auf unterschiedliche Eingangsvoraussetzungen zu Beginn der Kursstufe und zum anderen auf weitere Leistungsdifferenzierungen während der Kursstufe zurückführen lassen (vgl. Köller et al., 2006). In empirischen Studien am Ende der gymnasialen Oberstufe zeigten sich sowohl in Mathematik (z.B. Baumert & Watermann, 2000; Watermann et al., 2004) als auch in Englisch (z.B. Köller & Trautwein, 2004; Köller, Baumert et al., 2004) deutliche mittlere Leistungsunterschiede in Abhängigkeit vom Kursniveau.

2. Die Bedeutung beruflicher Interessen für schulische Leistungen

2.1 Die Person-Gegenstands-Konzeption des Interesses

Ein seit langem in der Pädagogischen Psychologie diskutiertes und breit erforschtes Feld betrifft die Einflussfaktoren schulischer Leistungen und Leistungsentwicklungen. Für erfolgreiches schulisches Lernen wurden in den vergangenen Jahrzehnten verschiedene Modelle formuliert und in ihnen individuelle, schulische und außerschulische Faktoren als direkte und/oder indirekte Determinanten des Lernerfolgs ausgewiesen (einen Überblick geben z.B. Helmke & Weinert, 1997). In neuerer Zeit wurde jedoch vom Versuch, umfassende Modelle schulischen Lernens aufzustellen, weitgehend Abstand genommen. Vielmehr bilden bereichsspezifische Theorien, in denen einzelne Determinanten der Schulleistung untersucht werden, die Grundlage empirischer Studien.

Im Bereich individueller Faktoren gehören Interessen zu den empirisch umfangreich erforschten und vielfach bestätigten Determinanten schulischen Lernens (z.B. Krapp, 1998, 2002). Davon ausgehend, dass sich Individuum und Umwelt in einem ständigen Austauschprozess befinden, wurde im letzten Quartal des 20. Jahrhunderts eine interessenbezogene Forschung von der „Münchener Gruppe“ (z.B. Krapp & Prenzel, 1992; Prenzel, Krapp & H. Schiefele, 1986) vorangetrieben und dabei das Interessenkonstrukt aus Sicht einer Person-Gegenstands-Konzeption beschrieben (vgl. Krapp, 1992). Zentral ist dabei die Gegenstandsspezifität des Interesses (Krapp, 2002). Es wird als eine durch bestimmte Merkmale herausgehobene „Beziehung“ einer Person zu einem Gegenstand bezeichnet (z.B. Krapp, 1992). Diese Gegenstände können bestimmte Objekte, Tätigkeiten oder auch Themengebiete sein, die als individuelle Sinn- und Bedeutungseinheiten kognitiv repräsentiert sind. Somit wird angenommen, dass die Person über ein mehr oder weniger ausdifferenziertes gegenstandsspezifisches Wissen verfügt (Krapp, 1998). Jeder Gegenstand kann für verschiedene Personen ganz unterschiedliche Bedeutung haben (persönliche Valenz, vgl. Pekrun, 1988). Die interessen-thematische Person-Gegenstands-Beziehung wird von zwei positiven emotionalen Zuständen begleitet: einer emotionalen und einer wertbezogenen Valenz (z.B. Krapp, 1992). Der emotionale Bezug im Kontext interessen-thematischer Auseinandersetzungen äußert sich in der Verknüpfung eines Gegenstands bzw. der auf ihn bezogenen Hand-

lungen mit positiven Gefühlen; der wertbezogene Aspekt lässt sich als hohe subjektive Wertschätzung des Interessengegenstands sowie einer Identifikation der Person mit dem Gegenstand des Interesses beschreiben (ebd.). Eine empirische Trennung beider Komponenten ist bislang jedoch nicht gelungen (vgl. Köller, Baumert et al., 2000). Die mit diesem Interessenkonstrukt verbundene pädagogisch-psychologische Forschung lässt sich in zwei Traditionen einteilen (Krapp, 1998). Aus der Perspektive aktueller Zustände werden vorrangig situationsspezifische Prozesse untersucht und Auslöse- und Entstehungsbedingungen für Interessen im Verlauf von Interessenhandlungen analysiert. Dazu kann entweder situationsspezifisches, durch aktuelle Anregungsbedingungen (Interessanztheit) entstandenes Interesse oder die Aktualisierung eines bestehenden individuellen (dispositionalen) Interesses untersucht werden (Krapp, 1992). Letzteres steht im Fokus einer zweiten Forschungstradition. Hier werden insbesondere thematische Interessen (bspw. ein Schulfach) im Kontext situationsübergreifender interindividueller Interessenunterschiede betrachtet und diese Dispositionen oftmals im Zusammenhang mit schulischem oder akademischem Lernen betrachtet. Dabei wird angenommen, dass das Interesse an einem Gegenstand eine handlungsleitende Funktion ausübt: Man möchte mehr über den Interessengegenstand erfahren und beschäftigt sich intensiver damit (Prenzel, 1988). Durch die Auseinandersetzung mit dem Gegenstand bildet sich das gegenstandsspezifische Wissen heraus (vgl. Krapp, 1992). Diese Zusammenhänge bestätigend fanden Schiefele, Krapp und Schreyer (1993) in einer Metaanalyse Korrelationen zwischen verschiedenen Indikatoren von Interesse und Leistungen in Höhe von $r = .28$ in Mathematik bzw. von $r = .33$ in Fremdsprachen. Deutlich niedriger fielen hingegen die Zusammenhänge im Fach Biologie ($r = .16$) und Literatur/Deutsch aus ($r = .14$).

Statt korrelativer Zusammenhänge werden in Erwartungs-mal-Wert-Modellen gerichtete Beziehungen spezifiziert, in denen Verhalten durch die subjektive Erwartung und den subjektiven Wert von Verhaltenskonsequenzen erklärt wird. Im Erwartungs-mal-Wert-Modell von Eccles[-Parsons] et al. (1983) lässt sich Interesse der Wert-Komponente zuordnen (z. B. Wigfield & Eccles, 2000). Dabei wird angenommen, dass das Interesse einen Effekt auf leistungsbezogenes Wahlverhalten sowie die Investition in eine Aufgabe und somit auf nachfolgende Leistung ausübt. Die Befundlage zu den Richtungen der Effekte zwischen Interesse und Leistungen ist uneinheitlich. Für Schülerinnen und Schüler der Sekundarstufe I belegten Marsh, Trautwein, Lüdtke,

Köller und Baumert (2005) in einem reziproken Effektmodell ausschließlich einen Effekt von Interesse auf Testleistungen in Mathematik, nicht aber umgekehrt. Umgekehrte Effektrichtungen zeigten hingegen Köller, Baumert et al. (2000) auf: In der Sekundarstufe I waren unter Kontrolle der Ausgangsleistung nur die Effekte von Mathematikleistungen auf nachfolgende Interessen bedeutsam. Gleichzeitig verwiesen die Autoren auf die Bedeutung des institutionellen Kontextes für gerichtete Effekte. In ihrer Studie ließ sich erst in der Sekundarstufe II ein Effekt des Interesses auf nachfolgende Leistungen bestätigen. Die Autoren argumentieren, dass erst in diesem institutionellen Schulabschnitt die Möglichkeit für Kurswahlen und somit für interessengeleitetes Lernen gegeben ist. Und auch Köller et al. (2006) wiesen anhand einer Längsschnittstudie in der 10. und 12. Jahrgangsstufe einen Effekt von Interesse auf Leistungen nach.

In der Metaanalyse von Schiefele et al. (1993) wurde zwischen Interesse und Leistungen in Mathematik eine ähnlich hohe Korrelation nachgewiesen wie für Interesse und Fremdsprachenleistungen. Es kann somit von einer vergleichbaren Bedeutung des Interesses für beide Bereiche schulischer Leistungen ausgegangen werden. Dennoch gibt es bislang kaum Längsschnittstudien zur Bedeutung von Interessen für Leistungsentwicklungen in Fremdsprachen. In einer neueren Studie zeigte Pae (2008) einen Effekt von intrinsischer Motivation auf Fremdsprachenleistungen von koreanischen Studenten. Abu-Rabia (2003) belegte einen Effekt von Interesse auf Englischleistungen als dritte Fremdsprache bei jugendlichen arabischen Schülerinnen und Schülern.

2.2 Die Theorie der Berufswahl nach Holland

Beim fachlichen Interesse ist im Rahmen einer Person-Gegenstands-Konzeption der Gegenstand primär durch die Inhalte oder Wissensgebiete eines Schulfachs definiert. „Im Prinzip können aber ebenso bestimmte Tätigkeitsklassen oder konkrete Dinge im Vordergrund eines Interesses stehen und den primären Bezug zu einem interessen-thematischen Gegenstand herstellen“ (vgl. Krapp, 1998, S. 186). Mit einer primären Ausrichtung auf Tätigkeiten haben berufliche Interessen eher einen praktischen Anwendungsbezug als fachliche Interessen (vgl. Bergmann & Eder, 1998). Des Weiteren wird angenommen, dass jede Person über ein mehr oder weniger ausdifferenziertes Wissen über die gesamte Berufslandschaft verfügt, berufliche Interessen also im Vergleich zu

thematischen Interessen nicht auf ein Themengebiet beschränkt, sondern umfassender kognitiv repräsentiert sind. Aufgrund dieser Repräsentationsart wurden Theorien formuliert, in denen meist die gesamte Berufslandschaft in verschiedene Berufsfelder bzw. berufliche Tätigkeitsklassen unterteilt wird (Rolf, 2001).⁴ In Anlehnung an die Person-Gegenstands-Konzeption (z. B. Krapp, 1992) lassen sich diese Klassen als Sinn- und Bedeutungseinheiten verstehen. Ihr Verhältnis zueinander wird oftmals in Strukturmodellen abgebildet (vgl. z. B. Savickas & Spokane, 1999). Hierbei interessieren zum einen das Ausmaß des Interesses an den einzelnen Klassen beruflicher Tätigkeiten, vor allem aber die individuellen Assoziationen zwischen den Tätigkeitsklassen, also die persönlichen Interessenkonfigurationen bzw. -profile.

2.2.1 Kernannahmen in Hollands Berufswahltheorie

Holland spezifiziert in seiner Theorie der Berufswahl (1997) verschiedene Klassen beruflicher Tätigkeiten. Seine erstmalig im Jahr 1959 veröffentlichte Theorie geht im Wesentlichen auf zwei Quellen zurück: die von Spranger (1913) verwendete typologische Beschreibung der menschlichen Persönlichkeit (der theoretische, ökonomische, ästhetische, soziale, machtorientierte oder religiöse Mensch) sowie die von Guilford, Christensen, Bond und Sutton (1954) identifizierten Faktoren menschlicher Interessen (Mechanical, Scientific, Aesthetic, Social Welfare, Business, Clerical). Vier Arbeitsthesen bilden den Kern der Theorie (vgl. Holland, 1997):

1. Die meisten Personen im westlichen Kulturkreis lassen sich einem von sechs Persönlichkeitstypen zuordnen: Realistic, Investigative, Artistic, Social, Enterprising oder Conventional (für eine nähere Beschreibung s. Tab. 1).

⁴ Krapp (2001) merkt an, dass die im Rahmen der Berufsinteressenforschung entwickelten Interessenerhebungsinstrumente häufig nur Einstellungen gegenüber sehr allgemein definierten Handlungsfeldern und Wissensgebieten erfassen und somit dem Prinzip der Gegenstandsspezifität nicht gerecht werden. In letzter Zeit wurde insbesondere im deutschsprachigen Raum eine engere Anbindung an die Münchner Interessentheorie angestrebt und berufliche Tätigkeitsklassen formuliert, die sich an der Definition eines Gegenstandsbezugs (Krapp, 1992) orientieren und sich aus Sicht einer Person-Gegenstands-Theorie beschreiben lassen (vgl. Bergmann & Eder, 2005).

2. Auch Umwelten lassen sich in diese sechs Bereiche einteilen und entsprechend beschreiben.
3. Personen wählen Umwelten aus, in denen sie ihre Fähigkeiten und Fertigkeiten trainieren sowie ihre Einstellungen und Werte ausdrücken können.
4. Das Verhalten einer Person ist durch ihre Interaktion mit der Umwelt bestimmt.

Hollands Annahmen führten während der 1970er und 1980er zu zahlreichen Studien, sodass sie sich inzwischen als eine der am sichtbarsten und am meisten beforschten Theorien etabliert hat (vgl. Hackett, Lent & Greenhaus, 1991; Savickas & Spokane, 1999). Insbesondere wurde das Verhalten von Personen in Abhängigkeit von ihren beruflichen Interessen sowie entsprechende Konsequenzen untersucht. So zeigte sich in Studien zu Studienfachwahlen (z.B. Nagy, 2006), dass Personen mit ihren Interessen übereinstimmende (kongruente) Umwelten aufsuchen. Selbstselektion und Sozialisation aufgrund einer Person-Umwelt-Interaktion führen wiederum dazu, dass Interessen von Schülerinnen und Schülern zunehmende Kongruenz mit den Merkmalen der Umwelt aufweisen (*Person-Environment Fit*). Eder (1988) belegte eine Übereinstimmung zwischen den Interessenorientierungen und der fachlichen Ausrichtung des gymnasialen Schultyps. Im Rahmen der TOSCA-Studie (*Transformation des Sekundarschulsystems und akademische Karrieren*; Köller, Watermann, Trautwein & Lüdtke, 2004) zeigte sich bspw., dass Schülerinnen und Schüler an Fachgymnasien mit Schwerpunkt Wirtschaft vergleichsweise hohe Interessen im Bereich Conventional angeben und solche in Fachgymnasien mit Schwerpunkt Sozialpädagogik vorrangig soziale Interessen haben. Dagegen lässt sich an allgemein bildenden Gymnasien im Durchschnitt über alle Interessenorientierungen hinweg ein relativ ausgeglichenes Profil in mittlerer Höhe aufzeigen (vgl. Lüdtke & Trautwein, 2004). Spokane, Meir und Catalano (2000) belegten in einer Metaanalyse, dass eine hohe Übereinstimmung zwischen Person und Umwelt in Zusammenhang mit Zufriedenheit, Leistungsmerkmalen sowie Stabilität der Interessen steht. Auch Eder und Reiter (2002) zeigten, dass eine höhere Person-Umwelt-Kongruenz tendenziell mit mehr Wohlbefinden und besseren schulischen Leistungen einhergeht.

Tabelle 1

Charakterisierung der sechs Orientierungen nach Holland (vgl. Bergmann & Eder, 1992)

Orientierung	Fähigkeiten und Fertigkeiten	Berufsbeispiele
Realistic (R)	mechanisch, technisch	Techniker/-in, Ingenieur/-in
Investigative (I)	mathematisch, naturwissenschaftlich	Chemiker/-in
Artistic (A)	sprachlich, gestalterisch, musisch	Autor/-in, Schauspieler/-in
Social (S)	sozial, zwischenmenschlich	Lehrer/-in, Therapeut/-in
Enterprising (E)	zwischenmenschlich, manipulativ	Verkäufer/-in, Manager/-in
Conventional (C)	rechnerisch, verwaltend	Banker/-in, Beamter/Beamtin

Im Kontext einer Person-Umwelt-Interaktion lassen sich auch die im Bereich beruflicher Interessen oftmals gefundenen deutlichen geschlechtsspezifischen Unterschiede interpretieren, die laut Gottfredson (1981) im Wesentlichen auf den Erwerb von Geschlechterrollen zurückgehen. In Hinblick auf berufliche Interessen lässt sich zeigen, dass es geschlechtsspezifische Differenzen insbesondere bei zwei Interessenorientierungen gibt: So haben in der Regel männliche Personen höhere Interessen im Bereich Realistic und weibliche Personen tendieren eher zur Interessenorientierung Social (z. B. Lippa, 1998; Mullis, Mullis & Gerwels, 1998; Nagy, 2006; Nagy, Trautwein & Lüdtko, in press; Tracey & Ward, 1998).

2.2.2 Charakteristika beruflicher Interessenprofile sensu Holland

Da im Rahmen beruflicher Interessentheorien insbesondere individuelle Interessenkonfigurationen im Vordergrund stehen, beschränkt man sich in der Regel nicht auf die Bestimmung eines Interessentyps, sondern ermittelt individuelle *Interessenprofile*. Nach Holland können solche Profile aus zwei bis sechs der von ihm spezifizierten Interessenvariablen bestehen. Zur Beschreibung von Interessenprofilen lassen sich wiederum drei zentrale Charakteristika heranziehen: Orientierung, Differenziertheit und Konsistenz. Die Interessenorientierung – ein zentraler Aspekt in Hollands Theorie –

entspricht derjenigen Tätigkeitsklasse, die eine Person am meisten präferiert. Die Differenziertheit ist ein Maß dafür, wie weit sich die von einer Person präferierte Interessenorientierung von den anderen Orientierungen abhebt bzw. wie eindeutig die Zuordnung zu einer bestimmten Interessenorientierung ist. Es gibt verschiedene Indizes zur Beschreibung der Differenziertheit. Holland (1997) verwendet dazu die Differenz zwischen dem höchsten und dem niedrigsten Skalenwert. Healy und Mourton (1983) schlagen mit der Streuung der Skalenwerte ein alle sechs Orientierungen umfassendes Maß vor. Die Ausdifferenzierung beruflicher Interessen wird als eine zentrale Entwicklungsaufgabe betrachtet (vgl. Eder, 1998). In empirischen Studien zeigte sich, dass Differenziertheit vor allem mit der Stabilität von beruflichen Wahlentscheidungen assoziiert ist (vgl. ebd.; Taylor, Kelso, Longthorp & Pattison, 1980). Konsistenz bezieht sich wiederum auf die dem Modell von Holland zugrunde liegende Struktur der sechs Interessenorientierungen. So lässt sich Hollands Theorie als ein Strukturmodell verstehen. Dieses lässt sich in Form eines Hexagons darstellen, bei dem die der Reihe nach angeordneten Orientierungen R-I-A-S-E-C die Eckpunkte bilden. Die räumliche Nähe der sechs Typen bildet deren inhaltliche Ähnlichkeit zueinander ab. Benachbarte Orientierungen (z.B. R-I) sind also einander ähnlich, indirekt benachbarte Orientierungen (z.B. R-A) eher unähnlich und gegenüberliegende Orientierungen (z.B. R-S) bilden sozusagen die Pole einer Dimension. Ein konsistentes Profil (bspw. vorrangig Interessen bei A-S, kaum Interesse an C) ist mit dieser Struktur vereinbar. Zum Erfassen von Konsistenz lassen sich erneut verschiedene Maße heranziehen (vgl. Holland, 1997; Strahan, 1987). Personen mit konsistenten Profilen zeichnen sich durch eine hohe Stabilität in ihren Karrierewegen aus (z.B. Holland, Gottfredson & Baker, 1990).

Die hexagonale Zusammenhangsstruktur beruflicher Interessen nach Holland wird als *Calculus*-Annahme bezeichnet; sie lässt sich über Korrelations- oder Kovarianzmatrizen erschließen und statistisch prüfen. Hierbei lässt sich zwischen einem perfekten *Circumplex* und einem *Quasi-Circumplex* bzw. *Circular-Order-Modell* unterscheiden (Guttman, 1954). Die von Holland verwendete Darstellung eines equilateralen Sechsecks (gleichmäßige Abstände zwischen den Orientierungen und gleiche Abstände der Orientierungen vom Figurmittelpunkt) entspricht einem perfekten *Circumplex*. Dessen Überprüfung geschieht in der Regel mit Hilfe konfirmatorischer Methoden (vgl. z.B. Gurtman & Pincus, 2003). In einer neueren Arbeit verwendete Nagy (2006) beispiels-

weise eine konfirmatorische Faktorenanalyse, in der das von Browne (1992) entwickelte Verfahren zur Modellierung von Circumplex-Strukturen weiterentwickelt wurde. Eine Prüfung des Circular-Order-Modells kann über (weniger restriktive) konfirmatorische Methoden oder auch mit dem *Randomization Test of Hypothesized Order Relationships* nach Hubert und Arabie (1987) erfolgen. Bei einem solchen Modell wird nicht von einem gleichmäßigen Hexagon ausgegangen, sondern von einem *misshapen polygon* (Holland & Gottfredson, 1992), bei dem sich die Variablen zwar in der angenommenen Reihenfolge, aber nicht gleichmäßig anordnen lassen. Zur Bildung und Interpretation der einzelnen Kennwerte beruflicher Interessen ist die Bestätigung eines Circular-Order-Modells eine notwendige Voraussetzung (vgl. Lent, Tracey, Brown, Soresi & Nota, 2006). Eine Vielzahl empirischer Arbeiten stützt die Annahme einer hexagonalen Struktur beruflicher Interessen (vgl. Meta-Analyse von Rounds & Tracey, 1993; Tracey & Rounds, 1993). Und auch für den deutschsprachigen Raum konnten Nagy et al. (in press) anhand von Strukturanalysen auf Basis des Allgemeinen-Interessen-Struktur-Tests (Bergmann & Eder, 2005), mit dem berufliche Interessen nach Hollands Modell erfasst werden sollen, die hexagonale Struktur validieren.

Dimensionalitätsanalysen zu Ergebnissen von Interesseninventaren, mit denen Hollands Modell erfasst werden soll, ergeben meist eine dreifaktorielle Lösung (vgl. Prediger, 1982; Prediger & Vansickle, 1992). Zum einen ein Generalfaktor, auf dem alle Orientierungen hoch laden. Zum anderen zwei orthogonale Achsen, die dem hexagonalen Modell zugrunde liegen und von Prediger als *Things/People* und *Data/Ideas*-Achsen bezeichnet wurden. Der Generalfaktor spiegelt interindividuelle Unterschiede in der grundsätzlichen Zustimmung zu den sechs beruflichen Interessenorientierungen wider und wird bspw. als *Level*, Antwortstil (vgl. Prediger, 1998) oder als Generalfaktor (Darcy & Tracey, 2003) bezeichnet. Letztere wiesen auf die zentrale Bedeutung des Levels in der Erfassung beruflicher Interessen hin. Sofern er als Persönlichkeitsstil (z. B. Antwortstil) substantielle Bedeutung habe, müsse er inhaltlich genauer beschrieben und in aktuelle Modelle und Theorien integriert werden. Sofern er die Erfassung beruflicher Interessen verfälsche, müsse er hingegen kontrolliert werden (vgl. ebd., S. 224).

Anhand der individuellen Ausprägungen aller sechs Interessenorientierungen lassen sich unter Berücksichtigung der hexagonalen Struktur die Positionen auf den beiden zugrunde liegenden orthogonalen Dimensionen berechnen und somit Personen im Kreis

positionieren. Über die Lokalisation auf den beiden Achsen ist es möglich, die individuelle Profilorientierung zu bestimmen, die sich als eine bestimmte Orientierung (z. B. R-Typ) oder auch als Mischtyp benachbarter Interessenorientierungen charakterisieren lässt (z. B. R–I-Typ). Im Rahmen der dimensional Positionierung ist der Abstand vom Mittelpunkt gleichzeitig ein Maß für die Differenziertheit eines Profils. Zusammen mit dem Level bilden die mit der Differenziertheit gewichteten Interessenorientierungen die zentralen Komponenten eines individuellen Profils, eines sogenannten *Structural Summary* nach Gurtman und Balakrishnan (1998). Mit diesem lassen sich berufliche Interessen umfassend (inkl. aller sechs Orientierungen) und unter Beachtung der hexagonalen Struktur darstellen sowie gleichzeitig sparsam (mit Hilfe zweier orthogonaler Achsen) modellieren. Durch das Einbeziehen des Levels werden darüber hinaus mögliche Antworttendenzen kontrolliert. Es mangelt bislang jedoch an Studien, in denen die Validität einer Reduktion von typologischen Profilen auf ein Structural Summary untersucht wird.

2.2.3 Berufliche Interessenprofile im Zusammenhang mit schulischen Leistungen

Wie im Abschnitt 2.2.1 dargestellt, basiert Hollands Theorie der Berufswahl auf vier Kernannahmen, die breite empirische Bestätigung erfahren haben. Neben der Untersuchung von Person-Umwelt-Kongruenz und ihrer Konsequenzen, liegt ein weiterer Schwerpunkt auf der Analyse des mit individuellen beruflichen Interessen einhergehenden Verhaltens. Holland (1997) sieht mit den Interessenorientierungen jeweils unterschiedliche Werte, Einstellungen und Fähigkeiten verbunden. In der *Prozess-, Persönlichkeits-, Interessen- und Wissenstheorie der intellektuellen Entwicklung* (*Process, Personality, Interests, and Knowledge*, Ackerman, 1996) wird u. a. angenommen, dass sich (berufliche) Interessen und Fähigkeiten in einer reziproken Beziehung zueinander entwickeln. Auf der einen Seite erhöhen sich bestimmte Fähigkeiten, weil das domänenspezifische Interesse eine verstärkte Auseinandersetzung mit dem Interessengegenstand bedingt. Andererseits verringern sich bestimmte Interessen, wenn ohne ausreichende Fähigkeiten wiederholte Misserfolgserfahrungen gemacht werden.

Wie im vorangegangenen Abschnitt dargestellt, lassen sich individuelle Interessenprofile über die Ausprägung der einzelnen Orientierungen oder auch sparsamer über

eine dimensionale Modellierung in Form eines Structural Summary nach Gurtman und Balakrishnan (1998) modellieren. Empirische Arbeiten, die den Zusammenhang zwischen beruflichen Interessen und Leistungen untersuchten, beziehen mehrheitlich einzelne Skalenwerte beruflicher Interessen in die Analysen mit ein. Die so ermittelten Assoziationen haben Ackerman und Heggestad (1997) in einer Metaanalyse zusammengefasst: Personen mit vorherrschenden Interessen im Bereich Investigative und Realistic weisen vergleichsweise gute Fähigkeiten im mathematischen Bereich auf, während spezifische Stärken von Personen mit hohen Investigative- und Artistic-Interessen im sprachlichen Bereich liegen. Für den deutschsprachigen Raum belegten Bergmann und Eder (1992) positive Zusammenhänge zwischen Interessenskalen im Bereich Investigative mit der Mathematiknote sowie Artistic-Interessen mit der Englisch- bzw. Deutschnote. Im Längsschnitt wurden bislang keine Effekte beruflicher Interessenskalen auf nachfolgende Testleistungen in den Fächern Mathematik und Englisch gefunden (Tracey et al., 2005).

Skalenwerte enthalten jedoch aufgrund ihrer Zusammenhangsstruktur redundante Informationen. Insbesondere in Hinblick auf die Vorhersage von Leistungen und Leistungsentwicklungen kann dies mit statistischen Problemen einhergehen. In Regressionsanalysen bei kleiner bis mittlerer Stichprobengröße mit einer Vielzahl an Variablen kann es zur Inflation des aufgeklärten Varianzanteils kommen (z.B. Wherry, 1931). Zum anderen birgt das simultane Einbeziehen von teils stark miteinander korrelierten Interessenskalen als Prädiktoren in multiplen Regressionsanalysen die Gefahr von Multikollinearität. Die Regressionskoeffizienten bzw. die Effekte einzelner Skalen auf Leistungen lassen sich dann nicht unverzerrt berechnen und im Extremfall ist die Regressionsanalyse selbst nicht mehr durchführbar.

Hollands Modell liegen jedoch zwei orthogonale Dimensionen (Things/People, Data/Ideas) zugrunde, anhand derer eine sparsamere Modellierung beruflicher Interessen möglich ist. Arbeiten, die sich auf Assoziationen zwischen dimensionalen Interessenorientierungen und Leistungen stützen, sind jedoch vergleichsweise rar. In einer aktuellen Querschnittstudie belegte Nagy (2006) Zusammenhänge zwischen Realistic- und Investigative-Orientierungen mit Testleistungen in Mathematik sowie einer Investigative und Artistic-Orientierung mit Englischleistungen. Längsschnittliche Analysen liegen bislang nicht vor.

2.3 Die Bedeutung beruflicher Interessen für den Zusammenhang von Selbstkonzept und Kursniveau mit schulischen Leistungsentwicklungen

In Hinblick auf die Vorhersage schulischer Leistungsentwicklungen in der gymnasialen Oberstufe lassen sich insbesondere das fachspezifische Selbstkonzept und das Kursniveau als weitere wichtige Determinanten einbeziehen, von denen gleichzeitig angenommen wird, dass sie in engem Zusammenhang mit Interessen stehen. Der Zusammenhang zwischen Kursniveau mit schulischen Leistungen und Leistungsentwicklungen in der gymnasialen Oberstufe wurde bereits im 1. Kapitel beschrieben und dazu empirische Befunde berichtet, in denen Schülerinnen und Schüler am Ende der gymnasialen Oberstufe in Kursen mit erhöhten Anforderungen bessere Leistungen zeigten als solche in Kursen mit grundlegenden Anforderungen (z.B. Baumert & Watermann, 2000; Köller & Trautwein, 2004; Watermann et al., 2004). Die deutlichen Unterschiede zwischen den Kursniveaus am Ende der gymnasialen Oberstufe lassen sich auf zunehmende Leistungsdivergenzen während der Kursstufe zurückführen (vgl. Köller et al., 2006). Hierfür spielen Interessen eine wichtige Rolle: Zum einen ist das Interesse ein entscheidendes Kriterium für die Kurswahlen der Schülerinnen und Schüler (vgl. z.B. Hodapp & Mißler, 1996; Köller, Daniels et al., 2000; Köller et al., 2006). Zum anderen sollte aus Sicht der Münchner Interessentheorie (z.B. Krapp, 1992) eine vertiefte Auseinandersetzung mit einem Gegenstand, so wie es in Kursen mit erhöhtem Anforderungsniveau zu erwarten ist, mit Kompetenzerleben einhergehen und förderlich für das Interesse an diesem Fach sein. Köller et al. (2006) bestätigten für das Fach Mathematik, dass sich Interessen bei Schülerinnen und Schülern im Kurs mit erhöhten, im Vergleich zu grundlegenden, Anforderungen günstiger verändern. Entsprechende Studien im Fach Englisch fehlen bislang; ebenso sind die postulierten Assoziationen nicht im Zusammenhang mit beruflichen Interessen untersucht worden.

Neben den bislang genannten Einflussgrößen spielt für die Entwicklung schulischer Leistungen auch das Selbstkonzept eine entscheidende Rolle. Das Selbstkonzept lässt sich als Vorstellung einer Person über sich selbst und die eigenen Fähigkeiten beschreiben (vgl. Shavelson, Hubner & Stanton, 1976). Ein höheres Selbstkonzept führt zu höheren Erfolgserwartungen, zu einer ausdauernderen Aufgabenbearbeitung und in

Folge dessen zu besseren Leistungen (vgl. Eccles[-Parsons], Adler & Meece, 1984; Fredricks & Eccles, 2002). Das fachspezifische Selbstkonzept wiederum ist enger auf ein bestimmtes Themengebiet bezogen. Im schulischen Kontext sind dies bspw. das Selbstkonzept in Mathematik oder Englisch. Zur Beschreibung der Prozesse, die zur Herausbildung eines fachspezifischen Selbstkonzepts führen, wird oftmals das *Internal/External Frame of Reference (I/E)* Modell nach Marsh (1986, 1990a) verwendet und dabei zwei wesentliche Vergleichsprozesse postuliert: zum einen kontrastieren Schülerinnen und Schüler ihre Leistungen in verschiedenen Fächern (internale Vergleiche), zum anderen führen sie fachinterne soziale Vergleiche durch, bei denen sie die eigenen Leistungen in einem Fach mit den fachspezifischen Leistungen der Mitschülerinnen und Mitschüler gegenüber stellen (externale Vergleiche). Zum Zusammenhang zwischen fachspezifischem Selbstkonzept und Leistungsindikatoren (Noten und Leistungstests) berichteten Möller und Köller (2004) in einem Überblicksartikel eine Median-Korrelation von $Md = .47$ im Fach Mathematik bzw. von $Md = .39$ für den sprachlichen Bereich. Schon früh tauchte die Frage nach den Richtungen der Effekte zwischen Selbstkonzept und Leistungen auf. Hierbei kontrastierten Calsyn und Kenny (1977) das *self-enhancement*- und das *skill development*-Modell. In ersterem wird das Selbstkonzept als eine den akademischen Leistungen vorausgehende Determinante betrachtet, in letzterem wird eine umgekehrte Einflussrichtung postuliert. In empirischen Studien zeigten sich meist beide Effektrichtungen (vgl. Marsh et al., 2005; Valentine, Dubois & Cooper, 2004). Marsh (1990b) hingegen belegte im Längsschnitt ausschließlich systematische Effekte vom Selbstkonzept auf nachfolgende Leistungen (Noten) unter Kontrolle der Ausgangsleistung, jedoch nicht umgekehrt.

Gleichermaßen stehen auch Selbstkonzept und Interessen in enger Verbindung zueinander (z.B. Krapp, 2001). So führt das Interesse an einem Gegenstand zu dem Wunsch, sich mit diesem auseinanderzusetzen und das eigene Wissen darüber zu aktualisieren. Mit diesem Wunsch geht oftmals einher, dass sich die Person mit dem Gegenstand identifiziert, was wiederum im Zusammenhang mit der Herausbildung des Selbstkonzepts steht (Hannover, 1998). Lent, Brown und Hackett (1994) zeigten in einer Meta-Analyse Korrelationen in Höhe von $r = .53$ bzw. $r = .52$ zwischen verschiedenen Maßen selbstbezogener Kognitionen und beruflichen Interessen. In einer neueren Studie belegten Marsh et al. (2005) reziproke Effekte zwischen Selbstkonzept und Interessen.

Im Kontext von Erwartungs-mal-Wert-Theorien wird dagegen gerichtet angenommen, dass das Selbstkonzept einen Effekt auf das Interesse ausübt, dass also ein Teil des Effekts des Selbstkonzepts auf nachfolgende Leistungen über Interessen mediiert ist. Erste empirische Hinweise zum Beziehungsgefüge in der gymnasialen Oberstufe geben Köller et al. (2006). Die Autoren zeigten für das Fach Mathematik zum einen, dass das fachspezifische Selbstkonzept einen Effekt auf das zeitlich nachgeordnete Interesse ausübte und zum anderen, dass bei simultaner Berücksichtigung sowohl das Selbstkonzept als auch das Interesse Effekte auf die Mathematikleistung unter Kontrolle der Ausgangsleistung hatten. Auch hier mangelt es an längsschnittlichen Untersuchungen, die sich auf das Fach Englisch beziehen. In der *Sozial-kognitiven Berufslaufbahntheorie* (*Social Cognitive Career Theory*, SCCT; Lent et al., 1994) wird ein den Erwartungs-mal-Wert-Theorien entsprechendes Beziehungsgefüge angenommen, jedoch vertreten hierbei berufliche Interessen das Interessenkonstrukt. Längsschnittliche Untersuchungen zum Zusammenspiel von Selbstkonzept, beruflichen Interessen und Leistungen liegen bislang nicht vor.

3. Fragestellungen

Zentrale Ziele des Unterrichts in der gymnasialen Oberstufe sind u. a. die Vermittlung einer vertieften Allgemeinbildung und einer allgemeinen Studierfähigkeit (vgl. KMK, 2006). Zur Umsetzung dieser Ziele wird den Fächern Deutsch, Mathematik und der ersten Fremdsprache eine entscheidende Rolle zugeschrieben. Entsprechend sind diese Fächer in der Vereinbarung zur Gestaltung der gymnasialen Oberstufe als Pflicht- und teilweise als Abitur-Prüfungsfächer festgelegt (vgl. ebd.). Die gymnasiale Oberstufe beginnt mit einer einjährigen Einführungsphase. Hier sollen unterschiedliche Voraussetzungen bei den Schülerinnen und Schülern ausgeglichen und ihr Wissensstand angeglichen werden, um sie möglichst optimal auf die anschließende Qualifikationsphase vorzubereiten (vgl. KMK, 2006; Niedersächsisches Kultusministerium, 1990, 2003).

Um Leistungen und Leistungsveränderungen möglichst objektiv zu erfassen, haben sich standardisierte Tests, die auf der Item Response Theory basieren, als Methode der Wahl herausgestellt. Diese gehen insbesondere bei längsschnittlichen Erhebungen mit entscheidenden ökonomischen Vorteilen einher (vgl. Hambleton & Swaminathan, 1985; Lord, 1980). Erste Ergebnisse aus Längsschnittstudien in der gymnasialen Oberstufe anhand standardisierter Tests liegen für die Fächer Mathematik und Englisch bereits vor (z. B. Köller, Baumert et al., 2004; Köller et al., 2006; Lehmann et al., 2004, 2006). Die bisherigen Studien geben jedoch wenig Aufschluss darüber, welche Leistungsentwicklungen sich in der Einführungsphase mit ihren speziellen curricularen Zielsetzungen verzeichnen lassen.

Neben einer möglichst objektiven Erfassung schulischer Leistungen und Leistungsentwicklungen stellt sich im Rahmen pädagogisch-psychologischer Forschung insbesondere die Frage nach entsprechenden Fördermöglichkeiten (vgl. Helmke & Weinert, 1997). Hierbei hat sich das Interesse als ein zentrales Konstrukt schulischen Lernens bewiesen (z. B. Krapp, 1998, 2002). Im Rahmen einer Person-Gegenstands-Konzeption (z. B. Krapp, 1992) wurde das fachspezifische Interesse intensiv erforscht und ein Zusammenhang mit schulischen Leistungen belegt (vgl. Schiefele et al., 1993). In ersten Längsschnittstudien lässt sich auch längsschnittlich ein Effekt von Interessen auf Leistungen in der gymnasialen Oberstufe nachweisen (z. B. Köller, Baumert et al., 2000; Köller et al., 2006). Berufliche Interessen sind in diesem Kontext hingegen kaum

berücksichtigt worden. Jedoch ist anzunehmen, dass insbesondere in der gymnasialen Oberstufe, also mit zunehmender zeitlicher Nähe zu Studium bzw. Beruf, berufliche Interessen eine handlungsleitende Funktion ausüben. Einen empirisch gut fundierten Hintergrund zur Erfassung beruflicher Interessen bildet die bereits beschriebene Theorie der Berufswahl von Holland (1997). Ackerman und Heggestad (1997) belegten in einer Metaanalyse Zusammenhänge zwischen Hollands Interessenorientierungen und Leistungen im mathematischen bzw. sprachlichen Bereich. In einer längsschnittlichen Studie waren bislang keine Effekte zwischen Interessenorientierungen und Testleistungen nachweisbar (Tracey et al., 2005). Im deutschsprachigen Raum sind entsprechende Studien vergleichsweise rar (aber: Bergmann & Eder, 1992; Eder & Reiter, 2002), eine längsschnittliche Prüfung der Zusammenhänge fand nicht statt.

Berufliche Interessen werden meist in Strukturmodellen beschrieben (vgl. z.B. Savickas & Spokane, 1999). Neben der Ausprägung einzelner Interessenorientierungen stehen deshalb die Assoziation zwischen den Orientierungen, also die individuellen Interessenkonfigurationen, im Vordergrund. Interessenkonfigurationen oder -profile werden meist typologisch in Form von Skalenwerten (vgl. Holland, 1997) dargestellt. Alternativ lassen sich Profile unter Berücksichtigung der dem RIASEC-Modell zugrunde liegenden dimensional Struktur beruflicher Interessen und in Anlehnung an ein Structural Summary nach Gurtman und Balakrishnan (1998) umfassend und sparsam modellieren. In einer entsprechenden Querschnittstudie belegte Nagy (2006) mit bisherigen Befunden übereinstimmende Zusammenhänge zwischen beruflichen Interessenorientierungen und Testleistungen in Mathematik und Englisch. Die Validität solch einer sparsamen Darstellung im Vergleich zu Skalenwerten wurde bislang nicht geprüft.

Insofern Interessenprofile eine valide Reduktion von Skalenwerten darstellen, ist es aufgrund ihrer sparsamen Modellierung möglich, sie in bestehende bereichsspezifische Theorien zu Determinanten von Leistungen zu integrieren und entsprechende Effekte auf Leistungen sowie Assoziationen zwischen den Einflussgrößen zu untersuchen. Im Kontext der gymnasialen Oberstufe lässt sich hierbei insbesondere das Kursniveau heranziehen. Schülerinnen und Schüler können im Rahmen der Vorgaben der KMK (2006) zwischen Kursen mit grundlegenden oder erhöhten Anforderungen wählen. Empirisch zeigten sich zunehmende Leistungsdifferenzierungen in Abhängigkeit vom Kursniveau im Verlauf der gymnasialen Oberstufe (vgl. Köller et al.,

2006) sowie deutliche Leistungsunterschiede am Ende der Gymnasialschulzeit (z.B. Baumert & Watermann, 2000; Köller & Trautwein, 2004; Watermann et al., 2004). Gleichermaßen ist davon auszugehen, dass die vertiefte Auseinandersetzung mit einem Fach förderlich für das Interesse an diesem Gegenstand ist (z.B. Köller et al., 2006; Krapp, 1992). Darüber hinaus stehen Interesse und Leistungen mit dem Selbstkonzept in Verbindung: In Erwartungs-mal-Wert-Modellen (z.B. Eccles[-Parsons] et al., 1984) wird angenommen, dass das Selbstkonzept einen Effekt sowohl auf Leistungen als auch auf Interessen ausübt. Somit mediiert Interessen (partiell) den Effekt des Selbstkonzepts. Diese Annahmen sind auch Bestandteil der Sozial-kognitiven Berufslaufbahtheorie (Lent et al., 1994), bei der explizit berufliche Interessen in das Beziehungsgefüge eingebunden werden. Erste Längsschnittstudien in der gymnasialen Oberstufe bestätigen diese Assoziationen am Beispiel des Fachs Mathematik (z.B. Köller et al., 2006). Entsprechende Befunde liegen weder für Englisch noch für berufliche Interessen vor.

Aufgrund der angeführten theoretischen Annahmen und bisheriger empirischer Befunde sollen in dieser Arbeit die folgenden drei übergeordneten Fragestellungen abgeleitet und expliziert werden:

1. Wie verlaufen die Leistungsentwicklungen in Mathematik und Englisch in der Einführungsphase der gymnasialen Oberstufe?
2. Sind berufliche Interessenprofile eine valide Reduktion von Skalenwerten im Rahmen der Theorie der Berufswahl nach Holland?
3. Welche Bedeutung haben berufliche Interessenprofile für Leistungsentwicklungen in Mathematik und Englisch in der Einführungsphase der gymnasialen Oberstufe? Und welche Zusammenhänge bestehen zwischen beruflichen Interessen, Selbstkonzept und Kursniveau in Hinblick auf Leistungsentwicklungen in den beiden Fächern?

Diesen Fragestellungen wird in den drei folgenden empirischen Teilen der Arbeit (Kapitel 4) differenzierter nachgegangen. Zentrale Ergebnisse werden im anschließenden Kapitel zusammengefasst und aufeinander Bezug nehmend diskutiert.

4. Leistungsentwicklungen und berufliche Interessen in der gymnasialen Oberstufe

4.1 Leistungsentwicklungen in Mathematik und Englisch im ersten Jahr an einem Fachgymnasium⁵

Zusammenfassung und Abstract

Aus Längsschnittstudien liegen bereits erste Ergebnisse zu Leistungsentwicklungen in Mathematik und Englisch in der gymnasialen Oberstufe vor. Hieraus zeigten sich für mathematische Grundbildung inkonsistente Befunde, während in Englisch überwiegend positive Leistungsentwicklungen resultierten. Des Weiteren lassen sich bessere Leistungen von Schülerinnen und Schülern in Kursen mit erhöhten Anforderungen im Vergleich zu denjenigen in Kursen mit grundlegenden Anforderungen am Ende der gymnasialen Oberstufe nachweisen. Ursächlich hierfür werden unterschiedliche Eingangsvoraussetzungen und weitere Leistungsdifferenzierungen im Verlauf des Kursystems angeführt. An diesen empirischen Befunden anknüpfend, wurden in der vorliegenden explorativen Studie erstmals Leistungsentwicklungen von Schülerinnen und Schülern innerhalb der 11. Jahrgangsstufe bzw. der Einführungsphase an einem Fachgymnasium untersucht. Zwei aufeinander folgende Jahrgänge ($N = 135$) nahmen an zwei Erhebungen mit standardisierten Leistungstests in Mathematik und Englisch teil. Es zeigte sich, dass mathematische Grundbildung im Mittel sank, die durchschnittliche Englischleistung hingegen anstieg. Bei einer differenzierten Betrachtung in Abhängigkeit vom angestrebten Kursniveau der Schülerinnen und Schüler ab der 12. Jahrgangsstufe resultierten sowohl fach- als auch messzeitpunktübergreifend erwartbare Unterschiede in den mittleren Leistungen beider Gruppen; die Leistungsschere ging in keinem der beiden Fächern weiter auf.

⁵ Dieser Abschnitt basiert auf einem zur Veröffentlichung angenommenen Manuskript in der Zeitschrift *Unterrichtswissenschaft* (Warwas, Watermann & Hasselhorn, im Druck).

Abstract. Several empirical studies have investigated the development of students' achievement in upper secondary level in Germany. Findings for mathematical literacy were heterogeneous, whereas achievement in English increased substantially. Furthermore, studies revealed higher achievement for students in advanced courses compared to those in basic courses. This can be traced back to different achievement levels at the beginning of upper secondary level as well as to an increasing differentiation during upper secondary level. We conducted an exploratory study to investigate the development of students' achievement within the first year after transition in a German academic-track school. Students of two consecutive cohorts ($N = 135$) participated twice in standardized achievement tests on mathematics and English. Results showed mathematical literacy decreasing, whereas English achievement significantly increased. We further examined students' achievement depending on which course level (basic or advanced) they choose in the 12th grade. Students in prospective advanced courses outperformed other students in both subjects each time. Yet, there was no interaction effect between course levels and achievement during 11th grade.

4.1.1 Theoretischer Hintergrund

Im deutschen Sekundarschulwesen hat in den letzten Jahrzehnten eine bemerkenswerte Öffnung von Wegen zur Allgemeinen Hochschulreife stattgefunden. Neben einer teilweisen Entkopplung von Schulabschluss und Schulform ist es zu institutionellen Erweiterungen gekommen, die Bildungswege offen halten und Schülerinnen und Schülern alternative Wege zum Abitur ermöglichen sollen (Baumert, Cortina & Leschinsky, 2003). Fachgymnasien stellen eine Form dieser Öffnung dar. Von allen Schülerinnen und Schülern, die die allgemeine Hochschulreife im Jahr 2005 erworben haben, erreichten diese im Bundesdurchschnitt 15% an Fachgymnasien. Dabei verzeichnete Baden-Württemberg mit 32% den größten Anteil und im Saarland gab es besonders niedrige Absolventenquoten (4%). Niedersachsen liegt mit 17% etwa im Bundesdurchschnitt (Niedersächsisches Kultusministerium, 2006, eigene Berechnungen). Neben den Aufgabefeldern des allgemein bildenden Gymnasiums werden in dieser Schulform berufsfeldspezifische Fächer, wie Wirtschaft oder Technik, unterrichtet. Der dreijährige Bildungsgang endet mit der allgemeinen Hochschulreife (vgl. KMK, 2006).

Als gymnasiale Oberstufe in Aufbauform richten sich Fachgymnasien vor allem an Schülerinnen und Schüler, die ihren Sekundarschulabschluss an Haupt- oder Realschulen erworben haben. Um ein gewisses Fähigkeitsniveau als Voraussetzung für den erfolgreichen Besuch der gymnasialen Oberstufe zu gewährleisten, wurden leistungsbezogene Übergangsvorschriften formuliert: Während für Absolventinnen und Absolventen allgemein bildender Gymnasien das Versetzungszeugnis in die 11. Jahrgangsstufe (Einführungsphase) als Nachweis hinreichender Voraussetzungen für den Eintritt in ein Fachgymnasium gilt, muss für einen direkten Übergang aus Haupt- oder Realschulen der erweiterte Sekundarabschluss I erworben worden sein. Im Abschlusszeugnis müssen dazu bestimmte Notenvoraussetzungen erfüllt sein (vgl. ebd.). Entsprechende Regelungen unterscheiden sich zwischen den Bundesländern, in Niedersachsen sind es u. a. im Durchschnitt befriedigende Leistungen in den Fächern Mathematik, erste Fremdsprache und Deutsch (Niedersächsisches Kultusministerium, 2000). Diese Fächer sind bundesweit in der Vereinbarung zur Gestaltung der gymnasialen Oberstufe (KMK, 2006) als Kernfächer festgelegt und werden von den Schülerinnen und Schülern durchgängig bis zum Abitur belegt. Somit sind Leistungen in den Kernfächern sowohl für den Übergang in ein Fachgymnasium als auch für einen erfolgreichen Verlauf an selbigem von hoher Relevanz. Trotz dieser positiven Selektion durch leistungsbezogene Übergangsvorschriften belegten Lehmann et al. (2004) anhand von Daten der LAU 11-Studie, dass die mittleren Leistungen der Schülerinnen und Schüler am Fachgymnasium direkt nach dem Übergang in die Sekundarstufe II niedriger sind als derjenigen an grundständigen bzw. allgemein bildenden Gymnasien. Die Differenz beträgt in Mathematik etwa eine halbe Standardabweichung, in Englisch – abhängig vom verwendeten Testbereich – eine halbe bis ganze Standardabweichung.

Schulische Leistungen werden in der Regel durch Noten beurteilt und an die Schülerinnen und Schüler zurückgemeldet. Takei, Johnson und Clark (1998) beschreiben verschiedene Prozesse zur Erklärung der begrenzten Validität von Noten als Leistungsindikatoren. So fließen zahlreiche Merkmale der Schülerinnen und Schüler in die Wahrnehmung und Benotung der Lehrkraft ein (Archer & McCarty, 1998). Zudem zeigte Ingenkamp (1995), dass Noten eher sozialnorm- als sachnormorientiert vergeben werden, und dass sie in hohem Maß vom Leistungsspektrum einer Bezugsgruppe (z. B. Schulklasse oder Schule) abhängig sind (vgl. auch Baumert, Trautwein et al., 2003).

Somit sind Objektivität und Vergleichbarkeit von Leistungen über Schulklassen hinweg deutlich eingeschränkt. Auch in Hinblick auf die objektive Erfassung von Leistungsveränderungen als Indikatoren schulischen Lernens sind Noten wenig geeignet. Die Benotung erfolgt meist in Hinblick auf die Klassenleistung und Veränderungen in den Noten spiegeln in hohem Maße eine veränderte Rangposition innerhalb der Klasse wider. Um der Objektivität bei der Erfassung von Leistungen und Leistungsveränderungen in größerem Maß gerecht zu werden und um Urteilsfehler zu reduzieren bzw. zu kontrollieren, wurden im Kontext (inter-)nationaler Vergleichsstudien standardisierte Leistungstests entwickelt und eingesetzt (z. B. PISA, OECD, 2001). Für Mathematik und Englisch gibt es entsprechende Tests, für die im folgenden Abschnitt Befunde berichtet werden. Im Fach Deutsch liegen noch keine erprobten Verfahren vor.

Zur Erfassung von Mathematikleistungen wurde im Rahmen von TIMSS (Mullis et al., 1998; Baumert et al., 1997; Baumert et al., 1999; Baumert et al., 2000) der Test zur mathematischen Grundbildung entwickelt. Für die Sekundarstufe II zeigten Ramseier et al. (1999) in einer Querschnittstudie in der Schweiz, dass sich die durchschnittlichen Leistungen in mathematischer Grundbildung bei berufsbildenden Schülerinnen und Schülern im letzten Jahr der Sekundarstufe II nicht bedeutsam von denjenigen im ersten Jahr unterscheiden. Die Interpretierbarkeit ist aber aufgrund des querschnittlichen Charakters der Studie eingeschränkt und der Abstand mit zwei Jahren relativ groß. Ebenfalls für einen zweijährigen Abstand berichteten Lehmann und Kollegen zu Längsschnittdaten aus den Hamburger LAU 11- und LAU 13-Studien (Lehmann et al., 2004, 2006) mittlere Leistungsanstiege in mathematischer Grundbildung – konstant über alle gymnasialen Schulformen (inkl. Fachgymnasien) hinweg – in Höhe einer Drittel bis einer halben Standardabweichung. Im Rahmen der TOSCA-Studie (Köller, Watermann et al., 2004) wurde mathematische Grundbildung am Ende der gymnasialen Oberstufe erfasst und dabei nicht als expliziter Inhalt der Sekundarstufe II, sondern als Eingangsvoraussetzung und somit Basisqualifikation für die Oberstufenmathematik im Allgemeinen, betrachtet (vgl. Watermann et al., 2004). Auch lt. Rahmenrichtlinien im Fach Mathematik (Niedersächsisches Kultusministerium, 1990) stehen in der Einführungsphase mit dem Tangentenproblem und Ableitungen zwar überwiegend für die Sekundarstufe II spezifische neue Themen an, mit denen die Schülerinnen und Schüler gezielt auf die Analysis zu Beginn der Kursstufe vorbereitet werden sollen, gleichzeitig

werden jedoch der Rückgriff auf aus der Sekundarstufe I bekannte Sachverhalte bzw. „auf Basis des Grundwissens“ (S. 6) sowie der Anwendungsbezug der Inhalte auf die Praxis gefordert. Befunde zu Leistungsentwicklungen in Mathematik innerhalb der 11. Jahrgangsstufe liegen bislang nicht vor.

Im Fach Englisch liegen sich im Bereich der Sekundarstufe II überwiegend Lernzuwächse verzeichnen: In LAU wurde für alle Gymnasialschulformen eine Verbesserung der Englischleistung um eine halbe bzw. zwei Drittel Standardabweichung(en) innerhalb von zwei Jahren (11. bis 13. Jahrgangsstufe) nachgewiesen (Lehmann et al., 2004, 2006). Im Rahmen der BIJU-Studie (Baumert et al., 1996) zeigten Köller, Baumert et al. (2004) mit Hilfe einer TOEFL-Kurzversion (*Test of English as a Foreign Language*; entwickelt vom *Educational Testing Service*, Version vom *Institutional Testing Program*), dass sich die Leistungen der Schülerinnen und Schüler an allgemein bildenden Gymnasien von der 12. zur 13. Jahrgangsstufe im Mittel verbesserten. Nur an Integrierten Gesamtschulen konnten keine mittleren Leistungssteigerungen nachgewiesen werden. In der Einführungsphase sind in den Rahmenrichtlinien für das Fach Englisch (Niedersächsisches Kultusministerium, 2003) Wiederholung und Vertiefung aus der Sekundarstufe I erlernter Inhalte durch „Grammatikarbeit“ und „Systematisierung und vertiefenden Betrachtung von Strukturen“ (S. 16) verzeichnet. Jedoch liegen für diese Jahrgangsstufe auch hier bislang keine empirischen Arbeiten über Leistungsentwicklungen in Englisch vor.

Eine differenzierte Betrachtung in Hinblick auf das jeweils angestrebte Kursniveau ermöglicht zusätzliche Informationen über die Leistungsentwicklungen der Schülerinnen und Schüler in den jeweiligen Fächern. In der gymnasialen Oberstufe belegen Schülerinnen und Schüler in den Kernfächern jeweils einen Kurs mit grundlegenden oder erhöhten Anforderungen. Mit dieser Struktur soll den Schülerinnen und Schülern eine individuelle Profilbildung ermöglicht und zugleich eine Sicherung von Standards einer vertieften Allgemeinbildung erzielt werden. Ein übergeordnetes Ziel der Unterrichtsorganisation in der Einführungsphase ist es lt. Rahmenrichtlinien für die gymnasiale Oberstufe im Fach Mathematik (Niedersächsisches Kultusministerium, 1990) „für alle Schülerinnen und Schüler dieser Jahrgangsstufe gleiche Lernvoraussetzungen für die Mitarbeit in der Kursstufe zu schaffen“ (S. 17). Gleichmaßen steht beim Unterricht in der Einführungsphase im Fach Englisch (Niedersächsisches Kultusministerium,

2003) der Ausgleich von Defiziten bei unterschiedlichen Eingangsvoraussetzungen im Vordergrund (vgl. ebd., S. 16). Hingegen sind deutliche Leistungsunterschiede am Ende der Sekundarstufe II zugunsten von Schülerinnen und Schülern im Kurs mit erhöhten Anforderungen im Vergleich zu denjenigen im Kurs mit grundlegenden Anforderungen empirisch gut belegt; die Leistungsunterschiede am Ende der Sekundarstufe II liegen meist zwischen zwei Drittel bis zu einer ganzen Standardabweichung(en) (vgl. Lehmann et al., 2006; Köller, Watermann et al., 2004). Diese Differenzen lassen sich zum einen auf unterschiedliche Eingangsvoraussetzungen zu Beginn der Kursstufe und zum anderen auf weitere Leistungsdifferenzierungen während der Kursstufe zurückführen (vgl. Köller et al., 2006).

4.1.2 Fragestellungen

Fachgymnasien (als gymnasiale Oberstufe in Aufbauform) richten sich vor allem an Schülerinnen und Schüler, die ihren Sekundarschulabschluss an Haupt- oder Realschulen erworben haben und über diesen alternativen Bildungsweg das Ziel verfolgen, einen Hochschulabschluss zu erwerben. Wesentliche Ziele des Unterrichts in der 11. Jahrgangsstufe, der Einführungsphase, sind lt. den curricularen Rahmenrichtlinien (Niedersächsisches Kultusministerium, 1990, 2003) die Vorbereitung aller Schülerinnen und Schüler auf die Kursstufe ab der 12. Jahrgangsstufe, der Ausgleich von Defiziten bei unterschiedlichen Eingangsvoraussetzungen sowie die Vorbereitung auf die Kursstufe unter Rückgriff auf Inhalte der Sekundarstufe I bzw. einer Wiederholung des Unterrichtsstoffes.

Erste Schulleistungsuntersuchungen in der gymnasialen Oberstufe zu Leistungsentwicklungen in Mathematik und Englisch kommen zu folgenden Ergebnissen: Die auf die mathematische Grundbildung bezogene Leistung blieb bei berufsbildenden Schülerinnen und Schülern innerhalb der Sekundarstufe II konstant (vgl. Ramseier et al., 1999) bzw. verbesserte sich (vgl. Lehmann et al., 2004, 2006). In TOSCA wird mathematische Grundbildung am Ende der gymnasialen Oberstufe als Maß für die Eingangsvoraussetzung betrachtet (vgl. Watermann et al., 2004). In Englisch lassen sich überwiegend positive Leistungsentwicklungen im Verlauf der gymnasialen Oberstufe nachweisen (vgl. Baumert et al., 1996; Lehmann et al., 2004, 2006). Die vorliegenden Befunde er-

lauben zurzeit keine Einschätzung von Leistungsveränderungen in den beiden Fächern in der Einführungsphase bzw. der 11. Jahrgangsstufe. Hauptziel der vorliegenden Studie war es daher, zu explorieren wie sich die Leistungen an einem Fachgymnasium innerhalb des ersten Jahres in der gymnasialen Oberstufe entwickeln. Aufgrund bisheriger Befunde ließen wir uns von folgenden Fragen leiten:

1. Verbessern sich die mittleren Leistungen in mathematischer Grundbildung bei Schülerinnen und Schülern im Fachgymnasium in der 11. Jahrgangsstufe oder bleiben sie konstant?
2. Verbessern sich die mittleren Leistungen in Englisch von Schülerinnen und Schülern im Fachgymnasium in der 11. Jahrgangsstufe?

Darüber hinaus war eine differenziertere Betrachtung der Leistungsentwicklungen in Abhängigkeit vom angestrebten Kursniveau (grundlegende oder erhöhte Anforderungen) von Interesse. Empirisch wurden am Ende der gymnasialen Oberstufe bislang deutliche kursniveauspezifische Leistungsunterschiede nachgewiesen (z.B. Köller, Watermann et al., 2004; Lehmann, 2006). Ursächlich hierfür werden sowohl unterschiedliche Eingangsvoraussetzungen zu Beginn als auch weitere Leistungsdivergenzen im Verlauf der gymnasialen Oberstufe angesehen (vgl. Köller et al., 2006). Es lässt sich erwarten, dass bereits während der Einführungsphase unterschiedliche Leistungen in Abhängigkeit vom angestrebten Kursniveau bestehen und sich die Leistungen zwischen den Gruppen weiter ausdifferenzieren. Gleichzeitig soll der Unterricht der 11. Jahrgangsstufe alle Schülerinnen und Schüler möglichst gut auf die anstehende Kursstufe vorbereiten und Defizite bei unterschiedlichen Eingangsvoraussetzungen ausgleichen (vgl. Niedersächsisches Kultusministerium, 1990, 2003). In welchem Ausmaß diese Unterrichtsziele einem gegenläufigen Trend einer Leistungsdivergenz entgegenzuwirken vermögen, soll in dieser explorativen Studie erstmals für die Einführungsphase untersucht werden:

3. Öffnet sich die Leistungsschere in den Fächern Mathematik und Englisch bereits in der Einführungsphase oder bleiben Leistungsunterschiede zwischen Schülerinnen und Schülern in Abhängigkeit vom angestrebten Kursniveau konstant?

4.1.3 Methode

Durchführung. Die Erhebungen fanden Mitte (Februar 2006, T1) und Ende (Juli 2006, T2) der 11. Jahrgangsstufe an einem Fachgymnasium Wirtschaft in Niedersachsen statt. Sie wurden im Gruppensetting und zeitgleich in den vier Schulklassen durchgeführt. Zu T1 dauerten die Tests in Mathematik und Englisch je ca. 60 Minuten mit einer halbstündigen Pause dazwischen. In jeder Klasse war eine Testleitung anwesend, die mit Hilfe eines ausführlichen Instruktionsmanuals dafür sorgte, dass die Testbedingungen über die Klassen hinweg vergleichbar waren. Zu T2 betrug die Testdauer in Mathematik 25 Minuten und in Englisch 20 Minuten. Es gab keine Pause zwischen den Tests. Diese Erhebungen waren Bestandteil einer weitere Schulfächer umfassenden Leistungsstanddiagnostik, die von geschulten Lehrkräften durchgeführt wurde. Zu beiden Messzeitpunkten waren die für die Testbearbeitung relevanten Instruktionen zusätzlich in den Testheften abgedruckt um Verständnisschwierigkeiten auszuschließen. Ein Jahr später (2007) wurde der anschließende Schuljahrgang an der gleichen Schule mit in die Erhebungen aufgenommen. Die Erhebungen fanden auch hier im Februar und Juli statt; Testmaterial und Vorgehen waren mit dem des vorherigen Jahrgangs vergleichbar.

Stichprobe. Die untersuchten Personen sind Schülerinnen und Schüler an einem niedersächsischen Fachgymnasium Wirtschaft. In die Auswertung wurden nur diejenigen einbezogen, für die Daten zu beiden Messzeitpunkten vorlagen ($N = 135$). Das mittlere Alter betrug zum ersten Messzeitpunkt 18.1 Jahre ($SD = 1.5$), das Geschlechterverhältnis war ausgewogen ($n = 68$ Schülerinnen). Ausgangsstichprobe ($N = 168$) und in dieser Studie einbezogene Stichprobe sind in ihren Leistungen zu T1 vergleichbar; die Mittelwertsdifferenzen und Effektgrößen (Cohen, 1988) zur Selektivität zu T1 sind für Mathematik ($M = 2.5$, $d = 0.04$) und Englisch ($M = 7.2$, $d = 0.07$) vernachlässigbar gering. In der Vereinbarung zur Gestaltung der gymnasialen Oberstufe (KMK, 2006) ist festgelegt, dass die Schülerinnen und Schüler für zwei der drei Kernfächer Mathematik, erste Fremdsprache und Deutsch einen Kurs mit erhöhten Anforderungen belegen. Eine Besonderheit der untersuchten Stichprobe ist, dass bereits zum Einstieg in das Fachgymnasium eine Gruppierung hinsichtlich der angestrebten Kursniveaus stattfand. Basis hierfür war eine ausführliche Eingangsdiagnostik zur Aufnahme am Fachgymnasium, in der u. a. fachspezifische Leistungen und Interessen an den drei Kernfächern erfasst wur-

den. Jede Klasse setzte sich somit aus Schülerinnen und Schülern zusammen, die die gleichen beiden Fächer auf erhöhtem Anforderungsniveau ab der 12. Jahrgangsstufe belegen wollten.

Abhängige Variablen. Leistungen in Mathematik wurden mit dem TIMSS-Test zur mathematischen Grundbildung erhoben. Die Testaufgaben beziehen sich auf zentrale Inhalte des mathematischen Unterrichts der Mittelstufe (Zahlenverständnis, Proportionalität, Algebra sowie Messen und Schätzen) und sind entsprechend der angelsächsischen Literacy-Debatte (*National Council of Teachers of Mathematics*, NCTM, 1989) in Alltagskontexte eingebunden. Der verwendete Mathematiktest ist dabei ein „Kompromiss zwischen Anwendungsorientierung und curricularer Anbindung“ (vgl. Baumert et al., 2001, S. 19). Die im Rahmen von TIMSS durchgeführten Validitätsstudien belegen deren Validität im Hinblick auf den Lehrplan wie auch bezogen auf die Lerngelegenheiten im Unterricht (Klieme, Baumert, Köller & Bos, 2000). Zugleich zeigten Re-Analysen von Daten aus TIMSS, dass die Leistungen weitgehend unabhängig von diesem konzeptionellen Aspekt sind (vgl. Mullis et al., 1998). Die Skalenreliabilität, berechnet auf Basis der TIMSS-III-Population, beträgt für die Items, die in diese Studie zu T1 eingesetzt wurden, $r = .83$ und für die Items zu T2 $r = .68$. Mit der Software *Conquest* (Wu, Adams & Wilson, 1998) wurden auf Basis der in TIMSS ermittelten Itemparameter Weighted-Likelihood-Estimates (Warm, 1989) als Personenparameter geschätzt. Diese wurden auf die in TIMSS verwendete Metrik ($M = 500$, $SD = 100$) transformiert. Zu T1 wurden 21 Aufgaben ausgewählt und zu T2 8, davon 4 wiederholt (Ankeritems) (s. Anhang A). Mit diesen Ankeritems ist es im Rahmen sogenannter Test Equating-Prozeduren (vgl. Hambleton & Swaminathan, 1985) möglich, zu verschiedenen Zeitpunkten erhobene Leistungen auf einer gemeinsamen Metrik abzutragen. Lord (1980) nimmt eine Zahl ab vier bis fünf Ankeritems als ausreichend an.

In Englisch wurde eine Kurzform des TOEFL (Educational Testing Service) eingesetzt. Dieser ITP-TOEFL (Institutional Testing Program) wurde von Köller und Trautwein (2004) im Rahmen der BIJU-Studie validiert. Mit dem Test sollen Leistungen erfasst werden, die für ein Studium an einer englischsprachigen Universität nötig sind. Curriculare Validität wird nicht beansprucht, es zeigen sich aber deutliche Korrelationen zwischen TOEFL-Wert und Englischnote ($r = .64$) (vgl. ebd.). Der Test setzt sich aus drei Skalen zusammen: Hörverständnis (*Listening Comprehension*, LC), Gram-

matik und Orthografie (*Structure and Written Expression*, SWE) sowie Vokabelwissen und Leseverständnis (*Vocabulary and Reading Comprehension*, VRC) (s. Anhang B). Die Kurzform enthält 80 Items und die im Rahmen der BIJU-Studie berechneten internen Konsistenzen (Cronbachs alpha) für die drei Skalen betragen $\alpha_{LC} = .86$ (29 Aufgaben), $\alpha_{SWE} = .81$ (23 Aufgaben) und $\alpha_{VRC} = .82$ (28 Aufgaben) sowie für den Gesamttstwert $\alpha_{TOEFL} = .92$. Die Skalen korrelierten untereinander jeweils zu über $r > .80$, sodass die Zusammenfassung zu einem Gesamtwert gerechtfertigt schien. Zu T1 wurde die Kurzversion vollständig eingesetzt. Zu T2 wurde aus organisatorischen Gründen die LC-Subskala nicht erneut mit berücksichtigt. Es wurden 21 Aufgaben aus den anderen beiden Skalen verwendet, davon 13 Ankeritems aus der Kurzform und 8 Aufgaben aus einem ITP-TOEFL aus dem Jahr 1995. Mit Conquest (Wu et al., 1998) wurden die Aufgaben zu T1 raschskaliert; die WLE-Personenparameter wurden z-standardisiert und auf eine zu Mathematik vergleichbare Metrik gebracht. Die WLEs zu T2 wurden am Mittelwert und der Standardabweichung zu T1 standardisiert, sodass die Werte zu beiden Messzeitpunkten auf einer gemeinsamen Metrik lagen.

4.1.4 Ergebnisse

Leistungsentwicklung in der Einführungsphase. Datengrundlage bildeten zwei aufeinander folgende Kohorten von Schülerinnen und Schülern an einem Fachgymnasium. Eine jahrgangsspezifische Betrachtung (s. Tab. 2) zeigte, dass die mittleren Leistungen in mathematischer Grundbildung zu T1 zwischen erstem und zweiten Jahrgang vergleichbar waren, während der Mittelwert in Englisch der ersten Kohorte bedeutsam über demjenigen der zweiten Kohorte lag.

Um zu prüfen, ob zwischen beiden Kohorten bedeutsame Entwicklungsunterschiede bestanden, wurden vorab die mittleren Leistungsentwicklungen innerhalb der Einführungsphase pro Jahrgang untersucht. Der Interaktionsterm aus Messzeitpunkt und Kohorte im Rahmen einer univariaten Varianzanalyse mit Messwiederholung war weder für Mathematik, $F(1, 133) = 2.40, p > .05, \eta^2 = .018$, noch für Englisch, $F(1, 133) = 3.39, p > .05, \eta^2 = .025$, signifikant. Da sich keine bedeutsamen Entwicklungsunterschiede zwischen den Kohorten zeigten, wurden im Folgenden die Leistungsentwicklungen für die Gesamtstichprobe betrachtet.

Tabelle 2

Ausgangsleistungen pro Kohorte (Kh.) – Standardabweichung in Klammern – in Mathematik und Englisch

	Kh. 1 <i>n</i> = 78	Kh. 2 <i>n</i> = 57	<i>t</i>	<i>p</i>	<i>d</i>
Mathematik	510.08 (67.10)	501.25 (68.70)	0.75	.46	–0.13
Englisch	523.98 (107.10)	484.26 (96.36)	2.22	.03	–0.39

Anmerkungen. *d* = Effektgröße nach Cohen (1988).

Ein Fokus dieser explorativen Studie lag auf der Beschreibung der Leistungsentwicklungen von Schülerinnen und Schülern in zwei Kernfächern während der Einführungsphase der gymnasialen Oberstufe. Tabelle 3 enthält die Ergebnisse entsprechender statistischer Analysen.

Tabelle 3

Mittlere Leistungsentwicklungen in mathematischer Grundbildung und Englisch – Standardabweichung in Klammern – nach dem Übergang (T1) und am Ende (T2) der Einführungsphase

	<i>M</i> _{T1}	<i>M</i> _{T2}	<i>r</i> _{TT}	<i>t</i>	<i>p</i>	<i>d</i>
Mathematik	506.35 (67.79)	491.71 (94.28)	.57	2.15	.03	–0.18
Englisch	507.21 (104.21)	556.67 (115.76)	.64	–7.21	<.01	0.53

Anmerkungen. *r*_{TT} = differentielle Stabilität, *d* = Effektgröße nach Cohen (1988).

Bei der fachspezifischen Betrachtung der Leistungsentwicklungen zeigten sich deutliche Unterschiede: Die mittleren Leistungen der Schülerinnen und Schüler in mathematischer Grundbildung sanken, während sie in Englisch anstiegen. Die differentielle Stabilität mit Abstand von einem halben Schuljahr lag sowohl für Mathematik mit *r* = .57 als auch für Englisch mit *r* = .64 im mittelhohen Bereich und somit in erwartbarer Höhe (vgl. Becker, Lüdtke, Trautwein & Baumert, 2006).

Kursniveauspezifische Leistungsentwicklung. Zusätzlich zu den mittleren Entwicklungsverläufen fand eine kursniveauspezifische Betrachtung statt (s. Tab. 4).

Tabelle 4

Kursniveauspezifische mittlere Leistungsentwicklungen in mathematischer Grundbildung und Englisch – Standardabweichung in Klammern – nach dem Übergang (T1) und am Ende (T2) der Einführungsphase

Kursniveau	Mathematik			Englisch		
	<i>n</i>	M_{T1}	M_{T2}	<i>n</i>	M_{T1}	M_{T2}
Grundlegende Anforderungen	66	487.88 (72.23)	462.30 (85.84)	36	450.65 (59.24)	505.21 (119.05)
Erhöhte Anforderungen	69	524.02 (58.49)	519.86 (93.97)	99	527.28 (109.51)	587.66 (106.92)

Sowohl in mathematischer Grundbildung als auch in Englisch zeichneten sich deskriptiv kursniveaubezogene Unterschiede in den mittleren Leistungen ab. Schülerinnen und Schüler, die ab der 12. Jahrgangsstufe einen Kurs mit erhöhten Anforderungen in einem Fach belegen wollten, waren zu beiden Messzeitpunkten und in beiden Fächern besser als diejenigen, die einen Kurs mit grundlegenden Anforderungen anstrebten. Ebenso resultierten im Fach Mathematik tendenziell gegenläufige Leistungsentwicklungen der beiden Gruppen: Die Leistungen der Schülerinnen und Schüler mit angestrebtem höheren Kursniveau blieben im Mittel unverändert, während die der anderen, auf deskriptiver Ebene, abnahmen.

Um zu überprüfen, ob die Leistungsschere bereits während der Einführungsphase aufging, wurde eine einfaktorielle Varianzanalyse mit Messwiederholung gerechnet. In Mathematik zeigte sich ein signifikanter Haupteffekt des angestrebten Kursniveaus, $F(1, 133) = 15.88, p < .01, \eta^2 = .107$. Wie erwartet unterschieden sich die Gruppen in ihren Leistungen zu beiden Messzeitpunkten bereits in der Einführungsphase. Der Interaktionseffekt zwischen angestrebtem Kursniveau und Messzeitpunkt war nicht signifikant, $F(1, 133) = 2.50, p > .05, \eta^2 = .018$, die Leistungsschere öffnete sich im Bereich mathematischer Grundbildung nicht weiter.

In Englisch waren die kursniveauspezifischen Gruppen unterschiedlich stark besetzt, was Konsequenzen für die Signifikanztestung haben kann, sofern die Varianzen unterschiedlich sind. Für den hier vorliegenden Fall, dass die Varianzen der zahlenmäßig größer besetzten Gruppe zum ersten Messzeitpunkt höher ausfallen als die Varianzen der kleineren Gruppe, ist die resultierende F -Statistik eher konservativ und Signifikanztestungen somit eher streng (vgl. Glass, Peckham & Saunders, 1972). Der Levene-Test bestätigte, dass die Fehlervarianzen zwischen den beiden Gruppen zum ersten Messzeitpunkt nicht homogen waren, $F(1, 133) = 8.40, p < .01$, und somit eine wesentliche Voraussetzung der Varianzanalyse nicht erfüllt war. Zur Prüfung des Haupteffekts des angestrebten Kursniveaus wurde deshalb eine Varianzanalyse mit dem Mittelwert der beiden Messzeitpunkte als abhängige Variable verwendet. Dieses Vorgehen ist äquivalent zur Testung des Haupteffekts im Rahmen einer Varianzanalyse mit Messwiederholung. Entsprechend dem Levene-Test waren die Varianzen der aus beiden Messzeitpunkten gemittelten Leistungsniveaus homogen, $F(1, 133) = 0.59, p > .05$. Es resultierte ein signifikanter Haupteffekt des angestrebten Kursniveaus, $F(1, 133) = 19.27, p < .01, \eta^2 < .127$. Wie erwartet unterschieden sich auch im Fach Englisch beide Schülergruppen in ihren Leistungen bereits in der Einführungsphase voneinander. Der Interaktionseffekt wurde mit einer Differenz der Leistungen beider Messzeitpunkte überprüft, mathematisch äquivalent zu einer Varianzanalyse mit Messwiederholung. Die Varianzen (Levene-Test) der Differenzen waren homogen, $F(1, 133) = 0.77, p > .05$, und der Interaktionseffekt nicht signifikant, $F(1, 133) = 0.08, p > .05, \eta^2 < .001$. In Englisch gab es während der Einführungsphase keine weitere Leistungsdifferenzierung zwischen den beiden kursniveauspezifischen Gruppen.

4.1.5 Diskussion

Anknüpfend an bisherige Längsschnittstudien zu Leistungsentwicklungen in der gymnasialen Oberstufe (vgl. z.B. Baumert et al., 1996; Lehmann et al., 2004, 2006), wurden in dieser explorativen Studie Schülerinnen und Schülern in der Einführungsphase der gymnasialen Oberstufe untersucht. In zwei curricularen Kernfächern der gymnasialen Oberstufe (Mathematik und Englisch) wurden Leistungsentwicklungen mit Hilfe standardisierter Leistungstests analysiert und fachspezifische Unterschiede aufgezeigt.

Entgegen bisheriger Studien, in denen die Veränderung der mathematischen Grundbildung innerhalb der gymnasialen Oberstufe untersucht worden war (z.B. Ramseier et al., 1999; Lehmann et al., 2004, 2006), resultierten in dieser Studie im Mittel abnehmende Leistungen im Verlauf der Einführungsphase. Ein möglicher Grund hierfür könnte in einer gewissen Diskrepanz zwischen den im Unterricht bereit gestellten Lerngelegenheiten und den durch den Test zur mathematischen Grundbildung erfassten Leistungen liegen. So ist in den Rahmenrichtlinien im Fach Mathematik für die Einführungsphase festgelegt, dass Oberstufenmathematik unterrichtet wird (Niedersächsisches Kultusministerium, 1990). Andererseits gilt mathematische Grundbildung als Voraussetzung der Oberstufenmathematik. Somit ist eine gewisse Überschneidung beider Inhalte erwartbar. Ebenso sollte lt. Rahmenrichtlinien im Unterricht ein Rückgriff auf Inhalte der Grundbildung und der Anwendungsbezug auf die Praxis stattgefunden haben. In Hinblick auf die Ergebnisse dieser Studie schienen diese Aufgaben und Ziele jedoch im zweiten Halbjahr der Einführungsphase keine hohe Unterrichtsrelevanz gehabt zu haben. Ob mathematische Grundbildung vor allem im ersten Schulhalbjahr bedeutsam ist und/oder im Verlauf der Kursstufe (erneut) relevant wird, sollte Gegenstand weiterer Studien sein. Auf Basis der Ergebnisse dieser Studie lässt sich die am Ende der gymnasialen Oberstufe erhobene mathematischer Grundbildung (z. B. Köller, Watermann et al., 2004) eher als ein Minimum an Eingangsvoraussetzungen ansehen.

In Englisch erhöhten sich, konform zu bisherigen Studien in der gymnasialen Oberstufe, die mittleren Leistungen. Ergänzend konnten im Rahmen dieser Studie positive Leistungsentwicklungen bereits in der Einführungsphase nachgewiesen werden – und dies in einem relativ engen Zeitfenster von einem halben Schuljahr. Eine mögliche Erklärung für den großen Leistungszuwachs lässt sich unter Rückgriff auf die Rahmenrichtlinien für die Einführungsphase im Fach Englisch (Niedersächsisches Kultusministerium, 2003) geben. So stehen im Wesentlichen Grammatik und eine vertiefende Betrachtung aus der Sekundarstufe I gelernter Inhalte an. Diese Aspekte sind auch zentrale Erfassungsbereiche der TOEFL-Skalen (insbesondere Grammatik und Orthografie sowie Vokabelwissen und Leseverständnis). Wie sich die Leistungsentwicklungen in Mathematik und Englisch in der Kursstufe fortsetzen und ob sich weitere fachspezifische Unterschiede in den Entwicklungsverläufen aufzeigen lassen, sollte in zusätzlichen Studien nachgegangen werden.

Eine differenziertere Beschreibung der Leistungsentwicklungen innerhalb der Einführungsphase erfolgte hier über eine kursniveauspezifische Betrachtung (Kurse mit grundlegenden oder erhöhten Anforderungen). In verschiedenen Schulleistungsstudien zeigten sich durchgängig große kursniveauspezifische Unterschiede am Ende der gymnasialen Oberstufe (vgl. Köller, Watermann et al., 2004; Lehmann et al., 2006). Erwartungsgemäß zeigten sich auch in dieser Studie für die Einführungsphase Leistungsunterschiede zwischen den Gruppen in beiden Fächern und zu beiden Erhebungen. Während einerseits vermutet werden konnte, dass sich die Leistungsschere bereits im Verlauf der Einführungsphase öffnete (vgl. Köller et al., 2006), war es andererseits auch ein in den Rahmenrichtlinien der Einführungsphase festgelegtes Ziel, alle Schülerinnen und Schüler auf die Kursstufe vorzubereiten und evtl. Vorwissensdefizite auszugleichen (vgl. Niedersächsisches Kultusministerium, 1990, 2003). In dieser Studie zeigte sich sowohl für Mathematik als auch für Englisch, dass die Leistungsentwicklungen innerhalb eines halben Schuljahres zwischen den Schülerinnen und Schülern mit angestrebtem höherem Kursniveau vergleichbar waren mit denjenigen, die einen Kurs mit grundlegenden Anforderungen ab der Kursstufe belegen wollten. Diese Entwicklungen entgegen der oftmals in der Kursstufe beobachteten Leistungsdifferenzierung sprechen zum einen für eine gewisse Umsetzung des in den Rahmenrichtlinien formulierten Ziels, alle Schülerinnen und Schüler möglichst gut auf die anstehende Kursstufe vorzubereiten. Während im Fach Englisch auf deskriptiver Ebene ein weitgehend paralleler Entwicklungsverlauf zwischen beiden Gruppen resultierte, deuteten sich in Mathematik in den deskriptiven Daten bereits für den Zeitraum eines halben Schuljahres gegenläufige Trends an. So sanken die mittleren Leistungswerte der Schülerinnen und Schüler bei angestrebtem Kurs mit grundlegenden Anforderungen bis zum Ende der Einführungsphase deutlich ab, wohingegen die durchschnittliche Leistung der anderen auf deskriptiver Ebene weitgehend konstant blieb. Ein Ausgleich von Defiziten bei unterschiedlichen Eingangsvoraussetzungen zeigte sich in Mathematik somit nicht. Vielmehr liegen mit Beginn der Qualifikationsphase weitere Leistungsdifferenzierungen im Verlauf der gymnasialen Oberstufe nahe (vgl. Köller et al., 2006). Deshalb erscheint es im Fach Mathematik wünschenswert, mögliche intrapersonelle und unterrichtliche Ursachen vermehrt zu berücksichtigen, um einen Leistungsabfall der Schülerinnen und Schüler mit angestrebtem Kurs mit grundlegenden Anforderungen bereits in der Einführungsphase entgegenzu-

wirken und damit auch die in den Rahmenrichtlinien formulierten Unterrichtsziele umfassender zu verfolgen.

Mit dieser explorativen Studie wurde ein erster Einblick in die fachspezifischen Leistungen und Leistungsentwicklungen von Schülerinnen und Schülern an einem Fachgymnasium innerhalb der Einführungsphase gegeben. Eine Einschränkung der Aussagekraft der Studie ist sicherlich darin zu sehen, dass lediglich zwei Jahrgänge an einem Fachgymnasium Wirtschaft in Niedersachsen einbezogen werden konnten. Trotz unterschiedlicher Lernausgangslage im Fach Englisch resultierten vergleichbare Leistungsentwicklungen innerhalb der Einführungsphase. Zum einen war es somit möglich, Entwicklungsverläufe zusammengefasst für beide Kohorten zu beschreiben. Zum anderen kann ein gewisser Anspruch auf Verallgemeinerbarkeit der Leistungsentwicklungen innerhalb der untersuchten Schule gestellt werden. Einschränkend ist hinzuzufügen, dass für die aufgezeigten Entwicklungsverläufe der Leistungen auch das für Fachgymnasien unübliche Vorgehen einer frühen Leistungs-/Interessengruppierung ab der 11. Jahrgangsstufe mit entscheidend gewesen sein könnte, da die Zusammensetzung der Klassen bereits einer Kursstufe ähnelte. Inwieweit das Befundmuster auf Besonderheiten der untersuchten Schule oder der beruflichen Ausrichtung (Wirtschaft) der Schulform zurückgeführt werden kann, ist offen. Von daher verbietet sich eine Generalisierung der Ergebnisse, solange die Befunde nicht an einer unabhängigen Stichprobe repliziert werden. Aufgrund der kleinen Stichprobe ist es zudem nicht möglich gewesen, Clustereffekte zu berücksichtigen. Clustereffekte treten dann auf, wenn sich Schülerinnen und Schüler in den Klassen ähnlicher sind als zwischen den Klassen. Infolge dessen werden Standardfehler unterschätzt und Signifikanzprüfung tendenziell zu liberal angesetzt. Insbesondere bei leistungsgruppierten Klassen wie in der Kursstufe als auch aufgrund der speziellen Zusammensetzung der Klassen an dieser Schule lassen sich Clustereffekte vermuten. In umfangreicheren Studien, die über den explorativen Charakter dieser Untersuchung hinausgehen, sollte diese Datenstruktur bei der Auswertung berücksichtigt werden. Da eine möglichst objektive Erfassung von Leistungsentwicklungen in der gymnasialen Oberstufe mit Hilfe standardisierter Tests erfolgen sollte, beschränkte sich die Auswahl auf die Fächer Mathematik und Englisch. Um Leistungsentwicklungen in weiteren Fächern zu erfassen, ist es wünschenswert, dass entsprechende Testverfahren in Zukunft entwickelt und erprobt werden.

4.2 The Relations of Vocational Interests and Mathematical Literacy: On the Predictive Power of Interest Profiles⁶

Abstract

This study examines the relations of vocational interests and mathematical literacy both cross-sectionally and longitudinally. It extends previous research by comparing results based on Holland's RIASEC scale scores with results from a reductionist approach based on individual interest profiles (level, differentiation, and orientation). Both sets of analyses found significant relations between interests and mathematical literacy. The scale score analyses revealed positive associations of Realistic interests and negative associations of Artistic interests with mathematical literacy. Interest profiles derived from a dimensional representation showed that individuals with interest orientations close to the Realistic domain scored highest on the test of mathematical literacy, and that those with interests between the Artistic and Social domains scored lowest. Results from profile analyses suggest that interest differentiation moderates the interest–ability relation. Interest profiles proved to be more robust predictors than the scale score models, as associations were still significant when covariates were controlled. Findings suggest that interest profiles are a valid reduction of the scale score model.

4.2.1 Theoretical Background

Vocational interests develop in adolescence (Tracey, 2001) and become increasingly stable over time (Low, Yoon, Roberts & Rounds, 2005). Holland (1959, 1997) proposed a model that characterizes vocational interests in terms of six orientations: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. Individuals can be classified to their dominant interest orientation (i.e., as a R, I, A, S, E, or C type) based on their responses to an interest inventory. According to Holland's RIASEC model, these interest orientations are associated with different values, attitudes, and

⁶ Dieser Abschnitt entspricht einem zur Veröffentlichung angenommenen Manuskript im *Journal of Career Assessment* (Warwas, Nagy, Watermann & Hasselhorn, in press).

abilities. For example, people with strong interests in the Artistic domain are expected to shine in languages and the arts, whereas people with strong interests in the Investigative and Realistic domains are expected to show high abilities in mathematics. Holland has suggested that these interest–ability relations develop as a result of the interaction of genetic and environmental factors. Ackerman's (1996) *Process, Personality, Interests, and Knowledge Theory of Intellectual Development* provides a theoretical framework that specifies the mechanisms underlying the postulated relationships. According to his theory, interests and abilities develop in a reciprocal relationship. Domain-specific interest prompts people to engage more intensely with the object of interest, leading to an increase in the corresponding skills and abilities. Should they lack the necessary skills and abilities, however, repeated experiences of failure will lead to a decrease in interest.

Several empirical studies have investigated the associations between Holland's vocational interest types and domain-specific abilities. Many findings were synthesized in the meta-analysis by Ackerman and Heggstad (1997). The authors found consistent positive relations between, for example, interests in the Realistic or Investigative domains and mathematical abilities, and interests in the Investigative or Artistic domains and verbal abilities. A more recent study investigated time-lagged relations between RIASEC scores and academic abilities in the domains of mathematics and English (Tracey et al., 2005). Contrary to expectations, however, no time-lagged relations were found between vocational interests and academic skills.

Strong (1955) pointed out that the correlations of interest and ability test scores may not measure the true association between interests and abilities and that the data are too complex to be adequately described by the methodologies previously employed (cf. Randahl, 1991). Therefore, instead of focusing on the relationship between single interest and ability *measures*, other authors have examined the relationship between interest and ability *profiles*. Randahl (1991) presented evidence to show that different interest profiles are associated with distinct ability profiles. In a more recent study, Reeve and Hakel (2000) identified significant intraindividual correlations that increased with the age of the respondents (see also Denissen, Zarret & Eccles, 2007). Although these studies provide evidence for associations between interest profiles and abilities, they do not consider the systematic pattern of RIASEC profiles proposed by Holland (1997).

RIASEC Structure and RIASEC Profiles. According to Holland (1997), individuals differ in their interest profiles, but all profiles are organized according to the same principles. In his hexagonal model, the six interests are represented in a two-dimensional space, with psychological similarity between the orientations determining their relative position and proximity on the hexagon: Interest domains are located next to those to which they are most similar and opposite those to which they are most dissimilar. The scale scores thus show a specific pattern of correlations: the relationships of adjacent types (e.g., R–I) should be higher than between alternating types (e.g., R–A), which in turn should be higher than those of opposite types (e.g., R–S). There is strong empirical support for this hexagonal structure (see Nagy et al., in press; Rounds & Tracey, 1993; Tracey & Rounds, 1993). Figure 1A provides a graphical representation of Holland’s structural model.

Holland proposed that three characteristics of individual RIASEC profiles be considered: consistency, orientation, and differentiation. First, *consistency* refers to the placement agreement of individual’s highest interest domains. Holland uses the two highest scale scores to define three degrees of consistency: Adjacent orientations (e.g., R–I) are most consistent, alternating orientations (e.g., R–A) have an intermediate level of consistency and opposite types (e.g., R–S) are least consistent. Second, an individual’s interest *orientation* is probably the most salient feature of a RIASEC profile. Interest orientation refers to an individual’s dominant or highest interest domain. Figure 1B gives an example of a hypothetical RIASEC profile with an interest orientation in the Investigative domain. Finally, interest *differentiation* indicates the extent to which a person’s dominant interests stand out from her or his other interests. In a highly differentiated interest profile, individual preferences are easily identified. In less differentiated profiles, preferences become more difficult to identify. Holland expected differentiation to function as a moderator, particularly of the association between outcome and person–environment congruence, with stronger relationships being expected for people with differentiated profiles than for those with undifferentiated profiles. Figure 1B illustrates two consistent individual scale score profiles, both with a dominant Investigative orientation, but one with low and the other with high differentiation.

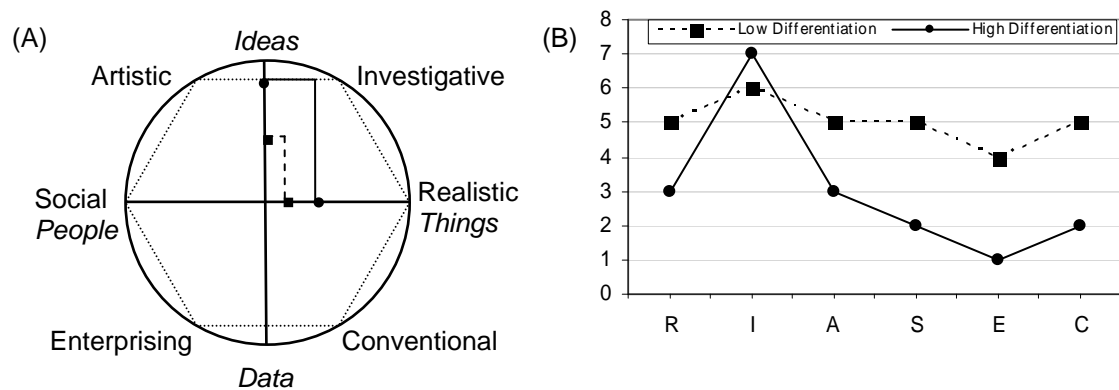


Figure 1. Interest profiles based on Holland's RIASEC model: (A) representation based on Prediger's (1982) Things/People and Data/Ideas dimensions and (B) scale score model.

Structural Summary of Individual RIASEC Profiles. According to Prediger's (1982) dimensional model, the spatial RIASEC structure can be represented by a three-factor solution. Factor analyses typically reveal a prominent general factor, with comparably high loadings on all six scales (see Rounds & Tracey, 1993; Tracey & Rounds, 1993). The structural aspect of Holland's model is reflected in two additional basic dimensions: Things/People (horizontal axis in Figure 1A) and Data/Ideas (vertical axis in Figure 1A). Plotting the loadings of the Things/People and Data/Ideas factors typically produces the graphical representation given in Figure 1A (Tracey, 2000).

As outlined by Nagy and colleagues (in press; see also Tracey & Robbins, 2005, 2006) Prediger's (1982) factors capture the most fundamental aspects of individuals' RIASEC profiles. The general factor reflects interindividual differences in the overall level of interests (i.e., the mean interest defined over all RIASEC domains). It has been given a number of labels, including "acquiescent style" (Holland, 1985, p. 5), "response style" (Prediger, 1998, p. 205), or "general interest factor" (Darcy & Tracey, 2003, p. 228).

Individuals' scores on the Things/People and Data/Ideas dimensions provide information about their interest orientation and interest differentiation (Nagy et al., in press). As shown by Nagy and colleagues, the Things/People and Data/Ideas factor scores can be regarded as a multiplicative function of individuals' interest differentiation and the sine or cosine of their interest orientation. This relation is visualized in

Figure 1A, which plots the scores of two hypothetical persons on Prediger's factors. Both persons are located the Investigative corner of the hexagon. However, they differ in their interest differentiation: one is located close to the origin of the Things/People and Data/Ideas coordinate system (i.e., low differentiation), whereas the other is located close to the edge of the coordinate system (i.e., high differentiation). For more details, see Nagy et al. (in press).

Taken together, the three dimensions proposed by Prediger (1982) provide a reductionist representation of RIASEC profiles according to the principles proposed by Holland (1997). Dimensional scores thus provide a structural summary (Gurtman & Balakrishnan, 1998) of individuals' interest profiles. As a consequence, interest profiles can be related to external variables, such as mathematical literacy, making it possible to evaluate the predictive power of each individual profile parameter. To our knowledge, however, research taking such an approach is currently scarce.

Prediger (1998) examined the role of interest level in predicting occupational choices, but found interest level to be unrelated to behavior. However, in a theoretical article, Darcy and Tracey (2003) interpreted the general interest factor as an indicator of interest *flexibility* and expected it to show moderate positive relationships to outcome measures such as achievement, satisfaction, and stability. Moreover, they hypothesized interest flexibility to be a moderator of the relations between person–environment congruence and outcome measures, postulating that individuals with a high profile level would exhibit a lower correlation between person–environment congruence and outcome measures than individuals with a low profile level. Tracey and Robbins (2006) confirmed this moderation hypothesis using the overall profile level, calculated as the mean score across all six RIASEC scales. However, they did not find profile level to have a direct effect on college success or persistence.

Recently, Nagy (2006) related individuals' profile orientation and interest differentiation as reflected in the Things/People and Data/Ideas factors to abilities in the mathematical and verbal domain and found the theoretically predicted correlations. Analyses showed that the closer an individual's interest orientation was to the Realistic and Investigative domains, the higher his or her mathematics achievement. Verbal abilities, in turn, were highest for individuals with dominant interests in the Artistic and Social domains.

4.2.2 Research Questions

This study investigates the associations between vocational interests and mathematical abilities from both a cross-sectional and a longitudinal perspective. In so doing, it follows up on the few studies to date that have examined the association between profiles of vocational interests and abilities (e.g., Nagy, 2006; Randahl, 1991; Tracey & Robbins, 2006). Although six scale score models use information from all RIASEC variables, the underlying dimensional arrangement means that the scale scores show a specific pattern of correlations, and thus include redundant information. Interest profiles based on a structural summary (Gurtman & Balakrishnan, 1998) allow vocational interests to be described more parsimoniously. This method seems particularly attractive and valuable because it further allows aspects of interest orientations that predict a person's abilities positively or negatively to be identified. The validity of this reductionist method remains to be confirmed, however.

In order to examine the robustness of the associations between interest profiles and mathematical abilities, we controlled for relevant covariates. Previous research has shown consistent gender differences in both mathematical abilities and vocational interests. A large body of research with young adults has shown that males outperform females in mathematics, especially in mathematical problem-solving tasks (Geary, 1996; Hosenfeld, Köller & Baumert, 1999). At the same time, previous studies on vocational interests have repeatedly found gender differences in interest orientations. As a rule, men show more interest in the Realistic domain, and women more interest in the Social domain (Lippa, 1998; Mullis et al., 1998; Nagy, 2006; Nagy et al., in press; Tracey & Ward, 1998). In our longitudinal analyses, moreover, we controlled for prior knowledge, which is known to be a strong predictor of future knowledge (Alexander & Judy, 1988; Dochy, 1992). Specifically, we expected to find the following patterns of results:

1. We expected the scale score models to show positive associations between interests in the Realistic and Investigative domains and mathematical literacy both cross-sectionally and longitudinally. Furthermore, because of the specific structure of Holland's RIASEC model, we expected to find negative associations of Artistic and Social interests.

2. We expected interest profiles to be associated with mathematical literacy in the cross-sectional and the longitudinal analyses. It was not possible to draw on an existing body of findings to formulate specific predictions for interest level. However, we expected to find higher mathematical literacy at both points of measurement with increasing proximity of individuals' dimensionally represented orientation to the Realistic and Investigative domains. We also assumed differentiation to be an important profile parameter that moderates the interest–ability relation. We therefore expected models in which profile orientation was weighted by differentiation to provide better predictions of mathematical literacy than models including unweighted profile orientation.
3. A main focus of our study was to examine the validity of the reductionist interest profile method. This method provides a valuable tool in career counselling and assessment because it allows aspects of interest orientations that predict a person's mathematical literacy positively or negatively to be specified. We expected the predictive power of interest profiles to be comparable to that of the scale score models; in other words, we expected the reduction to summarize the most relevant interest information.
4. In order to examine the robustness of the results, we analyzed whether there was still a positive association of vocational interests and mathematical literacy when gender and prior abilities in mathematical literacy were controlled. We further investigated the incremental effect of vocational interests beyond these covariates.

4.2.3 Method

Sample and Procedures. All participants were 11th graders from a German academic-track school specializing in economics. The students were administered a mathematics test in the middle and at the end of grade 11, at an interval of approximately 6 months. Students also completed a questionnaire tapping their vocational interests at the first point of measurement. The sample comprised 168 students at the first point of measurement (t1) and 128 students at the second point of measurement (t2). The decrease in participants was caused by illness, students dropping out of school or moving to new

schools, and refusal to participate at t2. Participants' average age at t1 was 18.15 years ($SD = 1.49$); 46% of participants were female.

Missing data caused by systematic dropout limit the generalizability of longitudinal findings (Allison, 2002). In recent research (Collins et al., 2001), a variety of algorithms have been proposed for dealing with missing data. There is growing consensus that the expectation-maximization algorithm and multiple imputation produce less biased estimates than do pairwise or listwise deletion. We therefore used multiple imputation methods to estimate missing values in the present study. The NORM 2.03 software (Schafer, 1999) was used to generate 20 datasets and Rubin's rules were applied to combine estimates and standard errors (cf. Rubin, 1987; Schafer & Graham, 2002).

Instruments. Mathematical literacy was assessed using the mathematical literacy test developed and administered in TIMSS (Mullis et al., 1998). In line with the literacy debate (NCTM, 1989), the mathematical content covered in the test items is embedded in everyday contexts. Although the literacy concept does not require curricular validity in the strict sense, validity studies have confirmed the TIMSS items' validity with respect to both the curriculum and the learning opportunities afforded in the classroom (Klieme et al., 2000). The mathematical literacy test is based on item response theory; the item parameters derived from TIMSS were used to scale the test scores. We used Conquest (Wu et al., 1998) to calculate weighted likelihood estimates (WLE, Warm, 1989) as person parameters for each student. The items used at both occasions (see Appendix A) were characterized by TIMSS experts as especially relevant for defining mathematical literacy and covered a broad range of item difficulty. This procedure resulted in a broad achievement test with good content validity. However, levels of internal consistency were lower than in more narrowly defined achievement tests. The reliabilities of the WLE scores were estimated by dividing the measures of "true" variance (variance of the mathematics factor) by the variance of the WLE scores (see Rost, 2004, Formula 6, p. 381). Reliability was $r_{tt} = .65$ at t1 and $r_{tt} = .50$ at t2.

The scores were standardized at $M = 0$ and $SD = 1$ at t1 and were transformed using this mean and standard deviation at t2 ($M = -0.21$, $SD = 1.42$). Mean abilities in mathematical literacy were thus found to decrease somewhat, but the difference was not statistically significant ($z = 1.73$, $p = .083$). The test scores showed a rank order stability of $r = .56$ ($p < .01$).

Vocational interests were assessed by means of the *Revised General Interest Structure Test* (GIST; Allgemeiner Interessen Strukturtest; Bergmann & Eder, 2005), an established German instrument based on Holland's model. The GIST is the best validated interest inventory in the German-speaking countries (i.e., Germany, Austria, and Switzerland). Scale score correlations between the GIST and an adaptation of Holland's Self-Directed Search instrument (Jörin, Stoll, Bergmann & Eder, 2004) range from $r = .60$ to $r = .75$. Nagy and colleagues (in press) recently presented evidence for the structural validity of the GIST, showing that the inventory produces the same rough hexagonal structure as other RIASEC measures. The 1-month retest reliability ranges from $r = .85$ to $r = .92$ (Bergmann & Eder, 2005). The test comprises 60 items, 10 for each of the 6 interest dimensions. Each item describes a school-related or occupational activity; respondents are asked to state how interested they are or would be in a specific activity on a 5-point Likert scale (1 = *not at all*; 5 = *very*). The internal consistencies (Cronbach's alpha) of the six scales in the present sample were $\alpha = .77$ or higher. The sample-specific means of the RIASEC scales were standardized against normative data secured from $N = 2,716$ students at traditional academic-track *Gymnasium* schools (Köller, Watermann et al., 2004).

Derived Variables. Vocational interests according to Holland's RIASEC model were calculated in two ways: one using the six scale scores and the other based on interest profiles as comprised in a structural summary (Gurtman & Balakrishnan, 1998). Profile *level* was calculated by computing the mean of the six scale scores. Profile *orientation* and *differentiation* were derived from Things/People and Data/Ideas dimensional scores (cf. Prediger, 1982; Prediger & Vansickle, 1992). To this end, we first transferred the standardized scale scores to a coordinate system described by Prediger's (1982) Things/People (cosine) and Data/Ideas (sine) axes. We assumed a regular hexagon, with the six interest scales being uniformly distributed around a circle. Each of the six RIASEC variables was given a fixed angular location (Realistic = 0° , Investigative = 60° , Artistic = 120° , and so on). To calculate the dimensional scores, we multiplied the six standardized scale scores by the cosine of the corresponding variable's angular position, and then summed them to a (weighted) *cosine score* (i.e., Things/People). The (weighted) *sine score* (i.e., Data/Ideas) was calculated in the same way, but using the sine of the corresponding variable's position. For instance, positive

scores on the Things/People (cosine) and Data/Ideas (sine) dimensions represent an Investigative orientation, whereas negative scores on these dimensions relate to a Conventional orientation (see Figure 1A). Both dimensional scores include information about profile orientation *and* differentiation.

As outlined by Nagy et al. (in press), an individual's i Things/People score can be denoted as $\alpha_i \times \cos(\delta_i)$, and his or her Data/Ideas score as $\alpha_i \times \sin(\delta_i)$, where α_i denotes the interest differentiation (amplitude) and δ_i the interest orientation. Note, that δ is an angular measure (ranging from 0° to 360°) that indicates the location of a person's dominant interest. The specified relations between dimensional Things/People and Data/Ideas scores make it possible to disentangle profile differentiation from profile orientation. We use this decomposition to examine the moderating role of profile differentiation on the interest–outcome relation (see below).

Analysis Strategy. Multiple regression analyses were performed to predict mathematical literacy. In regression models with scale scores, we used the six RIASEC variables as predictors of mathematical literacy. In the reductionist interest profile method, the predictor variables were interest level and the dimensional scores on the Things/People and Data/Ideas axes.

The Things/People and Data/Ideas scores were used as predictors in linear regression models (see also Batschelet, 1981). The resulting regression coefficients associated with the Prediger dimensions b_0 (intercept), b_l (level), b_c (Things/People), and b_s (Data/Ideas) describe the relation between individual interest profiles and outcome measures. The first parameter b_0 is a conventional regression intercept. The level effect (b_l) captures change in mathematical literacy as a function of individuals' overall interest level. Finally, the parameters b_c and b_s reflect the relations of individuals' interest orientations. These regression coefficients can be used to define two key parameters. First, the peak θ represents the interest orientation (in degree units) that is most positively associated with a given criterion. It is calculated by taking the arc tangent of the regression coefficient of the Data/Ideas (b_s) scores divided by the regression coefficient of the Things/People (b_c) scores, $\theta = \arctan\left(\frac{b_s}{b_c}\right)$. The parameter θ stands for the interest region in which the highest level of mathematical literacy is expected to

occur. Based on previous research, we expect an estimate of θ close to 0° , the angular location of realistic interests.

Second, the parameter A stands for the effect amplitude. This parameter reflects the difference in the expected outcome measures between persons with interest orientations equal to θ (i.e., the peak) and the mean value of the outcome across all possible interest orientations. A is calculated as the root of the sum of the squared coefficient of the Things/People (b_c) and Data/Ideas (b_s) dimensional scores, $A = \sqrt{b_c^2 + b_s^2}$. When an individual's profile orientation equals the peak, the amplitude has reached its maximum. The amplitude follows a sinusoidal path, decreasing until it reaches its maximal distance—in terms of a two-dimensional representation—from the peak and then increasing again with proximity to the peak.

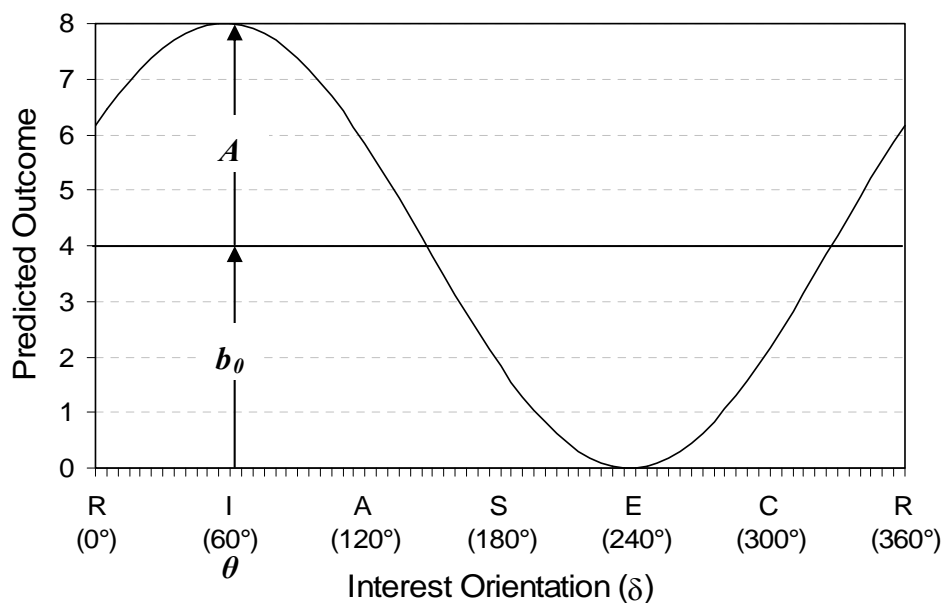


Figure 2. Representation of model parameters in interest profiles. The mean outcome value across the entire range is $b_0 = 4$. The peak θ (about 60° , I) reflects the orientation that is most positively associated with the outcome. When an individual's interest orientation δ equals the peak, the amplitude has reached its maximum with $A = 4$. The amplitude and thus the outcome follows a sinusoidal path, decreasing until it reaches its maximal distance—in terms of a two-dimensional representation—from the peak and then increasing again with proximity to the peak.

Figure 2 gives an example of an outcome of the regression model described. Here, individuals' interest orientations (ranging from 0° to 360°) are given on the x-axis, and the outcome measure is represented on the y-axis. The mean outcome value in the example is $b_0 = 4$. The highest outcome value is expected to occur at $\theta = 60^\circ$, with an effect amplitude of $A = 4$. Note that Figure 2 does not visualize the effect of the interest level (b_i), because this would unduly complicate the diagram. See the Appendix C for more details on the method used.

The dimensional scores used derive from the reductionist method. In order to contrast the predictive power of this method with that of the scale score model, we compared the R^2 values exploratively in terms of confidence intervals. We derived nonsymmetrical 95% confidence intervals using the bootstrap methods implemented in Mplus 5.1 (Muthén & Muthén, 1998–2008). Specifically, we computed the 95% intervals for each imputation, and then averaged the results over 20 imputations.

Finally, we included additional predictors in our profile analyses to examine the robustness of the associations between interests and mathematical literacy. We used the ΔR^2 statistic to evaluate whether interests contribute to the prediction of mathematical literacy above and beyond the covariates gender and prior abilities. Statistical significance of the derived ΔR^2 was evaluated by estimating regression models including covariates in which the effects of interest were fixed to zero. We converted the derived χ^2 values into F statistics (cf. Allison, 2002, p. 68). Significant F values indicate non-zero incremental effects of interests beyond the covariates.

4.2.4 Results

Table 5 documents the correlations between the RIASEC scale scores, mathematical literacy at the first (t1) and the second point of measurement (t2), and gender. For the most part, the intercorrelations of the RIASEC scale scores were in line with the pattern typical of the hexagonal structure. Adjacent scales, such as Realistic and Investigative ($r = .61$) or Realistic and Conventional ($r = .33$), showed higher correlations than nonadjacent scales, such as Realistic and Artistic ($r = .26$) or Realistic and Enterprising ($r = .21$); the lowest correlations were found between opposite scales, such as Realistic and Social ($r = .08$).

Table 5
Means, Standard Deviations, and Correlations Between Interest Scale Scores, Mathematical Literacy (ML) at t1 and t2, and Gender

Variable	1.	2.	3.	4.	5.	6.	7.	8.	<i>M</i>	<i>SD</i>
1. R									0.21	0.86
2. I	.61**								-0.11	0.82
3. A	.26**	.35**							-0.13	0.96
4. S	.08	.17*	.50**						-0.21	0.86
5. E	.21**	.29**	.26**	.51**					0.18	1.06
6. C	.33**	.36**	.12	.27**	.43**				0.48	0.97
7. ML t1	.18*	.10	-.18*	-.12	.01	.02			0.00	1.00
8. ML t2	.17*	.08	-.24**	-.23**	-.07	.03	.58**		-0.21	1.42
9. Gender	.30**	.15*	-.27**	-.33**	.07	-.02	.31**	.28**		

Note. Higher coding number for males.

* $p < .05$. ** $p < .01$.

As shown in Table 5, the RIASEC scale scores showed a systematic pattern of correlations with mathematical literacy and with gender. Mathematical literacy correlated positively with a Realistic orientation at t1 ($r = .18$) and t2 ($r = .17$) and negatively with Artistic interests at both times (t1: $r = -.24$, t2: $r = -.18$). Furthermore, mathematical literacy at t2 was negatively associated with a Social orientation ($r = -.23$). As expected, the correlations between gender and scale scores revealed particularly high interests for men in the Realistic domain ($r = .31$) and for women in the Social ($r = -.33$) and Artistic ($r = -.27$) domain. The means of the scale scores reflected the school's focus on economics. The students showed higher interests in the Realistic, Conventional, and Economic domains, whereas their interests in Investigative, Artistic, and Social activities were lower than those of students in traditional academic-track schools.

In order to test for the structural pattern of the RIASEC scale scores, we tested the fit of our measures to the hypothesized RIASEC structure using RANDALL (Tracey,

1997). This program conducts a randomization test and provides a correspondence index (CI). The randomization test (Hubert & Arabie, 1987) yields a significance level for the number of order predictions met by the data compared against a null conjecture of random ordering. The CI is an interpretative aid and has a numeric range from -1 to $+1$, with values close to 0 indicating a random fit of the data and a value of $+1$ indicating a perfect fit. The results revealed a very good fit ($p = .0167$, $CI = .81$).

Table 6

Correlations Between Interest Profile Parameters (Level, Unweighted Interest Orientation, and Interest Orientation Weighted by Differentiation) and Mathematical Literacy (ML) at t1 and t2, and with Gender

	ML t1	ML t2	Gender
Level	.00	-.08	-.03
Things/People (cosine) unweighted	.18*	.26**	.42**
Data/Ideas (sine) unweighted	-.00	-.01	-.13†
Things/People (cosine) weighted	.23**	.32**	.41**
Data/Ideas (sine) weighted	-.07	-.06	-.11

Note. Higher coding number for males.

† $p < .10$. * $p < .05$. ** $p < .01$.

Table 6 reports the correlations of the interest profile parameters derived from dimensional scores with mathematical literacy and with gender. The Things/People orientation—unweighted and weighted by differentiation—was positively associated with mathematical literacy and with gender at both times. In contrast, neither interest level nor the Data/Ideas dimension was found to be significantly correlated with mathematical literacy.

Relations of Vocational Interests and Mathematical Literacy. The aim of our study was to investigate the associations between vocational interests and mathematical literacy from both a cross-sectional and a longitudinal perspective, drawing on both interest scale scores and interest profiles. Table 3 presents the results of the multiple re-

gression analyses using scale scores. At both times, mathematical literacy was positively related to a Realistic orientation and negatively related to an Artistic orientation. The amount of variance explained was $R^2 = .092$ in the cross-sectional and $R^2 = .137$ in the longitudinal analysis.

Table 7 also includes the covariate models. As shown, gender had a substantial effect on mathematical literacy at t1. At t2, only prior abilities were significantly associated with mathematical literacy. Regression analyses revealed that only Artistic orientation showed a weak relation to mathematical literacy when gender was controlled, with a R^2 of .128 for this model. In the longitudinal analysis controlling for gender and prior abilities, no association of interest scale scores and mathematical literacy could be confirmed. The amount of variance explained by this model was $R^2 = .384$. As indicated by the ΔR^2 statistics in Table 3, the scale scores contributed only weakly to mathematical literacy above and beyond the covariates.

The simple variant of the interest profile model considered interest level and unweighted profile orientation (see Appendix C, Equation 1) as predictors of mathematical literacy. Interest level did not significantly predict mathematical literacy at either point of measurement (see Table 8). In other words, whether a student generally tended to endorse or to reject vocational activities did not directly contribute to her or his mathematical literacy. In the cross-sectional and the longitudinal analysis, consistent with our findings for the scale scores, the Things/People dimension emerged to be a significant predictor of mathematical literacy. The proportion of variance explained was $R^2 = .035$ at t1 and $R^2 = .076$ at t2.

As shown in Table 8, the angular peak was estimated to be $\theta = 353^\circ$ and $\theta = 351^\circ$ in the cross-sectional and longitudinal analysis, respectively. The effect amplitude was $A = .28$ at t1 and $A = .56$ at t2, meaning that persons with interest orientations equal to θ scored .28 and .56 points above the mean on the literacy test at t1 and t2, respectively. Figure 3 (panels A and C) gives a graphical description of the derived associations. The amplitude follows a sinusoidal path, with a maximum close to the location of Realistic interests, decreasing until it reaches its maximal distance from the peak (Artistic/Social orientation) and then increasing again with proximity to the peak.

Table 7
Multiple Regression of Mathematical Literacy (ML) at t1 and t2 on Interest Scale Scores and Covariates

	Scale Scores			Covariates			Scale Scores with Covariates		
	ML t1	ML t2	ML t1	ML t2	ML t1	ML t2	ML t1	ML t2	
	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	
R	0.24(0.11)*	0.34(0.16)*					0.15(0.12)	0.15(0.15)	
I	0.06(0.13)	0.11(0.18)					0.07(0.12)	0.05(0.16)	
A	-0.26(0.10)**	-0.37(0.14)**					-0.19(0.10)†	-0.18(0.14)	
S	-0.04(0.12)	-0.22(0.18)					0.06(0.12)	-0.18(0.18)	
E	0.05(0.09)	0.01(0.14)					-0.02(0.09)	0.04(0.13)	
C	-0.05(0.09)	-0.01(0.13)					-0.01(0.09)	0.03(0.12)	
Gender			0.62(0.15)**	0.33(0.20)			0.46(0.18)**	0.08(0.24)	
ML t1				0.77(0.11)**				0.73(0.11)**	
Constant	-0.07(0.09)	0.36(0.15)*	-0.34(0.11)**	-0.39(0.14)**			-0.28(0.12)*	-0.34(0.16)*	
R ²	.092	.137	.096	.344			.128	.384	
<i>F</i> (<i>df1</i> , <i>df2</i>)	<i>F</i> (6, 24.52) = 8.95**	<i>F</i> (6, 17.55) = 7.99**	<i>F</i> (1, 166) = 17.695**	<i>F</i> (1, 18.34) = 79.84**	<i>F</i> (7, 42.22) = 30.55**	<i>F</i> (8, 14.63) = 18.94**			
ΔR^2					.032				
<i>F</i> (<i>df1</i> , <i>df2</i>)							<i>F</i> (6, 43.21) = 1.96†	<i>F</i> (6, 20.71) = 2.25†	

Note. ΔR^2 and the corresponding *F* statistic refer to the comparison of the “Covariates” models with the corresponding “Scale Scores with Covariates” models.

† $p < .10$. * $p < .05$. ** $p < .01$.

Table 8

Multiple Regression of Mathematical Literacy (ML) at t1 and t2 on Interest Profile Parameters (Level and Unweighted Interest Orientation) and Covariates

	Interest Profiles		Interest Profiles with Covariates	
	ML t1	ML t2	ML t1	ML t2
	<i>b</i> (<i>SE</i>)	<i>b</i> (<i>SE</i>)	<i>b</i> (<i>SE</i>)	<i>b</i> (<i>SE</i>)
Level	0.00(0.13)	-0.19(0.19)	0.02(0.13)	-0.19(0.17)
T/P	0.28(0.12)**	0.56(0.17)**	0.09(0.13)	0.30(0.16)†
D/I	-0.03(0.12)	-0.08(0.17)	0.05(0.12)	-0.04(0.15)
Gender			0.58(0.17)**	0.16(0.22)
ML t1				0.76(0.10)**
Constant	-0.07(0.09)	-0.33(0.13)*	-0.32(0.11)**	-0.38(0.14)**
<i>Derived Angular Parameters</i>				
Amplitude	0.28(0.12)*	0.56(0.16)**	0.11(0.12)	0.30(0.17)†
Peak	353°	351°	29°	353°
<i>R</i> ²	.035	.076	.101	.367
<i>F</i> (<i>df1</i> , <i>df2</i>)	<i>F</i> (3, 62.91) = 5.18**	<i>F</i> (3, 20.64) = 5.86**	<i>F</i> (4, 850.99) = 69.10**	<i>F</i> (5, 15.70) = 27.34**
ΔR^2			.005	.023
<i>F</i> (<i>df1</i> , <i>df2</i>)			<i>F</i> (3, 888.52) = 0.12	<i>F</i> (3, 23.51) = 1.32

Note. T/P = Things/People score; D/I = Data/Ideas score. ΔR^2 and the corresponding *F* statistic refer to the comparison with the “Covariates” model in Table 7.

† $p < .10$. * $p < .05$. ** $p < .01$.

Table 8 also includes the covariate models. The results were similar to those reported for the scale score models. When gender and prior abilities were controlled, relations between Things/People and Data/Ideas dimensional scores and mathematical literacy at t1 were not significant. As a consequence, the effect amplitude was no longer statistically significantly different from zero. Predicting t2 mathematical literacy yielded a significant *A* parameter at the $p < .10$ level, although the ΔR^2 statistic was not significant.

Finally, we examined the power of interest profiles including level and orientation weighted by differentiation to predict mathematical literacy (see Table 9 and Appendix C, Equation 2). Note that the absolute values of the regression coefficients derived for the unweighted and weighted Things/People (cosine) and Data/Ideas (sine) orientations are not directly comparable because weighting changes the metric of the predictor variables. Interest level did not play a significant role in predicting mathematical literacy directly at either time. However, the Things/People dimension again proved to be a powerful predictor of mathematical literacy at t1 and t2. The amount of variance explained was $R^2 = .059$ at t1 and $R^2 = .114$ at t2. The angular peak was comparable at t1 ($\theta = 339^\circ$) and t2 ($\theta = 343^\circ$). These parameter estimates were again closely in line with our hypotheses. Figure 3 (panels B and D) summarizes the relations estimated between orientation, differentiation, and mathematical literacy. The highest outcomes were expected at interest orientations close to the Realistic domain, and the lowest at locations between the Artistic and Social domains. The moderating function of the amplitude is represented by including regression lines at the 25th, 50th, and 75th percentile of the amplitude distribution (see panels B and D). The higher the amplitude, the higher the change in predicted mathematical literacy as a student's profile orientation diverges from the peak. This moderating role of interest differentiation is most evident at t2.

In the last step, we included the covariates gender and prior abilities. As shown in Table 9, controlling for gender resulted in nonsignificant relations between interest profiles and mathematical literacy at t1. The same did not apply to the longitudinal analysis. Here, interest orientations predicted mathematical literacy even when prior abilities and gender were controlled. The estimated relation revealed a peak quite similar to that emerging from the analysis without covariates (i.e., $\theta = 347^\circ$), but a somewhat smaller effect amplitude. Most importantly, conceptualizing interests as weighted profiles proved to be the most robust approach because these variables significantly increased the predictive power of the baseline measures gender and prior abilities, as indicated by the significant ΔR^2 statistic and the significant effect amplitude $A = 0.16$.

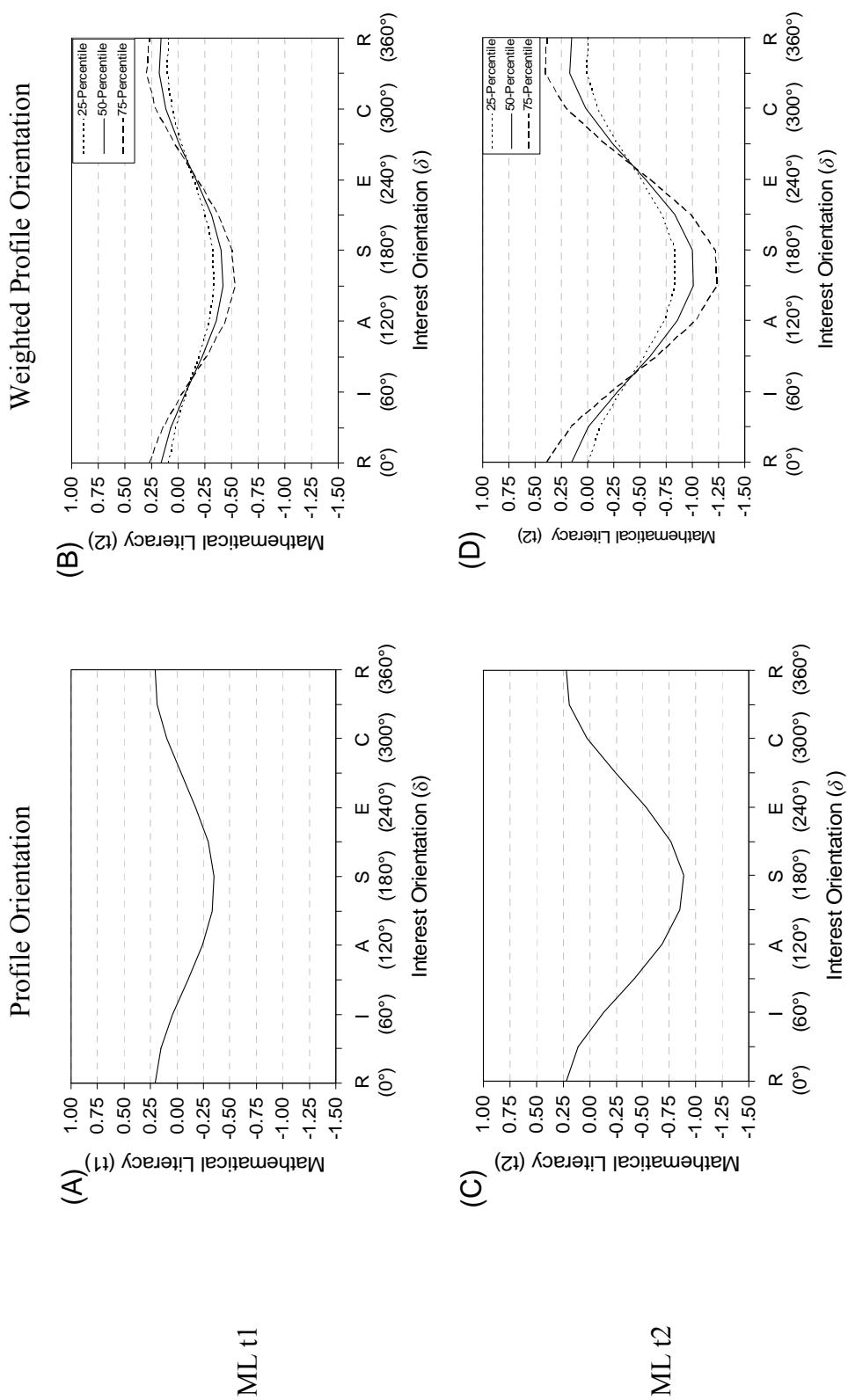


Figure 3. Predicted mathematical literacy (ML) at t1 and t2 (rows) by unweighted profile orientation and profile orientation weighted by amplitudes (columns) given by the periodic regression models. The moderating function of profile amplitudes is represented by including regression lines at the 25th, 50th, and 75th percentile of the amplitude distribution.

Table 9

Multiple Regression of Mathematical Literacy (ML) at t1 and t2 on Interest Profile Parameters (Level and Interest Orientation Weighted by Differentiation) and Covariates

	Interest Profiles		Interest Profiles with Covariates	
	ML t1	ML t2	ML t1	ML t2
	<i>b</i> (<i>SE</i>)	<i>b</i> (<i>SE</i>)	<i>b</i> (<i>SE</i>)	<i>b</i> (<i>SE</i>)
Level	0.01 (0.13)	-0.15 (0.19)	0.02 (0.13)	-0.16 (0.17)
T/P	0.14 (0.05)**	0.27 (0.07)**	0.08 (0.05)	0.15 (0.06)*
D/I	-0.05 (0.05)	-0.08 (0.08)	-0.03 (0.05)	-0.04 (0.06)
Gender			0.51 (0.16)**	0.12 (0.21)
ML t1				0.74 (0.10)**
Constant	-0.12 (0.09)	-0.42 (0.13)**	-0.45 (0.11)**	-0.38 (0.14)**
<i>Derived Angular Parameters</i>				
Amplitude	0.15 (0.05)**	0.28 (0.07)**	0.08 (0.05)	0.16 (0.06)**
Peak	339°	343°	337°	347°
R^2	.059	.114	.112	.379
$F(df1, df2)$	$F(3, 30.10) = 43.21^{**}$	$F(3, 19.29) = 10.67^{**}$	$F(4, 78.78) = 53.93^{**}$	$F(5, 15.66) = 28.01^{**}$
ΔR^2			.016	.035
$F(df1, df2)$			$F(3, 82.25) = 1.16$	$F(3, 23.32) = 3.84^*$

Note. T/P = Things/People score; D/I = Data/Ideas score. ΔR^2 and the corresponding F statistic refer to the comparison with the “Covariates” model in Table 7.

† $p < .10$. * $p < .05$. ** $p < .01$.

Validity of Interest Profiles. A main focus of our study was to examine the validity of reducing the scale score model to interest profiles. To this end, we compared the R^2 statistics provided by the different prediction models without the covariates gender and prior abilities. Figure 4 summarizes the mean R^2 estimates as well as their 95% confidence bands. The scale score models explained a greater amount of variance than the interest profiles. However, the 95% confidence regions for the models employed overlapped to a large degree. Thus, the generally higher R^2 statistics do not

necessarily mean that utilizing the reductionist profile approach results in a less valid representation of interests. For the t1 outcomes, the mean R^2 s were estimated at .092 (scale score model), .035 (unweighted profiles), and .059 (weighted profiles). A closer look at Figure 4 suggests meaningful differences in the R^2 estimates given by the scale score model and the unweighted profile models. The mean estimate given by the unweighted profile model is outside the 95% confidence region of the scale score model (.057 to .217). However, the picture looks quite different when the scale score model is compared with the weighted profile model. Here, the mean R^2 estimate given by the weighted profile model is within the confidence region of the scale score model.

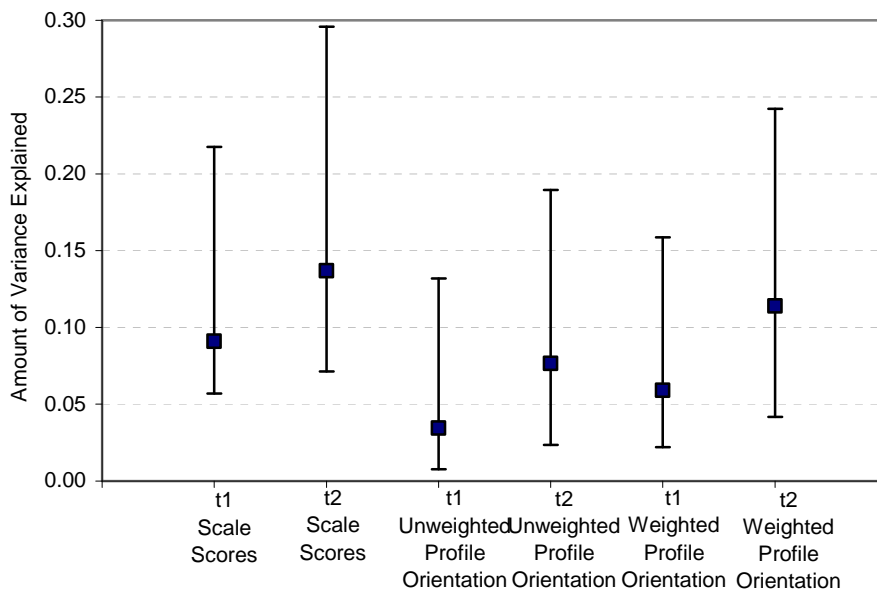


Figure 4. Amount of variance explained, R^2 , in the regression analyses with vocational interest as predictors of mathematical literacy; vertical lines depict nonsymmetrical 95% confidence intervals for the R^2 statistics.

A quite similar picture emerges when considering the R^2 statistics for the longitudinal analyses. The mean R^2 of .076 provided by the unweighted profile model was only just within the 95% confidence range of the R^2 derived from the scale score model (.071 to .296), but the mean R^2 given by the weighted profile model of .114 was well within the confidence region provided by the scale score model. This pattern suggests that the reductionist approach of using weighted profile orientations is likely to retain the most important information coded in the RIASEC scale scores. However,

disregarding the profile differentiation in the unweighted model is likely to result in significantly lowered predictive power, meaning that disregarding profile differentiation excludes an important piece of information.

Figure 4 illustrates another important point. The confidence bands derived from the scale score models are much wider than the corresponding bands given by the profile models. This means that the estimated prediction equations based on scale scores are less accurate than those of the (weighted) profile models. Furthermore, it is well known that application of regression models in small to medium-sized samples with many variables as employed here is likely to result in inflated R^2 statistics (e.g., Wherry, 1931). Hence, the results reported indicate that the weighted profile model achieves a good balance between reducing bias in R^2 estimates and enhancing the accuracy of prediction (i.e., narrowing the confidence bands).

4.2.5 Discussion

This study investigated the relations of vocational interests and mathematical literacy from both a cross-sectional and a longitudinal perspective using two different approaches: a scale score model based on Holland's RIASEC variables and interest profiles derived from a dimensional representation of Holland's model. Both methods yielded similar results. As expected, Realistic orientations were positively associated with mathematical literacy (e.g., Ackerman & Heggstad, 1997) whereas Artistic orientations were negatively associated. Findings were stable across both times of measurement. Our results thus emphasize the relevance of Realistic interests for students' mathematical literacy, further validating the focus of mathematical literacy on practical aspects that are important at work and in daily life.

One of our research objectives was to compare the scale score model with interest profiles as presented in a structural summary (cf. Gurtman & Balakrishnan, 1998) and thus to examine the validity of reducing scale scores to interest profiles. Interest profiles derived from a dimensional representation of Holland's RIASEC model seem an attractive and valuable approach, particularly because these profiles allow aspects of interest orientations that predict a person's abilities positively or negatively to be identified. As mentioned above, the interest profile and scale score methods generated

comparable results in terms of predictions of mathematical literacy. Nevertheless, when covariates were included in the regression analyses, only interest profiles were predictive for mathematical literacy above and above gender prior abilities, indicating that interest profiles are more robust predictors than scale score models. Furthermore, weighted profile models resulted in R^2 values that were well within the 95% confidence intervals provided by the scale score models. Additionally, the R^2 statistics derived from the weighted profile models appeared to be much more precise than the corresponding values determined by the scale score models.

When the six scales scores are analyzed simultaneously, the structural aspect of Holland's RIASEC model is not considered. Because the scale scores show a specific correlation pattern, multicollinearity problems may occur in multiple regression analyses, and some associations may go undetected. For instance, when inspecting the scales individually, we found a negative correlation between interests in the Social domain and mathematical literacy in the longitudinal analysis. This prediction could not be confirmed in the results of the regression analyses including scale scores as predictors of mathematical literacy. Furthermore, it is well known that application of regression models in small to medium-sized samples with many variables as employed here is likely to result in inflated R^2 statistics (e.g., Wherry, 1931). Hence, the results reported indicate that interest profiles as presented in a structural summary provide a valid representation of interests with less biased R^2 estimates.

Additional profile analyses highlighted the relevance of differentiation. Weighted profile models explained a larger amount of variance than their unweighted counterparts. This result is perfectly in line with Holland's prediction that differentiation functions as a moderating variable. As our results show (see Figure 3), relations between interest orientations are amplified in persons with a high level of interest differentiation, and greatly reduced in individuals with undifferentiated profiles. Thus, findings suggest that disregarding differentiation excludes an important piece of information. Weighted profile orientations are likely to retain the most important information coded in the RIASEC scale scores. In our study, we examined the moderating function of differentiation for the interest–ability relation in mathematical literacy. Further research is needed to examine the moderating role of differentiation for other outcome measures and interest–outcome relations.

Likewise, Tracey and Robbins (2006) highlighted the need for further examinations of moderating variables within Holland's model. For example, they investigated direct and moderating effects of profile level on different outcome measures. Profile level was the third parameter in our profile analyses. Darcy and Tracey (2003) suggested that interest level (i.e., the general factor) indicates flexibility of interests. The results of our study are consistent with the findings of Prediger (1998) and Tracey and Robbins (2006), who could not confirm a direct association between level and outcomes. The interest level parameter did not predict mathematical literacy in any of our regression analyses. Nevertheless, as proposed by Darcy and Tracey (2003) and confirmed by Tracey and Robbins (2006), the overall profile level should not be neglected in career counseling because it moderates the relationship between person–environment congruence and outcome measures.

Limitations and Outlook for Future Research. In this study, we examined the relations of vocational interests and mathematical literacy in a small sample of students from an academic-track school in Germany. Mathematical literacy is known to be a key skill both at work and in daily life (NCTM, 1989). Yet a meta-analysis reports positive associations between Realistic/Investigative interests and mathematical abilities for participants at various ages (Ackerman & Heggestad, 1997). These findings suggest that the association between Realistic interests and mathematical literacy may apply not only to school students, but to people of different ages and in different work environments. Further research is thus needed to test the stability of these findings in larger student populations as well as in non-student populations. Furthermore, the test of mathematical literacy implemented at our second point of measurement consisted of only eight items, and WLE-reliability was not optimal. However, we believe that—given its good content validity—the achievement test used provides a reasonable assessment of the concept of mathematical literacy. Indeed, the TIMSS mathematical literacy test has been proved to be a valid measure (see Mullis et al., 1998) and has been frequently used in international and national large-scale assessments. We believe that the quite low test reliability resulted in rather conservative estimates. Hence, research aiming to replicate the present results should use other and longer test forms.

As the empirical results suggest, interest profiles derived from a dimensional representation of Holland's structural model allow aspects of interest orientations that

predict a person's abilities positively or negatively to be identified. According to Ackerman (1996), a more intense engagement with the object of interest leads to an increase in the corresponding skills and abilities. Therefore, it is necessary to investigate the processes mediating students' engagement in mathematical literacy. Likewise, studies have not yet examined interests in conjunction with other constructs relevant to the enhancement of abilities, such as self-concept or self-efficacy; further research is needed here.

We examined the relations between interests and mathematical literacy using two different approaches: a scale score model and interest profiles. A promising avenue for future research would be to investigate how vocational interests predict other abilities thought to be important at work and in daily life, such as communication or reading skills. Although scale score models and interest profiles revealed quite similar results in our study, further evidence for the validity of reducing scale scores to profiles with other outcome measures is required.

This study was the first to demonstrate that vocational interests predict mathematical literacy both cross-sectionally and longitudinally, while additionally providing evidence for the validity of reducing Holland's scale score model to interest profiles.

4.3 Zum Effekt von beruflichen Interessenprofilen, Selbstkonzept und Kursniveau auf Leistungsentwicklungen in Mathematik und Englisch⁷

Zusammenfassung und Abstract

In dieser Längsschnittstudie wurde die Rolle beruflicher Interessenprofile für Leistungsentwicklungen in zwei Kernfächern der gymnasialen Oberstufe analysiert. Berufliche Interessen wurden in Anlehnung an Holland (1997) erfasst und in Profilform mit den Parametern Level sowie Things/People- und Data/Ideas-Orientierung nach Prediger (1982) modelliert. Darüber hinaus wurden mit dem fachspezifischen Selbstkonzept und dem Kursniveau weitere für die Leistungsentwicklungen als bedeutsam angenommene Variablen in den Analysen berücksichtigt und mediierende Effekte beruflicher Interessenprofile untersucht. Es zeigte sich ein positiver Effekt der Things/People-Dimension auf Leistungsentwicklungen in Mathematik, aber wider Erwarten kein Effekt beruflicher Interessenorientierungen auf die Entwicklung von Englischleistungen. Für die Veränderung der Leistungen in beiden Fächern resultierten die theoretisch erwartbaren direkten Effekte von Selbstkonzept und Kursniveau. Im Fach Mathematik mediieren berufliche Interessen die Effekte von Selbstkonzept und Kursniveau. Insgesamt zeigen sich fachspezifische Unterschiede im Zusammenspiel der Variablen, die weiterer Forschung bedürfen.

In this longitudinal study, we examined the effect of vocational interest profiles on the development of achievement in two core subjects in upper secondary level. Vocational interests according to Holland (1997) were modelled by individual profiles including the parameter level as well as the dimensions Things/People and Data/Ideas (Prediger, 1982). Furthermore, we included self-concept and course level in our analyses and investigated the mediating effects of interest profiles on the self-concept-achievement-relation and the course level-achievement-relation respectively. Our results revealed a positive effect of the Things/People dimension on the development of

⁷ Dieser Abschnitt basiert auf einem in der *Schweizerischen Zeitschrift für Bildungswissenschaften* eingereichten Manuskript (Warwas, Watermann, Nagy & Hasselhorn, eingereicht).

mathematics achievement. Contrary to our assumptions, we found no association between vocational interests and the development of achievement in English. We confirmed significant positive effects of self-concept and course level on the change in achievement in both subjects. Only in mathematics vocational interests turned out to mediate these effects. In sum, this study revealed subject-specific differences in the effects of vocational interest profiles on the development of achievement.

4.3.1 Theoretischer Hintergrund

Im Bereich individueller Faktoren gehören Interessen zu den empirisch umfangreich erforschten und vielfach bestätigten Determinanten schulischen Lernens (z. B. Krapp, 1998, 2002). Erste Längsschnittstudien belegen etwa einen Effekt von Interessen auf Mathematikleistungen in der gymnasialen Oberstufe unter Kontrolle der Ausgangswerte (z. B. Köller, Baumert et al., 2000; Köller et al., 2006). Allerdings sind Längsschnittstudien zum Zusammenspiel von Interesse, Selbstkonzept, Kurswahl und Schulleistung noch immer Mangelware (vgl. Köller et al., 2006). So ist bislang wenig darüber bekannt, wie sich das für Studienfach- und Berufswahlen als zentral angesehene berufliche Interesse in die Befunde zu fachlichen Interessen integrieren lässt. Außerdem fällt auf, dass in den vorliegenden Studien meist das Fach Mathematik untersucht wurde. So ist bislang unklar, ob Befunde auch auf andere Fächer, wie Englisch, übertragbar sind.

Interessen und Leistungen. Das individuelle Interesse wird als relativ stabile Disposition angesehen, sich bestimmten Gegenständen und Aktivitäten zu widmen (z. B. Hidi & Ainley, 2002). Das Interesse einer Person an verschiedenen – inhaltlich mehr oder weniger breit gefassten – Gegenständen wird in der Münchner Interessentheorie als zentrales Merkmal hervorgehoben und als Gegenstandsspezifität bezeichnet (vgl. Krapp, 1992). Interesse ist assoziiert mit positiven Affekten, Persistenz und Lernerfolg (Hidi & Ainley, 2002; Krapp, 2000; Renninger, 2000). Schiefele et al. (1993) fanden in einer Metaanalyse Assoziationen zwischen Leistungen und Interesse in Höhe von $r = .28$ in Mathematik bzw. von $r = .33$ in Englisch. Gemäß der Erwartungswert-Theorie (Eccles[-Parsons] et al., 1983; Wigfield & Eccles, 2000) sollte der subjektive Aufgabenwert einen Effekt auf leistungsbezogenes Wahlverhalten sowie die Investition in eine Aufgabe und somit auf nachfolgende Leistung ausüben. Das Interesse

wird dabei als ein wesentlicher Bestandteil des subjektiven Aufgabenwerts angesehen (vgl. Eccles[-Parsons] et al., 1983; Renninger, 2000). Köller, Baumert et al. (2000) berücksichtigten den institutionellen Kontext in ihren Analysen und zeigten für das Fach Mathematik, dass wesentliche Voraussetzungen für einen Effekt von Interesse auf Leistungsentwicklungen erst in der Sekundarstufe II und der damit einhergehenden Möglichkeit für Kurswahlen gegeben sind. Für Englisch liegen vergleichsweise wenige Studien vor. Im internationalen Bereich zeigte Pae (2008) einen Effekt von intrinsischer Motivation auf Fremdsprachenleistungen von koreanischen Studenten und Abu-Rabia (2003) belegte einen Effekt von Interesse auf Englischleistungen als dritte Fremdsprache bei jugendlichen arabischen Schülerinnen und Schülern.

Mit zunehmender zeitlicher Nähe zur Berufs- oder Studienwahl gewinnen neben den fachlichen auch berufliche Interessen an Bedeutung. Berufliche Interessen bilden sich im Jugendalter heraus (Tracey, 2001) und werden zunehmend stabil (Low et al., 2005). Insbesondere der Besuch der gymnasialen Oberstufe und somit der Erwerb der allgemeinen Hochschulreife wurde von Studienberechtigten ein halbes Jahr vor Schulabschluss dahingehend begründet, dass diese eine „unerlässliche Voraussetzung für ein Studium“ sowie „Voraussetzung für jede anspruchsvolle Berufsausbildung“ seien (Heine, Spangenberg & Willich, 2007).

Berufliche Interessen gehen über die von Krapp (1992) formulierte Person-Gegenstands-Konzeption des Interesses hinaus, indem sie weniger eng auf einen bestimmten Gegenstand beschränkt sind, sondern sich auf mehrere Klassen beruflicher Aktivitäten beziehen und somit eine Repräsentation der gesamten Berufslandschaft beinhalten (Nagy, 2006). Darüber hinaus sind berufliche Interessen eher noch als akademische Interessen auf praktische Tätigkeiten ausgerichtet (vgl. Bergmann & Eder, 1998). Die Theorie der Berufswahl nach Holland (1997) wurde primär zur Berufs- und Laufbahnberatung entwickelt. In dem sogenannten RIASEC-Modell nimmt Holland sechs Interessenorientierungen an: Realistic (mechanisch-technisch), Investigative (mathematisch-naturwissenschaftlich), Artistic (sprachlich-musisch), Social (sozialzwischenmenschlich), Enterprising (zwischenmenschlich-manipulativ) und Conventional (rechnerisch-verwaltend). Holland (1997) sieht mit diesen Interessenorientierungen unterschiedliche Werte, Einstellungen und Fähigkeiten verbunden. Einen theoretischen Rahmen, in dem u. a. angenommen wird, dass sich berufliche Interessen und Fähigkei-

ten in einer reziproken Beziehung zueinander entwickeln, bietet die Prozess-, Persönlichkeits-, Interessen- und Wissenstheorie der intellektuellen Entwicklung von Ackerman (1996). Eine Metaanalyse von Ackerman und Heggestad (1997) kommt zu dem Ergebnis, dass Personen mit vorherrschenden Interessen im Bereich Investigative und Realistic vergleichsweise gute Leistungen im mathematischen Bereich aufweisen, während spezifische Stärken von Personen mit hohen Investigative- und Artistic-Interessen im sprachlichen Bereich liegen. In einer neueren Längsschnittstudie zum Einfluss einzelner Skalen beruflicher Interessen auf nachfolgende Leistungen in Mathematik und Englisch (Tracey et al., 2005) wurden jedoch keine Assoziationen gefunden.

Strong argumentierte bereits 1955, dass Zusammenhänge zwischen Interessensskalen und Leistungen möglicherweise nicht die wahren Assoziationen erfassen würden und beide Bereiche komplexer seien. So untersuchte Randahl (1991) statt einzelner Interessen- und Leistungsmaße den Zusammenhang zwischen Interessen- und Leistungsprofilen. Individuelle Profile bestehen aus einer einfachen Rangordnung von Interessenorientierungen, die sich aus zwei bis sechs der RIASEC-Variablen zusammensetzen (vgl. Holland, 1997). Die Rangordnung der Variablen ist in der Regel nicht zufällig, da die sechs Orientierungen in spezifischer Weise miteinander assoziiert sind. Diese Zusammenhangsstruktur lässt sich in Form eines Hexagons darstellen. Die räumliche Nähe entspricht der inhaltlichen Ähnlichkeit: benachbarte Orientierungen (z.B. R-I) sind einander ähnlich, indirekt benachbarte Orientierungen (z.B. R-A) eher unähnlich und gegenüberliegende Orientierungen (z.B. R-S-Typen) bilden sozusagen die Pole einer Dimension (s. Abb. 5A).

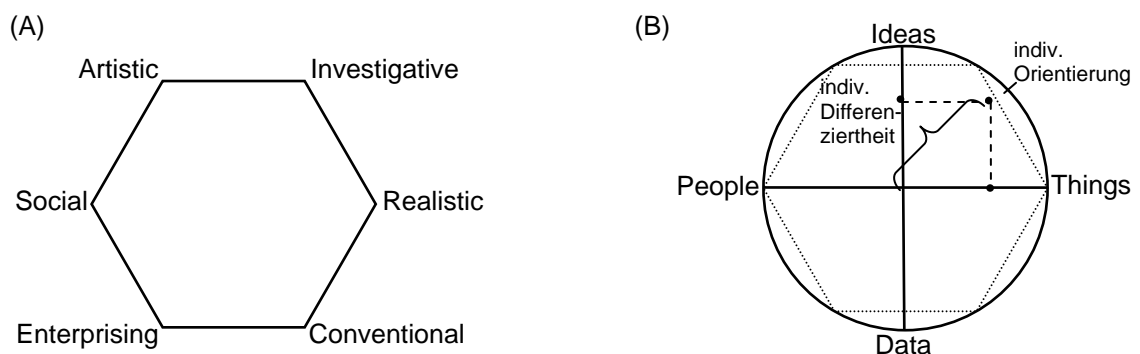


Abbildung 5. RIASEC-Modell nach Holland (1997): A) Typologisch als Hexagon und B) Dimensionale Repräsentation mit Beispiel eines individuellen Profils.

Prediger (1982) zeigte, dass dem hexagonalen Modell zwei orthogonale Achsen zugrunde liegen, die er als Things/People (horizontal) und Data/Ideas (vertikal) charakterisierte. In einer Meta-Analyse bestätigten Tracey und Rounds (1993; Rounds & Tracey, 1993) diese den Skalen zugrunde liegende zweidimensionale Struktur. Mit Hilfe der beiden Achsen lässt sich die individuelle *Orientierung* beruflicher Interessenprofile bestimmen (z.B. Eder, 1998; Nagy, 2006; Prediger, 1982). In die Berechnung der individuellen Achsenpositionen fließen jeweils alle sechs Orientierungen mit ein, sodass ein umfassendes und gleichzeitig ökonomisches Profil beruflicher Interessenorientierungen modelliert wird. Vorherrschende Interessen von Personen mit positiven Werten auf beiden Achsen liegen bspw. im Bereich Investigative, während negative Werte auf beiden Achsen eine Enterprising-Orientierung nahe legen (s. Abb. 5B). Dabei können Personen nicht nur idealtypisch einer bestimmten Interessenorientierung zugeordnet werden, sondern sie lassen sich anhand ihrer Ausprägung auf den beiden Achsen beliebig im Kreis positionieren und somit auch als Mischtypen benachbarter Interessensbereiche darstellen (z.B. R-I-Typen). Diese Positionierung beinhaltet gleichzeitig Informationen über die Differenziertheit eines Profils, was sich über den Abstand der individuellen Position vom Kreismittelpunkt veranschaulichen lässt. Die Differenziertheit ist ein Maß dafür, wie weit sich die von einer Person präferierte Interessenorientierung von den anderen Orientierungen abhebt bzw. wie eindeutig die Zuordnung zu einer bestimmten Interessenorientierung ist. Die dimensionale Bestimmung der Interessenorientierung enthält also Informationen über die Orientierung, gewichtet mit der Differenziertheit eines Interessenprofils. Arbeiten, die sich auf Assoziationen zwischen dimensionalen Interessenorientierungen und Leistungen stützen, sind vergleichsweise rar. In einer aktuellen Querschnittstudie belegte Nagy (2006) Zusammenhänge zwischen einer dimensionalen Ausrichtung beruflicher Interessen entsprechend einer Realistic-/Investigative-Orientierung und Testleistungen in Mathematik sowie einer Investigative-/Artistic-Orientierung mit Englischleistungen.

Die Erfassung beruflicher Interessen erfolgt in der Regel über Inventare, in denen Personen die Höhe ihres Interesses an bestimmten beruflichen Tätigkeiten angeben sollen. Faktorenanalysen zu diesen Interesseninventaren ergeben einen ersten Faktor mit vergleichbar hohen Ladungen aller sechs Skalen (z.B. Prediger, 1982). Entsprechend spiegelt dieser Faktor interindividuelle Unterschiede in der grundsätzlichen Zustim-

mung zu den sechs beruflichen Interessenorientierungen wider und wird bspw. als Level, Antwortstil (vgl. Prediger, 1998) oder Generalfaktor (vgl. Darcy & Tracey, 2003) bezeichnet. Die Interpretation des Levels variiert; während Prediger (1998) davor warnt, den Level als Stärke von Interessen anzusehen, vermuten Darcy und Tracey (2003), dass er ein Indikator für die Flexibilität von Interessen sei und somit in moderat positiver Beziehung zu Outcome-Maßen, wie Leistung, Zufriedenheit und Stabilität, stehen sollte. Zusätzlich nehmen die Autoren an, dass der Level die Beziehung zwischen Person-Umwelt-Kongruenz und Outcome moderiere. In der Empirie ließ sich ein bedeutsamer direkter Effekt des Levels auf Leistungsmaße bislang nicht bestätigen (vgl. Tracey & Robbins, 2006; Warwas, Nagy, Watermann & Hasselhorn, in press). Hingegen konnte ein moderierender Effekt des Levels auf die Kongruenz-Persistenz-, nicht aber auf die Kongruenz-Noten-Beziehung aufgezeigt werden (vgl. Tracey & Robbins, 2006). Darcy und Tracey (2003) wiesen darüber hinaus auf die zentrale Bedeutung des Levels in der Erfassung beruflicher Interessen hin. Sofern er als Persönlichkeitsstil (z. B. Antwortstil) substantielle Bedeutung habe, müsse er inhaltlich verstanden und in aktuelle Modelle und Theorien integriert werden. Sofern er die Erfassung beruflicher Interessen verfälsche, müsse er hingegen kontrolliert werden (vgl. ebd., p. 224).

Zusammen mit dem Level bilden die mit der Differenziertheit gewichteten Interessenorientierungen die zentralen Komponenten eines individuellen Profils entsprechend eines sogenannten Structural Summary nach Gurtman und Balakrishnan (1998). Die Validität der Reduktion des hexagonalen Skalenmodells auf Interessenprofile wurde im Quer- und Längsschnitt belegt (Warwas, Nagy et al., in press). Die Autoren zeigten exemplarisch in einem ersten empirischen Befund einen positiven Effekt beruflicher Interessenprofile, speziell der Things/People-Orientierung, auf Mathematikleistungen in der gymnasialen Oberstufe. Dieser blieb auch nach Kontrolle von Geschlecht und Ausgangsleistung erhalten. Domänenübergreifende Analysen sowie empirische Prüfungen zur Bedeutung beruflicher Interessen im Zusammenhang mit weiteren für Leistungsentwicklungen als bedeutsam angenommenen Variablen stehen bislang aus.

Selbstkonzept und Leistungen. Für die Entwicklung schulischer Leistungen spielt auch das Selbstkonzept eine entscheidende Rolle. Das Selbstkonzept lässt sich als Vorstellung einer Person über sich selbst und die eigenen Fähigkeiten beschreiben (vgl. Shavelson et al., 1976). Ein höheres Selbstkonzept führt zu höheren Erfolgserwartun-

gen, zu ausdauernderer Aufgabenbearbeitung und in Folge dessen zu besseren Leistungen (vgl. Eccles[-Parsons] et al., 1984; Fredricks & Eccles, 2002). Fachspezifisch berichteten Möller und Köller (2004) in einem Überblicksartikel eine Median-Korrelation von $Md = .47$ zwischen Leistungsindikatoren (Noten und Leistungstests) und mathematischem Selbstkonzept bzw. von $Md = .39$ für diesen Zusammenhang im sprachlichen Bereich. Längsschnittliche Zusammenhanganalysen bestätigten systematische Effekte des Selbstkonzepts auf nachfolgende Leistungen unter Kontrolle der Ausgangsleistung (Valentine et al., 2004), nicht jedoch von Schulleistungen (Noten) auf das zeitlich nachgeordnete Selbstkonzept (Marsh, 1990b). Köller et al. (2006) geben in ihrer Längsschnittstudie in der gymnasialen Oberstufe am Beispiel des Fachs Mathematik erste empirische Hinweise zum Beziehungsgefüge zwischen Selbstkonzept, Interesse und Leistungen. So belegten die Autoren einen Effekt des Selbstkonzepts auf das zeitlich nachgeordnete Mathematik-Interesse. Zusätzlich zeigten sie, dass bei gleichzeitiger Modellierung sowohl das Selbstkonzept als auch das Interesse jeweils Effekte auf die Mathematikleistung unter Kontrolle der Ausgangsleistung ausübten. Längsschnittliche Untersuchungen, die sich auf das Fach Englisch beziehen, fehlen bislang.

Erwartungs-mal-Wert-Theorien gehen im Übrigen davon aus, dass ein Teil des Effekts des Selbstkonzepts auf nachfolgende Leistungen über die Interessen mediiert ist. In der *Sozial-kognitiven Berufslaufbahntheorie* (Lent et al., 1994) wird etwa ein solches richtungsbezogenes Beziehungsgefüge angenommen, dabei aber explizit Bezug auf berufliche Interessen genommen. Die Autoren fanden in einer Meta-Analyse Korrelationen in Höhe von $r = .53$ bzw. $r = .52$ zwischen verschiedenen Maßen selbstbezogener Kognitionen und beruflichen Interessen. Längsschnittliche Untersuchungen zum Zusammenspiel von Selbstkonzept und beruflichen Interessen in Hinblick auf Leistungsentwicklungen in Mathematik und Englisch liegen bislang nicht vor.

Kursniveau und Leistungen. In der gymnasialen Oberstufe wählen die Schülerinnen und Schüler in der Regel ab der 12. Klasse (Kursstufe) in den curricularen Kernfächern zwischen einem Kurs mit grundlegenden oder erhöhten Anforderungen (auch: Grund- oder Leistungskurs). Mit einem erhöhten Kursniveau geht in der Regel ein kognitiv anspruchsvollerer Unterricht in dem Fach einher, der zu einem tieferen Verständnis der Inhalte beitragen und somit positive Effekte auf den Wissenserwerb haben sollte. Empirisch zeigten sich am Ende der gymnasialen Oberstufe deutlich

bessere Leistungen von Schülerinnen und Schülern in Kursen mit erhöhten Anforderungen als solche in Kursen mit grundlegenden Anforderungen (z.B. Baumert & Watermann, 2000; Köller & Trautwein, 2004; Watermann et al., 2004), die auf unterschiedliche Eingangsvoraussetzungen zu Beginn der gymnasialen Oberstufe (Warwas, Watermann & Hasselhorn, im Druck) und auf weitere Leistungsdifferenzierungen während der Kursstufe zurückzuführen sind (z.B. Köller et al., 2006).

Dem individuellen Interesse kommt dabei eine wichtige Bedeutung zu. Zum einen ist Interesse ein entscheidendes Kriterium beim Kurswahlverhalten der Schülerinnen und Schüler (vgl. z.B. Hodapp & Mißler, 1996; Köller, Daniels et al., 2000; Köller et al., 2006). Gleichwohl sollte aus Sicht der Interessentheorie (z.B. Krapp, 1992) eine vertiefte Auseinandersetzung mit einem Gegenstand, so wie es in Kursen mit erhöhtem Anforderungsniveau zu erwarten ist, mit Kompetenzerleben einhergehen und förderlich für das Interesse an diesem Fach sein. Köller et al. (2006) bestätigten, dass sich das Interesse in Kursen mit erhöhten Anforderungen günstiger veränderte als in solchen mit grundlegenden Anforderungen. Dieses Interesse sollte wiederum einen Effekt auf schulische Leistungen haben. Auch hierzu mangelt es allerdings an längsschnittlichen Studien, die den Zusammenhang zwischen Kursniveau, (beruflichen) Interessen und Leistungsveränderungen in Mathematik und Englisch untersuchen.

4.3.2 Fragestellungen

Die vorliegende Studie wurde geplant, um die Entwicklungsdynamik der Effekte von beruflichen Interessen auf nachfolgende Leistungen in zwei Kernfächern der gymnasialen Oberstufe (Mathematik und Englisch) bei Kontrolle der Ausgangsleistung zu prüfen. Hollands Modell beruflicher Interessen liegt eine zweidimensionale Struktur zugrunde. Diese berücksichtigend, wurden berufliche Interessenprofile entsprechend eines Structural Summary nach Gurtman und Balakrishnan (1998) modelliert. Nagy (2006) belegte im Querschnitt Assoziationen zwischen dimensional modellierten beruflichen Interessenorientierungen im Bereich Realistic/Investigative mit Mathematikleistungen sowie Investigative-/Artistic-Interessen mit Englischleistungen. In einer ersten Längsschnittstudie zeigten Warwas, Watermann et al. (im Druck) einen positiven Effekt der Things/People-Dimension auf Mathematikleistungen, der auch nach Kontrolle von

Geschlecht und Ausgangsleistung erhalten blieb. Der Level als weiterer Profilparameter erwies sich bislang nicht als bedeutsam für Leistungen (vgl. ebd.; Tracey & Robbins, 2006). Dennoch ist es nach Darcy und Tracey (2003) von zentraler Bedeutung den vom Antwortstil bereinigten Effekt der Interessenorientierungen zu erfassen. Berufliche Interessenprofile sollen nun, in Anlehnung an erste längsschnittliche Studien (z. B. Köller et al., 2006), mit Selbstkonzept, Kursniveau und Leistungen in Verbindung gebracht werden. Verglichen mit Mathematik sind Fremdsprachenleistungen trotz ihrer anerkannten Bedeutung eher selten Gegenstand empirischer Untersuchungen (Möller & Zaunbauer, 2008). Deshalb liegt ein wesentlicher zusätzlicher Beitrag dieser Studie in der Prüfung, ob Befunde für Mathematik vergleichbar mit denen in Englisch sind.

Insgesamt werden die nachstehenden Hypothesen untersucht:

1. *Effekt beruflicher Interessenprofile auf Leistungsentwicklungen:* Für das Fach Mathematik wird ein positiver Effekt der Things/People-Dimension beruflicher Interessenprofile (insgesamt einer Realistic-Orientierung entsprechend) auf nachfolgende Leistungen erwartet. In Englisch wird ein positiver Effekt der Data/Ideas-Dimension (entsprechend einer Investigative-/Artistic-Orientierung) auf nachfolgende Leistungen erwartet. Der Level sollte ohne Bedeutung für Leistungsentwicklungen in beiden Fächern sein.
2. *Effekte von fachspezifischem Selbstkonzept und beruflichen Interessenprofilen auf Leistungsentwicklungen:* Es wird angenommen, dass das fachspezifische Selbstkonzept in Mathematik [Englisch] einen positiven Effekt auf nachfolgende Leistungen in Mathematik [Englisch] ausübt. Zusätzlich wird erwartet, dass berufliche Interessenprofile einen Teil des Selbstkonzept-Effekts auf Leistungsveränderungen mediieren.
3. *Effekte von Kursniveau und beruflichen Interessenprofilen auf Leistungsentwicklungen:* Das Kursniveau in Mathematik [Englisch] sollte einen positiven Effekt auf nachfolgende Leistungen in Mathematik [Englisch] ausüben. Dabei wird angenommen, dass berufliche Interessenprofile den Effekt des Kursniveaus auf Leistungsentwicklungen mediieren.

4.3.3 Methode

Stichprobe und Durchführung. Die Untersuchungen fanden am Ende des ersten (T1) sowie zweiten Schulhalbjahres (T2) der 11. Klasse an einem Fachgymnasium mit der Ausrichtung Wirtschaft statt. Die Schülerinnen und Schüler bearbeiteten jeweils einen Mathematik- und einen Englischtest, zusätzlich füllten sie einen Fragebogen aus, der u. a. Fragen zu ihren beruflichen Interessen enthielt. Zu T1 nahmen $N = 168$ Schülerinnen und Schüler teil (46% weiblich, Durchschnittsalter 18.1 Jahre), zu T2 waren es $N = 128$. Der Ausfall ist begründet durch Krankheit, Schulabbruch oder -wechsel, teilweise auch durch Ablehnung einer Teilnahme. Mit fehlenden Werten gehen jedoch bestimmte Probleme einher (vgl. Lüdtke et al., 2007). Hier wurden deshalb fehlende Werte mit Hilfe von Multiple Imputation ersetzt (Rubin, 1987). Mit NORM 2.03 (Schafer, 1999) wurden $m = 20$ vollständige Datensätze erzeugt. In das Hintergrundmodell für die Mehrfachschätzungen der fehlenden Werte wurden zusätzliche Informationen (Geschlecht, Alter, Selbstkonzept zu T2 und berufliche Interessen zu T2) berücksichtigt. Die statistischen Kennwerte der einzelnen Datensätze wurden nach den Regeln von Rubin (1987) in Mplus (Muthén & Muthén, 1998–2008) aggregiert. Laut Vereinbarung zur Gestaltung der gymnasialen Oberstufe (KMK, 2006) müssen sich die Schülerinnen und Schüler in den Kernfächern Mathematik, Englisch und Deutsch für ein Kursniveau (grundlegende vs. erhöhte Anforderungen) entscheiden. Diese Wahl wird üblicherweise in Hinblick auf die anschließende Qualifikationsphase getroffen. In dieser Stichprobe legten sich die Schülerinnen und Schüler bereits zu Beginn der gymnasialen Oberstufe auf bestimmte Anforderungsniveaus in den Fächern fest. Etwa die Hälfte der Schülerinnen und Schüler ($n = 85$) befanden sich in Klassen, die auf ein erhöhtes Kursniveau in Mathematik hin unterrichtet wurden; in Englisch waren dies etwa drei Viertel ($n = 128$). Dieses Vorgehen ist eine Besonderheit der untersuchten Schule. Die hierarchische Struktur der Daten (Schülerinnen und Schüler in Klassen) wurde berücksichtigt, indem in Mplus die Analysen für komplexe Daten mit robustem Maximum-Likelihood-Schätzer (Muthén & Satorra, 1995) durchgeführt wurden. Hierbei wird eine größere Ähnlichkeit von Schülerinnen und Schülern innerhalb von Klassen im Vergleich zu natürlichen Gruppen bei der Berechnung von Standardfehlern mit einbezogen; Standardfehler werden in geringerem Ausmaß unterschätzt und Signifikanzprüfungen weniger

liberal angesetzt. Bei diesem Vorgehen berechnet Mplus nur unstandardisierte Parameter. Damit die Ergebnisse anhand standardisierter Parameter interpretierbar sind, wurden alle Variablen z-standardisiert.

Messinstrumente. In Mathematik wurde der TIMSS-Test zur mathematischen Grundbildung verwendet (z. B. Mullis et al., 1998). Die Testaufgaben beziehen sich auf zentrale Inhalte des mathematischen Unterrichts der Mittelstufe (Zahlenverständnis, Proportionalität, Algebra sowie Messen und Schätzen) und sind entsprechend der angelsächsischen Literacy-Debatte (NCTM, 1989) allesamt in Alltagskontexte eingebunden. Wenngleich der Test nicht speziell für die 11. Klasse konstruiert ist, zeigt sich in seiner Zielsetzung eine hohe Übereinstimmung mit den im Curriculum für die 11. Klasse spezifizierten Inhalten (Rückgriff auf mathematisches Grundwissen, Anwendungsbezug auf die Praxis, vgl. Niedersächsisches Kultusministerium, 1990). Mullis et al. (1998) belegten zudem, dass sich die Testergebnisse nicht wesentlich veränderten, wenn Analysen ausschließlich auf lehrplankonformen Items basierten. Beide Tests enthielten vier gemeinsame Aufgaben (Ankeritems, s. Anhang A). Diese deckten einen möglichst breiten Schwierigkeitsbereich ab und wurden in TIMSS jeweils als besonders repräsentativ für den Bereich der mathematischen Grundbildung angesehen. Mit Hilfe der Ankeritems war es im Rahmen sogenannter Test Equating-Prozeduren (vgl. Hambleton & Swaminathan, 1985; Lord, 1980) möglich, die zu zwei Zeitpunkten erhobenen Leistungen auf einer gemeinsamen Metrik abzutragen. In Mathematik und in Englisch (s. u.) wurden die Ergebnisse zu T1 (21 Items) z-standardisiert und die Werte zu T2 (8 Items) anschließend am Mittelwert und der Standardabweichung der Werte zu T1 standardisiert. Die mittleren Leistungen in mathematischer Grundbildung verringerten sich deskriptiv in der 11. Jahrgangsstufe; die Differenz ($M = -0.20$, $SD = 1.46$) war nicht signifikant ($z = 1.75$, $p = .082$). Die differentielle Stabilität lag bei $r = .58$ ($p < .01$).

Die Englischleistungen wurden zu T1 mit dem TOEFL (entwickelt vom Educational Testing Service) erhoben. Diese von Köller und Trautwein (2004) validierte Kurzversion eines ITP-TOEFL (Institutional Testing Program) enthielt 80 Items. Curriculare Validität wird zwar nicht beansprucht, es zeigten sich aber eine deutliche Korrelation zwischen TOEFL-Wert und Englischnote in Höhe von $r = .64$ (ebd.). Aus der Kurzform wurden 13 Items als Ankeritems verwendet, 8 neue Aufgaben zu T2 stammen aus einem ITP-TOEFL aus dem Jahr 1995. Die Items wurden im Rahmen dieser Studie rasch-

skaliert. Die mittleren Leistungen in Englisch verbesserten sich von Mitte bis Ende der 11. Jahrgangsstufe ($M = 0.73$, $SD = 1.14$) deutlich ($z = 8.26$, $p < .001$) bei einer differentiellen Stabilität von $r = .52$ ($p < .01$).

Die beruflichen Interessen wurden mit der revidierten Version des Allgemeinen-Interessen-Struktur-Tests (AIST-R; Bergmann & Eder, 2005) erhoben. Der AIST-R ist entsprechend den sechs Interessendimensionen nach Holland (1997) konstruiert, die strukturelle Validität des Instruments haben Nagy et al. (in press) belegt. Im Test werden verschiedene berufliche Tätigkeiten beschrieben und anhand einer fünfstufigen Likert-Skala (1 = gar nicht, ..., 5 = sehr) soll die befragte Person angeben, wie sehr sie sich dafür interessiert bzw. interessieren würde. Die Reliabilität (Cronbachs alpha) der sechs Skalen betrug jeweils $\alpha = .79$ und mehr. Die stichprobenspezifischen Mittelwerte der RIASEC-Skalen wurden an einer Normstichprobe (TOSCA; Köller, Watermann et al., 2004) von $N = 2.716$ allgemein bildenden Gymnasiasten standardisiert. Zur sparsamen Modellierung beruflicher Interessenprofile wurde eine Reduktion des RIASEC-Modells (Warwas, Nagy et al., in press) mit den Komponenten Level, Things/People- und Data/Ideas-Dimension verwendet. Der Interessenlevel entspricht dem Durchschnitt der sechs Skalenmittelwerte. Die individuellen Positionen auf den Achsen Things/People und Data/Ideas wurden mit Hilfe trigonometrischer Regeln berechnet. Ausgangspunkt ist ein Hexagon, über das sich die sechs Interessenskalen gleichmäßig (im 60°-Abstand) entlang eines Kreises verteilen. Jeder standardisierte Skalenwert wird mit dem Kosinus bzw. dem Sinus aus der entsprechenden Winkelposition der Skala multipliziert und zu einem Things/People- bzw. Data/Ideas-Wert aufsummiert. Diese Berechnung setzt also eine bestimmte Reihenfolge der Variablen entsprechend ihrer Anordnung im hexagonalen Modell voraus. Um die Passung der Daten mit der RIASEC-Struktur gegenüber einem Nullzusammenhang zu prüfen, wurde das Programm RANDALL (Tracey, 1997) verwendet. Die Nullhypothese wurde verworfen, die Daten waren mit dem hexagonalen Modell vereinbar ($p = .0167$).

Das Selbstkonzept in Mathematik [Englisch] wurde mit vier Items erfasst, für die ein vierstufiges Antwortformat vorgegeben wurde (1 = trifft überhaupt nicht zu, ..., 4 = trifft völlig zu). Zwei Items stammen aus dem von Schwanzer (2002) übersetzten *Self Description Questionnaire (SDQ) III* von Marsh (1992) und lauten „Ich bin gut in Mathematik [Englisch]“, „Ich habe Verständnisschwierigkeiten bei allem, für das man

Mathematik [Englisch] braucht“. Zwei weitere Items „Für Mathematik [Englisch] habe ich einfach keine Begabung“, „Mathematik [Englisch] liegt mir nicht besonders“ sollen das Mathematik-[Englisch-]Selbstkonzept der Begabung erfassen und stammen aus einer Studie von Köller, Schnabel und Baumert (2000). Die Reliabilität ist in Mathematik ($\alpha = .79$) und in Englisch ($\alpha = .78$) vergleichbar und als gut zu bewerten.

4.3.4 Ergebnisse

In Tabelle 10 sind die Korrelationen aller berücksichtigten Variablen angeführt.

Tabelle 10

Korrelationen zwischen Mathematik- (Ma) und Englisch-Leistungen (En) sowie beruflichen Interessen, Selbstkonzept (SK) und Kursniveau (KN)

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Ma T1										
2. Ma T2	.58**									
3. En T1	.19*	.29**								
4. En T2	.09	.09	.52**							
5. Level T1	.01	-.06	-.13	-.08						
6. T/P T1	.21**	.30**	.17*	.10	-.09					
7. D/I T1	-.05	-.07	.14†	.08	-.14†	.08				
8. MaSK T1	.19*	.24**	-.10	-.13	-.12	.36**	-.20*			
9. En SK T1	-.03	-.02	.32**	.31**	.11	-.09	.11	-.34**		
10. Ma KN	.29**	.35**	-.12	-.14†	-.06	.36**	-.19*	.47**	-.20*	
11. En KN	-.06	.04	.32**	.29**	-.06	-.21**	.17*	-.24**	.25**	-.55**

Anmerkungen. T/P = Things/People, D/I = Data/Ideas, Kursniveau mit erhöhten Anforderungen höher kodiert.

† $p < .10$, * $p < .05$, ** $p < .01$.

Es zeigten sich die erwarteten Korrelationen zwischen Mathematikleistungen und beruflichen Interessenorientierungen. Der positive Zusammenhang mit der Things/People-Dimension lässt sich, beide Dimensionen berücksichtigend, als Realistic-Orientierung interpretieren. Die Englischleistung ist ausschließlich zu T1 bedeutsam mit beruflichen Interessenprofilen assoziiert (positiv mit der Things/People-Dimension sowie schwach positiv mit der Data/Ideas-Dimension, entsprechend einer Realistic-/Investigative-Orientierung). Zwischen Selbstkonzept und Leistungen sowie Kursniveau und Leistungen resultieren die erwartbaren domäneninternen positiven Assoziationen.

Zur Prüfung der Fragestellungen dieser Studie wurden schrittweise Regressionsanalysen durchgeführt. Im Rahmen der Veränderungsperspektive dieser Studie wurde in einem ersten Modell der Effekt der Ausgangsleistung auf nachfolgende Leistung in zwei Kernfächern der gymnasialen Oberstufe geprüft (s. Modelle 1). In Übereinstimmung mit Renkl (1996) zeigten sich substantielle Effekte des Vorwissens sowohl für Mathematik als auch für Englisch. Nachfolgende Modelle beziehen zusätzliche Prädiktoren in die Regressionsanalysen mit ein; deren Effekte wurden als Veränderungseffekte interpretiert. Berufliche Interessenprofile wurden mit den Prädiktoren Level sowie Things/People- und Data/Ideas-Dimension (vgl. Prediger, 1982) in die Regressionsanalysen aufgenommen. Für das Fach Mathematik bestätigten Warwas, Nagy et al. (in press) in einem ersten empirischen Ansatz einen positiven Effekt der Things/People-Orientierung auf nachfolgende Leistungen unter Kontrolle von Ausgangsleistung und Geschlecht. Daran anknüpfend wurden Effekte beruflicher Interessenprofile auf Leistungsentwicklungen in Mathematik und darüber hinaus auch in Englisch geprüft (s. Modelle 2). Wie erwartet, war der Interessenlevel nicht mit Leistungsveränderungen in Mathematik oder Englisch assoziiert. Weitere Ergebnisse vorwegnehmend, zeigten sich in keiner der durchgeführten Regressionsanalysen Effekte des Levels auf Leistungsveränderungen in den beiden Domänen. Die im Folgenden berichteten Befunde zur Bedeutung beruflicher Interessenprofile für die Entwicklung schulischer Leistungen beziehen sich deshalb ausschließlich auf die beiden Dimensionen beruflicher Interessenorientierungen. In Mathematik resultierte erwartungsgemäß ein positiver Effekt der Things/People-Dimension (entsprechend einer Realistic-Orientierung) auf Leistungsveränderungen. Wider Erwarten ließ sich kein positiver Effekt der Data/Ideas-Dimension auf die Entwicklung der Englischleistungen nachweisen.

Tabelle 11

Befunde zur Vorhersage von Mathematikleistungen zu T2; standardisierte Regressionskoeffizienten und in Klammern Standardfehler

Prädiktoren zu T1	Mathematikleistung zu T2					
	Modell 1M	Modell 2M	Modell 3M	Modell 4M	Modell 5M	Modell 6M
Mathematikleistung	.58 (0.07)**	.54 (0.07)**	.55 (0.08)**	.53 (0.07)**	.52 (0.07)**	.51 (0.07)**
Interessenlevel		-.06 (0.07)		-.05 (0.06)		-.05 (0.07)
Things / People		.19 (0.08)*		.17 (0.07)*		.14 (0.07)*
Data / Ideas		-.06 (0.07)		-.05 (0.07)		-.03 (0.06)
Selbstkonzept Mathe			.13 (0.06)*	.06 (0.06)		
Kursniveau Mathe					.20 (0.09)*	.15 (0.08)†
R^2	.335**	.378**	.354**	.382**	.375**	.396**

Anmerkungen. † $p < .10$, * $p < .05$, ** $p < .01$.

Tabelle 12
Befunde zur Vorhersage von Englischleistungen zu T2; standardisierte Regressionskoeffizienten und in Klammern Standardfehler

Prädiktoren zu T1	Englischleistung zu T2					
	Modell 1E	Modell 2E	Modell 3E	Modell 4E	Modell 5E	Modell 6E
Englischleistung	.54 (0.10)**	.53 (0.10)**	.48 (0.12)**	.47 (0.11)**	.49 (0.10)**	.48 (0.09)**
Interessenlevel		-.02 (0.08)		-.04 (0.08)		-.01 (0.08)
Things / People		.01 (0.08)		.03 (0.08)		.06 (0.08)
Data / Ideas		.01 (0.08)		-.01 (0.07)		-.02 (0.07)
Selbstkonzept Englisch			.15 (0.08)†	.17 (0.08)*		
Kursniveau Englisch					.15 (0.06)**	.17 (0.06)**
<i>R</i> ²	.283**	.285**	.303**	.309**	.303**	.309**

Anmerkungen. † $p < .10$, * $p < .05$, ** $p < .01$.

Als nächstes wurde das fachspezifische Selbstkonzept als weitere zur Erklärung von Leistungsentwicklungen relevante angenommene Variable in den Analysen berücksichtigt. Zunächst wurden direkte Effekte des Selbstkonzepts auf Leistungsveränderungen in beiden Fächern geprüft (s. Modelle 3). Sowohl in Mathematik als auch in Englisch zeigten sich die erwartbaren Effekte des fachspezifischen Selbstkonzepts auf die Entwicklung der Leistungen. Um das Zusammenspiel von Selbstkonzept und beruflichen Interessen zu untersuchen, wurden die Parameter beruflicher Interessenprofile als weitere Prädiktoren in die Regressionsanalysen mit aufgenommen (s. Modelle 4). Als Voraussetzung für eine Mediation beruflicher Interessenprofile auf den Selbstkonzept-Leistungs-Zusammenhang wurden zusätzlich Effekte des Selbstkonzepts auf die Dimensionen beruflicher Interessenorientierung geprüft. In Mathematik zeigte sich, dass das Selbstkonzept mit *beiden* Dimensionen, entsprechend einer Conventional-Orientierung, assoziiert war (Things/People: $\beta = .36$, $SE = 0.08$, $p < .01$; Data/Ideas: $\beta = -.20$, $SE = 0.06$, $p < .01$). Bei Hinzunahme beruflicher Interessenprofile zeigte sich, dass diese einen über das Selbstkonzept hinaus gehenden Effekt auf Leistungsveränderungen ausübten und den Effekt des Selbstkonzepts vollständig mediieren. In Englisch ergab sich ein zum Modell 2E übereinstimmendes Bild: Das Selbstkonzept übte einen schwachen Effekt auf die Data/ Ideas-Dimension, vergleichbar mit einer Investigative-/ Artistic-Orientierung, aus ($\beta = .11$, $SE = 0.06$, $p < .10$; Things/People: $\beta = -.09$, $SE = 0.09$, $p > .10$). Der Effekt des Selbstkonzepts auf Leistungsentwicklungen war nicht über berufliche Interessen mediiert.

In einem letzten Analysebereich wurde das Kursniveau als bedeutsam für die Leistungsveränderungen in Mathematik und Englisch angenommene Variable mit einbezogen (s. Modelle 5) und daran anschließend das Zusammenspiel mit beruflichen Interessenprofilen untersucht (s. Modelle 6). Es zeigten sich die erwartbaren bedeutsamen Effekte des Kursniveaus auf Leistungsentwicklungen in beiden Fächern. Zur Analyse des Zusammenspiels wurden auch hier vorab Effekte der Kursniveaus auf die Dimensionen beruflicher Interessenorientierungen geprüft. In Mathematik wurde, wie auch beim Selbstkonzept, ein positiver Effekt des Kursniveaus auf die Things/People-Dimension ($\beta = .36$, $SE = .08$, $p < .01$) sowie ein negativer Effekt auf die Data/Ideas-Dimension nachgewiesen ($\beta = -.19$, $SE = 0.08$, $p < .05$). Dazu entgegengesetzt übte das Kursniveau in Englisch einen schwachen negativen Effekt auf die Things/People-

Orientierung ($\beta = -.21$, $SE = 0.12$, $p < .10$) und einen positiven Effekt auf die Data/Ideas-Orientierung aus ($\beta = .17$, $SE = 0.06$, $p < .01$), was zusammengenommen einer Artistic-Orientierung entsprach. Im Modell 6M zeigte sich im Fach Mathematik ein Effekt der Things/People-Dimension, der auch nach Kontrolle des Kursniveaus signifikant blieb und letzteres medierte. In Englisch resultierte in Übereinstimmung mit bisherigen Befunden dieser Studie kein (über das Kursniveau hinausgehender) Effekt beruflicher Interessenprofile auf Leistungsentwicklungen.

4.3.5 Diskussion

In dieser Längsschnittstudie wurde der Effekt beruflicher Interessenprofile auf Leistungsveränderungen in zwei Kernfächern der gymnasialen Oberstufe geprüft sowie deren Zusammenspiel mit dem fachspezifischen Selbstkonzept bzw. dem Kursniveau untersucht. Zur umfassenden und sparsamen Modellierung beruflicher Interessenprofile wurde eine valide Reduktion des RIASEC-Modells (Holland, 1997) entsprechend eines Structural Summary nach Gurtman und Balakrishnan (1998) verwendet. Die Befunde dieser Studie reihen sich ein in bisherige Arbeiten (z. B. Köller et al., 2006) zum Zusammenspiel von Interesse, Selbstkonzept, Kursniveau und Leistungen im Fach Mathematik und ergänzen diese um die Rolle *beruflicher* Interessen. Darüber hinaus lassen sich mit dieser Studie deutliche fachspezifische Unterschiede aufzeigen. In Englisch, als sprachliche Domäne, resultierten erwartbare Effekte von Selbstkonzept und Kursniveau auf Leistungsentwicklungen, jedoch ließen sich berufliche Interessen nicht in das Beziehungsgefüge integrieren.

Die beruflichen Interessenprofile setzten sich zusammen aus dem Level als generelle Höhe beruflicher Interessen sowie den dimensional Interessensorientierungen Things/People und Data/Ideas (vgl. Prediger, 1982). Mit der Berücksichtigung des Levels sollte zum einen geprüft werden, ob dieser für die Entwicklung schulischer Leistungen in Mathematik und Englisch bedeutsam ist. Zum anderen sollte der Effekt von Interessensorientierungen um den des Levels bzw. Antwortfaktors kontrolliert werden. In dieser Studie zeigten sich keine direkten Effekte des Levels auf Leistungsveränderungen in Mathematik und Englisch. Schülerinnen und Schüler, die generell ein sehr hohes oder sehr niedriges Interesse an beruflichen Tätigkeiten angeben, unterscheiden

sich in ihren schulischen Leistungsentwicklungen in Mathematik und Englisch nicht voneinander. Auch Tracey und Robbins (2006) fanden keine direkten Effekte des Levels auf Outcome-Maße wie Durchschnittsnoten oder Persistenz, jedoch belegten sie gleichzeitig, dass der Level die Beziehung zwischen Person-Umwelt-Kongruenz und Persistenz (nicht aber zwischen Kongruenz und Noten) moderierte und somit nicht generell als bedeutungslos für Outcome-Maße angesehen werden darf. Inwieweit der Level einen moderierenden Effekt auf den Zusammenhang von Person-Umwelt-Kongruenz und schulischen Leistungsentwicklungen ausübt, sollte in weiteren Studien geprüft werden.

Im Fokus beruflicher Interessenprofile standen die beiden Interessendimensionen Things/People und Data/Ideas nach Prediger (1982). In Anlehnung an die Querschnittstudie von Nagy (2006) sowie einer ersten längsschnittlichen Studie von Warwas, Nagy et al. (in press) wurden positive Assoziationen der Things/People-Dimension mit der Entwicklung von Mathematikleistungen sowie der Data/Ideas-Dimension mit der Entwicklung von Englischleistungen erwartet. In den Ergebnissen dieser Studie zeigten sich dabei deutliche fachspezifische Unterschiede: Im Fach Mathematik entsprachen die Ergebnisse unseren Erwartungen. Es zeigte sich, dass die Things/People-Dimension einen positiven Effekt auf Leistungsveränderungen ausübte, der zusammengenommen einer Realistic-Orientierung entsprach. Darüber hinaus resultierten positive direkte Effekte des fachspezifischen Selbstkonzepts und des Kursniveaus auf Leistungsentwicklungen, die jeweils durch berufliche Interessenorientierungen mediiert wurden. Es ließen sich nach Kontrolle von Selbstkonzept oder Kursniveau weiterhin Effekte beruflicher Interessenprofile auf Leistungsveränderungen in Mathematik aufzeigen. Bemerkenswert ist, dass die Assoziationen zwischen der dimensional orientierten Interessen mit Leistungen, Selbstkonzept und Kursniveau innerhalb dieser Domäne weitgehend vergleichbar waren. So zeigten sich durchwegs positive Assoziationen zwischen der Things/People-Orientierung und den Mathematikleistungen, dem Selbstkonzept sowie dem Kursniveau. Interessanterweise übten Selbstkonzept und Kursniveau darüber hinaus einen negativen Effekt auf die Data/Ideas-Dimension aus, sodass dieser Effekt bei Berücksichtigung der Things/People-Dimension der benachbarten Conventional-Orientierung entsprach. Hieran schließt sich die Frage nach der Bedeutung der beruflichen Ausrichtung des Fachgymnasiums (hier: Wirtschaft) für den Zusammen-

hang von Selbstkonzept bzw. Kursniveau mit beruflichen Interessen an. Wirtschaftswissenschaften lassen sich lose mit Interessen in den Bereichen Conventional/Enterprising assoziieren. Da Mathematik ein den Wirtschaftswissenschaften verbundenes Fach ist, könnte sich hierin eine tendenzielle Verschiebung in Richtung Conventional im Sinne einer Person-Umwelt-Kongruenz niederschlagen. Eine Prüfung dieser Annahme ist jedoch nur möglich, wenn diese Assoziationen für die verschiedenen beruflichen Ausrichtungen der Fachgymnasien vergleichend untersucht werden.

Wie in Mathematik, lassen sich auch in Englisch direkte Effekte von Selbstkonzept und Kursniveau auf Leistungsveränderungen aufzeigen. Die Orientierung beruflicher Interessenprofile (hier: Realistic/Investigative) ist zwar im Querschnitt mit Englischleistungen assoziiert und Ergebnisse dieser Studie somit weitgehend konform mit bisherigen Studien (z.B. Ackerman & Heggstad, 1997; Nagy, 2006); sie üben jedoch keinen Effekt auf *Leistungsentwicklungen* in Englisch aus, sodass sie sich auch nicht in ein Zusammenhangsgefüge mit dem fachspezifischen Selbstkonzept oder Kursniveau zur Vorhersage von Leistungsentwicklungen integrieren lassen. Für eine mangelnde Assoziation zwischen beruflichen Interessenorientierungen und Leistungsveränderungen in Englisch lassen sich verschiedene hypothetische Erklärungsansätze aufzeigen, deren Validität in weiteren Studien geprüft werden sollte. Zum einen ist nach Köller, Baumert et al. (2000) ein Effekt von Interessen auf nachfolgende Leistungen erst dann möglich, wenn institutionelle Rahmenbedingungen, wie Kurswahlen, Freiräume für interessengesteuertes Lernen ermöglichen. Ergänzend lässt sich vermuten, dass nicht nur institutionelle, sondern auch curriculare Rahmenbedingungen für die Bedeutung von Interessen in Hinblick auf Leistungsveränderungen entscheidend sein könnten. Zwischen beiden Messungen könnten im Fach Englisch Inhalte unterrichtet worden sein, die kaum Freiräume für interessengesteuertes Lernen ließen. Ebenso könnte die Richtung des Effekts umgekehrt sein und Leistungsveränderungen in Englisch einen Effekt auf die beruflichen Interessenorientierungen ausüben. Dagegen spricht jedoch, dass berufliche Interessen vergleichsweise stabil sind (Low et al., 2005). Schließlich lässt sich vermuten, dass die fachspezifischen Unterschiede in der Bedeutung beruflicher Interessen zu einem Teil auf die Stichprobe, nämlich Schülerinnen und Schüler an einem Fachgymnasium Wirtschaft, zurückzuführen sein könnte. Der wirtschaftswissenschaftli-

che Schwerpunkt der Schulform könnte noch eher zu einer Assoziation von beruflichen Interessen mit Leistungsveränderungen in Mathematik statt in Englisch geführt haben.

Bisherige Längsschnittstudien in der gymnasialen Oberstufe, in denen Assoziationen zwischen fachlichen Interessen, Selbstkonzept, Kursniveau und Leistungen untersucht wurden, beschränken sich bislang auf das Fach Mathematik (z.B. Köller et al., 2006). Die Ergebnisse dieser Studie im Fach Englisch können deshalb nicht mit Studien zu fachlichen Interessen in Beziehung gesetzt werden. Es deutet sich jedoch an, dass Befunde zum Zusammenhang zwischen Interessen, Selbstkonzept, Kursniveau und Leistungen nur teilweise von der mathematischen auf die sprachliche Domäne übertragbar sein könnten. Hier eröffnet sich ein breites Forschungsfeld, in dem fachspezifische Unterschiede eingehender untersucht werden sollten.

Grenzen der Studie. In dieser Studie wurde die Bedeutung beruflicher Interessen für Leistungsveränderungen in zwei Kernfächern der gymnasialen Oberstufe untersucht. Es bleibt jedoch offen, inwieweit sich die Befunde auf Besonderheiten der untersuchten Schule (Leistungsgruppierung nach angestrebtem Kursniveau bereits in der 11. Klasse) oder Schulform zurückführen lassen. Aufgrund der von allgemein bildenden Gymnasien zu unterscheidenden Schülerschaft an Fachgymnasien (vermehrt ehemalige Realschülerinnen und -schüler) könnten die Inhalte des (für die gymnasiale Oberstufe allgemein gültigen) Curriculums an Fachgymnasien unterschiedliche Schwerpunktsetzungen (z. B. mehr Wiederholungen von Inhalten aus der Sekundarstufe I) erfahren haben. Freiheitsgrade in der gymnasialen Oberstufe für interessengesteuertes Lernen und somit die Größe von Interesseneffekten würden dann systematisch zwischen den Schulformen variieren. Des Weiteren ist im Rahmen dieser Studie nicht prüfbar, inwieweit in Fachgymnasien mit spezifischer beruflicher Schwerpunktsetzung tendenziell Effekte beruflicher Interessen auf Leistungen begünstigen. Aufgrund dieser Aspekte verbietet sich eine Generalisierung der Ergebnisse, solange die Befunde nicht an einer unabhängigen Stichprobe repliziert wurden.

5. Zusammenfassende Schlussdiskussion und Ausblick

Den Kern dieser Arbeit bilden drei übergeordnete Forschungsfragen (Kapitel 3), denen in drei empirischen Studien (Kapitel 4) jeweils differenzierter nachgegangen wurde. Hieran anschließend werden im folgenden Abschnitt 5.1 zentrale Ergebnisse dieser Studien zusammengefasst und in den gegenwärtigen Forschungsstand eingeordnet. Bevor die Befunde im Abschnitt 5.3 integrativ diskutiert werden, ist es notwendig, zunächst auf wesentliche methodische Entwicklungen einzugehen und Konsequenzen für die Interpretation der Ergebnisse zu besprechen (Abschnitt 5.2). Abschließend werden die Grenzen dieser Arbeit in Bezug auf die betrachteten Beziehungen zwischen den berücksichtigten Variablen sowie die Selektivität der für die empirischen Studien einbezogenen Stichprobe aufgezeigt und diskutiert (Abschnitt 5.4).

5.1 Empirische Erträge und Einordnung in den aktuellen Forschungsstand

Studie 1 (Abschnitt 4.1) erfolgte unter der Fragestellung, welche Leistungsentwicklungen sich in Mathematik und Englisch in der Einführungsphase der gymnasialen Oberstufe zeigen lassen. Entgegen bisheriger Studien, in denen die abschließende Erhebung bei der Erfassung von Leistungsveränderungen am Ende der gymnasialen Oberstufe lag (z.B. Lehmann et al., 2004, 2006; Ramseier et al., 1999), wurde ein Absinken der mittleren Leistungen in mathematischer Grundbildung innerhalb der Einführungsphase aufgezeigt. Im Fach Englisch hingegen wurde in Übereinstimmung mit bisherigen Studien in der gymnasialen Oberstufe (vgl. Köller, Baumert et al., 2004; Lehmann et al., 2004, 2006) eine im Mittel bedeutsame Leistungszunahme bei den Schülerinnen und Schülern festgestellt. Differenzierter, in Hinblick auf das jeweils ab der Qualifikationsphase (12. und 13. Jahrgangsstufe) angestrebte Kursniveau, bestätigten sich die vermuteten Leistungsunterschiede zwischen den Schülerinnen und Schülern in beiden Fächern zu beiden Erhebungen in der Einführungsphase. Dieser Befund aus der Jahrgangsstufe 11 ergänzt bisherige Studien, in denen deutlich bessere Leistungen in Kursen mit erhöhten Anforderungen im Vergleich zu Kursen mit grundlegenden Anforderungen am Ende der gymnasialen Oberstufe nachgewiesen wurden (vgl. Köller, Watermann et al., 2004; Lehmann et al., 2006). Laut den Rahmenrichtlinien für die Fächer Mathematik und Eng-

lisch ist das Ausgleichen von Defiziten bei unterschiedlicher Eingangsniveaus zur Vorbereitung auf die Qualifikationsphase eine zentrale Zielsetzung des Unterrichts in der Einführungsphase (vgl. Niedersächsisches Kultusministerium, 1990, 2003). Andererseits ließen sich im Verlauf der gymnasialen Oberstufe zunehmende Leistungsdifferenzierungen durch den Unterricht auf unterschiedlichem Niveau sowie systematischer kognitiver und motivationaler Differenzen zwischen den Schülerinnen und Schülern vermuten (vgl. Köller et al., 2006). Die Daten dieser Studie zeigten für die Einführungsphase, dass es weder zu einer mittleren Angleichung der Leistungen, noch zu einer statistisch nachweisbaren zunehmenden Leistungsdifferenzierung kommt.

In Studie 2 (Abschnitt 4.2) wurde ein methodischer Zwischenschritt vorgenommen, um der anschließenden dritten Fragestellung nachgehen zu können. Am Beispiel des Fachs Mathematik wurde geprüft, ob berufliche Interessenprofile in Anlehnung an ein Structural Summary nach Gurtman und Balakrishnan (1998) als eine valide Reduktion von Skalenwerten im Rahmen der Theorie der Berufswahl nach Holland (1997) angesehen werden können. Insbesondere für längsschnittliche Analysen in Form von Regressionsanalysen lassen sich Interessenprofile als wertvolle Alternative zu Skalenwerten erachten, da letztere mit statistischen Problemen (wie Inflation des aufgeklärten Varianzanteils und Multikollinearität) einhergehen können. Ergebnisse dieser Arbeit bestätigten die Validität der Reduktion: Interessenprofile wiesen eine zu Skalenwerten vergleichbare Vorhersagekraft für Mathematikleistungen und -leistungsveränderungen auf. Über beide Methoden ließen sich inhaltlich vergleichbare Assoziationen zwischen beruflichen Interessenorientierungen und Leistungen bzw. Leistungsveränderungen im Fach Mathematik nachweisen. Darüber hinaus waren die Befunde beim Verwenden von Interessenprofilen robuster: Im Längsschnitt zeigte sich nach Kontrolle der Kovariaten Geschlecht und Ausgangsleistung, dass nur Interessenprofile weiterhin einen bedeutsamen Effekt auf Mathematikleistungen ausübten.

Auf Basis dieser Validitätshinweise war es möglich, im Abschnitt 4.3 der dritten Fragestellung nachzugehen. Es wurde geprüft, welchen Effekt berufliche Interessenprofile auf die Leistungsentwicklung in Mathematik und Englisch in der Einführungsphase der gymnasialen Oberstufe haben und welche Zusammenhänge zwischen beruflichen Interessen, Selbstkonzept und Kursniveau in Hinblick auf Leistungsentwicklungen in den beiden Fächern bestehen. Vorab sei auf die von Darcy und Tracey (2003) als zentral

bei der Erfassung beruflicher Interessen angesehene Rolle des Levels bzw. des Antwortstils (vgl. Prediger, 1998) hingewiesen. In dieser Studie ließ sich für zwei Fächer und zwei Erhebungszeitpunkte kein direkter Effekt des Levels auf Leistungen und deren Veränderung nachweisen. Dieser Befund stimmt überein mit Ergebnissen von Prediger (1998); er widerspricht aber einer Hypothese von Darcy und Tracey (2003), nach der der Level ein Indikator für Flexibilität sei und somit in positivem Zusammenhang mit Leistungen stehen sollte. Weiterhin zeigte sich, dass die Things/People-Dimension – entsprechend der Orientierung Realistic – einen positiven Effekt auf Leistungsveränderungen in Mathematik ausübte. Damit wurden Assoziationen aus Querschnittstudien (Ackerman & Heggestad, 1997; Bergmann & Eder, 1992; Nagy, 2006) erstmalig längsschnittlich für das Fach Mathematik bestätigt. Tracey et al. (2005) konnten im Längsschnitt bislang keine zeitversetzten Effekte zwischen den RIASEC-Skalenwerten und Testleistungen in Mathematik und Englisch nachweisen. Unabhängig von der verwendeten Modellierung ließ sich auch in dieser Arbeit die Bedeutung beruflicher Interessen in Hinblick auf Leistungsveränderungen im Fach Englisch nicht nachweisen. Die erwartbaren Korrelationen zwischen beruflichen Interessenorientierungen und Englischleistungen zeigten sich nur im Querschnitt.

Ein weiterer Fokus der dritten Studie lag auf dem Wechselspiel zwischen beruflichen Interessen, Selbstkonzept, Kursniveau und Leistungsentwicklungen. Die Ergebnisse bestätigten die erwarteten Effekte des fachspezifischen Selbstkonzepts (vgl. z. B. vgl. Eccles[-Parsons] et al., 1984; Fredricks & Eccles, 2002; Marsh, 1990a) und des Kursniveaus (vgl. z. B. Baumert & Watermann, 2000; Köller & Trautwein, 2004; Köller et al., 2006; Watermann et al., 2004) auf Leistungsentwicklungen in beiden Fächern. Für Mathematik zeigte sich darüber hinaus, dass die Effekte von Selbstkonzept und Kursniveau jeweils durch die beruflichen Interessenorientierungen mediiert waren. Dieser Befund stimmt somit im Fach Mathematik mit ersten Längsschnittstudien zum Zusammenhang von Interesse, Selbstkonzept und Kursniveau (z. B. Köller et al., 2006) überein und ergänzt diese um den Aspekt beruflicher Interessen.

5.2 Methodische Entwicklungen im Verlauf der empirischen Studien

Oftmals gibt es bei der längsschnittlichen Auswertung von im schulischen Kontext erhobenen Daten insbesondere zwei methodische Herausforderungen: der Umgang mit fehlenden Werten und mit der hierarchischen Struktur der Daten. Im Verlauf der empirischen Studien wurden unterschiedliche methodische Ansätze zum Umgang mit diesen Herausforderungen verwendet. Im Folgenden werden die Erträge dieser Arbeit im Kontext der jeweils verwendeten Methoden diskutiert.

5.2.1 *Missing Data*

Wie in Abschnitt 1.2 beschrieben, stehen in Längsschnittstudien häufig nicht alle Schülerinnen und Schüler, die an der ersten Messung teilgenommen haben, auch für weitere Erhebungen zur Verfügung. Dies kann insbesondere in der gymnasialen Oberstufe durch einen vorzeitigen Weggang aus der Schule oder die Wiederholung einer Jahrgangsstufe begründet sein. Ebenso kommen generell bei längsschnittlichen Erhebungen auftretende Ausfallprozesse in Frage (bspw. mangelnde Bereitschaft für eine weitere Teilnahme). In dieser Arbeit liegt mit knapp 20% eine vergleichsweise hohe Ausfallquote vor. Diese geht vorrangig auf die Schulform, das Fachgymnasium, zurück. Fachgymnasien richten sich primär an eine nicht-gymnasiale Schülerklientel. Sie unterliegen gleichzeitig der Vereinbarung zur Gestaltung der gymnasialen Oberstufe mit ihren Zielen einer vertieften Allgemeinbildung und einer allgemeinen Studierfähigkeit (vgl. KMK, 2006). Zusätzlich findet eine berufliche Spezialisierung entsprechend des Profils des Fachgymnasiums statt. Der Unterricht an Fachgymnasien ist für die Schülerinnen und Schüler somit kognitiv relativ anspruchsvoll, was eine hohe Abbruch- und Wiederholerquote mit sich bringt.

In Studie 1 wurden Personen mit fehlenden Werten fallweise aus den Analysen ausgeschlossen (*Listwise Deletion*). Anhand von Selektivitätsanalysen wurde gezeigt, dass sich Schülerinnen und Schüler, die an beiden Erhebungen teilgenommen haben, in ihren Mathematik- und Englischleistungen zum ersten Messzeitpunkt nicht systematisch von denjenigen, die nur an der ersten Erhebung teilgenommen haben, unterscheiden. Mit dieser Information ließ sich ausschließen, dass Leistungsunterschiede zu einem späteren Zeitpunkt möglicherweise bereits auf systematische Differenzen zwischen den

Teilnehmenden zu Beginn der Studie zurückzuführen waren. Dennoch musste bei fallweisem Ausschluss aufgrund der hohen Ausfallquote von einem Verlust an Effizienz und statistischer Power ausgegangen werden (vgl. Lüdtke et al., 2007). Darüber hinaus besteht, sofern der Ausfallprozess systematisch mit den abhängigen Variablen zusammenhängt, die Gefahr verzerrter Parameterschätzungen. Hiervon war auch in dieser Arbeit auszugehen, denn besonders in Schuluntersuchungen scheiden insbesondere Schülerinnen und Schüler aufgrund unzureichender Leistungsentwicklungen aus der Jahrgangsstufe, und damit aus dem Längsschnitt, aus. Um der Verzerrung in der Parameterschätzung entgegen zu wirken, wäre es in Zukunft wünschenswert, dass die Ursachen für den Ausfall systematisch erfasst werden und sich daran anschließend insbesondere bei Wiederholern oder Abbrechern um eine kontinuierliche Längsschnitterhebung bemüht wird.⁸

Ein solches Vorgehen konnte im Rahmen dieser Arbeit nicht realisiert werden, weshalb in den anschließenden beiden empirischen Studien eine andere Methode zum Umgang mit fehlenden Werten gewählt wurde. Entsprechend des aktuellen Forschungsstands in diesem Bereich (vgl. z. B. Lüdtke et al., 2007) wurde das *Multiple Imputation*-Verfahren, das mehrfache Schätzen fehlender Werte, verwendet (Rubin, 1987). Hierfür wird ein Hintergrundmodell verwendet, in das alle verfügbaren Variablen (Hilfsvariablen) einbezogen werden, von denen angenommen wird, dass sie im Zusammenhang mit dem Ausfallprozess oder den in anschließenden Analysen verwendeten Variablen stehen. Simulationsstudien belegen, dass sich mit diesen Hilfsvariablen die Parameterschätzung verbessert und dass die Aufnahme von Variablen, die nicht mit dem Ausfallprozess oder den zu analysierenden Variablen assoziiert sind, zu keiner Verschlechterung der Parameterschätzungen führt (vgl. Collins et al., 2001).

Ein zunächst unerwartetes Ergebnis war, dass beim fallweisen Ausschluss in Studie 1 die Mathematikleistung in der Einführungsphase im Mittel sank ($p < .05$). Hingegen blieb sie konstant, wenn für die Analysen vollständige Datensätze auf Basis von Multiple Imputation – und somit erwartbarer größerer Effizienz und Power –

⁸ In Schuluntersuchungen, in denen eine Verallgemeinerung auf Populationsparameter angestrebt wird, würden Schätzungen nicht nur auf Basis von Schülerinnen und Schülern mit regulären Schulkarrieren basieren.

verwendet wurden. Bevor nachfolgend eine mögliche Erklärung vorgestellt wird, muss zunächst das Ausmaß beschrieben werden, in dem sich die Ergebnisse voneinander unterscheiden. So lag das Signifikanzniveau für die Veränderung der Mathematikleistung in Studie 1 bei $p = .03$ und in den Studien 2 und 3 nach Multiple Imputation bei $p = .08$. So resultierten zwar bei Verwenden dieses Cutt-Off-Kriteriums unterschiedliche Signifikanz-Aussagen, jedoch handelt es sich um Werte innerhalb eines statistischen Grenzbereichs um ein Fehlerniveau von $\alpha = .05$ herum. Aufgrund erwartbar höherer Effizienz und Power bei Multiple Imputation müssen die Befunde dennoch eingehender betrachtet werden. Bei Multiple Imputation geschieht die Schätzung fehlender Werte unter Berücksichtigung zusätzlicher, in einem Hintergrundmodell bereit gestellter, Informationen über die zu schätzenden Variablen oder den Ausfallprozess. Aufgrund von Ergebnissen aus Simulationsstudien (vgl. Collins et al., 2001) lässt sich ein geringer Bias in den Parameterschätzungen erwarten; diese sollten also tendenziell den tatsächlichen Leistungsveränderungen in Mathematik entsprechen. Die in Studie 1 verwendeten Selektivitätsanalysen zum ersten Erhebungszeitpunkt können systematische Unterschiede zwischen beobachteten und fehlenden Werten nur in bedingtem Ausmaß erfassen und Parameterschätzungen nach fallweisem Ausschluss verzerrt sein.

5.2.2 Clustereffekte

In der vorliegenden Arbeit liegt eine hierarchische Struktur der Daten vor: Schülerinnen und Schüler aus acht Klassen bilden die Datengrundlage. Sofern sich diese hinsichtlich zentral erachteter Merkmale innerhalb der Klassen ähnlicher sind als eine zufällig zusammengesetzte Stichprobe, liegen Clustereffekte vor. Damit geht eine Unterschätzung des Standardfehlers einher, sodass Signifikanzprüfungen tendenziell zu liberal angesetzt werden.

In den ersten beiden Studien wurde die hierarchische Struktur der Daten in den statistischen Analysen nicht berücksichtigt, sondern auf mögliche Clustereffekte in der Diskussion als einen limitierenden Faktor hingewiesen. In Studie 3 wurde dieser möglichen Einschränkung durch eine in Mplus (Muthén & Muthén, 1998–2008) implementierte Korrektur der Standardfehler entgegnet. Da in dieser Studie alle im Rahmen dieser Arbeit interessierenden Variablen enthalten waren, wurden für diesen Diskussionspunkt

die entsprechenden Analysen noch einmal unter Vernachlässigung möglicher Clustereffekte durchgeführt. Während der Standardfehler für berufliche Interessen, Selbstkonzept und Mathematikleistungen meist unverändert blieb, lag die größte Abweichung für die Englischleistung zur ersten Erhebungen als Prädiktor vor ($\Delta SE = .04$). Schülerinnen und Schüler waren sich innerhalb der Klassen in ihren Englischleistungen somit vergleichsweise ähnlicher als zufällig ausgewählte Personen. Insgesamt waren die Vergrößerungen der Standardfehler durch Berücksichtigung der hierarchischen Datenstruktur als gering zu betrachten; sie führten nicht zu unterschiedlichen Signifikanzaussagen. Aufgrund dieser Ergebnisse lässt sich auch in Hinblick auf die ersten beiden Studien vermuten, dass Clustereffekte die Aussagekraft dieser Arbeit insgesamt nicht wesentlich einschränken.

5.3 Integrative Diskussion zur Bedeutung von beruflichen Interessen für schulische Leistungsentwicklungen

Auf Basis der empirischen Befunde zu Leistungsentwicklungen und beruflichen Interessen wird zunächst auf den fachspezifischen Zusammenhang von beruflichen Interessen mit schulischen Leistungen und Leistungsentwicklungen in der gymnasialen Oberstufe eingegangen. Anschließend werden Befunde zum Effekt beruflicher Interessenprofilparameter auf Leistungen und deren Veränderung beschrieben, Implikationen für den Unterricht abgeleitet und diskutiert sowie daran anknüpfende Forschungsfelder aufgezeigt.

5.3.1 Fachspezifische Unterschiede bei der Bedeutung beruflicher Interessen

Berufliche Interessen bilden sich im Jugendalter heraus (z.B. Tracey, 2001). Ihnen liegt eine kognitive Repräsentation der gesamten Berufslandschaft zugrunde, die sich wiederum in bestimmte Berufsfelder bzw. Tätigkeitsklassen einteilen lässt (Rolf, 2001). Im Rahmen einer Person-Gegenstands-Konzeption wird dem Interesse eine handlungsleitende Funktion zugeschrieben (Prenzel, 1998). Diese lässt sich auch für berufliche Interessen annehmen. Sie sind einflussreiche Faktoren bei Studienfachwahlen (Nagy, 2006) und stehen in Zusammenhang mit Leistungen. Ackerman (1996) nimmt

hierzu an, dass sich berufliche Interessen und Fähigkeiten in einer wechselseitigen Beziehung zueinander entwickeln. Zusammen mit Heggstad (1997) wies er Korrelationen zwischen beruflichen Interessenorientierungen und mathematischen bzw. sprachlichen Leistungen in vergleichbarer Höhe nach. Da die Erforschung von Determinanten schulischen Lernens einen zentralen Stellenwert in der pädagogisch-psychologischen Forschung besitzt, ist es bemerkenswert, dass im deutschsprachigen Raum vergleichsweise wenig Querschnittstudien (Bergmann & Eder, 1992; Eder & Reiter, 2002; Nagy, 2006) und bislang keine längsschnittlichen Untersuchungen vorliegen. Ergebnisse bisheriger Arbeiten bestätigend und daran anknüpfend, wurden in dieser Arbeit im Querschnitt Zusammenhänge zwischen beruflichen Interessenorientierungen und Mathematik- bzw. Englischleistungen in der gymnasialen Oberstufe nachgewiesen. Köller, Baumert et al. (2000) betonten die Bedeutung des institutionellen Kontextes für die Bedeutung von Interessen in Hinblick auf Leistungen und Leistungsentwicklungen. Ein Effekt von Interesse auf Leistungen sei erst in der gymnasialen Oberstufe und damit einhergehender Freiräume, wie Kurswahlen, möglich. Dieses wurde in der vorliegenden Arbeit für berufliche Interessen am Beispiel des Fachs Mathematik längsschnittlich belegt. Wider Erwarten ließen sich zeitversetzte Effekte von beruflichen Interessen auf Englischleistungen nicht nachweisen. Dieser fachspezifische Unterschied könnte ein Hinweis darauf sein, dass der institutionelle Kontext zwar eine notwendige, aber möglicherweise keine hinreichende Bedingung für die Bedeutung von Interessen in Hinblick auf Leistungsentwicklungen ist. Darüber hinaus belegten Marsh et al. (2005) Effekte von Interessen auf Leistungen bereits in der Sekundarstufe I. Für ein Verständnis des fachspezifischen Zusammenhangs bedarf es weiterer Forschung, in der den individuellen und kontextuellen Merkmalen nachgegangen wird, die für die Bedeutung beruflicher Interessen in Hinblick auf Leistungsentwicklungen relevant sein könnten.

5.3.2 Implikationen für den Unterricht

Aus den in dieser Arbeit aufgezeigten Zusammenhängen zwischen beruflichen Interessenorientierungen und Leistungen lassen sich Implikationen für den Unterricht aufzeigen. Sofern berufliche Interessen(-profile) verwendet werden, sind Implikationen aufgrund damit einhergehender Beschreibungsparameter facettenreicher als bei regulä-

ren Skalenwerten. So werden Theorien beruflicher Interessen meist als Strukturmodelle beschrieben (vgl. z.B. Savickas & Spokane, 1999). Der empirisch gut bestätigten Theorie der Berufswahl nach Holland (1997) liegt die Calculus-Annahme zugrunde, wonach sich berufliche Interessen zweidimensional in Form einer Quasi-Circumplex-Struktur (Guttman, 1954) anordnen lassen. Damit geht eine typische Gestalt individueller Interessenprofile einher (Gurtman & Pincus, 2003): eine Person hat ein hohes Interesse an einem bestimmten Tätigkeitsbereich (dominante Interessenorientierung); je unähnlicher andere Tätigkeitsbereiche diesem dominierenden Bereich sind, desto weniger interessiert sie sich dafür. Hieraus lassen sich Implikationen für den Zusammenhang von beruflichen Interessenorientierungen und Leistungen ableiten, die sich auch in der Datenstruktur dieser Arbeit zeigten: hohe Interessen im Bereich Realistic gingen einher mit vergleichsweise positiven Leistungen und Leistungsentwicklungen in mathematischer Grundbildung, während bei vorherrschenden Interessen in den räumlich (weitgehend) entgegen gesetzten Bereichen Artistic und Social vergleichsweise negative Leistungen und Leistungsveränderungen aufgezeigt wurden. Instruktionen im Unterricht und die Unterrichtsgestaltung in Mathematik sollten deshalb möglichst auch Schülerinnen und Schüler mit vergleichsweise hohen Interessen in den Bereichen Artistic und Social ansprechen, um einem Leistungsabfall in mathematischer Grundbildung in der Einführungsphase entgegenzuwirken. So könnten Anwendungen mathematischen Wissens auf Tätigkeiten im sozialen und künstlerischen Bereich aufgezeigt werden (z.B. Formeln zur Flächen- und Volumenberechnung zur Vermessung und Ausgestaltung eines Kunstobjekts). Die Förderung bestimmter Interessenorientierung scheint hingegen wenig angemessen. Zum einen werden berufliche Interessen als relativ stabil angesehen (vgl. Low et al., 2005). Zum anderen könnte aufgrund der zugrunde liegenden Assoziationsstruktur beruflicher Interessenprofile die spezielle Förderung eines Bereichs mit negativen Veränderungen in anderen Bereichen, und somit möglicherweise auch Leistungsveränderungen in anderen Fächern, einhergehen.

Neben der Interessenorientierung lassen sich weitere Charakteristika beruflicher Interessenprofile ableiten und deren Zusammenhang mit schulischen Leistungen und Leistungsentwicklungen untersuchen. So beschreibt beispielsweise die Differenziertheit das Ausmaß, in dem sich die am stärksten präferierte Orientierung von den anderen hervorhebt. Anhand von Profilanalysen in Studie 2 wurde veranschaulicht, dass die Dif-

ferenziertheit eines Profils den Zusammenhang von beruflichen Interessenorientierungen und Leistungen moderiert (Figure 3): je differenzierter ein Profil, desto stärker der Interessen-Leistungs-Zusammenhang. Die alleinige Betrachtung der Interessenorientierung führt somit zu einer unvollständigen Beschreibung des Zusammenhangs; vielmehr muss die relative Stärke in Hinblick auf die Ausprägung der anderen Orientierungen berücksichtigt werden. Zusammengefasst bedeutet dies, dass sich bei Schülerinnen und Schülern mit einem differenzierten Profil und hohen Interessen im Bereich Realistic besonders positive Leistungen und Leistungsveränderungen auffinden lassen, während diese bei differenzierten Profilen mit Interessen im Bereich Artistic und Social vergleichsweise negativ ausfallen. Letztere stellen also eine „Risikogruppe“ für die Entwicklung mathematischer Grundbildung dar. Eine Ausdifferenzierung beruflicher Interessen wird dabei als zentrale Entwicklungsaufgabe angesehen (vgl. Eder, 1998). Entsprechende Implikationen setzen somit auch hier nicht an der Veränderung beruflicher Interessen an. Vielmehr scheint es für die Unterrichtsgestaltung und Themenbearbeitung in Mathematik wichtig zu sein, insbesondere Schülerinnen und Schüler mit differenzierten Profilen im Bereich Social und Artistic zu erreichen.

Ein drittes Merkmal ist schließlich die generelle Zustimmung oder Ablehnung beruflicher Tätigkeiten, der Level bzw. Antwortfaktor. In dieser Studie zeigt sich konstant über beide Messzeitpunkte und Fächer hinweg, dass der Level in keiner direkten Beziehung zu Leistungen und Leistungsentwicklungen in Mathematik und Englisch steht. In Kombination mit den beiden zuvor berichteten Charakteristika (Orientierung und Differenziertheit) bedeutet dies für das Fach Mathematik, dass es für Leistungen und Leistungsveränderungen nicht wichtig ist, *wie viele* Interessen eine Schülerin oder ein Schüler hat, sondern vielmehr *welche* und *wie stark* ausgeprägt diese im Vergleich zu anderen Orientierungen sind.

Abschließend lässt sich anhand dieser Implikationen festhalten, dass es bei beruflichen Interessen im Gegensatz zu fachlichen Interessen weniger darum geht, Auslöse- und Entstehungsbedingungen für Interessen zu analysieren (Krapp, 1992), sondern vielmehr um die Beschreibung des mit beruflichen Interessen einhergehenden Verhaltens. Dennoch weisen beide Interessenbereiche zentrale konzeptuelle Gemeinsamkeiten auf (vgl. Abschnitt 2.2). Insbesondere in den letzten Jahren wurde im deutschsprachigen Raum eine engere Anbindung beruflicher Interessen an die Person-Gegenstands-Kon-

zeption des Interesses angestrebt und spezifische berufliche Tätigkeitsklassen formuliert (vgl. Bergmann & Eder, 2005). Hieraus resultieren nicht nur theoretische, sondern auch empirische Parallelen: Im Querschnitt lassen sich Assoziationen zwischen beruflichen und fachlichen Interessen mit Leistungen in Mathematik und Englisch belegen (vgl. z.B. Nagy, 2006; Schiefele et al., 1993). Eine Bedeutung fachlicher (Köller, Baumert et al. 2000; Köller et al., 2006) oder beruflicher Interessen (Warwas, Nagy et al., in press; Warwas et al., eingereicht) ließ sich auch im Längsschnitt in Hinblick auf die Veränderung von Mathematikleistungen in der gymnasialen Oberstufe nachweisen. Entsprechende Studien für das Fach Englisch liegen bezüglich des fachlichen Interesses nicht vor. Für berufliche Interessen ließen sich keine Effekte auf Leistungsentwicklungen aufzeigen (Warwas et al., eingereicht). Insgesamt erscheint deshalb eine vergleichende Betrachtung der Bedeutung von fachlichen und beruflichen Interessen für weitere schulische Fächer lohnenswert.

5.4 Grenzen der Studie

In dieser Arbeit wurden Leistungsentwicklungen in Mathematik und Englisch in der gymnasialen Oberstufe untersucht und die Bedeutung beruflicher Interessen für Leistungsentwicklungen geprüft. Mit diesen Zielsetzungen gehen bestimmte Forschungsdesigns und Analysemethoden einher, deren Befunde im Folgenden einer kritischen Betrachtung unterzogen werden. Abschließend wird auf die Datenbasis dieser Studie und entsprechender Grenzen in der Generalisierbarkeit der Befunde hingewiesen.

5.4.1 Effektrichtungen

Entsprechend den Fragestellungen dieser Arbeit war es ein zentrales Ziel, die Bedeutung beruflicher Interessen für schulisches Lernen zu analysieren. Den Hintergrund bildeten bereichsspezifische Modelle und Theorien, in denen gerichtete Beziehungen von Interessen (und weiteren damit zusammenhängenden Variablen) in Hinblick auf Leistungen beschrieben werden (wie Erwartungs-mal-Wert-Theorien). Diese Annahmen haben bislang, und auch in dieser Studie, breite empirische Bestätigung erfahren. Ungeachtet dessen ist es notwendig darauf hinzuweisen, dass mit der Prüfung

unidirektionaler Effekte von verschiedenen Determinanten auf Leistungsentwicklungen umgekehrt gerichtete Effekte oder reziproke Beziehungen nicht berücksichtigt wurden.

Bereits bei der Beschreibung des Zusammenhangs von Kursniveau und Interessen wurde darauf hingewiesen, dass sich nicht nur das Kursniveau positiv auf Interessen auswirkt, sondern Interessen eine zentrale Rolle für die Kurswahl haben, und dass beide somit in wechselseitiger Beziehung zueinander stehen. Da in der vorliegenden Stichprobe die (u. a. interessengeleitete) Wahl eines Kurses bereits mit Eintritt in das Fachgymnasium stattfand, erschien es plausibel, ausschließlich den Effekt vom Kursniveau auf das Interesse zu untersuchen. In Hinblick auf die gerichteten Beziehungen zwischen Interesse und Leistungen wiesen bereits Köller, Baumert et al. (2000) auf die Bedeutung des Kontextes hin. Diese Studie bezog sich, wie der Großteil längsschnittlicher Untersuchungen in der gymnasialen Oberstufe, auf das Fach Mathematik. Und auch Marsh et al. (2005) prüften anhand zweier Datensätze mit Schülerinnen und Schülern der Sekundarstufe I reziproke Effektmodelle zwischen Interessen, Selbstkonzept und Leistungen in Mathematik: In dieser Studie zeigten sich wechselseitige Effekte zwischen Selbstkonzept und Interesse, ein unidirektionaler Effekt von Interesse auf Testleistungen sowie größere Effekte vom Selbstkonzept auf Testleistungen als umgekehrt. Diese Befunde sprechen in gewisser Hinsicht für die Validität der in dieser Studie spezifizierten unidirektionalen Beziehungen, weisen jedoch auch auf mögliche weitere – wenngleich schwächere – Effekte zwischen Interessen, Selbstkonzept und Leistungen hin. Deren Berücksichtigung könnte in zukünftigen Studien ergänzende Informationen darüber liefern, welche bislang nicht berücksichtigten Assoziationen zwischen den Variablen für Leistungsentwicklungen bedeutsam sind, bzw. welche Konsequenzen diese Leistungsentwicklungen für die Entwicklung von Interessen und Selbstkonzept haben. Als ein wichtiges Forschungsdesiderat ist hinzuzufügen, dass die Prüfung reziproker Beziehungen zwischen Interesse, Selbstkonzept, Kursniveau und Leistungsentwicklungen weitere Fächer mit einbeziehen sollte.

5.4.2 Selektivität der Stichprobe

Abschließend sei auf die Selektivität der Stichprobe und entsprechender Einschränkungen in der Generalisierbarkeit der Befunde hingewiesen. Im Folgenden wer-

den Leistungsentwicklungen und die Bedeutung beruflicher Interessen für diese Leistungsentwicklungen unter Berücksichtigung dreier Besonderheiten der Datenbasis diskutiert: die Schulform in der gymnasialen Oberstufe (Fachgymnasium), das berufliche Profil des Fachgymnasiums (Wirtschaft) und das spezielle Auswahlverfahren zur Zulassung von Schülerinnen und Schülern am untersuchten Fachgymnasium.

Die gymnasiale Oberstufe kann an ganz unterschiedlichen Schulformen absolviert werden. Hierzu gehören im Wesentlichen allgemein bildende Gymnasien, Fachgymnasien und Gesamtschulen. All diese Schulformen unterliegen einer bundesweit einheitlichen Vereinbarung zur Gestaltung der gymnasialen Oberstufe (vgl. KMK, 2006). Insofern erfahren die Kernfächer auf struktureller Ebene eine über die Schulformen hinweg vergleichbare Bedeutung. Auch die jeweils für die Fächer spezifizierten Curricula der Einführungsphase gelten für die gymnasiale Oberstufe im Allgemeinen und somit für alle entsprechenden Schulformen (vgl. Niedersächsisches Kultusministerium, 1990, 2003). Hingegen muss die unterschiedliche Zusammensetzung der Schülerschaft in Abhängigkeit von den Schulformen berücksichtigt werden. Während an allgemein bildenden Gymnasien die Oberstufe mehrheitlich von einer primär gymnasialen Schülerschaft (seit Beginn der Sekundarstufe I) absolviert wird, sind Fachgymnasien in der Regel Schulen in Aufbauform. Als eine Form der Öffnung von Bildungswegen richten sie sich primär an Schülerinnen und Schüler, die ihren Sekundarschulabschluss an Real- oder Hauptschulen erworben haben. In einer Querschnittstudie in der Schweiz belegten Ramseier et al. (1999), dass die Leistungen von Schülerinnen und Schülern in mathematischer Grundbildung im Verlauf der gymnasialen Oberstufe bedeutsam anstiegen, während Leistungen bei berufsbildenden Schülerinnen und Schülern im Mittel konstant blieben. Da diese Ergebnisse zum einen auf einer Querschnittstudie basieren und zum anderen die verschiedenen Stichproben im Abstand von zwei Jahren betrachtet wurde, können diese Befunde jedoch nicht ohne weiteres auf die hier vorgenommene Studie übertragen werden. In den LAU 11- und 13-Studien (Lehmann et al., 2004, 2006) zeigten sich sowohl in mathematischer Grundbildung als auch in Englisch vergleichbare Leistungsveränderungen zwischen den Schulformen. Da sich diese Befunde auf einen Zeitraum von zwei Jahren beziehen, sollte dennoch ergänzend geprüft werden, ob die Leistungsentwicklungen innerhalb der Einführungsphase der gymnasialen Oberstufe zwischen den Schulformen vergleichbar sind.

Innerhalb der Schulform Fachgymnasium gibt es unterschiedliche Schwerpunktsetzungen bzw. Profile, wie das der Wirtschaft, Technik oder Sozialpädagogik. In Hinblick auf diese unterschiedlichen Schwerpunktsetzungen (berücksichtigt wurden Technik und Wirtschaft) zeigten sich in den LAU-Studien im Mittel bessere Leistungsentwicklungen in mathematischer Grundbildung und in Englisch zugunsten von Fachgymnasien mit Schwerpunkt Technik. Ob diese differenziellen Leistungsentwicklungen in Abhängigkeit von der Ausrichtung des Fachgymnasiums auch für die Einführungsphase zu finden sind, muss in zusätzlichen Studien untersucht werden.

Als dritte Besonderheit soll die Leistungsentwicklung unter Beachtung des speziellen Auswahlverfahrens zu Beginn am untersuchten Fachgymnasium diskutiert werden. Hintergrund des Auswahlverfahrens ist, dass die Zahl der Bewerberinnen und Bewerber die verfügbaren Plätze an Fachgymnasien oftmals überschreitet. Fachgymnasien müssen also gewisse Aufnahmekriterien (wie Schulnoten) festlegen, nach denen die Zulassung erfolgt. Um eine möglichst optimale Auswahl an Schülerinnen und Schülern in Hinblick auf einen erfolgreichen Verlauf am Fachgymnasium zu gewährleisten, fand an der untersuchten Schule eine Eingangsdiagnostik statt, die verschiedene fachbezogene Leistungstests, kognitive und motivationale Variablen enthielt⁹ und auf deren Basis eine Geeignetheits-Rangreihe erstellt wurde (vgl. Dittmar, 2007). Anzumerken ist, dass nicht alle Schülerinnen und Schüler in der untersuchten Stichprobe anhand der Tests ausgewählt wurden, sondern einige bereits vorab aus verschiedenen Gründen eine feste Zusage bekommen hatten. Darüber hinaus war die Ablehnungsquote auf Basis der Testungen mit unter 15% relativ gering. Ob durch diese Selektion vergleichsweise positive Leistungsentwicklungen resultierten, muss in weiteren Studien an Schulen geprüft werden.

Ein zweiter, in Bezug auf die Selektivität der Stichprobe wichtiger Aspekt, betrifft die Generalisierbarkeit der aufgefundenen Bedeutung von beruflichen Interessen für Leistungsentwicklungen. Hintergrund für die Vermutung differenzieller Zusammenhänge in Abhängigkeit von der betrachteten Schulform ist die mit Fachgymnasien

⁹ Berufliche Interessen wurden in der Eingangsdiagnostik dieser Schule nicht berücksichtigt. Die Eingangsdiagnostik wird deshalb im folgenden Absatz, in dem es um die Bedeutung beruflicher Interessen für Leistungsentwicklungen geht, nicht erneut thematisiert.

einhergehende berufsbezogene Schwerpunktsetzung, mit der ein direkter Übergang in entsprechende berufliche Weiterqualifikation eröffnet werden soll (Niedersächsisches Kultusministerium, 1998). So zeigten sich im Mittel differenzielle Interessenprofile von Schülerinnen und Schülern in Abhängigkeit von der besuchten Schulform: An Fachgymnasien gab es meist eine vorherrschende Interessenorientierung, die mit der beruflichen Schwerpunktsetzung des jeweiligen Fachgymnasiums übereinstimmte. An allgemein bildenden Gymnasien war das Interessenprofil vergleichsweise undifferenziert: Es gab keine eindeutig vorherrschende Interessenorientierung, vielmehr lagen die sechs Orientierungen jeweils im mittleren Bereich (vgl. Köller, Watermann et al., 2004). Aufgrund dieses mittleren Profils kann nicht per se auf eine mangelnde Bedeutung beruflicher Interessen für Leistungsentwicklungen an allgemein bildenden Gymnasien geschlussfolgert werden. Da an dieser Schulform keine spezifische berufliche Schwerpunktsetzung erfolgt, ist vielmehr davon auszugehen, dass dem mittleren Profil ganz unterschiedliche individuelle Profile mit vergleichsweise heterogenen bevorzugten Interessen zugrunde liegen. Auch an allgemein bildenden Gymnasien könnten somit spezifische Orientierungen mit Leistungsveränderungen in bestimmten Fächern einhergehen. Hinzu kommt, dass die Assoziationen zwischen bestimmten beruflichen Interessenorientierungen mit Leistungen im mathematischen und sprachlichen Bereich über verschiedene Studien hinweg bestätigt worden sind (vgl. Metaanalyse von Ackerman & Heggstad, 1997). Eine empirische Prüfung, ob die Bedeutung beruflicher Interessen für Leistungsentwicklungen zwischen den Gymnasialschulformen vergleichbar ist, steht noch aus.

Gleichermaßen ist zu hinterfragen, ob es zwischen Fachgymnasien mit verschiedenen Schwerpunktsetzungen unterschiedliche Zusammenhänge zwischen beruflichen Interessenorientierungen und Leistungen gibt, diese also in der vorliegenden Arbeit wesentlich durch die fachliche Ausrichtung der Wirtschaft beeinflusst werden. Zwischen dem Schwerpunkt eines Fachgymnasiums und den beruflichen Interessen der Schülerinnen und Schüler gibt es eine hohe Übereinstimmung bzw. Kongruenz (vgl. z. B. Eder, 1988; Köller, Watermann et al., 2004). Beispielsweise geben Schülerinnen und Schüler an Fachgymnasien mit dem Profil Wirtschaft vergleichsweise hohe Interessen im Bereich Conventional an, solche an Fachgymnasien mit dem Profil Sozialpädagogik hingegen hohe Interessen im Bereich Social. Ein schwerwiegendes methodisches Problem solch

einer homogenen Stichprobe liegt vor, wenn sich die von Holland postulierte Calculus-Annahme nicht in den Daten auffinden und somit nicht die hexagonale Struktur zeigen ließe. Infolge dessen wäre eine Interpretation der Parameter beruflicher Interessen nicht zulässig. Mit dem RANDALL Programm von Tracey (1997) wurde hingegen nachgewiesen, dass die Verteilung beruflicher Interessen der Schülerinnen und Schüler mit der Calculus-Annahme vereinbar war; es zeigte sich eine sehr gute Passung der empirischen Daten mit Hollands Strukturmodell. Unabhängig davon bleibt die Homogenität der beruflichen Interessen bei den untersuchten Schülerinnen und Schülern (mehrheitlich Interessen im Bereich Conventional) nicht ohne Konsequenzen für die berichteten Zusammenhänge mit Selbstkonzept, Kursniveau und Leistungsentwicklungen. In den Ergebnissen der Studie 3 (Abschnitt 4.3) zeigten sich positive Effekte des fachspezifischen Selbstkonzepts bzw. Kursniveaus auf berufliche Interessen im Bereich Conventional. Darüber hinaus war auch der Zusammenhang zwischen beruflichen Interessen und Veränderungen in der Mathematikleistung tendenziell in Richtung Conventional-Orientierung verschoben. Vorangehend wurde argumentiert, dass ein Zusammenhang zwischen bestimmten Interessenorientierungen und Leistungen in einer Metaanalyse bestätigt wurde und deshalb nicht von differenziellen Assoziationen auszugehen sei. Diese Vermutung lässt sich für den Vergleich von allgemein bildenden Gymnasien vs. Fachgymnasien über alle unterschiedlichen Ausrichtungen hinweg weiterhin aufrecht halten. In Hinblick auf die unterschiedlichen Schwerpunkte der Fachgymnasien muss jedoch von spezifischen, in bestimmte Richtungen einwirkenden, Umwelten und entsprechendem Verhalten der Schülerinnen und Schüler ausgegangen werden. Empirisch zeigte sich, dass eine hohe Übereinstimmung bzw. Kongruenz von Person und Umwelt tendenziell mit höherer Zufriedenheit und besseren Leistungen einhergeht (Eder & Reiter, 2002; Spokane, Meir & Catalano, 2000). Entsprechend ist es notwendig, in zukünftigen Studien zu prüfen, ob die Bedeutung beruflicher Interessenorientierungen für Leistungsentwicklungen systematisch zwischen den Profilen der Fachgymnasien variiert.

Zusammenfassung

Wesentliche Ziele des Unterrichts in der gymnasialen Oberstufe sind die Vermittlung einer vertieften Allgemeinbildung und einer allgemeinen Studierfähigkeit. Zum Erreichen dieser Ziele kommt positiven Leistungsentwicklungen in den curricularen Kernfächern, wie Mathematik und Englisch, eine zentrale Bedeutung zu. Für ein Verständnis der Leistungsentwicklungen stellt sich im Rahmen pädagogisch-psychologischer Forschung die Frage nach ihren Bedingungen. In diesem Kontext haben sich Interessen als aussagekräftige Prädiktoren erwiesen. Die Bedeutung beruflicher Interessen für schulische Leistungsentwicklungen ist hingegen bislang weitgehend vernachlässigt worden. In der vorliegenden Arbeit wurde dieses Forschungsdesiderat aufgegriffen.

Die im Folgenden berichteten Erträge basieren auf einer Längsschnittstudie mit Schülerinnen und Schülern zweier Jahrgänge der 11. Jahrgangsstufe an einem Fachgymnasium – zur ersten Erhebung $N = 168$, davon zur zweiten $N = 135$. In einem ersten Teil wurden Leistungsentwicklungen in zwei Kernfächern der gymnasialen Oberstufe untersucht und hierbei erstmalig der Zeitraum nach dem Übergang in die Sekundarstufe II betrachtet. Es zeigten sich fachspezifische Unterschiede in den Entwicklungsverläufen: die mittleren Leistungen in mathematischer Grundbildung nahmen ab, die durchschnittlichen Englischleistungen stiegen an. Weitere Analysen zeigten, dass sich insbesondere bei Schülerinnen und Schülern mit angestrebtem Kurs mit grundlegenden Anforderungen ein Leistungsabfall in mathematischer Grundbildung bereits in der Einführungsphase verzeichnete. In einem zweiten Teil wurde ein methodischer Zwischenschritt vorgenommen: Berufliche Interessen nach Hollands Theorie der Berufswahl wurden in Form von Interessenprofilen modelliert. Hierbei wurde erstmals nachgewiesen, dass berufliche Interessenprofile eine valide Reduktion von Skalenwerten beruflicher Interessenorientierungen darstellen. Hieran anschließend wurde in einem dritten Teil die Bedeutung beruflicher Interessenprofile für Leistungsentwicklungen in Mathematik und Englisch analysiert. Neben der bislang vernachlässigten Untersuchung der Bedeutung beruflicher Interessen für Leistungsveränderungen in der gymnasialen Oberstufe, wurde hierbei ein zweites Forschungsdesiderat aufgegriffen: die Bedeutung von Interessen für Leistungsveränderungen im sprachlichen Bereich. In dieser Arbeit konnte erstmalig aufgezeigt werden, dass berufliche Interessen im Bereich Realistic

(mechanisch-technische Tätigkeiten) positiv mit Leistungsveränderungen in mathematischer Grundbildung assoziiert waren, während sich vergleichsweise negative Veränderungen für Schülerinnen und Schüler mit Interessen in den Bereichen Social (sozialzwischenmenschlich) und Artistic (sprachlich-musisch) erwarten ließen. Darüber hinaus zeigte sich in diesem Fach, dass die positiven Effekte von Selbstkonzept und Kursniveau auf Leistungsentwicklungen durch berufliche Interessen mediiert waren. Berufliche Interessen ließen sich somit hinsichtlich ihrer Bedeutung für Leistungsentwicklungen in einen zu fachlichen Interessen übereinstimmenden Zusammenhang mit Selbstkonzept und Kursniveau bringen. Gleichzeitig wurden fachspezifische Unterschiede in der Rolle beruflicher Interessen für Leistungsentwicklungen deutlich: Ein Zusammenhang mit der Veränderung von Englischleistungen ließ sich nicht nachweisen. In weiteren Studien sollte eine vergleichende Betrachtung von beruflichen und fachlichen Interessen hinsichtlich ihrer Bedeutung für Leistungsentwicklungen erfolgen. Weiterhin bedarf es einer eingehenden Untersuchung fachspezifischer Unterschiede.

Literaturverzeichnis

- Abu-Rabia, S. (2003). Cognitive and social factors affecting Arab students learning English as a third language in Israel. *Educational Psychology, 23*(4), 347–360.
- Ackermann, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence, 22*(2), 227–257.
- Ackerman, P. L. & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*(2), 219–245.
- Alexander, P. A. & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research, 58*(4), 375–404.
- Allison, P. D. (2002). *Missing data: Sage University Papers Series on quantitative applications in the social sciences, 07–136*. Thousand Oaks, CA: Sage.
- Archer, J. & McCarthy, B. (1988). Personal biases in student assessment. *Educational Research, 30*(2), 142–145.
- Batschelet, E. (1981). *Circular statistics in biology*. London: Academic Press.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (2003). *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*. Opladen: Leske + Budrich.
- Baumert, J., Bos, W. & Lehmann, R. (2000). *Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 1: Mathematische und naturwissenschaftliche Bildung am Ende der Pflichtschulzeit*. Opladen: Leske+ Budrich.
- Baumert, J., Bos, W. & Watermann, R. (1999). *TIMSS-III – Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich; Zusammenfassung deskriptiver Ergebnisse* (2. Aufl.). Berlin: Max-Planck-Institut für Bildungsforschung, Studien und Berichte 64.
- Baumert, J., Cortina, K. S. & Leschinsky, A. (2003). Grundlegende Entwicklungen und Strukturprobleme im allgemein bildenden Schulwesen. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (S. 52–147). Reinbek: Rowohlt.

- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O. & Neubrand, J. (1997). *TIMSS. Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen: Leske+ Budrich.
- Baumert, J., Roeder, P. M., Gruehn, S., Heyn, S., Köller, O., Rimmel, R., Schnabel, K. & Seipp, B. (1996). Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). In K.-P. Treumann, G. Neubauer, R. Möller & J. Abel (Hrsg.), *Methoden und Anwendungen empirischer pädagogischer Forschung* (S. 170–180). Münster: Waxmann.
- Baumert, J., Trautwein, U. & Artelt, C. (2003). Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000: Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 261–331). Opladen: Leske+ Budrich.
- Baumert, J., Stanat, P. & Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 15–68). Opladen: Leske+ Budrich.
- Baumert, J. & Watermann, R. (2000). Institutionelle und regionale Variabilität und die Sicherung gemeinsamer Standards in der gymnasialen Oberstufe. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *Dritte internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (Band 2, S. 317–372). Opladen: Leske+ Budrich.
- Becker, M., Lüdtke, O., Trautwein, U. & Baumert, J. (2006). Leistungszuwachs in Mathematik. Evidenz für einen Schereneffekt im mehrgliedrigen System?. *Zeitschrift für Pädagogische Psychologie*, 20(4), 233–242.
- Bergmann, C. & Eder, F. (1992). *Allgemeiner Interessen-Struktur-Test (AIST)/Umwelt-Struktur-Test (UST)*. Göttingen: Beltz.
- Bergmann, C. & Eder, F. (1998). Berufs- und Laufbahnberatung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 34–38). Weinheim: PVU.

- Bergmann, C. & Eder, F. (1999). *Allgemeiner Interessen-Struktur-Test (AIST)/Umwelt-Struktur-Test (UST)* (2. Aufl.). Göttingen: Beltz.
- Bergmann, C. & Eder, F. (2005). *AIST-R. Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test. Revision*. Göttingen: Beltz.
- Brauckmann, S. & Neumann, M. (2004). Berufliche Gymnasien in Baden-Württemberg: Geschichte und heutige Ausgestaltung. In O. Köller, R. Watermann, U. Trautwein & O. Lüdtke (Hrsg.), *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien* (S. 69–111). Opladen: Leske+Budrich.
- Browne, M.W. (1992). Circumplex models for correlation matrices. *Psychometrika*, 57(4), 469–497.
- Bryk, A. S. & Raudenbush, S.W. (1989). Toward a more appropriate conceptualization of research on school-effects: A three-level hierarchical linear model. In R.D. Bock (Ed.), *Multilevel analysis of educational data* (pp. 159–204). San Diego, CA: Academic Press.
- Calsyn, R. & Kenny, D. (1977). Self-concept of ability and perceived evaluations by others: Cause or effect of academic achievement?. *Journal of Educational Psychology*, 69(2), 136–145.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Collins, L.M., Schafer, J.L. & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6(4), 330–351.
- Darcy, M. & Tracey, T.J.G. (2003). Integrating abilities and interests in career choice: Maximal versus typical assessment. *Journal of Career Assessment*, 11(2), 219–237.
- Denissen, J.J.A., Zarret, N.R. & Eccles, J.S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development*, 78(2), 430–447.
- Dittmar, M. (2007). *Zur Feststellung der Eignung von Schülern für das Fachgymnasium Wirtschaft: Entwicklung und Validierung eines Auswahlinstrumentariums*. Dissert-

- tation, Georg-August-Universität Göttingen. Zugriff am 1. Juli 2008, von <http://webdoc.sub.gwdg.de/diss/2007/dittmar/dittmar.pdf>
- Dochy, F. J. R. C. (1992) *Assessment of prior knowledge as a determinant of future learning: The use of knowledge state tests and knowledge profiles*. London: Jessica Kingsley.
- Eccles[-Parsons], J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece J. & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco: Freeman.
- Eccles[-Parsons], J. S., Adler, T. F. & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, 46(1), 26–43.
- Eder, F. (1988). Die Auswirkungen von Person-Umwelt-Kongruenz bei Schülern: Eine Überprüfung des Modells von J. L. Holland. *Zeitschrift für Pädagogische Psychologie*, 2(4), 259–270.
- Eder, F. (1998). Differenziertheit der Interessen als Prädiktor der Interessenentwicklung. In J. Abel & Ch. Tarnai (Hrsg.), *Pädagogisch-psychologische Interessenforschung in Studium und Beruf* (S. 63–77). Münster: Waxmann.
- Eder, F. & Reiter, C. (2002). Interessen und Schullaufbahn. In C. Wallner-Paschon & G. Haider (Hrsg.), *PISA Plus 2000. Thematische Analysen nationaler Projekte* (S. 111–116). Innsbruck: Studien Verlag.
- Educational Testing Service (2001). *Institutional testing program. Manual for supervisors*. Princeton, NJ: Educational Testing Service.
- Fredricks, J. & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology*, 38(4), 519–533.
- Geary, D. C. (1996). Sexual selection and sex differences in mathematical abilities. *Behavioral and Brain Sciences*, 19(2), 229–284.
- Glass, G. V., Peckham, P. D. & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. *Review of Educational Research*, 42(3), 237–288.

- Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, 28(6), 545–579.
- Guilford, J.P., Christensen, P.R., Bond, N.A., Jr. & Sutton, M.A. (1954). A factor analysis study of human interests. *Psychological Monographs*, 68(4 Whole No. 375).
- Gurtman, M.B. & Balakrishnan, J.D. (1998). Circular measurement redux: The analysis and interpretation of interpersonal circle profiles. *Clinical Psychology: Science and Practice*, 5(3), 344–360.
- Gurtman, M.B. & Pincus, A.L. (2003). The circumplex model: Methods and research applications. In I.B. Weiner (Series Ed.) & I.A. Schinka & W.F. Velicer (Vol. Eds.), *Handbook of psychology: Vol. 2. Research methods in psychology* (pp. 407–428). New York: Wiley.
- Guttman, L. (1954). A new approach to factor analysis: The radex. In P.F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258–348). Glencoe, IL: Free Press.
- Hackett, G., Lent, R. W. & Greenhaus, J. H. (1991). Advances in Vocational Theory and Research: A 20-Year Retrospective. *Journal of Vocational Behavior*, 38(1), 3–38.
- Hambleton, R.K. & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.
- Hannover, B. (1998). The development of self-concepts and interests. In J. Baumert, L. Hoffmann, A. Krapp & A. Renninger (Eds.), *Interest and learning* (pp. 105–125). Kiel: Institut der Pädagogik der Naturwissenschaften.
- Healy, C.C. & Mourton, D.L. (1983). Derivatives of the self-directed search: Potential clinical and evaluative uses. *Journal of Vocational Behavior*, 23(3), 318–328.
- Heine, Ch., Spangenberg, H. & Willich, J. (2007). Studienberechtigte 2006 ein halbes Jahr vor Schulabgang Studierbereitschaft und Bedeutung der Hochschulreife. HIS-Forum, Nr. F2/2007, Zugriff am 9. Mai 2008, von http://www.his.de/pdf/pub_fh/fh-200702.pdf
- Helmke, A. & Weinert, F.E. (1997). Bedingungsfaktoren schulischer Leistungen. In F.E. Weinert (Hrsg.), *Enzyklopädie der Psychologie D/I/3. Psychologie des Unterrichts und der Schule* (S. 71–176). Göttingen: Hogrefe.

- Hidi, S. & Ainley, M. (2002). Interest and adolescence. In F. Pajares & T. Urdan (Eds.), *Academic motivation of adolescents* (pp. 247–275). Greenwich, CT: Information Age.
- Hodapp, V. & Mißler, B. (1996). Determinanten der Wahl von Mathematik als Leistungs- bzw. Grundkurs in der 11. Jahrgangsstufe. In R. Schumann-Hengsteler & H.M. Trautner (Hrsg.), *Entwicklung im Jugendalter* (S. 143–164). Göttingen: Hogrefe.
- Holland, J.L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, 6(1), 35–45.
- Holland, J.L. (1985). *Vocational Preference Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Holland, J.L. (1997). *Making vocational choices: A theory of work personalities and work environments*. Odessa, FL: Psychological Assessment Resources.
- Holland, J.L. & Gottfredson, G.D. (1992). Studies of the hexagonal model: An evaluation (or, the perils of stalking the perfect hexagon). *Journal of Vocational Behavior*, 40(2), 158–170.
- Holland, J.L., Gottfredson, G.D. & Baker, H.G. (1990). Validity of vocational aspirations and interest inventories: Extended, replicated, and reinterpreted. *Journal of Counseling Psychology*, 37(3), 337–342.
- Hosenfeld, I., Köller, O. & Baumert, J. (1999). Why sex differences in mathematics achievement disappear in German secondary schools: A reanalysis of the German TIMSS data. *Studies in Educational Evaluation*, 25(2), 143–161.
- Hubert, L. & Arabie, P. (1987). Evaluating order hypotheses within proximity matrices. *Psychological Bulletin*, 102(1), 172–178.
- Ingenkamp, K. (1995). *Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte*. Weinheim: Beltz.
- Jörin, S., Stoll, F., Bergamnn, C. & Eder, F. (2004). *EXPLORIX – das Werkzeug zur Berufswahl und Laufbahnplanung. Deutschsprachige Adaptation und Weiterentwicklung des Self-Directed Search (SDS) nach John Holland (Testmanual)*. Bern, Switzerland: Huber.
- Klieme, E., Baumert, J., Köller, O. & Bos, W. (2000). Mathematische und naturwissenschaftliche Grundbildung: Konzeptuelle Grundlagen und die Erfassung und

- Skalierung von Kompetenzen. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (Band 1, S. 85–133). Opladen: Leske+Budrich.
- Kultusministerkonferenz – KMK (1995). *Weiterentwicklung der Prinzipien der gymnasialen Oberstufe und des Abiturs. Abschlussbericht der von der Kultusministerkonferenz eingesetzten Expertenkommission*. Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.
- KMK – Kultusministerkonferenz (2006). *Vereinbarung zur Gestaltung der gymnasialen Oberstufe in der Sekundarstufe II*. Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.
- Köller, O. & Baumert, J., Cortina, K. S., Trautwein, U. & Watermann, R. (2004). Öffnung von Bildungswegen in der Sekundarstufe II und die Wahrung von Standards. *Zeitschrift für Pädagogik*, 50(5), 679–700.
- Köller, O., Baumert, J. & Schnabel, K. (2000). Zum Zusammenspiel von schulischem Interesse und Lernen im Fach Mathematik: Längsschnittanalysen in den Sekundarstufen I und II. In U. Schiefele & K. Wild (Hrsg.), *Interesse und Lernmotivation* (S. 163–181). Münster: Waxmann.
- Köller, O., Daniels, Z., Schnabel, K. U. & Baumert, J. (2000). Kurswahlen von Mädchen und Jungen im Fach Mathematik: Zur Rolle von fachspezifischem Selbstkonzept und Interesse. *Zeitschrift für Pädagogische Psychologie*, 14(1), 26–37.
- Köller, O., Schnabel, K. U. & Baumert, J. (2000). Der Einfluss der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 32(2), 70–80.
- Köller, O. & Trautwein, U. (2004). Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien. In O. Köller, R. Watermann, U. Trautwein & O. Lüdtke (Hrsg.), *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien* (S. 285–326). Opladen: Leske+Budrich.

- Köller, O., Trautwein, U., Lüdtke, O. & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für Pädagogische Psychologie*, 20(1/2), 27–39.
- Köller, O., Watermann, R., Trautwein, U. & Lüdtke, O. (2004). *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien*. Opladen: Leske+Budrich.
- Krapp, A. (1992). Das Interessenkonstrukt. Bestimmungsmerkmale der Interessenhandlung und des individuellen Interesses aus der Sicht einer Person-Gegenstands-Konzeption. In A. Krapp & M. Prenzel (Hrsg.), *Interesse, Lernen, Leistung: Neuere Ansätze der pädagogisch-psychologischen Interessenforschung* (S. 297–329). Münster: Aschendorff.
- Krapp, A. (1998). Entwicklung und Förderung von Interessen im Unterricht. *Psychologie und Erziehung im Unterricht*, 44(3), 185–201.
- Krapp, A. (2000). Interest and human development during adolescence: An educational-psychological approach. In J. Heckhausen (Ed.), *Motivational psychology of human development* (pp. 109–128). London: Elsevier.
- Krapp, A. (2001). Interesse. In D.H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 286–294). Weinheim: PVU.
- Krapp, A. (2002). Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12(4), 383–409.
- Krapp, A. & Prenzel, M. (1992). *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung*. Münster: Aschendorff.
- Lehmann, R., Hunger, S., Ivanov, S. & Gänsfuß, R. (2004). *LAU 11. Aspekte der Lernausgangslage und der Lernentwicklung Klassenstufe 11*. Zugriff am 2. Februar 2007, von <http://www.hamburger-bildungserver.de/welcome.php?unten=/schulentwicklung/lau/lau11/index.htm>
- Lehmann, R., Vieluf, U., Nikolova R. & Ivanov, S. (2006). *LAU 13. Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 13*. Zugriff am 5. Dezember 2006, von <http://www.hamburger-bildungserver.de/welcome.php?unten=/schulentwicklung/lau/>

- Lent, R. W., Brown, S. D. & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interests, choice, and performance. *Journal of Vocational Behavior*, 45(1), 79–122.
- Lent, R. W., Tracey, T. J. G., Brown, S. D., Soresi, S. & Nota, L. (2006). Development of interests and competency beliefs in Italian adolescents: An exploration of circumplex structure and bidirectional relationships. *Journal of Counseling Psychology*, 53(2), 181–191.
- Lippa, R. (1998). Gender-related individual differences and the structure of vocational interests: The importance of the People–Things dimension. *Journal of Personality and Social Psychology*, 74(4), 996–1009.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Low, D. K., Yoon, M., Roberts, B. W. & Rounds, J. (2005). The stability of interests from early adolescence to middle adulthood: A quantitative review of longitudinal studies. *Psychological Bulletin*, 131(5), 713–737.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, 58(2), 103–117.
- Lüdtke, O. & Trautwein, U. (2004). Die gymnasiale Oberstufe und psychische Ressourcen. In O. Köller, R. Watermann, U. Trautwein & O. Lüdtke (Hrsg.), *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien* (S. 367–401). Opladen: Leske+Budrich.
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23(2), 129–149.
- Marsh, H. W. (1990a). Influence of internal and external frame of reference on the formation of math and English self-concepts. *Journal of Educational Psychology*, 82(1), 107–116.

- Marsh, H.W. (1990b). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology*, 82(4), 646–656.
- Marsh, H.W. (1992). *Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: An interim test manual and a research monograph*. Macarthur, New South Wales, Australia: University of Western Sydney, Faculty of Education.
- Marsh, H.W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76(2), 397–416.
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte: Effekte dimensionaler und sozialer Vergleiche. *Psychologische Rundschau*, 55(1), 19–27.
- Möller, J. & Zaunbauer, A.C.M. (2008). Erwerb fremdsprachlicher Kompetenzen. In M. Hasselhorn & W. Schneider (Hrsg.), *Handbuch Pädagogische Psychologie* (S. 587–596). Göttingen: Hogrefe.
- Mullis, I.V.S., Martin, M.O., Beaton, A.E., Gonzalez, E.J., Kelly, D.L. & Smith, T.A. (1998). *Mathematics and science achievement in the final year of secondary school: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College.
- Mullis, R.L., Mullis, A.K. & Gerwels, D. (1998). Stability of vocational interests among high school students. *Adolescence*, 33, 699–708.
- Muthén, L.K. & Muthén, B.O. (1998–2008). *Mplus* (Version 5.1) [Computer software]. Los Angeles: Muthén & Muthén.
- Muthén, B.O. & Satorra, A. (1995). Complex sample data in structural equation modeling. In P.V. Marsden (Ed.), *Sociological Methodology* (pp. 267–316). Washington, DC: The American Sociological Association.
- Nagy, G. (2006). *Berufliche Interessen, kognitive und fachgebundene Kompetenzen: Ihre Bedeutung für die Studienfachwahl und die Bewährung im Studium*. Dissertation, Freie Universität Berlin. Zugriff am 5. Dezember 2006, von <http://www.diss.fu-berlin.de/2007/109/>

- Nagy, G., Trautwein U. & Lüdtke, O. (in press). The structure of vocational interests in Germany: Different methodologies, different conclusions. *Journal of Vocational Behavior*.
- NCTM – National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: NCTM.
- Niedersächsisches Kultusministerium (1990). *Rahmenrichtlinien für das Gymnasium gymnasiale Oberstufe. Mathematik*. Zugriff am 25. Oktober 2007, von http://db2.nibis.de/1db/cuvo/datei/go_mathe.pdf
- Niedersächsisches Kultusministerium (2000). *Verordnung über die Abschlüsse im Sekundarbereich I*. Zugriff am 8. November 2007, von http://cdl.niedersachsen.de/blob/images/C597887_L20.doc
- Niedersächsisches Kultusministerium (2003). *Rahmenrichtlinien für das Gymnasium gymnasiale Oberstufe. Englisch*. Zugriff am 25. Oktober 2007, von http://www.nibis.de/nli1/gohrgs/rrl/rrlengl_go.pdf
- Niedersächsisches Kultusministerium (2006). *Statistikbroschüre. Die niedersächsischen allgemein bildenden Schulen in Zahlen*. Zugriff am 28. Mai 2008, von http://cdl.niedersachsen.de/blob/images/C23847893_L20.pdf
- OECD – Organisation for Economic Co-operation and Development (2001). *Knowledge and Skills for Life. First Results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris: OECD.
- Pae, T.-I. (2008). Second language orientation and self-determination theory: A structural analysis of the factors affecting second language achievement. *Journal of Language and Social Psychology*, 27(1), 5–27.
- Pekrun, R. (1988). *Emotion, Motivation und Persönlichkeit*. Weinheim: Beltz.
- Prediger, D.J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations?. *Journal of Vocational Behavior*, 21(3), 259–287.
- Prediger, D.J. (1998). Is interest profile level relevant to career counseling?. *Journal of Counseling Psychology*, 45(2), 204–211.
- Prediger, D.J. & Vansickle, T.R. (1992). Locating occupations on Holland's hexagon: Beyond RIASEC. *Journal of Vocational Behavior*, 40(2), 111–128.
- Prenzel, M., Krapp, A. & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik*, 32, 163–173.

- Ramseier, E., Keller, C. & Moser, U. (1999). *Bilanz Bildung. Eine Evaluation am Ende der Sekundarstufe II auf der Grundlage der „Third International Mathematics and Science Study“*. Zürich: Rüegger.
- Randahl, G.J. (1991). A typological analysis of the relations between measured vocational interests and abilities. *Journal of Vocational Behavior*, 38(3), 333–350.
- Reeve, C.L. & Hakel, M.D. (2000). Toward an understanding of adult intellectual development: Investigating within-individual convergence of interest and knowledge profiles. *Journal of Applied Psychology*, 85(6), 897–908.
- Renkl, A. (1996). Vorwissen und Schulleistung. In J. Möller & O. Köller (Hrsg.), *Emotionen, Kognitionen und Schulleistung* (S. 175–190). Weinheim: Beltz.
- Renninger, K.A. (2000). How might the development of individual interest contribute to the conceptualization of intrinsic motivation? In C. Sansone & J.M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 373–404). New York: Academic Press.
- Rolfs, H. (2001). *Berufliche Interessen: Die Passung zwischen Person und Umwelt in Beruf und Studium*. Göttingen: Hogrefe.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2. Auflage). Bern: Huber.
- Rounds, J. & Tracey, T.J. (1993). Prediger's dimensional representation of Holland's RIASEC circumplex. *Journal of Applied Psychology*, 78(6), 875–890.
- Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Savickas, M.L. & Spokane, A.R. (1999). *Vocational Interests: Meaning, Measurement, and Counseling Use*. Palo Alto, CA: Davies-Black.
- Schafer, J.L. (1999). NORM: Multiple imputation of incomplete data under a normal model (Version 2.3) [Computer software]. Verfügbar unter: <http://www.stat.psu.edu/~jls/>
- Schafer, J.L. & Graham, J.W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177.
- Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25(2), 120–148.

- Schwanzer, A. (2002). *Entwicklung und Validierung eines deutschsprachigen Instruments zur Erfassung des Selbstkonzepts junger Erwachsener*. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung Nr. 74.
- Shavelson, R.J., Hubner, J.J. & Stanton, G.C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46, 407–441.
- Spokane, A. R., Meir, E. I. & Catalano, M. (2000). Person–environment congruence and Holland’s theory: A review and reconsideration. *Journal of Vocational Behavior*, 57(2), 137–187.
- Spranger, E. (1966, Erstveröffentlichung 1913). *Lebensformen: Geisteswissenschaftliche Psychologie und Ethik der Persönlichkeit*. Tübingen: Niemeyer.
- Strahan, R. F. (1987). Measures of consistency for Holland-type codes. *Journal of Vocational Behavior*, 31(1), 37–44.
- Strong, E. K., Jr. (1955). *Vocational interests 18 years after college*. Minneapolis, MN: University of Minnesota Press.
- Takei, Y., Johnson, M. P. & Clark, M. E. (1998). Academic achievement and impression management as factors in the grading of white junior high pupils. *Sociological Perspectives*, 41(1), 27–48.
- Taylor, K. F., Kelso, G. I., Longthorp, N. E. & Pattison, P. E. (1980). *Differentiation as a construct in vocational theory and a diagnostic sign in practice*. Melbourne: Department of Psychology, University of Melbourne.
- Tenorth, H.-E. (2001). Englisch: Ein Kerncurriculum, seine Notwendigkeit und seine Gestalt – Zusammenfassung. In H.-E. Tenorth (Hrsg.), *Kerncurriculum Oberstufe. Mathematik – Deutsch – Englisch. Expertisen im Auftrag der Ständigen Konferenz der Kultusminister* (S. 156–161). Weinheim: Beltz.
- Tracey, T. J. G. (1997). RANDALL: A Microsoft FORTRAN program for a randomization test of hypothesized order relations. *Educational and Psychological Measurement*, 57(1), 164–168.
- Tracey, T. J. G. (2000). Analysis of circumplex models. In H. E. A. Tinsley & S. Browne (Eds.), *Handbook of applied multivariate statistics and mathematical modelling* (pp. 641–664). San Diego: Academic Press.
- Tracey, T. J. G. (2001). The development of structure of interests in children: Setting the stage. *Journal of Vocational Behavior*, 59(1), 89–104.

- Tracey, T.J.G. & Robbins, S.B. (2005). Stability of interests across ethnicity and gender: A longitudinal examination of grades 8 through 12. *Journal of Vocational Behavior*, 67(3), 335–364.
- Tracey, T.J.G. & Robbins, S.B. (2006). The interest–major congruence and college success relation: A longitudinal study. *Journal of Vocational Behavior*, 69(1), 64–89.
- Tracey, T.J.G., Robbins, S.B. & Hofsess, C.D. (2005). Stability and change in interests: A longitudinal study of adolescents from grades 8 through 12. *Journal of Vocational Behavior*, 66(1), 1–25.
- Tracey, T.J. & Rounds, J. (1993). Evaluating Holland’s and Gati’s vocational-interest models: A structural meta-analysis. *Psychological Bulletin*, 113(2), 229–246.
- Tracey, T.J.G. & Ward, C.C. (1998). The structure of children’s interests and competence perceptions. *Journal of Counseling Psychology*, 45(3), 290–303.
- Valentine, J.C., DuBois, D.L. & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39(2), 111–133.
- Warm, T.A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.
- Warwas, J., Nagy, G., Watermann, R. & Hasselhorn, M. (in press). The Relations of vocational interests and mathematical literacy: On the predictive power of interest profiles. *Journal of Career Assessment*.
- Warwas, J., Watermann, R. & Hasselhorn, M. (im Druck). Leistungsentwicklungen in Mathematik und Englisch im ersten Jahr an einem Fachgymnasium. *Unterrichtswissenschaft*.
- Warwas, J., Watermann, R., Nagy, G. & Hasselhorn, M. (eingereicht). Zum Effekt von beruflichen Interessenprofilen, Selbstkonzept und Kursniveau auf Leistungsentwicklungen in Mathematik und Englisch.
- Watermann, R. & Klieme, E. (2002). Reporting results of large-scale assessment in psychologically and educationally meaningful terms: Construct validity and proficiency scaling in TIMSS. *European Journal of Psychological Assessment*, 18(3), 190–203.

- Watermann, R., Nagy, G. & Köller, O. (2004). Mathematikleistungen in allgemein bildenden und beruflichen Gymnasien. In O. Köller, R. Watermann, U. Trautwein & O. Lüdtke (Hrsg.), *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien* (S. 205–284). Opladen: Leske+Budrich.
- Wherry, R.J. (1931). A new formula for predicting the shrinkage of the coefficient of multiple correlation. *Annals of Mathematical Statistics*, 2(4), 440-457.
- Wigfield, A. & Eccles, J.S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81.
- Wolter, K.M. (1985). *Introduction to variance estimation*. New York: Springer.
- Wu, M., Adams, R. & Wilson, M. (1998). *ACER ConQuest: Generalised item response modelling software* [Computer software]. Melbourne, Victoria: ACER Press.

Anhang

- Anhang A Übersicht über die TIMSS-Testaufgaben zur mathematischen Grundbildung zur ersten (T1) und zweiten Erhebung (T2)
- Anhang B Beispielaufgaben aus der von Köller und Trautwein (2004) validierten TOEFL-Kurzversion
- Appendix C Prediction Equations for Interest Profiles
- Anhang D SPSS-Syntax zur Berechnung von Interessenprofilen sensu Holland

Anhang A

Übersicht über die TIMSS-Testaufgaben zur mathematischen Grundbildung zur ersten (T1) und zweiten Erhebung (T2)

	T1		T1/T2		T2
Item	TIMSS- Itemparameter	Ankeritem	TIMSS- Itemparameter	Item	TIMSS- Itemparameter
B14	-0.86	B26	0.69/1.82	B19	-0.95
B15	-0.16	C02	-0.72	B23	0.16
B16	-1.34	C07	0.19	B25	0.95
B17	0.35	C12	1.74	C08	-0.52
B18	0.94				
B20	0.34				
B21	1.01				
B22	-0.75				
B24	0.73				
C01	-0.64				
C03	-0.28				
C04	-0.56				
C05	-0.80				
C06	-1.59				
C09	-0.71				
C11	-0.61				
C13	1.91				

Quelle: Macaskill, G., Adams, R.J. & Wu, M.L. (1998). Scaling Methodology and Procedures for the Mathematics and Science Literacy, Advanced Mathematics, and Physics Scales. In M.O. Martin & D.L. Kelly (Eds.), *Third International Mathematics and Science Study, Technical Report. Volume III: Implementation and Analysis. Final Year of Secondary School* (pp. 91–120).

Anhang B

Beispielaufgaben aus der von Köller und Trautwein (2004) validierten TOEFL-Kurzversion

Hörverständnis. (Ausschnitt aus einem transkribierten Text; Quelle: Köller & Trautwein, S. 296): “Welcome to the largest retrospective exhibit this art museum has ever devoted to an architect. The architect chosen for this honor is Frank Lloyd Wright, probably the greatest United States architect of the twentieth century. Wright had the reputation of being arrogant and insensitive to his client’s needs, but his work is based on a set of principles rather than a style. Because of his belief in unity of design and the elimination of unnecessary detail, he resisted his client’s wishes to introduce, what he called, ‘foreign objects’ to his carefully designed interiors.”

Grammatik und Orthographie. (ebd., S. 269):

Vegetables are an excellent source _____ vitamins.

- (A) of
- (B) has
- (C) where
- (D) that

Vokabelwissen und Leseverständnis. (ebd., S. 297): “Impressed by the amazing circulation of the new medium, advertisers of the 1830’s began to take a special interest in the new mass-consumption penny newspapers. They saw that readership of the cheap daily papers cut across political interest, so that the paper reached a broad base of people. These large circulations made it feasible to offer articles for sale that formerly would not have warranted advertising expense. In addition, advertising revenue made it possible for editors and publishers to expand and to experiment with news methods of new gathering.”

The word ‘feasible’ in line 4 is closest in meaning to

- (A) difficult,
- (B) more expensive,
- (C) necessary,
- (D) possible.

Appendix C

Prediction Equations for Interest Profiles

The regression models used in our study belong to a distinct class of regression analyses, namely periodic regression (e.g., Batschelet, 1981). As outlined by Nagy et al. (in press), an individual's *i* Things/People score can be denoted as $\alpha_i \times \cos(\delta_i)$, and his or her Data/Ideas score as $\alpha_i \times \sin(\delta_i)$. In periodic regression, the criterion variable is predicted by a linear combination of *sine* and *cosine* variables.

First of all, we disentangled individual interest orientation δ from differentiation α and tested a simple profile model that considered only individual interest level τ and profile orientation. It can be formalized as follows:

$$(1) y_i = b_0 + b_l \tau_i + b_c \cos(\delta_i) + b_s \sin(\delta_i) + e_i = b_0 + b_l \tau_i + A \cos(\delta_i - \theta) + e_i.$$

where b_0 denotes the regression intercept and thus the mean value of the outcomes across the entire range of the profile orientation δ at a value of $\tau = 0$, b_l is the regression coefficient of τ and captures change in the outcome as a function of individual interest level, b_c is the coefficient of the cosine dimension $\cos(\delta_i)$, b_s is the coefficient of the sine dimension $\sin(\delta_i)$, and e is a regression residual. Equation 1 can be modified by a transformation of the model parameters, such that the derived parameters can be interpreted directly. The peak θ represents the interest dimension that is most positively associated with the given criterion variable. A reflects the effect amplitude and thus the difference in the expected outcome measures between persons with interest orientations equal to the peak θ and the mean value of the outcome across all possible outcomes.

Equation 1 can be extended by the parameter α . Weighting $\sin(\delta_i)$ and $\cos(\delta_i)$ by differentiation α corresponds to the Things/People and Data/Ideas scores:

$$(2) y_i = b_0 + b_l \tau_i + b_c \alpha_i \cos(\delta_i) + b_s \alpha_i \sin(\delta_i) + e_i = b_0 + b_l \tau_i + A \alpha_i \cos(\delta_i - \theta) + e_i.$$

The extended Equation 2 makes it possible to examine the extent to which this weighting improves the prediction of a criterion variable. The parameters of the model above can be interpreted analogously to those of the model in Equation 1. However, the model given in Equation 2 includes profile differentiation as a moderator. As a consequence, A reflects the effect amplitude at a value of $\alpha = 1$.

Anhang D

SPSS-Syntax zur Berechnung von Interessenprofilen sensu Holland¹⁰

Berechnung von Level, Things/People-Dimension (tp) und Data/Ideas-Dimension (di) anhand z-Standardisierter Mittelwerte der RIASEC-Skalen nach Holland (1997).

```
COMPUTE level = mean(rz, iz, az, sz, ez, cz).
COMPUTE tp =   cos(0)*(rz - level_i) +
               cos(1.04719755)*(iz - level) +
               cos(2.09439510)*(az - level) +
               cos(3.14159265)*(sz - level) +
               cos(4.18879020)*(ez - level) +
               cos(5.23598776)*(cz - level).
COMPUTE di =   sin(0)*(rz - level) +
               sin(1.04719755)*(iz - level) +
               sin(2.09439510)*(az - level) +
               sin(3.14159265)*(sz - level) +
               sin(4.18879020)*(ez - level) +
               sin(5.23598776)*(cz - level).
EXE.
```

Berechnung der Parameter Peak (theta) und Amplitude (alpha) über die Regressionskoeffizienten der Things/People- (b_pt) sowie Data/Ideas-Dimension (b_di).

```
COMPUTE theta = atan(b_di/b_pt).
COMPUTE alpha = sqrt((b_di)**2 + (b_pt)**2).
EXE.
```

¹⁰ Diese Syntax wurde freundlicherweise von Gabriel Nagy, MPI Berlin, zur Verfügung gestellt.