

# Die Entwicklung des Interesses an naturwissenschaftlichen Tätigkeiten

Quer- und längsschnittliche Analysen zu Verlauf und Einflussfaktoren der Interessenentwicklung im Chemieunterricht der Sekundarstufen I und II

Dissertation  
zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Lars Höft

Kiel, Februar 2020

1. Gutachterin: Prof. Dr. Ilka Parchmann  
2. Gutachterin: Prof. Dr. Sabine Fechner  
Betreuer: Dr. Sascha Bernholt

Tag der mündlichen Prüfung: 6. April 2020

## Zusammenfassung

Die Etablierung eines Interesses an der Chemie, ihren Inhalten, ihrer Sprache sowie ihren spezifischen Denk- und Arbeitsweisen ist ein zentrales Ziel des Chemieunterrichts. Diese pointierte Stellung des Interesses aufgrund normativer Zielsetzungen ist durch positive Effekte des Interesses auf Lernprozesse und -erfolge sowie Wahlentscheidungen bedingt. Zudem ist das Interesse ein wichtiger Faktor der Entwicklung des Selbstkonzepts, das sich hinsichtlich des Chemieunterrichts idealiter in der Identifikation von Schülerinnen und Schülern mit der Domäne der Chemie äußert. Das Interesse beeinflusst somit nicht nur die schulische Laufbahn, sondern gleichfalls bspw. über Karriereaspirationen die berufliche Vita. Empirische Forschungsartikel zeigen jedoch wiederholt, dass das Interesse der Schülerinnen und Schüler an den Naturwissenschaften mit steigender Jahrgangsstufe fällt. Hinzu kommt, dass die Schülerinnen und Schüler sich überwiegend gegen den Unterricht sowie ein subsequentes Studium bzw. einen Beruf mit naturwissenschaftlichem Schwerpunkt entscheiden.

Eine entscheidende Stellschraube, um Schülerinnen und Schüler für das langfristige, professionelle Engagement in naturwissenschaftlichen Bereichen zu gewinnen, ist die Förderung des Interesses schon während der Schulzeit. Den Bedarf an einer gezielten Interessensförderung griffen Dierks et al. (2014b) und Blankenburg, Höffler und Parchmann (2016) in ihren Forschungsarbeiten unter der Prämisse auf, dass für eine erfolgreiche Steigerung des Interesses differenzierte Kenntnisse über die Interessenprofile der Schülerinnen und Schüler vorliegen müssen. So könnten gezielt vorhandene Interessen angesprochen bzw. neue Interessen hervorgerufen werden. Um die notwendigen differenzierten Einblicke in die Interessenstrukturen der Schülerinnen und Schüler erzielen zu können, wurde das RIASEC+N Interessenmodell entwickelt, das zwischen dem Interesse an Kontexten, Inhalten und naturwissenschaftlichen Tätigkeiten diskriminiert.

Die vorliegende Arbeit knüpft hieran an und widmet sich offen geliebten Fragestellungen nach (a) der Interessenentwicklung im Chemieunterricht der Sekundarstufe I und II, (b) den Zusammenhängen zum Verständnis fundamentaler chemischer Konzepte (Konzeptverständnis) sowie (c) den Zusammenhängen zur Lernfreude und der wahrgenommenen Nützlichkeit. Der Fokus der Forschungsarbeit lag auf der Analyse des Zusammenhangs zwischen dem Interesse und dem Konzeptverständnis, wobei ausgehend von dem generischen Interessenmodell sensu Hidi und Renninger (2006) und dem Erwartungs-Wert-Modell sensu Eccles et al. (2015) von beidseitigen, positiven (i.e., reziproken) Effekten zwischen Interesse und Konzeptverständnis ausgegangen wurde. Zur Beantwortung der Forschungsfragen wurde auf Daten des Forschungsprojekts Development of Learning in Science rekurriert und quer- und längsschnittliche Analysen vorgenommen. Die Ergebnisse indizieren, unabhängig von der fokussierten naturwissenschaftlichen Tätigkeit, eine Abnahme des chemiebezogenen Interesses mit steigender Jahrgangsstufe, wobei sowohl die Ausprägungen des initial gemessenen Interesses als auch der nachfolgenden Interessenabfälle tätigkeitsspezifisch waren. In Hinblick auf die Zusammenhänge zwischen dem Interesse und dem Konzeptverständnis indizieren die Befunde, dass in der Sekundarstufe I keine reziproken Effekte zwischen den Konstrukten vorliegen, in der Oberstufe jedoch das Interesse an Tätigkeiten, die mit kognitiv aktivierenden Prozessen und der Wissenskommunikation assoziiert sind, reziprok mit dem Konzeptverständnis zusammenhängen. Weiterhin konnten, bei Kontrolle des Konzeptverständnisses, positive Effekte der Lernfreude auf das subsequente Interesse sowie des Interesses auf die wahrgenommene Nützlichkeit gezeigt werden.

Diese Befundlage stützt die Notwendigkeit differenzierter Analysen des Interesses, um seine spezifischen Wirkmechanismen aufzudecken und passende Fördermaßnahmen zu konzipieren. Zeitgleich betonen die Ergebnisse die Bedeutung des Interesses für den Wissenserwerb und die wahrgenommene Nützlichkeit für die Gesellschaft und das Individuum. Das Interesse kann als wichtiger Faktor für die Vermittlung von Kenntnissen und für die Entwicklung des Selbstkonzepts angesehen werden und sollte entsprechend bei der Unterrichtsgestaltung miteinbezogen werden.

Auf der anderen Seite zeigen die Ergebnisse, dass noch weitere Forschung notwendig ist, um die Entwicklung des Interesses an naturwissenschaftlichen Tätigkeiten in Abhängigkeit der Quantität und Qualität korrespondierender Lernangebote besser zu verstehen.



## Abstract

Promoting and maintaining students' interest in chemistry, as well as its content, language and the characteristic ways in which knowledge is acquired, is a central goal of chemistry education. The prominent role of interest in education, e.g. in standards and curricula, can be justified by the positive effects of interest on learning processes, achievements and achievement related choices. In addition, interest plays a crucial role in the development of students' self-concept, which, in the case of chemistry education, ideally manifests itself in students' identification with chemistry. Interest does not only influence students' educational trajectory, but also influence their vocational vita.

However, Empirical findings repeatedly indicate a decline of students' interest in science subjects across grade levels. Furthermore, students predominantly decide against taking up classes as well as subsequent studies or a vocation in some form of science.

A determining factor to attract more students to scientific professions is the promotion of interest even during schooling. The need for purposefully fostering students' interest was stated by Dierks et al. (2014b) and Blankenburg, Höffler, and Parchmann (2016), on the assumption that for a significant and sustainable increase of students' interest sophisticated insights into their interest profile are indispensable. In such way, existing interests could be targeted and new interests triggered by teachers. To gain a more precise picture of students' interest, the RIASEC+N model of interest was developed, which discriminates between interest in contexts, contents and students' school science activities.

The present research takes up existing findings and addresses open questions concerning (a) the development of students' interest in chemistry across lower and upper secondary school, (b) the relation between students' interest and their understanding of fundamental concepts (conceptual understanding) and (c) the relation between interest and enjoyment as well as interest and utility values. The main focus of this research lies on the analysis of the relation between interest and conceptual understanding. Based on the generic model of interest development by Hidi and Renninger (2006) and Eccles et al.'s Expectancy-Value Theory (Eccles et al., 2015) bidirectional, positive (i.e., reciprocal) relations between interest and conceptual understanding were hypothesized. In order to answer the research questions, data of the research project Development of Learning in Science was utilized for cross and longitudinal analysis. The results indicated a decrease in chemistry interest with increasing grade level across all seven interest dimensions of the RIASEC+N model whereby initial interest levels and corresponding subsequent decreases varied significantly between interest dimensions. Results of latent cross-lagged models suggested no reciprocal effects between interest and conceptual understanding for lower secondary students. However, for upper secondary students the findings suggested that interests in activities associated with cognitive activation or the communication of knowledge may be reciprocally related. Furthermore, positive effects of students' prior enjoyment on subsequent interest as well as positive effects of interest on utility values were shown.

In summary, the results support the need of fine-grained analysis of students' interest to reveal specific impacts of interest on learning and to design fitting student enrichment programs. At the same time the results underline the importance of interest for students' knowledge acquisition and the development of utility values, i.e., the perceived usefulness for society and oneself. Interest can be considered as an important factor for the transmission of knowledge and skills and the development of one's self-concept and should therefore be considered an important factor in any academic situation. Future research can build on our results to create a better and more complete picture for understanding the development of interest in school science activities depending on the quantity and quality of corresponding learning opportunities.



# Inhaltsverzeichnis

Tabellenverzeichnis	v
Abbildungsverzeichnis	ix
I Theoretische Grundlagen	1
1 Einleitung	3
1.1 Motivation und Hintergrund der Arbeit	3
1.2 Gliederung der Arbeit	5
2 Theoretischer Hintergrund	7
2.1 Interesse als Bildungsziel	8
2.2 Interessentheorien	9
Situationales Interesse	10
Individuelles Interesse	10
Interessenentwicklung	11
Entwicklung schulfachbezogener Interessen	13
Das RIASEC+N Interessenmodell	14
2.3 Kovariaten	16
Konzeptverständnis	16
Lernfreude	17
Wahrgenommene Nützlichkeit	17
2.4 Die Erwartungs-Wert Theorie zur Generierung eines Rahmenmodells	18
2.5 Fazit und Hypothesen	22
Zusammenhänge des Interesses zu Kovariaten	23
3 Studiendesign	25
II Empirischer Teil	29
4 Das RIASEC+N Modell	33
4.1 Konfirmatorische Faktorenanalyse	33
Generische Modelle M1 und M2	33
Generische Modelle M4 bis M6	34
4.2 Deskriptive Statistik	35
Reliabilitäten	36
Intraklassenkorrelationen	36
Mittelwertvergleiche	37
4.3 Latente Profilanalyse	39
4.4 Fazit	42
5 Knowing more about things you care less about	45
5.1 Introduction	45
5.2 Theoretical background	46
Interest	46
Interplay between interest and academic achievement	48

	Conceptual understanding . . . . .	48
5.3	Research Questions . . . . .	49
5.4	Design and Analysis Procedure . . . . .	50
	Instruments . . . . .	50
	Data Collection . . . . .	52
	Data Analysis . . . . .	53
5.5	Results . . . . .	54
	Trends over time . . . . .	54
5.6	Discussion . . . . .	58
	Validity and reliability . . . . .	58
	Development of interest in school science activities . . . . .	58
	Development of conceptual understanding . . . . .	59
	Interplay between interest in school science activities and conceptual understanding . . . . .	59
	Limitations . . . . .	61
	Implications . . . . .	61
6	Longitudinal couplings between interest and conceptual understanding . . . . .	63
6.1	Introduction . . . . .	63
6.2	Theory . . . . .	64
	Academic interest and interest in school science activities . . . . .	64
	Academic achievement . . . . .	65
	Academic interest and academic achievement . . . . .	65
	Interest impacts on achievement . . . . .	66
	Achievement impacts on interest . . . . .	66
6.3	Research aims . . . . .	67
6.4	Method . . . . .	67
	Sample and procedure . . . . .	67
	Measures . . . . .	68
	Data analysis . . . . .	69
6.5	Results . . . . .	71
	Descriptive statistics . . . . .	71
	Latent growth models of interest in school science activities and conceptual understanding . . . . .	71
	Cross-lagged models of interest in school science activities and conceptual understanding . . . . .	71
6.6	Discussion . . . . .	72
6.7	Limitations and future research . . . . .	74
7	Probing Precursors of Students' Long-term Engagement in Chemistry . . . . .	77
7.1	Introduction . . . . .	77
7.2	Background . . . . .	78
	Interest . . . . .	78
	Achievement . . . . .	78
	Interest and Achievement . . . . .	79
	Utility Values . . . . .	79
	The Present Study . . . . .	80
7.3	Method . . . . .	80
	Sample and Procedure . . . . .	80
	Measures . . . . .	80
7.4	Results . . . . .	83
	Preliminary Analyses . . . . .	83
	Bivariate Analysis . . . . .	84



7.5	Discussion . . . . .	85
7.6	Limitations . . . . .	86
7.7	Conclusions . . . . .	87
8	Ich mag, was ich kann oder kann ich, was ich mag? . . . . .	89
8.1	Einleitung . . . . .	90
8.2	Theoretischer Hintergrund . . . . .	90
	Interesse . . . . .	90
	Freude . . . . .	92
	Konzeptwissen . . . . .	93
	Forschungsfragen . . . . .	94
8.3	Methode . . . . .	94
	Durchführung und Stichprobe . . . . .	94
	Instrumente . . . . .	95
	Analysestrategie . . . . .	97
8.4	Ergebnisse . . . . .	100
	Deskriptive Befunde . . . . .	100
	Analyseschritt 1: Univariate Befunde . . . . .	100
	Analyseschritt 2: Bivariate Zusammenhänge . . . . .	102
	Analyseschritt 3: Trivariate Zusammenhänge . . . . .	104
8.5	Diskussion . . . . .	105
8.6	Limitationen . . . . .	107
III Gesamtdiskussion und Schlussfolgerungen . . . . .		109
9	Diskussion . . . . .	111
9.1	Zusammenfassung der Ergebnisse . . . . .	111
	Ergebnisse bezüglich Forschungsfrage F1 . . . . .	111
	Ergebnisse bezüglich Forschungsfrage F2 und F3 . . . . .	113
	Ergebnisse bezüglich Kovariaten . . . . .	114
9.2	Limitationen . . . . .	115
9.3	Implikationen . . . . .	115
	Schule . . . . .	115
	Forschung . . . . .	116
Literatur . . . . .		121
IV Anhang . . . . .		135
A	Das RIASEC+N Modell: Deskriptive Befunde und Validierung . . . . .	137
B	Knowing more about things you care less about . . . . .	141
C	Longitudinal couplings between interest and conceptual understanding . . . . .	145
C.1	Supplemental Materials for Instruments . . . . .	145
	RIASEC+N Questionnaire . . . . .	145
	Achievement test . . . . .	146
C.2	Supplemental Materials for Data analysis . . . . .	149
	Missing data and method effect . . . . .	149
C.3	Supplemental Materials for Results – Latent growth models . . . . .	151
C.4	Supplemental Materials for Results – Cross-lagged models . . . . .	154

D	Probing Precursors of Students' Long-term Engagement in Chemistry	157
D.1	Supplemental Materials for Method – Measurement Invariance . . . . .	157
D.2	Supplemental Materials for Method – Instruments . . . . .	158
D.3	Supplemental Materials for Method – Measurement Invariance . . . . .	162
D.4	Supplemental Materials for Results – Latent Change Score Models . . . . .	164

## Tabellenverzeichnis

3.1	Samplezusammensetzung des DoLiS-Projekts je Messzeitpunkt und Land. Für das deutsche Sample sind in Klammern die Anteile der Schülerinnen und Schüler mit Chemieunterricht angegeben. . . . .	26
4.1	Fit-Statistik der konfirmatorischen Faktorenanalysen. . . . .	35
4.2	Interkorrelationen zwischen den sieben RIASEC+N Interessendimensionen des Sieben-Faktoren-Modells (M3). Alle Korrelationen sind hoch signifikant ( $p < .001$ ). . .	35
4.3	Mittlertsvergleiche zwischen den Ausprägungen der RIASEC+N Interessendimensionen im Querschnitt und Längsschnitt, die bedeutsame Unterschiede indizieren ( $d > .30$ , $p < .05$ ). . . . .	38
4.4	Multinomiale logistische Regression zu Vorhersage der latenten Klasse. Die Ergebnisse für die Referenzklasse CL 1 entsprechen der ersten Spalte (CL 1) und sind daher nicht redundant geführt. Der Regressionskoeffizient ist in log odds angegeben. . . . .	42
5.1	Reliabilities (Cronbach's $\alpha$ , McDonald's $\omega$ , reliability coefficient $\rho$ of the LTRM) and Descriptive Statistics for Interest. M and SD are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each reliability coefficient. . . . .	53
5.2	Average annual gains per grade transition, based on an ANOVA with planned contrasts, and corresponding effect sizes compared to benchmarks from a meta-analysis. . . . .	57
6.1	Dimensions and example items of the RIASEC+N Questionnaire to the question asking „How much are you interested in doing the following activities in chemistry lessons?“ . . . . .	68
7.1	Reliabilities, descriptive statistics, and correlation coefficients within grades. . . . .	83
7.2	Results of the bivariate latent change score models (LCSM) for lower secondary students (LS5) and upper secondary students (US9) including utility values. Displayed are standardized $\beta$ -regression and, correlation coefficients. . . . .	85
8.1	Schrittweise Prüfung der Messinvarianz (MI) über die Zeit für die Dimensionen Freude am Fach und individuelles Interesse (Meredith, 1993) anhand der Änderungen $\Delta$ der Fit-Indizes CFI, RMSEA, SRMR für Samplegrößen $N > 300$ . Metrische MI ist nicht gegeben bei $\Delta CFI \geq .010$ , in Kombination mit $\Delta RMSEA \geq .015$ , $\Delta SRMR \geq .030$ , skalare und strikte MI sind nicht gegeben bei $\Delta CFI \geq .010$ , bei zeitgleichem Auftreten von $\Delta RMSEA \geq .015$ , $\Delta SRMR \geq .010$ (S. 501 Chen, 2007). Da alle berichteten Änderungen des CFI kleiner als .010 sind, kann strikte Messinvarianz angenommen werden. Abkürzungen: IN = individuelles Interesse, ENJ = Freude. . . . .	97
8.2	Deskriptive Befunde, Reliabilität und Korrelationen der analysierten Konstrukte. . .	100
8.3	Mittelwerte (M), Regressionskoeffizienten und Fit-Indizes der univariaten True Individual Change Score und Latent Change Score Modelle (LCSM). Berichtet werden vollständig standardisierte Korrelationen, $\beta$ -Regressionskoeffizienten und in Klammern deren Standardfehler. . . . .	101
8.4	Ergebnisse der bi- (Modelle BI) und trivariaten Latent Change Score Modelle (Modelle TRI). Das Modell TRI 2 basiert auf einem reduzierten Datensatz, aus dem Informationen über Schülerinnen und Schüler mit Chemie als Profil gebendes Fach exkludiert wurden. Berichtet werden vollständig standardisierte Korrelationen, $\beta$ -Regressionskoeffizienten und in Klammern deren Standardfehler. . . . .	103

9.1	Interventionsblock der Anschlussstudie zur Analyse der Entwicklung des Interesses an naturwissenschaftlichen Tätigkeiten (RIASEC+N) durch vier Schülerexperimente im Bereich Säure-Basen. . . . .	118
A.1	Deskriptive Befunde zur Beschreibung des RIASEC+N Konstrukts. Es wurden nur Schülerinnen und Schüler berücksichtigt, die zum Erhebungszeitpunkt den Chemieunterricht nicht abgewählt hatten. Klassen, aus denen $n < 5$ Schülerinnen und Schüler partizipierten, wurden exkludiert. Zur Berechnung der psychometrischen Werte wurden alle durchgehend eingesetzten Items ( $n_{\text{Items}} = 28$ ) genutzt. . . . .	137
A.2	Fit-Statistik für fünf Modelle (M1-M5) mit zunehmender Anzahl Freiheitsgraden und jeweils zunehmender Anzahl an latenten Klassen (CL). . . . .	140
B.1	RIASEC+N Questionnaire: dimensions and items to the question asking „How much are you interested in doing the following activities in chemistry lessons?“ . . . . .	141
B.2	Sample and age distribution across grades and type of course covering chemistry content. . . . .	141
B.3	Fit indices for the multiple-group confirmatory factor analysis (MGCFA) and grade-specific CFAs . . . . .	142
B.4	Goodness-of-Fit Indices for Measurement Invariance . . . . .	142
C.1	Scales and items to the question asking „How much are you interested in doing the following activities in chemistry lessons?“ Items that were excluded in the reported analysis are crossed out. . . . .	145
C.2	Item CRCO2 . . . . .	146
C.3	Item ENCO4 . . . . .	146
C.4	Item MACO4 . . . . .	147
C.5	Item CR064 . . . . .	147
C.6	Item MA092 . . . . .	147
C.7	Item EN113 . . . . .	148
C.8	Reliabilities, descriptive statistics, and correlation coefficients for the RIASEC+N (Realistic, Investigative, Artistic, Social, Enterprising, Conventional, Networking) model and WLE Scores in Grades 9 to 11. . . . .	149
C.9	Goodness-of-Fit indices for testing measurement invariance per dimension of the RIASEC+N model. Model fits were compared by CFI and RMSEA difference testing. Hereby, a change of $< .01$ in CFI supplemented by a change of $< .015$ in RMSEA between subsequent models is generally considered to indicate invariance (G. W. Cheung & Rensvold, 2002). As this criterion was slightly exceeded in the invariance testing of the social dimension ( $\Delta\text{CFI} = .011$ ), McDonald's NCI and gammaHat were calculated. These global fit indices were found to be below recommended thresholds (cf. Meade et al., 2008; West et al., 2012), so that strong measurement invariance was assumed for the social dimension as well. . . . .	150
C.10	Fit indices of the first order latent growth curve models. . . . .	151
C.11	Fit indices of the True Individual Change Models (Steyer et al., 2000). . . . .	151
C.12	Fit indices and parameters of the first order latent growth curve models (nonlinear for Artistic and WLE Score, otherwise linear). The LGCM of the investigative dimension with a freely estimated shape didn't fit the data. The slope variance of the WLE scores was constrained to be non-negative. . . . .	152
C.13	Cross-lagged models for interest in school science activities and WLE scores: Standardized factor loadings, path coefficients, residual variances, and model fit indices. . . . .	155
D.1	Goodness-of-Fit indices for testing measurement invariance over time of the bivariate latent change score models for lower secondary students (LS5) and upper secondary students (US9). . . . .	157

D.2	RIASEC+N Questionnaire. Scales and items to the question asking „How much are you interested in doing the following activities in chemistry lessons?“ Items that were excluded in the reported analysis are crossed out (LS5 = lower secondary students, US9 = upper secondary students). Item R5 was not assessed at the first measurement occasion. . . . .	158
D.3	Achievement test. Exemplary item CRCO2 with coding scheme (in squared brackets behind each answer option) and item fit indices . . . . .	159
D.4	Reliabilities, descriptive statistics, and correlation coefficients within grades for lower secondary school students (LS5). . . . .	160
D.5	Reliabilities, descriptive statistics, and correlation coefficients within grades for lower secondary school students (US9). . . . .	161
D.6	Goodness-of-Fit indices for testing measurement invariance per dimension of the RIASEC+N model for lower secondary students (LS5). . . . .	162
D.7	Goodness-of-Fit indices for testing measurement invariance per dimension of the RIASEC+N model for upper secondary students (US9). . . . .	163
D.8	Results of the bivariate latent change score models (LCSM) for lower secondary students (LS5) including utility values. Displayed are standardized $\beta$ -regression coefficients, correlation and in brackets standard errors (SE). . . . .	165
D.9	Results of the bivariate latent change score models (LCSM) for lower secondary students (LS5) including utility values. Displayed are standardized $\beta$ -regression coefficients, correlation and in brackets standard errors (SE). . . . .	166



## Abbildungsverzeichnis

1.1	Multiple Unterrichtsziele und dazugehörige psychologische Konstrukte nach Kunter (2005). In dieser Arbeit findet eine Fokussierung auf die Interessenentwicklung statt, die Kunter (2005) der Persönlichkeitsbildung zuschreibt. Ferner wird auf der Seite der Vermittlung von Kenntnissen das konzeptuelle Verständnis der Schülerinnen und Schüler untersucht sowie auf der Seite der Persönlichkeitsbildung die Entwicklung der Lernfreude analysiert. Ebenfalls werden die hier indizierten Wechselwirkungen zwischen diesen drei psychologischen Konstrukten ausführlich beschrieben und diskutiert. . . . .	3
2.1	Relationale Struktur des Interessenkonstrukts entsprechend der Person-Gegenstand-Theorie nach Krapp und Prenzel (1992) . . . . .	10
2.2	Utility-Value-Intervention zur Steigerung des Interesses und der Performanz (Hulleman et al., 2010). . . . .	18
2.3	Reduziertes Modell der Erwartungs-Wert-Theorie sensu Eccles et al. (2015). . . . .	19
2.4	Reciprocal Effects Model (REM) für die Modellierung des Zusammenhangs zwischen dem individuellen Interesse und dem Konzeptverständnis. Die angedeutete Mediation des Effekts des Konzeptverständnisses auf das individuelle Interesse wurde in dieser Forschungsarbeit nicht untersucht. . . . .	21
3.1	Änderungen in der mittleren Lesekompetenz, der mathematischen und der naturwissenschaftlichen Kompetenzen schwedischer sowie deutscher Schülerinnen und Schüler in den OECD PISA-Studien von 2006 – 2012 ausgehend von der jeweilig ersten Testung als Untersuchungsschwerpunkt der jeweiligen Erhebung (OECD, 2007, 2011, 2014). . . . .	25
4.1	Darstellung des mittleren Interesses ( $M \pm SE$ ) der Schülerinnen und Schüler im Quer- (grau) und Längsschnitt (schwarz) differenziert je Dimension (R = Realistic, I = Investigative, A = Artistic, S = Social, E = Enterprising, C = Conventional, N = Networking). Es wurden nur die Daten von Schülerinnen und Schülern berücksichtigt, die zu den jeweiligen Erhebungszeitpunkten das Fach Chemie belegten. . . . .	38
4.2	Darstellung des mittleren Interesses ( $M \pm SE$ ) der Schülerinnen und Schüler (SuS) im Längsschnitt differenziert je Dimension (R = Realistic, I = Investigative, A = Artistic, S = Social, E = Enterprising, C = Conventional, N = Networking) und Zeitpunkt der Abwahl des Chemieunterrichts. Es wurden nur die Daten von Schülerinnen und Schülern berücksichtigt, die zu den jeweiligen Erhebungszeitpunkten das Fach Chemie belegten. . . . .	39
4.3	Mittelwertausprägungen der klassischen latenten Profilanalysen. Dargestellt sind die Lösungen für zwei bis sieben latente Klassen. . . . .	41
5.1	Example of an ordered multiple choice (OMC) item for the basic concept „chemical reaction“. Each option of the item corresponds to a specific level of students' conceptual understanding and was scored accordingly based on a partial credit model.	51
5.2	Mean factor scores and corresponding standard errors of students' interest in school science activities across Grades 5 to 11 based on the MG-CFA. Values in Grade 5 (reference group) are set to 0 due to identification reasons. Significant differences between grades are reported for comparisons between grades within two years of each other (* $p < .05$ ; ** $p < .01$ ; *** $p < .001$ ). . . . .	56

5.3	Latent correlations between the seven dimensions of students' interest in school science activities according to the RIASEC+N model (MG-CFA) across Grades 5 to 11 (R = Realistic, I = Investigative, A = Artistic, S = Social, E = Enterprising, C = Conventional, N = Networking). All correlation coefficients are significantly different from 0, with the exception of the correlation coefficients between the artistic and social as well as between the artistic and networking dimensions in Grade 10. . . . .	56
5.4	Latent correlations between students' conceptual understanding (WLE parameter of the MG-GPCM) and students' interest in school science activities according to the RIASEC+N model (MIMIC model). Correlation coefficients were tested to be significantly different from 0 (* $p < .05$ ; ** $p < .01$ ; *** $p < .001$ ). . . . .	57
6.1	Bivariate Cross-lagged Model used to analyse the relation between students' interest in school science activities and their test scores while controlling for students' sex, age and cognitive abilities (KFT). . . . .	71
7.1	Bivariate latent change score model for lower and upper secondary students. . . . .	82
8.1	Generisches bivariates Latent Change Score Modell der latenten Variablen X und Y für die drei Messzeitpunkte T1 bis T3; Latente Differenzwerte zwischen zwei Messzeitpunkten sind durch $\Delta$ s gekennzeichnet; Regressionspfade, die kein Label aufweisen, sind auf den Wert 1 fixiert; Entsprechend Grimm et al. (2012) wurden autoregressive Pfade zwischen latenten Differenzwerten modelliert. . . . .	99
8.2	Darstellung der mittels True Individual Change Score Modellen berechneten Personenparameter des individuellen Interesses, der Freude und des Konzeptwissens sowie deren Mittelwertsverläufe mit Standardabweichung. . . . .	101
9.1	Abstraktion der empirischen Befunde zum Modell der reziproken Effekte. Zusammenhänge innerhalb von Interessendimensionen, die mit einem Asterisk * markiert sind, wurden nur bedingt signifikant. . . . .	113
9.2	Studiendesign der Anschlussstudie bestehend aus zwei Interventionsblöcken sowie der Erfassung von Lernangeboten. . . . .	117
9.3	Strukturgleichungsmodell zur Analyse des Einflusses des situationalen Interesses an networking und realistic Tätigkeiten auf das individuelle Interesse. . . . .	119
B.1	RIASEC+N model for interest in science activities. The CFA model was estimated for the whole sample (without grouping of grades). . . . .	143
B.2	Scatter plot of students' WLE parameters across Grades 5 to 11, including mean values and standard deviations per grade and significant differences in mean values between adjacent grades (* $p < .05$ ; ** $p < .01$ ; *** $p < .001$ ). Only students attending chemistry courses are included. . . . .	143
C.1	Mean factor scores and corresponding standard deviations of students' interest in school science activities (Realistic, Investigative, Artistic, Social, Enterprising, Conventional, Networking) and conceptual understanding (WLE Score) across Grades 9 to 11 based on the True Individual Change Model (for interest variables; Steyer et al., 2000) respectively manifest means (for the WLE score). Significant differences between grades are reported for comparisons between grades within two years of each other (* $p < .05$ ; ** $p < .01$ ; *** $p < .001$ ). . . . .	153



# Teil I

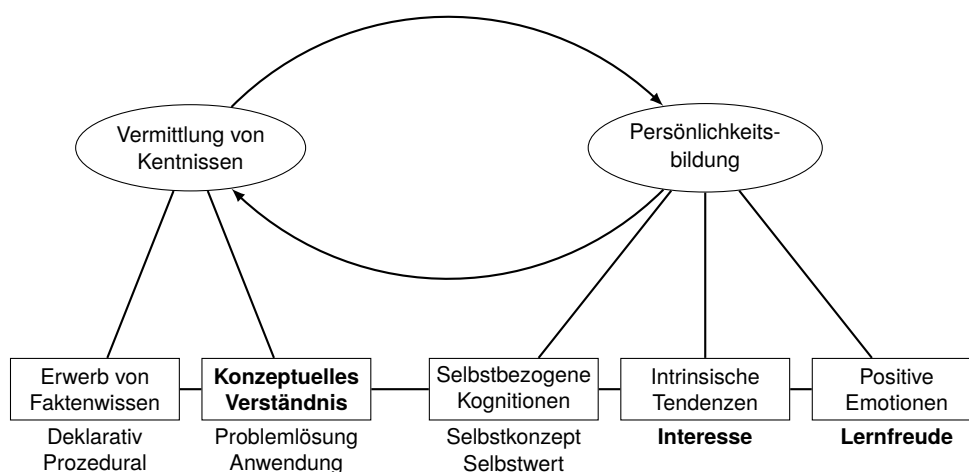
## Theoretische Grundlagen



## Einleitung

### 1.1 Motivation und Hintergrund der Arbeit

Die vorliegende Forschungsarbeit widmet sich der Analyse der Interessenentwicklung von Schülerinnen und Schülern der gymnasialen Sekundarstufen im Kontext des Chemieunterrichts, wobei ein besonderes Augenmerk auf dem Zusammenspiel zwischen dem Interesse und dem Konzeptverständnis liegt. Die Forschungsarbeit konzentriert sich somit auf Aspekte zweier wesentlicher Ziele des Schulunterrichts: (a) das fachliche Lernen im Sinne der Vermittlung eines sophistizierten und anschlussfähigen Fachwissens (Vermittlung von Kenntnissen) sowie (b) die motivational-emotionale Entwicklung (Persönlichkeitsbildung) im Sinne der Interessenentwicklung (s. Kunter, 2005, S. 12–17 und Abbildung 1.1). Den Rahmen dieser Arbeit bildet das Projekt Development of Learning in Science (DoLiS), das die Entwicklung kognitiver, motivationaler und affektiver Personenmerkmale fokussiert. Die hier gewonnenen Erkenntnisse sollen zum Verständnis der Interessenentwicklung und -ausdifferenzierung im Fach Chemie beitragen.



**Abb. 1.1:** Multiple Unterrichtsziele und dazugehörige psychologische Konstrukte nach Kunter (2005). In dieser Arbeit findet eine Fokussierung auf die Interessenentwicklung statt, die Kunter (2005) der Persönlichkeitsbildung zuschreibt. Ferner wird auf der Seite der Vermittlung von Kenntnissen das konzeptuelle Verständnis der Schülerinnen und Schüler untersucht sowie auf der Seite der Persönlichkeitsbildung die Entwicklung der Lernfreude analysiert. Ebenfalls werden die hier indizierten Wechselwirkungen zwischen diesen drei psychologischen Konstrukten ausführlich beschrieben und diskutiert.

Die theoretischen Grundlagen der Interessenforschung, die den Rahmen dieser Forschungsarbeit bilden, entstammen der Motivationspsychologie, wobei vorwiegend auf Theorien aus schulpädagogischen Kontexten rekurriert wird. Als zentral kann die Person-Gegenstands-Theorie (person-object theory of interest, POI), die maßgeblich von Krapp entwickelt wurde (bspw. Krapp, 1999a, 2002a, 2002b; Prenzel et al., 1986), sowie das Vier-Phasen-Modell der Interessenentwicklung (Hidi & Renninger, 2006) angesehen werden. Allgemein erfahren und erfuhren Studien zum Interesse, u. a. im Zuge der Beforschung der intrinsischen Motivation (bspw. Prenzel et al., 1986; Renninger et al., 1992), über verschiedene Forschungsdisziplinen eine hohe Aufmerksamkeit (bspw. Su, 2018). Das Interesse kann als eine Komponente der intrinsischen Motivation aufgefasst (OECD, 2016; Ryan & Deci, 2000) und genutzt werden, um Verhalten bei Absenz von externen Anreizen zu erklären. Ein Grund für die intensive Beforschung des Interesses ist, dass auch in Lern- und Leistungssituationen eine positive Wirkung des Interesses auf Tätigkeitsergebnisse belegt werden konnte. Entsprechend konnten im schulischen Kontext positive Zusammenhänge zwischen den Interessen und Kompetenzen der Schülerinnen und Schüler empirisch abgesichert werden (z. B. Schiefele et al., 1992).

Dies ist insbesondere von Bedeutung, da aufgrund positiver kognitiv-affektiver Erfahrungen die interessengeleitete Ausübung einer Tätigkeit selbst als Belohnung wahrgenommen wird (Renninger & Hidi, 2011). Vor diesem Hintergrund wird das Interesse im Schulkontext als besonders förderungswürdig und der interessengeleitete Wissenserwerb als erstrebenswert angesehen. Die positive Beurteilung des Interesses schlägt sich in den Formulierungen von Kompetenzdefinitionen (bspw. Weinert, 2001), Bildungsstandards (KMK, 2004), Lehrplänen und Fachanforderungen nieder (bspw. Kunter, 2005; BIMl, 2016; KMK, 2010). In der Forschung wird neben diesen normativen Zielstellungen ebenso der prädiktive Charakter des Interesses für bildungsbezogene Wahlentscheidungen des Individuums betont, die Ausbildung der Identität als akzentuierender Grund zur Untersuchung der Entwicklung des Interesses hervorgehoben und im Zusammenhang hierzu das lebenslange, selbstbestimmte Lernen sowie das Wohlbefinden diskutiert (bspw. Krapp, 2000). In Hinblick auf den beruflichen Werdegang von Individuen sei hier das Modell der beruflichen Interessen nach Holland (1997) genannt, das ein Streben nach Kongruenz zwischen der Präferenz für bestimmte Tätigkeiten und den Berufsanforderungen postuliert, welche maßgeblich das Wohlbefinden und die berufliche Performanz des Individuums mitbestimmt (z. B. Etzel & Nagy, 2015). Wie skizziert sind die Entwicklung und die Ausdifferenzierung der akademischen und beruflichen Interessen sowie die Ausbildung von Berufsaspirationen wichtige Aufgaben der Persönlichkeitsbildung während der Schulzeit. Gleichzeitig sind diese Personenmerkmale aber auch bestimmend für den zukünftigen Werdegang der Schülerinnen und Schüler (z. B. Schoon, 2001; Stoll et al., 2017).

Eine Analyse des derzeitigen Forschungsstands deutet auf einen Bedarf an längsschnittlichen Studien hin, die die Interessenentwicklung sowie das Zusammenwirken von Interesse und Fachwissen über ausgedehntere Zeitspannen analysieren. Hier soll die vorgenommene Forschung anknüpfen und zusätzliche empirische Evidenz liefern. Hierzu wird das Fach Chemie fokussiert und als Zeitspanne die gesamte gymnasiale Sekundarstufe von Jahrgangsstufe 5 bis 12 in den Fokus betrachtet. Ausgehend von diesen Überlegungen konzentriert sich das Forschungsvorhaben auf drei zentrale Fragestellungen:

- F1. Wie entwickeln sich die individuellen Interessen der Schülerinnen und Schüler im Fach Chemie über die Sekundarstufen I und II?
- F2. Wirkt sich ein elaboriertes Konzeptverständnis positiv auf die Entwicklung des individuellen Interesses aus?
- F3. Wirkt sich das individuelle Interesse positiv auf den Erwerb eines elaborierten Konzeptverständnisses aus?

Die Diskussion dieser drei Hauptfragestellungen bildet den Kern der vorliegenden Forschungsarbeit und zieht sich als roter Faden durch die folgenden Abschnitte. Neu für die Analyse dieser drei Kernfragen gegenüber bisherigen Studien ist, dass für die längsschnittliche Untersuchung des Interesses durch die Operationalisierung des RIASEC+N Interessenmodells ein differenzierteres Modell (Dierks et al., 2014b) genutzt wurde, das das Potential für einen nuancierten Einblick in die Interessenentwicklung und bei dieser involvierte Wirkmechanismen offeriert.

Das RIASEC+N Modell lehnt sich an die Konzeptualisierung des Modells der beruflichen Interessen (RIASEC Modell) sensu Holland (1997) an, richtet den Blick jedoch auf Schülerinnen und Schüler und ihr Interesse an prototypischen Tätigkeiten des naturwissenschaftlichen Unterrichts. Korrespondierend zu den naturwissenschaftlichen Unterrichtstätigkeiten nimmt es eine Facettierung des Interesses in sieben Dimensionen vor, die analog zu den Berufsinteressen des RIASEC Modells sensu Holland (1997) benannt sind, jedoch einen wesentlich engeren Fokus vornehmen, da sich alle Interessendimensionen und Tätigkeiten hier auf die Domänen der Naturwissenschaften beziehen. Die Bezeichnung der beiden Interessenmodelle (RIASEC und RIASEC+N) sind Akronyme der Bezeichnungen der jeweils konzeptualisierten sechs bzw. sieben Interessendimensionen. Das RIASEC+N Modell enthält drei Interessendimensionen, die Tätigkeiten der Sozialformen Partner- und Gruppenarbeit fokussieren (S, E, N) und zwei Interessendimensionen, die das Experimentieren in den Vordergrund stellen (R, I). Die beiden weiteren Dimensionen greifen das Interesse an kreativ-künstlerischen (A) bzw. ordnend-dokumentierenden (C) Tätigkeiten auf, die häufig weniger im Vor-

dergrund der Unterrichtsgestaltung stehen (Eine umfassendere Diskussion des RIASEC+N Modells findet sich in Kapitel 2.).

Die gewählte Analysestrategie dieser Forschungsarbeit trug dem Studiendesign des DoLiS-Projekts (kombinierte Querschnitt- und 2-Kohorten-Längsschnittstudie) Rechnung. Zunächst wurde als Basis für die späteren Analysen die Querschnitterhebung in einer ersten Studie ausgewertet. In dieser ersten Studie wurden querschnittliche Trends in den Interessenausprägungen je Dimensionen des RIASEC+N Modells [F1], in dem mittleren Konzeptverständnis und in den korrelativen Zusammenhängen [F2, F3] zwischen den beiden Konstrukten analysiert. Die Ergebnisse deuten auf eine Interessenabnahme, einen Anstieg des Konzeptverständnisses sowie auf korrelativen Zusammenhänge zwischen dem Interesse und dem Konzeptverständnis hin. Die Erkenntnisse aus der ersten Studie wurden zur Generierung der Forschungsfragen und der Hypothesenbildung für die anschließende Auswertung der Längsschnitterhebung in zwei Studien verwandt. In der zweiten und in der dritten Studie wurde die längsschnittliche Entwicklung des Interesses [F1] und des Konzeptverständnisses sowie Zusammenhänge zwischen den Konstrukten [F2, F3] untersucht, welche die querschnittlich gefundenen Trends der Interessenabnahme, des Anstiegs des Konzeptverständnisses sowie die stärkere Bedeutung der Wechselwirkungen zwischen Interesse und Konzeptverständnis in der Sekundarstufe II stützen. Abschließend wurde in einer vierten Studie die Entwicklung des individuellen Interesses an der Chemie [F1], der Freude im Chemieunterricht und dem Konzeptverständnis in der Sekundarstufe II sowie Wechselwirkungen zwischen den Konstrukten [F2, F3] analysiert, wobei die Befunde zusätzliche Evidenz für die Relevanz emotional-motivationaler Variablen für das Lernen und für das Potential der Unterrichtsgestaltung für die Interessenförderung liefern.

## 1.2 Gliederung der Arbeit

Die vorliegende Arbeit gliedert sich in drei wesentliche Bereiche: In dem ersten Abschnitt Theoretische Grundlagen werden wesentliche Interessentheorien vorgestellt, der RIASEC+N Ansatz in diesen verankert und anschließend Modelle zur Verknüpfung des Interesses mit den Kovariaten Konzeptverständnis, Lernfreude und wahrgenommene Nützlichkeit beschrieben, wobei der Fokus auf dem Zusammenwirken zwischen Interesse und Konzeptverständnis liegt, für das ein Modell der reziproken Zusammenhänge abgeleitet wird. Ausgehend von den Modellbeschreibungen werden Schlussfolgerungen für die vorliegende Forschungsarbeit konjiziert. Zuletzt wird das DoLiS-Projekt umfassender vorgestellt und das implementierte Studiendesign skizziert. In dem zweiten Abschnitt Empirischer Teil werden deskriptive Befunde zum Interessenmodell sowie Validitätsanalysen präsentiert. Konkret wird die Konstruktvalidität anhand von konfirmatorischen Faktorenanalysen geprüft und in einem personenzentrierten Ansatz auf Interessenprofile geschlossen. Daraufhin werden die vier Fachpublikationen in oben beschriebener Reihenfolge eingebunden. In dem dritten Abschnitt Gesamtdiskussion und Schlussfolgerungen erfolgt schließlich die zusammenfassende Diskussion der Ergebnisse der Forschungsarbeit. Hier werden unter Rückgriff auf die Theorie die Befunde der vier Publikationen zusammengeführt und synthetisiert, wobei Stärken und Schwächen der Studie aufgezeigt und resultierende Forschungsdesiderate konstruktiv abgeleitet werden.



## Theoretischer Hintergrund

In diesem Kapitel wird der theoretische Hintergrund der Forschungsarbeit dargelegt. Der inhaltliche Schwerpunkt des Kapitels liegt entsprechend auf der Diskussion von Interessenkonzeptionen, im Zuge derer zudem Verknüpfungen zu den Konstrukten Konzeptverständnis, wahrgenommene Nützlichkeit und Lernfreude vollzogen werden. Zunächst wird hierzu das RIASEC+N Interessenmodell sensu Dierks et al. (2014b) in der Interessenforschung verankert und in generelleren, motivationalen Rahmenmodellen der pädagogischen Psychologie verortet. Vorweg sei darauf hingewiesen, dass in der pädagogisch-psychologischen Forschung nicht die eine Interessenkonzeption existiert, sondern eine Vielzahl an Interessenkonzepten, -konstrukten und -modellen vorliegt (bspw. Renninger & Hidi, 2011). Umfassendere Aufarbeitungen der Historie der Interessenforschung, Übersichten über bisherige Konzepte und Ansätze der Zusammenführung und Synthese von Modellen können bspw. in Z. Daniels (2008), Krapp (1992b), Krapp und Prenzel (2011), Reeve et al. (2015), Renninger und Hidi (2011) und Schiefele (2009) oder Su (2018) nachgelesen werden. In dem folgenden Abschnitt wird hingegen eine Reduktion auf für die vorliegenden Forschungsfragen relevanten Konzepte und Modelle vorgenommen. Diese sind vornehmlich die Person-Gegenstands-Theorie sensu Krapp (1992a) sowie das Vier-Phasen-Modell der Interessenentwicklung sensu Hidi und Renninger (2006), die daher beide expliziert werden, wobei die Begriffe situationales Interesse, individuelles Interesse definiert und voneinander abgegrenzt werden. Beide Konzeptionen weisen große konzeptuelle Überlappungen auf und sind daher gut geeignet, um einen Arbeitsbegriff des Interesses zu deduzieren.<sup>1</sup> Zudem werden in diesem Abschnitt theoretische Annahmen und empirische Befunde zur Entwicklung der schulfachbezogenen Interessen beschrieben und somit die Grundlage für die Forschungsfrage F1 gelegt:

- F1. Wie entwickeln sich die individuellen Interessen der Schülerinnen und Schüler im Fach Chemie über die Sekundarstufen I und II?

Die Person-Gegenstands-Theorie und das Vier-Phasen-Modell der Interessenentwicklung bilden außerdem die theoretische Fundierung des RIASEC+N Modells sensu Dierks et al. (2014b), welches im Anschluss an diesen Abschnitt ausführlich beschrieben wird. Daraufhin wird mithilfe des erweiterten Erwartungs-Wert-Modells sensu Eccles (bspw. Eccles et al., 2015; Wigfield & Eccles, 2000) der bisherige Blickwinkel vergrößert, indem der Fokus von den „genuinen“ Interessenkonzeptionen auf Mechanismen des Zusammenwirkens von Personenmerkmalen und Umweltfaktoren für die Erklärung von Handeln und Leistung gelenkt wird. In diesem Abschnitt wird weiterhin ein Modell der reziproken Zusammenhänge zwischen Interesse und Konzeptverständnis aus der Literatur abgeleitet und die Hypothesen für die beiden Forschungsfragen F2 und F3 aufgestellt:

- F2. Wirkt sich ein elaboriertes Konzeptverständnis positiv auf die Entwicklung des individuellen Interesses aus?
- F3. Wirkt sich das individuelle Interesse positiv auf den Erwerb eines elaborierten Konzeptverständnisses aus?

Der theoretische Rahmen kann somit Orientierung für die folgenden Arbeiten bieten, indem eine Verknüpfung der Forschungsfragen mit dem bisherigen Forschungsstand vollzogen wird und Beziehungen zu den drei Kovariaten diskutiert werden.

---

1. Laut Hidi und Renninger (2006) erweitert und verfeinert das Vier-Phasen-Modell der Interessenentwicklung die Interessenkonzeption sensu Krapp (2002b).

## 2.1 Interesse als Bildungsziel

Das Konstrukt Interesse dient dem Verständnis, warum bzw. in welchem Maße Schülerinnen und Schüler sich mit einem Gegenstand, bspw. chemischen Fachinhalten, auseinandersetzen, auch wenn keine direkten externen Anreize bspw. in Form von Noten oder Klausuren vorliegen. So ist das Interesse ein Grund, aus dem Schülerinnen und Schülern das Lernen Spaß machen kann und sie es als bedeutsam empfinden können (OECD, 2016). Es unterstützt direkt (bspw. aufgrund erhöhter Aufmerksamkeit) und indirekt (bspw. aufgrund der Wahl der Lernaufgabe) den Prozess des Kompetenzerwerbs. In diesem Sinne kann das erlebte Interesse, Krapp spricht von dem Interessiertsein (Krapp & Prenzel, 1992; Prenzel et al., 1986), während des Unterrichts die Intensität der Auseinandersetzung der Schülerinnen und Schüler mit Unterrichtsaufgaben, wie der Erstellung von Plakaten oder dem Ausarbeiten von Kurzreferaten, beeinflussen und die Qualität und Quantität der Bearbeitungen prägen. Gleichfalls beeinflusst das Interesse die Aufnahme und Ausführung außerschulischer Tätigkeiten, wie z. B. das selbstbestimmte Lesen oder das Wiederholen von Lehrbuchinhalten, die Recherche von Internetartikeln, das Sehen von Videos mit naturwissenschaftlichen Inhalten oder das Besuchen von Naturkundemuseen. Aber auch für die Erklärung von Wahlentscheidungen, die innerhalb des Schulsystems getroffen werden, wie die Wahl der weiterführenden Schule, die An- und Abwahl von Schulfächern, die Wahl des Betriebs für das Berufspraktikum sowie Wahlentscheidungen, die prägend für das berufliche Leben sind, die Wahl des Studiengangs oder des Arbeitsorts, wird das Interesse herangezogen. Demzufolge kann Interesse als ein Personenmerkmal angesehen werden, das einerseits das Verhalten der Schülerinnen und Schüler im Schulunterricht beeinflusst und somit ihre soziale Umwelt mitgestaltet. Andererseits wirkt die Schule als spezifisches soziales Umfeld ebenfalls auf die Interessenentwicklung der Schülerinnen und Schüler ein. Dies zeigt sich auf der einen Seite darin, dass auf bildungsadministrativer Ebene die Strukturierung des Unterrichts in Fächer zur Ausbildung domänenspezifischer Interessen beiträgt (bspw. Jansen et al., 2019; Schiefele, 2009) und auf Fachebene der Unterricht spezifische Interessen an bestimmten Fachinhalten und Tätigkeiten fördern oder zu einem Desinteresse an diesen führen kann, Lehrerinnen und Lehrer aber auch Peers bei diesem Prozess als Rollenvorbilder fungieren und das Interesse entfachen können.

Als implizites Ziel der Schulbildung sieht Kunter (2005) die emotional-motivationale Persönlichkeitsentwicklung an (vgl. Ziel des lebenslangen Lernens bspw. in Krapp, 2000). Dieses Ziel kann als solches ebenfalls in den während der Durchführung des DoLiS-Projekts gültigen Bildungsstandards und Lehrplänen<sup>2</sup> gefunden werden. In Kompetenzdefinitionen lässt sich das Interesse in der Facette Motivation situieren und findet somit auch Niederschlag in den Bildungsstandards (KMK, 2004). Weinert (2001, S. 27) definiert Kompetenzen als die „bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.“

Explizit wird gefordert, dass auf Grundlage der Bildungsstandards das Interesse der Schülerinnen und Schüler gezielt unterstützt werden soll (KMK, 2010, S. 10). Ebenso wird die Förderung der Interessenentwicklung in den Lehrplänen und Fachanforderungen des Fachs Chemie berücksichtigt. Hier stellt die Interessenorientierung eine Facette des Verständnisses von Grundbildung dar, dessen Ziel es sei, Schülerinnen und Schüler zur Entfaltung individueller „Begabungen und Neigungen zu verhelfen“ (BIMI, 2005, S. 5). Die hieran aufbauende vertiefte Allgemeinbildung „zielt ab auf die vielseitige Entwicklung von Interessen und Fähigkeiten“ (BIMI, 2005, S. 9). Darüberhinausgehend sollen Schülerinnen und Schüler im Sinne der Studier- und Berufsfähigkeit dazu befähigt werden, eigene „Interessen im Hinblick auf die Berufswahl und die Berufsausübung zu erkennen, zu prüfen und zu artikulieren“ (BIMI, 2002, S. 4). Auch in den seit 2018 geltenden Fachanforderungen findet sich das Interesse in einer kompetenzorientierten Formulierung wieder. Die hier als Teil der Lernkompetenz spezifizierte Selbstkompetenz, die die Schülerinnen und Schüler im Zuge ihres Fachunterrichts erwerben sollen, beinhaltet die Artikulation und Differenzierung sowie die kritische Selbstreflexion

2. Die derzeit gültigen Lehrpläne werden in Schleswig-Holstein als Fachanforderungen bezeichnet (BIMI, 2016).



von Interessen (BIMI, 2016, S. 8).

## 2.2 Interessentheorien

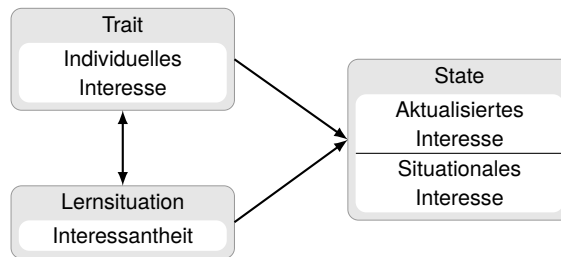
Die gegenwärtig etablierten pädagogisch-psychologischen Interessentheorien konzeptualisieren Interesse als eine motivationale Variable, die eine affektive und eine kognitive Komponente aufweist, deren Bewertung (gleichnamig positive Ausprägung) im Falle eines Interesses übereinstimmt, deren Wirken jedoch unabhängig voneinander ist (Hidi & Renninger, 2006; Krapp & Prenzel, 2011; Renninger & Hidi, 2011). Das Interesse einer Person ist immer auf einen spezifischen Gegenstand bzw. Inhalt gerichtet. Diese Gegenstands- bzw. Inhaltsspezifität ermöglicht die Abgrenzung des Interesses von anderen motivationalen Variablen (Hidi & Renninger, 2006; Krapp & Prenzel, 2011; Renninger & Hidi, 2011; Schiefele, 2009) und bedingt, dass die Aussage, dass eine Person interessiert sei, sich nur sinnstiftend interpretieren lässt, wenn zugleich die Person und der Interessengegenstand benannt werden (bspw. M. Ainley, 2017). Die Gegenstands- bzw. Inhaltsspezifität des Interesses wird in der Person-Gegenstands-Theorie und dem Vier-Phasen-Modell der Interessenentwicklung akzentuiert, die das Interesse als Relation zwischen einer Person und einem Gegenstand bzw. Inhalt definieren (s. u. a. Krapp, 1992a, 2002a, 2002b, 2005; Krapp & Prenzel, 2011; Prenzel et al., 1986). Der Gegenstand bzw. Inhalt wird konstruktivistisch als subjektiv bestimmter „Umweltausschnitt“ interpretiert (Krapp, 1992c), der eine „strukturierte Einheit“ in dem mentalen Repräsentationssystem einer Person bildet, welche infolge der wiederholten Auseinandersetzung mit dem Gegenstand einem permanenten Rekonstruktionsprozess unterliegt.

Merkmale des Interesses bzw. der Relation zwischen Gegenstand und Person sind entsprechend der Person-Gegenstand-Theorie eine epistemisch-kognitive (Bestreben des Erwerbs von Fähigkeiten und Fertigkeiten), eine emotionale (positive Erlebensqualitäten) und eine wertbezogene Facette (intrinsische Bedeutsamkeit, Wichtigkeit) (Krapp & Prenzel, 2011; Prenzel et al., 1986). Schiefele und Krapp (Krapp, 2002b; Schiefele, 2009; Schiefele & Krapp, 1996) sprechen in diesem Zusammenhang auch von einer gefühls- und einer wertbezogenen Valenz des Interesses, wobei Valenzen mentale Relationen zwischen dem Interessengegenstand und evaluativen Attributen bezeichnen.<sup>3</sup> Neben den gefühls- und wertbezogenen Valenzen des Interesses wird zudem der Stellenwert von differenzierten Wissensstrukturen über den Interessengegenstand (Hidi & Renninger, 2006), prozeduralen und meta-kognitiven Wissen (Krapp, 1992c; Schiefele, 2009) betont, welches während der wiederholten Auseinandersetzung mit diesem erworben wird. Laut Krapp und Prenzel (2011) verfügen interessierte Personen zudem über das Wissen, „of what else there is to know and to explore in ‚the zone of proximal development‘ (Vygotskij & Cole, 1978)“ (Krapp und Prenzel, 2011, S. 31).

Aufgrund der Stabilität des Interesses wird zwischen dem Interesse als einem psychischen Zustand (state-ähnlich, Prozessgeschehen) und einem relativ stabilen Personenmerkmal (trait-ähnliche Disposition bzw. Prädisposition) unterschieden (s. Hidi und Renninger, 2006; Krapp und Prenzel, 2011; Renninger und Hidi, 2011; Schiefele, 2009 und für eine generellere Diskussion bspw. M. Ainley, 2006, S. 393 oder Knogler, 2017). Das Interesse im Sinne eines traits wird als individuelles Interesse bezeichnet, das Interesse im Sinne eines states als Zustand des Interessiertseins. Das Interessiertsein umfasst zwei Komponenten, das aktualisierte individuelle Interesse und das situationale Interesse. Diese beiden Komponenten des Interessiertseins können theoretisch aufgrund der Ursache des Zustands, nicht jedoch aufgrund des Erlebenszustands differenziert werden. Von einem aktualisierten Interesse wird gesprochen, wenn die Ursache des Interessenzustands vorwiegend innerhalb der Person liegt, diese über ein ausgeprägtes individuelles Interesse verfügt. Von einem situationalen Interesse wird gesprochen, wenn der Interessenzustand vorwiegend durch externe Anreize der Lernsituation, die

3. Im Falle des situationalen Interesses wird von Valenzkognitionen (valence cognitions) und im Falle des individuellen Interesses von Valenzüberzeugungen (valence beliefs) gesprochen. Krapp bezeichnet in (vorwiegend älteren) deutschsprachigen Publikationen die Person-Gegenstands-Relation in Abhängigkeit der Interessenform auch als Person-Gegenstands-Beziehung (situationales Interesse) bzw. als Person-Gegenstands-Bezug (individuelles Interesse) (bspw. Krapp, 1992a; Prenzel et al., 1986). In dieser Arbeit wird der Begriff Valenz genutzt, um die Relation zwischen Person und Gegenstand zu beschreiben.

subjektiv empfundene Interessantheit der Situation, getriggert wurde (s. auch empirische Befunde von Knogler et al., 2015). Dies bedeutet, dass das situationale und das (aktualisierte) individuelle Interesse sich den gleichen Zustand (state) teilen. Die relationale Struktur zwischen individuellem Interesse (trait), Anreizbedingungen der Lernsituation und situationalem sowie aktualisiertem Interesse (state) ist in Abbildung 2.1 dargestellt.



**Abb. 2.1:** Relationale Struktur des Interessenkonstrukts entsprechend der Person-Gegenstand-Theorie nach Krapp und Prenzel (1992)

Als Bedingungen für das Interessiertsein (state) sieht Krapp (2005) neben Anreizen der Lernsituation bzw. dem individuellen Interesse unter Berufung auf die Selbstbestimmungstheorie sensu Deci und Ryan (1985a) das Kompetenz-, das Autonomierleben und die soziale Eingebundenheit an. Zudem definiert (Krapp, 1992c) in der Person-Gegenstands-Theorie aufbauend auf der Selbstbestimmungstheorie die interessengeleitete, wiederholte Auseinandersetzung mit dem Gegenstand als zielorientiert, d. h. selbstbestimmt und intentional (Selbstintentionalität).

### Situationales Interesse

Das situationale Interesse beschreibt einen temporären psychischen Prozesszustand (state) in einer konkreten Situation, der durch spezifische Eigenschaften der Umwelt (Lernumgebung), bspw. der Aufgabe oder des Objekts (Anregungsbedingungen), ausgelöst wurde. Der psychische Zustand ist gekennzeichnet durch eine positive emotionale Färbung, Affektivität (affektive Komponente, bzw. gefühlsbezogene Valenzkognitionen), Neugier, gesteigerte Aufmerksamkeit, Persistenz und erhöhte Lernbereitschaft (kognitiv-rationale Komponenten, bzw. wertbezogene Valenzkognitionen) (bspw. Hidi & Harackiewicz, 2000; Krapp, 1992a; Krapp et al., 1992). Krapp (1999a) charakterisiert diesen Zustand zudem durch das Vorliegen einer optimalen Spannung, das Fehlen von Angst und Zwängen von innen oder außen und dem Erleben von Kompetenz. Teils wird in der Forschung aber auch ausschließlich die gefühlsbezogene Valenz zur Definition des situationalen Interesses herangezogen und dieses als (Basis-)Emotion konzeptualisiert (bspw. M. Ainley, 2006; Reeve et al., 2015; Silvia, 2005, 2008). Die Person-Gegenstands-Theorie und das Vier-Phasen-Modell der Interessenentwicklung unterscheiden zwischen zwei Phasen des situationalen Interesses: (a) hervorgerufenen (triggered) und (b) aufrechterhaltenen (maintained) situationalen Interesses, die in der Interessenentwicklung sequentiell aufeinander folgen.<sup>4</sup> Aus dem situationalen Interesse kann durch die wiederholte Auseinandersetzung mit dem Interessengegenstand ein individuelles Interesse erwachsen, sofern die Auseinandersetzung als bedeutsam (wertbezogene Valenzkognition) erlebt wird und positive Gefühle (affektive Valenzkognition) überwiegen.

### Individuelles Interesse

Das individuelle Interesse beschreibt eine relativ stabile Person-Gegenstands-Relation im Sinne einer Disposition bzw. Prädisposition (Hidi & Renninger, 2006). Es ist charakterisiert durch persönlichkeitspezifische Wertvorstellungen und Handlungsbereitschaften (Krapp, 1992c), d. h. durch die habituelle Tendenz bzw. Neigung einer Person zu der wiederholten, freudvollen Auseinandersetzung

4. Vgl. empirische Arbeiten zur catch and hold Komponente des situationalen Interesses von Mitchell (1993), wobei diese laut Renninger et al. (2019) nicht synonym zu verstehen seien

mit dem Interessengegenstand, und erwächst aus der Interaktion der Person mit ihrer Umwelt. Diese interessenbezogenen Handlungen werden als „selbstintentional“ empfunden, da die Handlungsziele mit den das Selbstkonzept konstituierenden Einstellungen, Erwartungen und Werten kongruent sind. Das individuelle Interesse beeinflusst insbesondere Handlungen, in denen die Person Wahlmöglichkeiten hat oder über ihre Zeit frei verfügen kann. Im Kontext Schule könnte dies bspw. die Bearbeitung von Schulaufgaben, Gruppenarbeiten, die Projektarbeit, die Wahl des Profils oder das Engagement in Arbeitsgemeinschaften betreffen. Aufgrund des individuellen Interesses kann das aktualisierte Interesse (state) evoziert werden, wodurch eine Person eine Tätigkeit als interessant erlebt. Knogler et al. (2015) konnten mithilfe eines latent-state-trait Ansatzes zeigen, dass die stabile Komponente des in mehreren Situationen gemessenen Prozesszustands des Interesses (Interessiertsein) durch das zuvor erhobene individuelle Interesse vorhergesagt werden konnte, die variable Komponente jedoch nicht (vgl. Vorhersage des situationalen Interesses durch das individuelle Interesse bspw. bei Tröbst et al., 2016).

Schiefele (2009) unterscheidet zwischen Interessenkonzeptionen, die das individuelle Interesse als affektiv-evaluative Orientierung beschreiben, bspw. die Person-Gegenstands-Theorie, und Konzeptionen, die einen stärkeren Fokus auf die Wert- und Wissenskomponente legen, wie bspw. das Modell des domänenspezifischen Lernens sensu Alexander et al. (1995) und das Vier-Phasen-Modell der Interessenentwicklung.

Individuelles Interesse als affektive-evaluative Orientierung. Krapp, Schiefele und Prenzel definieren das individuelle Interesse anhand von gefühls- (bspw. Aufgeregtheit, Flow, Stimulation) und wertbezogenen Valenzüberzeugungen (bspw. Selbstverwirklichung, Identifikation, Zentralität der Relation in der Selbstdefinition), die in Form von relativ stabilen Gegenstands-Attribut-Relationen im Langzeitgedächtnis gespeichert werden (Krapp & Prenzel, 2011; Schiefele, 2001; Schiefele & Krapp, 1996). Krapp (2002; 2005) führt in seiner Person-Gegenstands-Theorie die wertbezogene Komponente auf ein explizites, bewusstes Regulationssystem und die gefühlsbezogene Valenz auf ein implizites, auf emotionale Erfahrungen beruhendes Regulationssystem zurück. Werden beide Valenzüberzeugungen als positiv erlebt, kann individuelles Interesse auftreten, wobei die Gewichtung der Valenzüberzeugungen je nach Situation und Person variieren kann. Die Entwicklung eines individuellen Interesses als affektiv-evaluative Orientierung wird demzufolge durch die Herausbildung der Valenzüberzeugungen getragen, wobei das Wissen als hiervon separierte Variable zu betrachten sei.

Individuelles Interesse als Prädisposition. Hidi und Renninger (2006) unterscheiden in ihrem Vier-Phasen-Modell der Interessenentwicklung zwischen (a) einem beginnenden (emerging) individuellen Interesse (Übergang zwischen state und trait) und (b) einem gut entwickelten (maintained) individuellen Interesse, welches eine das Wissen beinhaltende Prädisposition darstelle statt wie bei Schiefele (2009) eine ausschließlich affektiv-evaluative Orientierung. Hidi (2006) und Renninger (2000) betonen, dass sich individuelles Interesse durch positive Gefühle, Wissen und Wert (abgeleitet von Kompetenzerfahrungen) auszeichnet, wobei die Phase des gut entwickelten individuellen Interesses sich von dem beginnenden individuellen Interesse durch den Zuwachs an Wissen und Wert abgrenzen lasse. Sie schreiben der Wissenskomponente somit eine größere Bedeutung zu als es die Person-Gegenstands-Theorie macht. Ein wesentliches Merkmal des individuellen Interesses ist die Generierung und Verfolgung von Fragestellungen aufgrund von Neugier (vgl. epistemische Merkmalskomponente in der Person-Gegenstand-Theorie).

## Interessenentwicklung

Das Vier-Phasen-Modell der Interessenentwicklung von Hidi und Renninger (2006) ist derzeit prävalent in der pädagogisch-psychologischen und fachdidaktischen Forschung. Das Modell baut auf den Annahmen von Krapp (2002b) auf und beschreibt die Entwicklung des Interesses aufgrund von Internalisierungs- und Identifikationsprozessen ausgehend von einem situationalen Interesse, das

durch externe Stimuli evoziert wurde, hin zu dem relativ stabilen individuellen Interesse im Sinne einer Prädisposition. Während die Person-Gegenstands-Theorie aufbauend auf den Befunden von Mitchell (1993) zwischen drei Interessenphasen (hervorgerufenes, aufrechterhaltenes situationales Interesse und individuelles Interesse) differenziert, findet in dem Vier-Phasen-Modell der Interessenentwicklung eine weitere Unterscheidung zwischen einem beginnenden und einem gut entwickelten individuellem Interesse statt. Das hervorgerufene situationale Interesse wird durch die Interaktion der Umgebung und der Person evoziert (Passung von Anreizbedingungen und Personenmerkmale) und durch als bedeutungsvoll empfundene Handlungen (Passung zu Zielen, Motiven und Werten) aufrechterhalten. Die Aneignung eines individuellen Interesses läuft anschließend über die Stabilisierung des situationalen Interesses sowie Prozesse der Internalisierung und Identifizierung ab, die eine Integration des Interesses in die Definition des Selbst einschließt (vgl. Cribbs et al., 2015). Als Bedingung für die Ausbildung des individuellen Interesses nennt Krapp (2005) die Erfüllung der drei Grundbedürfnisse der Selbstbestimmungstheorie sensu Deci und Ryan (1985b).

Nachstehend ist der Ablauf über die vierstufige Phasierung der Interessengenese skizziert und für das individuelle Interesse typische Charakteristika von Lernenden (Indikatoren) gemäß Hidi und Renninger (2006) und Renninger et al. (2015, S. 4) aufgeführt.

1. Hervorgerufenes (triggered) situationales Interesse
  - Psychischer Zustand (state), der auf kurzzeitigen Änderungen affektiver und kognitiver Prozesse basiert, die durch externale Faktoren hervorgerufen werden.
2. Aufrechterhaltenes (maintained) situationales Interesse
  - Psychischer Zustand (state), der durch als bedeutsam empfundene Auseinandersetzungsprozesse aus dem hervorgerufenen situationalen Interesse erwächst, durch höhere wertbezogene Valenzen gekennzeichnet ist und über einen längeren Zeitraum als die erste Phase aufrechterhalten wird, sofern eine externale Unterstützung vorliegt.
3. Beginnendes (emerging) individuelles Interesse
  - Übergang zwischen state und Prädisposition, charakterisiert durch positive Gefühle (Affektivität); Wertzuschreibungen und Wissen
  - Indikatoren: teilweise Wiederaufnahme der Auseinandersetzung mit Inhalten; Stellen von Fragen (aufgrund von Neugier)
4. Gut entwickeltes (well-developed) individuelles Interesse
  - Prädisposition; charakterisiert durch positive Gefühle, höhere Wertzuschreibungen und mehr Wissen
  - Indikatoren: selbstbestimmte Auseinandersetzung mit Inhalten; Stellen und Beantworten von Fragen (aufgrund von Neugier)

Die Indikatoren des individuellen Interesses nach Hidi und Renninger (2006) weisen eine starke Assoziation zu im Langzeitgedächtnis gespeicherten Wissensstrukturen und dem Wissenserwerb zuträglichen Verhalten auf und lassen hinsichtlich Forschungsfrage F2 positive Zusammenhänge zwischen dem individuellen Interesse und dem Konzeptverständnis vermuten.

Kritik am Vier-Phasen-Modell der Interessenentwicklung. Das Vier-Phasen-Modell der Interessenentwicklung baut auf einer umfassenden empirischen Befundlage zum Interesse auf und wird oft als Grundlage für Forschungsarbeiten genutzt. Eine erschöpfende empirische Prüfung des Modells sowie die Validierung von Erhebungsinstrumenten steht jedoch noch aus. Eine stärkere empirische Prüfung erscheint erstrebenswert, zumal Hidi und Renninger (2006) vor einem heterogenen Hintergrund argumentieren und sich in ihren Schlussfolgerungen auf sehr divergente Interessenkonzeptionen, Testinstrumente und Stichproben beziehen. Hinsichtlich der Interessenkonzeptionen ist z. B. die Einbeziehung des Wissens (Schiefele, 2009) als Komponente des Interesses umstritten. Dies zeigt sich u. a. anhand der differierenden Konzeptualisierungen der Wissenskomponente während der

Entwicklung der Person-Gegenstands-Theorie und so wird auch von Hidi und Renninger (2006) die Rolle des Wissens für die Definition des individuellen Interesses kritisch hinterfragt. Zudem bleibt die strenge Sequenzierung anhand der vier Phasen empirisch zu untersuchen (Hidi & Renninger, 2006). Schiefele (2009) wirft die Frage auf, ob nicht das wiederholte Evozieren des hervorgerufenen situationalen Interesses (Phase I) über einen längeren Zeitraum ebenfalls ein individuelles Interesse erzeugen könnte, ohne das die Stufe des aufrechterhaltenen situationalen Interesses (Phase II) durchlaufen wird (vgl. Rotgans & Schmidt, 2017b). Auch ob zwischen den beiden Phasen des situationalen Interesses anhand des Zeitraumes des Interessiertseins oder aber anhand der Ausprägung der wertbezogenen Valenz unterschieden werden sollte, ist diskutierbar (Schiefele, 2009). Ebenfalls bleibt zu analysieren, ob idealiter eine lineare Entwicklung des Interesses mit „glatten“ Übergängen zwischen den vier Interessenphasen zu erwarten ist oder ob es spezifische, unterschiedlich geartete Schwellenwerte gibt, anhand derer die Phasen klar zu trennen sind (z.B. die Trennung zwischen hervorgerufenem und aufrechterhaltenem situationalen Interesse; vgl. Knogler et al., 2015; oder zwischen beginnendem und gut entwickeltem individuellen Interesse; s. Rotgans und Schmidt, 2017b). Rotgans und Schmidt (2017a) verweisen darauf, dass von Beginn an Schülerinnen und Schüler zu einem geringen Maße über ein individuelles Interesse verfügten, dies in dem Interessenmodell sensu Hidi und Renninger (2006) aber derzeit unzureichend abgebildet sei.

### Entwicklung schulfachbezogener Interessen

Die Ergebnisse aktueller Literaturreviews (bspw. Krapp & Prenzel, 2011; Potvin & Hasni, 2014b; Wigfield & Cambria, 2010) und Forschungsbeiträge (bspw. Potvin & Hasni, 2014a) deuten darauf hin, dass im Mittel das Interesse von Schülerinnen und Schülern an mathematisch-naturwissenschaftlichen Fächern in der Sekundarstufe I abnimmt (bspw. Dotterer et al., 2009; Schiefer et al., 2018). Eine Studie von Gaspard et al. (2017) zeigte zudem, dass die eng mit dem Interesse verknüpften Konstrukte intrinsischer Wert (z. B. „I like doing ...“, „I simply like ...“) und die persönliche Wichtigkeit (z. B. „...is not meaningful to me.“, „...is very important to me personally.“) in den Fächern Mathe, Biologie und Physik substantiell in der Sekundarstufe I abnehmen, dieser Trend sich aber partiell in der Sekundarstufe II abschwächt und hinsichtlich der Biologie in höheren Jahrgangsstufen diese beiden motivationalen Konstrukte im Mittel von den Schülerinnen und Schülern sogar wieder günstiger bewertet wurden. Die Annahme eines Abfalls des Fachinteresses in den Naturwissenschaften und der Mathematik in der Sekundarstufe I wird auch durch intraindividuelle Analyseansätze im Bereich der Mathematik gestützt, die zudem eine Verschiebung der Gewichtung von der gefühlsbezogenen hin zu der wertbezogenen Valenz des Interesses nahelegen (Frenzel et al., 2010; Frenzel et al., 2012). Dem Gegenüber gibt es Evidenz dafür, dass zwar das mittlere Fachinteresse mit steigender Jahrgangsstufe abnimmt, einzelne (neue) Fachinhalte und Tätigkeiten jedoch auch in höheren Jahrgangsstufen als hochinteressant erlebt werden (Häussler & Hoffmann, 2000).

Das Modell der Interessendifferenzierung versucht die Abnahme des mittleren Interesses mit steigender Jahrgangsstufe vorwiegend auf Grundlage innerschulischer Faktoren (s. u. a. Z. Daniels, 2008; Köller et al., 2001; Schiefele, 2009) zu erklären. Entsprechend des Modells der Interessendifferenzierung (Z. Daniels, 2008; Schiefer et al., 2018; Schurtz et al., 2014; Todt & Schreiber, 1998) liegt zum Zeitpunkt des Schuleintritts noch ein nahezu universelles Interesse der Schülerinnen und Schüler vor. Von einem individuellen Interesse an bestimmten Fächern im Sinne von Hidi und Renninger (2006) kann jedoch meist nicht gesprochen werden, da aufgrund geringer Vorerfahrungen nur wenig ausgeprägte Wissensstrukturen über das Fach vorliegen. Im Laufe der Schulzeit kommt es dann zu einer wiederholten Auseinandersetzung mit dem Fach und die Schülerinnen und Schüler lernen neue Fachinhalte und Tätigkeiten kennen, von denen sie einige als interessant, andere als weniger interessant erleben. Aufgrund dieser Erfahrungen bilden sich Vorlieben für bestimmte Kontexte, spezifische Inhalte und Tätigkeiten aus (z.B. Häussler & Hoffmann, 2000; Lewalter et al., 2001). Infolgedessen findet eine Interessenfokussierung auf diejenigen Fächer statt, in denen sich die Schülerinnen und Schüler vorwiegend engagieren. Währenddessen nimmt ihr Interesse an den übrigen Fächern ab, wodurch eine Differenzierung des Interesses stattfindet (Selbstdifferenzierung),

welche bspw. durch abnehmende Korrelationen zwischen den Fachinteressen gekennzeichnet sein sollte.<sup>5</sup> Dies bedeutet, dass im Mittel zwar das Interesse der Schülerinnen und Schüler sinkt, jedoch ebenso Personengruppen vorliegen müssten, die durch ein konstantes oder gar steigendes Interesse zu charakterisieren sind, nur ihr Vorhandensein in interindividuellen Ansätzen nicht aufgedeckt wird.

Ein zweiter Erklärungsansatz für die Differenzierung der schulfachbezogenen Interessen adressiert das akademische Selbstkonzept und das Kompetenzerleben, welche beide wiederum eng mit dem Interesse verknüpft sind (Weidinger et al., 2016). Dieser Ansatz baut auf dem internal/external frame of reference (I/E) Model von Marsh (1986) auf.<sup>6</sup> Schülerinnen und Schüler vergleichen ihre erbrachten Leistungen und erhaltenen Leistungsrückmeldungen zwischen verschiedenen Schulfächern (dimensionale Vergleiche Gaspard et al., 2018; Wolff et al., 2019 und Schurtz et al., 2014), mit denen ihren Mitschülerinnen und Mitschülern (soziale Vergleiche Köller, Schnabel et al., 2000 und Schurtz et al., 2014) und mit eigenen bereits zuvor erbrachten Leistungen und erhaltenen Leistungsrückmeldungen (temporalen Vergleichen bspw. Wolff et al., 2019). Diese Vergleiche können positive und negative Auswirkungen auf das fachspezifische Selbstkonzept haben. Insbesondere dimensionale Vergleiche bedingen aufgrund von Kontrast- und Assimilationseffekten eine Ausdifferenzierung des akademischen Selbstkonzepts, welches wiederum die Ausprägungen der Fachinteressen beeinflusst (Z. Daniels, 2008; Schiefer et al., 2018; Schurtz et al., 2014). Neben den dimensional Vergleichen wirken sich auch soziale Vergleiche auf das Interesse der Schülerinnen und Schüler aus. So konnte bspw. ein negativer Effekt der durchschnittlichen Performanz von Schülerinnen und Schülern in einer Klasse auf das Fachinteresse (Big-Fish-Little-Pond-Effekt), mediiert über das akademische Selbstkonzept, gezeigt werden (Köller et al., 2001). Generell sind die Entwicklung des Selbstkonzepts und die Entwicklung der Fachinteressen eng miteinander verwoben und tragen wesentlich zur Ausbildung der „Identität“ bzw. der Selbstdefinition bei. Zwischen beiden Konzepten liegen hohe Korrelationen und kreuzverzögerte Effekte vor (bspw. Marsh et al., 2005), sodass in höheren Jahrgangsstufen ein enger Zusammenhang zwischen Selbstkonzept, Interessen und Schulleistung existieren kann (Denissen et al., 2007).

Weitere Erklärungsansätze für die langfristige Entwicklung und Abnahme des Fachinteresses in der Sekundarstufe I fokussieren auf konkurrierende außerschulische Interessenbereiche, die fehlende Passung zwischen Bedürfnissen und Opportunitätsstrukturen bzw. Lernangeboten (Stage-Environment-Fit-Ansatz nach Eccles et al., bspw. Eccles und Midgley, 1989 oder Eccles et al., 1993) und der Orientierung an Berufsanforderungen (vgl. Interessenskongruenzen Etzel und Nagy, 2015, und Entwicklungsfragen nach Todt und Schreiber, 1998). Ansätze zur Förderung des Interesses können das situationale Interesse (bspw. kognitive Dissonanz, hands-on Tätigkeiten), das individuelle Interesse (bspw. Wahl der Kontexte) sowie Klassenraumvariablen, die bspw. im Zusammenhang zum Autonomieerleben (z. B. Eckes et al., 2018; Lazarides & Raufelder, 2017; Tsai et al., 2008), dem Kompetenzerleben (z. B. Rakoczy et al., 2018) und der soziale Eingebundenheit (Minnaert et al., 2011) stehen und somit Bedingungen für die Interessenentwicklung bereitstellen (Krapp, 2005), in den Fokus nehmen.

### Das RIASEC+N Interessenmodell

Das RIASEC+N Interessenmodell wurde entwickelt, um differenzierte Einblicke in die Interessen der Schülerinnen und Schüler an Naturwissenschaften zu erzielen und präzise Interessenprofile erstellen zu können. Der Anreiz solcher differenzierten Interessenprofile wurde in einer spezifischen, zielgerichteten und somit effektiven Förderung des Interesses durch Schaffung von Kongruenz zwischen den Anreizen einer Intervention bzw. Fördermaßnahme und den Interessenprofilen der Schülerinnen und Schüler gesehen (Dierks, 2014). Bei der Konzeptualisierung des RIASEC+N Modells fand eine starke Orientierung zum einen an dem Modell der Berufsinteressen sensu Hol-

5. Krapp (2002b) und Fink (1992) beschreiben drei mögliche Modelle zur ontogenitischen Entwicklung des individuellen Interesses, die jedoch nur unzureichend empirisch abgesichert sind.

6. Verallgemeinerungen des I/E Modells wurden bspw. von Wolff et al. (2019) vorgenommen.

land (1997) sowie den pädagogisch-psychologischen Interessenmodellen nach Todt und Schreiber (1998) und Hoffmann et al. (1998) statt, die eine Untergliederung des domänenspezifischen Interesses in Kontexte (Beruf, Freizeit, Schule), Inhalte sowie Interessenkomponenten (state- und trait-Komponenten) vornahm. So unterscheidet das RIASEC+N Modell in Anlehnung an das topologische Interessenkonstrukt von Hoffmann et al. (1998) zwischen Kontexten (bspw. Schulfächer, Fördermaßnahmen, Freizeitangebote und Berufsfelder), Inhalten (bspw. Säure-Base-Konzept) und naturwissenschaftlichen Tätigkeiten, die in einem spezifischen Kontext anhand eines spezifischen Inhalts ausgeübt werden. Eine große Änderung gegenüber dem topologischen Interessenkonstrukt ist die Ausdifferenzierung des Interesses durch eine tätigkeitsspezifische Konzeptionalisierung. Das RIASEC+N Modell unterscheidet zwischen Interessen der Schülerinnen und Schüler an verschiedenen Bereichen von naturwissenschaftlichen Tätigkeiten, die durch diese während des Unterrichts aktiv ausgeübt werden, wobei eine Orientierung an dem RIASEC Modell sensu Holland (1997) vorgenommen wurde. Das RIASEC+N Modell unterscheidet in naturwissenschaftlichen Schulfächern (Chemie, Physik, Biologie) zwischen den sieben Interessendimensionen realistic (R), investigative (I), artistic (A), social (S), enterprising (E), conventional (C) und networking (N).<sup>7</sup> Die sieben Interessendimensionen des RIASEC+N Modells nach Dierks et al. (2014b) lassen sich für den Chemieunterricht in Kürze wie folgt charakterisieren:

1. Die Interessendimension realistic adressiert praktische, handwerkliche Tätigkeiten des Chemieunterrichts wie bspw. die Durchführung von Experimenten nach einer vorgegebenen Anleitung.
2. Die Interessendimension investigative adressiert analytische, forschende Tätigkeiten wie bspw. das Planen von Experimenten.
3. Die Interessendimension artistic adressiert kreative, künstlerische Tätigkeiten des Chemieunterrichts wie bspw. das Zeichnen von experimentellen Versuchsaufbauten.
4. Die Interessendimension social adressiert Tätigkeiten des Chemieunterrichts, bei denen sich Schülerinnen und Schüler gegenseitig unterstützen, wie bspw. das Helfen von Mitschülerinnen und -schülern bei Chemieaufgaben.
5. Die Interessendimension enterprising adressiert leitende Tätigkeiten im Chemieunterricht wie bspw. die Führung einer Gruppenarbeit.
6. Die Interessendimension conventional adressiert systematisierende Tätigkeiten des Chemieunterrichts wie bspw. das Sammeln und Ordnen chemischer Produkte.
7. Die Interessendimension networking adressiert austauschende und vernetzende Tätigkeiten des Chemieunterrichts wie bspw. die gemeinsame Erarbeitung von Aufgaben mit Mitschülerinnen und Mitschülern.

Die hexagonale (circumplex) Struktur des RIASEC Modells, die eine konzeptionelle und empirisch prüfbar nahe der Interessendimensionen beschreibt, wurde aufgrund der Fokussierung auf das Feld der Naturwissenschaften aufgegeben. Diese Fokussierung schlägt sich auch in der korrelativen Struktur zwischen den Interessendimensionen des RIASEC+N Modells und denen des RIASEC Modells sensu Holland (1997) nieder, welches naturwissenschaftliche Berufe vorwiegend im realistic und investigative Bereich ansiedelt (Dierks et al., 2016). Empirische Evidenz, dass die Tätigkeiten, die den sieben Interessendimensionen zugrunde liegen, für den Unterricht sowie das berufliche Arbeiten in den Naturwissenschaften relevant sind, konnte mithilfe von Interviews mit Lehrkräften (Dierks, 2014) und Wissenschaftlerinnen und Wissenschaftlern (Stamer et al., 2019; Wentorf et al., 2015) gewonnen werden. Mittlerweile wurde das RIASEC+N Modell in verschiedenen schulischen und außerschulischen Kontexten genutzt (Blankenburg, 2015; Blankenburg et al., 2015, 2016; Blankenburg, Höffler, Peters et al., 2016; Dierks, 2014; Dierks et al., 2016; Dierks et al., 2014a, 2014b; Habig, Blankenburg et al., 2018; Höffler et al., 2019), vorwiegend um mittlere Interessenausprägungen zu bestimmen und diese zwischen Gruppen von Schülerinnen und Schülern zu vergleichen. Für die

7. Blankenburg, Höffler und Parchmann (2016) vergleichen die Operationalisierung von naturwissenschaftlichen Tätigkeiten in den Forschungsarbeiten von Häussler und Hoffmann (2000), Swarat et al. (2012) und Dierks et al. (2014b).

vorliegende Forschungsarbeit waren insbesondere die Arbeiten von Dierks et al. (2016), Dierks et al. (2014b) und Blankenburg, Höffler und Parchmann (2016) relevant, da in diesen die empirischen und theoretischen Grundlagen für diese Arbeit gelegt wurden. Dierks et al. (2014b) beschreiben die explorative Analyse des RIASEC+N Modells mittels explorativer Faktorenanalysen in den Kontexten Schule, Freizeit und Fördermaßnahmen (Enrichment), wobei für den schulischen Kontext eine Beschreibung der Daten anhand von sieben anstatt, wie erwartet, sechs latenten Faktoren vorteilhaft war, weswegen eine zusätzliche Interessendimension (networking) in das Modell integriert wurde. Hieran anschließend konnte mittels konfirmatorischen Faktorenanalysen weitere Evidenz für die RIASEC+N Struktur gewonnen werden, indem Bi-Faktor-Modelle gegen weitere plausible Modelle getestet wurden. In den beiden Studien von Dierks et al. (2014b) und Blankenburg, Höffler und Parchmann (2016) wurden in den Bi-Faktor-Modellen jeweils zwei Faktorladungen je Interessenitem zugelassen. So lud in den konfirmatorischen Faktorenanalysen jedes Interessenitem auf einer Interessendimension und einem Kontext (Schule, Enrichment, Beruf bei Dierks et al., 2016 und Chemie, Physik, Biologie bei Blankenburg, Höffler und Parchmann, 2016). Diese Studien umfassten jedoch lediglich einzelne Jahrgangsstufen, weshalb als ein Forschungsdesiderat die Prüfung der Interessenstruktur im Fach Chemie über die Sekundarstufen I und II abgeleitet wurde. Hierzu wurde ein Messinstrument entwickelt und im DoLiS-Projekt eingesetzt, das als Kontext das Fach Chemie annimmt, jedoch keinen spezifischen Fachinhalt beinhaltet, aber weiterhin die RIASEC+N Dimensionen abbildet. Die Aussparung des Fachinhaltes erfolgte zum einen aufgrund testökonomischer Aspekte (Testdauer) und zum anderen, um eine Messinvarianz über die Jahrgangsstufen zu gewährleisten und somit eine bedeutsame Vergleichbarkeit zu ermöglichen. Die querschnittliche Prüfung des RIASEC+N Modells erfolgte in Publikation 1 (Höft et al., 2019) und in Kapitel 4.

### 2.3 Kovariaten

#### Konzeptverständnis

Die Performanzen der Schülerinnen und Schüler wurden in dem DoLiS-Projekt mithilfe eines Chemiefachtests erfasst, der das Verständnis der in Bildungsstandards und Lehrplänen formulierten zentralen Konzepte (Basiskonzepte) fokussierte (BIMI, 2016; KMK, 2004). Als Basiskonzepte werden für den Chemieunterricht das Stoff-Teilchen-Konzept und das Konzept der Struktur-Eigenschaftsbeziehungen, bezogen auf den atomaren Aufbau und die Zusammensetzung der Materie, das Konzept der chemischen Reaktion und das Energiekonzept spezifiziert. Die Strukturierung des Fachunterrichts anhand zentraler Konzepte der Domäne soll das kumulative Lernen und den Erwerb von vernetzten Wissensstrukturen fördern und den Aufbau von kompartimentalisierten Wissensseinheiten mindern (Duschl et al., 2011; Osborne & Collins, 2000; Osborne & Dillon, 2008). Wissensstrukturen, die entlang der zentralen Konzepte vernetzt sind, sollen es Schülerinnen und Schülern erlauben, die Relevanz eines Konzepts in einer Vielzahl von Problemsituationen zu erkennen, diese mithilfe der Basiskonzepte erklären und bewerten und darauf aufbauend persönlich und gesellschaftlich relevante Entscheidungen treffen zu können. Bisherige empirische Befunde deuten darauf hin, dass es Schülerinnen und Schülern schwerfällt, ein sophistiziertes Verständnis der zentralen Konzepte aufzubauen (bspw. Hadenfeldt et al., 2016; Löfgren & Helldén, 2009; Neumann et al., 2013; Stains & Talanquer, 2008; Taskin & Bernholt, 2014; Yao et al., 2017). So konnten hinsichtlich des Konzepts der Energie Neumann et al. (2013) in einer querschnittlichen Studie in den Klassenstufen 6, 8 und 10 zeigen, dass es Schülerinnen und Schülern insbesondere schwerfällt, ein vertieftes Verständnis des Prinzips der Energieerhaltung zu erwerben. Hinsichtlich des Konzepts der chemischen Reaktion schlussfolgern Taskin und Bernholt (2014) aus einem Literaturreview, dass Schülerinnen und Schüler meist dazu in der Lage sind, bekannte Reaktionsgleichungen zu lösen bzw. auszugleichen, es ihnen jedoch schwerfällt, diese mit submikroskopischen Deutungen und stöchiometrischen Berechnungen zu verbinden, weswegen der Transfer auf neue, weniger bekannte Beispiele häufig misslingt. Bislang gibt es jedoch nur wenige Studien, die eine längsschnittliche Analyse des Konzeptverständnisses vornehmen und dabei differenzierte Einblicke in die durch zentrale Konzepte suggerierte



Dimensionalität des Konzeptverständnisses in den Naturwissenschaften liefern.

Der erfolgreiche Erwerb eines sophistizierten Verständnisses zentraler Konzepte der Chemie scheint mit dem individuellen Fachinteresse der Schülerinnen und Schüler positiv verknüpft zu sein. Dies ist zum einen indirekt durch das Kurswahlverhalten bedingt (bspw. Köller et al., 2001; Köller, Daniels et al., 2000; Kriegbaum et al., 2018; Nagy et al., 2007), zum anderen deuten empirische Befunde wiederholt auch auf direkte positive Effekte des individuellen Interesses auf die Performanz von Schülerinnen und Schülern hin (Jansen et al., 2016). Gleichfalls konnte gezeigt werden, dass der Wissenserwerb ebenfalls die Interessenentwicklung fördern kann (bspw. Rotgans & Schmidt, 2017c). Ein Modell zur Analyse des Zusammenhangs zwischen dem individuellen Interesse und dem Konzeptverständnis der Schülerinnen und Schüler erfolgt im nachstehenden Abschnitt.

### Lernfreude

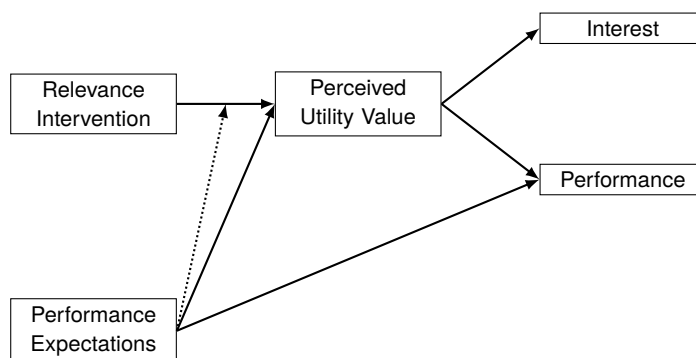
Die Lernfreude wurde in der vorliegenden Arbeit als Leistungsemotion im Rahmen der Kontroll-Wert-Theorie (KWT) aufgefasst (Pekrun, 2006). Die Kontroll-Wert-Theorie liefert einen theoretischen Rahmen zur Erklärung von Antezedenzen (Kontroll- und Wertappraisals; s. bspw. Putwain et al., 2018) und Effekten von Leistungsemotionen. Sie konzeptualisiert die Freude als eine positive, aktivierende Leistungsemotion. Dies äußert sich bspw. in einer erhöhten Aufmerksamkeitssteuerung und der höheren Wahrscheinlichkeit des „tiefen“ Lernens. Entsprechend konnten in längsschnittlichen Studien positive Zusammenhänge zwischen der Lernfreude und der Leistung (bspw. Ahmed et al., 2013; Pekrun et al., 2017; Pinxten et al., 2014; Putwain et al., 2017; Tibubos et al., 2019) sowie positive Effekte der Lernfreude auf das Interesse gezeigt werden (bspw. Schukajlow & Rakoczy, 2016). Aufgrund der positiven Eigenschaften der Lernfreude auf das Lernen sind die Befunde aus den PISA Studien 2006 und 2015 als negativ zu bewerten, dass in Deutschland die Freude der getesteten Schülerinnen und Schüler an Naturwissenschaften im Jahr 2015 signifikant unter dem Wert der PISA Studie im Jahr 2006 lag und zudem Schüler eine höhere Freude angaben als Schülerinnen (OECD, 2016, p. 121).

Für diese Forschungsarbeit ist die Einbeziehung der Lernfreude in die Analysen der Entwicklung des individuellen Interesses aufgrund der konzeptuellen Überlappung zur gefühlsbezogenen (affektiven) Valenz des Interesses spannend. Das Testinventar des DoLiS-Projekts sah neben der Erfassung des Interesses der Schülerinnen und Schüler an naturwissenschaftlichen Tätigkeiten im Chemieunterricht keine weitere systematische Erfassung des Fachinteresses vor, sodass keine Informationen über die gefühls- und wertbezogene Komponente des Fachinteresses vorlagen. Daher wurde die Lernfreude genutzt, um die affektiven Aspekte des Lernens während des Chemieunterrichts zu modellieren und Einflüsse auf die Ausbildung eines individuellen Interesses an der Domäne Chemie zu untersuchen.

### Wahrgenommene Nützlichkeit

Ein weiteres Ziel des Chemieunterrichts ist es, den Schülerinnen und Schülern die Relevanz und Nützlichkeit des Faches aufzuzeigen. Die wahrgenommene Nützlichkeit (utility value) eines Objekts (hier des Faches Chemie) umfasst mehrere Facetten, bspw. die Nützlichkeit für den angestrebten Beruf oder die Gesellschaft (bspw. Gaspard, Dicke, Flunger, Schreier et al., 2015). Die Nützlichkeit im Sinne des Erwartungs-Wert-Modells ist als Nützlichkeit bzw. Relevanz eines Handlungsergebnisses für das Erreichen von kurz- und langfristigen Zielen definiert und weist somit einen starken extrinsischen Charakter auf (Jiang et al., 2018; Wigfield & Eccles, 2000). Dementsprechend ist die wahrgenommene Nützlichkeit ein guter Prädiktor für das Verhalten von Schülerinnen und Schülern in Leistungskontexten (bspw. Fredricks et al., 2018). Ähnlich wie beim Fachinteresse nimmt die wahrgenommene Nützlichkeit von naturwissenschaftlichen Fächern in der Sekundarstufe I ab. Die wahrgenommene Nützlichkeit der Biologie steigt in der Sekundarstufe II insbesondere für Schülerinnen an, während sie für das Fach Physik für Jungen in der Sekundarstufe II stagniert und für Mädchen weiter fällt (Gaspard et al., 2017; Einflussfaktoren auf die Entwicklung der wahrgenommenen Nützlichkeit

werden bspw. bei Bathgate und Schunn, 2017a, diskutiert). Die wahrgenommene Nützlichkeit gilt als ein Faktor, um das Interesse und die Performanz von Schülerinnen und Schülern nachhaltig zu steigern (bspw. Canning et al., 2019; Fryer & Ainley, 2019; Gaspard, Dicke, Flunger, Brisson et al., 2015). Sogenannte Utility-Value-Interventionen zielen darauf ab, die Relevanz des Fachinhalts bspw. durch eigenständig gezogene oder durch die Lehrperson introjizierte Verknüpfungen zwischen dem Lerngegenstand und Alltagssituationen, bzw. außerschulischen Erfahrungen, oder späteren Berufen aufzuzeigen, zum Reflektieren dieser Verknüpfung anzuregen und somit die wahrgenommene Nützlichkeit für kurz- und langfristige Ziele zu fördern (bspw. Akcaoglu et al., 2018; Johnson & Sinatra, 2013). Hierbei ist zu beachten, dass das akademische Selbstkonzept den Einfluss von Instruktionsmaßnahmen auf die wahrgenommene Nützlichkeit moderiert (s. Abbildung 2.2), sodass davon auszugehen ist, dass zur Steigerung der wahrgenommenen Nützlichkeit der Chemie eine Binnendifferenzierung durch die Lehrkräfte vorgenommen werden müsste (vgl. Hulleman et al., 2010).



**Abb. 2.2:** Utility-Value-Intervention zur Steigerung des Interesses und der Performanz (Hulleman et al., 2010).

Hinsichtlich des naturwissenschaftlichen Unterrichts bilden kontext-basierte Unterrichtsansätze eine gut untersuchte Methode zur Förderung der wahrgenommenen Nützlichkeit des Fachinhaltes (Sevian et al., 2018). Diese Unterrichtsansätze zielen darauf ab, wissenschaftliches Wissen bspw. mit Alltagskontexten zu verknüpfen und das Interesse sowie das Konzeptverständnis zu fördern (Fechner, 2009; Ültay & Çalik, 2012). Hierbei ist es entscheidend, für die Schülerinnen und Schüler bedeutungsvolle Kontexte zu identifizieren, die die Relevanz des Fachinhalts für die persönlichen Ziele der Schülerinnen und Schüler in gesellschaftlichen Kontexten aufzeigen. Zeitgleich versuchen Ansätze wie „Chemie im Kontext“ die wahrgenommene Autonomie der Schülerinnen und Schüler zu steigern, um eine weitere Bedingung für die Entwicklung eines individuellen Interesses zu adressieren (Overman et al., 2014; Parchmann et al., 2006). Dies kann gelingen, wenn die Schülerinnen und Schüler zwar Autonomie (bspw. durch die Mitgestaltung des Unterrichts, Bestimmung des Kontexts und der Festlegung der Ziele, etc.) erleben, sich zeitgleich aber bei der Aufgabenbearbeitung ausreichend unterstützt fühlen (Broman et al., 2018; Habig, Blankenburg et al., 2018).

## 2.4 Die Erwartungs-Wert Theorie zur Generierung eines Rahmenmodells

Als Rahmenmodell zur Erklärung des Zusammenspiels von Interesse, Leistung, Lernfreude und wahrgenommener, subjektiver Nützlichkeit wurde für diese Forschungsarbeit das erweiterte Erwartungs-Wert-Modell der Leistungsmotivation sensu Eccles et al. gewählt (Eccles, 2009; Eccles et al., 2015; Eccles & Jacobs, 1986; Wigfield & Eccles, 2000). Das Erwartungs-Wert-Modell beschreibt das Handeln von Personen in Leistungssituationen, weswegen es vermehrt in der pädagogisch-psychologischen Forschung zur Untersuchung des schulischen Leistungsverhaltens zum Einsatz kommt und vor dessen Hintergrund es im Folgenden diskutiert wird. Im schulischen Kontext hat es sich insbesondere für die Vorhersage von Wahlentscheidungen (bspw. Nagy et al., 2007), Berufsaspirationen (Lauer-mann et al., 2017) und der Performanz (Guo, Marsh, Parker et al., 2015; Guo et al., 2016) bewährt. Das Erwartungs-Wert-Modell nimmt an, dass die nächsten Determinanten von leistungsthematischem

Verhalten (bspw. Aufgabenwahl, Ausdauer, Leistung) die tätigkeitsspezifischen Erfolgserwartungen (Erfolgszuversicht) und Wertzuschreibungen (u. a. mit der Tätigkeit assoziierte, positive Anreize) einer Person sind. Aufgrund der Tätigkeitspezifität der Erwartungs- und der Wertkomponente eignet sich diese Theorie, um intraindivuelle Unterschiede der Motivationsausprägung in differierenden Leistungssituationen zu modellieren. Beide Komponenten, Erwartung und Wert, sind in dem Modell direkt und unabhängig voneinander wirkende Determinanten der Leistungsmotivation. Verfügt eine Person über hohe aufgabenspezifische Erfolgserwartungen und Wertzuschreibungen, so sollte sie eher dazu neigen, fordernde Aufgaben zu bearbeiten (Wahl), die Bearbeitung eher als positiv zu bewerten (bspw. Kompetenzerleben, Freude), elaboriertere Lernstrategien anzuwenden, bei auftretenden Schwierigkeiten nicht abzubrechen (Persistenz) und als Ergebnis einen höheren Lernzuwachs verzeichnen (Leistung) sowie die Aufgabenbearbeitung retrospektiv positiv bewerten. Letzteres heißt, dass die Erfolgserwartungen und Wertzuschreibungen einer Person wiederum durch bisherige Erfahrungen in vergleichbaren Kontexten und deren Interpretation sowie Attribution (Weiner, 2010) geprägt werden. Weiterhin diskutiert die Erwartungs-Wert-Theorie den Einfluss der kulturellen und sozialen Umwelt und der Persönlichkeit auf die Kernvariablen Erwartung und Wert (bspw. Häfner et al., 2017; Kriegbaum & Spinath, 2016; Muntoni & Retelsdorf, 2019), sodass insgesamt ein sehr komplexes Wirkungsgefüge in der Theorie beschrieben wird (Eccles & Jacobs, 1986). In dieser Forschungsarbeit wird der Fokus auf die Personenmerkmale sowie die tätigkeitsspezifische Wertkomponente und die aus dem schulischen Leistungsverhalten resultierenden Lernresultate diskutiert (s. Abbildung 2.3). Dieses reduzierte Modell modelliert eine längsschnittliche Schleife, in der

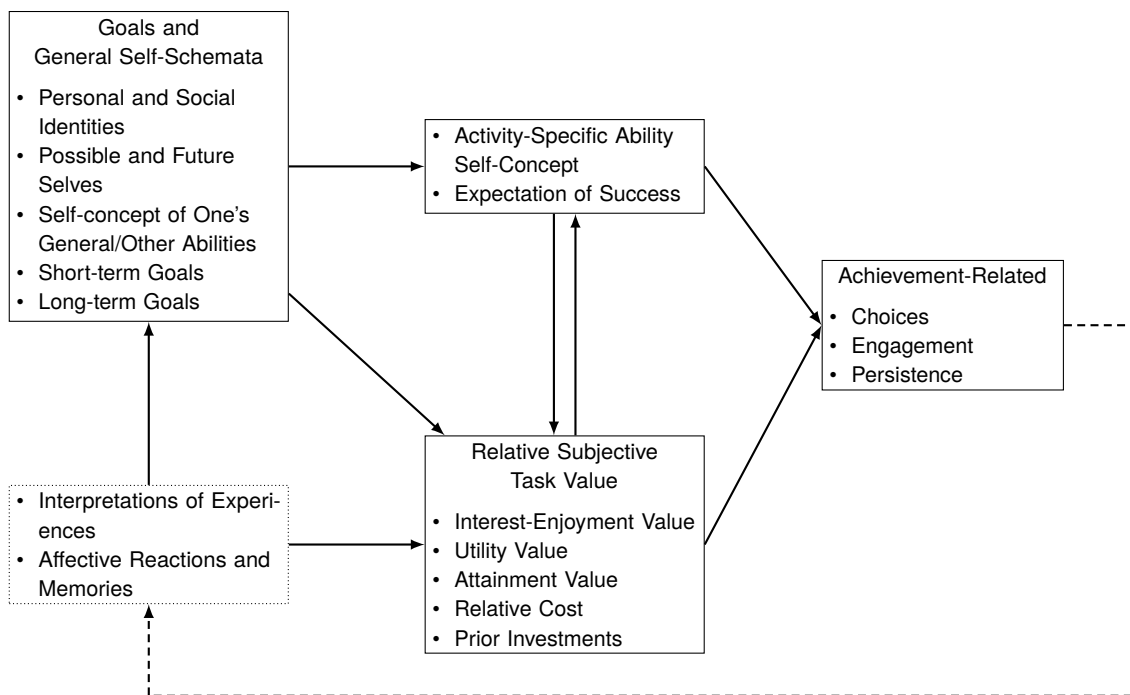


Abb. 2.3: Reduziertes Modell der Erwartungs-Wert-Theorie sensu Eccles et al. (2015).

die Interpretation und Attribution vorheriger Leistung und Leistungsrückmeldungen (bspw. durch Peers, Lehrkräfte und Noten) sowohl Personenmerkmale (bspw. Ziele, das Selbstkonzept) als auch die tätigkeitsspezifische Wertkomponente beeinflussen (s. z. B. Gaspard et al., 2018; Weiner, 2010). Frühere Erfahrungen in vergleichbaren Situationen, bspw. eine ähnliche Aufgabe auf eine bestimmte Art erfolgreich lösen zu können bzw. Misserfolg bei der Bewältigung ähnlicher Aufgaben, sind Determinanten der Entwicklung von Personenmerkmalen, zu denen bspw. das generelle Selbstkonzept zählt. Die Personenmerkmale wiederum beeinflussen in einer spezifischen Situation direkt die tätigkeitsspezifische Erwartungs- und Wertkomponente von Schülerinnen und Schülern und haben somit gleichfalls direkte Auswirkungen auf das Verhalten in Leistungssituationen (Lernverhalten) sowie auf die erbrachte Leistung. In diesem Beziehungsgeflecht nimmt demnach das Selbstkonzept

die Rolle eines Mediators zwischen den Leistungserfahrungen und der Wertkomponente und dem Leistungs/Lernverhalten ein.<sup>8</sup> So wirkt sich ein niedriges akademisches Selbstkonzept ungünstig auf das Lernverhalten und das Lernresultat aus, da die Wertkomponente eine negativere Ausprägung als bei Schülerinnen und Schülern mit einem höheren tätigkeitsspezifischen Selbstkonzept aufweist, was sich z. B. in einer emotional negativeren Empfindung der Lernsituation und der Wichtigkeit widerspiegelt (bspw. Weidinger et al., 2016). Im Folgenden wird diese sich selbstverstärkende Wirkungskette mithilfe des Interessenkonstrukts ausstaffiert und erläutert.

Integration des Interesses in das Erwartungs-Wert-Modell. Das individuelle Interesse entsprechend der Person-Gegenstands-Theorie (Krapp, 2002b; Schiefele, 2009) bzw. des Vier-Phasen-Modells der Interessengenese (Hidi, 2006) kann mit Aspekten der Wertkomponenten des Erwartungs-Wert-Modells assoziiert werden, unter denen Aufgabenanreize verstanden werden, die sich zum einen auf den angestrebten Ergebniszustand und dessen Folgen, zum anderen auf tätigkeitszentrierte Anreize beziehen. Die Wertkomponente des Erwartungs-Wert-Modells umfasst demnach im Gegensatz zum Interesse sowohl intrinsische<sup>9</sup> als auch extrinsische Aufgabenanreize. Sie setzt sich nach Eccles et al. (2015) aus einer intrinsischen (enjoyment, interest), Nützlichkeits- (utility), Wichtigkeits- (attainment) und Kostenvariable (cost und prior investments) zusammen.<sup>10</sup> Die intrinsische Variable beschreibt die Freude und die positiven Gefühle, die während der Tätigkeitsausführung erfahren werden. Dieser Aspekt der Wertkomponente weist demnach eine große konzeptuelle Überlappung mit der gefühlsbezogenen Valenz des Interesses auf. Die Wichtigkeitsvariable deckt Bereiche der wertbezogenen Valenz des Interesses ab (bspw. Köller, Daniels et al., 2000; Tsai et al., 2008).

Eine Assoziation des Interesses mit der Wertkomponente des Erwartungs-Wert-Modells liefert einen theoretischen Rahmen, um Einflussfaktoren auf die Interessenentwicklung sowie die Wirkung des Interesses in Leistungskontexten zu diskutieren. Ein Beispiel für die Diskussion der Interessenentwicklung vor dem Hintergrund des Erwartungs-Wert-Modells ist das oben angesprochene Modell der Interessendifferenzierung, das Wirkmechanismen zwischen Leistung, Selbstkonzept und Interesse infolge von dimensionale, soziale und temporale Vergleiche beschreibt.

Modellierung des Zusammenhangs zwischen Interesse und Leistung. Zur Modellierung des Zusammenhangs zwischen Interesse und Leistung wurde in den vier subsequent berichteten Publikationen dieser Forschungsarbeit auf Grundlage des Erwartungs-Wert-Modells auf ein Reciprocal Effects Model (REM, s. Pekrun et al., 2017; Sewasew et al., 2018) zurückgegriffen, das beidseitige, positive, kreuzverzögerte Effekte zwischen dem individuellen Interesse und der Leistung annimmt, die z. T. über das Selbstkonzept mediiert werden.

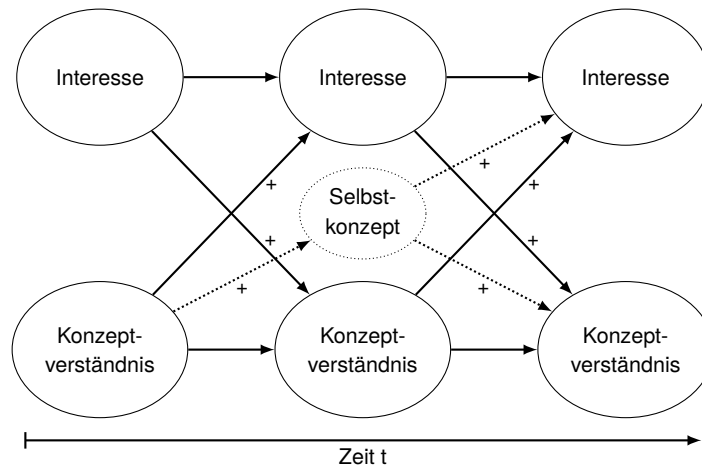
Gemäß dem Modell der reziproken Effekte nehmen interessierte Schülerinnen und Schüler eher lernförderliche Angebote im und außerhalb des Unterrichts wahr (bspw. Köller et al., 2001), zeigen bei der Bearbeitung von Aufgaben eher lern- und leistungsfördernde psychische Prozesse (bspw. gesteigerte Aufmerksamkeit z. B. M. Ainley, Hillman et al., 2002, Selbstregulation z. B. Lee et al., 2014) und Verhaltensweisen (bspw. Lernstrategien z. B. Berger und Karabenick, 2011, Anstrengung z. B. Trautwein et al., 2015; Xu, 2018) empfinden die Aufgabenbearbeitung zudem eher als positiv (gefühlsbezogene Valenz) und erleben sich selbst eher als kompetent (wertbezogene Valenz), sodass sie mit einer höheren Wahrscheinlichkeit höhere Lernzuwächse und bessere Testergebnisse erzielen. Aufgrund der positiven affektiven Tönung der Tätigkeitsdurchführung sowie der eher positiven

8. vgl. Ausbildung spezifischer Fachinteressen aufgrund der Ausdifferenzierung des akademischen Selbstkonzepts, die auf sozialen, dimensional und temporalen Vergleich basiert, u. a. bei Z. Daniels, 2008; Gaspard et al., 2018; Wolff et al., 2019

9. Krapp (1999b) verwendet eine gegenstandsbezogene Definition von intrinsisch anstelle einer tätigkeitsbezogenen Definition. Demzufolge kann auch das Streben nach Lernzuwachs intrinsisch sein, sofern die Schülerin oder der Schüler das Lernen als selbstintentional erfährt.

10. Die intrinsische Variable umfasst in dem Erwartungs-Wert-Modell affektive Facetten, jedoch keine kognitiv-epistemischen Aspekte, sodass das individuelle Interesse nicht vollständig unter dieser Variable subsumiert werden kann (s. Eccles & Wigfield, 2002; Lazarides et al., 2019).

Interpretation und Attribution der Erfahrungen bei Erfolg, kann die Leistung wiederum einen direkten positiven Einfluss auf die Entwicklung des Interesses haben (sowohl auf die gefühlsbezogene als auch auf die wertbezogene Valenz), weiterhin aber auch indirekt die wertbezogene Valenz des Interesses mediiert über das Selbstkonzept positiv beeinflussen (bspw. Sewasew et al., 2018). Dies kann dazu führen, dass Schülerinnen und Schüler sich als „Chemikerin“ bzw. „Chemiker“ erleben, sich mit den assoziierten Tätigkeiten identifizieren und diese mentale Relation in ihre Selbstdefinition integrieren (Hosbein & Barbera, 2020). Ebenso könnte der Verständniszuwachs dazu führen, dass Schülerinnen und Schüler vermehrt Unterrichtsinhalte als relevant erleben, Neugier und infolgedessen Interesse an diesen Inhalten und zugehörigen Tätigkeiten entwickeln.



**Abb. 2.4:** Reciprocal Effects Model (REM) für die Modellierung des Zusammenhangs zwischen dem individuellen Interesse und dem Konzeptverständnis. Die angedeutete Mediation des Effekts des Konzeptverständnisses auf das individuelle Interesse wurde in dieser Forschungsarbeit nicht untersucht.

Das synergistische Zusammenwirken von Interesse und Leistung äußert sich entsprechend dieser sich positiv verstärkenden Schleife in reziproken Effekten zwischen den beiden Konstrukten (s. Abbildung 2.4). Ebenso kann es aber bspw. im Falle von wiederholten Misserfolgen, negative Attributionen oder dimensionalen Vergleichen jederzeit zu einer Unterbrechung dieser Wirkungskette oder einer negativen Spirale aus sich abschwächendem Interesse und schwächeren Performanzen kommen.

Aus dem Modell der reziproken Effekte lassen sich Konsequenzen für die Unterrichtsgestaltung ableiten. Liegen positive reziproke Effekte zwischen dem Interesse an bestimmten Tätigkeiten sowie dem Konzeptverständnis vor, so könnten diese Tätigkeiten genutzt werden, um sowohl das Interesse als auch den Wissenserwerb simultan verstärkt zu fördern. Liegen innerhalb einer Dimension hingegen nur einseitig positive Effekte des Interesses an einer spezifischen Dimension auf das Konzeptverständnis vor, so würden sich diese Tätigkeiten für eine verstärkte Interessenförderung mit dem Ziel der Förderung des Wissenserwerbs anbieten. Liegen hingegen einseitig positive Effekte des Konzeptverständnisses auf das Interesse einer Dimension vor, so würden die Ergebnisse andeuten, dass das individuelle Interesse durch eine Fokussierung auf den kumulativen Wissenserwerb gestärkt werden könnte.

Alternative Modelle zur Verknüpfung von Interesse und Leistung. Neben dem Erwartungs-Wert-Modell gibt es alternative Modelle zur Beschreibung des Zusammenhangs zwischen dem Interesse und der Leistung. Zwei Beispiele, die diesen Zusammenhang direkt adressieren, sind das Model of Domain Learning sensu Alexander et al. (1995) und das Epistemic Model of Interest in Education sensu Rotgans und Schmidt (2017d) und H. G. Schmidt und Rotgans (2017). Das Model of Domain Learning beschreibt den Zusammenhang zwischen Leistung, situationalem und individuellem Interesse auf einem grobkörnigen Level über einen sehr großen Zeitraum anhand von drei Phasen, die primär durch den Grad der bereichsspezifischen Kompetenz gekennzeichnet sind. Rotgans und Schmidt (2017d) und H. G. Schmidt und Rotgans (2017) hingegen beschreiben in ihrem Epistemic

Model of Interest in Education den Zusammenhang zwischen Interesse und Leistung anhand relativ kurzer Zeitspannen, bspw. einer „Lerneinheit“ in der Schule und berücksichtigen dabei interessenfördernde Instruktionen der Lehrkraft. Das Modell kontuiert die unterschiedlichen Wirkungsweisen von situationalem und individuellem Interesse beim Wissenserwerb und beschreibt die Entwicklung des individuellen Interesses durch wiederholtes Hervorrufen des situationalen Interesses.

Im DoLiS-Projekt wurde lediglich das individuelle Interesse der Schülerinnen und Schüler erfasst. Eine komplexe Modellierung, die das situationale Interesse miteinbezieht, war somit nicht möglich, weswegen diese beiden Modelle in dieser Forschungsarbeit nicht zur Analyse des Zusammenhangs zwischen dem Interesse und der Leistung genutzt wurden.

## 2.5 Fazit und Hypothesen

In diesem Kapitel wird der theoretische Rahmen dieser Forschungsarbeit vorgestellt. Die Schwerpunkte lagen hierbei auf der Integration des RIASEC+N Modells in die aktuellen Interessentheorien sensu Krapp (2002b) und Hidi und Renninger (2006) sowie dem Zusammenspiel von individuellem Interesse und Fachwissen, wozu aus den Interessentheorien und dem Erwartungs-Wert-Modell ein Modell der reziproken Zusammenhänge deduziert wurde. Ausgehend von den diskutierten theoretischen Grundlagen und den empirischen Befund lassen sich für die drei zentralen Forschungsfragen Arbeitshypothesen formulieren, die nachstehend aufgeführt werden:

F1. Wie entwickeln sich die individuellen Interessen der Schülerinnen und Schüler im Fach Chemie über die Sekundarstufen I und II?

Bezüglich der Forschungsfrage F1, die die Entwicklung der individuellen Interessen der Schülerinnen und Schüler im Fach Chemie über die Sekundarstufen I und II thematisiert, wird eine Abnahme des individuellen Interesses angenommen, die zum Teil auf einer Interessendifferenzierung basieren kann. Die Abnahme der individuellen Interessen sollte alle sieben Interessendimensionen des RIASEC+N Modells sowie das Interesse an der Domäne Chemie betreffen und vorwiegend zu Beginn des Chemieunterrichts zur Entfaltung kommen.

F2. Wirkt sich ein elaboriertes Konzeptverständnis positiv auf die Entwicklung des individuellen Interesses aus?

F3. Wirkt sich das individuelle Interesse positiv auf den Erwerb eines elaborierten Konzeptverständnisses aus?

Bezüglich der Forschungsfrage F2 und F3, die den Zusammenhang zwischen dem individuellen Interesse und dem Konzeptverständnis thematisieren, wird grundsätzlich ein reziproker Zusammenhang (REM) zwischen Interesse und Leistung angenommen. Dieser positive Zusammenhang sollte sich in Cross-Lagged-Panel-Analysen in positiven, beidseitigen kreuzverzögerten Pfaden zwischen dem individuellen Interesse an der Domäne Chemie und dem Konzeptverständnis manifestieren.

Neben dem individuellen Interesse an der Domäne Chemie wurde in dieser Forschungsarbeit das Interesse mittels des RIASEC+N Modells erhoben, das das individuelle Interesse der Schülerinnen und Schüler an verschiedenen naturwissenschaftlichen Unterrichtstätigkeiten erfasst. Auch für das individuelle Interesse an naturwissenschaftlichen Tätigkeiten und dem Kompetenzerwerb in den avisierten Tätigkeitsbereichen der jeweiligen RIASEC+N Dimensionen werden beidseitige, positive Zusammenhänge angenommen. Jedoch sind die sieben Interessendimensionen, zwischen denen das RIASEC+N Modell differenziert, durch Tätigkeiten gekennzeichnet, die das fachliche Lernen zu unterschiedlichen Anteilen adressieren und auf den Erwerb verschiedener Kompetenzen (bspw. Fachwissen, Erkenntnisgewinnung, Kommunikation, Bewertung; s. KMK, 2004) fokussieren. So steht bei den durch die Dimensionen investigative und networking beschriebenen Tätigkeiten das Problemlösen im Vordergrund, welches mit Lernzuwächsen assoziiert ist. Daher sollte für diese Dimensionen ein bedeutsamer Zusammenhang zwischen dem tätigkeitsspezifischen individuellen

Interesse und dem Konzeptverständnis vorliegen. Die Dimensionen social und enterprising hingegen rücken zusätzlich die Kommunikationskompetenz in den Vordergrund, die Dimension artistic spricht eher Tätigkeiten an, die zumeist nicht vorwiegend mit dem Lernziel des Fachwissenserwerbs verknüpft sind. So kann das Zeichnen von chemischen Apparaturen auch aus einer künstlerischen Perspektive als interessant erachtet werden, ohne dass eine Auseinandersetzung mit den fachlichen Inhalten stattfindet. Die Dimension realistic fokussiert hands-on Tätigkeiten, die stark experimentelle Fähigkeiten betonen. Diese Tätigkeiten sind prinzipiell für den Wissenserwerb geeignet, jedoch kann dieser, wenn er nicht explizit durch Instruktionen angeregt wird, gegenüber der experimentellen Durchführung in den Hintergrund treten (Abrahams & Reiss, 2012). Im Mittel sollten daher kreuzverzögerte Effekte zwischen dem Interesse der Dimension realistic und dem Konzeptverständnis schwach ausfallen.

Zusammenfassend lässt sich sagen, dass vorwiegend die Interessendimensionen investigative und networking direkt mit Tätigkeiten, die das Lösen von Lernaufgaben beschreiben, verknüpft sind, und hier der positive Zusammenhang zwischen dem Interesse und dem Konzeptverständnis am stärksten hervortreten sollte.

#### Zusammenhänge des Interesses zu Kovariaten

Im Zuge der Forschungsarbeit wurde weiteren Fragestellungen nachgegangen, die u. a. die Zusammenhänge des Interesses zu Kovariaten adressieren. Hierbei sind vorwiegend die Fragestellungen nach Zusammenhängen zu der Lernfreude und der wahrgenommenen Nützlichkeit hervorzuheben.

**Lernfreude.** Zwischen der Lernfreude im Chemieunterricht und dem individuellen Interesse werden positive Zusammenhänge angenommen. Die Arbeitshypothese lautet, dass Schülerinnen und Schüler mit einer höheren Lernfreude eher dazu geneigt sind, sich selbstbestimmt in ihrer Freizeit mit Chemiekontexten auseinanderzusetzen und dabei ein stabiles Interesse entwickeln könnten, insbesondere wenn im Unterricht Alltagsbezüge aufgezeigt werden. Ebenso erschien es plausibel, dass Schülerinnen und Schüler, die sich in ihrer Freizeit mit Chemiekontexten auseinandersetzen, ein höheres individuelles Interesse aufweisen und den Chemieunterricht ebenfalls eher als freudvoll erfahren, da sie hier ihren Neigungen nachgehen und diese vertiefen können.

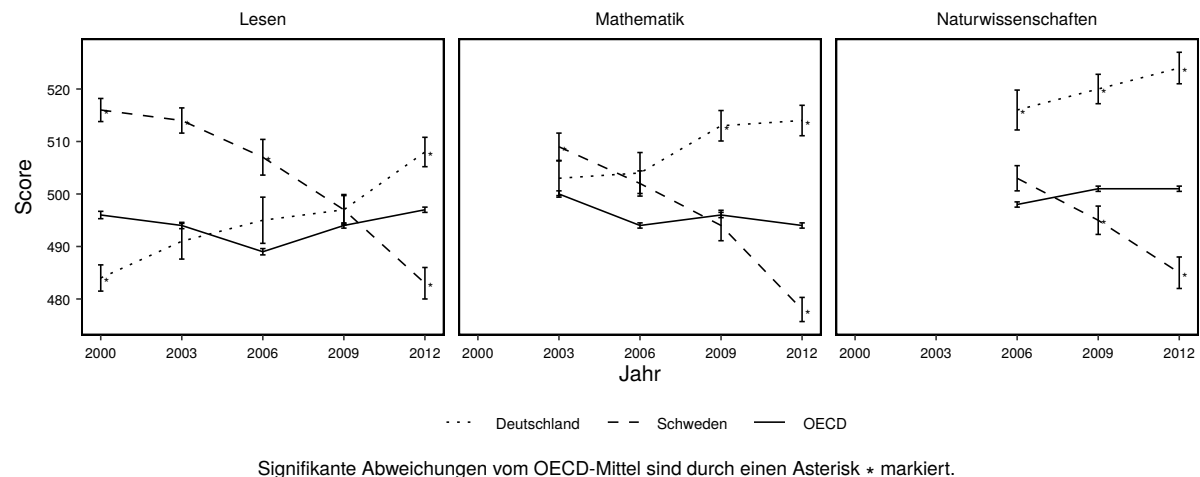
**Wahrgenommene Nützlichkeit.** Zwischen dem individuellen Interesse und der wahrgenommenen Nützlichkeit werden positive Zusammenhänge angenommen. Zum einen kann durch die Steigerung der wahrgenommenen Relevanz und Nützlichkeit der Fachinhalte zur Realisierung kurzfristiger und langfristiger Ziele das Interesse gefördert werden (bspw. Hulleman et al., 2010), zum anderen zeigen empirische Befunde ebenso positive Effekte der intrinsischen Wertkomponente auf die wahrgenommene Nützlichkeit (Lauermaun et al., 2017).





## Studiendesign

Die Grundlage der empirischen Arbeiten dieses Forschungsvorhabens bilden Daten des binationalen DoLiS-Projekts, das Erhebungen mit vergleichbaren Testinstrumenten und Samples in Schweden und Deutschland vorsah. Kern des DoLiS-Projekts war auf deutscher Seite, auf die sich diese Arbeit kapriziert, eine Kombination aus einer quer- und längsschnittlichen Fragebogenerhebung.



**Abb. 3.1:** Änderungen in der mittleren Lesekompetenz, der mathematischen und der naturwissenschaftlichen Kompetenzen schwedischer sowie deutscher Schülerinnen und Schüler in den OECD PISA-Studien von 2006 – 2012 ausgehend von der jeweilig ersten Testung als Untersuchungsschwerpunkt der jeweiligen Erhebung (OECD, 2007, 2011, 2014).

**Studienziele.** Das DoLiS-Projekt verfolgte das Ziel mithilfe eines binationalen, längsschnittlichen Studiendesigns die Entwicklungsverläufe von Schülerinnen und Schülern im Chemieunterricht der Sekundarstufen in Schweden und Deutschland (Gelegenheitsstichprobe) zu charakterisieren, nationale Befunde zwischen den beiden Stichproben zu vergleichen und mit unterrichtsbezogenen sowie strukturellen Unterschieden in Beziehung zu setzen. Motiviert ist das binationale Design des DoLiS-Projekts u. a. durch divergierende Trends in internationalen Vergleichsstudien wie der PISA Studie (s. Abbildung 3.1 für Trends in den gemessenen Kompetenzen).

Während die regulären PISA Ergebnisse die Naturwissenschaften im Allgemeinen fokussieren, nur Vergleiche zwischen Kohorten erlauben und somit keine Aufschlüsse über kausale Zusammenhänge liefern können, ist die Intention des DoLiS-Projekts, einen genaueren Einblick in die Entwicklung von chemiebezogenen kognitiven (Konzeptverständnis) sowie motivationalen Personenmerkmalen (u. a. Interesse, wahrgenommene Nützlichkeit) zu gewinnen, um Zusammenhänge über die Zeit und Trends aufzuzeigen. Die im DoLiS-Projekt gewonnenen Daten könnten somit dazu beitragen, die Entwicklung von chemiebezogenen Personenmerkmalen bis zum Schulabschluss aufzuzeigen, Diskrepanzen zwischen Personenmerkmalen schwedischer und deutscher Schülerinnen und Schüler zu erklären.

**Studiendesign und Durchführung.** Das Studiendesign des DoLiS-Projekts ist eine Kombination aus einer Querschnitt- und einer 2-Kohorten-Längsschnittstudie. Im Jahr 2015 fanden in Deutschland und Schweden die querschnittlichen Erhebungen in den Jahrgangsstufen 5 bis 11 bzw. 12 statt. Ausgehend von der querschnittlichen Erhebung wurden zudem längsschnittliche Daten erfasst. Hierzu wurden die Schülerinnen und Schüler der fünften (Kohorte LS5) und neunten Klassen (Kohorte LS9) der Querschnitterhebung über einen Zeitraum von drei Jahren wiederholt im jährlichen Abstand

getestet, infolgedessen die Schülerinnen und Schüler in Deutschland idealiter an vier Erhebungen teilnahmen. Auch Schülerinnen und Schüler, die sich entschieden, den Chemieunterricht in der Oberstufe abzuwählen, wurden wiederholt befragt. An der Studie nahmen in Deutschland sechs Gymnasien und zwei Gemeinschaftsschulen teil. Ein Gymnasium partizipierte an der Studie mit einer Verzögerung von einem Jahr und eine Gesamtschule stieg nach dem zweiten Erhebungszeitpunkt aus der Studie aus. Insgesamt konnten  $N = 6555$  Schülerinnen und Schüler befragt werden ( $N_{\text{SWE}} = 2959$ ;  $N_{\text{GER}} = 3596$ ), wobei in Deutschland die längsschnittliche Kohorte der Sekundarstufe I (LS5)  $N = 724$ , die Kohorte der Sekundarstufe II (LS9)  $N = 769$  Gymnasiastinnen und Gymnasiasten umfasst. Eine genauere Auflistung der Probandenzahlen ist in Tabelle 3.1 ersichtlich. Die Tabelle gibt zudem einen Überblick darüber, wie viele Schülerinnen und Schüler des deutschen Subsamples zur Zeit der Erhebung Chemieunterricht erhielten. Hier ist abzulesen, dass über ein Drittel der Schülerinnen und Schüler, die in Jahrgangsstufe 12 getestet wurde, Chemie bereits ausgewählt hatte. Da durch die Abwahl u. a. schulische Lerngelegenheiten für chemisches Fachwissen stark reduziert wurden, muss das Abwahlverhalten der Schülerinnen und Schüler in den Datenanalysen berücksichtigt werden. Um eine Abschätzung der Abwahl des Faches Chemie auf die Analyseergebnisse treffen zu können, werden in Kapitel 4 differenzierte Voruntersuchen dargestellt.

**Tab. 3.1:** Samplezusammensetzung des DoLiS-Projekts je Messzeitpunkt und Land. Für das deutsche Sample sind in Klammern die Anteile der Schülerinnen und Schüler mit Chemieunterricht angegeben.

Messzeitpunkt	Jahrgangsstufe							
	5	6	7	8	9	10	11	12
Schweden								
1	219	244	401	326	536	318	244	
2		276				495		
3			286				304	
Gesamt	219	520	687	326	536	813	548	
Deutschland								
1	614 (0)	374 (0)	299 (0.74)	309 (1)	571 (1)	332 (0.80)	514 (0.66)	126 (0.61)
2		524 (0.21)				543 (0.88)		
3			532 (0.85)				484 (0.73)	
4				447 (1)				386 (0.62)
Gesamt	614 (0)	898 (0.13)	831 (0.81)	756 (1)	571 (1)	875 (0.85)	998 (0.69)	512 (0.62)

Alle Daten des DoLiS-Projekts wurden mittels Fragebögen erhoben und reflektieren die subjektive Wahrnehmung der Schülerinnen und Schüler. Zur Gewinnung der Daten wurden drei Fragebögen eingesetzt:

**Fragebogen I** Der erste Fragebogen bestand aus einem Fachtest zum Konzeptverständnis der für die Chemie zentralen Konzepte Materie, Chemische Reaktion und Energie in Anlehnung an die Basiskonzepte der Bildungsstandards (KMK, 2004). Jedem Proband lagen 30 geschlossene Aufgaben mit vier Antwortoptionen zur Beantwortung vor, von denen lediglich eine richtig ist. Der Fachtest umfasst sowohl klassische Multiple-Choice-Aufgaben als auch Ordered-Multiple-Choice-Aufgaben. Um einen Vergleich zwischen den Jahrgangsstufen herstellen zu können, waren nur zwölf Aufgaben (vier je Basiskonzept) jahrgangsstufenspezifisch. Die restlichen 18 Aufgaben (fünf Aufgaben je Basiskonzept) waren in allen Fachtests identisch (Ankeraufgaben). Insgesamt bestand die Testbatterie somit aus 114 Aufgaben. Die Fachtests wurden über alle vier Erhebungszeitpunkte gleich gehalten, lediglich Rechtschreibfehler wurden behoben.

Zur Beantwortung des Fachtests hatten die Schülerinnen und Schüler 45 Minuten Zeit (Dauer einer typischen Schulstunde).

**Fragebogen II** Der zweite Fragebogen erfasste demographische Daten sowie sozial-kognitive Variablen wie motivationale, emotionale und kognitive Personenmerkmale. Vorwiegend wurden die Daten mittels Likert-Skalen erfasst. Die eingesetzten Skalen blieben überwiegend über alle vier Erhebungszeitpunkte konstant, um längsschnittliche Entwicklungen abbilden zu können. Zudem

wurden Teile des Fragebogens je Erhebungszeitpunkt variiert, um jahrgangsstufenspezifische Aspekte wie Wahlentscheidungen und Berufsaspirationen analysieren zu können. Zur Beantwortung des Fachtests hatten die Schülerinnen und Schüler 45 Minuten Zeit (Dauer einer typischen Schulstunde).

**Fragebogen III** Der dritte Fragebogen konstituierte sich aus der figuralen Subskala des kognitiven Fähigkeitstests (KFT; Heller und Perleth, 2000). Dieser Fragebogen wurde nur zum ersten Erhebungszeitpunkt eingesetzt. Die Bearbeitungszeit betrug acht Minuten.

Die Testdurchführung erfolgte durch geschulte Mitarbeiterinnen und Mitarbeiter des Forschungsinstituts im Klassenverbund. Der Erhebungszeitpunkt war typischerweise im Frühjahr eines Kalenderjahres. Die Partizipation an der Studie durch die Schülerinnen und Schüler geschah auf freiwilliger Basis. Ein Ausstieg aus der Studie war zu jedem Zeitpunkt möglich.



## Teil II

### Empirischer Teil



## Übersicht über die empirischen Untersuchungen dieser Forschungsarbeit

In diesem Abschnitt werden die empirischen Untersuchungen des Forschungsvorhabens vorgestellt. Diese bestehen aus umfassenden Voruntersuchungen sowie vier eigenständigen Fachpublikationen. Um Verknüpfungen zwischen den Publikationen zu erleichtern, die Orientierung zu fördern und die Lesbarkeit zu erhöhen, wird einleitend eine kurze Zusammenfassung der Voruntersuchungen sowie der vier Publikationen vorgenommen.

### **Voruntersuchungen** Das RIASEC+N Modell: Deskriptive Befunde und Validierung

Für die vier empirischen Publikationen (Publikation 1–4) wurde in den Voruntersuchungen die Struktur des RIASEC+N Modells mittels konfirmatorischer Faktorenanalysen geprüft, Auswirkungen der Abwahl des Faches Chemie auf die Interessenentwicklung analysiert und das Vorliegen von latenten Klassen mit charakteristischen Interessenprofilen mittels latenten Profilanalysen geprüft. Die Voruntersuchungen deuten auf das Vorliegen von sieben positiv korrelierten Interessendimensionen hin und indizieren einen bedeutsamen Zusammenhang zwischen dem Interesse und dem Wahlverhalten der Schülerinnen und Schüler. Die latenten Profilanalysen weisen auf vier latente Klassen hin, die sich lediglich durch gleichgerichtete Verschiebungen der mittleren Interessenausprägungen auszeichnen, sodass für diesen Auswertungsansatz kein substantieller Mehrertrag gegenüber variablenzentrierten Verfahren gesehen wurde.

**Publikation 1.** Knowing more about things you care less about: Cross-sectional analysis of the opposing trend and interplay between conceptual understanding and interest in secondary school chemistry.

Die querschnittliche Studie untersucht Zusammenhänge zwischen dem Interesse an naturwissenschaftlichen Tätigkeiten (RIASEC+N) und dem Konzeptverständnis. Hier wurde zunächst mittels einer konfirmatorischen Multi-Gruppen Faktorenanalyse (MG-CFA) die Struktur des RIASEC+N Modells geprüft und anschließend ein manifester Score als Indikator für das Konzeptverständnis in das Modell aufgenommen. Die Resultate indizierten, dass die Struktur des Interesses an naturwissenschaftlichen Schultätigkeiten über die Sekundarstufen invariant ist und erwartungskonform das Interesse in allen sieben Dimensionen abnimmt, wohingegen das Konzeptverständnis von Klassenstufe 5 bis 11 zunimmt. Hinsichtlich korrelativer Zusammenhänge zwischen dem Interesse und der Leistung konnte gezeigt werden, dass das Interesse an Tätigkeiten, die potentiell mit einem hohen Grad der kognitiven Aktivierung verknüpft sind (investigative, networking) oder die Kommunikation von Wissen beinhalten (social, enterprising und networking), zunehmend mit dem Konzeptverständnis assoziiert sind. Dieser Befund traf insbesondere für Schülerinnen und Schüler in der Oberstufe zu. Außerdem konnte gezeigt werden, dass das Interesse an „hands-on“ Tätigkeiten (realistic), welche als typisch für den Chemieunterricht gelten, nur geringfügige Korrelationen zum Konzeptverständnis aufwies und das Interesse an artistic und conventional Tätigkeiten nicht bedeutsam mit dem Konzeptverständnis zusammenhängt.

Höft, L., Bernholt, S., Blankenburg, J. S. & Winberg, M. (2019). Knowing more about things you care less about: Cross-sectional analysis of the opposing trend and interplay between conceptual understanding and interest in secondary school chemistry. *Journal of Research in Science Teaching*, 56(2), 184–210

**Publikation 2.** Longitudinal couplings between interest and conceptual understanding in secondary school chemistry: An activity-based perspective.

Die längsschnittliche Studie untersucht die Entwicklung des Interesses an naturwissenschaftlichen Tätigkeiten (RIASEC+N) und des Konzeptverständnisses in den Jahrgangsstufen 9 bis 11 sowie das Zusammenspiel beider Konstrukte über die Zeit. Hinsichtlich der Entwicklung der beiden Konstrukte kann ein Interessenabfall bei gleichzeitigem Lernzuwachs gezeigt werden. In Hinblick auf Zusammenhänge zwischen dem Interesse und dem Konzeptverständnis (Prüfung des Modells der reziproken Effekte) indizieren die Analysen geringe reziproke Zusammenhänge zwischen

dem Interesse an investigative, social und networking Tätigkeiten und dem Konzeptverständnis. Zudem war das Interesse an enterprising Aktivitäten prädiktiv für das spätere Konzeptverständnis. Die übrigen drei Interessendimensionen wiesen keine systematischen Zusammenhänge zum Konzeptverständnis auf. Zusammengefasst deuten die Resultate daraufhin, dass Tätigkeiten, die mit kognitiv herausfordernden Aufgaben assoziiert sind, die Chance bieten, das Interesse und das Konzeptverständnis der Schülerinnen und Schüler simultan zu fördern.

Höft, L. & Bernholt, S. (2019b). Longitudinal couplings between interest and conceptual understanding in secondary school chemistry: An activity-based perspective. *International Journal of Science Education*, 19(4), 1–21

**Publikation 3.** Probing Precursors of Students' Long-term Engagement in Chemistry: Individual Interest and Achievement Development and their Relation with Utility Values.

Die längsschnittliche Studie untersucht Zusammenhänge zwischen dem Interesse an investigative, social und networking Tätigkeiten und dem Konzeptverständnis in der Sekundarstufe I (Jahrgangsstufen 6–8) und II (Jahrgangsstufen 9–11). Hierzu wurden mittels bivariater Latent-Change-Score-Modelle kreuzverzögerte Zusammenhänge zwischen dem Interesse und der Änderung des Konzeptverständnisses, dem Konzeptverständnis und der Änderung des Interesses untersucht und zudem Zusammenhänge zu den subjektiven Wert der Chemie für die Schülerinnen und Schüler analysiert (Prüfung des Modells der reziproken Effekte). Die Befunde deuten darauf hin, dass in der Sekundarstufe I nur geringe Zusammenhänge zwischen dem Interesse und dem Konzeptverständniserwerb sowie zwischen dem Konzeptverständnis und Interessenänderungen bestehen. In der Sekundarstufe II liegt ein ausgeprägteres Befundmuster vor. So finden sich interessengetriebene Lernzuwächse sowie reziproke Zusammenhänge zwischen dem Interesse an networking Tätigkeiten und dem Konzeptverständnis.

Höft, L. & Bernholt, S. (in Rev.). Probing Precursors of Students' Long-term Engagement in Chemistry: Individual Interest and Achievement Development and Their Relation with Utility Values.

**Publikation 4.** Ich mag, was ich kann oder kann ich, was ich mag? Über das Zusammenspiel von Interesse, Freude und Konzeptwissen im Fach Chemie.

Die längsschnittliche Studie untersucht anhand von Gymnasiastinnen und Gymnasiasten in der Sekundarstufe II, inwieweit sich die Freude am Chemieunterricht, das Konzeptwissen und das individuelle Interesse (Prüfung des Modells der reziproken Effekte) an der Domäne Chemie gegenseitig beeinflussen. Hinsichtlich des Zusammenhangs zwischen den drei Konstrukten deuten die bivariaten Analysen auf beidseitige Wechselwirkungen zwischen den drei Konstrukten hin. Bei gleichzeitiger Berücksichtigung aller drei Konstrukte mittels trivariater Analysen zeigen sich weiterhin reziproke Zusammenhänge zwischen der Freude und dem Konzeptverständniserwerb bzw. dem Konzeptverständnis und der Interessenänderung, jedoch nur einseitig positive Effekte der Freude und des Konzeptverständnisses auf die Veränderung des individuellen Interesses. Die Befunde unterstreichen die Relevanz motivationaler Variablen für das Lernen und weisen auf Potentiale einer gezielteren Integration von Interessen der Schülerinnen und Schüler in den Chemieunterricht für die Förderung des Interesses hin.

Höft, L. & Bernholt, S. (2019a). Ich mag, was ich kann oder kann ich, was ich mag? Über das Zusammenspiel von Interesse, Freude und Konzeptwissen im Fach Chemie. *Zeitschrift für Didaktik der Naturwissenschaften*, 25(1), 161–180



## Das RIASEC+N Modell: Deskriptive Befunde und Validierung

In diesem Kapitel wird eine deskriptive Datenanalyse der RIASEC+N Interessenskalen vorgenommen und zusätzliche Evidenz für die Validität des Interessenkonstrukts gewonnen. Diese zusätzliche statistische Beschreibung der erhobenen Daten soll die Verknüpfung der einzelnen Publikationen unterstützen, eine fundierte Bewertung der gefundenen Ergebnisse fördern sowie der Kohärenz und Vollständigkeit der Gesamtdiskussion dienen. Hierzu werden zunächst konfirmatorische Faktorenanalysen zur Prüfung der Struktur des Interessenkonstrukts dargestellt (vgl. Ergebnisse der konfirmatorischen Faktorenanalysen in Kapitel 5). Flankiert werden die konfirmatorischen Faktorenanalysen durch deskriptive Maße der querschnittlichen und längsschnittlichen Daten. Hier werden Reliabilitäten und Mittelwertsverläufe berichtet. Abschließend wird querschnittlich die Prädiktion von Interessenprofilen durch die Jahrgangsstufe und das Geschlecht analysiert. Zur Bestimmung der Interessenprofile wurde als Methode die latente Profilanalyse gewählt, welche eine personenzentrierte Perspektive einnimmt und versucht, Subpopulationen aufgrund von Antwortmustern, in diesem Fall Interessenausprägungen der Probanden, empirisch abzubilden.

### 4.1 Konfirmatorische Faktorenanalyse

In diesem Abschnitt wird die theoretisch erwartete Dimensionalität des RIASEC+N Modells mittels konfirmatorischer Faktorenanalyse (CFA, Fehlerkorrektur mittels „Type Is Complex“ Option in Mplus) geprüft. Die Analysen basieren auf den Daten der 28 durchgehend eingesetzten Items zur Erfassung des Interesses an naturwissenschaftlichen Tätigkeiten (vier Items je Skala). Aus der Gesamtstichprobe des DoLiS-Projekts wurde ein Subsample gezogen, das nur Schülerinnen und Schüler der Querschnitterhebungen umfasst, die den Chemieunterricht zum Zeitpunkt der Testung nicht abgewählt hatten. Es werden im Folgenden sechs plausible Modelle gegenübergestellt und verglichen. Hierbei handelt es sich um

- M1. ein General-Faktor-Modell,
- M2. ein Zwei-Faktoren-Modell (Einzelarbeit und Partner/Gruppenarbeit),
- M3. ein Sieben-Faktoren-Modell (RIASEC+N),
- M4. ein Higher-Order-Modell (RIASEC+N mit Higher-Order-Factor),
- M5. ein Bi-Faktor-Modell mit unkorrelierten spezifischen Faktoren (RIASEC+N und General-Faktor) sowie
- M6. ein Bi-Faktor-Modell mit korrelierten spezifischen Faktoren (RIASEC+N und General-Faktor).

Die Modelle M1 und M2 sind einfache, generische Modelle mit einem bzw. zwei latenten Faktoren. Die Modelle M3 – M6 gehen von der bereits in vorherigen Studien gefundenen RIASEC+N Struktur mit sieben korrelierten, latenten Faktoren aus. Zusätzlich zu diesen sieben latenten Faktoren, im nachfolgenden Text auch Interessendimensionen genannt, nehmen die Modelle M4 – M6 noch einen zusätzlichen achten Faktor an, welcher das tätigkeitsunspezifische Chemiefachinteresse abbilden soll.

#### Generische Modelle M1 und M2

In dem General-Faktor-Modell (M1) laden alle 28 Items auf einem einzelnen latenten Faktor, der das Interesse der Schülerinnen und Schüler an naturwissenschaftlichen Schultätigkeiten als ein eindimensionales, komplexes Konstrukt annimmt. Die Varianz in den Antworten kann demzufolge alleine durch den General-Faktor sowie die Residualvarianz der einzelnen Items erklärt werden. Würde dieses Modell die Daten adäquat beschreiben, könnte dies bspw. dafür sprechen, dass Schülerinnen und Schüler die beschriebenen Tätigkeiten nur schwer diskriminieren können, die

Zusammenhänge zwischen den Interessen an den angenommenen Tätigkeiten sehr hoch korreliert sind, weswegen die Befunde als fachspezifisches Interesse gedeutet werden könnten.

In dem Zwei-Faktoren-Modell (M2) laden die Items auf einem der zwei spezifischen, miteinander korrelierten Faktoren. Das Modell differenziert aufgrund der in den Itemformulierungen konnotierten Sozialform zwischen zwei latenten Faktoren. Die Interessenitems der Tätigkeiten realistic, investigative, artistic und conventional laden zusammen auf einem latenten Faktor, der Tätigkeiten in der Sozialform Einzelarbeit beschreibt (Die Items der Dimension realistic wurden diesem latenten Faktor zugeordnet, da sie zwar hands-on Tätigkeiten adressieren, in den Itemformulierungen aber keine Mitschülerinnen und Mitschüler aufgenommen wurden.). Die Items der Tätigkeiten social, enterprising und networking laden auf dem zweiten latenten Faktor, der Tätigkeiten der Sozialformen Partner- und Gruppenarbeit beschreibt. Dem Modell liegt somit die Annahme zugrunde, dass das Interesse an naturwissenschaftlichen Tätigkeiten nicht durch die Tätigkeit per se, sondern durch die Interaktion mit den Mitschülerinnen und Mitschülern geprägt ist.

#### Generische Modelle M4 bis M6

Diese Modellannahmen der Modelle M4 – M6 gehen auf explorative (Dierks et al., 2014b) sowie konfirmatorische Forschungsergebnisse zurück (Blankenburg, Höffler & Parchmann, 2016).

In dem Sieben-Faktoren-Modell (M3) laden die Items auf einem der sieben spezifischen, miteinander korrelierten Faktoren. Das Modell unterscheidet zwischen sieben latenten Faktoren, die das Interesse an differenzierbaren Tätigkeitsbereichen beschreiben. Die Annahme des Modells ist, dass Schülerinnen und Schüler gut zwischen den postulierten Tätigkeitsbereichen des Chemieunterrichts unterscheiden können und während der Ausübungen dieser voneinander abgrenzbare Ausprägungen des situationalen Interesses empfinden, infolgedessen sich intraindividuelle Interessendifferenzen ausbilden, die sich in interindividuellen Interessenprofilen manifestieren sollten. So könnte es bspw. Schülerinnen und Schüler geben, die sich vorwiegend für investigative Tätigkeiten interessieren, social Tätigkeiten aber eine geringere Attraktivität zuschreiben.

Das Higher-Order-Modell (M4) greift die Struktur des Sieben-Faktoren-Modells auf. Anstelle der Korrelationen zwischen den sieben spezifischen latenten Faktoren tritt ein Faktor höherer Ordnung, der sich aus den Varianzen der sieben spezifischen Faktoren berechnet und somit die Korrelationen zwischen diesen repräsentiert. Es liegt dementsprechend ein hierarchisches Modell mit sieben tätigkeitsspezifischen Faktoren niedriger Ordnung und ein allgemeiner Faktor höherer Ordnung vor. Die Annahme hinter dem Modell ist, dass sich aus den tätigkeitsspezifischen Interessen der Schülerinnen und Schüler ein allgemeines Interesse an dem Unterrichtsfach Chemie entwickelt, wobei alle sieben Dimensionen substantiell zu dem erlebten Fachinteresse beitragen.

Die Bi-Faktor-Modelle M5 und M6 nehmen sieben tätigkeitsspezifische latente Faktoren sowie einen hierzu orthogonalen General-Faktor an. Jedes Item lädt in diesen Modellen auf zwei Faktoren; auf einem tätigkeitsspezifischen Faktor und auf dem General-Faktor. Das Modell M5 spezifiziert keine Korrelationen zwischen den tätigkeitsspezifischen Faktoren, wodurch die Varianz in additive Komponenten zerlegt wird. Das Modell M6 erlaubt Korrelationen zwischen den tätigkeitsspezifischen Faktoren. Die Interpretation der Faktoren kann ähnlich zu dem Higher-Order-Modell (M4) erfolgen.

Die Passung der Modelle M1 – M3 sowie die Passung des Higher-Order-Modells (M4) und der Bi-Faktor-Modelle (M5, M6) auf die Daten kann direkt mittels  $\chi^2$ -Test miteinander verglichen werden, da diese Modelle ineinander verschachtelt (nested models) sind (vgl. L. M. Daniels et al., 2008). Der Vergleich zu den restlichen Modellen kann über gängige Fit-Indizes (CFI, RMSEA, SRMR) erfolgen. Die Modelle wurden, um der Mehrebenenstruktur der Daten sowie der Testung in unterschiedlichen Jahrgangsstufen Rechnung zu tragen, als CFAs mit der Option „Type Is Complex“ in Mplus gerechnet. Für die Schätzung der Bi-Faktor-Modelle musste die Anzahl der zu schätzenden Parameter reduziert werden, was durch die Gleichsetzung der Faktorladungen je Interessenfaktor (RIASEC+N) bei variablen Faktorenladungen auf den General-Faktor erreicht wurde. Gängige Gütekriterien zur Bewertung der CFAs sind in Tabelle 4.1 gelistet.

Die Fit-Indizes deuten darauf hin, dass nur das Sieben-Faktoren- (M3) sowie das Bi-Faktor-Modell

**Tab. 4.1:** Fit-Statistik der konfirmatorischen Faktorenanalysen.

Modell	Label	$\chi^2$	$\chi^2_{DF}$	$\chi^2_p$	CFI	RMSEA	SRMR
M1	General-Faktor-Modell	4051.03	189	.000	0.696	0.115	0.086
M2	Zwei-Faktoren-Modell	3759.41	188	.000	0.719	0.111	0.085
M3	Sieben-Faktoren-Modell	1278.39	168	.000	0.913	0.065	0.061
M4	Higher-Order-Modell	1613.70	183	.000	0.887	0.071	0.072
M5	Bi-Faktor-Modell (unkorrelierte spezifische Faktoren)	1734.21	182	.000	0.878	0.074	0.071
M6	Bi-Faktor-Modell (korrelierte spezifische Faktoren)	1150.61	161	.000	0.922	0.063	0.053

mit korrelierten Faktoren (M6) die Daten adäquat beschreiben (Hu & Bentler, 1999). Gegen das Bi-Faktor-Modell mit korrelierten Faktoren spricht, dass dieses negative Faktorenladungen von Items, die realistic, artistic und conventional Tätigkeiten nennen, auf den General-Faktor enthält ( $\lambda_G = [-.511, .375]$ ). Hierdurch ist eine plausible Interpretation des General-Faktors als allgemeines Interesse am Fachunterricht impraktikabel. Aus diesem Grund wurde das Bi-Faktor-Modell mit korrelierten Faktoren (M6) für die folgenden Analysen trotz der guten Fit-Werte verworfen und stattdessen auf das Sieben-Faktoren-Modell (M3) zurückgegriffen. Die Interkorrelation zwischen den sieben latenten Faktoren liegt im Mittel bei  $\bar{r} = .64$ , wobei insbesondere die Dimension investigative hoch mit den Dimensionen realistic ( $r = .89$ ,  $p < .001$ ) und networking ( $r = .86$ ,  $p < .001$ ) korreliert (s. Tabelle 4.2). Die Dimension artistic weist im Mittel die kleinsten Korrelationskoeffizienten auf.

**Tab. 4.2:** Interkorrelationen zwischen den sieben RIASEC+N Interessendimensionen des Sieben-Faktoren-Modells (M3). Alle Korrelationen sind hoch signifikant ( $p < .001$ ).

Latenter Faktor	R	I	A	S	E	C
Realistic (R)						
Investigative (I)	.89					
Artistic (A)	.43	.38				
Social (S)	.38	.61	.30			
Enterprising (E)	.78	.85	.35	.57		
Conventional (C)	.45	.69	.52	.57	.62	
Networking (N)	.66	.86	.40	.70	.79	.79

Eine genauere Analyse der Struktur des Interesses der Schülerinnen und Schüler an naturwissenschaftlichen Tätigkeiten auf Grundlage des Sieben-Faktoren-Modells wird in Publikation 1 (Höft et al., 2019) vorgenommen. Hier werden die querschnittlichen Daten mithilfe einer Multi-Group CFA untersucht, wobei die einzelnen Gruppen den Jahrgangsstufen 5–12 entsprechen. Dieses Verfahren ermöglicht es, Trends in den mittleren Interessenausprägungen aufzudecken und jahrgangsspezifische Korrelationen zu schätzen.

Die Befunde deuten auf diskriminierbare Interessendimensionen hin und motivieren eine differenzierte Analyse der Entwicklung des Interesses an naturwissenschaftlichen Tätigkeiten. Hier erscheinen nicht nur Analysen mit inkludierten Kovariaten gewinnbringend, auch die Analyse von Interessenprofilen oder der Konvergenz bzw. Divergenz der Interessendimensionen könnten fruchtbar sein, um präzisere Informationen über Stellschrauben für interesselördernde Interventionen gewinnen zu können.

## 4.2 Deskriptive Statistik

Zur Einschätzung der psychometrischen Güte des eingesetzten RIASEC+N Instruments werden im Folgenden als Ausgangspunkt Reliabilitäten und deskriptive Maße berichtet; eine detaillierte Auflistung der Kennwerte befindet sich im Anhang (s. Tabelle A.1). Als Grundlage für die nachfolgenden Analysen dienen ebenfalls die 28 durchgehend eingesetzten Items zur Erfassung des Interesses an naturwissenschaftlichen Tätigkeiten (vier Items je Skala). Aus der Gesamtstichprobe des DoLiS-Projekts wurde ein Subsample gezogen, das nur Schülerinnen und Schüler der Quer-

sowie der Längsschnitterhebungen umfasst, die den regulären Chemieunterricht zum Zeitpunkt der Testung nicht abgewählt hatten. Weitergehend werden schließlich basierend auf dem reduzierten längsschnittlichen Datensatz die mittleren Interessenausprägungen von Schülerinnen und Schülern der Oberstufe, die das Fach Chemie während des Testzeitraums abwählten, gegenübergestellt. Der Vergleich von Schülerinnen und Schülern der Sekundarstufe II mit und ohne Chemieunterricht ermöglicht es, den Einfluss der Zusammensetzung des Subsamples auf die gefundenen Ergebnisse abzuschätzen und liefert somit wichtige Hinweise für weiterführende längsschnittliche Analysen der Publikationen 2–4.

### Reliabilitäten

Die Reliabilität der Skalen wurde mittels Cronbachs  $\alpha$  und McDonalds  $\omega$  bestimmt, wobei beide Kennwerte zumeist auf zufriedenstellende bis gute Messgenauigkeiten hindeuteten (Cronbachs  $\alpha$ :  $\alpha_R = [.64, .77]$ ,  $\alpha_I = [.73, .87]$ ,  $\alpha_A = [.66, .78]$ ,  $\alpha_S = [.81, .91]$ ,  $\alpha_E = [.66, .86]$ ,  $\alpha_C = [.76, .87]$ ,  $\alpha_N = [.74, .84]$ ; McDonalds  $\omega$ :  $\omega_R = [.65, .77]$ ,  $\omega_I = [.73, .86]$ ,  $\omega_A = [.66, .78]$ ,  $\omega_S = [.81, .92]$ ,  $\omega_E = [.68, .86]$ ,  $\omega_C = [.75, .87]$ ,  $\omega_N = [.75, .86]$ ), sodass nachfolgende Ergebnisse als hinreichend reliabel angenommen wurden (s. Tabelle A.1).

### Intraklassenkorrelationen

Neben den Maßen der internen Konsistenz ist weiterhin der Einfluss der Mehrebenenstruktur der Daten auf die Messwerte und Messgenauigkeit abzuschätzen. Die Mehrebenenstruktur liegt aufgrund der Schachtelung der Schülerinnen und Schüler in Klassen und Schulen vor. Es lässt sich in diesem Fall sinnvoll zwischen der Personenebene (Level 1) und der Klassenebene (Level 2) unterscheiden. Die Berücksichtigung der Schulebene entfällt ob der geringen Anzahl an in der Studie inkludierten Schulen ( $N = 6$  Gymnasien). Ein Grund für eine mögliche Verzerrung der Ergebnisse durch die Mehrebenenstruktur ist, dass die Angaben der Schülerinnen und Schüler einer Klasse nicht als statistisch unabhängig voneinander zu betrachten sind. So können die Unterrichtswahrnehmungen und -empfindungen der Schülerinnen und Schüler, aber auch die Testbedingungen maßgeblich durch den Unterricht bzw. die Testdurchführung im Klassenverbund geprägt sein und diese Einflüsse sich in den subjektiven Einschätzungen der Schülerinnen und Schüler bei Beantwortung der Fragebögen widerspiegeln, weswegen es bei Nichtbeachtung der Nestung der Schülerinnen und Schüler in Klassen zu Verzerrungen der statistischen Ergebnisse kommen kann. Der Einfluss von Faktoren auf Klassenraumbene auf die angegebenen Interessenwerte lässt sich durch die Intraklassen-Korrelationskoeffizient ICC (intraclass correlation coefficient) ausdrücken. Es wird zwischen dem ICC(1), der prozentuale Anteil der Gesamtvarianz, der auf Unterschiede zwischen den Klassen basiert, und dem ICC(2), die Reliabilität des Klassenmittelwertes der Urteile, unterschieden (Lüdtke et al., 2006, 2007). Hohe ICC(1)-Werte indizieren folglich, dass eine hohe Varianz des gemessenen Merkmals durch Variablen auf Klassenraumbene erklärt werden und eine Nichtbeachtung der Mehrebenenstruktur zu fehlerhaften Interpretation führen kann. Hierbei können ICC(1)-Werte von .01 als „klein“, von .10 als „mittel“ und Werte von .25 als „groß“ in ihrer Bedeutung interpretiert werden (s. Murphy et al., 1998, S. 47, zitiert nach LeBreton und Senter, 2008). Die ICC(1)-Werte der RIASEC+N Skalen liegen zwischen  $-.02 < ICC(1) < .30$ , d. h. maximal 30% der Varianz des individuellen Interesses ist auf Variablen auf Klassenraumbene zurückzuführen, wobei hier von einem großen Effekt auszugehen ist ( $R = [-.01, .16]$ ,  $I = [-.02, .26]$ ,  $A = [-.01, .10]$ ,  $S = [.01, .19]$ ,  $E = [-.01, .20]$ ,  $C = [.00, .21]$ ,  $N = [-.02, .30]$ ). Der Einfluss auf komplexere Analysen lässt sich mit Hilfe des Designeffekts abschätzen, welcher eine Funktion der Intraklassenkorrelation sowie der Stichprobengröße ist und gemäß B. O. Muthén und Satorra (1995) berechnet und klassifiziert werden kann. Überschreitet der Designeffekt den Wert 2, so sollte die Mehrebenenstruktur und die Clusterung in Klassen berücksichtigt werden. Eine Berechnung des Designeffekts wurde bspw. im Zuge der Analysen des Artikels *Longitudinal couplings between interest and conceptual understanding* vollzogen (Höft & Bernholt, 2019b). Für diese Studie wurde ein mittlerer Designeffekt über drei

Messzeitpunkte von 1.8 spezifiziert, weswegen hier von einem vertretbaren Effekt auf die Ergebnisse ausgegangen und eine explizite Berücksichtigung der Mehrebenenstruktur vernachlässigt wurde (s. Anhang C.2). Trotzdem bleibt festzuhalten, dass der Einbezug der Mehrebenenstruktur zum einen bspw. durch Korrektur der Standardfehler zu präziseren Signifikanzschätzungen führen und zum anderen die Analyse von Kompositions- und Kontexteffekten erlauben würde. Jedoch stellen diese zunehmend komplexer werdende Analysen auch höhere Anforderungen an die Daten (bspw. Anzahl an Klassen, Anzahl der Schülerinnen und Schüler je Klasse), die mit dem vorliegenden Sample für die in den Publikationen gestellten Fragestellungen nicht ausreichend erfüllt werden konnten.

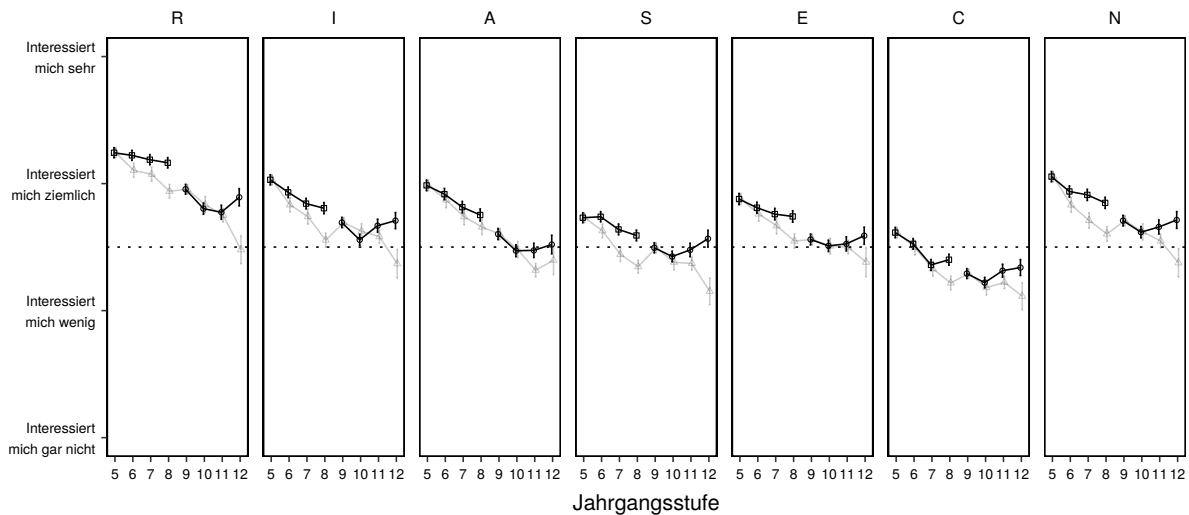
### Mittelwertvergleiche

Zur weiteren deskriptiven Analyse der Daten wurden die Mittelwerte und Standardabweichungen/-fehler je Messzeitpunkt und Jahrgangsstufe bestimmt, in Abbildung 4.1 graphisch aufgetragen und statistische Unterschiede in den Interessenausprägungen zwischen Längs- und Querschnitt sowie zwischen erstem und viertem Erhebungszeitpunkt berechnet. Bedeutsame Effekte ab einer moderaten Effektstärke ( $d > .30$ ,  $p < .05$ ) sind in Tabelle 4.3 dargestellt. Zudem wird eine differenzierte Betrachtung der längsschnittlichen Interessenverläufe der Schülerinnen und Schüler der Sekundarstufe II in Abhängigkeit ihrer Wahlentscheidung bzgl. des Faches Chemie vorgenommen.

Bei der Interpretation der Ergebnisse ist zu beachten, dass in die Bestimmung der Mittelwerte nur die Daten von Schülerinnen und Schülern eingingen, die zu den jeweiligen Erhebungszeitpunkten das Fach Chemie belegten, weswegen es sich bei diesem Subsample um eine Positivauswahl hinsichtlich des mittleren Interesses handelt. Diese Positivauswahl ist durch interessengeleitete Wahlentscheidungsprozesse begründet, die durch dimensionale Vergleiche beeinflusst werden (bspw. Nagy et al., 2007). Zur Erklärung des Mechanismus, der hinter dem Zusammenhang zwischen Wahlverhalten und Interesse der Schülerinnen und Schüler steht, kann Eccles et al.'s Erwartungswert-Modell der Motivation (s. Wigfield & Eccles, 2000) genutzt werden, das aussagt, dass proximale Determinanten des Wahlverhaltens Erfolgs- und Wertüberzeugungen sind. Dementsprechend entscheiden sich Schülerinnen und Schüler mit einem hohen Fachinteresse an Chemie eher, Chemie weiter zu belegen, während Schülerinnen und Schüler mit einem geringen Fachinteresse an Chemie eher dazu neigen, Chemie abzuwählen, sodass schließlich das Subsample verglichen mit der vollständigen Stichprobe Schülerinnen und Schüler mit im Mittel höheren Interessenausprägungen umfasst (Höft et al., 2017). Der in der Forschungsliteratur berichtete Trend der Fachinteressenabnahme (bspw. Krapp & Prenzel, 2011; Potvin & Hasni, 2014a) ist somit in dieser Analyse nicht bzw. nur in geringerem Umfang zu erwarten.

Bei Betrachtung der Mittelwertverläufe des Interesses an naturwissenschaftlichen Tätigkeiten (s. Abbildung 4.1) zeigt sich, dass in der Sekundarstufe I sowohl im Quer- als auch im Längsschnitt eine Abnahme des Interesses von Jahrgangsstufe 5 zu 8 in allen sieben Interessendimensionen zu verzeichnen ist, von denen im Längsschnitt drei statistisch bedeutsam sind (investigative:  $t(288) = 5.16$ ,  $p < .001$ ,  $d = 0.36$ ; artistic:  $t(287) = 5.84$ ,  $p < .001$ ,  $d = 0.39$ ; networking:  $t(287) = 4.46$ ,  $p < .001$ ,  $d = 0.33$ ). Die Interessendifferenzen zwischen Jahrgangsstufe 5 und 8 fallen in den Dimensionen realistic, investigative, social und networking für die Querschnitterhebung höher als für die Längsschnitterhebung aus ( $d > .35$ ,  $p < .001$ , s. Tabelle 4.3).

Ebenso lässt sich eine Divergenz zwischen den mittleren Interessenverläufen der Quer- und Längsschnittdaten der Schülerinnen und Schüler aus der oberen Kohorte der Jahrgangsstufen 9 bis 12 erkennen, da der im Querschnitt gefundene Interessenabfall in dem hier betrachteten Subsample des Längsschnitts nicht gleichermaßen zu identifizieren ist. Vielmehr scheinen die Schülerinnen und Schüler, die tatsächlich noch am Chemieunterricht partizipieren, im Mittel eher eine konstante Interessenausprägung aufzuweisen. So unterscheiden sich die Schülerinnen und Schüler der Querschnitterhebung von den Schülerinnen und Schülern der Längsschnitterhebung in Jahrgangsstufe 12 in den Dimensionen realistic, investigative, social, conventional und networking ( $d > .33$ ,  $p \leq .03$ , s. Tabelle 4.3). Das Ausbleiben des negativen Trends könnte durch die vorliegende Positivauswahl und differierende Probandenzusammensetzungen je Erhebung begründet sein, die Divergenz zum



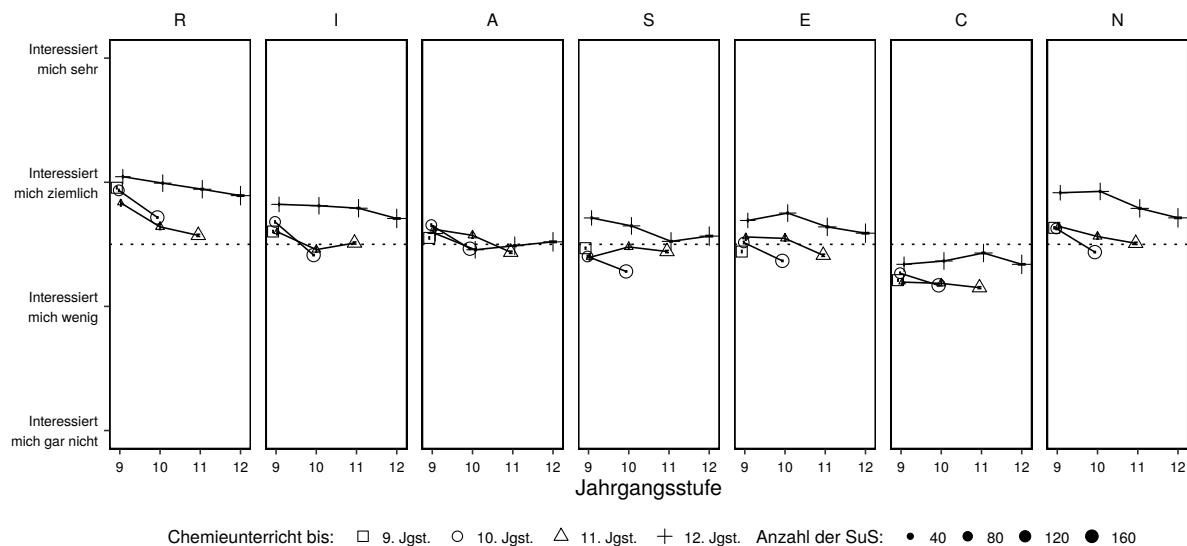
**Abb. 4.1:** Darstellung des mittleren Interesses ( $M \pm SE$ ) der Schülerinnen und Schüler im Quer- (grau) und Längsschnitt (schwarz) differenziert je Dimension (R = Realistic, I = Investigative, A = Artistic, S = Social, E = Enterprising, C = Conventional, N = Networking). Es wurden nur die Daten von Schülerinnen und Schülern berücksichtigt, die zu den jeweiligen Erhebungszeitpunkten das Fach Chemie belegten.

**Tab. 4.3:** Mittelwertvergleiche zwischen den Ausprägungen der RIASEC+N Interessensdimensionen im Querschnitt und Längsschnitt, die bedeutsame Unterschiede indizieren ( $d > .30$ ,  $p < .05$ ).

Jgst.	Skala	<i>t</i> -Wert	<i>df</i>	<i>p</i>	<i>d</i>
7	N	-3.700	485.967	.000	.303
8	R	-4.500	508.031	.000	.360
8	I	-4.421	513.728	.000	.352
8	S	-4.663	577.312	.000	.359
8	N	-4.710	531.236	.000	.372
12	R	-4.441	134.828	.000	.613
12	I	-3.197	110.303	.002	.489
12	S	-3.783	132.671	.000	.526
12	C	-2.194	116.675	.030	.326
12	N	-3.236	122.756	.002	.468

Querschnitt jedoch nicht. Gründe für Unterschiede zwischen Ergebnissen des Quer- und Längsschnitts können u. a. Selektionseffekte, Kohorteneffekte, Änderungen auf Schulebene (geänderte Curricula im Zuge des neuen Lehrplans), aber z. B. auch Kontext- und Kompositionseffekte, die auf die Lehrkraft oder Klassenraumvariablen zurückzuführen sind, sein. Letztlich könnten aber gleichfalls Testeffekte (wiederholtes Testen, Einfluss des Testleiters, etc.) aufgetreten sein, die die Ergebnisse verzerren. Kausale Gründe für die Differenz zwischen den Daten des Quer- und des Längsschnitts lassen sich basierend auf den hier analysierten Skalen nicht treffen.

Um den Einfluss der Probandenzusammensetzung bzw. der Positivauswahl durch die Abwahl des Faches Chemie auf die mittlere Interessenentwicklung aufzudecken, wurde eine Differenzierung der Schülerinnen und Schüler in vier Gruppen aufgrund des Abwahlzeitpunkts des Chemieunterrichts getroffen (Gruppe 1: Schülerinnen und Schüler, die Chemie nach der 9. Jahrgangsstufe abwählten; Gruppe 2: Schülerinnen und Schüler, die Chemie nach der 10. Jahrgangsstufe abwählten; Gruppe 3: Schülerinnen und Schüler, die Chemie nach der 11. Jahrgangsstufe abwählten; Gruppe 4: Schülerinnen und Schüler, die Chemie durchgehend belegten). Die Ergebnisse sind graphisch in Abbildung 4.2 dargestellt. Paarweise durchgeführte *t*-Tests ( $\alpha$ -Fehler-Adjustierung nach Holm, 1979) zeigen signifikante Unterschiede in den Interessenausprägungen der Schülerinnen und Schüler, die Chemie nach der 9. Jahrgangsstufe abwählten (Gruppe 1), zu denen, die Chemie durchgehend belegten, in den Dimensionen investigative ( $d = .39$ ,  $p < .01$ ), enterprising ( $d = .39$ ,  $p < .01$ ) und networking ( $d = .46$ ,  $p < .01$ ). Hinsichtlich der längsschnittlichen Interessenentwicklung



**Abb. 4.2:** Darstellung des mittleren Interesses ( $M \pm SE$ ) der Schülerinnen und Schüler (SuS) im Längsschnitt differenziert je Dimension (R = Realistic, I = Investigative, A = Artistic, S = Social, E = Enterprising, C = Conventional, N = Networking) und Zeitpunkt der Abwahl des Chemieunterrichts. Es wurden nur die Daten von Schülerinnen und Schülern berücksichtigt, die zu den jeweiligen Erhebungszeitpunkten das Fach Chemie belegten.

der Schülerinnen und Schüler, die durchgehend Chemieunterricht in der Sekundarstufe II hatten (Gruppe 4), wurden die Interessenausprägungen in Jahrgangsstufe 9 und 12 miteinander verglichen (Vergleich zwischen Schülerinnen und Schüler, die an Testzeitpunkten 1 und 4 teilnahmen). Für dieses Subsample liegen bedeutsame Interessenabnahmen in den Interessendimensionen realistic ( $t(112) = 3.38, p < .01, d = .36$ ), investigative ( $t(112) = 2.83, p < .01, d = .32$ ) und networking ( $t(112) = 3.07, p < .01, d = .33$ ) vor. Die Differenzierung in vier Gruppen kann somit den Interessenverlauf präziser beschreiben und wird weniger durch die Samplezuschneidung verzerrt. Um in den komplexeren längsschnittlichen Analysen der Publikationen 2–3 diese Verzerrung zu vermeiden und gleichzeitig keine Informationen zu verlieren, wurden die fehlenden bzw. durch Abwahl des Faches Chemie invaliden Daten der Schülerinnen und Schüler durch Imputationsverfahren bzw. FIML geschätzt.

### 4.3 Latente Profilanalyse

Ziel der Latenten Profilanalyse (LPA) ist es, in der heterogenen Stichprobe homogene, voneinander abgrenzbare Subgruppen von Personen (Klassen) zu identifizieren und somit die Beschreibung der Stichprobe anhand der Subgruppen zu ermöglichen. Hierbei werden Wahrscheinlichkeiten geschätzt, mit welchen Probanden den gefundenen latenten Klassen zugeordnet werden können. Die in die Schätzung eingehenden Skalen werden genutzt, um die gefundenen Gruppen aufgrund mittlerer Ausprägungen auf den Skalen zu charakterisieren. Insofern verfolgt die latente Profilanalyse einen personenzentrierten Ansatz. Die Bestimmung der Anzahl an latenten Klassen erfolgt modellbasiert aufgrund von informationstheoretischen Passungsmaßen und Signifikanztests.

Die Identifikation von Subgruppen von Schülerinnen und Schülern könnte einen höheren Grad der Binnendifferenzierung des Unterrichts aufgrund der Reduzierung des gesamten Samples auf Subgruppen und der Entwicklung passender Lernangebote erleichtern oder das Verständnis komplexer Zusammenhänge bspw. zwischen Personenmerkmalen und Interventionsmaßnahmen unterstützen, indem die Aufdeckung von zwischen den Subgruppen differierenden Wirkungsweisen ermöglicht wird. Bisherige, auf latenten Profilanalysen und latenten Transitionsanalysen basierende Untersuchungen zu affektiv-motivationalen Variablen zeigten vorwiegend, dass die extrahierten Gruppen sich meist nur aufgrund von Mittelwertverschiebungen unterscheiden ließen (Lazarides et al., 2019,

S. 164) und dass die gefundenen Profile für junge Erwachsene relativ stabil sind (Hayenga & Corpus, 2010). Werden zudem stärker kognitiv geprägte Variablen wie die Leistung und das Selbstkonzept in die Analysen integriert, können zwischen den Gruppen Verschiebungen in der Stärke der relativen Mittelwertausprägungen gefunden werden (Lazarides et al., 2019; Linnenbrink-Garcia et al., 2018; Seidel & Prenzel, 2006; Viljaranta et al., 2017). So lassen sich z. B. Gruppen von Schülerinnen und Schülern finden, die über ein hohes Selbstkonzept und hohe kognitive Fähigkeiten verfügen, jedoch uninteressiert am Unterricht sind (Seidel & Prenzel, 2006) bzw. invertiert lassen sich ebenso Schülerinnen und Schüler identifizieren, die ein hohes Interesse, niedrige Kompetenzen und ein geringes akademisches Selbstkonzept aufweisen (Viljaranta et al., 2017).

In diesem Abschnitt soll untersucht werden, ob sich (a) Klassen von Schülerinnen und Schülern mit differierenden Profilen des Interesses an naturwissenschaftlichen Tätigkeiten identifizieren lassen und ob sich (b) die Klassenzugehörigkeit durch das Geschlecht, Konzeptverständnis sowie die Jahrgangsstufe vorhersagen lässt.

Aufgrund bisheriger empirischer Befunden ist zu vermuten, dass sich mindestens drei Klassen von Schülerinnen und Schülern diskriminieren lassen, von denen anzunehmen ist, dass sie sich lediglich durch Mittelwertverschiebungen in den Interessenausprägungen unterscheiden. Genauer gesagt ist es wahrscheinlich, dass sich in dem Sample eine Gruppe von interessierten, eine Gruppe von desinteressierten Schülerinnen und Schülern sowie eine Gruppe von Schülerinnen und Schülern mit mittleren Interessenausprägungen finden lassen. Charakteristische Interessenprofile wie sie mittels Circumplex-Modellen für Berufsinteressen gefunden werden konnten (bspw. Nagy et al., 2019), sind aufgrund der relativ homogenen Stichprobenszusammensetzung (Beschränkung auf Gymnasiastinnen und Gymnasiasten) sowie der relativ hohen, positiven Zusammenhänge zwischen den sieben Interessendimensionen nicht zu erwarten.

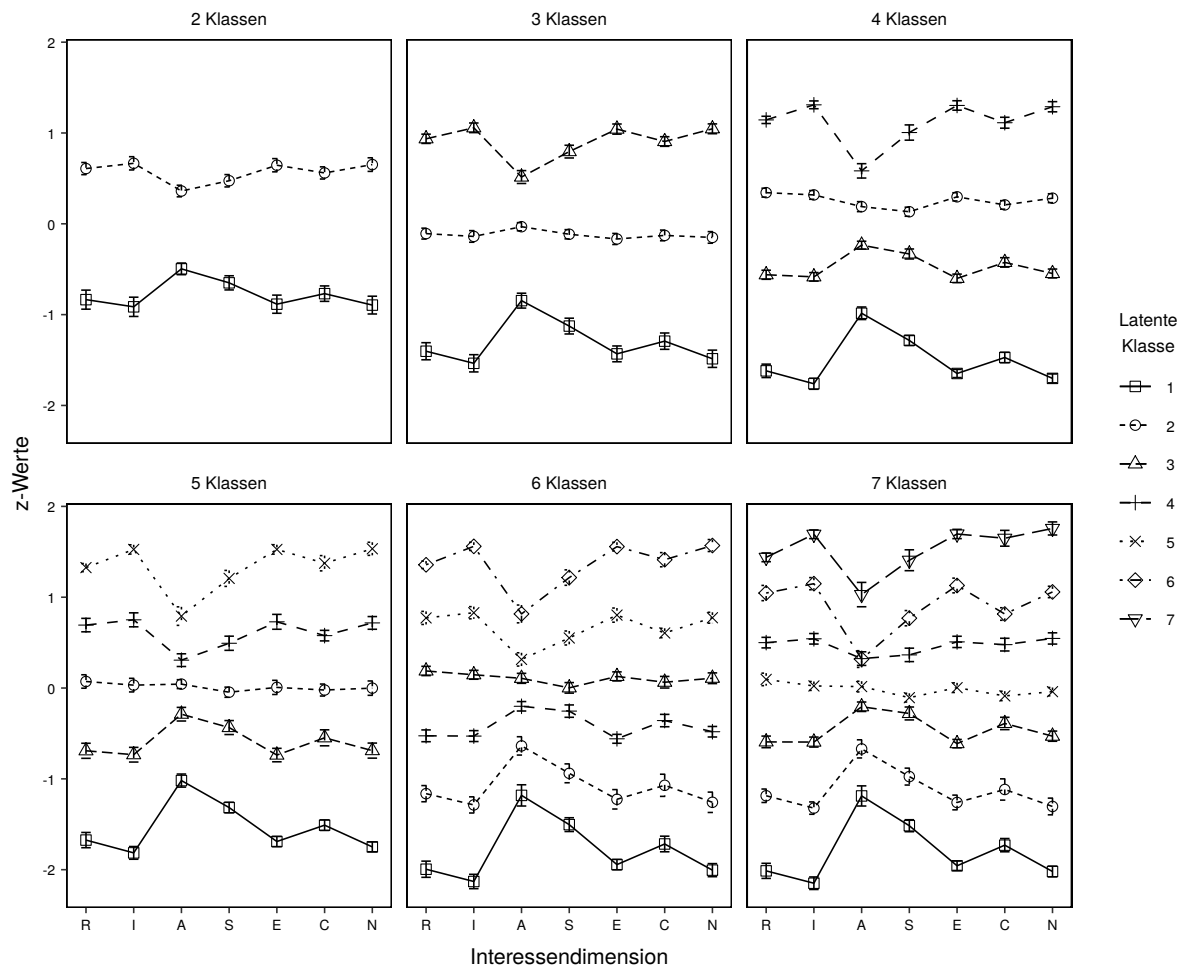
Hinsichtlich der Vorhersage der Klassenzugehörigkeit durch das Geschlecht, das Konzeptverständnis sowie die Jahrgangsstufe ist zu erwarten, dass in einer Klasse, die sich durch geringe Interessenausprägungen auszeichnet, mehrheitlich Schülerinnen sowie Schüler und Schülerinnen aus höheren Jahrgangsstufen zu finden sein sollten, da bisherige empirische Befunde darauf hindeuten, dass erstens Mädchen im Mittel ein geringeres Interesse an naturwissenschaftlichen Fächern angeben und dass zweitens Schülerinnen und Schüler im Laufe ihrer Schullaufbahn im Mittel das Interesse an naturwissenschaftlichen Fächern verlieren (Krapp & Prenzel, 2011; Potvin & Hasni, 2014b; Wigfield & Cambria, 2010). Bei Kontrolle der Jahrgangsstufe sollten zudem Schülerinnen und Schüler mit einem niedrigeren Konzeptverständnis mit einer höheren Wahrscheinlichkeit in Klassen mit im Mittel geringen Interessenausprägungen zu finden sein (Denissen et al., 2007).

Als Datengrundlage für die LPAs wurde auf die factor scores der konfirmatorischen Faktorenanalyse des Sieben-Faktoren-Modells (M3) zurückgegriffen. Dieses Vorgehen bietet gegenüber der Berechnung von manifesten Mittelwerten den Vorteil, dass die beobachtbaren Antworten gewichtet in die factor scores eingehen, Residuen der Items exkludiert werden und fehlende Werte bereits auf Itemebene mittels FIML geschätzt werden konnten.

Zur Bestimmung der Anzahl der latenten Klassen wurde ein klassisches, restriktives Modell (M1) geschätzt, das fixierte Varianzen zwischen den Klassen und keine Kovarianz zwischen den Interessendimensionen zulässt (lokale Unabhängigkeit; Marsh et al., 2009). Der Vollständigkeit halber wurden zudem vier Modelle mit schrittweise abnehmenden Restriktionsgraden definiert, die durch das Erlauben von über die latenten Klassen varianten bzw. invarianten Kovarianzen und über die Klassen varianten Varianzen schrittweise mehr Freiheitsgrade aufweisen (s. Tabelle A.2). Die Relaxation der lokalen Unabhängigkeit entspricht laut Marsh et al. (2009) nicht klassischen LPA Anwendungen und kann zu der Detektion einer geringen Anzahl von latenten Klassen führen. Zudem kann die Erhöhung der zu schätzenden Parameter das Finden der Lösung mit dem größten Loglikelihood-Wert erschweren. So konnte in der vorliegenden Untersuchung für die Modelle M2 bis M5 teils keine Lösung mit dem größten Loglikelihood-Wert repliziert werden. Um die Mehrebenenstruktur zu berücksichtigen, wurde in Mplus die Option „Type Is Complex“ gewählt.

Für die Bestimmung der Anzahl der latenten Klassen wurde modellbasiert vorgegangen. Es wurde das klassische Modell M1 mit einer steigenden Anzahl an latenten Klassen geschätzt und die Infor-





**Abb. 4.3:** Mittelwertausprägungen der klassischen latenten Profilanalysen. Dargestellt sind die Lösungen für zwei bis sieben latente Klassen.

mationskriterien AIC, BIC und der adjustierte BIC (sample-size-adjusted BIC, SSA-BIC) betrachtet sowie die Lo-Mendell-Rubin Likelihood-Ratio- (LMR) und Vuong-Lo-Mendell-Rubin-Signifikanztests (VLMR) durchgeführt. Hierbei wird davon ausgegangen, dass das Hinzufügen einer zusätzlichen Klasse zu einer Abnahme der Informationskriterien führen und die Anpassungsgüte signifikant besser ausfallen sollte (LMR, VLMR). Eine Einschätzung der Passung erfolgte anhand einer Kombination aller fünf Kriterien, da die Betrachtung eines einzelnen Kriteriums zu falschen Schlüssen führen kann; bspw. kann die alleinige Betrachtung der BIC-Werte zu einer Überschätzung der Anzahl der latenten Klassen führen. Anschließend wurde für die Einschätzung der Güte der gefundenen Lösung zudem die Entropie, die als Maß für die Zuverlässigkeit der Klassifikation über .70 liegen sollte, sowie die Größe der bestimmen latenten Klassen (mindestens 10 % der Probanden) bewertet.

Die Ergebnisse der LPAs (s. vollständige Auflistung der Kennwerte in Tabelle A.2) deuten auf eine Lösung mit vier latenten Klassen hin. Die Hinzunahme einer fünften latenten Klasse führt zwar zu einer weiteren Reduktion der Werte der Informationskriterien, jedoch werden der LMR und VLMR-Test für die Fünf-Klassen-Lösung nicht signifikant. Zudem unterschreitet die fünfte Klasse die Mindestgröße von 10 % der Schülerinnen und Schülern. Bei Betrachtung der mittleren Interessenausprägungen der latenten Klassen (s. Abbildung 4.3) zeigt sich, dass sich die einzelnen Klassen erwartungskonform nur durch die Verschiebung der Mittelwerte voneinander unterscheiden. Entsprechend der Lösung mit vier latenten Klassen, findet man eine Klasse mit einem hohen Interesse über alle sieben Faktoren hinweg (CL 4  $\hat{=}$  hohes Interesse), eine Klasse von uninteressierten Schülerinnen und Schülern (CL 3  $\hat{=}$  kein Interesse) und dazwischen noch zwei weitere latente Klassen mit einem mittleren Interesse an naturwissenschaftlichen Tätigkeiten im Chemieunterricht

(CL 2  $\hat{=}$  geringes Desinteresse, CL 1  $\hat{=}$  geringes Interesse). Der Erkenntnisgewinn der LPA erscheint gering, da keine charakteristischen Interessenprofile gefunden wurden, die als Ansatzpunkt für weitere Studien genutzt werden könnten. Trotzdem ist berichtenswert, dass im Gegensatz zu Analysen fußend auf dem RIASEC Modell der Berufsinteressen nach Holland, an welches das RIASEC+N Modell angelehnt ist, keine charakteristischen Interessenprofile vorliegen.

**Tab. 4.4:** Multinomiale logistische Regression zu Vorhersage der latenten Klasse. Die Ergebnisse für die Referenzklasse CL 1 entsprechen der ersten Spalte (CL 1) und sind daher nicht redundant geführt. Der Regressionskoeffizient ist in log odds angegeben.

	CL 1		CL 2		CL 3	
	B	SE	B	SE	B	SE
Referenzklasse CL 2 (geringes Desinteresse)						
Jgst.	0.653***	0.150				
Geschlecht	-0.119	0.152				
WLE Score	-0.069	0.113				
Referenzklasse CL 3 (kein Interesse)						
Jgst.	1.091***	0.178	0.438**	0.168		
Geschlecht	-0.815***	0.196	-0.696***	0.193		
WLE Score	0.316	0.148	0.384**	0.143		
Referenzklasse CL 4 (hohes Interesse)						
Jgst.	-0.710***	0.137	-1.363***	0.174	-1.801***	0.192
Geschlecht	0.190	0.160	0.309	0.161	1.005***	0.200
WLE Score	0.044	0.121	0.113	0.122	-0.271	0.138

Die Vorhersage der Zugehörigkeit von Schülerinnen und Schülern zu latenten Klassen durch die drei Kovariaten Geschlecht, Konzeptverständnis und Jahrgangsstufe wurde mit multinominalen logistischen Regressionen basierend auf dem 3-Steps-Verfahren berechnet, um eine Beeinflussung der Klassenzuordnung durch die Kovariaten zu verhindern (Asparouhov & Muthén, 2014). Die Ergebnisse der multinominalen logistischen Regressionen sind in Tabelle 4.4 geführt. Die Regressionskoeffizienten (Angabe in log odds) geben hierbei den Zusammenhang zwischen der Wahrscheinlichkeit, dass ein Proband in der interessierenden Klasse situiert im Vergleich zu der Wahrscheinlichkeit, dass der Proband in die Referenzklasse eingeordnet wird. Die Ergebnisse legen nahe, dass Schülerinnen und Schüler, die sich in einer höheren Jahrgangsstufe befinden, sowie Schülerinnen im Vergleich mit Schülern eher latenten Klassen mit geringeren Interessenausprägungen zugeordnet werden. Diese Befunde unterfüttern variablenzentrierte Verfahren, die ebenso ein Interessenabfall über die Schulzeit sowie ein geringeres Interesse von Schülerinnen verglichen mit Schülern nahelegen.

#### 4.4 Fazit

In diesem Kapitel wurden zunächst konfirmatorische Faktorenanalysen zur Prüfung der Struktur des Interessenkonstrukts durchgeführt, darauf aufbauend deskriptive Maße berichtet und abschließende Ergebnisse latenter Profilanalysen vorgestellt. Die konfirmatorischen Faktorenanalysen deuten daraufhin, dass das individuelle Interesse der Schülerinnen und Schüler an naturwissenschaftlichen Tätigkeiten am adäquatesten durch sieben miteinander korrelierte, latenten Faktoren beschrieben werden kann. Die sieben latenten Faktoren entsprechen hierbei den angenommenen sieben Interessendimensionen des RIASEC+N Modells. Die anschließende Prüfung der Reliabilitäten je Jahrgangsstufe deutet auf überwiegend akzeptable bis teilweise sehr gute Reliabilitäten der Skalen hin. Die Analyse der Intraklassenkorrelationen zeigen an, dass teilweise ein moderater Anteil der Varianz innerhalb der Interessendimensionen durch die Klassenzuordnungen erklärt werden kann. Eine Berücksichtigung der Klassenzuordnung (Nestung) in den nachfolgenden längsschnittlichen Analysen durch Mehrebenenmodelle bzw. Messfehlerkorrektur mittels „Type Is Complex“ Option in Mplus ist jedoch

aufgrund einer zu geringen Anzahl an Klassen (Einheiten der zweiten Ebene) und sich ändernder Klassenkomposita nicht möglich. Die abschließend durchgeführten latenten Profilanalysen konnten keine charakteristischen Interessenprofile aufzeigen. Jedoch bekräftigen diese Analysen eine generelle Abnahme des Interesses an naturwissenschaftlichen Tätigkeiten über alle sieben Interessendimensionen des RIASEC+N Modells hinweg.



# Knowing more about things you care less about: Cross-sectional analysis of the opposing trend and interplay between conceptual understanding and interest in secondary school chemistry

## Abstract

The development of students' interest in school science activities, their understanding of central chemical concepts and the interplay between both constructs across Grades 5 to 11 were analyzed in a cross-sectional paper-and-pencil study ( $N = 2510$ , mean age 11 to 17 years). Previous empirical findings indicate that students' knowledge increases over the time of secondary school while students' interest, especially in natural science subjects, tends to decrease. Concomitantly, there is evidence for an increase in the positive coupling between interest and knowledge across time. However, previous studies mainly rely on rather global measures, e.g., school grades or general subject-related interest, and focus on science as an integrated subject instead of specific disciplines, e.g. chemistry. For the present paper, more proximal and differentiated measures for students' understanding of three chemical concepts (Chemical Reaction, Energy, Matter) and interest in seven dimensions of school science activities according to the RIASEC+N model (Realistic, Investigative, Artistic, Social, Enterprising, Conventional, and Networking; cf. Dierks et al., 2014b) were applied. The results are in line with previous research indicating a general increase in conceptual understanding and a decline in students' interest for all school science activities. However, the interplay between conceptual understanding and interest differs across the seven dimensions. Interest in activities which are likely to promote cognitive activation (investigative, networking) or involving the communication of knowledge (social, enterprising, networking) are increasingly connected to conceptual understanding, especially in upper secondary grades. Interest in guided hands-on activities (realistic) which are typical in secondary science teaching, however, shows only small positive correlations to students' conceptual understanding across all grades. Hence, in upper-secondary school, investigative, social, enterprising and networking activities seem to provide opportunities to benefit most from the interrelation between students' interests and their understanding.

Keywords: conceptual understanding, interest, chemistry education

Höft, L., Bernholt, S., Blankenburg, J. S., & Winberg, M. (2019). Knowing more about things you care less about: Cross-sectional analysis of the opposing trend and interplay between conceptual understanding and interest in secondary school chemistry. *Journal of Research in Science Teaching*, 56(2), 184–210

## 5.1 Introduction

Scientific knowledge and interest (in science) are decisive factors for our understanding of the world and, hence, our action. The foundations for both knowledge about and interest in science are laid in school as a meaningful engagement with scientific ideas is likely to occur here first. Unsurprisingly, the acquisition of a sound scientific knowledge is a crucial goal of science education, not least because it is a prerequisite for the much needed participation of citizens in the modern society that is shaped by scientific inventions and technical achievements (Chopyak & Levesque, 2002). Interest has been proposed to play a central role in knowledge development, affecting students' learning processes (Hidi & Renninger, 2006; Krapp & Prenzel, 2011; Renninger & Hidi, 2011), quality of learning outcomes (Schiefele, 2001), and sustaining learning over time (Cordova et al., 2014) as well as study success (Bengmark et al., 2017). Although several empirical studies corroborate the theoretically assumed positive relationship between interest and knowledge development ambiguities do exist. Correlations vary substantially between studies, from near zero to strong (Schiefele et al., 1992). Furthermore, uncertainty remains on how the association between interest and knowledge develops over the time of schooling (Denissen et al., 2007; Rotgans & Schmidt, 2017d). One reason for this ambiguities could be that students' interest varies not only between subjects and topics, but also between different activities within the process of doing science (Azevedo, 2013; Luce & Hsi, 2015; Swarat et al., 2012). A focus on scientific activities also establishes a basis for comparing students' interest across longer time periods

(i.e., several years of schooling) as these activities are rather stable and comparable while topics and contexts usually differ considerably between different grades. Hence, to generate meaningful results on the long-term relationship between interest and knowledge development, measurements of interest must account for the existence of several possible types of activities within a subject instead of merely considering an averaged (general) subject-related interest or comparing students' interest in topics which vary over time. Taking scientific activities into account when examining students' interests in chemistry is particularly reasonable because a characteristic of chemistry education is its emphasis on scientific activities such as devising, performing, and interpreting experiments (National Research Council, 2012; NGSS Lead States, 2013).

However, for chemistry education, there is, to our best knowledge, a lack of studies examining the relation between interest and knowledge over a longer time span. Therefore, in the current research, we examine the development of and the relationship between students' interest and their chemistry knowledge across Grades 5 to 11. In order to provide more detailed insights into the interrelation between both constructs, we applied more proximal and differentiated measures for students' understanding of three chemical concepts (Chemical Reaction, Energy, and Matter) and students' interest in seven dimensions of school science activities. Students' interest was assessed based on the RIASEC+N model which distinguishes between students' interest in realistic (performing given lab experiments), investigative (solving theoretical problems), artistic (emphasizing linguistic and visual aspects), social (explaining something to classmates), enterprising (managing group works), conventional (organize the chemicals storage), and networking (debating with classmates) school science activities (Dierks et al., 2016). Insights based on these more differentiated measures could help to identify science activities which support students to gain a more sophisticated conceptual understanding and are therefore crucial for teaching.

## 5.2 Theoretical background

### Interest

Interest is an important outcome variable in its own right and additionally decisive for the development of other outcome variables, such as scientific literacy. Thus, ample studies examine ways and means of fostering students' interest. Commonly, interest is defined as a motivational and multi-dimensional construct (Knogler et al., 2015; Krapp & Prenzel, 2011; Renninger & Hidi, 2011) and is conceptualized as being „content or object specific“, involving a „relation between a person and the environment“, having „both cognitive and affective components“ with a physiological/neurological foundation and as being potentially unrecognized by the person (Renninger & Hidi, 2011, p. 169). The specificity of interest, i.e., being focused on a specific subject or activity („what am I motivated to engage in?“) and not on the goals and reasons for individuals' engagement („why am I doing this?“), is often accentuated when distinguishing interest from other related constructs (Krapp & Prenzel, 2011).

The psychological state of interest is characterized by focused attention and increased cognitive and affective functioning which can facilitate integration of information with prior knowledge and enhance learning (M. Ainley, Hidi, & Berndorff, 2002; Hidi & Renninger, 2006). Furthermore, the process of becoming interested involves appraisals of the alignment between the object or activity and the individual's personal goals, ideals, and values. For example, students who are interested in a subject may be so because they have identified the value of learning the content for their further studies or they find the activity of learning the content emotionally rewarding (M. Ainley & Ainley, 2011b; Fredricks et al., 2018).

The process of interest development is described, for instance, by the four-phase model of interest development (Hidi & Renninger, 2006) and the person-object theory of interest (Krapp, 2002b). Both models differentiate between a situational (interest-as-state) and an individual interest (interest-as-trait). Situational interest is a short-term, psychological state evoked by the interaction between the learner and some aspect of a specific situation, e.g., personal relevance, novelty, vividness and

comprehensibility (Wigfield & Cambria, 2010). Individual interest, in turn, is considered to be a more long-lasting predisposition to reengage in a certain activity, subject, or a content of interest as well as to emerge from an increase in content knowledge (Hidi & Renninger, 2006; Krapp & Prenzel, 2011). Krapp (2002a) discusses the development of interest in terms of the self-determination theory (Deci & Ryan, 1985a) and suggests that this predisposition will only be realized if a person satisfies the need of competence, autonomy (e.g. Tsai et al., 2008), and social-relatedness.

With that in mind, we argue that, regardless of the type of interest, the higher the attractiveness of the activity is, the higher the perceived scope for self-determined behavior will be, ultimately leading to increased engagement, concentration, and persistence of the individual in the specific situation (cf. Deci & Ryan, 1985a; Guay et al., 2008; Jang et al., 2010). Although interest may initially be triggered by situational features that appeal to the individual, e.g., by perceived goal congruency or novelty of the activity (Palmer, 2009; Renninger & Bachrach, 2015), if sustained, it will lead to increased effort (Trautwein et al., 2015) and, eventually, increased content knowledge. Higher content knowledge will in turn feedback on the individual's interest by an increased perception of both competence and ability to pursue own „curiosity questions“ in relation to the material to be learned (Hidi & Renninger, 2006). Although the causes of interest might differ during the development from situational interest to individual interest, actual competence should increase during the course of interest development (Renninger & Hidi, 2011).

With regard to the specificity of interest, i.e. pertaining to what a person is interested in, the use of school subjects for characterizing the content structure of students' interest in science is often considered too rough from a theoretical perspective. Additional aspects such as subject contents and themes, contexts (Häussler et al., 1998), or activities (J. A. Schmidt et al., 2018; Swarat et al., 2012) may be indispensable for describing profiles of interest in science. Such an integrated approach is provided by the adapted version of Holland's RIASEC model, allowing more precise insights into the development of interest and its connection to learning (Blankenburg, Höffler, & Parchmann, 2016; Dierks et al., 2016; Dierks et al., 2014b). The adapted version of the RIASEC model takes up the multifaceted structure of the original model, i.e. the categorization of interests into different personality types (according to which scientists would be allocated to the dimension investigative solely), but instead of focusing on vocational interests, it describes interest in different types of school science activities. In this vein, the adapted RIASEC model differentiates between interest in realistic (e.g., „doing experiments guided by an instruction“), investigative (e.g., „plan experiments to investigate something“), artistic (e.g., „draw an observation“), social (e.g., „explaining something to fellow students“), enterprising (e.g., „organizing a small chemistry project“), conventional (e.g., „tabling the results of an experiment“), and networking (e.g., „talking to fellow students about chemical topics“) school science activities (Dierks et al., 2016). Compared to the original model (Holland, 1997), networking was added as a new dimension due to the results of factor analyses, hence, the model has been extended to the seven RIASEC+N dimensions of interest in school science activities (Dierks et al., 2014b). The RIASEC+N model acknowledges all seven dimension as meaningful elements of science instruction and thus provides a more multifarious picture of (school) science than the rather stereotypical categorization of science as an exclusively investigative endeavor in the original model (Holland, 1997). So far, evidence for the hypothesized structure has been found (Blankenburg, Höffler, & Parchmann, 2016), but no findings in connection to achievement have been reported. Due to its differentiation and decontextualized operationalization, the model allows precise comparisons of students' interests over longer time spans.

There is ample evidence for a decline of interest over the years of schooling. Particularly, students tend to lose their interest in natural science subjects such as chemistry and physics, (Anderhag et al., 2016; Krapp & Prenzel, 2011; Wigfield & Cambria, 2010). Krapp and Prenzel (2011) emphasize three main reasons for this decline, namely, (a) the quality and type of instruction (Kunter et al., 2013; Schiefele & Schaffner, 2015; Tröbst et al., 2016); (b) developmental issues (students in adolescence prioritize new developmental tasks); and (c) differentiation processes of interest in adolescence.

Additionally, findings by Anderhag et al. (2016) indicate a change of students' object of interest in science subjects during their school life, from general enjoyment of being in class to an increasing

focus on scientific objects, underlining the need for a more discriminated approach for assessing interest.

### Interplay between interest and academic achievement

The interplay between interest and academic achievement is commonly discussed against the background of theories focusing solely on interest, e.g. the four-phase model of interest development (Hidi & Renninger, 2006; Renninger & Hidi, 2011), or broader frameworks such as the social cognitive career theory (Lent et al., 1994) and the expectancy-value theory of achievement motivation (Eccles, 2009; Wigfield & Eccles, 2000). The latter two seek to explain the pursuit of academic activities as well as career aspirations and choices. The social cognitive career theory posits favorable beliefs about one's academic performance and outcome expectations as prerequisites for interest development and stresses the importance of interest for the engagement in academic activities and consequently knowledge acquisition (Sheu et al., 2010). Likewise, according to Eccles et al. (2015), interest can be interpreted „as both the outcome of activity choice processes and as one of the inputs into activity choice“ (p. 326). Hence, both framework suggests that interest is associated with an increased likelihood to engage in more challenging tasks as well as a higher persistence in performing a given task, resulting in higher achievement which in turn is related to interest development, so that reciprocal relations between interest and achievement can be concluded.

In general, studies document a positive relation between interest and academic achievement in different domains. Empirical findings indicate that students with a higher interest are more likely to adopt effective learning strategies (Berger & Karabenick, 2011; Krapp, 2002b) and to invest more effort (Trautwein et al., 2015) than students with lower interest. Moreover, the positive effects of interest on learning seem to be mediated by attentional processes, affect, and persistence (M. Ainley, Hidi, & Berndorff, 2002) as well as self-regulation (Lee et al., 2014). In a meta-analysis, Schiefele et al. (1992) found a mean correlation of  $r = .30$  (95% CI: .04 to .56) between domain-specific interest and achievement. Other studies indicate that this correlation seems to increase over time (Denissen et al., 2007; Kim et al., 2015; Köller et al., 2001). More recent studies revealed that interest was the best predictor of class grade, whereas test scores were more strongly related to cognitive abilities – with interest playing a less important, but still significant role (Jansen et al., 2016). There is evidence, that the interest-achievement relation is higher in chemistry and physics as well as math when compared to biology (perceived to be more verbal) and first-language (L1, Jansen et al., 2016; Jansen et al., 2015; Krapp & Prenzel, 2011; Wigfield & Cambria, 2010). Longitudinal studies indicate small reciprocal effects between interest and achievement (Garon-Carrier et al., 2016; Harackiewicz et al., 2008; Marsh et al., 2005), but results regarding the causal relation are still inconclusive (Jögi et al., 2015; Rotgans & Schmidt, 2017d; Viljaranta et al., 2014).

### Conceptual understanding

In this study, academic achievement is conceptualized in terms of students' understanding of fundamental chemical concepts. The acquisition of a sound scientific understanding is a pivotal aim of science education. With regard to school science, a central question is how to support the growth of understanding over time. In recent years, researchers have recommended to focus science education on core concepts that are fundamental to the discipline, instead of teaching a broad range of different topics and content aspects (Bransford et al., 2000). In many countries, novel educational standards have adopted this view by incorporating disciplinary core concepts (cf. Bernhold et al., 2012; NGSS Lead States, 2013). Hence, modern curricula underlie the assumption that „learning ought to be coordinated and sequenced along conceptual trajectories (Driver et al., 1994), developmental corridors (A. L. Brown, 1997) and learning progressions“ (Duschl et al., 2011, p. 123). The focus on core knowledge and practices which are fundamental to the domain intends to overcome the problems with the often partitioned teaching of independent topics, modules or units (Duschl et al., 2007; Osborne & Collins, 2000). Although the relevance of fundamental concepts have been emphasized



for many years, research on students' understanding has consistently shown that students fail to obtain a deeper understanding even of the most fundamental concepts, for instance the particulate nature of matter (cf. Löfgren & Helldén, 2009; Morell et al., 2017; Stefani & Tsaparlis, 2009), chemical reaction (cf. Talanquer, 2008; Taskin & Bernholt, 2014), or energy (cf. Herrmann-Abell & DeBoer, 2018; Neumann et al., 2013). Thus, it remains an open question to which extent school science enables students to develop a sophisticated understanding of fundamental concepts in the domain and which learning gains can be expected from year to year (Bloom et al., 2008). Studies have revealed much about the conditions for learning in various domains and how the thinking of high achievers differs from that of novices (Bransford et al., 2000; Chi et al., 1981). „High performers have acquired extensive stores of knowledge and skill in a particular domain. But perhaps most significant, their minds have organized this knowledge in ways that make it highly retrievable and useful“ (p. 84 Pellegrino, 2012). With regard to the current study, conceptual understanding is thus considered to be reflected by a successful application of scientific knowledge in a variety of contextual settings related to a particular concept that is fundamental to the domain (cf. Bransford et al., 2000; Weinert, 2001). Consequently, high achievers should differ from low achievers by demonstrating a more sophisticated understanding of a particular concept, but also in a higher chance of recognizing and applying a particular concept successfully in a given problem context (diSessa & Wagner, 2005; Pellegrino et al., 2001).

### 5.3 Research Questions

As studies addressing the interrelation of interest and achievement often make use of rather global measures, both with regard to interest (e.g., interest in chemistry) and achievement (e.g., course grades or results from standardized tests), the present study intends to provide more differentiated insights into the interplay between both constructs. With regard to students in lower and upper secondary school (i.e., Grades 5 to 11) in Germany, students' interest in school science activities was measured with the adapted version of the RIASEC+N model. While the original RIASEC model (Holland, 1997) focuses on students' vocational interests, the adapted version used in this study aims at capturing students' interest in school science activities in seven dimensions (Dierks et al., 2014b). In addition, students' conceptual understanding was assessed with regard to three fundamental concepts that also provide the structure of the curriculum students are taught to in school (KMK, 2004). Besides the implementation of more differentiated measures, the objects of both constructs (fundamental chemistry concepts and school science activities) are in focus of chemistry teaching across the entire grade span covered in this study, i.e. from grade 5 to 11. While the assessment of students' understanding of or interest in different science topics or contexts is also a plausible account (cf. Häußler et al., 1998), these features generally vary between grades. Even students' perception of a specific subject (and thus the measured construct when operationalized as, for example, interest in chemistry) is considered to change over time (Frenzel et al., 2012). Consequently, the measures applied in this study were selected to provide a sound basis for comparing the results of students' conceptual understanding and interests across the entire time span of secondary school. In summary, both measures were used in a cross-sectional design to address the following three research questions:

1. To what extent do both instruments provide valid and reliable measures for comparing students' interests in school science activities and their conceptual understanding across Grades 5 to 11?

The instrument based on the adapted version of the RIASEC model has been used in different grade levels in different studies (Blankenburg, Höffler, & Parchmann, 2016; Dierks et al., 2014b), but not across the entire grade span of secondary school chemistry. With respect to the first research question pertaining to the validity of the RIASEC+N model, we expect that tests against plausible, alternative models will speak in favor for the assumed seven-dimensional factor structure of the adapted model. Additionally, we expect measurement invariance across grade levels indicating that

the construct is perceived equally by students of different grades and thus comparisons across grades to be meaningful. Comparably, we expect that a rigorous test design which is closely aligned to the curriculum students are taught to will provide a reliable and valid measure of students' understanding of three central concepts from introductory courses to upper secondary Chemistry education.

2. How do both constructs (interest in school science activities and conceptual understanding) develop over the different grades of lower and upper secondary school?

Based on findings from studies focusing on students' interest in the domain of science, generally, it is expectable that the cross-sectional analysis in this study will also indicate an overall decline across grades. With regard to the more differentiated measure based on the seven dimensional RIASEC+N model, however, it is so far unclear whether this decline pertains to all seven dimensions of students' interest in school science activities or whether different patterns can be observed. We expect an increase in students' conceptual understanding across grades. However, due to a lack of studies focusing on the development of students' conceptual understanding with regard to central and curricular anchored concepts on longer timescales (in general but also concerning the concepts of the German educational standards), there are only vague insights concerning expectable learning gains from year to year.

3. How does the relation between interest and conceptual understanding develop over the different grades?

Based on the literature, a positive relationship between interest and conceptual understanding is the expected finding in this study. However, across studies and thus across different conceptualizations of both constructs, the strength of the relation between interest and conceptual understanding varies substantially (Schiefele et al., 1992). So far, the adopted version of the RIASEC+N instrument was not used in combination with an achievement test, so it is unclear whether there will be a common pattern across grades between students' interest in school science activities and their performance in the conceptual test or whether the results suggest a tighter link between interest in specific activities and students' conceptual understanding. Across grades, different studies indicate that the correlation between interest and achievement increases over time (Denissen et al., 2007; Kim et al., 2015; Köller et al., 2001), but differences in the interrelation between interest in specific activities and students' conceptual understanding might alter this overall trend for specific combinations, thus providing a more diverse picture across the different dimensions of school science activities.

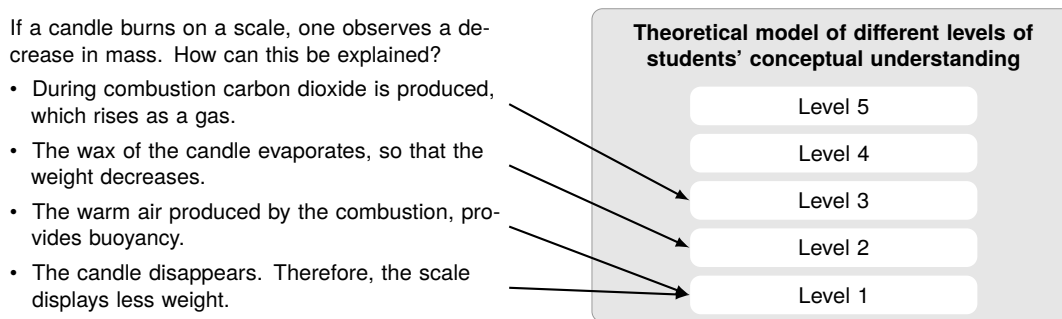
## 5.4 Design and Analysis Procedure

### Instruments

**Interest.** To capture students' interest in school science activities, we used the RIASEC+N instrument developed by Dierks et al. (2014b) and adapted by Blankenburg, Höffler, and Parchmann (2016) and (Dierks et al., 2016). The RIASEC+N model distinguishes between seven dimensions of interest in school science activities (realistic, investigative, artistic, social, enterprising, conventional, and networking activities). Each of the seven dimensions was represented by four items asking for students' interest in the specific activity (see Table B.1). In sum, 28 RIASEC+N items were administered that had to be answered on a four-point rating scale (from „I am not interested at all in doing this“ (1) to „I am very interested in doing this“ (4)). As the instrument has been used in previous studies, including pretests to assure item comprehensibility and an understandable answer format with regard to the target population no further changes or adaptations of the published version of the questionnaire were made for the present study. In terms of validity (Joint Committee, 2014; Messick, 1995), the items were developed to address authentic school scenarios. This was done in close collaboration with teachers from different school types. In addition, think-aloud interviews were performed with students of different grade levels during filling in the questionnaire to ensure that students' understanding and

their answer process matched the researchers' intentions (Dierks et al., 2014b). In consequence, the items cover authentic and representative activities in school science lessons and are phrased in a manner to ensure that students' cognitive processes are in line with the theoretical target construct (cf. content and substantive aspects of construct validity; Messick, 1995).

**Conceptual understanding.** As we intended to investigate the conceptual understanding of a diverse sample, ranging from Grade 5 to Grade 11, we attempted to survey the complexity of students' understanding by combining classical multiple choice items with a rather new item type – so called ordered multiple-choice items (OMC). These OMC items are designed similarly to regular multiple-choice items by providing several response options (see Figure 5.1). One response option is interpreted as correct (i.e., a scientifically acceptable understanding) and corresponds to the highest level of understanding that can be assessed with this item. The remaining options either represent a scientifically inappropriate understanding, usually reflecting common alternative conceptions, or are incomplete, but correspond to specific lower levels of understanding (based on theoretical or empirically validated models of understanding a particular concept; cf. Hadenfeldt et al., 2013). OMC items are assumed to be a valid replacement for open-ended item formats and are additionally less time and money consuming (Alonzo & Steedle, 2009; Briggs et al., 2006).



**Figure 5.1:** Example of an ordered multiple choice (OMC) item for the basic concept „chemical reaction“. Each option of the item corresponds to a specific level of students' conceptual understanding and was scored accordingly based on a partial credit model.

Based on the German educational standards, three concepts (denoted as basic concepts) were selected for constructing the conceptual test: the structure and composition of matter, chemical reaction, and energy (KMK, 2004). Besides newly developed items, about half of the test is based on items that focused on the target concepts and had been empirically validated in previous studies for assessing the conceptual understanding of students across multiple grade levels (Matter and Chemical Reaction: Hadenfeldt et al., 2016; Energy: Sebastian Tobias Opitz et al., 2017). The test construction followed a common item anchor design with grade specific items. In order to populate this design, six items were intended to constitute the anchor to link the results from students in different grades. Besides, four items for each concept and each grade of the target population were selected to address content aspects that are usually taught in a particular grade. Consequently, each participant received the six anchor items and four grade-specific items for each concept, resulting in 30 items per student. An expert panel of four chemistry education researchers and three chemistry teachers reviewed the final item pool (114 items) to select the most suitable ones (in terms of content validity), to check the response options for correctness, ambiguities, or alternative plausible interpretations that could affect between-options relationships, to confirm the rating of the student's level of understanding, and to check the assignment of items to specific grades. To provide additional evidence for a valid interpretation of the test scores, think-aloud interviews were conducted with a sample of items used in this study to ensure that the items are understandable and evoke the intended cognitive processes for solving these items (Hadenfeldt, Repenning, et al., 2014). Besides interest in school science activities and conceptual understanding, the final questionnaire included questions regarding demographic variables (gender, age, first language, SES) and several other constructs (self-concept,

motivation, epistemic beliefs) that are not in focus of this article.

### Data Collection

The data were collected in five secondary schools in Northern Germany. All students were attending the Gymnasium (the highest track of secondary schools in Germany, preparing students for higher studies). The sample consisted of  $N = 2510$  students (56% female; 12% speaking predominantly a foreign language at home) from 136 classes, ranging from Grade 5 to 11. The five selected schools are a convenience sample of schools that were willing to participate in this study. Within these schools, all classes were included in the data collection. Although the sample was not drawn as a representative sample for the federal state or the country, basic demographic information indicate a comparable composition of our sample (in terms of gender, migration background, and SES) and sample characteristics reported in representative studies on state level (cf. Leucht et al., 2016). Science education in Germany normally commences in fifth grade, with science subjects being taught as a combination of distinct science subjects (e.g., physics and biology) or as an integrated science instruction in Grades 5 and 6 (Möller, 2014). Thereafter, science subjects are taught in a differentiated manner, with biology, physics, and chemistry as separate subjects. After Grade 9, students have the option to drop-out from taking chemistry courses. In our sample, 78% of the students in Grade 10 and 68% of the students in Grade 11 attended chemistry classes (see Table B.2).

As the constitution and also the quality of school lessons depends on multiple factors, e.g., the teacher's background, class and school composition, or school climate, the teaching usually differs between classes and schools. Although there is certainly variance between teachers, video studies indicate that „a typical German chemistry lesson can be characterized by a short initial (1.8% [of the teaching time]) and repetition phase (3.2%) and a longer consolidation phase (6.0%). The development phase (83.4%) represents the longest part of the lessons, in which new content is taught. It is mostly dominated by independent seatwork (41.4%) and class discussion (32.2%). Frontal teaching social forms (class discussion (32.2%) and teacher talk (18.6)) occur most of the time. In general, class discussion (35.8%) represents a very dominant teaching method with the German chemistry lessons“ (Björkman and Tiemann, 2013, p. 3; Stiller, 2016). Although different projects tried to accomplish a shift in communication opportunities and agency for the benefit of students (e.g., context-based learning approaches that emphasize student-centered activities and decisions during the learning process; (Menthe & Parchmann, 2015)), the role of the teacher is still rather dominant. Moreover, summative and content-oriented assessment approaches are prevalent with an emphasis on factual knowledge, both within class as well as in centrally administered tests (Kühn, 2016). More formative approaches of performance assessment and feedback are seldom applied in school practice (OECD, 2005). With regard to lab work, science lessons in Germany can generally be characterized by a regular implementation of experiments, even when these occasions are dominated by carrying out prescribed, recipe-like experiments and seldom by developing own experiments. Overall, science education in Germany is usually characterized by social activities, such as introducing one's own ideas and discussing scientific questions or experiments (Schiepe-Tiska et al., 2016).

The study was conducted in February to April 2015 during regular lessons based on a standardized protocol by trained personnel, i.e. by the research team, not by teachers of the school. After the students received a brief introduction into the study, questionnaires were administered, instructions given and clarification questions taken. The students had sufficient time for filling in the questionnaire, which took about 120 minutes for all constructs. Students participated on a voluntary basis, with the parents' written consent, and were informed that they could drop-out from the study at any time. The mean participation rate per school was .81, ranging from .55 to .90. All collected data were anonymized and no individual results were provided to the participants, their teachers, or the schools, so that neither participation nor non-participation would have immediate consequences for the students.

## Data Analysis

Interest. To analyze the construct validity of the RIASEC+N model, confirmatory factor analysis (CFA) models with a maximum likelihood estimator (see Brunner et al., 2012; MacCallum & Austin, 2000) were estimated in Mplus 7.4 (L. K. Muthén & Muthén, 1998-2015). In order to ensure construct validity we applied goodness-of-fit and absolute-fit indices to evaluate the fit of the CFA model to the data. Here, values of the comparative fit index (CFI) and Tucker-Lewis index (TLI) greater than .90 and values of the root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) not exceeding .08 were interpreted as indicating an acceptable to good fit (Browne & Cudeck, 1993; Hu & Bentler, 1999; Marsh et al., 2010). We neglected the chi-square difference test as it is overly sensitive to sample size (Marsh et al., 1988). The questionnaire data met the normality assumption with absolute values for skewness and kurtosis of less than 1 for all variables. The investigation of internal consistency as an estimate of reliability for all latent variables showed acceptable to excellent Cronbach's alpha values ( $\alpha = .67 - .87$ ) after deleting one item per dimension because of poor item-scale correlation (Table 5.1). Additionally, score reliabilities ( $\rho$ ) defined as the ratio of true variance to observed variance were calculated for each scale following the approach described by Raykov (2009). The obtained coefficients ( $\rho = .68 - .87$ ) corroborate the assumption of sufficient to good reliability of the seven interest scales.

**Table 5.1:** Reliabilities (Cronbach's  $\alpha$ , McDonald's  $\omega$ , reliability coefficient  $\rho$  of the LTRM) and Descriptive Statistics for Interest. M and SD are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each reliability coefficient.

Scale	n Items <sup>a</sup>	$\alpha$	$\omega$	$\rho$	95% CI	M <sup>b</sup>	SD
Realistic	3	.67	.67	.68	[.65, .71]	2.91	0.91
Investigative	3	.74	.74	.74	[.71, .76]	2.77	0.92
Artistic	3	.72	.78	.74	[.71, .76]	2.55	0.97
Social	3	.87	.87	.87	[.86, .88]	2.40	0.87
Enterprising	3	.77	.77	.76	[.74, .78]	2.64	0.95
Conventional	3	.77	.77	.76	[.74, .79]	2.29	0.90
Networking	3	.78	.79	.78	[.76, .81]	2.55	0.93

a One item per scale was deleted because of poor fit.

b Higher values indicate higher interest.

To investigate whether the items provide valid measures of students' interest in school science activities, a theory-based CFA model with seven RIASEC+N activities first-order factors (realistic, investigative, artistic, social, enterprising, conventional, and networking) was constructed (Blankenburg, Höffler, & Parchmann, 2016; Dierks et al., 2016; Dierks et al., 2014b). This model had the following characteristics: (a) each activity measure had a nonzero loading on one first-order RIASEC+N factor but zero loadings on the others; (b) correlations among the factors were allowed, and (c) factor loadings were freely estimated and factor variances were constrained to 1. In this model, higher scores on a particular RIASEC+N factor are connected with higher scores on all activities related to that factor. The resulting model showed a good fit to the data ( $\chi^2 = 1458.187$ ,  $p < .001$ ,  $df = 167$ , CFI = .938, TLI = .922, RMSEA = .058, SRMR = .047). This indicates that the instrument is sensitive to the different activity types and adds support for a valid assessment of the assumed structure of students' interest in school science activities.

The correlations between the RIASEC+N factors vary from  $r = .34$  to  $.87$ . The factor loadings on the RIASEC+N factors varied from  $\lambda = .48$  (artistic) to  $\lambda = .85$  (social) with a median loading of Mdn  $\lambda = .73$  (all factor loadings were significant). Thus, the latent variables exhibit strong effects on most subscales meaning that interest in an activity of a certain RIASEC+N dimension leads to higher interest in all activities connected to this dimension (see Figure B.1). To compare students from different grades by analyzing latent values, strong measurement invariance had to be ensured (Wu et al., 2007). Firstly, seven CFAs were conducted for each grade separately, indicating an appropriate fit of the model in each grade (see Table B.3).

Next, we tested for measurement invariance to ensure that the items measure the same theoretical

construct across all grades. Multiple-group confirmatory factor analyses were sequentially evaluated for models of configural, metric, scalar and strict invariance, whereby each step adds more restrictions to the baseline model (Wang & Wang, 2012): Configural invariance tests for equal factor loading patterns across groups; metric invariance requires equal factor loadings; scalar invariance requires both invariant factor loadings and item intercepts; and strict invariance requires equal factor loadings, item intercepts, and error variances across groups. The goodness-of-fit indices for the first three models are consistently acceptable to good whereas the strict invariance model exhibits a poor fit (see Table B.4). For testing measurement invariance, the first three model fits were compared by CFI and RMSEA difference testing. Hereby, a change of  $< .01$  in CFI supplemented by a change of  $< .015$  in RMSEA between subsequent models would indicate invariance (G. W. Cheung & Rensvold, 2002). Using this criterion, we could demonstrate scalar measurement invariance, as  $\Delta\text{CFI}$  was smaller than  $.01$  and  $\Delta\text{RMSEA}$  was smaller than  $.015$  between all consecutive models (see Table B.4). However, testing for strict invariance failed due to a poor fit of the fourth model and a change in CFI of  $-.025$ , which is above the threshold. Therefore, we have used latent, rather than absolute, means in our comparisons across groups (Steinmetz, 2011; Vandenberg & Lance, 2000).

Conceptual understanding. In order to account for the matrix design of the test (i.e., not all items having been administered to all students), we utilized Item Response Theory (IRT) to analyze the data. Students' answers to multiple-choice items were coded as incorrect (0) or correct (1). In case of OMC items, answers were assigned partial credits from 0 to 3 in correspondence to the theoretical model. For instance, answer options that were assigned to the lowest level of the model were coded 0, the next level was coded 1, and so on.

A one-dimensional multi-group generalized partial credit model (MG-GPCM, as implemented in the R package TAM; Test Analysis Modules; (Kiefer et al., 2017)) was used for analyzing the coded results of student answers, with students' grade (from 5 to 11) as grouping variable. The expected a posteriori based on plausible values reliability (EAP/PV) was calculated and found to be 0.88, which indicates a high reliability of the measured construct (Field, 2013). The MG-GPCM analysis then yields so called weighted likelihood estimates (WLE) which reflect the persons' conceptual understanding, as measured by the test, on an interval scale. With respect to item fit, Weighted Mean Square (WMNSQ) estimates and standardized mean-square fit statistics (i.e., T values) were considered. One item that did not meet typical cut-off values (WMNSQ from 0.7 to 1.3; Wright et al., 1994) was excluded from further analysis. Taken together, the results suggest that the test provides an adequate measure for students' understanding of three central chemical concepts across Grades 5 to 11.

Interplay between interest and conceptual understanding. To investigate the relationship between students' interests in the seven RIASEC+N dimensions and students' conceptual understanding as measured by our test, we included students' WLE parameter from the multiple-group generalized partial credit model as a manifest variable into the multiple-group CFA model and correlated the seven latent variables (RIASEC+N) with the WLE score (MIMIC approach; cf. Fan, 1997). The resulting model showed a good fit to the data ( $\chi^2 = 3198.324$ ,  $p < .001$ ,  $df = 1476$ , CFI = .917, TLI = .909, RMSEA = .060, SRMR = .079); the Mplus syntax of the MIMIC approach is included in the supplementary materials of the paper.

## 5.5 Results

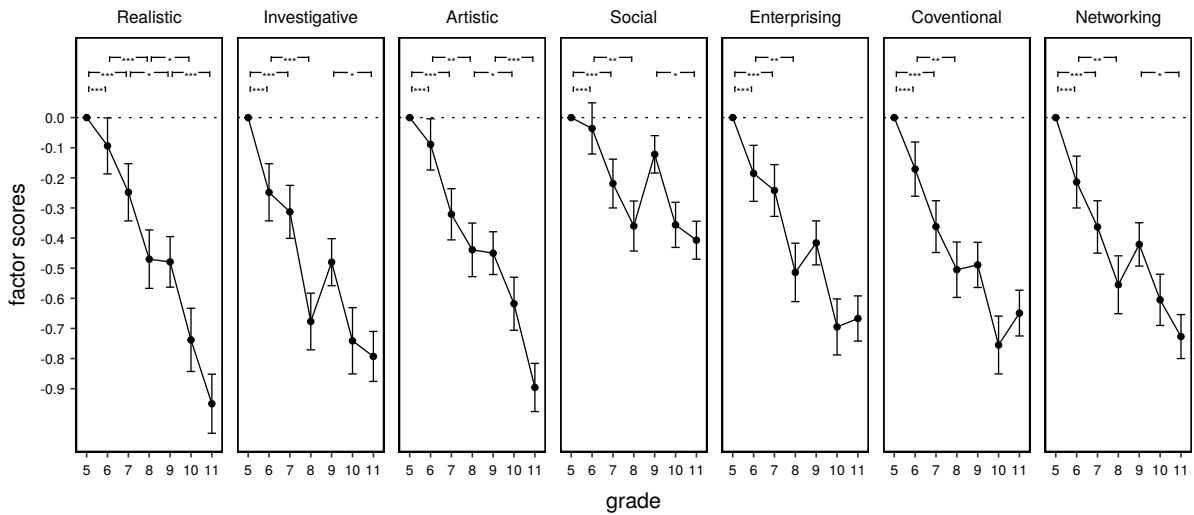
### Trends over time

As the main goal of this study is to compare the interrelation of interest and conceptual understanding across the Grades 5 to 11, it is important to ensure the comparability of the constructs across this grade span. Students in Grades 10 and 11 generally have the option to drop-out from chemistry (cf. Table 5.1). With regard to the conceptual test, the results of students who dropped out

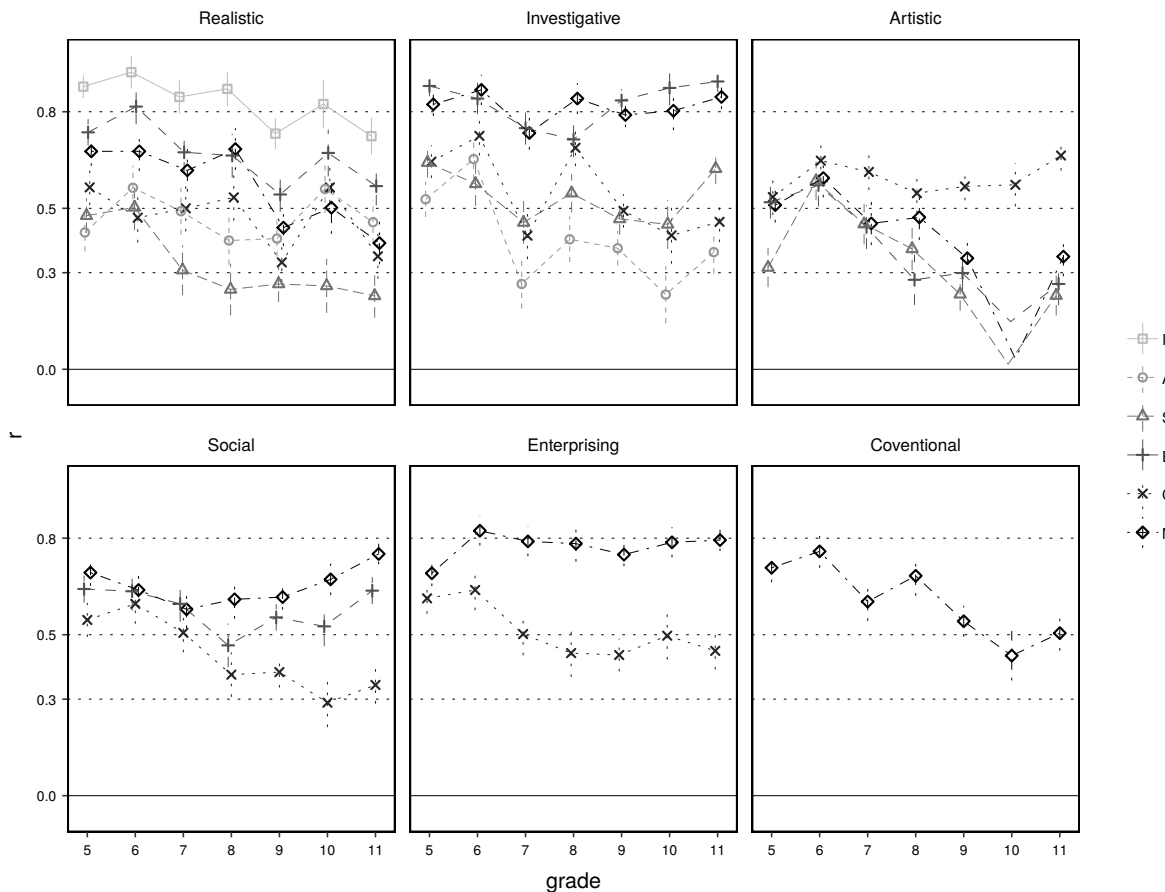
of chemistry courses are on the same level as the average result of students in Grade 9. Hence, no learning gains can be observed when comparing the results of students in Grade 9 to students who do not attend chemistry courses in Grade 10 (mean WLE difference 0.06,  $t(91.85) = -0.86$ ,  $p = .39$ ) or Grade 11 (mean WLE difference 0.08,  $t(300.35) = -1.55$ ,  $p = .12$ ). The results of these students are cumulatively falling behind when compared to their peers attending chemistry courses (Grade 10: mean WLE difference 0.28,  $t(126.73) = -3.86$ ,  $p < .001$ ; Grade 11: mean WLE difference 0.42,  $t(336.57) = -7.68$ ,  $p < .001$ ). As attending or not attending chemistry courses in upper secondary obviously has an effect on students' results on the conceptual test, the following steps of the analysis focus on students who are provided with adequate learning opportunities in Grades 10 and 11 ( $N = 2294$ ), thus excluding students who decided to drop out from chemistry.

**Interest.** As to students' interest in school science activities, the results follow the pattern known from many studies: students' interest declines more or less monotonically from grade to grade (see Figure 5.2). Although the instrument in this study allows for a more fine-grained differentiation of students' interest in seven dimensions of different school science activities based on the RIASEC+N model, the overall trend is comparable in all dimensions. While students' evaluations of their interests, with regard to the different school science activities, decline significantly on a year-to-year rate only sporadically, substantial drops can be observed in all seven dimensions about every two years. Overall, students in Grade 5 indicate a significantly higher interest in all seven dimensions than students in upper secondary. Regarding the individual dimensions, there are two types of deviations from a monotonic decline present in almost all dimensions: first, students' interest in Grade 9 is unexpectedly high compared to the other grades (pertaining to the realistic, social, enterprising, conventional, and networking dimensions), resulting in an increase of students' interest from Grade 8 to 9. Second, eighth grade students' interest in investigative activities is overly low. Only in the artistic dimension, students' interest declines from grade to grade. With regard to the interrelations of students' interest between the seven dimensions of the RIASEC+N model, certain patterns emerge (see Figure 5.3). First, the correlation coefficients between the seven latent factors start relatively high in Grade 5, with values between .5 and .9. Only the correlations between artistic and social as well as artistic and realistic start with a value below .5 ( $r(AS) = .35$ ;  $r(AR) = .48$ ). Almost all correlation coefficients increase from Grade 5 to 6, before dropping again in Grade 7 (when chemistry teaching usually starts). While correlations between the realistic, investigative, enterprising, and networking dimensions stay on a high level throughout the grades, these four dimensions exhibit progressively lower correlations to the social, artistic, and conventional dimensions.

**Conceptual understanding.** In a first step, we investigated the progression of students' understanding of the three chemical concepts covered in the test. As a first indicator, we calculated the correlation between students' grade and their WLE parameter, resulting in a high correlation between both variables ( $r(2476) = .68$ ,  $p < .001$ ). This suggests that independent of differences in school curricula and classroom instruction, there is, as expected, a general trend towards higher levels of understanding with higher grades (see Figure B.2). We then compared the mean WLE parameter across the Grades 5 to 11. Based on an ANOVA, there was a significant effect of the grade level on students' WLE parameter,  $F(6, 890.46) = 387.56$ ,  $p < .001$ ,  $\eta^2 = .50$ . This result is supported by planned contrasts, indicating that students' understanding significantly progresses between Grades 6 to 11, while the mean difference between Grades 5 and 6 is non-significant (Table 5.2). Concerning the effect sizes for the average annual gains on the WLE parameter, large effects can be found for grade transitions from 6 to 7 ( $d = .84$ ) and 8 to 9 ( $d = .73$ ), respectively, followed by the transition from 7 to 8 with an (almost) middle-sized effect ( $d = .43$ ). The gains from Grade 9 to 10 and 10 to 11 are only small effects, while transitioning from 5 to 6 is not associated with a significant increase in gains on the conceptual test. To facilitate the interpretation of the obtained effect sizes for the different grade transitions, we compared them to the results of a meta-analysis based on average annual gains on nationally normed science tests (cf. Table 5.2; Bloom et al., 2008). The calculated effect sizes in this



**Figure 5.2:** Mean factor scores and corresponding standard errors of students' interest in school science activities across Grades 5 to 11 based on the MG-CFA. Values in Grade 5 (reference group) are set to 0 due to identification reasons. Significant differences between grades are reported for comparisons between grades within two years of each other (\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ ).



**Figure 5.3:** Latent correlations between the seven dimensions of students' interest in school science activities according to the RIASEC+N model (MG-CFA) across Grades 5 to 11 (R = Realistic, I = Investigative, A = Artistic, S = Social, E = Enterprising, C = Conventional, N = Networking). All correlation coefficients are significantly different from 0, with the exception of the correlation coefficients between the artistic and social as well as between the artistic and networking dimensions in Grade 10.

study are substantially larger for transitions between Grades 6 and 10, while the effect size for the transition from Grade 10 to 11 is comparable. The effect size for the transition from Grade 5 to 6

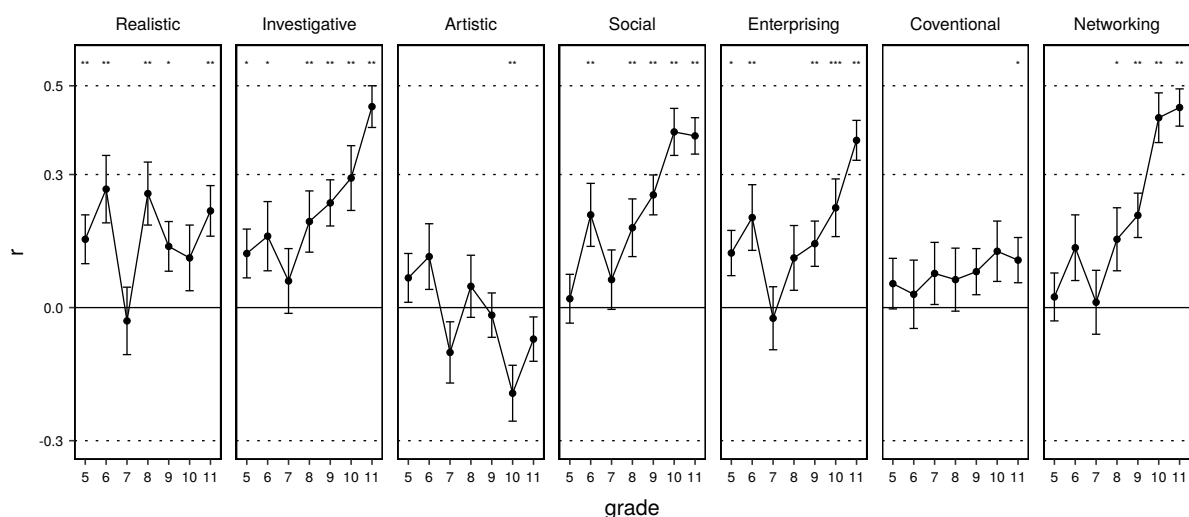


is substantially lower. Overall, both the effect sizes in our study and the reported benchmarks in Bloom et al. (2008) are lower in higher grades (when disregarding the transition from Grades 5 to 6 in this study). However, the reported benchmark effect sizes follow a quite constant trend in a small range of values ( $0.15 < d < 0.28$ ), while the effect sizes in our study escalate from Grade 6 to 7 (when teaching chemistry as a distinct subject usually begins in Germany), remain on a high level until Grade 10, before decreasing substantially when students transition to upper secondary. In addition, the range of the obtained effect sizes is much larger ( $0.08 < d < 0.84$ ) when compared to the benchmarks in Bloom et al. (2008).

**Table 5.2:** Average annual gains per grade transition, based on an ANOVA with planned contrasts, and corresponding effect sizes compared to benchmarks from a meta-analysis.

Grade transition	Mean difference (WLE)	$p$	Cohen's $d$	Effect size benchmarks <sup>a</sup>
5 – 6	0.08	.058	0.18	0.27
6 – 7	0.39	< .001	0.84	0.28
7 – 8	0.21	< .001	0.43	0.26
8 – 9	0.40	< .001	0.73	0.22
9 – 10	0.16	< .001	0.35	0.19
10 – 11	0.09	.016	0.19	0.15

a Average annual gains in effect size for science from nationally-normed tests (Bloom et al., 2008, p. 305)



**Figure 5.4:** Latent correlations between students' conceptual understanding (WLE parameter of the MG-GPCM) and students' interest in school science activities according to the RIASEC+N model (MIMIC model). Correlation coefficients were tested to be significantly different from 0 (\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ ).

Interplay between interest in school science activities and conceptual understanding. When including the WLE parameter from the IRT analysis as a covariate into the MG-CFA model (MIMIC approach), it was possible to estimate the latent correlations between students' conceptual understanding and students' interest in school science activities in the different grades. Overall, the correlation between conceptual understanding and interest increases with grade but these interrelations differ much between the different dimensions (see Figure 5.4). In general, the course of the correlation coefficients across grades subdivides the seven dimensions into two groups. In one group (realistic, artistic, conventional), correlation coefficients are generally lower than .3, indicating an only weak connection between students' conceptual understanding and their interest in these school science activities. While the correlation between students' conceptual understanding and their interest in realistic and conventional items remains on a quite constant level (despite fluctuations across grades), the correlation between conceptual understanding and students' appraisal of artistic activities slightly

declines. In the second group (investigative, social, enterprising, and networking), a weak connection between conceptual understanding and interest can also be observed in lower grades. However, values for the correlation coefficients gradually increase from lower to higher grades, resulting in substantial correlations in Grades 10 (social and networking) and 11 (additionally investigative and enterprising). Here, correlation coefficients fall into the range of .3 to .5, indicating a middle-sized connection between both constructs. Across the grades and with the exception of the conventional dimension, all correlation coefficients undergo a substantial decline in Grade 7, when chemistry teaching usually starts.

## 5.6 Discussion

The goal of the present study was: (a) to investigate the validity and reliability of the measures for students' interests in school science activities as well as students' conceptual understanding across Grades 5 to 11, (b) to analyze the development of students' interest in school science activities and their conceptual understanding across this grade span, and (c) to scrutinize the relationship between these two constructs in the different grades. Therefore, a cross-sectional study was conducted collecting data about students' conceptual understanding of three basic concepts of chemistry (energy, matter, and chemical reaction) and their interests in school science activities, based on the RIASEC+N model, an adapted version of Holland's RIASEC model of vocational interests (Holland, 1997).

### Validity and reliability

With regard to the overall design and analysis procedure, multiple evidence and arguments with regard to the content of test and questionnaire items, item phrasing, cognitive processes, internal structure, and consequences of testing were collected and discussed (Joint Committee, 2014; Messick, 1998). In addition, further validity aspects, such as the relation of the measured constructs to external criteria, were discussed in the results section. In summary, we consider the utilized instruments to provide reliable and valid measures of students' interests in school science activities and conceptual understanding over the school years 5 to 11 and, consequently, to be well suited to describe almost the entire period of chemical education in secondary schools with regard to these constructs (cf. Research Question 1). Scalar measurement invariance allowed us to make comparisons across grades and to investigate correlations between conceptual understanding and interest at this more precise level, providing insights into the interplay between the two constructs. Hence, this study adds further evidence for the suitability of the RIASEC+N model to describe the multidimensional structure of students' interests in school science activities (cf. Blankenburg, Höffler, & Parchmann, 2016) and, moreover, enhances previous findings through the conjunction with learning outcomes.

### Development of interest in school science activities

Similar to findings reported in the literature, students' interest in school science activities significantly decreases in all interest dimensions. This is particularly notable as our sample represents a positively biased selection of students neglecting students who elected not to follow chemistry courses in higher grade levels. With regard to the RIASEC+N model, high intercorrelations ( $r > .8$ ) exists between i) the investigative and the realistic, enterprising, and networking dimensions as well as between ii) the enterprising and networking dimensions. Based on the underlying theoretical model (Dierks et al., 2016; Holland, 1997) these dimensions address different, but highly interrelated, aspects of scientific inquiry Rönnebeck et al., 2015, which in turn might be reflected by the strong ties between these activities. A more detailed consideration of the observed decline of interests reveals striking deviations in Grade 9 (realistic, social, enterprising, conventional and networking) and Grade 8 (investigative). The setback of students' interest in the investigative dimension in Grade 8 might be partly explained by the beginning of chemistry education in Grade 7 and a shift of focus towards scientific ideas and reasoning incorporating demanding and tedious tasks in the subsequent grade.

Besides developmental effects, for instance, the gender-intensification and differentiation of interests hypotheses (Krapp & Prenzel, 2011), variations of class compositions and the customary assignment of new teachers in Grade 9 are factors likely to contribute to this pattern. New class compositions and new teachers are accompanied by new social interactions, deviating styles of instruction and potential changes in class climates which has been shown to impact on achievement and interest development Harks et al., 2013; Keller and Reintjes, 2016; Kunter et al., 2013.

#### Development of conceptual understanding

While students' interests in school science activities decreases across the grades, their conceptual understanding increases. These learning gains are primarily achieved in Grades 6–9 and levels out from Grade 9 and onwards, which is in accordance with recent findings (Hadenfeldt et al., 2016; Liu & Lesniak, 2005). Based on Cohen's *d* as effect size to evaluate students' learning gains, we compared the average increase in students' conceptual understanding for each grade transitions between Grades 5 to 11. The effect sizes were compared with the results of a meta-analysis based on average annual gains on nationally normed science tests (cf. Table 5.2; Bloom et al., 2008). Although the meta-analysis refers to science, and not to chemistry as an individual disciplinary subject, as well as to a different country (US instead of Germany), the benchmarks reported in (Bloom et al., 2008) provide a reference to interpret the learning gains observed in our study. Here, only the effect size for the transition from Grade 10 to 11 is on a comparable level. The low effect size for the transition from Grade 5 to 6 in this study is probably due to the fact that teaching chemistry regularly starts in Germany in Grade 7. Although students in Grades 5 and 6 are attending either science or biology and/or physics courses, depending on the timetable of their school, these courses seem to provide only limited learning opportunities with regard to the concepts covered in the test. From the onset of chemistry teaching in Grade 7, effect sizes escalate to large effects for the transitions from Grade 6 to 7 (probably due to having dedicated chemistry courses for the first time) and from Grade 8 to 9 (probably due to a strong emphasis on theoretical models for describing the structure, composition, and change of matter, hence, providing intensive learning opportunities for the concepts covered in the test). In general, the effect sizes for the annual learning gains for transitions between Grades 6 and 10 are generally higher than effect sizes reported in Bloom et al. (2008; Table 5.2). This result that can probably be traced back to specific discrepancies between the meta-analysis and this study, e.g. students' learning opportunities in a disciplinary subject chemistry (instead of science), the test design focusing on core chemical concepts (instead of a broader test on concepts relevant in different science disciplines; cf. Bloom et al., 2008), and the alignment between curriculum and test (Liu, 2012). However, the results obtained in this cross-sectional study also indicate that the test is sensitive to students' learning in chemistry across the Grades 5 to 11, supporting the assumption that the developed instrument provides a thorough measure of students' conceptual understanding across this grade span.

#### Interplay between interest in school science activities and conceptual understanding

The main findings obtained in this study are summarized in Figure 5.4, depicting correlations between students' conceptual understanding and the distinct interest dimensions. The findings reveal small to middle-sized positive correlations between conceptual understanding and interest in school science activities, with the exception of one negative correlation for the artistic dimension in Grade 10. At the beginning of secondary school, interest is mainly shaped by expectancies and experiences in early science (primary school) and integrated science courses (Grades 5 and 6). From Grade 7 onwards, interest in school science activities is affected by education in biology, physics, and increasingly by experiences in chemistry lessons. Consequently, there may be a re-evaluation process of interests in Grade 7 which leads, at first, to lower correlations between interests and conceptual understanding. Due to incipient chemistry education, students encounter new and additional learning opportunities and, hence, the acquisition of conceptual understanding is more and more taken over by instruction

in chemistry, resulting in increasing correlations between interest and conceptual understanding. After Grade 9, students have the option to drop-out from taking chemistry courses. With regard to our sample, only students attending chemistry courses in upper secondary are included, which constitutes a selection of students being more interested and achieving higher (although the results are almost identical when including the full sample into the analysis).

On closer inspection, Figure 5.4 also reveals differing interrelations between conceptual understanding and interest:

1. small to middle-sized, positive relations, increasing from Grade 7 to 11, between interest in investigative, social, enterprising, and networking activities and conceptual understanding;
2. small positive, but almost stable relations between interest in realistic activities and conceptual understanding; and
3. negligible correlations between conceptual understanding and interest in the artistic and conventional dimensions. Taking into account findings from previous research and a closer examination of the wording of the items, it seems plausible to assume multiple factors contributing to this differentiation.

With regard to 1., items of these scales (investigative, social, networking, enterprising) reflect activities associated with cognitive activation (e.g., involving cognitive conflict, or problem-solving) and fostering knowledge acquisition (Kunter et al., 2013). Interest in these activities manifests itself in superior learning strategies, a higher degree of persistence, effort and attention (Hidi & Harackiewicz, 2000; Hidi & Renninger, 2006; Renninger, 2000; Trautwein et al., 2015). Students being interested in these kinds of activities thus tend to develop a more sophisticated understanding, as reflected by the escalating correlations between interest and conceptual understanding. Likewise, students' with a more sophisticated conceptual understanding are more likely to experience competence and to act autonomously on the basis of their knowledge. Furthermore, it may enable them to appreciate the personal relevance of the current topic and their acting in class and, hence, both subject and acting are valued more highly so that altogether the probability to develop a more stable interest in associated activities increases (M. Ainley & Ainley, 2011a; Harackiewicz et al., 2016). Additionally, enterprising activities refer to leading roles which might be more appealing to students with a high academic self-concept and a more sophisticated understanding (Wigfield & Eccles, 2000). A growing importance, a higher frequency of and spending more time on such tasks (e.g., group presentations) in higher grades, might be auxiliary factors contributing to the increasing coupling between conceptual understanding and interest in enterprising activities. In addition, social learning, perceived competence and social relatedness, especially when regarding the social, networking and enterprising dimensions, might contribute to the greater coupling in higher grades.

As for 2., the realistic dimension, the findings indicate small positive correlation between interest and conceptual understanding. Items of the realistic scale merely focus on performing given lab experiments, instead of designing and interpreting own experiments (these kinds of activities belong to the investigative scale). Links between the experiment and scientific ideas are not always explicit, neither in instruction or in students actions (Abrahams & Reiss, 2012). Without scaffolding, students may struggle to establish these links, thus, just doing appears meaningless for the students (Abrahams & Reiss, 2012; Schwichow et al., 2016; van Duzor, 2016) and may be neither appealing nor effective for their learning. Indeed, the contribution of laboratory work to students' conceptual understanding (Abrahams & Reiss, 2012) as well as to students' interest (Holstermann et al., 2010) has been questioned, making the weak correlation between interest in laboratory work and conceptual understanding less surprising. However, it is important to notice that items of the realistic dimension of our questionnaire solely focus on the hands-on part of experiments, not the „minds-on“ aspect of laboratory work (Hofstein & Lunetta, 2004; Pickering, 1980) which has been shown to impact on conceptual understanding (Duit et al., 2008; Hofstein & Lunetta, 2004).

With regard to 3., neither the artistic nor the conventional scale include activities highly associated with cognitively demanding tasks or feelings of relatedness; instead, the focus is on hands-on or, in addition, individual work without enriching elements of communication.

## Limitations

Although all of these different factors that have been reviewed from the literature to explain the findings in this study, it is not possible to investigate their relative importance with regard to the obtained results. The study design of a cross-sectional study does only provide insights of trends in averaged results across different cohorts. Hence, the results obtained in this study cannot be interpreted in terms of developmental trends on the level of individual students.

In addition, the correlational nature of the statistical analysis that is applicable to the type of data collected in this study provides no insights into the mechanism behind the reported interrelations between interest and conceptual understanding. Hence, it is not possible to decide on the basis of our study whether interest fosters conceptual understanding or vice versa. In fact, longitudinal studies indicate small reciprocal effects between interest and achievement (Garon-Carrier et al., 2016; Harackiewicz et al., 2008; Marsh et al., 2005), indicating a back and forth between both constructs.

Finally, the sample selected for this study entails specific limitations. Students were not drawn as a representative sample for the federal state or the country, so the results are not representative in this respect as well. In addition, all students included in this study were attending the German gymnasium, the higher of two tracks of secondary schools in Germany. Consequently, students' performance on the achievement test as well as probably their interest as indicated by the questionnaire do not cover the full variance that could be observed in the student population.

## Implications

With regard to teaching practice, the cross-sectional results underline the important link between students' interest and achievement in chemistry. Substantive correlations between interest and achievement were only found in higher grades and only for interests in school science activities which are associated with cognitive activation or the communication of knowledge (i.e., investigative, social, enterprising, and networking activities). These activities seem to provide productive starting points to address students' interests as well as fostering their understanding. Increasingly using related teaching methods, e.g. project-based learning (cf. Harris et al., 2015) or peer-tutoring (cf. Ding & Harskamp, 2011), could offer the opportunity to profit most from this close interrelation between students' interest in specific activities and their achievement. With regard to science lessons in Germany, but probably also with regard to other countries, these teaching approaches cannot be considered as standard repertoire (Risch, 2010; Schiepe-Tiska et al., 2016)).

However, it does not seem to be solely the inclusion of specific activities in chemistry lessons that impacts on students' interest as, for instance, the different pattern between students' conceptual understanding and their interests in realistic and investigative activities illustrates. The correlation between students' conceptual understanding and their interest in the hands-on part of experiments (i.e., realistic activities) is negligible in all grades, while medium to high correlations can be found for the minds-on part (i.e., investigative activities). This is in line with previous findings questioning the unconditional contribution of laboratory work to students' conceptual understanding (Abrahams & Reiss, 2012) or interest (Holstermann et al., 2010). While „minds-on“ laboratory work has been shown to impact on conceptual understanding (Duit et al., 2008; Hofstein & Lunetta, 2004), it seems to be generally important to consider a minds-on component in all school science activities when students' understanding is an issue. While some activities might be more appealing for students, higher engagement, concentration, and persistence will only be sustained on longer timescales by a perception of self-determined behavior (Deci & Ryan, 1985a) and of competence (cf. Eccles, 2009; Marsh et al., 2016). Consequently, students need to perceive the activity as productive and rewarding, either personally for themselves or for experiencing themselves as competent agents with regard to the demands of the course.

The assumption that implementing specific activities might provide a productive starting point for fostering both students' interests and understanding is supported by the results obtained in this study as well as findings from the literature reviewed in the discussion. However, the cross-sectional design

and the correlational nature of the obtained results provide no insights into the actual mechanism between interest and conceptual understanding. While some studies indicate small reciprocal effects between interest and achievement (Garon-Carrier et al., 2016; Harackiewicz et al., 2008; Marsh et al., 2005), there is a need for further longitudinal studies that provide insights into the development of both constructs as well as their interplay on the individual level.

In addition, further research needs to address the impact of classroom variables and practices and how these influence students' interest in specific school science activities. Intervention studies could provide more evidence with regard to the question of whether the implementation of specific activities actually fosters students' interest, students' understanding, or even both. This kind of studies seems to be particularly important in crucial grade levels (e.g., when students are making choices for upper secondary school).

# Longitudinal couplings between interest and conceptual understanding in secondary school chemistry: An activity-based perspective

## Abstract

This paper presents the findings of a three-year longitudinal study (Grades 9–11;  $N = 756$  students in German secondary schools; mean starting age = 14.7 years) that explores the development of and the interplay between upper-secondary students' individual interest and conceptual understanding in chemistry. Students' individual interest was assessed in seven dimensions of school science activities (realistic, investigative, artistic, social, conventional, and networking) according to Dierks et al.'s RIASEC+N model. The data was collected annually using paper-and-pencil questionnaires. In order to analyse the longitudinal relation between students' academic interest and conceptual understanding, Eccles et al.'s Expectancy-Value Theoretical Model of Achievement Choices (EEVT) was utilised and seven bivariate cross-lagged models were performed to investigate the assumed reciprocal relations between both constructs. Regarding the developmental progression of individual interest and conceptual knowledge, latent growth curve models indicate a small decrease in all of the seven dimensions of students' interest in school science activities and concomitant a moderate growth of students' conceptual understanding. Moreover, structural equation modelling suggests weak reciprocal relations between students' conceptual understanding and their interest in investigative, social and networking school science activities, respectively. Furthermore, students' interest in enterprising activities predicted their subsequent conceptual understanding. No systematic pattern of cross-lagged effects between the three remaining interest dimensions and conceptual understanding were found. The results indicate that school science activities, which provide the potential for cognitively activating learning opportunities, could enhance the relation between students' interest and conceptual understanding. Implications for teaching practice and further research are discussed.

**Keywords:** Interest, conceptual understanding, achievement, upper-secondary, chemistry, longitudinal, cross-lagged effects, school science activities, RIASEC+N

Höft, L., & Bernholt, S. (2019b). Longitudinal couplings between interest and conceptual understanding in secondary school chemistry: An activity-based perspective. *International Journal of Science Education*, 19(4), 1–21

## 6.1 Introduction

Fostering interest and the acquisition of a sound conceptual understanding are crucial goals of chemistry education. Both interest and knowledge lay the foundations for a meaningful engagement with socio-scientific issues and for pursuing a science-related career after graduating from school (Stoll et al., 2017). Hence, numerous investigations focus on these two constructs in the field of science education research and abundant studies have revealed a close connection between interest and achievement. In general, students with a higher interest in a subject are more likely to perform better in this specific subject, particularly in higher grades (cf. Krapp & Prenzel, 2011; Potvin & Hasni, 2014b; Schiefele et al., 1992). With regard to the underlying mechanism of the interrelation between interest and achievement, i.e. the direction of the interrelation, previous findings are still inconclusive and scarce. Furthermore, both constructs were often measured at a rather distal level, e.g. students' interest was captured as interest in a subject and achievement in terms of school grades in this subject (e.g. Köller et al., 2001; Marsh et al., 2005; Rotgans & Schmidt, 2017c). In addition, there is little longitudinal research examining the relation between students' interest and achievement in upper-secondary science courses with more than two measurement points. However, the need for longitudinal data is a frequent desideratum put forward in the conclusion section of research in this field as sound results are expected to provide teachers with crucial hints for lesson planning when aiming to foster students' interest and knowledge. Basically, three constellations regarding the long-term interplay between interest and knowledge acquisition are discussed in the literature, namely: (a) prior achievement fosters future interest (Ganley & Lubienski, 2016) and thus teachers should

mainly focus on promoting students' knowledge or (b) prior interest fosters future achievement (Jansen et al., 2016) and therefore it is more advisable to firstly concentrate on triggering students' interest or (c) reciprocal relations exist (Köller et al., 2001; Marsh et al., 2005) and hence is it necessary to take into account both the promotion of interest and knowledge at the same time.

In the present study, we aimed to examine the underlying mechanism of the relationship between students' academic interest and their conceptual knowledge. We conducted a three-year longitudinal investigation of students' development in chemistry that follows up students' from Grade 9 to Grade 11. To gain deeper insights into the nature of the interrelation between interest and knowledge and to identify types of activities that are suitable to simultaneously support learning and interest development, we applied a more differentiated measure focusing on students' interest in seven dimensions of school science activities while students' conceptual knowledge was surveyed with regard to three concepts central to chemistry: the structure and composition of matter, chemical reaction, and energy. We estimated bivariate cross-lagged models (CLM) to test the hypothesised reciprocal relations between interest and understanding.

## 6.2 Theory

### Academic interest and interest in school science activities

Interest is typically defined as being content specific, i.e. specific to an object, domain or activity, and having a multicomponent structure, i.e. a cognitive and an affective component (e.g. Krapp & Prenzel, 2011; Renninger & Hidi, 2011). Moreover, interest is conceptualised as comprising both a state-like component, denoted as situational interest, and a trait-like component, referred to as individual interest. Situational interest is described as a psychological state that is predominantly evoked by situation specific environmental stimuli (Knogler et al., 2015), for example hands-on experiments that are novel or striking (Lin et al., 2013; Palmer, 2009), and that has to be supported externally by the environment. At the other end of the spectrum, individual interest is defined as 'a relatively enduring predisposition to reengage particular contents over time' (Hidi & Renninger, 2006, p. 111). Hence, according to Knogler et al. (2015) and Hidi and Renninger (2006), situational interest varies greatly between different situations (low cross-situational consistency) whereas individual interest is more consistent across different situations (high-level of cross-situational consistency). Moreover, the affective component is prevalent for situational interest (M. Ainley, 2017) and the cognitive component, containing a thirst of knowledge (Rotgans & Schmidt, 2014) and the autonomous pursuit of self-generated curiosity questions, is more pronounced for individual interest (Hidi & Renninger, 2006). In order to develop individual interest in specific school science activities opportunities to engage and reengage in such activities have to be provided and to be supported by the teacher (M. Ainley & Ainley, 2015).

With regard to students' subject-specific individual interest, numerous studies report a decrease throughout the entire secondary school, which is especially prevalent in math and science (Frenzel et al., 2010; Frenzel et al., 2012; Krapp & Prenzel, 2011). The decline of individual interest is often ascribed to developmental issues, differentiation processes of interest in adolescence, and the quality and type of instruction, for instance, the shifting focus of science education from phenomenological driven teaching styles to rather theory-loaded discussions of the underlying scientific concepts and principles (Krapp & Prenzel, 2011). Accordingly, interest in elementary school and lower secondary school is rather shaped by affective components whereas in upper secondary school the cognitive component of interest becomes increasingly important (Anderhag et al., 2016; Frenzel et al., 2012). In order to scrutinise students' interest at a finer grain size, Dierks et al. (2014b) devised a more proximal instrument that allows to assess students' interest in different types of activities in science classrooms. The newly devised instrument is adapted from Holland's RIASEC model of vocational interest (Holland, 1997), but intends to overcome the rather stereotypical categorisation of science as an exclusively investigative endeavour in the original model. The instrument retains the multifaceted structure of the RIASEC model, but shifts the focus from vocational interest to students'



interest in different types of school science activities, thus providing a more multifaceted picture of (school) science. For this purpose, the RIASEC+N model differentiates between interest in realistic (performing given lab experiments), investigative (solving theoretical problems), artistic (emphasising linguistic and visual aspects), social (explaining something to classmates), enterprising (managing group works), and conventional (organise the chemicals storage) school science activities (Dierks et al., 2016). In comparison to the original model, interest in networking school science activities (debating with classmates) was added as a new dimension based on empirical findings (Dierks et al., 2014b). All seven dimension are considered to be meaningful core elements of science instruction. Additionally, the focus on school science activities (in contrast to a focus on students' interest in chemistry in general) intends to provide more specific information for teachers about feasible ways to devise, manipulate, and implement such activities in class. So far, evidence for the hypothesised seven dimensional structure (Blankenburg, Höfler, & Parchmann, 2016) and a general decrease of students' interest across all seven dimensions throughout grade 5–11, albeit with a different pace (Höft et al., 2019) have been reported.

### Academic achievement

In order to participate in the world of the twenty-first century, students need to develop a deeper and integrated understanding of fundamental concepts of the discipline (Krajcik et al., 2012). This recommendation takes up the longstanding criticism that science teaching often only provides students with rather superficial knowledge of a broad range of discrete facts and distinct topics whose interconnections might not be obvious for the students (Bransford et al., 2000; Duschl et al., 2011). Novel educational standards have been structured around fundamental concepts emphasising their importance for teaching and learning. However, research on students' understanding shows that students struggle to acquire a sophisticated understanding even of the most fundamental science concepts, for instance the structure and composition of matter (Hadenfeldt et al., 2016; Löfgren & Helldén, 2009), chemical reactions (Taskin & Bernholt, 2014; Yan & Talanquer, 2015), and energy (Herrmann-Abell & DeBoer, 2018; Neumann et al., 2013; Yao et al., 2017). In the present study, we scrutinise the conceptual understanding of these three fundamental concepts of chemistry education. For this purpose, we conceptualised students' conceptual understanding as the successful application of scientific knowledge in a variety of contexts related to each of the three concepts (cf. Bransford et al., 2000; Weinert, 2001).

### Academic interest and academic achievement

Three hypotheses for the relation between interest and achievement are plausible: First, prior academic interest positively influences future achievement; Second, prior achievement fosters future interest (interest as an „affective by-product of learning“, Rotgans and Schmidt, 2017c, p. 350); Third, prior interest positively influences future achievement and vice versa, i.e. a reciprocal relation between both constructs, accentuating the intertwinement of interest development and knowledge acquisition.

In this study, we build on Eccles et al.'s Expectancy-Value Theoretical Model of Achievement Choices (EEVT) to analyse the mechanism of the interrelation between students' interest in school science activities and their conceptual understanding (Eccles, 2009; Eccles et al., 2015; Wigfield & Eccles, 2000). The EEVT seeks to explain the pursuit of academic activities and achievement choices whereby it emphasises the crucial role of students' expectancies and values of success for the decision-making process. Expectancies refer to students' 'beliefs about how one will do on different tasks or activities' and values contain 'incentives or reasons for doing the activity' (Eccles & Wigfield, 2002, p. 110). Value beliefs consists of four components: (a) an intrinsic, (b) an attainment, (c) a utility, (d) a cost component. The intrinsic and attainment component can be related to different parts of interest. The intrinsic component is closely associated with feeling-related valences like enjoyment whereas the trait-like, personal interest, as scrutinised in this study, is best covered by the attainment value

(Eccles et al., 2015; Gaspard, Dicke, Flunger, Schreier, et al., 2015; Marsh et al., 2005). The underlying mechanism as described by the EEVT which posits that students' interest can be interpreted „as both the outcome of activity choice processes and as one of the inputs into activity choice“ (Eccles et al., 2015, p. 326), potentially resulting in a synergistic process of reciprocal causations. In the following, we elucidate both directions of such a model.

### Interest impacts on achievement

According to the EEVT, students with higher interest, i.e. a higher subjective task value, are more likely to choose and succeed in an achievement activity. A higher interest positively affects learning strategies (Berger & Karabenick, 2011; Krapp, 2002b), goal setting (Hidi & Harackiewicz, 2000), effort (Trautwein et al., 2015), engagement and persistence (M. Ainley, Hidi, & Berndorff, 2002) as well as self-regulation (Lee et al., 2014). Hence, students with a higher interests are, on the long run, more likely to acquire a more sophisticated understanding. For instance, students having a higher interest in networking activities are more likely to actively participate in such activities, to share and debate their thoughts about a given problem with their peers and to enjoy themselves. Hence, these students' are more likely to acquire new knowledge in this specific situation and to reengage in future activities akin.

### Achievement impacts on interest

Several studies suggest that interest can be interpreted as a consequence of performance. A more sophisticated understanding is more likely to entail success in a given task, the perception of competence and autonomy, the appreciation of personal relevance and subjective task value, ultimately leading to a higher interest (M. Ainley & Ainley, 2011b; Harackiewicz et al., 2016). According to Eccles's 2009 identity approach to the EEVT, both performing an activity and perceiving oneself as competent at the activity are internalised into one's identity. For instance, students' with a more sophisticated understanding are more likely to devise chemistry experiments suited to solve a given problem (investigative activities) and are thus more likely to succeed. Consequently these students' are more likely to experience positive emotions, feel competent and hence have a higher willingness to reengage in activities akin.

Empirical findings support the assumption that interest promotes learning. With regard to upper secondary chemistry education, interest meaningfully predicts achievement between and within persons (Jansen et al., 2016; Kriegbaum et al., 2015). In order to analyse the causal relation between interest and academic achievement, multiple longitudinal studies have been conducted. Findings (primarily based on cross-lagged models) suggest that in elementary school (Jögi et al., 2015; Rotgans & Schmidt, 2017c; Viljaranta et al., 2014) and lower-secondary school (Ganley & Lubienski, 2016), prior achievement predicts later interest but no effects of interest on learning have been found. Only Marsh et al. (2005) reported reciprocal relations between students' interest in chemistry and their achievement. One plausible explanation could be that learning in lower-secondary school is predominantly driven by external motivation and school grades (Köller et al., 2001). In higher grades of secondary school, findings of cross-lagged models suggest positive relations from interest to achievement (Köller et al., 2001) and for older students even reciprocal relations (Guo, Marsh, Morin, et al., 2015; Harackiewicz et al., 2008). In summary, reciprocal causations comprising the synergistic process of interest development and knowledge acquisition can be assumed for students in upper secondary grades.

Pertaining to students' interest in school science activities (RIASEC+N), cross-sectional but yet no longitudinal findings in connection to achievement have been reported (Höft et al., 2019). These findings suggest that interest in investigative, social, enterprising, and networking activities are substantially connected to conceptual understanding, especially in upper secondary grades. Interest in realistic activities shows only small positive, interest in artistic and conventional activities exhibit no substantial correlations to students' conceptual understanding across all grades.

### 6.3 Research aims

In the present study, we aimed to scrutinise the longitudinal relation between students' individual interest in school science activities and their conceptual understanding in upper-secondary chemistry classes.

Research questions:

1. How does students' individual interest in school science activities and their conceptual understanding develop over Grades 9–11 of secondary school?
2. Is the relation between individual interest in school science activities and conceptual understanding directed or reciprocal?

With regard to Research Question 1., we assumed that students' individual interest declines in all seven dimensions (cf. Krapp & Prenzel, 2011; Potvin et al., 2018) while their conceptual understanding steadily increases (Hadenfeldt et al., 2016; Höft et al., 2019; Neumann et al., 2013).

With regard to Research Question 2., we expected moderate bidirectional relations between students' interest in investigative, social, enterprising, and networking school science activities and their conceptual understanding, based on findings from a preceding cross-sectional study with students in Grades 5–11 (Höft et al., 2019; Köller et al., 2001; Marsh et al., 2005; Pekrun et al., 2017). Furthermore, we expected no substantial and systematic effects between interest in artistic and conventional as well as realistic school science activities and conceptual understanding. In this prior study, interest in investigative, social, enterprising, and networking activities was substantially connected to students' conceptual understanding, especially in upper secondary grades (Höft et al., 2019). Based on these cross-sectional findings, we assumed to find a similar pattern with regard to the longitudinal interrelations between students' interests in the different school science activities and their conceptual understanding over time.

### 6.4 Method

#### Sample and procedure

Students from five secondary schools in Northern Germany participated in this study ( $N = 756$ , 52% female). All students were attending the Gymnasium, the higher of two tracks of secondary schools in this region (attended by about 66% of all students in this federal state). According to findings of the PISA 2015 study (Schiepe-Tiska et al., 2016), chemistry education at the Gymnasium in Grade 9 is generally ( $n = 66\%$ ) characterised by patterns of instruction that rarely provides opportunities to perform chemistry experiments or to solve highly cognitive activating tasks. Only approximately 15% of the students reported that they frequently conducted their own chemistry experiments (investigative activities), respectively carried out hands-on experiments (realistic activities), and meanwhile perceived chemistry education as being cognitive activating.

The five schools are a convenient sample of schools that were willing to participate in this longitudinal study. Within each school, all classes were included in the study to cover the full range of abilities and interests of students attending this type of school. As no random sampling of students and/or schools was possible, the generalizability of the results reported here is restricted and can only be done with caution. However, basic demographic information (in terms of gender, migration background, language background, and SES) indicated that our sample features a composition that was comparable to samples drawn for representative studies on state level (cf. Leucht et al., 2016).

Data were collected in three annual waves from 2015 to 2018, in February to March of each year. Written information about the study was sent to all students and their parents in advance and students participated on a voluntary basis after written consent was provided by the parents. Students were also informed that they could drop-out from the study at any time. Students were asked to fill out the tests and questionnaires during regular lessons based on a standardised protocol, carried out by trained personnel. Answering tests and questionnaires took about 120 minutes per wave. All

collected data were anonymized and no individual results were provided to the participants, their teachers, or the schools, so that neither participation nor non-participation would have immediate consequences for the students. The participation rate ranged from .46 to .85 between schools and waves, with a mean of .74. Across the three measurement waves, the amount of missing data ranged from 30% to 37%.

## Measures

Interest in school science activities. Students' individual interest in different school science activities was measured using the RIASEC+N instrument developed by Dierks et al. (2014b). The original phrasing of the items was devised together with teachers from the respective types of school to ensure that the items address common school science activities. Afterwards the phrasing was validated through cognitive interviews to ensure that students of the targeted age group both understand and interpret the different as intended (Dierks et al., 2014b). The RIASEC+N instrument categorises activities into seven dimensions: realistic, investigative, artistic, social, enterprising, conventional, and networking activities. The instrument conceptualises students' interest as a trait-like concept, i.e. students' enduring and reoccurring interest in school science activities. The high-level of cross-situational consistency of individual interest was taken into account by a decontextualised operationalisation (e.g. there are no references to specific topics) of the interest items. Additionally, due to the instruments' differentiation and decontextualised operationalisation, it allows precise comparisons of students' interests over longer time spans. In the RIASEC+N instrument, each of the seven dimensions is addressed by four items (cf. Table 6.1; all items are provided in Electronic Supplementary Materials Appendix C.1), resulting in 28 items total. One item per scale (except for the social scale) was deleted because of poor item-scale correlation. For each item, students were asked to rate their interest in each school science activity on a four-point Likert scale, ranging from „I am not interested at all in doing this“ (1) to „I am very interested in doing this“ (4).

**Table 6.1:** Dimensions and example items of the RIASEC+N Questionnaire to the question asking „How much are you interested in doing the following activities in chemistry lessons?“

Dimension	Example item
Realistic	To do experiments guided by an instruction
Investigative	To plan experiments to investigate something
Artistic	To draw an observation
Social	To explain chemical topics to fellow students
Enterprising	To organise a small chemistry project
Conventional	To table the results of an experiment
Networking	To investigate a chemical topic together with fellow students

Conceptual understanding. Students' conceptual understanding was measured by a written test covering three central chemical concepts: The structure and composition of matter, chemical reaction, and energy (in accordance to the German educational standards; KMK, 2004. For the test, classical multiple choice items were combined with so called ordered multiple-choice items (OMC; cf. Briggs et al., 2006). OMC items are assumed to be a valid replacement for open-ended item formats and are considered to be less time and money consuming (cf. Alonzo & Steedle, 2009; Briggs et al., 2006). These OMC items are designed similarly to regular multiple-choice items by providing several response options, but these options correspond to specific levels of understanding based on theoretical or empirically validated models of understanding a particular concept (cf. Appendix C.1 for exemplary items to each concept). With regard to the three chemical concepts that are in focus in this study, the assumed levels of students' understanding when learning about matter and chemical reactions are characterised by increasingly differentiated models of the structure and composition of matter as well as theories of chemical bonding. More specifically, students' understanding of these concepts is

assumed to start off with interpreting observable phenomena and recognising the formation of new substances; with further instruction, students are then confronted with Dalton's atomic model and its application to interpret chemical reactions as rearrangements of atoms; with the transition from lower to upper secondary school, students learn about Bohr's atomic model and its implication to conceptualise chemical reactions in terms of changes in the electron shell (cf. Emden et al., 2018; Hadenfeldt, Liu, et al., 2014; Hadenfeldt et al., 2016). With regard to the energy concept, students' understanding is assumed to progress through a sequence of four energy aspects: forms/sources, transfer and transformation, energy degradation and dissipation, and finally energy conservation (cf. Liu & Lesniak, 2005; Neumann et al., 2013; S. T. Opitz et al., 2017). In each OMC item, one answer option reflects the highest level of understanding that can be tested with this item. The other answer options reflect (scientifically) incorrect or incomplete ideas and thus correspond to lower levels of understanding. The design process of these answer options is typically based on the structure of the curriculum as well as research on students' alternative conceptions of a particular concept (Briggs et al., 2006; Hadenfeldt et al., 2013; Taber, 2002; Talanquer, 2009).

Based on a common item anchor design, students received six anchor items and four grade-specific items (addressing content aspects that are usually taught in a particular grade) for each of the three concepts, resulting in 30 items for each measurement wave. All items were reviewed by an expert panel of four chemistry education researchers and three chemistry teachers to assure the content validity of the items, the correctness of the response options, to eliminate ambiguities or alternative plausible interpretations that could affect between-options relationships, to confirm the rating of the student's level of understanding, to check the assignment of items to specific grades, and to select items to anchor the test forms across grade levels, i.e. items which answer options prototypically reflect the progress of students' understanding across grades for a particular concept.

Besides interest in school science activities and conceptual understanding, the final questionnaire included questions regarding demographic variables (gender, age, first language, SES) and several other constructs (self-concept, motivation, epistemic beliefs) that are not in focus of this article.

### Data analysis

Missing data and method effect. All longitudinal modelling were conducted with Mplus 8.0 (L. K. Muthén & Muthén, 1998-2015) using the MLR algorithm (robust to non-normality issues) and applying the full information maximum likelihood method (FIML) to make full use of the data from students with missing data (Enders, 2010). Correlations between residuals for identical interest items over time were included in longitudinal models (correlated uniqueness approach) to account for the shared method variance. It is widely recognised that students' achievement and interests do not develop independently from the context they are learning in. For instance, students in different classes might potentially be confronted with different kinds of instructional practices which in turn impact students' learning. To systematically analyse the effect of this kind of nested data structure, a multilevel approach would be necessary. However, this was not possible in the present study due to the limited number of classes participating in this study ( $N_{\text{classes}} = 25$ ) which is below common recommendations for multilevel approaches.

Interest in school science activities. Cronbach's  $\alpha$  was calculated for each of the seven dimensions as a measure of internal consistency, indicating a sufficient reliability across all dimensions and measurement waves ( $\alpha = [.65, .90]$ , Appendix C.2). In order to compare the latent values of students' interest in the different school science activities across grades (i.e. across measurement waves), scalar measurement invariance had to be ensured (Wu et al., 2007). For this purpose, confirmatory factor analyses were sequentially evaluated for models of configural, metric, scalar, and strict invariance, whereby each step adds more restrictions to the baseline model (Wang & Wang, 2012). Again, goodness-of-fit and absolute-fit indices were applied, demonstrating scalar measurement invariance of all interest constructs across all three measurement waves (Appendix C.2).

Conceptual understanding. Students' answers to the conceptual test were analysed by using Item Response Theory (IRT) in order to account for the matrix design of the test (i.e. not all items having been administered to all students). Answers to multiple choice items were coded as incorrect (0) or correct (1), while answers to OMC items were assigned partial credits ranging from 0 to 3, in concordance to the theoretical model. This coded data was analysed using a one-dimensional multi-group generalised partial credit model for each wave (MG-PCM, as implemented in the R package TAM; Test Analysis Modules; Kiefer et al., 2017) with student grade as grouping variable. The expected a posteriori based on plausible values reliability (EAP/PV) were calculated and found to range from 0.81 to 0.86, indicating a high reliability of the measured construct. With respect to item fit, Weighted Mean Square (WMNSQ) estimates and standardised mean-square fit statistics (i.e. T values) were considered. Items that did not meet typical cut-off values ( $WMNSQ < 1.3$ ,  $T < 2.0$ ; Wright et al., 1994) were excluded from further analysis ( $n_{T1} = 2$ ,  $n_{T2} = 3$ ,  $n_{T3} = 2$ ). Finally, weighted likelihood estimates (WLE) were calculated as a measure of students' conceptual understanding (WLE reliability = [.74, .77]), as measured by the test. IRT parameters estimates of the three measurement waves were linked based on a generalisation of log-mean-mean linking in accordance with Haberman (2009).

Longitudinal analysis. With regard to analysing trends over time, two separate lines of analysis were conducted. First, we were interested in the development of students' interest and understanding over time. For this purpose, first order latent growth curve models (LGCM) were employed to test whether students' interest in school science activities and students' conceptual understanding decreased respectively increased from grade 9 to 11. We specified two models per scale: (a) a linear growth model and (b) a non-linear model with a freely estimated loading at T2 (e.g. Retelsdorf et al., 2011). To test if the second model fitted better than a linear growth model, likelihood ratio (LR) tests were estimated. Second, we aimed to examine the longitudinal associations between students' interest in specific school science activities and their conceptual understanding, i.e. how both interest and understanding impact each other from grade to grade. Based on the literature, the association between both constructs could be driven by interest, by understanding, or bidirectionally. Here, bivariate latent cross-lagged models (see Figure 6.1) were applied to analyse the interplay between students' interest in a specific school science activity and conceptual understanding over time, controlling for gender differences (with coding female as 1, male as 2), age and general intelligence (as measured by the KFT; Heller and Perleth, 2000). In each model, the influence of interest in a specific school science activity on conceptual understanding and from conceptual understanding on interest (cross-lagged path) is controlled for 1-year stability paths (e.g. interest in investigative activities in grade 9 predicting interest in investigative activities in grade 10) and within-time correlations among all variables. Controlling for prior variance in outcomes is important because otherwise this autoregressive influence can lead to spurious associations between variables. We estimated four versions for all of the seven CLMs. In the first version cross-lagged paths were freely estimated. In the second version, the effects of interest on subsequent achievement were constrained to be invariant. In the second version, the effects of achievement on subsequent interest were constrained to be invariant and in the final model, both constraints were added simultaneously. Afterwards, LR tests were performed between the first version and the constrained versions, to test if the more parsimonious versions described the data equally well. Model fit was evaluated using the comparative fit index ( $CFI > .95$ ) and the root-mean-square error of approximation ( $RMSEA < .08$  acceptable;  $< .05$  good) (Byrne, 2012). In both lines of analysis, we used separate models for each of the seven school science activities to minimise complexity (cf. L. M. Daniels et al., 2009).

## 6.5 Results

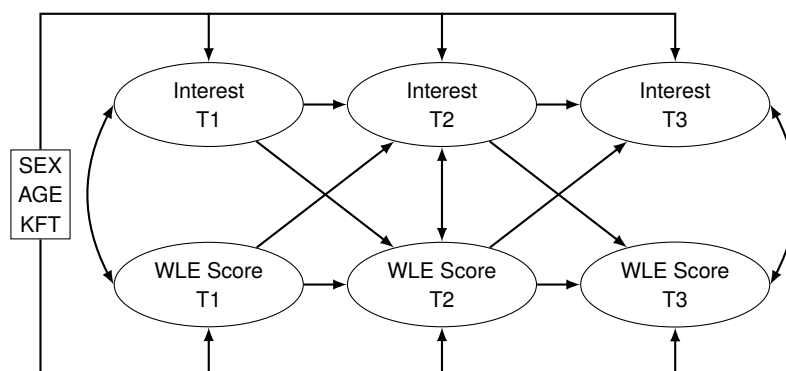
### Descriptive statistics

All dimensions of students' interests in school science activities were positively correlated ( $r = [.25, .70]$ ) at all three time points (see Appendix C.2). The highest correlation with conceptual understanding was found for students' interest in investigative activities, while students' interests in artistic activities were not correlated to conceptual understanding.

### Latent growth models of interest in school science activities and conceptual understanding

We computed first order LGCMs to analyse the progression of students' interests in school science activities and students' conceptual understanding from Grade 9 to 11 (for comparison see Appendix C.3). For six out of seven scales, the linear growth model fitted the data acceptably ( $CFI > .973$ ,  $TLI > .9181$ ,  $RMSEA < .083$ ,  $SRMR < .032$ ) and was superior to a model with freely estimated shape (Appendix C.3). For the artistic scale, a model with freely estimated shape was superior to a linear growth model. With regard to students' interests in school science activities, the results of the LGCM indicate a small decrease ( $d = [.15, .34]$ ) in all but one case: students' interest in realistic activities, i.e. activities students liked most, decreased moderately ( $d = .53$ ). Intercepts and slopes of six dimensions of students' interest were uncorrelated, indicating that the extent to which students' interest in science activities decreased was independent from students' initial interest at Grade 9. This does not apply in case of students' interest in social activities ( $r = .32$ ,  $p < .05$ ). That means, students with a higher initial interest level lost their interest at a faster rate. With regard to students' conceptual understanding, a model with freely estimated shape fitted the data best indicating a moderate increase ( $d = .62$ ). The gain in conceptual understanding over time is again uncorrelated to their initial understanding at Grade 9, indicating that the gap between high- and low-performing students is not getting wider during the three year period covered in this study (but low-performing students are also not catching up).

Overall, the variance of the initial levels but not of the growth factors were significant for all constructs. Thus, the results indicated momentous inter-individual variation at each measurement point, but still comparable trends regarding students' individual development over the time.



**Figure 6.1:** Bivariate Cross-lagged Model used to analyse the relation between students' interest in school science activities and their test scores while controlling for students' sex, age and cognitive abilities (KFT).

### Cross-lagged models of interest in school science activities and conceptual understanding

The fit indices ( $CFI > .943$ ,  $TLI > .914$ ,  $RMSEA < .037$ ,  $SRMR < .058$ ) provided support for the CLMs with free cross-lagged paths (version 1) for all seven dimensions of students' interest in school science activities. When constraining cross-lagged paths to be invariant time (version 3), LR tests showed a nonsignificant change for all of the models and additionally the loss of fit was acceptable ( $\Delta CFI \leq .001$ ,  $\Delta RMSEA \leq .001$ , and  $\Delta SRMR \leq .002$ ). Therefore, the results shown hereafter bear on

the most parsimonious version of the CLMs (version 3) with invariant cross-lagged effects across time intervals. Factor loadings, path coefficients, and residual variances for the bivariate CLMs (version 3) are listed in Appendix C.4. Both the interest variables and students' conceptual understanding were considerably stable over time, as indicated by the moderate to high autoregressive effects for these variables ( $\beta = [.08, .70]$ ,  $p < .001$ ). Moreover, there were significant, positive correlations between interest and conceptual understanding ( $r = [.13, .30]$ ,  $p < .05$ ) at Grade 9, except for the artistic and realistic activities.

With regard to cross-lagged effects between interest and conceptual understanding, 14 out of 28 cross-paths were significant ( $\beta = [.09, .18]$ ,  $p < 0.05$ ): four dimensions (investigative, social, enterprising, networking) of interests in school science activities predicted subsequent achievement, likewise achievement predicted subsequent interest in three dimensions (investigative, social, networking). Hence, the results revealed three distinct patterns of longitudinal couplings between students' interest in school science activities and students' achievement: 1) no cross-lagged effects existed between students' interest in realistic, artistic and conventional activities and achievement; 2) students' prior interest in enterprising activities predicted subsequent achievement (achievement was driven by interest); 3) bidirectional relations emerged between students' interest in investigative, social and networking activities and achievement.

In order to gain more precise insights into the bidirectional relations between interest in investigative, social as well as networking activities and achievement, the extent of the cross-lagged effects were compared for the four bivariate CLMs. Therefore, we constrained the regression coefficients of all four cross-lagged paths to be equal and performed LR tests. The results indicated equal regression coefficients for the investigative ( $\chi^2(1) = .15$ ,  $p = .70$ ) and the social scale ( $\chi^2(1) = .39$ ,  $p = .53$ ), but unequal regression coefficients for the networking scale ( $\chi^2(1) = 4.49$ ,  $p < .05$ ), i.e. prior achievement predicted subsequent interest to a higher degree than prior interest predicted subsequent achievement.

## 6.6 Discussion

The aims of the study were to investigate (a) how interest in school science activities and conceptual understanding develop for chemistry students' from Grade 9 to Grade 11 and (b) whether systematic cross-lagged relations between interest and conceptual understanding exist. Therefore, we carried out a three-year longitudinal study (Grade 9–11) and performed LGCM and CLM analysis to examine the developmental progression of and the interplay between both constructs.

With regard to the first research question, we found the expected general but small decline of students' interest in school science activities and concomitant a moderate increase of students' conceptual understanding from Grade 9 to 11. Students' interest decreased in all seven dimensions. Reasons for the decline of students' interest are often discussed with regard to developmental effects (e.g. students alter their interests according to social contexts), effects of internal differentiation, an increasing degree of cognitive demanding tasks and the need for successive learning. The underlying process of students' internal differentiation of interests comprises effects of social, temporal and dimensional comparisons and is related to the development of their academic self-concept (Gaspard et al., 2017; Goetz et al., 2008; Sewasew et al., 2018; Wolff et al., 2018). In addition, a low fit between students' interest and the content taught in school as well as disadvantageous ways of embedding the activities into teaching could negatively influence students' interests. However, the decline in students' interest in all seven dimensions was only associated with small effect sizes. With regard to results based on longitudinal data, this is in line with other findings indicating that a considerable loss of interest takes place between grades 5 and 7, which then plateaued in grades 8 and 9 (Frenzel et al., 2010). In general, the decline of students' interests is regrettable because interest is, for instance, assumed to be crucial for successful and sustained learning over longer periods of time, career aspirations, and engaging in scientific activities out of school.

An opposing trend was found for students' conceptual understanding. Here, our results showed the expected moderate increase of students' conceptual knowledge. In terms of learning gains, the



effect size for this increase in understanding was  $d = 0.62$  (Appendix C.3), i.e. a medium effect. Compared with the results of a meta-analysis based on average annual gains on nationally normed science tests (Bloom et al., 2008), the effect size reported here is substantially higher. This difference can of course partially be traced back to specific discrepancies between the meta-analysis and this study. First, our study analysed students' learning in a single, disciplinary subject (i.e. chemistry) instead of integrated science instruction. Second, the test design focused on core chemical concepts instead of a broader test on concepts relevant in different science disciplines (cf. Bloom et al., 2008). Third, the development of items aimed at reaching a high alignment between curriculum and test (Liu, 2012). However, the results obtained in this study also indicate that the test is sensitive to students' learning in chemistry across the Grades 9–11, supporting the assumption that the developed instrument provides a thorough measure of students' conceptual understanding across this grade span.

The finding that students' gains in their conceptual understanding is uncorrelated to their initial understanding at Grade 9 is somewhat surprising. Often, a cumulative advantage effect (also called Matthew effect) is expected, however, there is rare empirical evidence that initial differences actually lead to growing inequalities. In addition, this finding is in line with findings in mathematics or reading: Baumert et al. (2012) also did not find an effect of initial ability on learning gains in either domain. As the level of students' understanding is highly dependent on prior performance, interindividual differences remain highly stable over time, indicating that the gap between high- and low-performing students is not getting wider during the three year period covered in this study. Baumert et al. (2012) argue that one possible explanation is that there is „a time lag in the development of weak and strong [learners] – both trajectories follow a similar logistic function but (due to the time lag) show different gradients over the period of investigation“ (p. 1364). Testing this assumption would probably require to monitor students' development more closely and over longer time spans. Overall, both the decrease of students' interests and the increase of their conceptual knowledge at the same time are also in line with recent cross-sectional findings (Höft et al., 2019).

With regard to the second research question, we expected substantive reciprocal relations between students' interest in investigative, social, enterprising, and networking school science activities and students' conceptual understanding. These assumptions were derived from a preceding cross-sectional study with students in Grades 5–11 (Höft et al., 2019). With regard to the seven dimensions of the RIASEC+N instrument, different interrelations between conceptual understanding and interest were observed: While the correlations between students' interests in investigative, social, enterprising, and networking activities and their conceptual understanding were small to middle-sized and increased from Grade 7 to 11, the correlations between interest in realistic activities and conceptual understanding was small and almost stable over time. In addition, correlations between conceptual understanding and interest in the artistic and conventional dimensions were negligible. Regarding the results presented here, we observed reciprocal cross-lagged relations between students' interest in investigative, social and networking school science activities and a positive effect of students' interest in enterprising activities on their conceptual understanding. To our perception, items of the investigative and the networking dimension best illustrate activities associated with cognitive activation (e.g. involving cognitive conflict, or problem-solving) and knowledge acquisition (Kunter et al., 2013), whereas enterprising and social activities pronounce group work characterised by expert-novice relations between the students. Hence, students' interests in scientific activities that are cognitively demanding and include the communication of knowledge could enhance conceptual understanding and simultaneously conceptual understanding contributes to interest development in this kind of activities, except for enterprising activities. Potentially, more complex analyses integrating students' academic self-concept could provide more insights into the relation between students' interest in enterprising activities and their knowledge acquisition. Consequently, investigative, social and networking activities could represent convenient starting points for fostering both students' interest and their conceptual understanding.

In addition, the fact that only four out of seven dimensions of students' interests were meaningfully linked to their knowledge acquisition indicates that interest does not enhance learning per se. Making

use of the more differentiated RIASEC+N model in this study resulted in a more differentiated pattern of interrelations. While some activities in science classrooms not only seem to be more appealing than others to students, some of these activities (given that opportunities to engage in such activities are available), additionally seem to result in a more intense processing of the content (or these activities require more content knowledge to perform these activities successfully). This is not to say that, for example, conventional (i.e. organise the chemicals storage) or artistic (i.e. emphasising linguistical and aesthetical aspects) activities are not cognitively demanding. However, the way how these activities are typically designed and implemented in the classroom might not require students to retrieve and deepen their understanding. Investigative (i.e. solving theoretical problems or designing own experiments), social (i.e. explaining something to peers) and networking (i.e. debating with peers) activities might enclose a relation to content in principle, resulting in a closer and reciprocal relation of these activities and students' conceptual understanding over time. Here, it is important to note that the findings provided here only pertain to the interrelation between students' conceptual understanding and their self-reported interest in performing the different school science activities. Students were not actually performing these activities as part of this study and no information about the frequency and quality of students' opportunities to perform these activities in school were included in the analysis (cf. the next section on limitations). In addition, the backlash of students' conceptual understanding on their interest in these activities also highlights the importance of extrinsic motivation for learning (cf. Köller et al., 2001), as not only doing an activity but also seeing oneself as competent at the activity is an important aspect of the interrelation between interest and understanding (M. Ainley & Ainley, 2011b; Eccles, 2009; Harackiewicz et al., 2016).

Regarding implications for teaching practice, investigative, social and networking activities could serve as productive starting points to simultaneously foster students' interests and conceptual understanding. To take advantage from the reciprocal relation between students' interest in these specific activities and their conceptual understanding using associated teaching methods such as inquiry learning (Jocz et al., 2014; Lau & Lam, 2017), project-based learning (cf. Harris et al., 2015), and collaborations (cf. Ding & Harskamp, 2011) might be suitable approaches. Additionally, formative assessment could provide students with opportunities to experience success and may help to foster interest and conceptual understanding, ideally resulting in positive feedback loops (Kingston & Nash, 2011; Pinger et al., 2016).

## 6.7 Limitations and future research

Despite the strong support that was found for the existence of reciprocal relations between students' interest in investigative, social and networking activities and their conceptual knowledge, this study has some limitations. As the scope of the study was to identify overall developmental relations between interest and conceptual understanding no information about the frequency and quality of school science activities were collected. Thus, our analysis failed to control for effects of the quality and quantity of school science activities as well as other classroom variables. To simultaneously incorporate both individual- (e.g. prior achievement or interest) and class-level (e.g. quality and quantity of school science activities) variables in an analysis, a multilevel approach would be the most appropriate. As students in the present study were also nested in different classes and thus might potentially be confronted with different kinds of instructional practices, a multilevel approach would also provide more insights about the impact of the nested data structure on the results presented here. However, a multilevel analysis was not possible due to the limited number of classes participating in this study and varying class compositions (especially at transition from Grade 9 to 10), which implies the need of a time variant cluster variable. Although the values for both the ICC and the design effect support the approach taken in this study, i.e. considering the development and interplay of achievement and interest as a predominantly individual and not a class level phenomenon, all causal interpretations of estimated effects need to be treated with caution.

In order to gain more detailed insights into the causal mechanism and derive trustworthy guidelines for practice, short-term longitudinal data acquired from observational studies with shorter

measurement intervals, interventions and experimental studies are needed. In addition, even though a decontextualised version of the RIASEC+N model was used, it is an open question, to which degree the instrument measures, besides students' individual interests, parts of students' situational interest (Knogler, 2017; Knogler et al., 2015). Once more, short-term longitudinal data could serve as a basis to disentangle state-like and trait-like components of students' interest and provide further evidence for the generalisability of the RIASEC+N model (cf. Habig, Blankenburg, et al., 2018; Höft et al., 2019). Our data suggest a close connection between interest development in investigative, social as well as networking activities and conceptual understanding. Given the limitations of our study, no conclusions with regard to how to best implement this kind of activities in classrooms can be drawn. Thus, further studies focusing on how to trigger and foster students' interest in investigative and networking activities while taking into account classroom variables are necessary to identify mediators between interest and conceptual knowledge. Intervention studies and short-term studies could provide more evidence with respect to the question of whether the implementation of specific activities can actually fosters students' interest and understanding simultaneously.

The study focused on bivariate relations between interest and conceptual understanding, disregarding potential interactions between different dimensions of interests. Synergistic as well as competitive effects between different types of school science activities are conceivable as cross-sectional findings showed high correlations between different types of school science activities. For instance, students' initial interest 'to do experiments guided by an instruction' (realistic) could lead to an interest „to plan experiments to investigate something“ (investigative) and ultimately indirectly promote their conceptual knowledge. Therefore, another interesting question for future research is how interest in different dimensions interact and how these interests relate to students' general interest in chemistry as subject. Furthermore, it would be intriguing to examine whether specific initial profiles of interests exist that are especially beneficial for learning. Similarly, an analysis of possible mediators that might influence the relation between students' interest and conceptual understanding, such as students' strategy use and perception of competence, would provide more insights into the interplay between interest and achievement. Finally, schools and students participating in this study represent a convenience sample. Hence, the generalizability of the presented results is restricted. To minimise this effect, all classes of each school were included to cover the full range of students attending these schools. In addition, we scrutinised whether basic demographic information (e.g. gender, migration background, language background, and SES) of our sample is comparable to samples drawn for representative studies on state level (cf. Leucht et al., 2016). Nevertheless, the variance in both students' performance and their interest might not cover the full range that could be observed in the student population.



# Probing Precursors of Students' Long-term Engagement in Chemistry: Individual Interest and Achievement Development and Their Relation with Utility Values.

7

## Abstract

Preparing students to pursue STEM careers is deemed critical in the 21st century. Therefore, the development of and the interplay between secondary school students' individual interest and achievement as well as associations to utility values were investigated using the example of chemistry education. In a two-cohort study with yearly measurement occasions (Cohort 1: Grade 6–8,  $N = 493$ ,  $M_{\text{age}} = 11.6$  years in Grade 6; Cohort 2: Grade 9–12;  $N = 382$ ,  $M_{\text{age}} = 14.7$  years in Grade 9), we assessed students' interest in three dimensions of school science activities emphasizing cognitive and social aspects of learning. Based on latent change score modeling, results indicate increasing couplings between students' interests and achievement from lower to upper secondary. Reciprocal relations were only partly found in upper secondary school. These findings extend research findings pertaining to the longitudinal interplay between interest and achievement and stress age-related differences.

Keywords: individual interest; achievement; utility values; chemistry education

Höft, L., & Bernholt, S. (in Rev.). Probing Precursors of Students' Long-term Engagement in Chemistry: Individual Interest and Achievement Development and Their Relation with Utility Values.

## 7.1 Introduction

The goal of science education in secondary school is twofold: it prepares students for university studies and primes them to take on constructive roles as citizens in a modern society (cf. Millar et al., 1998; Osborne & Dillon, 2008; Sadler, 2009). In achieving these objectives, motivation plays a crucial role, as variables such as interest and utility values, have been found to be strong determinants of future academic (Köller et al., 2001) and vocational choices (Eccles, 2009; Lauermann et al., 2017). Empirical findings suggest that students' interests, utility values and achievement are interrelated, on account of which research has drawn particular attention to the development of these student characteristics, and the relations between them (Krapp & Prenzel, 2011; Potvin & Hasni, 2014b; Wigfield & Cambria, 2010). With regard to the longitudinal interplay between individual interests and achievement, four different patterns are plausible: (a) no relations between individual interests and achievement exist, (b) prior achievement promotes subsequent interests (Ganley & Lubienski, 2016; Rotgans & Schmidt, 2017c), (c) previous interests facilitates subsequent achievement (Jansen et al., 2016; Kriegbaum et al., 2015), or (d) reciprocal relations between interests and achievement exist (Köller et al., 2001; Marsh et al., 2005).

The development of interest and achievement is also closely associated with students' utility values, as manipulations of utility values have been repeatedly found an effective means to simultaneously foster interest and achievement (Lazowski & Hulleman, 2016). These findings indicate the opportunity to counteract the often reported decline of interest, especially in STEM subjects, and the associated early drop-out from optional school subjects such as chemistry in upper secondary education, and could thereby lead to the increase of students pursuing a science career (e.g., Frenzel et al., 2012; Potvin & Hasni, 2014a).

In the present study, we sought to contribute to the growing knowledge base concerning the longitudinal coupling of students' interests, achievement, and utility values by contrasting findings of two cohorts of a three-year longitudinal study. We annually assessed students' individual interest and achievement from lower (Grade 5 to 8) and upper secondary chemistry students (Grade 9 to 12) and their utility values at the last measurement wave. Thereby, we focused on three dimensions of students' interests in specific school science activities. These activities differ in their emphasis of cognitive and social aspects of learning: 1) investigative activities concern individual problem-solving processes, 2) networking activities concern the collaboration on tasks in small groups, and 3) social

activities concern helping peers in their understanding or problem-solving. All three activities are implemented in the German educational standards for science and typical for chemical education (Björkman & Tiemann, 2013; KMK, 2004; Stiller, 2016).

## 7.2 Background

In this paper, we utilize Eccles et al.'s expectancy-value theory of achievement motivation (EVT) (cf. Eccles, 2009; Wigfield & Eccles, 2000) as a framework for our longitudinal analyses, as it represents a comprehensive approach to explain students' achievement choices (such as task and course choices), effort, engagement, and future performance (Eccles, 2009; Wigfield & Eccles, 2000). The EVT states that students' expectancies of success and value beliefs, consisting of (1) an intrinsic, (2) an attainment, (3) a utility, and (4) a cost component, are important predictors in decision-making processes. Therefore, researchers working with the EVT commonly posit that students with higher expectancies of success („beliefs about how one will do on different tasks or activities“) and higher values of success („incentives or reasons for doing the activity“; Eccles and Wigfield, 2002, p. 110) are more likely to engage in achievement activities and perform better, consequently leading to the achievement of higher learning gains.

Regarding value beliefs, individual interest is closely related to the intrinsic and the attainment component of value beliefs (Eccles et al., 2015; Schiefele, 1999; Tsai et al., 2008). In this vein, the intrinsic component covers feeling-related valences of interest whereas the attainment component covers value-related valences of interest. Instead of focusing on one single decision-making process, the EVT takes the influence of attribution processes in response to the outcome of prior achievement choices on future choices into consideration. Consequently, utility values and interests are construed as „both the outcome of activity choice processes and one of the inputs into activity choice“ (Eccles et al., 2015, p. 326). In the following sections, we detail the three constructs examined in the present paper and elucidate relations between them.

### Interest

Interest is commonly defined as a dynamic relation between a person and an object, such as a topic or an activity (Krapp, 1999a). Recent interest theories conceptualize interest as comprising both a psychological state and a predisposition (cf. Renninger & Hidi, 2011). For the purposes of the present study, we focus on the trait-like interest, denoted as individual interest. Such an individual interest is defined to be inherent in the person, i.e., as an individual's predisposition to re-engage in certain activities or contents over time (Hidi & Renninger, 2006). Empirical findings indicate that students' individual interest in math and science declines across secondary school (cf. Frenzel et al., 2012; Krapp & Prenzel, 2011). Therefore, Dierks et al. (2014b) devised the RIASEC+N model to gain more differentiated insights into students' individual interests and to find more specific levers to foster interest development. The RIASEC+N model focuses on students' interests in common school science activities (cf. Holstermann et al., 2010; Swarat et al., 2012) and distinguishes between seven dimensions of students' individual interest in realistic (e.g., conducting hands-on experiments), investigative (e.g., designing investigations and problem-solving), artistic (e.g., drawing and designing), social (e.g., supporting peers by explaining problems), enterprising (e.g., organizing group work), conventional (e.g., sorting and categorizing) and networking (e.g., working in teams) school science activities.

### Achievement

In science education, students' achievement is commonly discussed against the background of scientific literacy which received wide acceptance as the overall aim of science education (DeBoer, 2000; Gräber & Bolte, 1997; Roberts, 2007). Albeit in different notions, scientific literacy is often considered a functional educational concept that aims to provide the basis for lifelong learning and for

life in a modern society (e.g. Bybee, 1997; Millar et al., 1998; Roberts, 2007). In addition, it advocates focusing on a deeper and more integrated understanding of the fundamental concepts of a discipline (Bransford et al., 2000; Duschl et al., 2011; Krajcik et al., 2012).

Focusing instruction on students' understanding of central concepts, instead of covering a broad range of distinct topics, provides the basis for making this knowledge accessible in different problem contexts and thus for acknowledging the relevance of a particular concept across a variety of situations. Purposefully selecting these problem contexts and situations could then provide the opportunity for students to also see the relevance of science concepts outside the science classroom (Sevian et al., 2018). However, empirical research repeatedly indicates that students often struggle in acquiring a sophisticated understanding of fundamental science concepts (Claesgens et al., 2009; Romine et al., 2016; Stamovlasis et al., 2005) and that they hardly relate the content learned in class to their everyday lives (Sjøberg & Schreiner, 2012).

### Interest and Achievement

Commonly, interest theories suggest a positive relation between interest and achievement (Alexander, 2003; Hidi & Renninger, 2006; Rotgans & Schmidt, 2017d). In line with this theoretical standpoint, empirical evidence suggests a positive link between interest and achievement, as compiled, for instance, in the meta-analysis by Schiefele et al. (1992). Here, the connection between the two variables appears closer in higher grades of secondary school (cf. Denissen et al., 2007). The positive inter- and intra-individual effect of interest on achievement was shown by Jansen et al. (2016) as well as Kriegbaum et al. (2015) for different domains (L1, math, biology, chemistry, and physics) based on representative samples of German students. Studies analyzing the interplay of interest and achievement by means of cross-lagged designs, i.e., by controlling for autoregressive effects, draw a more complex picture. With regard to primary and lower secondary school students, empirical findings demonstrate positive effects of prior achievement on subsequent individual interest (Ganley & Lubienski, 2016; Grigg et al., 2018; Jögi et al., 2015; Köller et al., 2001; Marsh et al., 2005; Rotgans & Schmidt, 2017c; Viljaranta et al., 2014) and underscore the important role of the academic self-concept as a mediator in this relationship (Marsh et al., 2005; Sewasew et al., 2018; Viljaranta et al., 2014). However, positive cross-lagged effects of prior individual interest on subsequent achievement when controlling for autoregressive effects were only partly found (e.g. Marsh et al., 2005; Schaffner et al., 2016) and were attributed to the positive influence of interest on course choices in upper secondary and college students (Harackiewicz et al., 2008; Köller et al., 2001). With regard to interest in school science activities, a cross-sectional study indicated that students' interests in investigative, networking, and social activities is connected to their achievement, notably in secondary school (Höft et al., 2019).

### Utility Values

Utility values cover extrinsic reasons for engaging in a task conducive to an individual's future plans. Hence, utility values focus on the outcome of a process and can be described as the perceived usefulness of an object (such as chemistry) to pursue a specific goal (Jiang et al., 2018; Wigfield & Eccles, 2000). These goals can be related to one's personal life as well as to societal and communal issues (Bathgate & Schunn, 2017b; E. R. Brown et al., 2015; Eccles & Wigfield, 2002; Hill & Tyson, 2009). Accordingly, utility values predict students' behavioral, cognitive and social engagement in science (Fredricks et al., 2018). Vice versa, students with a more sophisticated understanding and with a higher self-perception of competence in a given domain are more likely to perceive this domain as relevant and to attribute a greater potential for utility (Durik, Hulleman, et al., 2015; Durik, Shechter, et al., 2015). Likewise, interest contains intrinsic values for a domain or an activity and should therefore be closely connected to students' utility values (Durik, Hulleman, et al., 2015; Hidi & Renninger, 2006).

## The Present Study

In the present study, we strived to extend research findings pertaining to the longitudinal interplay between interest and achievement. Furthermore, we analyzed students' utility values as outcomes of students' interests and their achievement.

We sought to answer the following research questions:

- Are students' interests associated with their achievement gains and is, vice versa, achievement associated with changes in students' interests?
- Are couplings between students' interests and achievement moderated by secondary school level?
- How are students' utility values associated to their interests and their achievement?

Generally, we expected a decline in students' interests across secondary school and a parallel increase of their achievements. Furthermore, we hypothesized positive reciprocal relations between changes of interests in investigative, networking, and social school science activities and gains in students' achievement, as these dimensions of interest are closely connected to cognitively challenging tasks and learning processes (Höft et al., 2019). We also expected stronger longitudinal couplings for upper secondary students than for lower secondary students (cf. Denissen et al., 2007). With regard to utility values, we expected positive relations to students' individual interests (cf. Lauermaun et al., 2017) and achievement (cf. Hulleman et al., 2010).

## 7.3 Method

### Sample and Procedure

The current study is based on longitudinal data obtained from five schools in Northern Germany (academic track, attended by about 66% of all students in this federal state) which voluntarily participated in the study (convenience sampling method). A three years (i.e., four annual measurement waves) two-cohort design was employed, following lower secondary students in Grade 5 up to Grade 8 (cohort denoted as LS5) and upper secondary students in Grade 9 up to Grade 12 (cohort denoted as US9), thus covering all grade levels at secondary school (Grade 5 to 12). From the original sample, we excluded responses of lower secondary students who participated only once and of upper secondary students who opted out of chemistry education or participated less than three times (US9). The reduced samples comprised  $n = 493$  lower secondary students (54.2% female,  $M_{\text{age}} = 10.6$  years in Grade 5) and  $n = 382$  upper secondary students (57% female,  $M_{\text{age}} = 14.7$  years in Grade 9), respectively. We used data on students' interest, achievement, and utility values gathered in Grade 6 to 8 (i.e., measurement waves 2 to 4 in this project) and Grade 9 to 12 (i.e., measurement waves 1 to 4 in this project), as chemistry education starts in grades 6–7 (as a disciplinary subject) and students are free to opt out of chemistry education after finishing Grade 10. Participation in questionnaires was voluntary after written consent was provided by the participant's caretakers. Students were also informed that they could dropout from the study at any time without any consequences.

### Measures

Interest in school science activities. We focus on the three dimensions of students' interest in school science activities (IN) of the respective RIASEC+N scales developed by Dierks et al. (2014b) as a cross-sectional study indicates that these dimensions seem to be closely associated with knowledge acquisition: 1) interest in investigative (I, e.g., „to find an explanation for an observation“), 2) networking (N, e.g., „to work on a chemical topic together with fellow students“), and 3) social (S, e.g., „to explain chemical topics to fellow students“) school science activities (results pertaining the other four dimensions of the RIASEC+N model will be reported in the supplementary materials). Each dimension was measured using three items (after one item per scale was dropped due to low factor



loadings) on a four-point Likert scale ranging from „I am not interested at all in doing this“(1) to „I am very interested in doing this“(4). Scores on these measures demonstrated acceptable to good internal reliabilities across all grades and measurement points ( $\alpha = [.66, .87]$ ; see Table 7.1).

Utility values. Two facets of utility values were assessed per cohort at the fourth measurement occasion; i.e., utility for future life (UF; OECD, 2009, p. 325) and social utility (SU; OECD, 2009, p. 324). Utility for future life was measured using four items (e.g., „Chemistry is useful for my future.“; one of the items was excluded in LS5 due to poor fit) and social utility (e.g., „Chemistry is valuable to society.“) using three items. Students' responded on a five-point Likert scale ranging from „totally disagree“(1) to „totally agree“(5). Scores on these measures demonstrated good internal reliabilities ( $\alpha = [.78, .88]$ ; see Table 7.1).

Achievement. In the present study, achievement (ACH) was assessed as students' conceptual understanding of three fundamental concepts of chemistry education: the structure and composition of matter, chemical reaction, and energy. This selection was based on the German educational standards (KMK, 2004). Test items comprised classical multiple-choice items and so called ordered multiple-choice items (OMC; Briggs et al., 2006). In this item format, the different answer options correspond to different levels of understanding. The description of these levels were based on empirically validated models for assessing students' conceptual understanding across multiple grade levels for each of the three concepts (Hadenfeldt, Liu, et al., 2014; Neumann et al., 2013; Sebastian Tobias Opitz et al., 2017). The test composition was based on a common item anchor design with both anchor (i.e., items are administered across grades) and grade-specific items (i.e., items are administered only in a particular grade). Each student answered 30 items (10 per concept) and in case of multiple-choice items students' answers were coded as correct (1) or incorrect (0), or, in case of OMC items, assigned partial credits (0 to 3) in correspondence to the underlying theoretical. Data was analyzed by a one-dimensional multiple-group generalized partial credit model (MG-GPCM) with students' grade (from 5 to 12) as grouping variable to account for the matrix design of the test (i.e., to take into account that each student received only a subset of all items and that this subset differed between grades). This analysis was done for each measurement wave individually and weighted likelihood estimates (WLE) were calculated (WLE reliability =  $[.71, .77]$ ; see Table 7.1). Finally, these estimates of students' achievement, as measured by the test, were linked across measurement waves based on a generalization of the log-mean-mean linking in accordance with Haberman (2009) to obtain the same metric across grade levels.

Covariates. Students' gender (1 = female, 2 = male), age, and cognitive abilities and the grade when chemistry education starts (LS5: 21.1% in Grade 6, 63% in Grade 7, 16.2% in Grade 8) were used as control variables. Students' cognitive abilities were assessed at the first measurement wave using the Figure Analogies subscale (LS5:  $\alpha = .87$ ; US9:  $\alpha = .80$ ) from the Cognitive Abilities Test 4–12+R (Heller & Perleth, 2000).

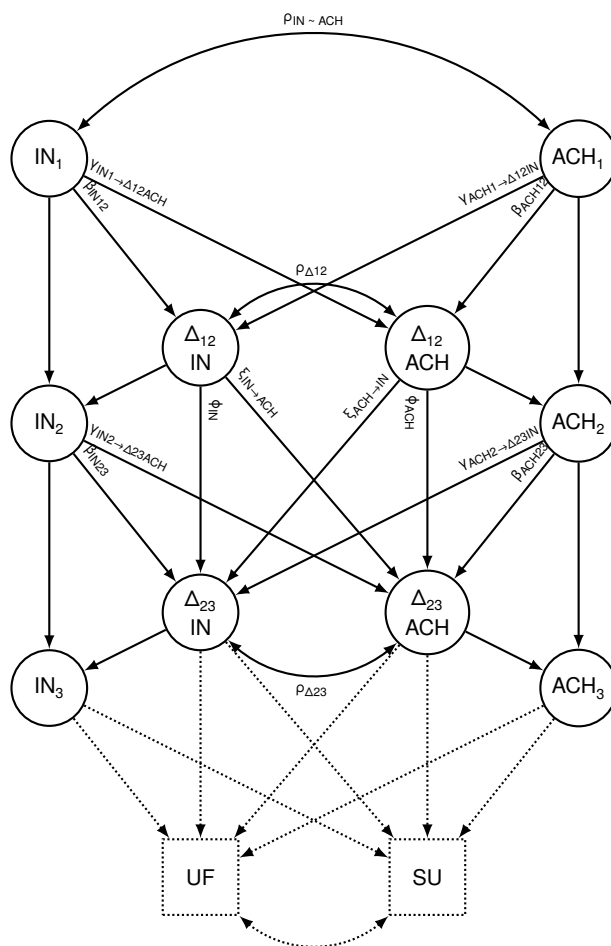
Statistical Analyses. Structural equation modelling (SEM) was applied to the data in order to examine the development and the interplay between students' interest in school science activities and their achievement as well as connections to students' value beliefs. All SEM analyses were conducted in Mplus 8 using the robust maximum-likelihood estimator (MLR) which accounts for non-normality by correcting test statistics and standard errors.

We assessed the model fit using the comparative fit index (CFI > .90 indicating acceptable, CFI > .95 excellent fit), the root-mean-square error of approximation (RMSEA < .08 indicating acceptable, RMSEA < .06 good fit), and the standardized root mean square residual (SRMR < .08 indicating good fit) (Browne & Cudeck, 1993; Hu & Bentler, 1999). Missing values ranged for LS5 from 0.20% to 29.21% (M = 17.13%, SD = 4.44%) and for US9 from 0% to 36.13% (M = 27.84%, SD = 6.13%) across all four measurement waves. We used multiple imputation as implemented in Mplus to replace

these missing values by generating  $m = 200$  datasets per cohort. We employed the correlated uniqueness approach to control for systematic measurement error.

Our data analysis approach consists of two main steps. First, we performed descriptive statistics and ensured measurement invariance across time (Chen, 2007, p. 501). Second, we conducted bivariate latent change score models (LCSM) to investigate the interplay between interest and achievement across time as well as their connections to students' utility values (see Figure 7.1). In Figure 7.1 and hereinafter, measurement waves are represented by indices and changes by  $\Delta$ s. For instance, students' achievement at the first measurement wave is denoted as  $ACH_1$  and learning gains between the first and the second measurement wave are referred to as  $\Delta_{12ACH}$ .

Here, we first performed LCSMs with freely estimated cross-lagged effects to capture effects from students' previous interest on subsequent learning gains, effects from students' previous achievement on subsequent changes of interest, and cross-lagged effects between previous change and subsequent change (change-change effects). Afterwards, we calculated more parsimonious LCSMs whose cross-lagged effects were constrained to be time invariant (interest on achievement,  $\gamma_{IN1} \rightarrow \Delta_{12ACH} = \gamma_{IN2} \rightarrow \Delta_{23ACH}$ , and achievement on interest  $\gamma_{ACH1} \rightarrow \Delta_{12IN} = \gamma_{ACH2} \rightarrow \Delta_{23IN}$ ), i.e., regression coefficients of cross-lagged effects of interest on learning (and vice versa) were fixed to be equal across time. Finally, we included students' value beliefs as endogenous variables into the LCSMs and evaluated connections to students' individual interest and achievement.



Note. IN = first latent variable, i.e., an interest dimension of the RIASEC+N model, ACH = second latent variable, i.e., achievement, US = utility for school, UF = utility for future life, SU = social utility. Indices represent measurement waves.

Figure 7.1: Bivariate latent change score model for lower and upper secondary students.

## 7.4 Results

## Preliminary Analyses

We calculated descriptive statistics on the means, their standard deviations, and correlations between variables within measurement waves (i.e., within school grades) as displayed in Table 7.1. Results indicate that, compared to the center of the respective scale, students on average express higher interest in investigative activities, and hold positive utility values (LS5 – SU:  $t(425) = 5.95$ ,  $p < .001$ , UF:  $t(424) = 3.04$ ,  $p < .001$ ; US9 – SU:  $t(301) = 3.86$ ,  $p < .001$ ), except for upper secondary students utility for future beliefs (UF:  $t(303) = -3.50$ ,  $p < .001$ ).

**Table 7.1:** Reliabilities, descriptive statistics, and correlation coefficients within grades.

Grade	Scale	n	M	SD	Reliability <sup>a</sup>	Correlations				
						I	S	N	SU	UF
lower secondary students (LS5)										
6	I	406	2.84	0.89	0.75					
6	S	403	2.80	0.83	0.81	.53***				
6	N	407	2.80	0.91	0.75	.70***	.59***			
6	ACH	444	-0.38	0.42	0.75	.21***	.05	.20***		
7	I	386	2.80	0.85	0.69					
7	S	386	2.71	0.86	0.85	.48***				
7	N	384	2.78	0.89	0.73	.58***	.59***			
7	ACH	446	-0.14	0.51	0.77	.16**	.04	.11*		
8	I	425	2.73	0.88	0.72					
8	S	425	2.65	0.83	0.86	.42***				
8	N	384	2.70	0.86	0.75	.63***	.57***			
8	SU	429	3.25	1.04	0.81	.38***	.29***	.33***		
8	UF	429	3.14	1.10	0.84	.45***	.29***	.41***	.63***	
8	ACH	430	0.04	0.46	0.77	.22***	.11*	.19***	.25***	.22***
upper secondary students (US9)										
9	I	307	2.76	0.88	0.66					
9	S	305	2.49	0.88	0.85	.45***				
9	N	306	2.59	0.89	0.76	.58***	.60***			
9	ACH	317	0.30	0.56	0.74	.21***	.23***	.16**		
10	I	296	2.61	0.88	0.71					
10	S	297	2.44	0.86	0.84	.51***				
10	N	296	2.52	0.90	0.79	.62***	.56***			
10	ACH	320	0.36	0.53	0.75	.19***	.17**	.23***		
11	I	252	2.62	0.88	0.71					
11	S	253	2.44	0.88	0.87	.44***				
11	N	252	2.51	0.93	0.84	.63***	.67***			
11	ACH	263	0.73	0.71	0.77	.31***	.32***	.37***		
12	SU	305	3.19	1.02	0.78					
12	UF	305	2.80	1.14	0.85				.54***	

a Reliability is equal to Cronbach's alpha (I, N, S, SU, UF) and WLE reliability (ACH) respectively.

Note. I = investigative, S = social, N = networking, ACH = achievement, SU = social utility; UF = utility for future.

\*\*\* $p \leq .001$ .

To ensure scalar measurement invariance across measurement waves, we separately conducted bivariate SEMs per interest dimension and cohort (including scores of one interest dimension and achievement). We sequentially constrained factor loadings (metric invariance) and intercepts (scalar invariance), i.e., each step increased the level of restrictions compared to the baseline model (Wang & Wang, 2012). Model fit indices were compared, suggesting scalar measurement invariance of all interest constructs across all three measurement waves (see Table D.1).

## Bivariate Analysis

We performed bivariate LCSMs to derive insights into the longitudinal interplay between interest in investigative, networking, and social school science activities and achievement and to analyze the effects of both constructs on students' value beliefs at Grade 8 and Grade 12, respectively. There was no significant deterioration of model fit when constraining cross-lagged effects to be time invariant (CFIs  $\geq .910$ , RMSEAs  $\leq .074$ , SRMRs  $\leq .057$ ). Next, we added students' utility values as manifest outcome indicators into the bivariate LCSMs (see Figure 7.1). After adding utility values into the LCSMs, the model fit was still acceptable (CFIs  $\geq .911$ , RMSEAs  $\leq .071$ , SRMRs  $\leq .052$ ). The fraction of missing information (FMI) calculated as described by Pan and Wei (2016) did not exceed .50 (LS5: FMI  $\leq .37$ ; US9: FMI  $\leq .46$ ). In the following sections, we will only present results of the final, more parsimonious models, i.e. models including time invariant cross-lagged effects and utility values as endogenous variables (see Table 7.2).

**Development of interest and achievement.** The results of the bivariate LCSMs indicate negative couplings of moderate to high magnitude between students' prior interest and their subsequent change of interest for all three interest dimensions. This finding implies that students with higher interest exhibit a higher decline of interest in the following year. Moreover, for lower secondary students (LS5), the changes of interests in social and networking activities negatively predicts subsequent changes suggesting that students with smaller losses of interests between the first two measurement occasions subsequently experienced a great decline of interests.

Likewise, students' prior achievement negatively predicts a moderate to high growth of knowledge, i.e. students with higher prior achievement exhibit lower achievement gains. Additionally, prior change of achievement negatively predicts subsequent change, i.e., students with lower knowledge acquisition between the first two measurement occasions subsequently exhibit greater learning gains.

**Interplay between interest and achievement.** Students' initial interests in investigative, networking, and social school science activities are positively correlated with their initial achievement ( $.14 \leq \rho_{12} \leq .25$ ,  $p < .05$ ), with the exception of lower secondary students' interest in social activities. Students' individual changes of interests and achievement gains were only significantly correlated in upper secondary school. Here, changes ( $\Delta_{23}$ ) of students' interests in investigative and social activities and achievement were positively correlated. Moreover, the analyses yielded positive cross-lagged effects of students' achievement on subsequent changes of students' interest in investigative (LS5) and networking school science activities (US9). Moreover, students' interests in investigative (US9), social (US9), and networking (LS5, US9) school science activities predicted subsequent learning gains. There were no significant change-change paths within any construct. All correlations and regression coefficients reported above can be interpreted as small ( $< .30$ ).

**Effect of interest and achievement on utility values.** Hereinafter, we solely report effects of interest and achievement on students' utility values as assessed at the last measurement wave. For upper secondary students (US9), there was a time-lag of one year between measures of students' interests and achievement on the one hand and students' utility values on the other hand, while for lower secondary students (LS5), measures of students' interest, achievement, and utility values were collected in the same year. The findings suggest that students' prior interests and achievement positively predict their utility values for future life ( $.16 \leq \beta \leq .47$ ,  $p < .01$ ), except for achievement when controlling for lower secondary interest in investigative activities, and societal utility value ( $.18 \leq \beta \leq .37$ ,  $p < .05$ ). When comparing impacts of individual interests and achievement on students' utility values, Wald tests indicate that only students' interest in investigative school science activities is closer associated with utility for future values than achievement (LS5:  $\chi^2(1) = 9.63$ ,  $p < .01$ , US9:  $\chi^2(1) = 5.66$ ,  $p < .05$ ).

**Table 7.2:** Results of the bivariate latent change score models (LCSM) for lower secondary students (LS5) and upper secondary students (US9) including utility values. Displayed are standardized  $\beta$ -regression and, correlation coefficients.

	Investigative		Social		Networking	
	LS5	US9	LS5	US9	LS5	US9
<b>Lagged Effects (Change <math>\leftarrow</math> Factor, <math>\beta</math>)</b>						
12IN $\leftarrow$ IN1	-0.525***	-0.285*	-0.489***	-0.493***	-0.570***	-0.398***
23IN $\leftarrow$ IN2	-0.369***	-0.404***	-0.399***	-0.511***	-0.367***	-0.428***
12ACH $\leftarrow$ ACH1	-0.367***	-0.544***	-0.361***	-0.553***	-0.375***	-0.539***
23ACH $\leftarrow$ ACH2	-0.390***	-0.147*	-0.380***	-0.153*	-0.384***	-0.148*
<b>Cross-lagged Effects (Change <math>\leftarrow</math> Factor, <math>\gamma</math>)</b>						
12IN $\leftarrow$ ACH1	0.079*	0.025	0.038	0.051	0.012	0.150**
23IN $\leftarrow$ ACH2	0.098*	0.021	0.054	0.045	0.017	0.111**
12ACH $\leftarrow$ IN1	0.058	0.106*	0.061	0.175***	0.108**	0.135**
23ACH $\leftarrow$ IN2	0.052	0.099	0.064	0.151***	0.106**	0.126**
<b>Stabilities (Change <math>\leftarrow</math> Change, <math>\phi</math>)</b>						
23IN $\leftarrow$ 12IN	-0.151	-0.163	-0.181*	-0.052	-0.307***	-0.154
23ACH $\leftarrow$ 12ACH	-0.304***	-0.259***	-0.301***	-0.263***	-0.300***	-0.253***
<b>Cross-lagged Effects (Change <math>\leftarrow</math> Change, <math>\xi</math>)</b>						
23IN $\leftarrow$ 12ACH	-0.119	0.078	-0.047	-0.044	-0.062	-0.046
23ACH $\leftarrow$ 12IN	0.061	0.021	0.008	0.048	0.005	-0.071
<b>Correlations (<math>\rho</math>)</b>						
IN1 $\sim$ ACH1	0.171**	0.245***	0.024	0.203***	0.139*	0.162**
12IN $\sim$ 12ACH	0.119	0.15	-0.013	0.095	-0.005	0.088
23IN $\sim$ 23ACH	0.043	0.261**	0.028	0.159*	0.058	0.133
<b>Regression Estimates (Utility Values <math>\leftarrow</math> Factor)</b>						
SU $\leftarrow$ IN3	0.368***	0.205**	0.265***	0.182*	0.297***	0.227**
SU $\leftarrow$ ACH3	0.210***	0.342***	0.258***	0.358***	0.250***	0.343***
UF $\leftarrow$ IN3	0.470***	0.467***	0.309***	0.289***	0.370***	0.428***
UF $\leftarrow$ ACH3	0.107	0.225**	0.173**	0.319***	0.160**	0.261**
<b>Intercorrelations (Utility Values)</b>						
SU $\sim$ UF	0.547***	0.502***	0.587***	0.524***	0.568***	0.502***
<b>Model Fit</b>						
CFI	0.949	0.911	0.973	0.949	0.938	0.939
RMSEA	0.049	0.071	0.042	0.065	0.056	0.065
SRMR	0.037	0.052	0.031	0.038	0.047	0.047

Note. IN = interest, ACH = achievement; Indices represent measurement waves; \*\*\*p < .001, \*\*p < .01, \*p < .05,  $\cdot$ p < .1.

## 7.5 Discussion

The goal of this study was to systematically analyze the development of important precursors of students' career choices, i.e., students' individual interest, their achievement, and their utility values. Based on the reported findings, a complex image emerges with regard to the interplay of individual interests and achievement. Prior interests in investigative (US9), networking (LS5 and US9), and social (US9) activities positively predict subsequent learning gains while prior achievement seems to exert at least small positive effects on students interests in investigative (LS5) and networking (US9) activities (cf. Table 7.1).

Across the two cohorts analyzed in this study, the findings also indicate a shift from rather weak couplings between interest and acquisition of achievement in lower secondary school to significant relations between students' interests in specific school science activities and their achievement in

upper secondary school, including reciprocal relations between interest in networking activities and achievement. These findings correspond to findings reporting positive effects of interest on achievement in upper secondary school (cf. Jansen et al., 2016) and could be explained by positive effects of interest on learning through maintaining attention as well as increased persistence, effort and positive effects on voluntary engagement (cf. M. Ainley, Hidi, & Berndorff, 2002; Tsai et al., 2008). For instance, students with higher interests in social activities may more frequently help peers in their understanding, e.g., while learning for exams or doing homework.

Complementary, higher achievements should increase the likelihood to succeed in tasks, to feel competent and to experience enjoyment, for instance, when collaborating in small groups or successfully solving a demanding problem. Consequently, higher achievement may facilitate interest development in investigative and networking activities (Bathgate & Schunn, 2017a). In sum, reciprocal, cross-lagged effects were only found between upper secondary students' interest in networking activities and achievement. One premise to explain this pattern across the three dimensions of interest is that networking activities provide both a specific potential for cognitively activating learning opportunities and to experience competence while working with peers, which in turn could enhance the inter-relation between students' interest and achievement (Höft et al., 2019).

With regard to interest in investigative activities, an inconclusive pattern emerged. While students in introductory chemistry courses might have problems understanding the status and the function of conducting inquiry-oriented experiments (Duschl & Grandy, 2011; Stamovlasis et al., 2005), the perception of these investigative activities and of what can be considered as investigative activities might change over the course of secondary school.

Finally, our analyses revealed that individual interests in school science activities and achievement predicted students' utility value beliefs. This finding is in alignment with EVT assuming positive associations between utility values, individual interest, and achievement (cf. Hulleman et al., 2010) as well as positive impacts from achievement on utility values (Durik, Hulleman, et al., 2015; Durik, Shechter, et al., 2015). Highest associations between students' societal utility values were found to students' interest in investigative (LS5) and networking (US9) activities while utility for future shows the strongest relation to investigative (LS5, US9) and networking (US9) school science activities. Hence again, students' interest in school science activities that are associated with cognitive activation or the communication of knowledge (investigative, networking) are especially predictive for students valuing Chemistry to be useful for their future life.

Consequently, fostering both individual interest and achievement seems to be important throughout secondary school. Here, achievement could play a protective role regarding the course of interest development, as the results indicate a smaller decline of students' interest when their understanding is high. Consequently, approaches to foster or, at least, to maintain students' interest should not drift to merely entertaining students, but should rather aim to foster both interest and understanding by meaningful and relevant contexts, which in turn might also impact students' utility values (cf. E. R. Brown et al., 2015).

## 7.6 Limitations

With regard to the theoretical background presented in the beginning, a major constraint is the absence of proximal information about teaching and learning processes in the classrooms of the participating students. This lack of information pertains to the quantity and quality of the implementation of the different school science activities addressed in the measures of students' interests, but also to other features of students' day-to-day experiences in these chemistry classes. Consequently, we were not able to include these crucial classroom variables into our analysis to investigate how the quantity and quality of these school science activities affect the development of students' interests, their achievement as well as the interplay between both constructs. Further limitations of the current study mainly pertain to the data and resulting restrictions for the data analysis process. Our data derived from a convenience sample, thus the generalizability of our findings are confined, as the sample might not represent the student population as a whole. However, basic

demographic information of our sample does not indicate substantial deviations from characteristics of a representative sample (cf. Leucht et al., 2016). A severe, but common problem in any longitudinal study are missing values due to panel mortality. We performed multiple imputation for missing data and also assessed all classes of each school to address this issue and to obtain reliable results. Moreover, we only assessed students' utility values at the last measurement wave. We therefore failed to control for prior utility values when estimating the effect of students' interests and conceptual knowledge on subsequent utility values. It would therefore be worthwhile to test if interest and conceptual knowledge predict subsequent utility values when controlling for auto-regressive effects in all three constructs.

## 7.7 Conclusions

Our results support the assumption that the development of students' individual interests, achievement, and utility values are interwoven. Here, it seems important to foster students' individual interests in school science activities and achievement in a combined and continuous manner to establish and profit from directional and reciprocal relations between both constructs as well as positive relations to utility values. Means to facilitate interest development and achievement simultaneously may involve emphasizing the relevance and usefulness of chemistry contents to enhance students' utility values. Depending on the class and age group, the same context may have different effects on interest development and learning (Habig, Blankenburg, et al., 2018). For example, Habig, Blankenburg, et al. (2018) stated that „at least in lower grades, students seem to need a clear relationship to their personal life in order to find a topic interesting. If the contextual framing is more global or societal, the students demonstrated lower interest“. These rather distal contexts, however, might be more interesting to students in higher grades (Broman & Simon, 2015).





# Ich mag, was ich kann oder kann ich, was ich mag? Über das Zusammenspiel von Interesse, Freude und Konzeptwissen im Fach Chemie

## Zusammenfassung

Die Förderung des Interesses und der Freude am Fach sowie der Erwerb eines grundlegenden Verständnisses chemischer Konzepte sind wesentliche Ziele des Chemieunterrichts. Diese motivationalen und kognitiven Personenmerkmale prägen nicht nur das Handeln der Schülerinnen und Schüler im Unterricht, sondern auch außerhalb der Schule. In dieser Studie wurde daher untersucht, inwieweit sich die Freude am Chemieunterricht, das Konzeptwissen und das individuelle Interesse der Schülerinnen und Schüler an der Domäne Chemie gegenseitig beeinflussen. Hierzu wurde in einer längsschnittlichen Fragebogenstudie das individuelle Interesse, die Freude und das Konzeptwissen von  $N = 756$  Gymnasiasten (52 % weiblich, Alter zum Studienbeginn = 14.7 Jahre) in der neunten, zehnten und elften Jahrgangsstufe erfasst sowie die Entwicklung der einzelnen Konstrukte und Wechselwirkungen zwischen den Konstrukten, d. h. Effekte auf die individuelle Veränderung über die Zeit, mittels latenter Strukturgleichungsmodelle untersucht. Univariate Analysen zeigen einen geringen Abfall des individuellen Interesses sowie einen moderaten Abfall der Freude und zeitgleich einen starken Anstieg des Konzeptwissens. Hinsichtlich des Zusammenhangs zwischen den Konstrukten deuten bivariate Latent Change Score Modelle auf beidseitige Wechselwirkungen zwischen den drei Konstrukten hin. Bei gleichzeitiger Berücksichtigung aller drei Konstrukte mittels trivariater Analysen zeigen sich weiterhin reziproke Zusammenhänge zwischen der Freude und dem Konzeptwissen, jedoch nur einseitig positive Effekte der Freude und des Konzeptwissens auf die Veränderung des individuellen Interesses. Die Befunde unterstreichen die Relevanz motivationaler Variablen für das Lernen und weisen auf Potentiale einer gezielteren Integration von Interessen der Schülerinnen und Schüler in den Chemieunterricht hin.

Schlüsselwörter: Freude, Interesse, Konzeptwissen, Chemie, Sekundarstufe I

## Abstract

Fostering students' interest and enjoyment as well as the acquisition of a profound understanding of chemical concepts are crucial goals of chemical education in secondary school. These motivational and cognitive variables do not only shape students behavior in the classroom, but also outside of school. Therefore, this study examines to what extent enjoyment, conceptual knowledge, and individual interest influence each other. For this purpose, we conducted a longitudinal survey study assessing individual interest, enjoyment, and conceptual knowledge of  $N = 756$  students in German secondary schools (52% female, age at the beginning of the study = 14.7) in Grades 9, 10, and 11. By means of latent structural equation modeling, we studied the development of and the interplay between these three constructs, i.e. effects on individual change over time. Results from univariate analyses show a slight decrease of students' individual interest, a moderate decrease of their enjoyment, and concomitantly a large increase of their conceptual knowledge. Pertaining to the interplay between these constructs, bivariate latent change score models indicate bidirectional interactions in all pairwise constellations. When simultaneously taking into account all three constructs, trivariate analyses indicate reciprocal relations between enjoyment and conceptual knowledge, whereas only directed effects of enjoyment and conceptual knowledge on individual interest persists. Overall, these results underscore the relevance of motivational variables for learning and suggest potential benefits of a more systematic integration of students' interests into chemistry education.

Keywords: enjoyment, interest, conceptual understanding, chemistry, secondary school

Höft, L., & Bernholt, S. (2019a). Ich mag, was ich kann oder kann ich, was ich mag? Über das Zusammenspiel von Interesse, Freude und Konzeptwissen im Fach Chemie. *Zeitschrift für Didaktik der Naturwissenschaften*, 25(1), 161–180

## 8.1 Einleitung

Dem Verständnis intrinsischer motivationaler Variablen, wie das Interesse und die Freude, wird in der pädagogisch-psychologischen (u. a. Hidi, 2016; Pekrun, 2006) sowie der fachdidaktischen Forschung (Brandtner & Hertel, 2018; Habig, van Vorst et al., 2018; Potvin & Hasni, 2014b; van Vorst et al., 2018) gesteigerte Aufmerksamkeit entgegengebracht, da sie als Determinanten für selbstbestimmtes, langfristiges Lernen und Engagement mit dem Fach identifiziert wurden (Cerasoli et al., 2014). Sowohl das Interesse als auch die Freude wirken in der Schule während konkreter Lernprozesse (Hidi & Renninger, 2006; Pekrun, 2006), gleichzeitig unterliegen sie aber auch prospektiven und retrospektiven Bewertungsprozessen, beeinflussen zukünftiges Handeln und die Ausbildung der Identität (Cribbs et al., 2015; Eccles, 2009; Rabe & Krey, 2018). Der Einfluss des Interesses und der Freude manifestieren sich weiterhin in Wahlentscheidungen wie Bildungswegentscheidungen (bspw. Bøe, 2012; Eitemüller & Walpuski, 2018; Köller et al., 2001; Nagy et al., 2007), Berufsaspirationen (bspw. Ahmed & Mudrey, 2018; Lauermann et al., 2017; Su, 2018) und der Freizeitgestaltung (Bergin, 1992, 1996). Um, im Kontext der Chemie, eine aktive gesellschaftliche Teilhabe und einen erfolgreichen Berufseinstieg zu initiieren, ist es daher ein immanentes Ziel chemischer Schulbildung, das Interesse und die Freude am Fach zu wecken und zu fördern (Hidi & Harackiewicz, 2000; Maltese & Tai, 2011).

Bisherige Befunde betonen jedoch, dass im Mittel das Fachinteresse der Schülerinnen und Schüler an mathematisch-naturwissenschaftlichen Fächern abnimmt (vgl. Krapp & Prenzel, 2011; Potvin & Hasni, 2014a), während gleichzeitig ein deutlicher Wissenserwerb zu verzeichnen ist (Höft et al., 2019). Hinsichtlich der Chemie ist die Abnahme des Fachinteresses insbesondere vor dem Hintergrund bedauerlich, dass die Fachinhalte zahlreiche Anknüpfungspunkte zur Lebens- und Erfahrungswelt der Schülerinnen und Schüler bieten und somit ein günstiger Ausgangspunkt für die Interessensförderung gegeben ist (Fechner, 2009; Parchmann et al., 2006; van Vorst et al., 2018). So können außerschulische Erfahrungen in den Unterricht integriert und wiederum Unterrichtserfahrungen in Bewertungsprozesse zur Freizeitgestaltung einfließen, sodass eine hohe Relevanzzuschreibung der Chemie für die Identität und eine starke positive Bindung zum Fach angeregt wird. Das Induzieren von Interesse durch den Unterricht auf die Freizeitgestaltung untersuchten bspw. Ciani et al. (2010) und Bergin (1992). Es zeigte sich, dass im Unterricht Interesse in dem Maße getriggert werden kann, dass Schülerinnen und Schüler Unterrichtsthemen autotelisch in ihrer Freizeit weiterverfolgen und dies sich wiederum positiv auf die Performanz im Unterricht auswirken kann. Diese Form des Interesses (das individuelle Interesse), die durch die freiwillige, intrinsische und wiederholte Auseinandersetzung mit Inhalten der Domäne und durch ihren habituellen Charakter gekennzeichnet ist, scheint daher besonders erstrebenswert und förderungswürdig zu sein (Harackiewicz et al., 2002; Hidi & Renninger, 2006; Rotgans, 2015).

Vor diesem Hintergrund beschäftigt sich der vorliegende Artikel mit der Fragestellung, inwiefern die Freude am Chemieunterricht sowie das im Chemieunterricht erworbene Konzeptwissen die Entwicklung des individuellen Interesses an der Domäne Chemie beeinflusst und zu welchen Wechselwirkungen es zwischen diesen Facetten über längere Zeiträume kommt. Ein besseres Verständnis der zeitlichen Entwicklungen und Wechselwirkungen soll dabei erste Hinweise darauf liefern, welche Ansatzpunkte potenziell für die Förderung dieser Personenmerkmale in Frage kommen und somit im Fokus nachfolgender Interventionen stehen könnten.

## 8.2 Theoretischer Hintergrund

### Interesse

Das Interesse kann im Sinne der Person-Gegenstands-Theorie (Krapp, 1999a; Prenzel et al., 1986) als Relation zwischen Person und Gegenstand definiert werden, wobei der Begriff Gegenstand durch die Person konstruierte mentale Repräsentationen, wie das Unterrichtsfach bzw. die Domäne Chemie, umfasst. Gegenwärtige Interessentheorien unterscheiden zwischen „zeit-situationsspezifischen

Person-Gegenstands-Beziehungen“, dem situationalen Interesse, und zeitlich stabilen „Person-Gegenstands-Bezügen“, dem individuellen Interesse (Hidi, 2001; Hidi & Renninger, 2006; Prenzel et al., 1986). Das situationale Interesse bezeichnet einen durch Anreizbedingungen erzeugten, situationsspezifischen motivationalen Zustand (Z. Daniels, 2008; Knogler et al., 2015), wohingegen das individuelle Interesse sowohl eine relativ stabile Neigung (Trait im Sinne einer Predisposition) zur wiederholten, selbstbestimmten, intrinsischen Auseinandersetzung mit einem Gegenstand als auch den während dieser Interaktion hervorgerufenen psychologischen Zustand (State) beschreibt (Hidi & Renninger, 2006). Die Entwicklung von einem situationalen Interesse hin zu einem individuellen Interesse beschreiben Hidi und Renninger (2006) anhand von Internalisierungsprozessen in einem vierphasigen Modell. Während das situative Interesse in der ersten Phase durch eine starke Wechselbeziehung zwischen spezifischen Merkmalen einer Lernumgebung und dem Interesse einer Schülerin bzw. eines Schülers geprägt ist, nimmt der Einfluss der Lernumgebung über die vier Phasen schrittweise ab. Gleichmaßen bildet sich ein zunehmend stabiles individuelles Interesse auf Seiten der Person aus, welches in der vierten Phase als eine dauerhafte Prädisposition charakterisiert wird. Übergreifend ist jeder Phase gemein, dass das Interesse in verschiedene, intrinsische Valenzen differenziert werden kann, bspw. nach Schiefele (1991) in eine affektive (u. a. Freude, Aktiviertheit) und eine gefühlsneutrale Valenz (u. a. Bedeutsamkeit), wobei angenommen wird, dass die affektive Valenz in frühen Formen des Interesses dominant ist (Frenzel et al., 2012; Knogler, 2017). Entsprechend sind dem individuellen Interesse die Wahrnehmung von positiven Emotionen und die Bedeutsamkeit, die einem Gegenstand zugeschrieben wird, zugehörig (Schaffner & Schiefele, 2015). Empirische Befunde zur Entwicklung des individuellen Interesses an mathematisch-naturwissenschaftlichen Fächern zeigen, dass dieses in der Sekundarstufe mit steigender Jahrgangsstufe abnimmt (bspw. Frenzel et al., 2010; Krapp & Prenzel, 2011; Potvin & Hasni, 2014a). Als Gründe für die Interessenabnahme werden in Bezug auf die Schulumgebung u. a. mangelnde Passung bspw. zwischen Lernumgebungen, -inhalten und Bedürfnissen sowie Interessen der Schülerinnen und Schüler und eine auf temporalen, dimensional sowie sozialen Vergleichen beruhende Ausdifferenzierung der Interessen im Zuge der Entwicklung des Selbstkonzept genannt (bspw. Z. Daniels, 2008; Krapp & Prenzel, 2011).

In Interessentheorien wird generell ein positiver Zusammenhang zwischen Interesse und Leistung diskutiert (Alexander, 2003; Hidi, 2006; Rotgans & Schmidt, 2017d). Diese positive Verknüpfung zwischen Interesse und Leistung konnte u. a. in einer Metaanalyse von Schiefele et al. (1992) gefunden werden, wobei der Zusammenhang zwischen den beiden Personenmerkmalen in höheren Jahrgangsstufen stärker ausgeprägt ist (vgl. Denissen et al., 2007; Höft et al., 2019; Kim et al., 2015). Als Grund hierfür wird häufig angeführt, dass ein höheres Interesse die Aufmerksamkeitslenkung (bspw. Hidi, 2001), Selbstregulationsprozesse (bspw. Lee et al., 2014) und Persistenz (M. Ainley, Hidi & Berndorff, 2002) während sowie die Qualität des Lernens (Schiefele, 1991) fördert, Conceptual Change erleichtert (Cordova et al., 2014) und somit zu einer besseren Performanz führen kann. Diese positiven inter- und intraindividuellen Effekte des Interesses auf die spätere Leistung konnten Jansen et al. (2016) sowie Kriegbaum et al. (2015) mithilfe eines für Deutschland repräsentativen Samples für unterschiedliche Domänen (Deutsch, Mathematik, Biologie, Chemie und Physik) nachweisen. Ebenso wird davon ausgegangen, dass bei positiv attribuerter Performanz und Kompetenzerfahrung, mediiert über das Selbstkonzept (D. Cheung, 2017; Fryer et al., 2016) bzw. die Selbstwirksamkeitserwartungen (Silvia, 2003), das Interesse steigt (vgl., Eccles & Wigfield, 2002; Krapp, 2005). Studien, die mithilfe von cross-lagged Designs den Zusammenhang zwischen dem individuellen Interesse und der Performanz untersuchten, d. h. kreuzverzögerte (cross-lagged) Effekte bei autoregressiver Kontrolle analysierten, zeichnen kein eindeutiges Bild. Für Schülerinnen und Schüler der Primarstufe und Sekundarstufe I indizieren empirische Befunde positive Effekte der Leistung auf das spätere individuelle Interesse (Ganley & Lubienski, 2016; Grigg et al., 2018; Jögi et al., 2015; Köller et al., 2001; Marsh et al., 2005; Rotgans & Schmidt, 2017c; Viljaranta et al., 2014), wobei der Effekt durch das Selbstkonzept mediiert zu sein scheint (s. Marsh et al., 2005; Sewasew et al., 2018; Viljaranta et al., 2014). Jedoch konnten in bivariaten cross-lagged Designs nur teilweise zeitgleich positive Wirkungen des Interesses auf die spätere Leistung gefunden werden (bspw. Marsh et al., 2005; Schaffner et al., 2016). So deuten die Ergebnisse von cross-lagged Analysen von Köller et al. (2001) daraufhin, dass zwischen dem

Mathematikinteresse von Schülerinnen und Schülern in der siebten Klasse und ihrer Testperformanz in Klassenstufe 10 kein signifikanter, kreuzverzögerter Effekt ( $\beta = .09$ ,  $p > .05$ ), simultan aber ein positiver Effekt der Testperformanz auf das spätere Interesse vorliegt ( $\beta = .20$ ). Marsh et al. (2005) hingegen fanden bei der Analyse zweier für Deutschland repräsentativer Stichproben, dass in bivariate cross-lagged Analysen das Mathematikinteresse von Schülerinnen und Schülern der siebten Klassenstufe die spätere Mathematiknote am Ende der Jahrgangsstufe 7 bzw. in Jahrgangsstufe 8 ( $\beta = .15$ ;  $\beta = .14$ ) sowie die Testperformanz ( $\beta = .09$ ;  $\beta = .10$ ) signifikant vorhersagten. Köller et al. (2001) führten das Ausbleiben positiver Effekte des Interesses auf die spätere Performanz auf die lange Zeitspanne zwischen den beiden Messzeitpunkten sowie auf das Wirken extrinsischer Lernanreize zurück, insbesondere hinsichtlich des Erreichens guter Noten. Diese extrinsischen Lernanreize könnten die intrinsischen Lernanreize (bspw. in Form individueller Interessen) in ihrer Wirkung auf die Leistungsentwicklung übertreffen und dadurch den Effekt des Interesses auf die spätere Leistung unterminieren. Für Probanden der Sekundarstufe II und der Universität konnte hingegen weitere Evidenz gefunden werden, dass das Interesse durch Wahlentscheidungen auf die spätere Performanz wirkt (Harackiewicz et al., 2008; Köller et al., 2001). Köller et al. (2001) berichteten in der o. g. Studie, dass das Mathematikinteresse der Schülerinnen und Schüler in der Jahrgangsstufe 10 die spätere Mathematikperformanz in Jahrgangsstufe 12 vorhersagt ( $\beta = .30$ ), und fanden, dass dieser Effekt durch die Kurswahl (Leistungskurs vs. Grundkurs) moderiert wurde. Gleichfalls war der kreuzverzögerte Effekt der Performanz auf das spätere Interesse nicht mehr bedeutsam.

## Freude

Die Freude kann im Schulkontext als Leistungsemotion im Rahmen der Kontroll-Wert-Theorie (KWT) konzeptualisiert werden (Pekrun, 2006). Die KWT dient der Analyse und Deutung von Antezedenzen und Effekten von Emotionen in Leistungs- und akademischen Kontexten. Die Freude wird hier als positive, aktivierende Leistungsemotion, die bei Fokussierung auf die Aktivität als (Lern-)Freude, bei Fokussierung auf das Ergebnis der Aktivität prospektiv als Vorfreude, retrospektiv als (Ergebnis-)Freude, erlebt werden kann. Ursachen für das Erleben von Leistungsemotionen sind emotionale Traits (habituelles Erleben der Leistungsemotion) und kognitive Bewertungsprozesse (Appraisals) von Situationen oder Tätigkeiten (Nett et al., 2017). In diesen Bewertungsprozessen sind zentral die Appraisaldimensionen der subjektiven Kontrolle (Facetten sind bspw. das Selbstkonzept, die Selbstwirksamkeit) sowie des subjektiven Werts (bspw. die Wichtigkeit des Ergebnisses). Die Vorfreude kann demzufolge bei einer Appraisalkonstellation, die durch hohe Kontrolle und positiven Wert charakterisiert ist, auftreten. Die Ergebnisfreude ist bei eingetretenem Erfolg abhängig von einem hohen Wert, jedoch unabhängig von Kontrollkognitionen. Für positive, aktivierende Leistungsemotionen wie die Vor- und Lernfreude wird angenommen, dass diese die Aufmerksamkeitssteuerung, das Interesse, die intrinsische Motivation sowie das „tiefe“ Lernen unterstützen (Pekrun, 2006). „Tiefes“ Lernen bezieht sich in diesem Zusammenhang vor allem auf Erkenntnisse aus der Textverständnisforschung (bspw. Kintsch, 1988; van Dijk & Kintsch, 1983), in der oftmals zwischen wörtlicher, propositionaler und situationaler Ebene des Textverständnisses differenziert wird. Dabei sind die wörtliche (d. h. grafisch-visuelle oder phonetisch-auditive Merkmale) und propositionale Ebene (d. h. unmittelbare Bedeutung des Texts) zwar relevant hinsichtlich des Textverständnisses, jedoch kennzeichnen sie eher reproduktives und oberflächliches Lernen. Demgegenüber ermöglicht die situationale Ebene die Anwendung der Informationen aus dem Text für Problemlösungen oder zur kritischen Reflexion, so dass erst diese Ebene „echtes“ Lernen (Schaffner und Schiefele, S. 756) darstellt, weil es über die Reproduktion von Textinhalten hinaus- und aus der Integration von Textinhalten und dem Vorwissen der Leserin bzw. des Lesers hervorgeht (Kintsch, 1988) Insofern beeinflussen die Vor- und Lernfreude sowohl wie als auch was gelernt wird. Die prospektiv sowie die während der Aktivität erlebte Freude wirkt sich somit auf das Ergebnis der Leistungsaktivität aus, welches wiederum bei eingetretenem Erfolg durch retrospektive Attribuierung Ergebnisfreude evoziert (Pekrun et al., 2017; Putwain et al., 2017).

Aufgrund dieses beidseitigen Zusammenhangs zwischen Freude und Ergebnis postuliert die KWT

reziproke Zusammenhänge zwischen Freude und Performanz, welche durch positive Feedback-Schleifen beschrieben werden. Für diese angenommenen reziproken Zusammenhängen konnte mittels bivariater Strukturgleichungsmodelle für das Fach Mathematik empirische Evidenz generiert (Pekrun et al., 2017; Pinxten et al., 2014) und überdies gezeigt werden, dass der Zusammenhang, im Einklang mit der KWT, durch das akademische Selbstkonzept mediert wird (Goetz et al., 2008). So untersuchten Pekrun et al. (2017) für  $N = 3.425$  Schülerinnen und Schüler den Zusammenhang zwischen den jährlich erhobenen Leistungsemotionen und Performanzmaßen im Verlauf der Jahrgangsstufen 5 bis 9 in der Mathematik mithilfe eines cross-lagged Designs und fanden kleine, positive, kreuzverzögerte Effekte der Mathematiknote ( $\beta = .11, p < .001$ ) bzw. der Testperformanz auf die Freude ( $\beta = .07, p < .001$ ) sowie der Freude auf die Mathematiknote ( $\beta = .13, < .001$ ) und die Testperformanz ( $\beta = .04, p < .001$ ).

Gleichfalls werden in der KWT positive, reziproke Zusammenhänge zwischen der Freude und dem Interesse postuliert. Die KWT nimmt dabei an, dass das individuelle Interesse, u. a. aufgrund selbstregulatorischer motivationaler Aspekte (Hidi, 2006), zu höheren Kontroll- und Wert-Appraisals führen kann (M. Ainley & Hidi, 2014; Pekrun et al., 2007), die ihrerseits positive Leistungsemotionen fördern (bspw. Ahmed et al., 2010; Buff, 2014), sich somit positiv auf die Performanz (bspw. Pekrun et al., 2017) und letztlich zudem positiv auf das Interesse auswirken (bspw. Eccles et al., 2015; Jansen et al., 2016; Kriegbaum et al., 2015). Zur Analyse der Zusammenhänge zwischen Appraisals und Emotionen (u. a. Freude) führten Ahmed et al. (2010) eine Tagebuchstudie ( $N = 120$ , Jahrgangsstufe 7) durch, deren Ergebnisse signifikante Zusammenhänge zwischen den intraindividuellen Varianzen der Appraisalkomponenten und Emotionen indizierte. Hierbei variierte der Zusammenhang zwischen den Wert-Appraisals sowie der berichteten Freude substanziell zwischen den Schülerinnen und Schülern. Weiterhin konnte der beidseitig positive Zusammenhang zwischen Interesse und Freude u. a. von Schukajlow und Rakoczy (u. a. 2016) im Fach Mathematik gezeigt werden. Hierzu erfassten Schukajlow und Rakoczy (2016) Mathematikinteresse und -performanz ( $N = 144$ , Jahrgangsstufe 9) vor und nach einer fünfstündigen Intervention sowie die Freude während der Intervention. Ihre Analysen deuten darauf hin, dass das Interesse positiv mit der während der Intervention erlebten Freude ( $\beta = .34, p < .05$ ) zusammenhängt, welche wiederum einen positiven Zusammenhang zum späteren Interesse aufwies ( $\beta = .30, p < .05$ ).

### Konzeptwissen

Lernen wird in den Naturwissenschaften als erfolversprechend eingeschätzt, wenn es systematisch, an zentralen Konzepten orientiert und kumulativ erfolgt. Diesen Einschätzungen wurde zum Beispiel bei der Erstellung der Bildungsstandards (KMK, 2004) und Neugestaltung der Fachanforderungen (BIMI, 2016) Rechnung getragen, infolge derer für die Chemie die Basiskonzepte Stoff-Teilchen, Struktur-Eigenschafts-Beziehung, chemische Reaktion und Energie für den mittleren Schulabschluss und die Oberstufe in Schleswig-Holstein als zentral definiert wurden. Empirische Studien deuten wiederholt und nachhaltig darauf hin, dass Schülerinnen und Schüler nur bedingt ein grundlegendes Verständnis der zentralen Fachkonzepte aufbauen. Als Möglichkeiten zur stärkeren Abstimmung von Lerninhalten und Lernvoraussetzungen wurden Learning Progressions (vgl. Bernholt & Sevian, 2018; Emden et al., 2018) untersucht, die die Entwicklung des Verständnisses von Konzepten entlang von definierten Verständnisniveaus beinhalten können. Solche Verständnisniveaus beschreiben Stufen, die durch relativ kohärente Interpretationsmuster geprägt sind. So wurden Learning Progressions für das Verständnis der Struktur von Materie (Smith et al., 2006), der Genetik (Duncan et al., 2009) oder des naturwissenschaftlichen Modellierens (Schwarz et al., 2009) und Argumentierens (Berland & McNeill, 2010) vorgeschlagen. Beispielsweise zum Materiekonzept konnten (Hadenfeldt et al., 2016) in einer querschnittlich angelegten Studie zeigen, dass Probanden aus unteren Jahrgangsstufen (wie erwartet) vermehrt Antworten auf unteren Verständnisniveaus und Probanden aus oberen Jahrgangsstufen vermehrt Antworten auf höheren Verständnisniveaus geben (vgl. Emden et al., 2018). Bisherige Studienbefunde für den Zusammenhang zwischen individuellem Interesse bzw. Freude und Performanz liegen vorwiegend für das Fach Mathematik, den Primar- und

frühen Sekundarbereich, vor. Studien, die den Zusammenhang zwischen dem individuellen Interesse bzw. der Freude und dem Konzeptwissen im Chemieunterricht im Längsschnitt untersuchen sind rar. Zudem fehlen Studien, die systematisch den Zusammenhang zwischen dem individuellen Interesse und der Freude am Fach über mehrere Messzeitpunkte hinweg analysieren. Dieses Defizit soll die vorliegende Studie aufgreifen, indem Wechselwirkungen zwischen dem individuellen Interesse an der Chemie, der Freude am Chemieunterricht sowie dem chemiebezogenen Konzeptwissen von Schülerinnen und Schüler der neunten bis zur elften Jahrgangsstufe im Längsschnitt untersucht werden. Neu ist zudem, dass nicht Zusammenhänge zwischen den interindividuellen Unterschieden der Personenmerkmalsausprägungen, sondern Effekte auf individuelle Veränderungen modelliert werden. Die Auswertungen zielen dabei darauf ab, ein besseres Verständnis der zeitlichen Entwicklungen und Wechselwirkungen zwischen den fokussierten Konstrukten zu erreichen. Dieses Verständnis soll letztlich dazu dienen, erste Hinweise auf potenzielle Ansatzpunkte für die Förderung der verschiedenen Personenmerkmale zu liefern.

### Forschungsfragen

In dem vorliegenden Artikel soll das Zusammenwirken des individuellen Interesses, der Freude an und des Konzeptwissens in dem Fach Chemie analysiert werden. Dieses Vorhaben wird durch die folgenden drei Forschungsfragen adressiert:

1. Wie entwickeln sich das individuelle Interesse, die Freude und das Konzeptwissen im Verlauf der Jahrgangsstufen 9 bis 11 in der Chemie?
2. Wie gestalten sich bivariate Zusammenhänge zwischen dem individuellen Interesse und der Freude, dem individuellen Interesse und dem Konzeptwissen sowie der Freude und dem Konzeptwissen?
3. Wie gestalten sich die trivariaten Zusammenhänge zwischen dem individuellem Interesse, der Freude und dem Konzeptwissen?

Ausgehend von den oben diskutierten empirischen Befunden wird bzgl. der ersten Forschungsfrage eine Abnahme des individuellen Interesses und der Freude am Fach bei simultanen Anstieg des Konzeptwissens erwartet. Hinsichtlich der zweiten und dritten Forschungsfrage werden positive, reziproke Zusammenhänge zwischen dem individuellem Interesse, der Freude am Fach und dem Konzeptwissen erwartet, wobei aufgrund der Operationalisierung (Freude bezieht sich auf das Fach, individuelles Interesse auf die Domäne), eine engere Verknüpfung zwischen Freude am Fach und Konzeptwissen als zwischen Konzeptwissen und individuellem Interesse angenommen wird.

## 8.3 Methode

### Durchführung und Stichprobe

Die Daten dieser Untersuchung wurden im Rahmen des längsschnittlich angelegten Projektes DoLiS an sechs Gymnasien aus Schleswig-Holstein, die freiwillig an der Erhebung teilnahmen, generiert. Die Studiendaten stellen somit eine Gelegenheitsstichprobe dar und sind nicht repräsentativ für das Bundesland. Das längsschnittliche Studiendesign sah vier Messzeitpunkte in einem einjährigen Abstand vor, von denen die Daten dreier Messzeitpunkte (T1, T2, T3) in die vorliegende Studie einfließen.<sup>1</sup> Die erste Erhebungswelle fand im Frühjahr 2015 in der neunten Jahrgangsstufe statt; die darauffolgenden Messzeitpunkte erfolgten in den Frühjahren 2016 und 2017. Das hier analysierte Subsample der DoLiS-Studie umfasst die Angaben von N = 756 Schülerinnen und Schülern (52 % weiblich) aus 26 initialen Klassen des neunten Jahrgangs an Gymnasien, deren Alter zum ersten Erhebungszeitpunkt im Mittel 14.7 Jahre betrug. In diese Untersuchung flossen Daten von

1. Der vierte Messzeitpunkt wurde nicht in diese Auswertung mitaufgenommen, da für die hier durchgeführten Analysen zu wenige Probanden akquiriert werden konnten und eine Beeinflussung der Ergebnisse durch das nahezu zeitgleich stattfindende Abitur zu vermuten war.

allen Probanden ein, die mindestens an einem Erhebungszeitpunkt teilnahmen. Hinsichtlich der Teilnahmequote über die drei Messzeitpunkte unterteilt sich die Stichprobe zu etwa jeweils einem Drittel in Schülerinnen und Schüler, die an einem (31,5 %), an zwei (32,7 %) bzw. an allen drei Messzeitpunkten (34,4 %) teilgenommen haben, wobei keine Systematik der Nicht-Teilnahme über die drei Messzeitpunkte festzustellen ist. Am Ende der 9. Jahrgangsstufe haben die Schülerinnen und Schüler die Möglichkeit, ein Profil für die Oberstufe zu wählen. In der vorliegenden Stichprobe haben dabei 55 Probanden (7,4 %) Chemie als Profil gebendes Fach gewählt und dadurch in größerem zeitlichem Umfang und auf erhöhtem Niveau Chemieunterricht in der Oberstufe erteilt bekommen. Für eine systematische Berücksichtigung als Moderatorvariable (bspw. als Gruppenvariable eines Multigruppenmodells) reicht die vorliegende Stichprobe jedoch nicht aus. Im Rahmen der Ergebnisdarstellung wird an entsprechender Stelle jedoch auf mögliche Unterschiede zwischen Schülerinnen und Schüler, die Chemie als Profil gebendes Fach gewählt haben, und Schülerinnen und Schüler aus anderen Profilen eingegangen. Angaben zur Freude am Chemieunterricht sowie Ergebnisse des Fachtests von Schülerinnen und Schülern, die zu dem jeweiligen Testzeitpunkt nicht am Chemieunterricht teilnahmen, wurden als fehlende Werte kodiert. Die im Abschnitt „Ergebnisse“ dargestellten Modelle wurden auf Basis der Angaben aller Personen unter Berücksichtigung der fehlenden Werte mit dem Full Information Maximum Likelihood (FIML)-Ansatz geschätzt. Die Befragung erfolgte durch instruierte Testleiterinnen und Testleiter mittels Fragebögen während der regulären Unterrichtszeit und dauerte je Messzeitpunkt zwei, zum ersten Messzeitpunkt zwischen zwei und drei Unterrichtsstunden. Alle Schülerinnen und Schüler partizipierten freiwillig und mit dem Einverständnis der Erziehungsberechtigten.

### Instrumente

Um die Entwicklung der Schülerinnen und Schüler möglichst detailliert charakterisieren zu können, wurden in der DoLiS-Studie demographische Daten erfragt und verschiedene etablierte Instrumente zur Erfassung kognitiver sowie motivationaler Personenmerkmale eingesetzt. In dieser Untersuchung wird der Zusammenhang zwischen den drei Personenmerkmalen individuelles Interesse, Freude am Fach und Konzeptwissen unter Berücksichtigung des Alters, des Geschlechts und der kognitiven Grundfähigkeiten analysiert.

Die Freude am Unterrichtsfach Chemie wird in dieser Studie als emotionaler Trait, d. h. als habituelles Erleben der Leistungsemotion Freude während des Unterrichts, konzeptualisiert und mithilfe von drei Items erfasst. Zwei Items („Chemie macht mir Spaß.“, „Ich mag Chemie.“; (Rakoczy et al., 2005), waren auf einer vier-stufigen Likert-Skala (1 = „stimmt gar nicht“, ..., 4 = „stimmt völlig“), das dritte Item („Wenn ich Aufgaben im Chemieunterricht bearbeite, dann mache ich das vor allem weil es mir Spaß macht“; (Ryan & Connell, 1989) auf einer fünf-stufigen Likert-Skala (1 = „stimmt gar nicht“, ..., 5 = „stimmt völlig“) zu beantworten. Die Diskrepanz in der Anzahl der Abstufungen ist darin begründet, dass diese Items (und auch weitere Items) und damit zusammenhängend auch das Antwortformat unverändert aus den originären Instrumenten übernommen wurden. Zur Berechnung von Skalenwerten (bspw. des Skalenmittelwerts) wurden die Werte des fünf-stufigen Items auf Werte im Bereich 1 bis 4 transformiert, um keine Verzerrungen der berechneten Werte zu erzeugen und die Skalenwerte sinnvoll interpretieren zu können. Die interne Konsistenz als Maß für die Reliabilität ist zu jedem der drei Messzeitpunkte ( $\alpha_1 = .92$ ,  $\alpha_2 = .90$ ,  $\alpha_3 = .90$ , s. Tabelle 8.2) als sehr gut anzusehen. Es liegt strikte Messinvarianz über die Zeit vor (s. Tabelle 8.1).

Das individuelle Interesse an der Domäne Chemie wurde distal mithilfe von fünf Items, die die Häufigkeit der außerschulischen Auseinandersetzung mit Chemiekontexten erfragen (Seidel 2003), auf einer fünf-stufigen Likert-Skala (1 = „nie“, ..., 5 = „sehr oft“) erfasst. Außerschulische (Freizeit)Aktivitäten eignen sich insbesondere als Indikator des individuellen Interesses, da sie durch geringere extrinsische Anreize wie Noten beeinflusst sowie durch einen hohen Autonomiegrad gekennzeichnet sind, d. h. das selbstbestimmte Handeln reflektieren ohne bspw. direkten Vorgaben durch Lehrkräften zu unterliegen. Ein Beispielitem der Skala lautet: „Wie oft liest du Bücher, die Themen aus dem Bereich der Chemie behandeln?“. Die Cronbachs  $\alpha$ -Werte deuten auf eine akzeptable interne Konsis-

tenz der Skala hin ( $\alpha_1 = .80$ ,  $\alpha_2 = .83$ ,  $\alpha_3 = .83$ , Tabelle 8.2). Zudem liegt strikte Messinvarianz über die Zeit vor (s. Tabelle 8.1).

Das chemiebezogene Konzeptwissen wurde in der vorliegenden Studie als Verständnis der drei Basiskonzepte Struktur-Eigenschafts-Beziehungen, chemische Reaktion und Energie operationalisiert, das sich durch die erfolgreiche Anwendung von Wissen über diese Konzepte in unterschiedlichen Kontexten manifestiert (bspw. Bransford et al., 2000; Weinert, 2001). Die Fokussierung auf drei Basiskonzepte ist dem Hintergrund des vorliegenden Projekts geschuldet, das aus einer schwedisch-deutschen Kooperation entstanden ist. Durch einen umfassenden Vergleich der Lehrplanvorgaben bilden diese drei Konzepte den gemeinsamen Kern der Vorgaben beider Länder ab. Unter dem hier verwendeten Begriff der Struktur-Eigenschafts-Beziehungen werden jedoch auch Aspekte subsumiert, die in den deutschen bzw. schleswig-holsteinischen Vorgaben dem Konzept Stoff-Teilchen zugeordnet werden, insbesondere Vorstellungen zum Aufbau der Materie oder Bindungsmodelle zur Interpretation von Teilchenaggregationen. Aus inhaltlicher Sicht wurden die beiden Basiskonzepte Stoff-Teilchen und Struktur-Eigenschaftsbeziehungen somit zu einem Konzeptbereich zusammengelegt. Damit geht aus empirischer Sicht jedoch auch eine entsprechende Gewichtung einher, da jeder der drei Konzeptbereiche durch dieselbe Aufgabenanzahl in die Testgestaltung und -auswertung repräsentiert ist.

Das Konzeptwissen der Schülerinnen und Schüler wurde mithilfe eines Fachtests gemessen (s. Höft et al., 2019), der sich aus klassischen Multiple-Choice- (MC) sowie Ordered-Multiple-Choice-Items (OMC) zusammensetzte. OMC-Aufgaben sind analog zu MC-Aufgaben aufgebaut, jedoch entsprechen alle Antwortoptionen theoretisch bzw. empirisch hergeleiteten, hierarchischen Verständnisniveaus (bspw. auf Grundlage einer Learning Progression) des interessierenden Konzepts. Das Ziel von OMC-Aufgaben ist es, das Verständnis der Schülerinnen und Schüler differenzierter erfassen und analysieren zu können, indem die Antwortoptionen entsprechend des angenommenen Verständnisniveaus bewertet und ausgewertet werden. Das in der DoLiS-Studie verwendete Rahmenmodell basiert auf der Metaanalyse von Hadenfeldt, Liu et al. (2014), in welcher ein fünfstufiges Modell des Materieverständnis (1. Alltagsvorstellungen, 2. Hybridvorstellungen, 3. einfache Teilchenvorstellungen, 4. differenzierte Teilchenvorstellungen, 5. Systematische Teilchenvorstellungen) erarbeitet wurde.

Für den Fachtest wurden im Rahmen der DoLiS-Studie insgesamt 114 Items eingesetzt, die zum Teil aus bisherigen Studien übernommen wurden (s. Hadenfeldt et al., 2013; Hadenfeldt et al., 2016; Hadenfeldt, Repenning et al., 2014). Je Jahrgangsstufe wurden, basierend auf einem Anker-Item-Design, 30 Items zur Erfassung des Konzeptwissens genutzt. Von den 30 Items waren wiederum 18 Items Anker-Items und 12 Items jahrgangsspezifisch, sodass eine Verlinkung der Performanz zu den verschiedenen Messzeitpunkten möglich ist. Sowohl die Anker- als auch die jahrgangsspezifischen Items verteilten sich paritätisch auf die drei zentralen Basiskonzepte. Alle eingesetzten Items wurden zuvor von einem Expertenpanel, bestehend aus drei Chemielehrkräften und vier Fachdidaktikerinnen bzw. Fachdidaktikern evaluiert, um Mehrdeutigkeiten und die Möglichkeit von Fehlinterpretationen durch die Probanden zu minimieren sowie um die Zuordnung der Antwortoptionen zu den theoretischen Verständnisniveaus zu bewerten. Der Fachtest wurde mittels eindimensionaler IRT-Analyse unter Verwendung eines generalisierten Multigruppen-Partial-Credit-Modells (multi-group generalized partial credit model; MG-PCM), ausgewertet, wobei die Jahrgangsstufe als Gruppierungsvariable fungierte. Die EAP/PV-Reliabilität indizierte eine gute Reliabilität des Fachtests je Messzeitpunkt (EAP/PV Reliabilität:  $T_1 = .83$ ,  $T_2 = .86$ ,  $T_3 = .81$ ). Hinsichtlich der Itempassung wurde als Itemselektionskriterien der Infit (weighted mean square [WMNSQ]) betrachtet und Items, die typische Ausschlusskriterien nicht erfüllten ( $WMNSQ < 1.3$ ; Wright et al., 1994) von anschließenden Analysen exkludiert ( $n_{T_1} = 2$ ,  $n_{T_2} = 3$ ,  $n_{T_3} = 2$ ). Schließlich wurden die Personenparameter mittels gewichteter Likelihood Schätzung (weighted likelihood estimates; [WLE]) als Maß für das Konzeptverständnis der Schülerinnen und Schüler bestimmt (WLE Reliabilität:  $T_1 = .74$ ,  $T_2 = .75$ ,  $T_3 = .77$ ). Die erzielten IRT Parameter der drei Messzeitpunkte wurden abschließend mithilfe des Haberman-Verfahrens auf eine einheitliche Metrik gebracht (Haberman, 2009).

Als Kovariaten wurden jährlich das Geschlecht (weiblich = 1, männlich = 2), das Alter sowie in



**Tab. 8.1:** Schrittweise Prüfung der Messinvarianz (MI) über die Zeit für die Dimensionen Freude am Fach und individuelles Interesse (Meredith, 1993) anhand der Änderungen  $\Delta$  der Fit-Indizes CFI, RMSEA, SRMR für Samplegrößen  $N > 300$ . Metrische MI ist nicht gegeben bei  $\Delta\text{CFI} \geq .010$ , in Kombination mit  $\Delta\text{RMSEA} \geq .015$ ,  $\Delta\text{SRMR} \geq .030$ , skalare und strikte MI sind nicht gegeben bei  $\Delta\text{CFI} \geq .010$ , bei zeitgleichem Auftreten von  $\Delta\text{RMSEA} \geq .015$ ,  $\Delta\text{SRMR} \geq .010$  (S. 501 Chen, 2007). Da alle berichteten Änderungen des CFI kleiner als .010 sind, kann strikte Messinvarianz angenommen werden. Abkürzungen: IN = individuelles Interesse, ENJ = Freude.

Skala	Messinvarianz	CFI	RMSEA	SRMR	CFI	RMSEA	SRMR
ENJ	konfigural	.998	.021	.023			
ENJ	metrisch	.997	.021	.031	.001	0	.008
ENJ	skalar	.995	.024	.030	.002	.003	.001
ENJ	strikt	.993	.027	.038	.002	.003	.008
IN	konfigural	.986	.026	.035			
IN	metrisch	.980	.029	.047	.006	.003	.012
IN	skalar	.973	.033	.049	.007	.004	.002
IN	strikt	.968	.034	.054	.005	.001	.005

Anmerkungen. IN = individuelles Interesse, ENJ = Freude.

Jahrgangsstufe 9 die kognitiven Grundfähigkeiten der Probanden erfasst. Zur Messung der kognitiven Grundfähigkeiten wurde der Test zu Figurenalogien aus dem Kognitiven Fähigkeitstest für 4.–12. Klassen (KFT 4–12 + R; Heller und Perleth 2000) eingesetzt. Die Reliabilität der Skala ist als gut einzuschätzen ( $\alpha_1 = .82$ ,  $\omega_1 = .80$  [.77, .84]).

#### Analysestrategie

Für die vorliegenden Daten wurden zunächst deskriptive Analysen durchgeführt und Voraussetzungen für die latenten, längsschnittlichen Analysen geprüft, bspw. Reliabilität und skalare Messinvarianz (Meredith, 1993). Zudem wurde mittels konfirmatorischer Faktorenanalyse je Erhebungszeitpunkt getestet, inwiefern sich die beiden Konstrukte Freude und individuelles Interesse empirisch trennen lassen. Dabei zeigte sich innerhalb jedes Messzeitpunkts ein signifikant besserer Modellfit zugunsten einer zweidimensionalen Struktur (T1:  $\Delta\chi^2(1) = 262.774$ ,  $p < .001$ ; T2:  $\Delta\chi^2(1) = 258.697$ ,  $p < .001$ ; T3:  $\Delta\chi^2(1) = 151.188$ ,  $p < .001$ ; Satorra und Bentler, 2010), sodass theoretische und empirische Argumente dafürsprechen, Freude und individuelles Interesse im Rahmen der Auswertung als separate Konstrukte anzunehmen.

Alle längsschnittlichen Strukturgleichungsanalysen wurden mittels MPlus (L. K. Muthén & Muthén, 1998-2015) mit einem robusten Maximum Likelihood (MLR) Schätzer kalkuliert und fehlende Werte unter Verwendung FIML Methode berücksichtigt, sodass alle vorhandenen Informationen in die Berechnungen einfließen. In den längsschnittlichen Analysen wurde skalare Messinvarianz vorausgesetzt (s. Tabelle 8.1) und indikatorspezifische Korrelationen zwischen Residuen erlaubt (correlated uniqueness approach). Zur Beurteilung der Passung der latenten Modelle wurden die Fit-Indizes CFI (comparative fit index), SRMR (standardized root mean square residual) und RMSEA (root mean square error of approximation) berechnet und gängige Passungskriterien angewandt (gute Passung bei Werten von CFI  $> .95$ , SRMR  $< .10$ , RMSEA  $< .08$ ; s. Hu und Bentler, 1999; Schermelleh-Engel et al., 2003).

Die längsschnittliche Analyse der Daten erfolgte in drei konsekutiven Schritten mit zunehmender Komplexität der Modelle. In der Darstellung bzw. Beschreibung werden im Folgenden individuelle Ausprägungen der einzelnen Konstrukte zu einem spezifischen Messzeitpunkt mit einem Kürzel des Konstrukts (IN = individuelles Interesse, ENJ = Freude, ACH = Konzeptwissen) und des Messzeitpunkts (T1, T2, T3) angegeben (also bspw. IN(T1) für das individuelle Interesse zum ersten Messzeitpunkt). Veränderungen der Ausprägung eines Konstrukts zwischen zwei Messzeitpunkten werden mit  $\Delta_{12}$ ,  $\Delta_{23}$  bzw.  $\Delta_{13}$  für Veränderungen zwischen dem ersten und zweiten, dem zweiten und dritten bzw. dem ersten und dritten Messzeitpunkt angegeben. Auch bei spezifischen Pfadkoeffizienten, die im Folgenden bezogen auf die jeweilige Modellklasse erläutert werden, beziehen sich

Indezahlen auf den jeweiligen Messzeitpunkt bzw. Effekte zwischen den jeweiligen Messzeitpunkten.

Im ersten Analyseschritt wurden zur Beantwortung der ersten Forschungsfrage drei latente, univariate Modelle für die Konstrukte individuelles Interesse und Freude sowie zwei latente, univariate Modelle für das Konzeptwissen geschätzt. Zunächst wurde die interindividuelle Stabilität der Personenmerkmale mithilfe konfirmatorischer Faktorenanalysen (CFA) für das individuelle Interesse und die Freude bestimmt, anschließend die längsschnittliche Entwicklung un konditionaler True Individual Change Modelle (TICM) sowie Latent Change Score Modelle (LCSM) analysiert. TICMs sind Reparametrisierungen von LCSMs, weswegen beide Modellierungen die gleichen Modellfit-Werte aufweisen. Die Modellierungen erlauben die Analyse der latenten Mittelwertsverläufe (TICM), die Berechnung von Zusammenhängen zwischen individuellen Veränderungen und interindividuellen Ausprägungen, sowie die Analyse der Stabilität der Veränderung der Personenmerkmale (LCSM).

Im zweiten Analyseschritt wurden zur Beantwortung der zweiten Forschungsfrage drei bivariate LCSMs geschätzt. Abbildung 8.1 zeigt eine generische Darstellung des Modells, in das paarweise zwei der drei interessierenden Konstrukte (X und Y) zu den drei Messzeitpunkten (T1 – T3) eingeführt wurden (dargestellt als X(T1) bis X(T3) für das erste und Y(T1) bis Y(T3) für das zweite Konstrukt; die univariaten LCSMs aus dem ersten Analyseschritt entsprechen analog einer Hälfte der Abbildung, ohne Querspfade zu einem zweiten Konstrukt). LCSMs erlauben durch die Berechnung von Regressionen (analog zu den univariaten Modellen im ersten Analyseschritt) die Untersuchung

- von Zusammenhängen zwischen interindividuellen Ausprägungen und individuellen Veränderungen (bspw. den Effekt von X(T1) auf die Veränderung von X zwischen T1 und T2, ( $\Delta_{12}X$ ), beschrieben als  $\beta_{X12}$  in Abbildung 8.1) und
- der Stabilität der Veränderung der Personenmerkmale (bspw. den Effekt der Veränderung von X zwischen T1 und T2 auf die Veränderung von X zwischen T2 und T3, beschrieben als  $\phi_X$ ),

mittels Berechnung von Korrelationen die Untersuchung

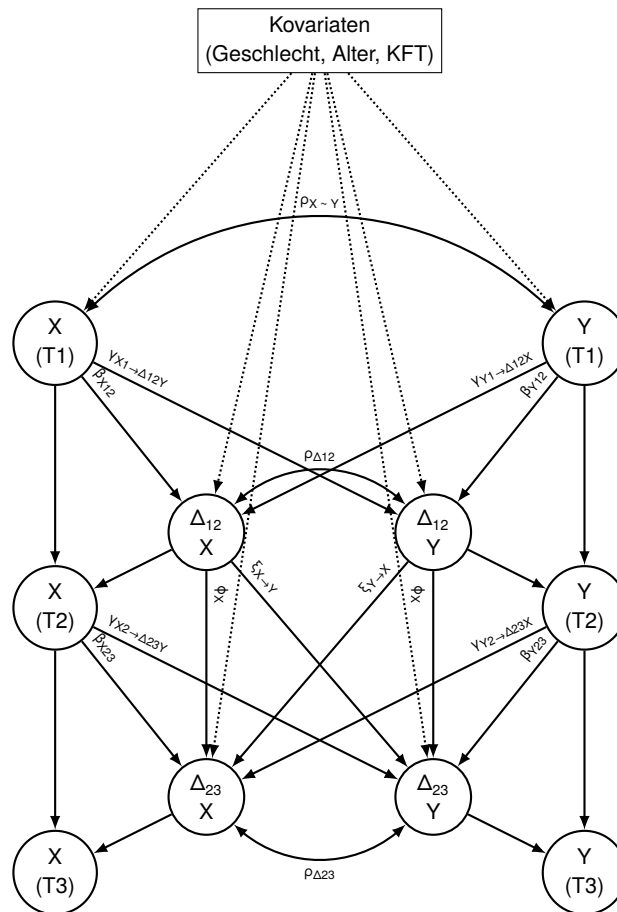
- von Zusammenhängen zwischen zeitgleichen Merkmalsausprägungen unterschiedlicher Konstrukte ( $\rho_{XY}$ ) und
- von Zusammenhängen zwischen Merkmalsänderungen unterschiedlicher Konstrukte ( $\rho_{\Delta 12}$ ,  $\rho_{\Delta 23}$ ),

sowie durch die Schätzung von kreuzverzögerten (cross-lagged) Effekten die Analyse

- von Zusammenhängen zwischen interindividueller Ausprägung des einen und den individuellen Veränderungen des zweiten Personenmerkmals (also bspw. den Effekt der Ausprägung von X(T1) auf die Veränderung von Y zwischen T1 und T2, beschrieben als  $\gamma_{X1 \rightarrow \Delta 12Y}$ ; gleichermaßen den Effekt der Ausprägung von X(T2) auf die Veränderung von Y zwischen T2 und T3 sowie die umgekehrten Effekte der Ausprägung von Y zu T1 bzw. T2 auf die Veränderungen von X zwischen T1 und T2 bzw. T2 und T3) sowie
- von Zusammenhängen zwischen individuellen Veränderungen (also den Effekt der Veränderung von X zwischen T1 und T2 auf die Veränderung von Y zwischen T2 und T3, beschrieben als  $\xi_{X \rightarrow Y}$  bzw. umgekehrt die Veränderung von Y auf die zeitversetzte Veränderung von X) (Grimm et al., 2012).

Zur Beantwortung der dritten Forschungsfrage wurde im dritten Analyseschritt ein trivariates LCSM geschätzt, in welches das zu den drei Messzeitpunkten erfasste individuelle Interesse, die Freude und das Konzeptwissen aufgenommen wurden, um die Wechselwirkungen zwischen allen drei Konstrukten über die Zeit zu analysieren. Die Modelldarstellung und Benennung der Variablen und Pfade ist analog zu Abbildung 8.1, jedoch um ein drittes Konstrukt erweitert.

Zur Simplifizierung der Modelle (Parsimonitätsprinzip) wurde mittels Log-Likelihood-(LL-)Tests zunächst geprüft, ob cross-lagged Pfade (bspw.  $\gamma_{X1 \rightarrow \Delta 12Y} = \gamma_{X2 \rightarrow \Delta 23Y}$ ) gleicher Wirkrichtung über die Zeit stabil sind (bspw. ob der Pfad vom individuellen Interesse zu T1 auf die Änderung des Konzeptwissens zu T2,  $\gamma_{I_{T1} \rightarrow \Delta 12ACH}$ , einen nahezu identischen Pfadkoeffizienten aufweist wie der



**Abb. 8.1:** Generisches bivariates Latent Change Score Modell der latenten Variablen X und Y für die drei Messzeitpunkte T1 bis T3; Latente Differenzwerte zwischen zwei Messzeitpunkten sind durch  $\Delta$ s gekennzeichnet; Regressionspfade, die kein Label aufweisen, sind auf den Wert 1 fixiert; Entsprechend Grimm et al. (2012) wurden autoregressive Pfade zwischen latenten Differenzwerten modelliert.

Pfad vom individuellen Interesse zu T2 auf die Änderung des Konzeptwissens zu T3,  $\gamma_{IN2 \rightarrow \Delta 23 ACH}$ ) und unter Bestätigung dieser Annahme für weitere Analysen als zeitinvariant gesetzt. Zudem wurde in einem weiteren Analyseschritt mittels LL-Tests die Wirkrichtung zwischen den jeweils untersuchten Konstrukten analysiert, indem verglichen wurde, ob die Höhe der Regressionskoeffizienten der cross-lagged Effekte in beide Wirkrichtungen gleich hoch waren.

Sowohl in den bivariaten als auch in dem tivariaten LCSM wurde für die Kovariaten Alter, Gender und kognitive Grundfähigkeit kontrolliert, da aus bisherigen Studien bekannt ist, dass diese Personenmerkmale die Ausprägungen und die Veränderungen von kognitiven und motivationalen Variablen beeinflussen können.

Im Folgenden werden zur Interpretation von Effekten neben Signifikanzwerten stichprobengrößenunabhängige Effektstärken, d. h. für Mittelwertsunterschiede die Effektstärke Cohens  $d$  (nach Cohen, 1988 kann die Effektstärke  $d$  im Einstichprobenfall für  $d \approx 0.14$  als kleiner Effekt, für  $d \approx 0.35$  als mittlerer Effekt und für  $d \approx 0.57$  als großer Effekt, beim Vergleich zweier Gruppen für  $d > 0.2$  als kleiner Effekt, für  $d > 0.5$  als mittlerer Effekt und für  $d > 0.8$  als großer Effekt interpretiert werden), für Zusammenhänge zwischen intervallskalierten Variablen vollständig standardisierte Korrelations- bzw. Regressionskoeffizienten (Beta-Gewichte) berichtet.

**Tab. 8.2:** Deskriptive Befunde, Reliabilität und Korrelationen der analysierten Konstrukte.

Jhg.	Skala	n	M	SD	Reliabilität <sup>a</sup>	Ω [95% CI]	Korrelationen	
							ENJ	IN
9	ENJ <sup>b</sup>	491	2.57	1.00	.92	.92 [.90, .93]		
9	IN	502	2.10	1.00	.80	.81 [.77, .84]	.53***	
9	ACH	503	0.27	0.60	.74		.25***	.21***
10	ENJ <sup>b</sup>	428	2.45	0.98	.90	.90 [.87, .92]		
10	IN	516	1.99	1.00	.83	.83 [.80, .85]	.61***	
10	ACH	456	0.33	0.54	.75		.34***	.29***
11	ENJ <sup>b</sup>	394	2.36	0.97	.90	.91 [.88, .92]		
11	IN	478	1.98	1.05	.83	.84 [.81, .86]	.67***	
11	ACH	409	0.67	0.71	.77		.38***	.45***

Anmerkungen. IN = individuelles Interesse, ENJ = Freude, ACH = Konzeptwissen. \*\*\*p < .001, \*\*p < .01, \*p < .05.

a Den Skalen entsprechend Angabe des Cronbachs  $\alpha$ -Werts (ENJ, IN) bzw. der WLE-Reliabilität (ACH).

b Das einzelne fünf-stufige Item dieser Skala wurde vor der Berechnung deskriptiver Statistiken auf Werte im Bereich 1 bis 4 transformiert, so dass auch hier der Skalenmittelwert bei 2.5 liegt.

## 8.4 Ergebnisse

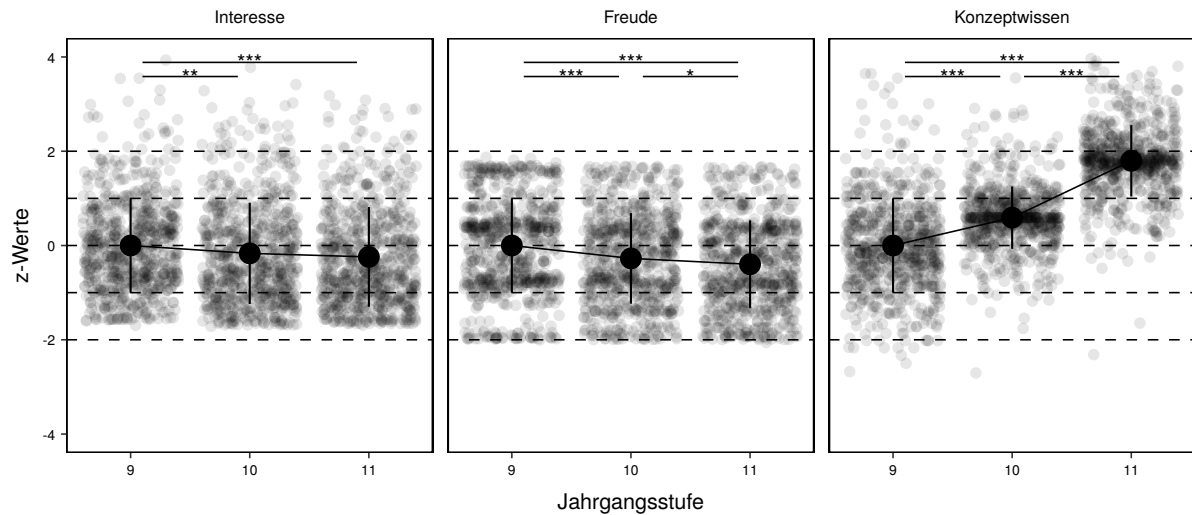
### Deskriptive Befunde

Die mittlere Ausprägung des individuellen Interesses liegt zu allen drei Messzeitpunkten signifikant unter dem theoretischen Skalenmittelwert von 3 (T1:  $t(501) = -27.50$ ,  $p < .001$ ,  $d = 1.74$ ; T2:  $t(515) = -30.68$ ,  $p < .001$ ,  $d = 1.91$ ; T3:  $t(477) = -28.29$ ,  $p < .001$ ,  $d = 1.83$ ) und nimmt mit steigender Jahrgangsstufe ab (s. Tabelle 8.2). Die mittlere Ausprägung der Freude entspricht in Jahrgangsstufe 9 dem Skalenmittelwert von 2.5, nimmt in den beiden folgenden Jahren ab und liegt in Jahrgangstuf 11 signifikant unter dem Skalenmittelwert (T1:  $t(490) = 1.80$ ,  $p = .073$ ,  $d = 0.11$ ; T2:  $t(427) = -1.2$ ,  $p = .226$ ,  $d = 0.08$ ; T3:  $t(393) = -3.22$ ,  $p < .01$ ,  $d = 0.23$ ). Das Konzeptwissen steigt hingegen kontinuierlich über die Jahrgangsstufen an. Effektgrößen dieser zeitlichen Veränderungen finden sich im nachfolgenden Abschnitt auf Basis univariater TICMs. Die Korrelation zwischen den beiden motivationalen Maßen Freude und individuelles Interesse ist je Messzeitpunkt bedeutsam ( $r > .53$ ,  $p < .001$ ) und ist zudem deutlich größer als die jeweiligen Korrelation mit dem Konzeptwissen (Freude:  $.25 < r < .28$ ; Interesse:  $.21 < r < .45$ ). Die Interkorrelationen zwischen allen Konstrukten nehmen mit steigender Jahrgangsstufe zu. Während somit über die Jahrgangsstufen die Ausprägungen der beiden motivationalen Variablen sinken und das Konzeptwissen ansteigt, sind die wechselseitigen Zusammenhänge zwischen den drei Personenmerkmalen über die Jahrgangsstufen also zunehmend stärker ausgeprägt.

### Analyseschritt 1: Univariate Befunde

Mithilfe univariater Analysen wurden die Stabilitäten der Merkmalsausprägung, der Mittelwertverlauf sowie die Veränderung zwischen zwei Messzeitpunkten je Konstrukt analysiert. Die Stabilitäten des individuellen Interesses und der Freude wurden mittels CFA bei strikter Invarianz als latente Korrelationen berechnet. Die Korrelationskoeffizienten deuten für das individuelle Interesse ( $\rho_{12} = .68$ ,  $\rho_{23} = .73$ ,  $\rho_{13} = .52$ ) und die Freude ( $\rho_{12} = .63$ ,  $\rho_{23} = .73$ ,  $\rho_{13} = .53$ ) auf eine hohe Stabilität der interindividuellen Unterschiede hin. Dies gilt auch für die Einschätzung der Stabilität des Konzeptwissens über die Zeit. Da für das Konzeptwissen auf den WLE-score der IRT-Skalierung zurückgegriffen wird und mit der vorgenommenen Verlinkung bereits eine messzeitpunktübergreifende, einheitliche Metrik vorliegt, werden die Zusammenhänge nicht als single indicator model latent geschätzt, sondern als manifeste Korrelationen berichtet (Widaman et al., 2010). Im Ergebnis sind für das Konzeptwissen ( $r_{12} = .74$ ,  $r_{23} = .72$ ,  $r_{13} = .67$ ) höhere Stabilitäten als für das individuelle Interesse und die Freude zu verzeichnen, insbesondere bezogen auf den Zusammenhang über einen

Zeitraum von zwei Jahren (also zwischen den Werten in Jahrgangsstufe 9 und 11).



**Abb. 8.2:** Darstellung der mittels True Individual Change Score Modellen berechneten Personenparameter des individuellen Interesses, der Freude und des Konzeptwissens sowie deren Mittelwertsverläufe mit Standardabweichung.

Die Entwicklung der drei Merkmale wurde mittels TICMs und LCSMs analysiert (s. Abbildung 8.2). Die Fit-Werte dieser Modelle deuten auf eine gute Modellpassung hin (CFIs  $\geq .973$ , SRMRs  $\geq .049$ , RMSEAs  $\geq .033$ ; s. Tabelle 8.3). Hinsichtlich des individuellen Interesses kann eine kleine, si-

**Tab. 8.3:** Mittelwerte (M), Regressionskoeffizienten und Fit-Indizes der univariaten True Individual Change Score und Latent Change Score Modelle (LCSM). Berichtet werden vollständig standardisierte Korrelationen,  $\beta$ -Regressionskoeffizienten und in Klammern deren Standardfehler.

	IN	ENJ	ACH
TICMs			
M (T1)	2.073	2.324	0.277
M (T2)	1.986	2.146	0.315
M (T3)	1.948	2.065	0.642
LCSMs			
$\Delta_{12}X \leftarrow X$ (T1)	-.363 (.072)***	-.465 (.035)***	-.496 (.039)***
$\Delta_{23}X \leftarrow \Delta_{12}X$	-.058 (.108)	-.126 (.095)	-.281 (.074)***
$\Delta_{23}X \leftarrow X$ (T2)	-.321 (.074)***	-.314 (.055)***	-.103 (.065)
Fit-Indizes			
CFI	.973	.995	1.000
RMSEA	.033	.024	0.000
SRMR	.049	.031	0.000

Anmerkungen. IN = individuelles Interesse, ENJ = Freude, ACH = Konzeptwissen, X = interessierendes Konstrukt. \*\*\*p < .001, \*\*p < .01, \*p < .05.

gnifikante Abnahme des außerschulischen Chemieinteresses zwischen Jahrgangsstufe 9 und 11, die hauptsächlich auf eine Interessenabnahme zwischen Jahrgang 9 und 10 zurückzuführen ist ( $d_{12} = .16$ ,  $d_{23} = .08$ ,  $d_{13} = .23$ ), beobachtet werden (s. Abbildung 8.2, TICM). Die Abnahme des individuellen Interesses  $\Delta_{12}(\text{IN})$  und  $\Delta_{23}(\text{IN})$  hängt negativ mit dem initialen Interesse bzw. dem Interesse der Schülerinnen und Schüler in Jahrgangsstufe 10 zusammen ( $\beta_{12} = -.36$ ,  $p < .001$ ;  $\beta_{23} = -.32$ ,  $p < .001$ ). Die Stabilität der Interessenänderung ist nicht signifikant ( $\phi = -.11$ ,  $p = .590$ ). Ein höheres individuelles Interesse hängt also mit einer höheren Abnahme des Interesses über die Jahrgangsstufen zusammen. Jedoch kann die Interessenabnahme  $\Delta_{12}(\text{IN})$  keine zusätzliche Varianz der folgenden Interessenabnahme  $\Delta_{23}(\text{IN})$  erklären, es findet also keine beschleunigte Abnahme statt. Hinsichtlich der Freude kann eine kontinuierliche, signifikante Abnahme je konsekutiver

Jahrgangsstufe konstatiert werden, die in einer moderaten Abnahme der Freude zwischen den Jahrgangsstufen 9 und 11 ( $d_{12} = .25$ ,  $d_{23} = .14$ ,  $d_{13} = .40$ ) kumuliert (s. Abbildung 8.2). Die Abnahme der Freude  $\Delta_{12}(\text{EN})$  und  $\Delta_{23}(\text{EN})$  hängt negativ mit der initialen Freude der Schülerinnen und Schüler in Jahrgangsstufe 9, respektive der Freude in Jahrgangsstufe 10 zusammen ( $\beta_{12} = -.47$ ,  $p < .001$ ;  $\beta_{23} = -.31$ ,  $p < .001$ ). Auch hier hängt eine höhere Freude am Chemieunterricht in Jahrgang 9 mit einer höheren Abnahme der Freude über die Jahrgangsstufen zusammen. Die Stabilität der Änderung der Freude ist, analog zum individuellen Interesse, nicht signifikant ( $\phi = -.13$ ,  $p = .186$ ). Hinsichtlich des Konzeptwissens kann eine kontinuierliche, positive Zunahme des Konzeptwissens verzeichnet werden, welche hauptsächlich auf einen Wissenserwerb zwischen Jahrgangsstufe 10 und 11 beruht ( $d_{12} = .07$ ,  $d_{23} = .62$ ,  $d_{13} = .62$ ) (s. Abbildung 8.2). Die Zunahme des Konzeptwissens  $\Delta_{12}(\text{ACH})$  hängt negativ mit dem vorherigen Konzeptwissen der Schülerinnen und Schüler zusammen ( $\beta_{12} = -.50$ ,  $p < .001$ ;  $\beta_{23} = -.10$ ,  $p = .114$ ). Ein höheres Konzeptwissen in Jahrgangsstufe 9 hängt demnach mit einer geringeren Zunahme des Konzeptwissens über von Jahrgang 9 zu 10 zusammen. Demnach scheinen vor allem leistungsschwächere Schülerinnen und Schüler beim Übergang von Jahrgang 9 zu 10 aufzuholen, wobei in der Gesamtstichprobe nur ein geringer Lernzuwachs zu verzeichnen ist. Dennoch scheint sich die Leistungsschere ein wenig zu schließen, was auch mit einer verringerten Varianz der Leistung in Jahrgang 10 einhergeht (vgl. Tabelle 8.2). Dieser Befund scheint beim Übergang von Jahrgang 10 zu 11 jedoch nicht mehr vorzuliegen. Hier findet ein deutlich höherer Lernzuwachs statt, dessen Höhe zudem nicht mehr mit dem individuellen vorherigen Konzeptwissen in Jahrgang 10 zusammenhängt. Die Stabilität des Wissenserwerbs ist gering ( $\phi = -.28$ ,  $p < .001$ ) und deutet auf unterschiedliche Leistungszuwächse zwischen den beiden Jahrgangübergängen bei leistungstärkeren und leistungsschwächeren Schülerinnen und Schülern hin.

Zusammenfassend lässt sich sagen, dass sich von Jahrgangsstufe 9 bis 11 ein kleiner Abfall des individuellen Interesses, ein moderater Abfall der Freude und gleichzeitig eine starke Zunahme des Konzeptwissens vollzieht. Es liegen demnach gegenläufige Trends zwischen den motivationalen Variablen und der kognitiven Variable vor.

## Analyseschritt 2: Bivariate Zusammenhänge

In dem zweiten Analyseschritt wurden mittels LCSMs die bivariaten Zusammenhänge zwischen dem individuellen Interesse und der Freude, dem individuellen Interesse und dem Konzeptwissen sowie der Freude und dem Konzeptwissen analysiert (s. Abbildung 8.1 und Tabelle 8.4). Die Ergebnisse der LL-Tests indizieren, dass die gleichgerichteten cross-lagged Effekte in allen drei Modellen zeitinvariant sind (Modell BI 1 (IN, ACH):  $\Delta\chi^2(2) = 1.12$ ,  $p = .571$ ; Modell BI 2 (ENJ, ACH):  $\Delta\chi^2(2) = 3.07$ ,  $p = .215$ ; Modell BI 3 (IN, ENJ):  $\Delta\chi^2(2) = .14$ ,  $p = .933$ ), weshalb im Folgenden nur die Ergebnisse der sparsameren Modelle mit zeitinvarianten (d. h. bezogen auf den Pfadkoeffizienten identischen) cross-lagged Pfaden beschrieben werden.

Zwischen dem individuellen Interesse und dem Konzeptwissen (s. Tabelle 8.4 Modell BI 1) liegen kleine Korrelationen zum ersten Messzeitpunkt ( $\rho = .20$ ,  $p < .001$ ) und zwischen den jeweiligen Änderungen von Jahrgangsstufe 10 und 11 ( $\rho_{\Delta 23} = .19$ ,  $p < .001$ ) vor. Zudem sind sowohl die cross-lagged Effekte von dem individuellen Interesse auf die Veränderung des Konzeptwissens ( $\gamma_s > .13$ ,  $p_s < .05$ ) als auch die Effekte vom Konzeptwissen auf die Veränderung des individuellen Interesses ( $\gamma_s > .22$ ,  $p_s < .001$ ) positiv und bedeutsam. Die Höhe der Regressionskoeffizienten unterscheidet sich nicht bedeutsam ( $\Delta\chi^2(1) = 2.47$ ,  $p_s = .116$ ), sodass von gleichstarken Wirkungen gesprochen werden kann. Zwischen den Veränderungen der Merkmalsausprägungen liegen keine signifikanten Effekte  $\xi$  vor. Somit liegen durchgehend reziproke Zusammenhänge zwischen dem individuellen Interesse und dem Konzeptwissen vor, die in vergleichbarem Umfang die Veränderung des jeweils anderen Konstrukts positiv beeinflussen. Die Ausprägungen der jeweiligen Veränderungen hängen jedoch nicht wechselseitig zusammen.

Zwischen der Freude und dem Konzeptwissen (s. Tabelle 8.4 Modell BI 2) liegen kleine bis moderate, positive Korrelationen zwischen den initialen Merkmalsausprägungen und den jeweiligen

**Tab. 8.4:** Ergebnisse der bi- (Modelle BI) und trivariaten Latent Change Score Modelle (Modelle TRI). Das Modell TRI 2 basiert auf einem reduzierten Datensatz, aus dem Informationen über Schülerinnen und Schüler mit Chemie als Profil gebendes Fach exkludiert wurden. Berichtet werden vollständig standardisierte Korrelationen,  $\beta$ -Regressionskoeffizienten und in Klammern deren Standardfehler.

	Vollständiges Sample			Reduziertes Sample	
	Modell BI 1: (IN, ACH)	Modell BI 2: (ENJ, ACH)	Modell BI 3: (IN, ENJ)	Modell TRI 1: (IN, ENJ, ACH)	Modell TRI 2: (IN, ENJ, ACH)
Korrelationen ( $\rho_{XY}$ , $\rho_{\Delta 12}$ , $\rho_{\Delta 23}$ )					
IN (T1) ~ ACH (T1)	.199 (.051)***			.213 (.050)***	.195 (.054)***
$\Delta_{12}$ IN ~ $\Delta_{12}$ ACH	.062 (.066)			.076 (.067)	.024 (.072)
$\Delta_{23}$ IN ~ $\Delta_{23}$ ACH	.188 (.076)*			.193 (.079)*	.230 (.080)**
ENJ (T1) ~ ACH (T1)		.242 (.039)***		.253 (.039)***	.212 (.044)***
$\Delta_{12}$ ENJ ~ $\Delta_{12}$ ACH		.191 (.064)**		.183 (.066)**	.154 (.071)*
$\Delta_{23}$ ENJ ~ $\Delta_{23}$ ACH		.175 (.082)*		.152 (.080)	.184 (.087)*
IN (T1) ~ ENJ (T1)			.627 (.036)***	.627 (.036)***	.601 (.040)***
$\Delta_{12}$ IN ~ $\Delta_{12}$ ENJ			.594 (.053)***	.577 (.056)***	.553 (.060)***
$\Delta_{23}$ IN ~ $\Delta_{23}$ ENJ			.593 (.066)***	.590 (.066)***	.614 (.068)***
cross-lagged Effekte (Veränderung $\leftarrow$ Ausprägung; $Y_{X1 \rightarrow \Delta 12Y}$ , $Y_{X2 \rightarrow \Delta 23Y}$ , $Y_{Y1 \rightarrow \Delta 12X}$ , $Y_{Y2 \rightarrow \Delta 23X}$ )					
$\Delta_{12}$ IN $\leftarrow$ ACH (T1)	.216 (.047)***			.210 (.046)***	.185 (.050)***
$\Delta_{23}$ IN $\leftarrow$ ACH (T2)	.231 (.055)***			.204 (.048)***	.169 (.048)***
$\Delta_{12}$ ACH $\leftarrow$ IN (T1)	.150 (.057)**			.075 (.068)	.010 (.071)
$\Delta_{23}$ ACH $\leftarrow$ IN (T2)	.132 (.052)*			.065 (.060)	.008 (.060)
$\Delta_{12}$ ENJ $\leftarrow$ ACH (T1)		.148 (.041)***		.142 (.040)***	.111 (.042)**
$\Delta_{23}$ ENJ $\leftarrow$ ACH (T2)		.160 (.044)***		.157 (.043)***	.122 (.044)**
$\Delta_{12}$ ACH $\leftarrow$ ENJ (T1)		.196 (.042)***		.132 (.060)*	.112 (.062)
$\Delta_{23}$ ACH $\leftarrow$ ENJ (T2)		.160 (.036)***		.106 (.049)*	.092 (.051)
$\Delta_{12}$ IN $\leftarrow$ ENJ (T1)			.285 (.067)***	.215 (.068)**	.169 (.071)**
$\Delta_{23}$ IN $\leftarrow$ ENJ (T2)			.289 (.071)***	.217 (.069)**	.163 (.068)**
$\Delta_{12}$ ENJ $\leftarrow$ IN (T1)			.069 (.065)	.091 (.061)	.055 (.064)
$\Delta_{23}$ ENJ $\leftarrow$ IN (T2)			.084 (.081)	.112 (.076)	.065 (.077)
cross-lagged Effekte (Veränderung $\leftarrow$ Veränderung; $\xi_{X \rightarrow Y}$ , $\xi_{Y \rightarrow X}$ )					
$\Delta_{23}$ IN $\leftarrow$ $\Delta_{12}$ ACH	-.045 (.074)			-.031 (.077)	-.053 (.082)
$\Delta_{23}$ ACH $\leftarrow$ $\Delta_{12}$ IN	.087 (.091)			.174 (.099)	.165 (.110)
$\Delta_{23}$ ENJ $\leftarrow$ $\Delta_{12}$ ACH		.003 (.077)		-.005 (.077)	-.010 (.088)
$\Delta_{23}$ ACH $\leftarrow$ $\Delta_{12}$ ENJ		-.099 (.096)		-.178 (.102)	-.159 (.111)
$\Delta_{23}$ IN $\leftarrow$ $\Delta_{12}$ ENJ			-.095 (.117)	-.034 (.078)	-.001 (.082)
$\Delta_{23}$ ENJ $\leftarrow$ $\Delta_{12}$ IN			-.019 (.113)	-.064 (.074)	-.097 (.075)
Fit-Indizes					
CFI	.995	.992	.964	.963	.963
RMSEA	.014	.024	.032	.031	.029
SRMR	.027	.028	.046	.045	.047

*Anmerkungen.* Angegebene Signifikanzniveaus entsprechen: Abkürzungen: IN = individuelles Interesse, ENJ = Freude, ACH = Konzeptwissen. Für die Benennung von Koeffizienten vgl. Abbildung 8.1. \*\*\*p < .001, \*\*p < .01, \*p < .05.

Merkmalsveränderungen vor ( $\rho_s > .18$ ,  $p_s < .05$ ). Die cross-lagged Effekte von der Freude auf die Veränderung des Konzeptwissens ( $\gamma_s = .16$ ,  $p_s < .001$ ) sowie von dem Konzeptwissen auf die Veränderung der Freude ( $\gamma_s > .15$ ,  $p_s < .001$ ) sind positiv und signifikant. Die Höhe der Regressionskoeffizienten unterscheidet sich nicht bedeutsam ( $\Delta\chi^2(1) = 2.00$ ,  $p = .157$ ), sodass von gleichstarken Wirkungen gesprochen werden kann. Zwischen den Veränderungen der Merkmalsausprägungen liegen keine signifikanten Effekte  $\xi$  vor. Entsprechend lässt sich das Befundmuster analog zum Befundmuster zwischen individuellem Interesse und Konzeptwissen zusammenfassen: durchgehend reziproke Zusammenhänge zwischen beiden Konstrukten, die im vergleichbaren Umfang die Veränderung des jeweils anderen Konstrukts positiv beeinflussen, deren Änderungsraten jedoch keinen Zusammenhang aufweisen.

Zwischen dem individuellen Interesse und der Freude (s. Tabelle 8.4 Modell BI 3) liegen hohe Korrelationen zwischen den initialen Merkmalsausprägungen und den jeweiligen Merkmalsveränderungen vor ( $\rho_s > .60$ ,  $p_s < .001$ ). Weiterhin sind die cross-lagged Effekte von der Freude auf die Veränderung des individuellen Interesses moderat, positiv und signifikant ( $\gamma_s = .29$ ,  $p_s < .001$ ), nicht jedoch die Effekte des individuellen Interesses auf die Veränderung der Freude. Dementsprechend unterscheidet sich die Höhe der Regressionskoeffizienten bedeutsam ( $\Delta\chi^2(1) = 1.096$ ,  $p = .295$ ), sodass von einseitig gerichteten Effekten von der Freude auf das individuelle Interesse gesprochen werden kann. Zwischen den Veränderungen der Merkmalsausprägungen liegen keine signifikanten Effekte  $\xi$  vor. Insgesamt zeigt sich somit ein durchgehender Effekt der Freude auf die Entwicklung des individuellen Interesses, nicht jedoch umgekehrt. Auch hier gibt es keinen Zusammenhang zwischen den Änderungsverläufen beider Konstrukte.

### Analyseschritt 3: Trivariate Zusammenhänge

Zur Beantwortung der dritten Forschungsfrage wurden trivariate LCSMs spezifiziert, in die die Angaben zu allen drei Konstrukten eingingen, um die gleichzeitigen Wechselwirkungen zu analysieren. Da der LL-Test ( $\Delta\chi^2(6) = 5.89$ ,  $p = .436$ ) ebenso wie in den bivariaten Modellen auf zeitinvariante cross-lagged Effekte hinweist, werden im Anschluss die Ergebnisse beschrieben, in denen alle gleichgerichteten cross-lagged Pfade gleichgesetzt wurden. Die Ergebnisse des trivariaten Modells präzisieren das Befundmuster der bivariaten Modelle (s. Tabelle 8.4 Modell TRI 1), zeigen jedoch auch Abweichungen. So sind die Korrelationen zwischen den Veränderung ( $\Delta_{23}$  und  $\Delta_{23}$ ) der Freude bzw. des individuellen Interesses und des Konzeptwissens nicht mehr bedeutsam. Gleichzeitig liegen weiterhin positive Korrelationen zwischen den drei latenten Faktoren zum ersten Messzeitpunkt und zwischen den Veränderungen ( $\Delta_{23}$ ) des individuellen Interesses und des Konzeptwissens, den Veränderungen ( $\Delta_{12}$  und  $\Delta_{23}$ ) der Freude und des Konzeptwissens sowie den Veränderungen ( $\Delta_{12}$  und  $\Delta_{23}$ ) des individuellen Interesses und der Freude vor. Analog zu den bivariaten Modellen deutet dies darauf hin, dass die Änderungsverläufe der beiden motivationalen Variablen und die des Konzeptwissens nicht gleichsinnig über die Zeit verlaufen. Insbesondere die unterschiedlichen Zuwachsraten beim Konzeptwissen zwischen den Übergängen der Jahrgangsstufen 9 und 10 bzw. 10 und 11 bestärken diesen Eindruck.

Hinsichtlich der gerichteten Zusammenhänge sind die cross-lagged Pfade des individuellen Interesses auf die Änderungen des Konzeptwissens bzw. auf die Änderungen der Freude im Vergleich zu den bivariaten Modellen nicht mehr bedeutsam. Bedeutsame cross-lagged Effekte liegen beidseitig zwischen der Freude und dem Konzeptwissen, sowie einseitig von der Freude bzw. dem Konzeptwissen auf das individuelle Interesse vor. Die Höhe der Regressionskoeffizienten zwischen den cross-lagged Effekten zwischen der Freude und dem Konzeptwissen differieren nicht bedeutsam ( $\Delta\chi^2(2) = 3.47$ ,  $p = .176$ ). Die cross-lagged Effekte zwischen den Veränderungen der Konstrukte sind weiterhin allesamt nicht bedeutsam. Hinsichtlich des Zusammenspiels der drei Konstrukte deuten diese Befunde darauf hin, dass die Veränderung des individuellen Interesses über die Zeit eher vom Konzeptwissen und der Freude beeinflusst wird als dass das individuelle Interesse die beiden anderen Konstrukte in ihrer Entwicklung bestimmt.

Der Einfluss der Kovariaten äußert sich in einem positiven Effekt der kognitiven Grundfähigkeiten



auf die initiale Freude ( $\beta_1 = .11$ ,  $p = .015$ ), das initiale Konzeptwissen ( $\beta_1 = .36$ ,  $p < .001$ ) und die Veränderung des Konzeptwissens von Jahrgangsstufe 10 zu 11 ( $\beta_{\Delta 23} = .36$ ,  $p < .001$ ). Das Geschlecht (codiert als 1 – weiblich, 2 – männlich) ist positiv prädiktiv für das initiale Konzeptwissen ( $\beta_1 = .11$ ,  $p = .004$ ), sowie die Veränderung des individuellen Interesses ( $\beta_{\Delta 23} = .18$ ,  $p = .002$ ) und Freude ( $\beta_{\Delta 23} = .15$ ,  $p = .007$ ) von Jahrgangsstufe 10 zu 11.

Zum ersten Messzeitpunkt, d. h. am Ende der 9. Jahrgangsstufe, haben die teilnehmenden Schülerinnen und Schüler ihr Profil für die Oberstufe gewählt. Die Stichprobengrößen der Teilgruppen sind allerdings zu klein, um systematische Vergleiche der Zusammenhänge zwischen unterschiedlichen Profilen oder zwischen Schülerinnen und Schülern mit Chemie als Profil gebendem Fach und anderen Profilmächern durchzuführen (bspw. in Form eines Multigruppenmodells). Da anzunehmen ist, dass alle drei im Rahmen der vorliegenden Auswertung berücksichtigten Variablen mit der Kurswahlentscheidung der Schülerinnen und Schüler im Zusammenhang stehen (Nagy et al., 2007; Taskinen et al., 2013), wurde das obige Modell nochmals auf Basis der größten Teilgruppe geschätzt, d. h. mit den Angaben der Schülerinnen und Schüler, die Chemie in der Oberstufe belegen, aber nicht als Profil gebendes Fach gewählt haben. Bei gleichbleibend guter Modellpassung (CFI = .963, RMSEA = .029, SRMR = .047) sind nahezu keine Änderungen des Befundmusters zu verzeichnen (s. Tabelle 8.4 Modell TRI 2). Einzig die beiden cross-lagged Pfade der Freude auf die Änderungen des Konzeptwissens werden nicht mehr signifikant ( $\gamma_{\Delta 12} = -.11$ ,  $p_{12} = .068$ ;  $\gamma_{\Delta 23} = -.09$ ,  $p_{23} = .065$ ). In der obigen Auswertung auf Basis der Gesamtstichprobe wurde das Signifikanzniveau jedoch auch nur knapp erreicht ( $\gamma_{\Delta 12} = -.13$ ,  $p_{12} = .029$ ;  $\gamma_{\Delta 23} = -.11$ ,  $p_{23} = .031$ ), sodass das Ausbleiben dieser beiden Effekte nicht überinterpretiert werden sollte und zum Teil sicherlich auch auf die geringere Power durch die geringere Stichprobengröße zurückzuführen ist. Die Höhe der Regressionskoeffizienten ist in beiden Analysen nahezu identisch. Diese Befunde reichen jedoch nicht aus, um unterschiedliche Zusammenhangsmuster in den einzelnen Teilgruppen der Schülerinnen und Schüler bzgl. Ihrer Kurswahlentscheidung bezogen auf das Fach Chemie auszuschließen. Es lässt sich aber zumindest annehmen, dass die in der obigen Analyse aufgedeckten Zusammenhänge nicht in erster Linie auf die besonders interessierten und besonders leistungsstarken Schülerinnen und Schüler (d. h. diejenigen mit Chemie als Profil gebendem Fach) zurückzuführen sind.

## 8.5 Diskussion

In der vorliegenden Studie wurde der Zusammenhang zwischen den drei Konstrukten individuelles Interesse, Freude am Fach und Konzeptwissen und deren Wechselwirkungen über die Zeit untersucht. Ziel dieser Untersuchung war es, durch ein besseres Verständnis der Entwicklung des individuellen Interesses, der Freude am Fach und des Konzeptwissens sowie deren Wechselwirkungen potenzielle Ansatzpunkte für eine Förderung dieser drei Personenmerkmale zu identifizieren. Entsprechend der Zielstellung wurden die Daten in drei Analyseschritten untersucht. In dem ersten Analyseschritt wurde die zeitliche Entwicklung der drei Personenmerkmale fokussiert, anschließend wurden bivariate zeitliche Zusammenhänge untersucht und abschließend in einem trivariaten Modell das Zusammenspiel aller drei Konstrukte über die Zeit modelliert. Die univariaten Befunde deuten erwartungskonform auf zwei gegenläufige Trends hin: Die Ausprägungen des individuellen Interesses und der Freude nehmen von Jahrgangsstufe 9 zu 11 ab, während gleichzeitig das Konzeptwissen der Schülerinnen und Schüler zunimmt (Höft et al., 2019). Die Ergebnisse der LCSMs zeigen zudem, dass insbesondere die Schülerinnen und Schüler, die eine zunächst hohe Ausprägung des individuellen Interesses bzw. der Freude hatten, diese im Anschluss in stärkerem Maße verlieren. Hinsichtlich der Entwicklung des Konzeptwissens lassen die Befunde darauf schließen, dass in dem vorliegenden Sample Schülerinnen und Schüler mit einem höheren Konzeptwissen anschließend einen geringeren Zuwachs des Konzeptwissens erfahren, die Varianz der Ausprägung nicht zunimmt und insofern kein Matthäuseffekt vorliegt.

Die hierauf aufbauenden bivariaten Befunde deuten ebenfalls erwartungskonform auf reziproke, sich gegenseitig positiv verstärkende Effekte zwischen dem individuellen Interesse und dem Konzeptwissen (s. Eccles et al., 2015) sowie der Freude am Fach und dem Konzeptwissen hin. Diese

bedeutsame, jedoch in der Stärke als gering zu interpretierende, reziproke Beziehung zwischen Freude und Konzeptwissen ist gemäß KWT theoriekonform und ist vergleichbar mit bisherigen Befunden für die Mathematik in Sekundarstufe I (vgl., Pekrun et al., 2017). Die Befunde bzgl. des Zusammenhangs zwischen individuellem Interesse und Leistung sind schwer mit bisherigen Befunden zu vergleichen, da sich die hier vorgestellte Studie gleich in mehreren Faktoren von bisherigen Studien unterscheidet (bspw. in Hinblick auf die Studien von Köller et al. (2001) und Marsh et al. (2005) im Fach, den Instrumenten zur Erfassung von Interesse bzw. Performanz, dem Zeitabstand zwischen Messungen oder dem Alter). Entgegen den Erwartungen konnten keine Effekte des individuellen Interesses auf die Freude gefunden werden, während die Freude sich wiewohl positiv auf die Interessenentwicklung auswirkte (bspw. Schukajlow & Rakoczy, 2016). Diese Ergebnisse indizieren, dass dem Prozess der Interessenabnahme durch während des Unterrichts erlebter Freude sowie durch den Konzeptwissenserwerb, d. h. auch durch die Gestaltung der Lernumgebung, entgegengesteuert werden kann. Positive Leistungsentwicklungen und damit verbundene positive Leistungsemotionen wirken sich demnach positiv auf das individuelle Interesse der Schülerinnen und Schüler aus. Dabei sei im Zusammenhang dieser Studie noch einmal betont, dass das Interesse als Häufigkeit der Auseinandersetzung mit Chemiekontexten außerhalb des Schulunterrichts operationalisiert und damit ein äußerst autonomes und selbstbestimmtes Engagement ohne institutionelle extrinsische (Leistungs-)Anreize wie bspw. Noten als Referenzrahmen gewählt wurde. Ein höheres Interesse der Schülerinnen und Schüler manifestiert sich entsprechend in einer vermehrten aktiven Auseinandersetzung mit Chemieinhalten bspw. dem Schauen von Lernvideos oder dem Lesen von Zeitschriften und Büchern. Die durch diese Medien transportierten Inhalte können von den in den Lehrpläne formulierten Inhalten und Kompetenzen abweichen, wodurch ggf. eine geringere Passung zu dem eingesetzten Fachttest vorliegt und das erhobene Interesse deshalb – im Gegensatz zu der Auseinandersetzung mit Unterrichtsinhalten – möglicherweise weniger stark zu einer erfolgreichen Performanz im vorliegenden Fachttest beiträgt (vgl. Rotgans und Schmidt, 2017c hinsichtlich der Operationalisierung). Eine Operationalisierung des individuellen Interesses, die stärker auf das Fach als auf die Domäne fokussiert, könnte ggfs. stärkere, positive Effekte des Interesses auf das Konzeptwissen indizieren. Die positiven Zusammenhänge der fachlichen Leistung sowie der Freude am Fach mit dem außerschulischen Engagement mögen für sich naheliegend erscheinen. Sie verdeutlichen jedoch auch noch einmal das Potenzial in den Händen der Lehrkräfte. So könnte die Wahl der Unterrichtskonzeption zur Steigerung des Interesses und der Leistung genutzt werden, bspw. indem durch kontextbasierte Unterrichtsverfahren die Relevanzwahrnehmung der Schülerinnen und Schüler gesteigert wird (Hulleman et al., 2010; Sevan et al., 2018). Aber auch Maßnahmen, die sich positiv auf das Unterrichtsklima auswirken können zur Förderung des individuellen Interesses, der Freude und der Leistung beitragen. So könnte ein unterstützendes Lernklima, das u. a. durch konstruktives, individualisiertes Feedback und entsprechende Formen der Leistungsattribution in diesem Feedback charakterisiert ist, sowohl die Leistungsentwicklung (Hattie & Timperley, 2007) als auch die Leistungsemotionen (Pekrun et al., 2014) und darüber mediiert auch das individuelle Interesse positiv beeinflussen. Die positiven Rückkopplungseffekte des individuellen Interesses auf den Erwerb des Konzeptwissens werden hingegen durch die Befunde des trivariaten Modells relativiert. Die Befunde dieses letzten Auswertungsschritts deuten drauf hin, dass der positive Effekt des individuellen Interesses auf den Konzeptwissenserwerb nicht über den Effekt der Freude am Fach auf das Wissen hinweg wirkt. Ein Grund hierfür könnte sein, dass individuelles Interesse dazu führt, dass Schülerinnen und Schüler sich Wissen aneignen, dass über die Lehrpläne bzw. die durch den Fachttest erfassten Inhalte hinausgeht (vgl., Pinxten et al., 2014). In diesem Sinne wäre zu prüfen, ob eine andere Operationalisierung des individuellen Interesses, die unmittelbar auf das Interesse am Unterrichtsfach Chemie abzielt, an dieser Stelle doch zu einem reziproken Zusammenhangsmuster führen würde.

Zusammenfassend indizieren die vorliegenden Ergebnisse, dass vorwiegend positive Effekte der unterrichtsbezogenen Konstrukte Freude am Fach und Konzeptwissen auf die Änderung des individuellen Interesses und weniger rückwirkende Effekte des individuellen Interesses auf die Freude am Fach und auf das Konzeptwissen vorliegen. Dies deutet insgesamt auf positive, motivationale Effekte

des Unterrichts auf außerschulische Kontexte hin. Relevant für die Forschung ist insbesondere der Fokus auf die Oberstufe, den Chemieunterricht (bisherige Befunde konzentrieren sich auf die Mathematik) sowie auf individuelle Veränderungen der Personenmerkmale über mehrere Messzeitpunkte. Hinsichtlich Implikationen für die Schule muss aufgrund der distalen Erfassung – die Freude am Chemieunterricht und das individuelle Interesse wurden als Trait erfasst, sie können somit keine interindividuellen Variationen zwischen Situationen (Schulstunden und Inhalten) abbilden – auf eher globaler Ebene verweilt werden. Ausgehend von den Befunden, dass sowohl die Freude am Fach als auch das individuelle Interesse zum Wissenserwerb beitragen kann, kann auf globale Interventionen bspw. zur Förderung der Freude der Lehrkräfte (Frenzel et al., 2017), Interventionen die den subjektiven Wert und die Relevanz betonen (Bernacki & Walkington, 2018; Habig, van Vorst et al., 2018; Hulleman et al., 2010; Reber et al., 2018; van Vorst et al., 2018) sowie den verstärkten Einsatz des formativen Assessments (Pinger et al., 2016; Rakoczy et al., 2018) verwiesen werden. So können formative Assessments aufgrund der regelmäßigen individuellen Leistungsrückmeldung zur Förderung von Kontroll-Appraisals bzw. Erfolgserwartungen führen, die wiederum positive Emotionen und Interesse wecken können und so zu besserer Leistung führen und positive Feedback-Schleifen könnten (vgl., Pekrun et al., 2007).

## 8.6 Limitationen

Die Auswertung im Rahmen der vorliegenden Studie fokussiert in erster Linie auf die Betrachtung inter- und intraindividuelle Unterschiede und deren Entwicklung über die Zeit. Dabei wurde in den Analysen die genestete Struktur der Daten nicht durch bspw. Mehrebenenmodelle berücksichtigt, z.B. zur Modellierung des Einflusses der Lehrkraft oder von Kompositionseffekten. Diese methodische Entscheidung ist zum einen darin begründet, dass die Klassenkomposita im Verlauf der Studie durch die Wahl von Profil- bzw. Prüfungskursen wechselten, und zum anderen die Stichprobengröße für die Schätzung der latenten Modelle nicht hinreichend war. Des Weiteren war das Sample nicht repräsentativ und die Samplegröße nahm aufgrund struktureller Rahmenbedingungen (insb. Abwahl des Faches Chemie) sowie der Probandenmortalität ab. Diese Rahmenbedingungen der Studie ließen eine Betrachtung weiterer institutioneller Ebenen aus methodischen wie inhaltlichen Gesichtspunkten nicht sinnvoll erscheinen. Auch konnte dadurch der Einfluss des Oberstufenprofils auf das Zusammenspiel der drei Konstrukte nicht berücksichtigt werden, welcher den Zusammenhang zwischen Interesse und Performanz moderieren kann (Köller et al., 2001). Eine weitere Limitation in Bezug auf das Zusammenspiel von individuellen Interesse, Freude am Fach und Konzeptwissen betrifft den Fachtest. Dieser bezieht sich stark auf Kerninhalte des Curriculums, während weitergehende Inhalte, von denen auszugehen ist, dass diese insbesondere von interessierten Schülerinnen und Schülern durch unterschiedlichste inner- und außerschulische Lerngelegenheiten erworben werden, nicht berücksichtigt werden konnten. Es verbleibt zu prüfen, ob bei einer engeren Kopplung der Operationalisierungen des Tests und des individuellen Interesses dementsprechend doch eine Rückwirkung des Interesses auf das Testergebnis bzw. die Freude am Fach konstatiert werden könnte. Des Weiteren könnte eine feinere Modellierung des Interesses (Berücksichtigung der Entwicklungsphase und der Valenzen), einen präziseren Einblick in das Wechselspiel der drei Konstrukte ermöglichen und genauere Ansatzpunkte für Interventionsstudien offerieren.



## Teil III

### Gesamtdiskussion und Schlussfolgerungen



## Diskussion

Die vorliegenden empirischen Arbeiten widmeten sich der Entwicklung des individuellen Interesses von Schülerinnen und Schülern der gymnasialen Sekundarstufen im Kontext des Chemieunterrichts. Hierzu wurde zum einen das individuelle Interesse der Gymnasiastinnen und Gymnasiasten an der Domäne Chemie untersucht, zum anderen konnten durch die Nutzung des RIASEC+N Modells zusätzlich differenzierte Einblicke in die tätigkeitsspezifischen Interessen sowie deren Entwicklungen gewonnen werden. Ferner wurden in dieser Forschungsarbeit Zusammenhänge zu Kovariaten analysiert. Ein besonderes Augenmerk lag hierbei auf dem Zusammenspiel zwischen dem individuellen Interesse und dem Konzeptverständnis, für das ein Modell der reziproken Zusammenhänge angenommen wurde. Weiterhin wurden Zusammenhänge des individuellen Interesses zur Lernfreude im Chemieunterricht und der wahrgenommenen Nützlichkeit der Chemie inspiziert. Die drei zentralen Fragestellungen der Forschungsarbeit lauten:

- F1. Wie entwickeln sich die individuellen Interessen der Schülerinnen und Schüler im Fach Chemie über die Sekundarstufen I und II?
- F2. Wirkt sich ein elaboriertes Konzeptverständnis positiv auf die Entwicklung des individuellen Interesses aus?
- F3. Wirkt sich das individuelle Interesse positiv auf den Erwerb eines elaborierten Konzeptverständnisses aus?

Aufschluss über diese Fragestellungen sollte die Auswertung einer querschnittlichen (Jahrgangsstufen 5 – 11) sowie zweier an den Querschnitt gekoppelten längsschnittlichen Erhebungen (LS5: Jahrgangsstufen 5–8; LS9: Jahrgangsstufen 9–11) geben. Hierzu wurden in den vier inkludierten Publikationen (a) querschnittliche Trends der individuellen Interessenausprägungen an naturwissenschaftlichen Tätigkeiten (RIASEC+N) und des Konzeptverständnisses sowie korrelative Zusammenhänge zwischen den Konstrukten in den Jahrgangsstufen 5–11 analysiert (Höft et al., 2019), (b) die längsschnittliche Entwicklung der individuellen Interessen an naturwissenschaftlichen Tätigkeiten (RIASEC+N) und die Entwicklung des Konzeptverständnisses sowie kreuzverzögerte Zusammenhänge zwischen diesen beiden Konstrukten in der Sekundarstufe II (Jahrgangsstufe 9–11) untersucht (Höft & Bernholt, 2019b), (c) die Vorhersage von Änderungen des individuellen Interesses an naturwissenschaftlichen Tätigkeiten (RIASEC+N) und des Konzeptverständnisses durch vorhergehendes Interesse und Konzeptverständnis durchgeführt sowie die Prädiktion der wahrgenommenen Nützlichkeit durch Interesse und Konzeptverständnis vorgenommen (Höft & Bernholt, in Rev.) und abschließend (d) die beidseitigen kreuzverzögerten Effekte zwischen dem individuellen Interesse an der Domäne Chemie, der Lernfreude im Chemieunterricht und dem Konzeptverständnis modelliert (Höft & Bernholt, 2019a).

Im Folgenden werden zunächst die Ergebnisse der vier Publikationen zusammengefasst und bestehende Limitationen diskutiert. Anschließend wird in den Implikationen als Forschungsdesiderat eine mögliche Anschlussstudie diskutiert, die eine präzisere Beschreibung der Interessenentwicklung über kürzere Zeitskalen in Angriff nehmen soll.

### 9.1 Zusammenfassung der Ergebnisse

#### Ergebnisse bezüglich Forschungsfrage F1

Hinsichtlich Forschungsfrage F1, die die Entwicklung des individuellen Interesses fokussiert, wurde eine Abnahme des chemiebezogenen individuellen Interesses mit steigender Jahrgangsstufe angenommen. Die Interessenabnahme sollte sich vorwiegend zu Beginn des Chemieunterrichts,

d. h. in den Jahrgangsstufen 7 – 9, vollziehen und sowohl das individuelle Interesse an naturwissenschaftlichen Tätigkeiten im Chemieunterricht (RIASEC+N) als auch das individuelle Interesse an der Domäne Chemie betreffen.

Die querschnittlichen Befunde deuten auf eine signifikante Interessenabnahme über die Sekundarstufe I und II in alle sieben Dimensionen des RIASEC+N Modells hin. Frappant bei den querschnittlichen Daten ist, dass das Interesse an social Tätigkeiten, also Tätigkeiten, bei denen sich Schülerinnen und Schüler gegenseitig beim Lernen unterstützen, ab der Jahrgangsstufe 8 nicht mehr bedeutsam fällt (Höft et al., 2019). Die längsschnittlichen Untersuchungen bekräftigen diese Befunde überwiegend. So konnte ein im Mittel linearer Interessenabfall von Jahrgangsstufe 9 zu 11 über alle sieben Interessendimensionen hinweg beobachtet werden. Hinsichtlich der Dimension social steht der längsschnittliche Befund (lineare Interessenabnahme in der Sekundarstufe II) somit im Widerspruch zu dem querschnittlichen Befund (keine signifikante Interessenabnahme in der Sekundarstufe II). Die längsschnittliche Interessenabnahme in der Sekundarstufe II fällt für realistic Tätigkeiten am stärksten aus (moderate Effektstärke). Die Schülerinnen und Schüler verlieren somit am stärksten ihr Interesse an Tätigkeiten, die sie in Jahrgangsstufe 9 (und auch zu Beginn des Chemieunterrichts) am ehesten als interessant bewerteten (Höft & Bernholt, 2019b). Zu korrespondierenden Ergebnissen kam die längsschnittliche Analyse des individuellen Interesses an der Domäne Chemie. Auch hier wurde ein signifikanter Abfall des Interesses von Jahrgangsstufe 9 zu 11 verzeichnet, wobei die Interessenabnahmen für interessiertere Schülerinnen und Schüler höher ausfielen (Höft & Bernholt, 2019a).

Die in der Literatur diskutierte U-Form der Interessenabnahme aufgrund der Stabilisierung der Interessen in der Sekundarstufe II oder gar der Zunahme des Interesses in der Oberstufe konnte in der vorliegenden Stichprobe für das Fach Chemie nicht gefunden werden (vgl. Dotterer et al., 2009; Gaspard et al., 2017). Auch eine Ausdifferenzierung des individuellen Interesses an naturwissenschaftlichen Tätigkeiten (RIASEC+N) konnte in den personenzentrierten Voruntersuchungen nicht festgestellt werden.

Anknüpfungspunkte für weiterführende Analysen. Die bisherigen längsschnittlichen Befunde nehmen eine separate Betrachtung der einzelnen Interessendimensionen vor. Es ist jedoch wahrscheinlich, dass die Entwicklung der tätigkeitsspezifischen Interessen nicht disjunkt voneinander erfolgt, sondern ein komplexes Wirkungsgefüge zwischen den Interessendimensionen des RIASEC+N Modells besteht. So könnte bspw. das Interesse an realistic Tätigkeiten das subsequente Interesse an investigative Tätigkeiten positiv beeinflussen. Einen Einblick in die Zusammenhänge zwischen den Interessen in verschiedenen Dimensionen können bivariate Cross-Lagged-Panel oder Latent-Change-Score Modelle liefern. Eine simultane Modellierung der Zusammenhänge zwischen alle sieben Interessendimensionen war mit den vorliegenden Daten des DoLiS-Projekts allerdings nicht möglich.

Zudem könnten die Trajektorien der Interessenentwicklung in verschiedenen Dimensionen in Zusammenhang gestellt werden. Hier könnten bspw. die Korrelationen zwischen Steigungsparametern von latenten Wachstumsmodellen berechnet werden. Neben diesen variablenzentrierten Ansätzen könnte sich, um eine weitere Perspektive auf die Entwicklung des individuellen Interesses gewinnen zu können, die Modellierung mittels personenzentrierter Analyseverfahren, bspw. mittels latenter Profilanalysen, lohnen. Hier könnten intraindividuelle Trajektorien der Interessenentwicklung je Interessendimension bestimmt und die gefundenen Trajektorien mittels Profilanalyse latenten Klassen zugeordnet werden. Dies ermöglicht die Aufdeckung von Gruppen von Schülerinnen und Schüler mit vergleichbaren Interessenentwicklungen. Ein sophisticatedes Verfahren stellt die Bildung von Profilen mithilfe von latenten Wachstumsmodellen dar (growth mixture modeling, GMM; s. Nylund-Gibson et al., 2014). Eine simplere Methode, die eine geringere Samplegröße erfordert, ist die Bildung von Profilen direkt über die manifesten Interessenausprägungen je Messzeitpunkt (Schiefer et al., 2018).

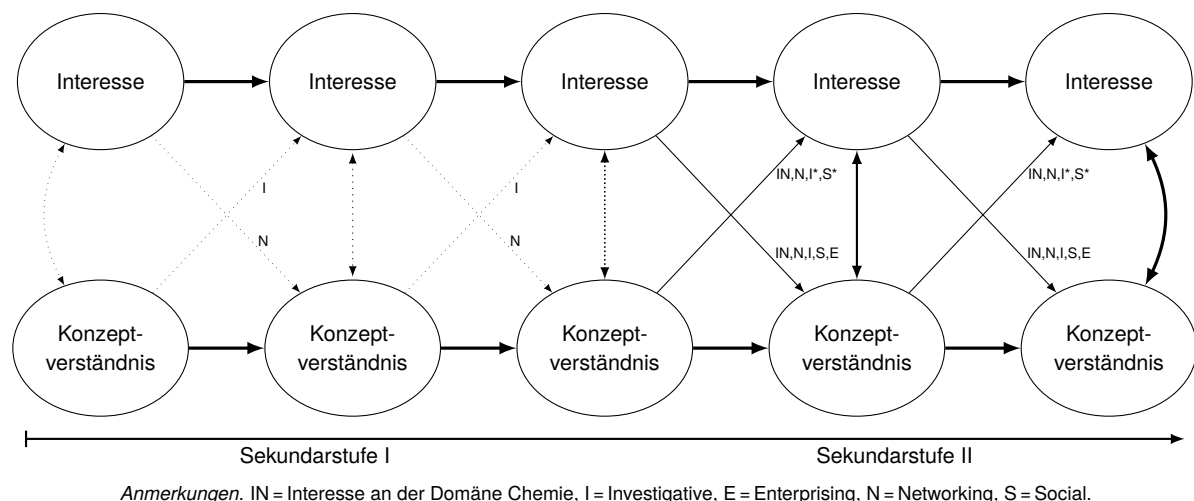


### Ergebnisse bezüglich Forschungsfrage F2 und F3

Hinsichtlich Forschungsfrage F2 und F3, die das Zusammenspiel von individuellem Interesse und Konzeptverständnis fokussieren, wurde aus der Literatur ein Modell der reziproken Effekte (REM) abgeleitet. Dieses Modell postuliert beidseitige, kreuzverzögerte, positive Effekte zwischen dem individuellen Interesse und dem Konzeptverständnis, welche idealiter einen synergistischen Prozess beschreiben. Es wurde angenommen, dass die reziproken Zusammenhänge zwischen dem Interesse und dem Konzeptverständnis vorwiegend für die Interessendimensionen investigative und networking auftreten, da die Items dieser Dimensionen wirkungsvolle Lerngelegenheiten für den Wissenserwerb adressieren. Eine weitere Annahme war, dass diese Zusammenhänge insbesondere in höheren Jahrgangsstufen zum Tragen kommen.

Die in Publikation I vorgenommenen querschnittlichen Analysen deuten auf bedeutungsvolle Zusammenhänge zwischen dem Interesse an investigative, enterprising, social sowie networking Tätigkeiten und dem Konzeptverständnis der Schülerinnen und Schüler der Oberstufe hin. Für Schülerinnen und Schüler der Sekundarstufe I konnten keine bedeutungsvollen korrelativen Zusammenhänge gefunden werden. Die querschnittlichen Analysen konnten erste Hinweise auf eine positive Assoziation zwischen den individuellen Interessen an diesen spezifischen vier Tätigkeitsbereichen und dem Konzeptverständnis liefern, für einen Einblick in das Zusammenwirken von Interesse und Verständniserwerb waren jedoch längsschnittliche Analysen notwendig. Aus diesem Grund wurde in Publikation II und III das Zusammenspiel zwischen dem Interesse an naturwissenschaftlichen Tätigkeiten und dem Konzeptverständnis in bivariaten Cross-Lagged-Panel sowie Latent-Change-Score Modellen untersucht. Daran anschließend wurden in Publikation IV wiederum mittels Latent-Change-Score Modellen Zusammenhänge zwischen dem individuellen Interesse an der Domäne Chemie und dem Konzeptverständnis exploriert.

Im Folgenden soll nun eine Zusammenführung der Ergebnisse der vier Publikationen stattfinden. Hierzu wurden die Ergebnisse der Studien in Abbildung 9.1 zusammengetragen, indem, wie bereits diskutiert, die mit zunehmender Jahrgangsstufe steigende Korrelationen zwischen dem Interesse und dem Konzeptverständnis angedeutet und signifikante kreuzverzögerte Effekte zwischen den beiden Konstrukten eingezeichnet wurden.<sup>1</sup>



**Abb. 9.1:** Abstraktion der empirischen Befunde zum Modell der reziproken Effekte. Zusammenhänge innerhalb von Interessendimensionen, die mit einem Asterisk \* markiert sind, wurden nur bedingt signifikant.

Für die Sekundarstufe I konnten positive Effekte des Konzeptverständnisses auf subsequentes Interesse an investigative Tätigkeiten sowie positive Effekte des Interesses an networking Tätigkeiten auf das subsequente Konzeptverständnis aufgedeckt werden. Es liegen für keine der sieben Interessendimensionen reziproke Effekte zum Konzeptverständnis vor.

1. Hier wurde auf die Ergebnisse der bivariaten cross-lagged Analysen zurückgegriffen.

Für die Sekundarstufe II indizieren die Befunde das Vorliegen von positiven, reziproken Effekten zwischen dem individuellen Interesse an der Domäne Chemie sowie an networking Tätigkeiten und dem Konzeptverständnis. Weiterhin liegen für die Dimensionen investigative und social nur eingeschränkte Hinweise auf reziproke Zusammenhänge zum Konzeptverständnis in der Sekundarstufe II vor.

Die Ergebnisse dieser Forschungsarbeit unterstützen somit bisherige Befunde, die indizieren, dass der Zusammenhang zwischen dem individuellen Interesse und der Performanz in höheren Jahrgangsstufen ausgeprägter ist und vornehmlich in der Sekundarstufe II zum Tragen kommt (bspw. Denissen et al., 2007; Ganley & Lubienski, 2016; Kim et al., 2015; Köller et al., 2001). Die differenzierte Modellierung des individuellen Interesses mithilfe des RIASEC+N Modells konnte zusätzlich zeigen, dass diese Befunde, das Vorliegen von positiven kreuzverzögerten Effekten und der Anstieg dieser Effekte mit steigender Jahrgangsstufe, nicht gleichermaßen für alle Tätigkeitsbereiche gelten. Aus diesem Grund motivieren die Ergebnisse dieser Forschungsarbeit dazu, bei zukünftigen Analysen des Zusammenhangs zwischen dem individuellen Interesse und der Performanz, das Interesse der Schülerinnen und Schüler an (vermutlich) lernwirksamen Tätigkeiten bei der Operationalisierung des jeweiligen Interessenkonstrukts miteinzubeziehen.

Anknüpfungspunkte für weiterführende Analysen. Die Zusammenhänge zwischen dem individuellen Interesse an der Domäne Chemie und dem Konzeptverständnis wurden nur für Schülerinnen und Schüler der Sekundarstufe II untersucht. Hier könnten Analysen anknüpfen, die diese Zusammenhänge für Gymnasiastinnen und Gymnasiasten der Sekundarstufe I modellieren und die Moderation des Effekts durch das Alter prüfen. Weiterhin wäre es spannend, zu untersuchen, ob es Schülerinnen und Schüler gibt, deren individuelles Interesse und Konzeptverständnis simultan mit steigender Jahrgangsstufe steigen, wie es das Model of Domain Learning sensu Alexander et al. (1995) und das Epistemic Model of Interest in Education sensu Rotgans und Schmidt (2017; 2017) nahelegen. Diese Annahme könnte wiederum mit dem bereits angesprochenen growth mixture modeling Verfahren untersucht werden (Nylund-Gibson et al., 2014).

### Ergebnisse bezüglich Kovariaten

In der dritten Publikation wurde die Vorhersage der wahrgenommenen Nützlichkeit der Chemie (erhoben in Jahrgangsstufe 8 bzw. 12) durch das Interesse an investigative, social und networking Tätigkeiten (erhoben in Jahrgangsstufe 8 bzw. 11) berechnet. Die wahrgenommene Nützlichkeit wurde als Relevanz bzw. Nützlichkeit für die eigene Zukunft sowie die Gesellschaft operationalisiert und als Indikator für zukünftiges Engagement in Chemiekontexten interpretiert (Höft & Bernholt, in Rev.). Das individuelle Interesse an investigative, social und networking Tätigkeiten war prädiktiv für die wahrgenommene Nützlichkeit, auch nach Kontrolle von Konzeptverständnis, Geschlecht und kognitiven Fähigkeiten. Für autoregressive Effekte konnte nicht kontrolliert werden, da die wahrgenommene Nützlichkeit nur zum letzten Messzeitpunkt erhoben wurde. Die Ergebnisse deuten an, dass die wahrgenommene Nützlichkeit sich nicht nur zur Förderung des Interesses eignet, sondern zeitgleich abhängig von vorherigen Interessen der Schülerinnen und Schüler ist (vgl. Lauermaun et al., 2017). Dies könnte bedeutend für die Planung von Utility-Value Interventionen sein.

Weiterhin wurde der Zusammenhang zwischen dem individuellen Interesse an der Domäne Chemie und der Lernfreude im Chemieunterricht mittels Latent-Change-Score Modellen untersucht. Diese Analysen deuten darauf hin, dass die Lernfreude im Chemieunterricht positiv auf die Interessenentwicklung wirkt, das Interesse hingegen nicht prädiktiv für die Lernfreude ist. Diese Befunde widersprechen der aufgestellten Hypothese, dass Schülerinnen und Schüler mit einem höheren Interesse an der Domäne Chemie den Unterricht eher als freudvoll erfahren, da sie hier ihren potentiellen Neigungen und Interessen nachgehen können. Das Ausbleiben dieses positiven Effekts könnte darauf zurückzuführen sein, dass möglicherweise die angebotenen Lernangebote nicht die Interessen der Schülerinnen und Schüler aufgreifen oder der Unterricht als zu restriktiv wahrgenommen wird. Weiterhin könnten sich differierende Zugänge zu den Fachinhalten, bspw. eine stärker

an der Wissenschaft orientierte Perspektive im Unterricht gegenüber populärwissenschaftlichen Herangehensweisen in Freizeitangeboten, negativ auf die Lernfreude auswirken.

## 9.2 Limitationen

In diesem Abschnitt werden Limitationen diskutiert, die für alle vier Publikationen gelten. Diese betreffen u. a. das Studiendesign, die Datengrundlage sowie die Methodik. Hinsichtlich des Studiendesigns ist eine starke Einschränkung des DoLiS-Projekts, dass keine längsschnittlichen Aussagen über den Übergang von Jahrgangsstufe 8 zu Jahrgangsstufe 9 gemacht werden können. Dieser Übergang ist zum einen aus struktureller Sicht interessant, da fünf der sechs an dem Projekt beteiligten Gymnasien mit dem Chemieunterricht in der 8. Jahrgangsstufe starteten (G8), zum anderen aber auch aus fachdidaktischer Sicht spannend, da in dieser Phase meist die Einführung der elektronischen Struktur der Atome stattfindet und Anstelle einer phänomenologischen eine modellbasierte, submikroskopische Deutung an Bedeutung gewinnt (Bernholt et al., 2019). Ebenso deuten die querschnittlichen Trends der tätigkeitsspezifischen Interessenausprägungen sowie die korrelativen Zusammenhänge zwischen Interesse und Konzeptverständnis auf bedeutsame Veränderungen in dieser Zeitspanne hin (Höft et al., 2019).

Hinsichtlich der Datengrundlage ist die geringe Samplegröße des DoLiS-Projekts zu nennen, die eine Berücksichtigung der Mehrebenenstruktur (Nestung von Schülerinnen und Schülern in Klassen) für komplexere Modelle nicht erlaubte, sodass weder eine Fehlerkorrektur (bspw. mittels „Type Is Complex“) noch eine Schätzung von Kompositionseffekten möglich war. Zudem ist hier zu beachten, dass sich die Klassenkomposita während der Erhebung bspw. beim Übergang von der Sekundarstufe I in die Sekundarstufe II änderten und somit eine wesentlich komplexere Modellierung der Mehrebenenstruktur notwendig wäre.

Bezüglich der Methodik ist eine starke Limitation, dass keine Informationen über die Opportunitätsstruktur im Unterricht, d. h. bspw. über die Art, die Qualität und Quantität der durch das RIASEC+N Testinstrument adressierten Lerngelegenheiten, erhoben wurden. Dies ist von besonderer Bedeutung, da Interessentheorien die (wiederholte) Auseinandersetzung mit dem Interessengegenstand als notwendige Bedingung für die Entwicklung eines individuellen Interesses artikulieren (bspw. Hidi & Renninger, 2006). Weiterhin kann die vorgenommene Operationalisierung des RIASEC+N Modells kritisiert werden. So fand aus testökonomischen Gründen weder eine Differenzierung der Interessenvalenzen statt, noch wurde auf spezifische Indikatoren des individuellen Interesses eingegangen (bspw. die selbstbestimmte Auseinandersetzung mit Inhalten oder das Stellen und Beantworten von eigenständig generierten Fragen; s. Kapitel 2).

Eine grundsätzliche Limitation der Studie ist zudem die ausschließliche Erfassung des Interesses mittels Likert-Skalen. Qualitative Herangehensweisen, bspw. mithilfe von Interviews oder Tagebüchern, könnten weitergehende Einblicke in die Interessenentwicklung der Schülerinnen und Schüler liefern. Zuletzt sei darauf hingewiesen, dass es sich bei der Studie um eine Feldstudie handelt. Dies bedingt, dass auch längsschnittliche Zusammenhänge nur bedingt kausal gedeutet werden dürfen, da nicht berücksichtigte Variablen die Ergebnisse konfundieren könnten.

## 9.3 Implikationen

In diesem Abschnitt werden aus den Forschungsarbeiten gezogene Implikationen ausgeführt. Der Schwerpunkt liegt hierbei auf der Entwicklung von Forschungsfragen und dem Skizzieren einer Anschlussstudie.

### Schule

Die Entwicklung des Interesses und der Erwerb des Fachwissens scheinen miteinander verquickt zu sein. Positive Effekte der Lehrperson auf die Interessenentwicklung können durch die Wahl der Unterrichtskonzeption (Savelsbergh et al., 2016) und die Aufbereitung der Fachinhalte (Fechner,

2009; Sevia et al., 2018), der Förderung des Kompetenz- (Pinger et al., 2016; Rakoczy et al., 2018) und Autonomieerlebens (Flunger et al., 2019) oder des eigenen Auftretens (Frenzel et al., 2017) bewirkt werden. Soll zeitgleich eine Förderung des individuellen Interesses und des Erwerbs eines elaborierten Konzeptverständnisses erfolgen, so legen die Forschungsbefunde dieser Arbeit eine Konzentration auf investigative, networking und social Tätigkeiten nahe. Es ist spannend zu sehen, dass Tätigkeiten, die das gemeinsame Lösen von Aufgaben in Gruppen sowie die Unterstützung von Mitschülerinnen und Mitschülern bei Lernschwierigkeiten betonen, potentiell positiv mit dem Interesse und der Leistung zusammenhängen könnten. Hier könnten beispielsweise Ansätze des Peer-Tutorings (Ding & Harskamp, 2011) greifen, entsprechend derer leistungsstarke Schülerinnen und Schüler Leistungsschwächeren bei den ersten Lösungsschritten helfen, bis diese die Lernaufgaben selbständig bewältigen können, bzw. diese als mittelschwer einschätzen (vgl. Schuster, 2015).

Da in dieser Studie jedoch keine systematische Erfassung von Klassenraumvariablen integriert war, ist eine Ableitung von Implikationen für den Schulunterricht nur basierend auf den berichteten Analysen und Befunden nicht valide möglich. Nichtsdestotrotz lässt sich schlussfolgern, dass Lehrerinnen und Lehrer über ein großes Potential verfügen, die Interessenentwicklung ihrer Schülerinnen und Schüler positiv zu beeinflussen, ohne den Lernerfolg zu vernachlässigen.

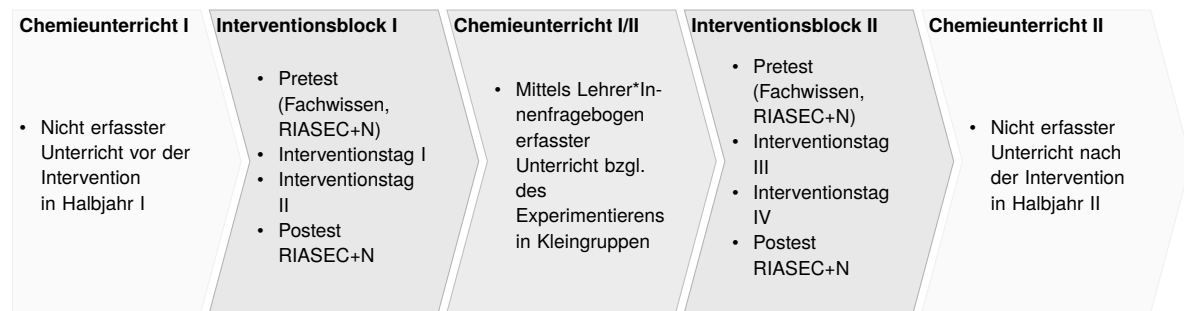
## Forschung

Für die Forschung lässt sich aus den gewonnenen Erkenntnissen und Limitationen der Studien u. a. ein Bedarf an (a) der Weiterentwicklung des RIASEC+N Instruments als auch (b) der Durchführung von Anschlussstudien, die die Verknüpfung von Lerngelegenheit und Interessenwicklung fokussieren, schlussfolgern. Hinsichtlich der Weiterentwicklung des RIASEC+N Erhebungsinstruments könnte eine stärkere Orientierung an den Interessentheorien sensu Krapp (2002b) und Hidi und Renninger (2006) fruchtbar sein. Hierzu könnten bereits bestehende Testinstrumente zur differenzierten Erfassung des situationalen Interesses (bspw. Knogler et al., 2015; Linnenbrink-Garcia et al., 2010; Linnenbrink-Garcia et al., 2013) und des individuellen Interesses (bspw. Krapp et al., 1993; Rotgans & Schmidt, 2017c) aufgegriffen werden und für die Erfassung des Interesses an naturwissenschaftlichen Tätigkeiten adaptiert werden. Ebenso ist es denkbar, das situationale Interesse mit tätigkeitsunspezifischen Testinstrumenten während der Ausübung verschiedener naturwissenschaftlichen Tätigkeiten zu erfassen, um die Vergleichbarkeit der erfassten Interessenmaße sicherzustellen.

Anschlussstudie. Nachfolgend wird der zweite Aspekt, die Analyse der Interessenentwicklung aufgrund spezifischer naturwissenschaftlicher Tätigkeiten sowie die Erfassung des situationalen Interesses mittels tätigkeitsunspezifischer Testinstrumente, expliziert. Hierzu werden Forschungsfragen und ein Studiendesign skizziert. Ziel der Anschlussstudie ist es, zu zeigen, wie sich das situationale Interesse während verschiedener Experimentierstationen entwickelt und inwiefern das Evozieren des tätigkeitspezifischen, situationalen Interesses zu einer Interessenänderung (RIASEC+N) beiträgt.

Die im Rahmen des DoLiS-Projekts durchgeführten Studien zeigten, dass Schülerinnen und Schüler zu Beginn des Chemieunterrichts an der Ausübung von hands-on Tätigkeiten (realistic) interessiert sind, ihr Interesse an diesen Tätigkeiten mit steigender Jahrgangsstufe jedoch verhältnismäßig stark sinkt. Zudem deuteten die Studien an, dass die Entwicklung des Interesses an realistic Tätigkeiten nicht systematisch mit dem Konzeptverständnis verknüpft ist. Demgegenüber indizierten die Studien, dass Schülerinnen und Schüler nur mäßig an networking Tätigkeiten interessiert sind, das Interesse in dieser Dimension aber in den vorliegenden Daten des DoLis-Projekts die stärksten Zusammenhänge zum Konzeptverständnis aufweist und die Zusammenhänge in der Sekundarstufe II einen reziproken Charakter annehmen. Vor dieser Ausgangslage soll die Interessenförderung an realistic und networking Tätigkeiten in der Sekundarstufe I (bspw. Klassenstufe 8 oder 9) durch Experimentierstationen untersucht werden. Die hypothetische Anschlussstudie besteht aus zwei geblockten Interventionen, ein Interventionsblock je Schulhalbjahr, die gleich aufgebaut sind und jeweils sieben Schulstunden umfassen. Der erste Interventionsblock wird zu Beginn des Schuljahrs, der zweite am Ende des Schuljahrs durchgeführt. Ein Interventionsblock besteht jeweils aus einem Einführungstag

mit einem Pretest (2 Schulstunden), zwei Experimentiertagen (jeweils 2 Schulstunden) und einem Posttest (eine Schulstunde). Das individuelle Interesse wird mittels des RIASEC+N Modells sowohl vor (Pretest) als auch nach einem Interventionsblock (Posttest) erfasst (s. Abbildung 9.2). Das Ziel



**Abb. 9.2:** Studiendesign der Anschlussstudie bestehend aus zwei Interventionsblöcken sowie der Erfassung von Lernangeboten.

der Unterteilung in zwei Interventionsblöcke ist die Abschätzung der langfristigen Wirkung derartig kurzer Interventionen auf die Entwicklung der individuellen Interessen. Hierzu kann die Erhebung des individuellen Interesses vor Beginn des zweiten Interventionsblocks als Follow-up Erhebung des ersten Interventionsblocks angesehen werden. Die Verteilung der Interventionsblöcke auf zwei Schulhalbjahre ermöglicht zudem einen Vergleich der Interessenänderung über diesen längeren Zeitraum mit den Ergebnissen des DoliS-Projekts, die auf Messungen in jährlichen Intervallen basieren. Um den Einfluss des Unterrichtsangebots, das zwischen den beiden Interventionsblöcke liegt, auf die Interessenentwicklung abschätzen zu können, soll mithilfe eines Lehrer\*innenfragebogens das Experimentieren in Kleingruppen erfasst werden (Zeitpunkt, Dauer, Thema, Ziel, Freiheitsgrad<sup>2</sup>).

Der Hauptbestandteil der Intervention ist das Arbeiten in den vier Experimentierstationen, die in Kleingruppen mit konstanter Zusammensetzung (Teams bestehend aus drei bis vier Lernenden) durchgeführt werden sollen und typische Aspekte von realistic und networking Tätigkeiten aufgreifen. In Abgrenzung zu unterrichtsüblichen Experimentierphasen soll in der Intervention durch übergeordnete Lernziele eine hohe Kohärenz über die vier Experimentierstationen erreicht und inhaltliche Brüche vermieden werden (vgl. Seidel et al., 2005). Die Lernziele der Unterrichtseinheit werden zu Beginn der Intervention vorgestellt und die Arbeitsphasen der Experimentierstationen expliziert und wiederholt (Problemstellung, Auswertung) mit den Lernzielen verknüpft.

Aufgrund der Durchführung in Kleingruppen sollte für die Nestung der Schülerinnen und Schüler in Gruppen statistisch kontrolliert werden, wofür eine entsprechend große Stichprobe anzustreben ist. Zudem sollte die Zusammensetzung der Kleingruppen über die gesamte Intervention möglichst konstant gehalten werden, um Kompositionseffekte statistisch schätzen zu können. Die Reihung der zu bearbeitenden Experimentierstationen wird je Interventionsblock randomisiert, um Reihenfolgeeffekte auszugleichen. Es bietet sich an, bereits erprobte Experimentierstationen einzusetzen, die sich als lernwirksam erwiesen haben. Beispielsweise könnte in der Sekundarstufe I das Säure-Base-Konzept mithilfe von bereits wissenschaftlich untersuchten Experimenten aufgegriffen werden (Rumann, 2004; Sumfleth et al., 2002; Walpuski & Sumfleth, 2007).

Die Durchführung des Interventionsblocks erfolgt in sieben konsekutiven Unterrichtsstunden des regulären Chemieunterrichts (s. Tabelle 9.1). Am ersten Studientag, dem Einführungstag, wird in einem Pretest das Vorwissen bspw. im Bereich Säuren und Basen, das individuelle Interesse an naturwissenschaftlichen Tätigkeiten (Dierks et al., 2014b), sozio-kulturelle Daten (bspw. sozio-ökonomischer Status) und Personenmerkmale (Gender, Alter) erhoben. Anschließend wird den Schülerinnen und Schülern eine Übersicht über den Ablauf der Studie gegeben, das Experimentieren in Kleingruppen an einer Experimentierstation geübt und die Bearbeitung der Aufgabenzettel und Testinstrumente zur Erhebung des situationalen Interesses besprochen (Experiment 0, s. Tabelle 9.1). An den anschließenden zwei Testtagen (Experimentiertag 1 und 2) werden jeweils zwei Experi-

2. Verortung anhand der Pole: forschend-entwickelndes Vorgehen (konstruktivistisches Vorgehen) und rezeptähnliche Experimentalvorschriften (transmissives Vorgehen).

**Tab. 9.1:** Interventionsblock der Anschlussstudie zur Analyse der Entwicklung des Interesses an naturwissenschaftlichen Tätigkeiten (RIASEC+N) durch vier Schülerexperimente im Bereich Säure-Basen.

Einführungstag Pretest	Experimentiertag 1	Experimentiertag 2	Posttest T2
<ul style="list-style-type: none"> <li>Fachwissen<sup>a</sup></li> </ul>			
<ul style="list-style-type: none"> <li>II RIASEC+N<sup>a</sup></li> <li>Gender, ...<sup>a</sup></li> </ul>			<ul style="list-style-type: none"> <li>II RIASEC+N<sup>a</sup></li> </ul>
-----			
<ul style="list-style-type: none"> <li>Experiment 0 als Einführung</li> </ul>	<ul style="list-style-type: none"> <li>Stundenbeginn</li> <li>Problemstellung</li> <li>SI<sup>a</sup></li> <li>Experiment 1</li> <li>SI realistic<sup>a</sup></li> <li>Auswertung</li> <li>SI networking<sup>a</sup></li> </ul>	<ul style="list-style-type: none"> <li>Stundenbeginn</li> <li>Problemstellung</li> <li>SI<sup>a</sup></li> <li>Experiment 3</li> <li>SI realistic<sup>a</sup></li> <li>Auswertung</li> <li>SI networking<sup>a</sup></li> </ul>	
-----			
	<ul style="list-style-type: none"> <li>Problemstellung</li> <li>SI<sup>a</sup></li> <li>Experiment 2</li> <li>SI Realistic<sup>a</sup></li> <li>Auswertung</li> <li>SI networking</li> <li>Stundenende</li> </ul>	<ul style="list-style-type: none"> <li>Problemstellung</li> <li>SI<sup>a</sup></li> <li>Experiment 4</li> <li>SI Realistic<sup>a</sup></li> <li>Auswertung</li> <li>SI networking<sup>a</sup></li> <li>Stundenende</li> </ul>	

<sup>a</sup> Erhebung mit Testinstrumenten

Anmerkungen. SI = Situationales Interesse, II = Individuelles Interesse

mentierstationen je Gruppe durchgeführt. Während dieser Experimentierstationen wird jeweils zu drei Zeitpunkten das hervorgerufene situationale Interesse gemessen: direkt nach dem Lesen und Besprechen der Problemstellung, direkt nach dem Experimentieren und direkt nach der Bearbeitung der Auswertungsaufgabe. Nach den beiden Experimentiertagen wird das individuelle Interesse an naturwissenschaftlichen Tätigkeiten mittels eines Posttests erfasst, um die Wirksamkeit der Intervention abschätzen zu können.

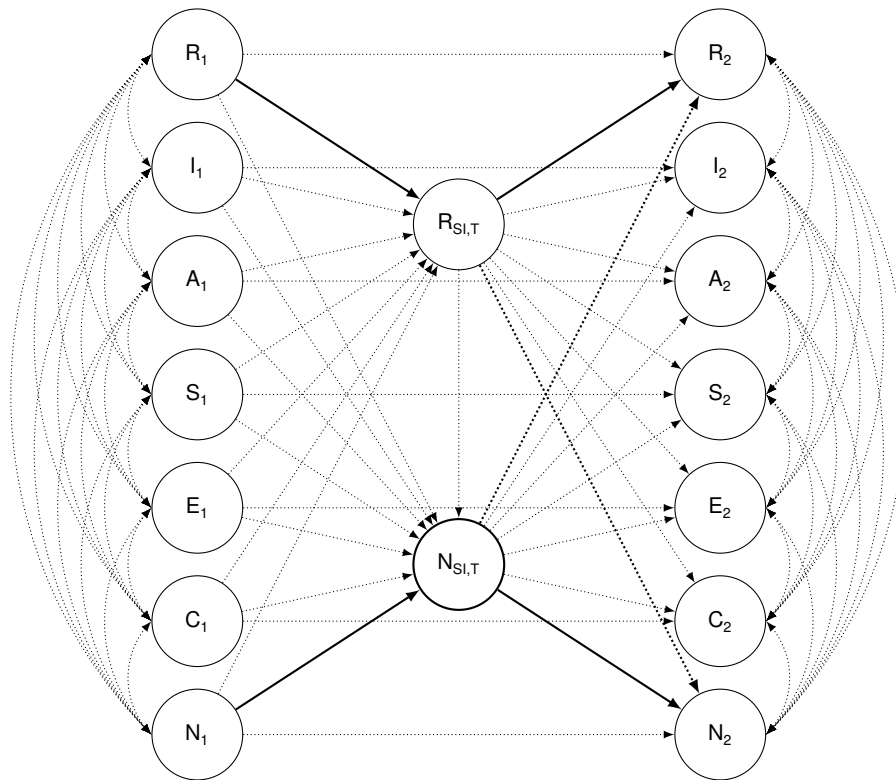
Mögliche Forschungsfragen, Hypothesen und Analysestrategien:

F1. Verändert sich das individuelle Interesse an naturwissenschaftlichen Tätigkeiten (RIASEC+N) zwischen dem Pre- und dem Posttest sowie zwischen dem Pre- und dem Follow-up Test?

Die Experimentierstationen betonen realistic und networking Tätigkeiten. Durch die wiederholte Ausübung dieser Tätigkeiten während der Intervention sollte sich das Interesse der Schülerinnen und Schüler an diesen Tätigkeiten über die Zeit verändern. Diese Veränderungen sollten zwischen dem Pre- und dem Posttest größer ausfallen als zwischen dem Pre- und dem Follow-Up Test, da nach dem Posttest von einer Interessenabnahme auszugehen ist. Dieser Fragestellung kann statistisch mittels True Individual Change Modellen untersucht werden, wobei die Standardfehler aufgrund der Nestung mittels „Type Is Complex“ kontrolliert werden sollten.

F2. Beeinflusst das situationale Interesse an realistic und networking Tätigkeiten die Entwicklung der individuellen Interessen an naturwissenschaftlichen Tätigkeiten?

Schülerinnen und Schüler, die während der Ausführung der realistic und networking Tätigkeiten ein höheres situationales Interesse angeben, sollten zum Post- und zum Follow-up Test ein höheres individuelles Interesse an realistic und networking Tätigkeiten angeben. Das individuelle Interesse an investigative, artistic, social, enterprising und conventional Tätigkeiten zum Postzeitpunkt sollte nicht durch das situationale Interesse an realistic und networking Tätigkeiten beeinflusst werden. Dieser Fragestellung kann mittels Strukturgleichungsmodellen



Anmerkungen. Die Indizes 1 und 2 bezeichnen den Messzeitpunkt, SI das situationales Interesse, T die stabile „trait“ Varianz.

**Abb. 9.3:** Strukturgleichungsmodell zur Analyse des Einflusses des situationalen Interesses an networking und realistic Tätigkeiten auf das individuelle Interesse.

nachgegangen werden. In Abbildung 9.3 ist eine mögliche Auswertungsmethode skizziert. Entsprechend dieser wird untersucht, ob das situationales Interesse (genauer die mittels Latent-State-Trait Ansatzes berechnete, über alle vier Experimentierstationen stabile Varianzkomponente) an realistic ( $R_{SI,T}$ ) sowie networking ( $N_{SI,T}$ ) Tätigkeiten das subsequeute individuelle Interesse an naturwissenschaftlichen Tätigkeiten zum Postzeitpunkt ( $R_2, \dots, N_2$ ) bei Kontrolle des initialen individuellen Interesses an naturwissenschaftlichen Tätigkeiten ( $R_1, \dots, N_1$ ) vorhersagt.

F3. Beeinflusst die Gruppenkomposition während der Experimentierstationen die Entwicklung des individuellen Interesses an naturwissenschaftlichen Tätigkeiten?

Schülerinnen und Schüler, die während der Ausführung der realistic und networking Tätigkeiten Gruppen zugehörten, deren Mitglieder im Mittel ein höheres situationales Interesse angaben, sollten zum Post- und zum Follow-up Test ein höheres individuelles Interesse an realistic und networking Tätigkeiten angeben. Diese Fragestellung könnte mittels Mehrebenenmodellen nachgegangen werden (Level 1 = Schülerinnen und Schüler, Level 2 = Kleingruppe).

Für die Erhebung des situationalen Interesses kann auf bereits bestehende Iteminventare zurückgegriffen werden. Für die hier skizzierte Anschlussstudie könnte sich die Adaption des Erhebungsinstruments von Knogler et al. (2015) eignen, mit welchem das situationales Interesse anhand einer vierstufigen Likert-Skala (1 = „gar nicht“, ..., 4 = „sehr“), der Fragestellung „Wenn du an den vorherigen Unterrichtsabschnitt (das vorherige Experiment, die vorherige Aufgabe) denkst, wie sehr ...“ und sechs Items (bspw. „...hat dieser deine Neugier geweckt?“) erhoben werden kann.

Ziel des RIASEC+N Interessenmodells ist die Gewinnung eines differenzierten Einblicks in die chemiebezogenen Interessenstrukturen von Schülerinnen und Schülern, um spezifische und zielgerichtete Interessenfördermaßnahmen zu realisieren (Dierks, 2014). Bisher wurden u. a. Interessenprofile in Abhängigkeit der Leistung, des Selbstkonzepts und des Geschlechts oder der Teilnahme an naturwissenschaftlichen Schülerwettbewerben (Blankenburg, Höffler & Parchmann, 2016; Dierks

et al., 2016; Höffler et al., 2019) gebildet und das RIASEC+N Modell genutzt, um mittels eines Projektstages (fiktiver Wettbewerbstag) Schülerinnen und Schüler für naturwissenschaftliche Schülerwettbewerbe zu begeistern (Blankenburg, Höffler, Peters et al., 2016). Die Wirkung der Intervention auf das Interesse an naturwissenschaftlichen Tätigkeiten (RIASEC+N) wurde nicht berichtet. An diese Forschungslücke könnte die vorgestellte Interventionsstudie anschließen, indem systematisch (und unter Kontrolle der Vorinteressen) die Förderung des Interesses an naturwissenschaftlichen Tätigkeiten im Regelunterricht untersucht würde. Dies könnte weitere Hinweise darauf liefern, ob die Planung von Interessenfördermaßnahmen mithilfe des RIASEC+N Modells spezifische Interessen der Schülerinnen und Schüler ansprechen. Genauer könnte die Verfolgung der obigen Fragestellungen mit dem vorgestellten Design zur Klärung der Frage beisteuern, wie Experimentierphasen die Interessenentwicklung an naturwissenschaftlichen Tätigkeiten beeinflussen. Zudem könnte Evidenz für die offene Frage generiert werden, ob es Wechselwirkungen zwischen den Interessendimensionen des RIASEC+N Modells gibt, bspw. ob das Evozieren des hervorgerufenen situationalen Interesses an realistic Tätigkeiten ebenfalls zur Entwicklung des individuellen Interesses an investigative Tätigkeiten beiträgt oder ob isoliert das spezifische Interesse an einzelnen Dimensionen gefördert werden kann.



## Literatur

- Abrahams, I. & Reiss, M. J. (2012). Practical work: Its effectiveness in primary and secondary schools in England. *Journal of Research in Science Teaching*, 49(8), 1035–1055.
- Ahmed, W. & Mudrey, R. R. (2018). The role of motivational factors in predicting STEM career aspirations. *International Journal of School & Educational Psychology*, 18(4), 1–14.
- Ahmed, W., van der Werf, G., Kuyper, H. & Minnaert, A. (2013). Emotions, self-regulated learning, and achievement in mathematics: A growth curve analysis. *Journal of Educational Psychology*, 105(1), 150–161.
- Ahmed, W., van der Werf, G., Minnaert, A. & Kuyper, H. (2010). Students' daily emotions in the classroom: Intra-individual variability and appraisal correlates. *The British journal of educational psychology*, 80(Pt 4), 583–597.
- Ainley, M. (2006). Connecting with Learning: Motivation, Affect and Cognition in Interest Processes. *Educational Psychology Review*, 18(4), 391–405.
- Ainley, M. Interest: Knowns, Unknowns, and Basic Processes (P. A. O'Keefe & J. M. Harackiewicz, Hrsg.). In: *In The Science of Interest* (P. A. O'Keefe & J. M. Harackiewicz, Hrsg.). Hrsg. von O'Keefe, P. A. & Harackiewicz, J. M. Cham: Springer International Publishing, 2017, S. 3–24. ISBN: 978-3-319-55507-2.
- Ainley, M. & Ainley, J. (2011a). A Cultural Perspective on the Structure of Student Interest in Science. *International Journal of Science Education*, 33(1), 51–71.
- Ainley, M. & Ainley, J. (2011b). Student engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science. *Contemporary Educational Psychology*, 36(1), 4–12.
- Ainley, M. & Ainley, J. Early Science Learning Experiences: Triggered and Maintained Interest (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). In: *In Interest in Mathematics and Science Learning* (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). Hrsg. von Renninger, K. A., Nieswandt, M. & Hidi, S. Washington, D.C.: American Educational Research Association, 2015, S. 17–32. ISBN: 9780935302424.
- Ainley, M. & Hidi, S. Interest and Enjoyment (R. Pekrun, Hrsg.). In: *In International Handbook of Emotions in Education* (R. Pekrun, Hrsg.). Hrsg. von Pekrun, R. Routledge, 2014. ISBN: 9780203148211.
- Ainley, M., Hidi, S. & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology*, 94(3), 545–561.
- Ainley, M., Hillman, K. & Hidi, S. (2002). Gender and interest processes in response to literary texts: Situational and individual interest. *Learning and Instruction*, 12(4), 411–428.
- Akcaoglu, M., Rosenberg, J. M., Ranellucci, J. & Schwarz, C. V. (2018). Outcomes from a self-generated utility value intervention on fifth and sixth-grade students' value and interest in science. *International Journal of Educational Research*, 87, 67–77.
- Alexander, P. A. (2003). The Development of Expertise: The Journey From Acclimation to Proficiency. *Educational Researcher*, 32(8), 10–14.
- Alexander, P. A., Jetton, T. L. & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology*, 87(4), 559–575.
- Alonzo, A. C. & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education*, 93(3), 389–421.
- Anderhag, P., Wickman, P.-O., Bergqvist, K., Jakobson, B., Hamza, K. M. & Säljö, R. (2016). Why Do Secondary School Students Lose Their Interest in Science? Or Does it Never Emerge? A Possible and Overlooked Explanation. *Science Education*, 100(5), 791–813.
- Asparouhov, T. & Muthén, B. O. (2014). Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using M plus. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(3), 329–341.
- Azevedo, F. S. (2013). The Tailored Practice of Hobbies and Its Implication for the Design of Interest-Driven Learning Environments. *Journal of the Learning Sciences*, 22(3), 462–510.
- Bathgate, M. E. & Schunn, C. D. (2017a). Factors that deepen or attenuate decline of science utility value during the middle school years. *Contemporary Educational Psychology*, 49, 215–225.
- Bathgate, M. E. & Schunn, C. D. (2017b). The psychological characteristics of experiences that influence science motivation and content knowledge. *International Journal of Science Education*, 39(17), 2402–2432.
- Baumert, J., Nagy, G. & Lehmann, R. (2012). Cumulative advantages and the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools? *Child development*, 83(4), 1347–1367.
- Bengmark, S., Thunberg, H. & Winberg, T. M. (2017). Success-factors in transition to university mathematics. *International Journal of Mathematical Education in Science and Technology*, 29, 1–14.
- Berger, J.-L. & Karabenick, S. A. (2011). Motivation and students' use of learning strategies: Evidence of unidirectional effects in mathematics classrooms. *Learning and Instruction*, 21(3), 416–428.
- Bergin, D. A. (1992). Leisure Activity, Motivation, and Academic Achievement in High School Students. *Journal of Leisure Research*, 24(3), 225–239.
- Bergin, D. A. (1996). Adolescents' Out-of-School Learning Strategies. *The Journal of Experimental Education*, 64(4), 309–323.

- Berland, L. K. & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, 94(5), 765–793.
- Bernacki, M. L. & Walkington, C. A. (2018). The role of situational interest in personalized learning. *Journal of Educational Psychology*, 110(6), 864–881.
- Bernhold, S., Neumann, K. & Nentwig, P. Making it Tangible – Specifying Learning Outcomes in Science Education (S. Bernhold, K. Neumann & P. Nentwig, Hrsg.). In: *In Learning outcomes in science education* (S. Bernhold, K. Neumann & P. Nentwig, Hrsg.). Hrsg. von Bernhold, S., Neumann, K. & Nentwig, P. Münster: Waxmann, 2012, S. 13–27. ISBN: 978-3-8309-2644-3.
- Bernholt, S., Höft, L. & Parchmann, I. (2019). Die Entwicklung fachlicher Basiskonzepte im Chemieunterricht – Findet ein kumulativer Aufbau im Kompetenzbereich Fachwissen statt? *Unterrichtswissenschaft*, 47(1), 151.
- Bernholt, S. & Sevian, H. (2018). Learning progressions and teaching sequences – old wine in new skins? *Chem. Educ. Res. Pract.*, 19(4), 989–997.
- Björkman, J. & Tiemann, R. (2013). Teaching Patterns of Scientific Inquiry: A Video Study of Chemistry Lessons in Germany and Sweden. *Science Education Review Letters*, 1–7.
- Blankenburg, J. S. (2015). Modellbasierte Entwicklung und Evaluation eines naturwissenschaftlichen Projekttages zur Heranführung an Schülerwettbewerbe (Dissertation). Christian-Albrechts-Universität zu Kiel. Kiel.
- Blankenburg, J. S., Höffler, T. N. & Parchmann, I. (2015). Naturwissenschaftliche Wettbewerbe – Was kann junge Schülerinnen und Schüler zur Teilnahme motivieren? *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 141–153.
- Blankenburg, J. S., Höffler, T. N. & Parchmann, I. (2016). Fostering Today What is Needed Tomorrow: Investigating Students' Interest in Science. *Science Education*, 100(2), 364–391.
- Blankenburg, J. S., Höffler, T. N., Peters, H. & Parchmann, I. (2016). The effectiveness of a project day to introduce sixth grade students to science competitions. *Research in Science & Technological Education*, 34(3), 342–358.
- Bloom, H. S., Hill, C. J., Black, A. R. & Lipsey, M. W. (2008). Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions. *Journal of Research on Educational Effectiveness*, 1(4), 289–328.
- Bøe, M. V. (2012). Science choices in Norwegian upper secondary school: What matters? *Science Education*, 96(1), 1–20.
- Brandtner, M. & Hertel, S. (2018). Naturwissenschaftlich interessierte Äußerungen 4- bis 6-jähriger Kinder. *Zeitschrift für Didaktik der Naturwissenschaften*, 96(1), 763.
- Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000). *How people learn: Brain, mind, experience and school*. National Academy Press.
- Briggs, D. C., Alonzo, A. C., Schwab, C. & Wilson, M. (2006). Diagnostic Assessment With Ordered Multiple-Choice Items. *Educational Assessment*, 11(1), 33–63.
- Broman, K., Bernholt, S. & Parchmann, I. (2018). Using model-based scaffolds to support students solving context-based chemistry problems. *International Journal of Science Education*, 40(10), 1176–1197.
- Broman, K. & Simon, S. (2015). Upper Secondary School Students' Choice and Their Ideas on How to Improve Chemistry Education. *International Journal of Science and Mathematics Education*, 13(6), 1255–1278.
- Brown, A. L. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist*, 52(4), 399–413.
- Brown, E. R., Smith, J. L., Thoman, D. B., Allen, J. M. & Muragishi, G. (2015). From Bench to Bedside: A communal utility value intervention to enhance students' biomedical science motivation. *Journal of Educational Psychology*, 107(4), 1116–1135.
- Browne, M. W. & Cudeck, R. Alternative ways of testing structural equation models (K. A. Bollen & S. J. Long, Hrsg.). In: *In Testing structural equation models* (K. A. Bollen & S. J. Long, Hrsg.). Hrsg. von Bollen, K. A. & Long, S. J. Sage focus editions. Newbury Park, Calif.: Sage, 1993, S. 136–162. ISBN: 9780803945074.
- Brunner, M., Nagy, G. & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of personality*, 80(4), 796–846.
- Buff, A. (2014). Enjoyment of learning and its personal antecedents: Testing the change–change assumption of the control-value theory of achievement emotions. *Learning and Individual Differences*, 31, 21–29.
- Bybee, R. W. Toward an Understanding of Scientific Literacy (W. Gräber & C. Bolte, Hrsg.). In: *In Scientific literacy* (W. Gräber & C. Bolte, Hrsg.). Hrsg. von Gräber, W. & Bolte, C. IPN. Kiel: IPN, 1997, S. 37–68. ISBN: 3-89088-109-2.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge Taylor & Francis Group.
- Canning, E. A., Priniski, S. J. & Harackiewicz, J. M. (2019). Unintended consequences of framing a utility-value intervention in two-year colleges. *Learning and Instruction*, 62, 37–48.
- Cerasoli, C. P., Nicklin, J. M. & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, 140(4), 980–1008.
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504.
- Cheung, D. (2017). The key factors affecting students' individual interest in school science lessons. *International Journal of Science Education*, 30(4), 1–23.
- Cheung, G. W. & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255.

- Chi, M. T. H., Feltovich, P. J. & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, 5, 121–152.
- Chopyak, J. & Levesque, P. (2002). Public participation in science and technology decision making: Trends for the future. *Technology in Society*, 24(1-2), 155–166.
- Ciani, K., Ferguson, Y., Bergin, D. A. & Hilpert, J. (2010). Motivational influences on school-prompted interest. *Educational Psychology*, 30(4), 377–393.
- Claesgens, J., Scalise, K., Wilson, M. & Stacy, A. M. (2009). Mapping student understanding in chemistry: The Perspectives of Chemists. *Science Education*, 93(1), 56–85.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Taylor and Francis.
- Cordova, J. R., Sinatra, G. M., Jones, S. H., Taasobshirazi, G. & Lombardi, D. (2014). Confidence in prior knowledge, self-efficacy, interest and prior knowledge: Influences on conceptual change. *Contemporary Educational Psychology*, 39(2), 164–174.
- Cribbs, J. D., Hazari, Z., Sonnert, G. & Sadler, P. M. (2015). Establishing an Explanatory Model for Mathematics Identity. *Child development*, 86(4), 1048–1062.
- Daniels, L. M., Haynes, T. L., Stupnisky, R. H., Perry, R. P., Newall, N. E. & Pekrun, R. (2008). Individual differences in achievement goals: A longitudinal study of cognitive, emotional, and achievement outcomes. *Contemporary Educational Psychology*, 33(4), 584–608.
- Daniels, L. M., Stupnisky, R. H., Pekrun, R., Haynes, T. L., Perry, R. P. & Newall, N. E. (2009). A longitudinal analysis of achievement goals: From affective antecedents to emotional effects and achievement outcomes. *Journal of Educational Psychology*, 101(4), 948–963.
- Daniels, Z. (2008). *Entwicklung schulischer Interessen im Jugendalter: Zugl.: Potsdam, Berlin, Univ., Diss., 2004 (Bd. 69)*. Waxmann.
- DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, 37(6), 582–601.
- Deci, E. L. & Ryan, R. M. (1985a). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US.
- Deci, E. L. & Ryan, R. M. (1985b). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2), 109–134.
- Denissen, J. J. A., Zarrett, 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.
- Dierks, P. O. (2014). *Das Interesse an Naturwissenschaften: Entwicklung eines RIASEC-basierten Interessenmodells zur differenzierten Analyse von Interessenprofilen bei Jugendlichen*.
- Dierks, P. O., Höffler, T. N., Blankenburg, J. S., Peters, H. & Parchmann, I. (2016). Interest in science: A RIASEC-based analysis of students' interests. *International Journal of Science Education*, 38(2), 238–258.
- Dierks, P. O., Höffler, T. N. & Parchmann, I. (2014a). Interesse von Jugendlichen an Naturwissenschaften. *CHEMKON*, 21(3), 111–116.
- Dierks, P. O., Höffler, T. N. & Parchmann, I. (2014b). Profiling interest of students in science: Learning in school and beyond. *Research in Science & Technological Education*, 32(2), 97–114.
- Ding, N. & Harskamp, E. G. (2011). Collaboration and Peer Tutoring in Chemistry Laboratory Education. *International Journal of Science Education*, 33(6), 839–863.
- diSessa, A. & Wagner, J. What Coordination has to say about Transfer (J. P. Mestre, Hrsg.). In: *In Transfer of learning from a modern multidisciplinary perspective* (J. P. Mestre, Hrsg.). Hrsg. von Mestre, J. P. Greenwich, CT: Information Age Publishing, 2005, S. 121–154. ISBN: 9781593111656.
- Dotterer, A. M., McHale, S. M. & Crouter, A. C. (2009). The Development and Correlates of Academic Interests from Childhood through Adolescence. *Journal of Educational Psychology*, 101(2), 509–519.
- Driver, R., Leach, J., Scott, P. & Wood-Robinson, C. (1994). Young People's understanding of science concepts: Implications of cross-age studies for curriculum planning. *Studies in Science Education*, 24(1), 75–100.
- Duit, R., Treagust, D. F. & Widodo, A. Teaching science for conceptual change: Theory and practice (S. Vosniadou, Hrsg.). In: *In International handbook of research on conceptual change* (S. Vosniadou, Hrsg.). Hrsg. von Vosniadou, S. Educational psychology handbook series. New York: Routledge, 2008, S. 629–646. ISBN: 0805860444.
- Duncan, R. G., Rogat, A. D. & Yarden, A. (2009). A learning progression for deepening students' understandings of modern genetics across the 5th-10th grades. *Journal of Research in Science Teaching*, 46(6), 655–674.
- Durik, A. M., Hulleman, C. S. & Harackiewicz, J. M. One Size Fits Some: Instructional Enhancements to Promote Interest (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). In: *In Interest in Mathematics and Science Learning* (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). Hrsg. von Renninger, K. A., Nieswandt, M. & Hidi, S. Washington, D.C.: American Educational Research Association, 2015, S. 49–62. ISBN: 9780935302424.
- Durik, A. M., Shechter, O. G., Noh, M., Rozek, C. S. & Harackiewicz, J. M. (2015). What if I can't? Success expectancies moderate the effects of utility value information on situational interest and performance. *Motivation and Emotion*, 39(1), 104–118.
- Duschl, R. A. & Grandy, R. E. Demarcation in science education: Toward an enhanced view of scientific method (R. S. Taylor & M. Ferrari, Hrsg.). In: *In Epistemology and science education* (R. S. Taylor & M. Ferrari, Hrsg.). Hrsg. von Taylor, R. S. & Ferrari, M. New York, N.Y.: Routledge, 2011, S. 3–19. ISBN: 9780203839638.
- Duschl, R. A., Maeng, S. & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education*, 47(2), 123–182.
- Duschl, R. A., Schweingruber, H. A. & Shouse, A. W. (2007). *Taking Science to School*. National Academies Press.

- Eccles, J. S. (2009). Who Am I and What Am I Going to Do With My Life? Personal and Collective Identities as Motivators of Action. *Educational Psychologist*, 44(2), 78–89.
- Eccles, J. S., Fredricks, J. A. & Epstein, A. Understanding Well-Developed Interests and Activity Commitment (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). In: *In Interest in Mathematics and Science Learning* (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). Hrsg. von Renninger, K. A., Nieswandt, M. & Hidi, S. Washington, D.C.: American Educational Research Association, 2015, S. 315–330. ISBN: 9780935302424.
- Eccles, J. S. & Jacobs, J. E. (1986). Social Forces Shape Math Attitudes and Performance. *Signs: Journal of Women in Culture and Society*, 11(2), 367–380.
- Eccles, J. S. & Midgley, C. Stage-Environment Fit: Developmentally Appropriate Classrooms for Young Adolescents (C. Ames & R. Ames, Hrsg.). In: *In Research on Motivation in Education* (C. Ames & R. Ames, Hrsg.). Hrsg. von Ames, C. & Ames, R. San Diego, California: Acad. Press, 1989, S. 139–186. ISBN: 0-12-056703-2.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C. & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, 48(2), 90–101.
- Eccles, J. S. & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology*, 53, 109–132.
- Eckes, A., Großmann, N. & Wilde, M. (2018). Studies on the effects of structure in the context of autonomy-supportive or controlling teacher behavior on students' intrinsic motivation. *Learning and Individual Differences*, 62, 69–78.
- Eitemüller, C. & Walpuski, M. (2018). Wahl- und Abwahlprofile im Fach Chemie: Ergebnisse einer Clusteranalyse zur Charakterisierung von Lernenden am Ende der Sekundarstufe I. *Zeitschrift für Didaktik der Naturwissenschaften*, 96(1), 83.
- Emden, M., Weber, K. & Sumfleth, E. (2018). Evaluating a learning progression on 'Transformation of Matter' on the lower secondary level. *Chem. Educ. Res. Pract.*, 19(4), 1096–1116.
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.
- Etzel, J. M. & Nagy, G. (2015). Students' Perceptions of Person–Environment Fit. *Journal of Career Assessment*, 24(2), 270–288.
- Fan, X. (1997). Canonical correlation analysis and structural equation modeling: What do they have in common? *Structural Equation Modeling: A Multidisciplinary Journal*, 4(1), 65–79.
- Fechner, S. (2009). Effects of context-oriented learning on student interest and achievement in chemistry education: Zugl.: Duisburg-Essen, Univ., Diss., 2009 (Bd. 95). Logos-Verl.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock 'n' roll* (4th edition). Sage.
- Fink, B. Interessenentwicklung im Kindesalter aus der Sicht einer Person-Gegenstands-Konzeption (A. Krapp & M. Prenzel, Hrsg.). In: *In Interesse, Lernen, Leistung* (A. Krapp & M. Prenzel, Hrsg.). Hrsg. von Krapp, A. & Prenzel, M. Arbeiten zur sozialwissenschaftlichen Psychologie. Münster: Aschendorff, 1992, S. 51–83. ISBN: 3-402-04596-6.
- Flunger, B., Mayer, A. & Umbach, N. (2019). Beneficial for some or for everyone? Exploring the effects of an autonomy-supportive intervention in the real-life classroom. *Journal of Educational Psychology*, 111(2), 210–234.
- Fredricks, J. A., Hofkens, T., Wang, M.-T., Mortenson, E. & Scott, P. (2018). Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *Journal of Research in Science Teaching*, 55(2), 271–298.
- Frenzel, A. C., Becker-Kurz, B., Pekrun, R., Goetz, T. & Lüdtke, O. (2017). Emotion Transmission in the Classroom Revisited: A Reciprocal Effects Model of Teacher and Student Enjoyment. *Journal of Educational Psychology*.
- Frenzel, A. C., Goetz, T., Pekrun, R. & Watt, H. M. G. (2010). Development of Mathematics Interest in Adolescence: Influences of Gender, Family, and School Context. *Journal of Research on Adolescence*, 20(2), 507–537.
- Frenzel, A. C., Pekrun, R., Dicke, A.-L. & Goetz, T. (2012). Beyond quantitative decline: Conceptual shifts in adolescents' development of interest in mathematics. *Developmental psychology*, 48(4), 1069–1082.
- Fryer, L. K. & Ainley, M. (2019). Supporting interest in a study domain: A longitudinal test of the interplay between interest, utility-value, and competence beliefs. *Learning and Instruction*, 60, 252–262.
- Fryer, L. K., Ainley, M. & Thompson, A. (2016). Modelling the links between students' interest in a domain, the tasks they experience and their interest in a course: Isn't interest what university is all about? *Learning and Individual Differences*, 50, 157–165.
- Ganley, C. M. & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences*, 47, 182–193.
- Garon-Carrier, G., Boivin, M., Guay, F., Kovas, Y., Dionne, G., Lemelin, J.-P., Seguin, J. R., Vitaro, F. & Tremblay, R. E. (2016). Intrinsic Motivation and Achievement in Mathematics in Elementary School: A Longitudinal Investigation of Their Association. *Child development*, 87(1), 165–175.
- Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B. & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental psychology*, 51(9), 1226–1240.
- Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U. & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology*, 107(3), 663–677.
- Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U. & Nagengast, B. (2017). Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology*, 48, 67–84.

- Gaspard, H., Wigfield, A., Jiang, Y., Nagengast, B., Trautwein, U. & Marsh, H. W. (2018). Dimensional comparisons: How academic track students' achievements are related to their expectancy and value beliefs across multiple domains. *Contemporary Educational Psychology*, 52, 1–14.
- Goetz, T., Frenzel, A. C., Hall, N. C. & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology*, 33(1), 9–33.
- Gräber, W. & Bolte, C. (Hrsg.). (1997). *Scientific literacy: An international symposium* (Bd. 154). IPN.
- Grigg, S., Perera, H. N., McIlveen, P. & Svetleff, Z. (2018). Relations among math self efficacy, interest, intentions, and achievement: A social cognitive perspective. *Contemporary Educational Psychology*, 53, 73–86.
- Grimm, K. J., An, Y., McArdle, J. J., Zonderman, A. B. & Resnick, S. M. (2012). Recent Changes Leading to Subsequent Changes: Extensions of Multivariate Latent Difference Score Models. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(2), 268–292.
- Guay, F., Ratelle, C. F. & Chanal, J. (2008). Optimal learning in optimal contexts: The role of self-determination in education. *Canadian Psychology/Psychologie canadienne*, 49(3), 233–240.
- Guo, J., Marsh, H. W., Morin, A. J. S., Parker, P. D. & Kaur, G. (2015). Directionality of the Associations of High School Expectancy-Value, Aspirations, and Attainment: A Longitudinal Study. *American Educational Research Journal*, 52(2), 371–402.
- Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S. & Yeung, A. S. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences*, 37, 161–168.
- Guo, J., Nagengast, B., Marsh, H. W., Kelava, A., Gaspard, H., Brandt, H., Cambria, J., Flunger, B., Dicke, A.-L., Häfner, I., Brisson, B. & Trautwein, U. (2016). Probing the Unique Contributions of Self-Concept, Task Values, and Their Interactions Using Multiple Value Facets and Multiple Academic Outcomes. *AERA Open*, 2(1), 233285841562688.
- Haberman, S. J. (2009). Linking Parameter Estimates Derived From an Item Response Model Through Separate Calibrations. *ETS Research Report Series*, 2009(2), i–9.
- Habig, S., Blankenburg, J. S., van Vorst, H., Fechner, S., Parchmann, I. & Sumfleth, E. (2018). Context characteristics and their effects on students' situational interest in chemistry. *International Journal of Science Education*, 40(10), 1154–1175.
- Habig, S., van Vorst, H. & Sumfleth, E. (2018). Merkmale kontextualisierter Lernaufgaben und ihre Wirkung auf das situationale Interesse und die Lernleistung von Schülerinnen und Schülern. *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 99–114.
- Hadenfeldt, J. C., Bernholt, S., Liu, X., Neumann, K. & Parchmann, I. (2013). Using Ordered Multiple-Choice Items To Assess Students' Understanding of the Structure and Composition of Matter. *Journal of Chemical Education*, 90(12), 1602–1608.
- Hadenfeldt, J. C., Liu, X. & Neumann, K. (2014). Framing students' progression in understanding matter: A review of previous research. *Studies in Science Education*, 50(2), 181–208.
- Hadenfeldt, J. C., Neumann, K., Bernholt, S., Liu, X. & Parchmann, I. (2016). Students' progression in understanding the matter concept. *Journal of Research in Science Teaching*, 53(5), 683–708.
- Hadenfeldt, J. C., Repenning, B. & Neumann, K. (2014). Die kognitive Validität von Ordered Multiple Choice Aufgaben zur Erfassung des Verständnisses von Materie. *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), 57–68.
- Häfner, I., Flunger, B., Dicke, A.-L., Gaspard, H., Brisson, B. M., Nagengast, B. & Trautwein, U. (2017). The Role of Family Characteristics for Students' Academic Outcomes: A Person-Centered Approach. *Child development*.
- Harackiewicz, J. M., Barron, K. E., Tauer, J. M. & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, 94(3), 562–575.
- Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L. & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology*, 100(1), 105–122.
- Harackiewicz, J. M., Smith, J. L. & Priniski, S. J. (2016). Interest Matters. *Policy Insights from the Behavioral and Brain Sciences*, 3(2), 220–227.
- Harks, B., Rakoczy, K., Hattie, J., Besser, M. & Klieme, E. (2013). The effects of feedback on achievement, interest and self-evaluation: The role of feedback's perceived usefulness. *Educational Psychology*, 34(3), 269–290.
- Harris, C. J., Penuel, W. R., D'Angelo, C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., Cheng, B. H. & Krajcik, J. S. (2015). Impact of project-based curriculum materials on student learning in science: Results of a randomized controlled trial. *Journal of Research in Science Teaching*, 52(10), 1362–1385.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112.
- Häussler, P. & Hoffmann, L. (2000). A curricular frame for physics education: Development, comparison with students' interests, and impact on students' achievement and self-concept. *Science Education*, 84(6), 689–705.
- Häussler, P., Hoffmann, L., Langeheine, R., Rost, J. & Sievers, K. (1998). A typology of students' interest in physics and the distribution of gender and age within each type. *International Journal of Science Education*, 20(2), 223–238.
- Hayenga, A. O. & Corpus, J. H. (2010). Profiles of intrinsic and extrinsic motivations: A person-centered approach to motivation and achievement in middle school. *Motivation and Emotion*, 34(4), 371–383.
- Heller, K. A. & Perleth, C. (2000). *KFT 4 – 12 + R: Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision ; Manual*. Beltz Test.

- Herrmann-Abell, C. F. & DeBoer, G. E. (2018). Investigating a learning progression for energy ideas from upper elementary through high school. *Journal of Research in Science Teaching*, 55(1), 68–93.
- Hidi, S. (2001). Interest, Reading, and Learning: Theoretical and Practical Considerations. *Educational Psychology Review*, 13(3), 191–209.
- Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review*, 1(2), 69–82.
- Hidi, S. (2016). Revisiting the Role of Rewards in Motivation and Learning: Implications of Neuroscientific Research. *Educational Psychology Review*, 28(1), 61–93.
- Hidi, S. & Harackiewicz, J. M. (2000). Motivating the Academically Unmotivated: A Critical Issue for the 21st Century. *Review of Educational Research*, 70(2), 151–179.
- Hidi, S. & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, 41(2), 111–127.
- Hill, N. E. & Tyson, D. F. (2009). Parental involvement in middle school: A meta-analytic assessment of the strategies that promote achievement. *Developmental Psychology*, 45(3), 740–763.
- Höffler, T. N., Köhler, C. & Parchmann, I. (2019). Scientists of the future: An analysis of talented students' interests. *International Journal of STEM Education*, 6(1), 726.
- Hoffmann, L., Krapp, A., Renninger, K. A. & Baumert, J. (Hrsg.). (1998). *Interest and learning: Proceedings of the Seon Conference on Interest and Gender* (Bd. 164). IPN.
- Hofstein, A. & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, 88(1), 28–54.
- Höft, L. & Bernholt, S. (in Rev.). Probing Precursors of Students' Long-term Engagement in Chemistry: Individual Interest and Achievement Development and Their Relation with Utility Values.
- Höft, L. & Bernholt, S. (2019a). Ich mag, was ich kann oder kann ich, was ich mag? Über das Zusammenspiel von Interesse, Freude und Konzeptwissen im Fach Chemie. *Zeitschrift für Didaktik der Naturwissenschaften*, 25(1), 161–180.
- Höft, L. & Bernholt, S. (2019b). Longitudinal couplings between interest and conceptual understanding in secondary school chemistry: An activity-based perspective. *International Journal of Science Education*, 19(4), 1–21.
- Höft, L., Bernholt, S. & Blankenburg, J. S. Zusammenhänge zwischen Interessens- und Verständnissentwicklung (C. Maurer, Hrsg.). In: *In Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis* (C. Maurer, Hrsg.). Hrsg. von Maurer, C. Universität Regensburg, 2017, S. 596–599.
- Höft, L., Bernholt, S., Blankenburg, J. S. & Winberg, M. (2019). Knowing more about things you care less about: Cross-sectional analysis of the opposing trend and interplay between conceptual understanding and interest in secondary school chemistry. *Journal of Research in Science Teaching*, 56(2), 184–210.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3. ed.). Psychological Assessment Resources.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65–70.
- Holstermann, N., Grube, D. & Bögeholz, S. (2010). Hands-on Activities and Their Influence on Students' Interest. *Research in Science Education*, 40(5), 743–757.
- Hosbein, K. N. & Barbera, J. (2020). Alignment of theoretically grounded constructs for the measurement of science and chemistry identity. *Chem. Educ. Res. Pract.*
- Hu, L.-.-t. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Hulleman, C. S., Godes, O., Hendricks, B. L. & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*, 102(4), 880–895.
- Jang, H., Reeve, J. & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology*, 102(3), 588–600.
- Jansen, M., Lüdtke, O. & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology*, 46, 116–127.
- Jansen, M., Schroeders, U., Lüdtke, O. & Marsh, H. W. (2015). Contrast and assimilation effects of dimensional comparisons in five subjects: An extension of the I/E model. *Journal of Educational Psychology*, 107(4), 1086–1101.
- Jansen, M., Schroeders, U., Lüdtke, O. & Marsh, H. W. (2019). The dimensional structure of students' self-concept and interest in science depends on course composition. *Learning and Instruction*, 60, 20–28.
- Jiang, Y., Rosenzweig, E. Q. & Gaspard, H. (2018). An expectancy-value-cost approach in predicting adolescent students' academic motivation and achievement. *Contemporary Educational Psychology*, 54, 139–152.
- Jocz, J. A., Zhai, J. & Tan, A. L. (2014). Inquiry Learning in the Singaporean Context: Factors affecting student interest in school science. *International Journal of Science Education*, 36(15), 2596–2618.
- Jögi, A.-L., Kikas, E., Lerkkanen, M.-K. & Mägi, K. (2015). Cross-lagged relations between math-related interest, performance goals and skills in groups of children with different general abilities. *Learning and Individual Differences*, 39, 105–113.
- Johnson, M. L. & Sinatra, G. M. (2013). Use of task-value instructional inductions for facilitating engagement and conceptual change. *Contemporary Educational Psychology*, 38(1), 51–63.
- Joint Committee. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Keller, S. & Reintjes, C. (Hrsg.). (2016). *Aufgaben als Schlüssel zur Kompetenz: Didaktische Herausforderungen, wissenschaftliche Zugänge und empirische Befunde* (1. Aufl.). Waxmann Verlag GmbH.
- Kiefer, T., Robitzsch, A. & Wu, M. (2017). TAM: Test Analysis Modules.

- Kim, S.-I., Jiang, Y. & Song, J. The Effects of Interest and Utility Value on Mathematics Engagement and Achievement (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). In: *Interest in Mathematics and Science Learning* (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). Hrsg. von Renninger, K. A., Nieswandt, M. & Hidi, S. Washington, D.C.: American Educational Research Association, 2015, S. 63–78. ISBN: 9780935302424.
- Kingston, N. & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice*, 30(4), 28–37.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological review*, 95(2), 163–182.
- Knogler, M. Situational Interest: A Proposal to Enhance Conceptual Clarity (P. A. O’Keefe & J. M. Harackiewicz, Hrsg.). In: *In The Science of Interest* (P. A. O’Keefe & J. M. Harackiewicz, Hrsg.). Hrsg. von O’Keefe, P. A. & Harackiewicz, J. M. Cham: Springer International Publishing, 2017, S. 109–124. ISBN: 978-3-319-55507-2.
- Knogler, M., Harackiewicz, J. M., Gegenfurtner, A. & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology*, 43, 39–50.
- Köller, O., Baumert, J. & Schnabel, K. U. (2001). Does Interest Matter? The Relationship between Academic Interest and Achievement in Mathematics. *Journal for Research in Mathematics Education*, 32(5), 448.
- 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.
- Krajcik, J. S., Sutherland, L. M., Drago, K. & Merritt, J. The Promise and Value of Learning Progression Research (S. Bernhold, K. Neumann & P. Nentwig, Hrsg.). In: *In Learning outcomes in science education* (S. Bernhold, K. Neumann & P. Nentwig, Hrsg.). Hrsg. von Bernhold, S., Neumann, K. & Nentwig, P. Münster: Waxmann, 2012, S. 283–306. ISBN: 978-3-8309-2644-3.
- Krapp, A. Das Interessenkonstrukt. Bestimmungsmerkmale der Interessenhandlung und des individuellen Interesses aus der Sicht einer Person-Gegenstands-Konzeption (A. Krapp & M. Prenzel, Hrsg.). In: *In Interesse, Lernen, Leistung* (A. Krapp & M. Prenzel, Hrsg.). Hrsg. von Krapp, A. & Prenzel, M. Arbeiten zur sozialwissenschaftlichen Psychologie. Münster: Aschendorff, 1992, S. 297–330. ISBN: 3-402-04596-6.
- Krapp, A. (1992b). Interesse, Lernen und Leistung. Neue Forschungsansätze in der Pädagogischen Psychologie. *Zeitschrift für Pädagogik*, 38(5), 747–770.
- Krapp, A. Konzepte und Forschungsansätze zur Analyse des Zusammenhangs von Interesse, Lernen und Leistung (A. Krapp & M. Prenzel, Hrsg.). In: *In Interesse, Lernen, Leistung* (A. Krapp & M. Prenzel, Hrsg.). Hrsg. von Krapp, A. & Prenzel, M. Arbeiten zur sozialwissenschaftlichen Psychologie. Münster: Aschendorff, 1992, S. 9–52. ISBN: 3-402-04596-6.
- Krapp, A. (1999a). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education*, 14(1), 23–40.
- Krapp, A. (1999b). Intrinsische Lernmotivation und Interesse. *Forschungsansätze und konzeptuelle Überlegungen. Zeitschrift für Pädagogik*, 45(3), 387–406.
- Krapp, A. Individuelle Interessen als Bedingung lebenslangen Lernens (F. Achtenhagen & W. Lempert, Hrsg.). In: *In Lebenslanges Lernen im Beruf — seine Grundlegung im Kindes- und Jugendalter* (F. Achtenhagen & W. Lempert, Hrsg.). Hrsg. von Achtenhagen, F. & Lempert, W. Wiesbaden: VS Verlag für Sozialwissenschaften, 2000, S. 54–75. ISBN: 978-3-8100-2749-8.
- Krapp, A. An Education-Psychological Theory of Interest and its Relation to SDT (E. L. Deci & R. M. Ryan, Hrsg.). In: *In Handbook of self-determination research* (E. L. Deci & R. M. Ryan, Hrsg.). Hrsg. von Deci, E. L. & Ryan, R. M. Rochester, NY: Univ. of Rochester Press, 2002, S. 405–427. ISBN: 9781580461085.
- Krapp, A. (2002b). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12(4), 383–409.
- Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instruction*, 15(5), 381–395.
- Krapp, A., Hidi, S. & Renninger, K. A. Interest, learning, and development (K. A. Renninger, A. Krapp & S. Hidi, Hrsg.). In: *In The Role of interest in Learning and Development* (K. A. Renninger, A. Krapp & S. Hidi, Hrsg.). Hrsg. von Renninger, K. A., Krapp, A. & Hidi, S. Hillsdale, N.J.: L. Erlbaum Associates, 1992, S. 3–25. ISBN: 9781317783534.
- Krapp, A. & Prenzel, M. (Hrsg.). (1992). *Interesse, Lernen, Leistung: Neuere Ansätze der pädagogisch-psychologischen Interessenforschung* (Bd. 26). Aschendorff.
- Krapp, A. & Prenzel, M. (2011). Research on Interest in Science: Theories, methods, and findings. *International Journal of Science Education*, 33(1), 27–50.
- Krapp, A., Schiefele, U., Wild, K.-P. & Winteler, A. (1993). Der Fragebogen zum Studieninteresse (FSI). *Diagnostika*, 39(4), 335–351.
- Kriegbaum, K., Becker, N. & Spinath, B. (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, 25, 120–148.
- Kriegbaum, K., Jansen, M. & Spinath, B. (2015). Motivation: A predictor of PISA’s mathematical competence beyond intelligence and prior test achievement. *Learning and Individual Differences*, 43, 140–148.

- Kriegbaum, K. & Spinath, B. (2016). Explaining Social Disparities in Mathematical Achievement: The Role of Motivation. *European Journal of Personality*, 30(1), 45–63.
- Kühn, S. M. Aufgaben in (zentralen) Abschlussprüfungen. Theoretische und empirische Perspektiven auf ein interdisziplinäres Forschungsfeld (S. Keller & C. Reintjes, Hrsg.). In: In Aufgaben als Schlüssel zur Kompetenz (S. Keller & C. Reintjes, Hrsg.). Hrsg. von Keller, S. & Reintjes, C. Münster, New York: Waxmann Verlag GmbH, 2016, S. 73–92. ISBN: 9783830984214.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht* (Zugl.: Berlin, Freie Univ., Diss., 2004). Waxmann. Münster, New York, München, Berlin.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820.
- Lau, K.-c. & Lam, T. Y.-p. (2017). Instructional practices and science performance of 10 top-performing regions in PISA 2015. *International Journal of Science Education*, 39(15), 2128–2149.
- Lauermann, F., Tsai, Y.-M. & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy-value theory of achievement-related behaviors. *Developmental psychology*, 53(8), 1540–1559.
- Lazarides, R., Dietrich, J. & Taskinen, P. H. (2019). Stability and change in students' motivational profiles in mathematics classrooms: The role of perceived teaching. *Teaching and Teacher Education*, 79, 164–175.
- Lazarides, R. & Raufelder, D. (2017). Longitudinal Effects of Student-Perceived Classroom Support on Motivation – A Latent Change Model. *Frontiers in psychology*, 8, 417.
- Lazowski, R. A. & Hulleman, C. S. (2016). Motivation Interventions in Education. *Review of Educational Research*, 86(2), 602–640.
- LeBreton, J. M. & Senter, J. L. (2008). Answers to 20 Questions About Interrater Reliability and Interrater Agreement. *Organizational Research Methods*, 11(4), 815–852.
- Lee, W., Lee, M.-J. & Bong, M. (2014). Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemporary Educational Psychology*, 39(2), 86–99.
- Lent, R. W., Brown, S. D. & Hackett, G. (1994). Toward a Unifying Social Cognitive Theory of Career and Academic Interest, Choice, and Performance. *Journal of Vocational Behavior*, 45(1), 79–122.
- Leucht, M., Kampa, N. & Köller, O. (Hrsg.). (2016). *Fachleistungen von Abiturienten: Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein : LISA 6*. Waxmann.
- Lewalter, D., Wild, K.-P. & Krapp, A. Interessenentwicklung in der beruflichen Ausbildung (K. Beck & V. Krumm, Hrsg.). In: In Lehren und Lernen in der beruflichen Erstausbildung (K. Beck & V. Krumm, Hrsg.). Hrsg. von Beck, K. & Krumm, V. Wiesbaden: VS Verlag für Sozialwissenschaften, 2001, S. 11–35. ISBN: 978-3-8100-3055-9.
- Lin, H.-S., Hong, Z.-R. & Chen, Y.-C. (2013). Exploring the Development of College Students' Situational Interest in Learning Science. *International Journal of Science Education*, 35(13), 2152–2173.
- Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A. & Harackiewicz, J. M. (2010). Measuring Situational Interest in Academic Domains. *Educational and Psychological Measurement*, 70(4), 647–671.
- Linnenbrink-Garcia, L., Patall, E. A. & Messersmith, E. E. (2013). Antecedents and consequences of situational interest. *The British journal of educational psychology*, 83(Pt 4), 591–614.
- Linnenbrink-Garcia, L., Wormington, S. V., Snyder, K. E., Riggsbee, J., Perez, T., Ben-Eliyahu, A. & Hill, N. E. (2018). Multiple Pathways to Success: An Examination of Integrative Motivational Profiles Among Upper Elementary and College Students. *Journal of Educational Psychology*, 110(7), 1026–1048.
- Liu, X. Using Learning Progression to Organize Learning Outcomes: Implications for Assessment (S. Bernhold, K. Neumann & P. Nentwig, Hrsg.). In: In Learning outcomes in science education (S. Bernhold, K. Neumann & P. Nentwig, Hrsg.). Hrsg. von Bernhold, S., Neumann, K. & Nentwig, P. Münster: Waxmann, 2012, S. 309–325. ISBN: 978-3-8309-2644-3.
- Liu, X. & Lesniak, K. M. (2005). Students' progression of understanding the matter concept from elementary to high school. *Science Education*, 89(3), 433–450.
- Löfgren, L. & Helldén, G. (2009). A Longitudinal Study Showing how Students use a Molecule Concept when Explaining Everyday Situations. *International Journal of Science Education*, 31(12), 1631–1655.
- Luce, M. R. & Hsi, S. (2015). Science-Relevant Curiosity Expression and Interest in Science: An Exploratory Study. *Science Education*, 99(1), 70–97.
- Lüdtke, O., Trautwein, U., Kunter, M. & Baumert, J. (2006). Analyse von Lernumwelten. *Zeitschrift für Pädagogische Psychologie*, 20(1/2), 85–96.
- Lüdtke, O., Trautwein, U., Kunter, M. & Baumert, J. (2007). Reliability and agreement of student ratings of the classroom environment: A reanalysis of TIMSS data. *Learning Environments Research*, 9(3), 215–230.
- MacCallum, R. C. & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual review of psychology*, 51, 201–226.
- Maltese, A. V. & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education*, 95(5), 877–907.
- Marsh, H. W. (1986). Verbal and Math Self-Concepts: An Internal/External Frame of Reference Model. *American Educational Research Journal*, 23(1), 129–149.
- Marsh, H. W., Balla, J. R. & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103(3), 391–410.



- Marsh, H. W., Lüdtke, O., Muthén, B. O., Asparouhov, T., Morin, A. J. S., Trautwein, U. & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological assessment*, 22(3), 471–491.
- Marsh, H. W., Lüdtke, O., Trautwein, U. & Morin, A. J. S. (2009). Classical Latent Profile Analysis of Academic Self-Concept Dimensions: Synergy of Person- and Variable-Centered Approaches to Theoretical Models of Self-Concept. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(2), 191–225.
- Marsh, H. W., Pekrun, R., Lichtenfeld, S., Guo, J., Arens, A. K. & Murayama, K. (2016). Breaking the double-edged sword of effort/trying hard: Developmental equilibrium and longitudinal relations among effort, achievement, and academic self-concept. *Developmental psychology*, 52(8), 1273–1290.
- 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.
- Meade, A. W., Johnson, E. C. & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *The Journal of applied psychology*, 93(3), 568–592.
- Menthe, J. & Parchmann, I. *Getting Involved: Context-Based Learning in Chemistry Education* (M. Kahveci & M. Orgill, Hrsg.). In: *In Affective Dimensions in Chemistry Education* (M. Kahveci & M. Orgill, Hrsg.). Hrsg. von Kahveci, M. & Orgill, M. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, S. 51–67. ISBN: 978-3-662-45084-0.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749.
- Messick, S. (1998). Test Validity: A Matter of Consequence. *Social Indicators Research*, 45(1/3), 35–44.
- Millar, R., Osborne, J. & Nott, M. (1998). Science education for the future. *School Science Review*, 80(291), 19–24.
- Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein. (2002). *Lehrplan für die Sekundarstufe II Gymnasium, Gesamtschule: Chemie* (Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein, Hrsg.).
- Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein. (2005). *Lehrplan für die Sekundarstufe I der weiterführenden allgemeinbildenden Schulen Hauptschule, Realschule, Gymnasium: Chemie* (Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein, Hrsg.).
- Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein. (2016). *Fachanforderungen Chemie: Allgemein bildende Schulen Sekundarstufe I – Gymnasium Sekundarstufe II* (Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein, Hrsg.).
- Minnaert, A., Boekaerts, M., de Brabander, C. & Opendakker, M.-C. (2011). Students' Experiences of Autonomy, Competence, Social Relatedness and Interest Within a CSCL Environment in Vocational Education: The Case of Commerce and Business Administration. *Vocations and Learning*, 4(3), 175–190.
- Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, 85(3), 424–436.
- Möller, K. (2014). Vom naturwissenschaftlichen Sachunterricht zum Fachunterricht – Der Übergang von der Grundschule in die weiterführende Schule. *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), 33–43.
- Morell, L., Collier, T., Black, P. & Wilson, M. (2017). A construct-modeling approach to develop a learning progression of how students understand the structure of matter. *Journal of Research in Science Teaching*, 54(8), 1024–1048.
- Muntoni, F. & Retelsdorf, J. (2019). At their children's expense: How parents' gender stereotypes affect their children's reading outcomes. *Learning and Instruction*, 60, 95–103.
- Murphy, K. R., Myron, B. & Wolach, A. (1998). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (2. ed.). L. Erlbaum.
- Muthén, B. O. & Satorra, A. (1995). Complex Sample Data in Structural Equation Modeling. *Sociological Methodology*, 25, 267.
- Muthén, L. K. & Muthén, B. O. (1998-2015). *Mplus User's Guide: Statistical Analysis with Latent Variables*. Muthén & Muthén.
- Nagy, G., Etzel, J. M. & Lüdtke, O. (2019). Integrating covariates into circumplex structures: An extension procedure for Browne's circular stochastic process model. *Multivariate behavioral research*, 54(3), 404–428.
- Nagy, G., Trautwein, U., Baumert, J., Köller, O. & Garrett, J. (2007). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, 12(4), 323–345.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academy Press.
- Nett, U. E., Bieg, M. & Keller, M. M. (2017). How Much Trait Variance Is Captured by Measures of Academic State Emotions? *European Journal of Psychological Assessment*, 33(4), 239–255.
- Neumann, K., Viering, T., Boone, W. J. & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, 50(2), 162–188.
- NGSS Lead States. (2013). *Next Generation Science Standards*. National Academies Press.
- Nylund-Gibson, K., Grimm, R., Quirk, M. & Furlong, M. (2014). A Latent Transition Mixture Model Using the Three-Step Specification. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(3), 439–454.
- OECD. (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. OECD Publishing and Centre for Educational Research and Innovation.
- OECD. (2007). *PISA 2006, science competencies for tomorrow's world* (Bd. 1). OECD.

- OECD. (2009). PISA 2006 Technical Report. OECD.
- OECD. (2011). PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (1. Aufl.). OECD.
- OECD. (2014). PISA 2012 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science (Bd. 1). OECD Publishing.
- OECD. (2016). Excellence and equity in education (1. Aufl.). OECD Publishing.
- Opitz, S. T. [S. T.], Neumann, K., Bernholt, S. & Harms, U. (2017). Students' Energy Understanding Across Biology, Chemistry, and Physics Contexts. *Research in Science Education*, 25(3), 209.
- Opitz, S. T. [Sebastian Tobias], Neumann, K., Bernholt, S. & Harms, U. (2017). How Do Students Understand Energy in Biology, Chemistry, and Physics? Development and Validation of an Assessment Instrument. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(7), 3019–3042.
- Osborne, J. & Collins, S. (2000). Pupils' & parents' views of the school science curriculum. Kings College.
- Osborne, J. & Dillon, J. (2008). Science education in Europe: Critical reflections (Bd. 13). London: The Nuffield Foundation.
- Overman, M., Vermunt, J. D., Meijer, P. C., Bulte, A. M. W. & Brekelmans, M. (2014). Students' Perceptions of Teaching in Context-based and Traditional Chemistry Classrooms: Comparing content, learning activities, and interpersonal perspectives. *International Journal of Science Education*, 36(11), 1871–1901.
- Palmer, D. H. (2009). Student interest generated during an inquiry skills lesson. *Journal of Research in Science Teaching*, 46(2), 147–165.
- Parchmann, I., Gräsel, C., Baer, A., Nentwig, P., Demuth, R. & Ralle, B. (2006). "Chemie im Kontext": A symbiotic implementation of a context-based teaching and learning approach. *International Journal of Science Education*, 28(9), 1041–1062.
- Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, 18(4), 315–341.
- Pekrun, R., Cusack, A., Murayama, K., Elliot, A. J. & Thomas, K. (2014). The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction*, 29, 115–124.
- Pekrun, R., Frenzel, A. C., Götz, T. & Perry, R. P. The control-value theory of achievement emotions: An integrative approach to emotions in education (P. A. Schutz & R. Pekrun, Hrsg.). In: *In Emotion in education* (P. A. Schutz & R. Pekrun, Hrsg.). Hrsg. von Schutz, P. A. & Pekrun, R. Amsterdam: Academic Press, 2007, S. 13–36.
- Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K. & Goetz, T. (2017). Achievement Emotions and Academic Performance: Longitudinal Models of Reciprocal Effects. *Child development*, 88(5), 1653–1670.
- Pellegrino, J. W. The Design of an Assessment System Focused on Student Achievement: A Learning Sciences Perspective on Issues of Competence, Growth, and Measurement (S. Bernhold, K. Neumann & P. Nentwig, Hrsg.). In: *In Learning outcomes in science education* (S. Bernhold, K. Neumann & P. Nentwig, Hrsg.). Hrsg. von Bernhold, S., Neumann, K. & Nentwig, P. Münster: Waxmann, 2012, S. 87–117. ISBN: 978-3-8309-2644-3.
- Pellegrino, J. W., Chudowsky, N. & Glaser, R. (Hrsg.). (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. National Academies Press.
- Pickering, M. (1980). Are lab courses a waste of time? *The Chronicle of Higher Education*, (19).
- Pinger, P., Rakoczy, K., Besser, M. & Klieme, E. (2016). Implementation of formative assessment – effects of quality of programme delivery on students' mathematics achievement and interest. *Assessment in Education: Principles, Policy & Practice*, 10(2), 1–23.
- Pinxten, M., Marsh, H. W., de Fraigne, B., van den Noortgate, W. & van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *The British journal of educational psychology*, 84(Pt 1), 152–174.
- Potvin, P. & Hasni, A. (2014a). Analysis of the Decline in Interest Towards School Science and Technology from Grades 5 Through 11. *Journal of Science Education and Technology*, 23(6), 784–802.
- Potvin, P. & Hasni, A. (2014b). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education*, 50(1), 85–129.
- Potvin, P., Hasni, A., Sy, O. & Riopel, M. (2018). Two Crucial Years of Science and Technology Schooling: A Longitudinal Study of the Major Influences on and Interactions Between Self-Concept, Interest, and the Intention to Pursue S&T. *Research in Science Education*, 100(5), 791.
- Prenzel, M., Krapp, A. & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik*, 32, 163–173.
- Putwain, D. W., Becker, S., Symes, W. & Pekrun, R. (2017). Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learning and Instruction*.
- Putwain, D. W., Pekrun, R., Nicholson, L. J., Symes, W., Becker, S. & Marsh, H. W. (2018). Control-Value Appraisals, Enjoyment, and Boredom in Mathematics: A Longitudinal Latent Interaction Analysis. *American Educational Research Journal*, 55(6), 1339–1368.
- Rabe, T. & Krey, O. (2018). Identitätskonstruktionen von Kindern und Jugendlichen in Bezug auf Physik – Das Identitätskonstrukt als Analyseperspektive für die Physikdidaktik? *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 201–216.
- Rakoczy, K., Buff, A. & Lipowsky, F. (2005). Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie. Unterrichtsqualität, Lernverhalten und mathematisches Verständnis. 1. Befragungsinstrumente (Bd. 13). GFPF u.a.

- Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B. & Besser, M. (2018). Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction*.
- Raykov, T. (2009). Evaluation of Scale Reliability for Unidimensional Measures Using Latent Variable Modeling. *Measurement and Evaluation in Counseling and Development*, 42(3), 223–232.
- Reber, R., Canning, E. A. & Harackiewicz, J. M. (2018). Personalized Education to Increase Interest. *Current Directions in Psychological Science*, 27(6), 449–454.
- Reeve, J., Lee, W. & Won, S. Interest as Emotion, as Affect, and as Schema (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). In: *In Interest in Mathematics and Science Learning* (K. A. Renninger, M. Nieswandt & S. Hidi, Hrsg.). Hrsg. von Renninger, K. A., Nieswandt, M. & Hidi, S. Washington, D.C.: American Educational Research Association, 2015, S. 79–92. ISBN: 9780935302424.
- Renninger, K. A. Individual interest and its implications for understanding intrinsic motivation (C. Sansone & J. M. Harackiewicz, Hrsg.). In: *In Intrinsic and Extrinsic Motivation* (C. Sansone & J. M. Harackiewicz, Hrsg.). Hrsg. von Sansone, C. & Harackiewicz, J. M. Educational Psychology. Elsevier, 2000, S. 373–404. ISBN: 9780126190700.
- Renninger, K. A. & Bachrach, J. E. (2015). Studying Triggers for Interest and Engagement Using Observational Methods. *Educational Psychologist*, 50(1), 58–69.
- Renninger, K. A., Bachrach, J. E. & Hidi, S. E. (2019). Triggering and maintaining interest in early phases of interest development. *Learning, Culture and Social Interaction*, 23, 100260.
- Renninger, K. A. & Hidi, S. (2011). Revisiting the Conceptualization, Measurement, and Generation of Interest. *Educational Psychologist*, 46(3), 168–184.
- Renninger, K. A., Krapp, A. & Hidi, S. (Hrsg.). (1992). *The Role of interest in Learning and Development*. L. Erlbaum Associates.
- Renninger, K. A., Nieswandt, M. & Hidi, S. (Hrsg.). (2015). *Interest in Mathematics and Science Learning*. American Educational Research Association.
- Retelsdorf, J., Köller, O. & Möller, J. (2011). On the effects of motivation on reading performance growth in secondary school. *Learning and Instruction*, 21(4), 550–559.
- Risch, B. (Hrsg.). (2010). *Teaching chemistry around the world*. Waxmann.
- Roberts, D. A. Scientific Literacy/Science Literacy (K. Appleton, D. L. Hanuscin, S. K. Abell & N. G. Lederman, Hrsg.). In: *Handbook of Research on Science Education* (K. Appleton, D. L. Hanuscin, S. K. Abell & N. G. Lederman, Hrsg.). Hrsg. von Appleton, K., Hanuscin, D. L., Abell, S. K. & Lederman, N. G. New York: Routledge, 2007, S. 743–794. ISBN: 9781136781216.
- Romine, W. L., Todd, A. N. & Clark, T. B. (2016). How Do Undergraduate Students Conceptualize Acid-Base Chemistry? Measurement of a Concept Progression. *Science Education*, 100(6), 1150–1183.
- Rönnebeck, S., Bernholt, S. & Ropohl, M. (2015). Searching for a common ground – A literature review of empirical research on scientific inquiry activities. *Studies in Science Education*, 52(2), 161–197.
- Rotgans, J. I. (2015). Validation Study of a General Subject-matter Interest Measure: The Individual Interest Questionnaire (IIQ). *Health Professions Education*, 1(1), 67–75.
- Rotgans, J. I. & Schmidt, H. G. (2014). Situational interest and learning: Thirst for knowledge. *Learning and Instruction*, 32, 37–50.
- Rotgans, J. I. & Schmidt, H. G. (2017a). How individual interest influences situational interest and how both are related to knowledge acquisition: A microanalytical investigation. *The Journal of Educational Research*, 1(1), 1–11.
- Rotgans, J. I. & Schmidt, H. G. (2017b). Interest development: Arousing situational interest affects the growth trajectory of individual interest. *Contemporary Educational Psychology*, 49, 175–184.
- Rotgans, J. I. & Schmidt, H. G. (2017c). The relation between individual interest and knowledge acquisition. *British Educational Research Journal*, 43(2), 350–371.
- Rotgans, J. I. & Schmidt, H. G. The Role of Interest in Learning: Knowledge Acquisition at the Intersection of Situational and Individual Interest (P. A. O'Keefe & J. M. Harackiewicz, Hrsg.). In: *In The Science of Interest* (P. A. O'Keefe & J. M. Harackiewicz, Hrsg.). Hrsg. von O'Keefe, P. A. & Harackiewicz, J. M. Cham: Springer International Publishing, 2017, S. 69–93. ISBN: 978-3-319-55507-2.
- Rumann, S. (2004). *Kooperatives experimentieren im Chemieunterricht* (Duisburg, Essen, Univ. Diss., 2004). Universität Duisburg – Essen. Fachbereich Chemie.
- Ryan, R. M. & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of personality and social psychology*, 57(5), 749–761.
- Ryan, R. M. & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67.
- Sadler, T. D. (2009). Situated learning in science education: Socio-scientific issues as contexts for practice. *Studies in Science Education*, 45(1), 1–42.
- Satorra, A. & Bentler, P. M. (2010). Ensuring Positiveness of the Scaled Difference Chi-square Test Statistic. *Psychometrika*, 75(2), 243–248.
- Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M. & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. *Educational Research Review*, 19, 158–172.
- Schaffner, E., Philipp, M. & Schiefele, U. (2016). Reciprocal effects between intrinsic reading motivation and reading competence? A cross-lagged panel model for academic track and nonacademic track students. *Journal of Research in Reading*, 39(1), 19–36.

- Schaffner, E. & Schiefele, U. (2015). The Contributions of Intrinsic and Extrinsic Reading Motivation to the Development of Reading Competence over Summer Vacation. *Reading Psychology*, 37(6), 917–941.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, 8(2), 23–74.
- Schiefele, U. (1991). Interest, Learning, and Motivation. *Educational Psychologist*, 26(3-4), 299–323.
- Schiefele, U. (1999). Interest and Learning From Text. *Scientific Studies of Reading*, 3(3), 257–279.
- Schiefele, U. The role of interest in motivation and learning (J. M. Collis & S. Messick, Hrsg.). In: In *Intelligence and personality* (J. M. Collis & S. Messick, Hrsg.). Hrsg. von Collis, J. M. & Messick, S. Mahwah, N.J.: L. Erlbaum, 2001, S. 163–194. ISBN: 9780805831665.
- Schiefele, U. Situational and Individual Interest (K. R. Wentzel & A. Wigfield, Hrsg.). In: In *Handbook of motivation at school* (K. R. Wentzel & A. Wigfield, Hrsg.). Hrsg. von Wentzel, K. R. & Wigfield, A. *Educational psychology handbook series*. New York: Routledge, 2009, S. 197–222. ISBN: 978-0-8058-6290-4.
- Schiefele, U. & Krapp, A. (1996). Topic interest and free recall of expository text. *Learning and Individual Differences*, 8(2), 141–160.
- Schiefele, U., Krapp, A. & Winteler, A. Interest as a predictor of academic achievement: A meta-analysis of research (K. A. Renninger, A. Krapp & S. Hidi, Hrsg.). In: In *The Role of interest in Learning and Development* (K. A. Renninger, A. Krapp & S. Hidi, Hrsg.). Hrsg. von Renninger, K. A., Krapp, A. & Hidi, S. Hillsdale, N.J.: L. Erlbaum Associates, 1992, S. 182–212. ISBN: 9781317783534.
- Schiefele, U. & Schaffner, E. (2015). Teacher interests, mastery goals, and self-efficacy as predictors of instructional practices and student motivation. *Contemporary Educational Psychology*, 42, 159–171.
- Schiefer, I. M., Becker, S. & Artelt, C. (2018). Eine personenzentrierte Betrachtung der Entwicklung des Fachinteresses Deutsch, Mathematik und Englisch von Jahrgangsstufe 4 bis 11. *Psychologie in Erziehung und Unterricht*, 65(2), 141.
- Schiepe-Tiska, A., Schmidtner, S., Müller, K., Heine, J., Neumann, K. & Lüdtke, O. Naturwissenschaftlicher Unterricht in Deutschland in PISA 2015 im internationalen Vergleich (K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller, Hrsg.). In: In *PISA 2015* (K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller, Hrsg.). Hrsg. von Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E. & Köller, O. Deutschland: Waxmann, 2016, S. 133–176. ISBN: 978-3-8309-3555-1.
- Schmidt, H. G. & Rotgans, J. I. (2017). Like it or not: Individual interest is not a cause but a consequence of learning. Rejoinder to Hidi and Renninger (2017). *British Educational Research Journal*, 43(6), 1266–1268.
- Schmidt, J. A., Rosenberg, J. M. & Beymer, P. N. (2018). A person-in-context approach to student engagement in science: Examining learning activities and choice. *Journal of Research in Science Teaching*, 55(1), 19–43.
- Schoon, I. (2001). Teenage job aspirations and career attainment in adulthood: A 17-year follow-up study of teenagers who aspired to become scientists, health professionals, or engineers. *International Journal of Behavioral Development*, 25(2), 124–132.
- Schukajlow, S. & Rakoczy, K. (2016). The power of emotions: Can enjoyment and boredom explain the impact of individual preconditions and teaching methods on interest and performance in mathematics? *Learning and Instruction*, 44, 117–127.
- Schurtz, I. M., Pfof, M., Nagengast, B. & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction*, 34, 32–41.
- Schuster, B. Motivation (E. Wild & J. Möller, Hrsg.). In: In *Pädagogische Psychologie* (E. Wild & J. Möller, Hrsg.). Hrsg. von Wild, E. & Möller, J. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, S. 23–74. ISBN: 978-3-642-41290-5.
- Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B. & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 632–654.
- Schwichow, M., Zimmerman, C., Croker, S. & Härtig, H. (2016). What students learn from hands-on activities. *Journal of Research in Science Teaching*, 53(7), 980–1002.
- Seidel, T. & Prenzel, M. (2006). Stability of teaching patterns in physics instruction: Findings from a video study. *Learning and Instruction*, 16(3), 228–240.
- Seidel, T., Rimmel, R. & Prenzel, M. (2005). Clarity and coherence of lesson goals as a scaffold for student learning. *Learning and Instruction*, 15(6), 539–556.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. Luchterhand Verlag GmbH.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2010). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Link.
- Sevian, H., Dori, Y. J. & Parchmann, I. (2018). How does STEM context-based learning work: What we know and what we still do not know. *International Journal of Science Education*, 40(10), 1095–1107.
- Sewasew, D., Schroeders, U., Schiefer, I. M., Weirich, S. & Artelt, C. (2018). Development of sex differences in math achievement, self-concept, and interest from grade 5 to 7. *Contemporary Educational Psychology*, 54, 55–65.
- Sheu, H.-B., Lent, R. W., Brown, S. D., Miller, M. J., Hennessy, K. D. & Duffy, R. D. (2010). Testing the choice model of social cognitive career theory across Holland themes: A meta-analytic path analysis. *Journal of Vocational Behavior*, 76(2), 252–264.

- Silvia, P. J. (2003). Self-efficacy and interest: Experimental studies of optimal incompetence. *Journal of Vocational Behavior*, 62(2), 237–249.
- Silvia, P. J. (2005). What is interesting? Exploring the appraisal structure of interest. *Emotion (Washington, D.C.)*, 5(1), 89–102.
- Silvia, P. J. (2008). Interest—The Curious Emotion. *Current Directions in Psychological Science*, 17(1), 57–60.
- Sjøberg, S. & Schreiner, C. Results and Perspectives from the Rose Project (D. Jorde & J. Dillon, Hrsg.). In: *Science Education Research and Practice in Europe: Retrospective and Prospective* (D. Jorde & J. Dillon, Hrsg.). Hrsg. von Jorde, D. & Dillon, J. Rotterdam: SensePublishers, 2012, S. 203–236. ISBN: 978-94-6091-900-8.
- Smith, C. L., Wiser, M., Anderson, C. W. & Krajcik, J. S. (2006). FOCUS ARTICLE: Implications of Research on Children's Learning for Standards and Assessment: A Proposed Learning Progression for Matter and the Atomic-Molecular Theory. *Measurement: Interdisciplinary Research & Perspective*, 4(1-2), 1–98.
- Stains, M. & Talanquer, V. (2008). Classification of chemical reactions: Stages of expertise. *Journal of Research in Science Teaching*, 45(7), 771–793.
- Stamer, I., Pönicke, H., Tirre, F., Laherto, A., Höffler, T. N., Schwarzer, S. & Parchmann, I. (2019). Development & validation of scientific video vignettes to promote perception of authentic science in student laboratories. *Research in Science & Technological Education*, 16(90), 1–17.
- Stamovlasis, D., Tsapalis, G., Kamilatos, C., Papaoikonomou, D. & Zarotiadou, E. (2005). Conceptual understanding versus algorithmic problem solving: Further evidence from a national chemistry examination. *Chem. Educ. Res. Pract.*, 6(2), 104–118.
- Stefani, C. & Tsapalis, G. (2009). Students' levels of explanations, models, and misconceptions in basic quantum chemistry: A phenomenographic study. *Journal of Research in Science Teaching*, 46(5), 520–536.
- Steinmetz, H. Estimation and comparison of latent means across cultures (E. Davidov, P. Schmidt & J. Billiet, Hrsg.). In: *In Cross-cultural analysis* (E. Davidov, P. Schmidt & J. Billiet, Hrsg.). Hrsg. von Davidov, E., Schmidt, P. & Billiet, J. European Association of Methodology Series. New York: Routledge, 2011, S. 85–116. ISBN: 9781848728226.
- Steyer, R., Shanahan, M. J. & Partchev, I. Modeling True Intraindividual Change in Structural Equation Models: The Case of Poverty and Children's Psychosocial Adjustment (T. D. Little, K. S. Cortina & J. Baumert, Hrsg.). In: *In Modeling longitudinal and multilevel data* (T. D. Little, K. S. Cortina & J. Baumert, Hrsg.). Hrsg. von Little, T. D., Cortina, K. S. & Baumert, J. New York, NY: Psychology Press, 2000, S. 109–126. ISBN: 978-0805830545.
- Stiller, J. (2016). *Scientific Inquiry im Chemieunterricht* (Diss.). Humboldt-Universität zu Berlin.
- Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U. & Roberts, B. W. (2017). Vocational interests assessed at the end of high school predict life outcomes assessed 10 years later over and above IQ and Big Five personality traits. *Journal of personality and social psychology*, 113(1), 167–184.
- Su, R. (2018). The three faces of interests: An integrative review of interest research in vocational, organizational, and educational psychology. *Journal of Vocational Behavior*.
- Sumfleth, E., Wild, E., Rumann, S. & Exeler, J. Wege zur Förderung der naturwissenschaftlichen Grundbildung im Chemieunterricht. Kooperatives Problemlösen im schulischen und familialen Kontext zum Themenbereich Säure-Base (M. Prenzel & J. Doll, Hrsg.). In: *In Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (M. Prenzel & J. Doll, Hrsg.). Hrsg. von Prenzel, M. & Doll, J. Zeitschrift für Pädagogik. Weinheim: Beltz, 2002, S. 207–221.
- Swarat, S., Ortony, A. & Revelle, W. (2012). Activity matters: Understanding student interest in school science. *Journal of Research in Science Teaching*, 49(4), 515–537.
- Taber, K. S. (2002). *Chemical Misconceptions: Prevention, diagnosis and cure: Theoretical background*, Volume 1 (Bd. 1). The Royal Society of Chemistry.
- Talanquer, V. (2008). Students' predictions about the sensory properties of chemical compounds: Additive versus emergent frameworks. *Science Education*, 92(1), 96–114.
- Talanquer, V. (2009). On Cognitive Constraints and Learning Progressions: The case of "structure of matter". *International Journal of Science Education*, 31(15), 2123–2136.
- Taskin, V. & Bernholt, S. (2014). Students' Understanding of Chemical Formulae: A review of empirical research. *International Journal of Science Education*, 36(1), 157–185.
- Taskinen, P. H., Schütte, K. & Prenzel, M. (2013). Adolescents' motivation to select an academic science-related career: The role of school factors, individual interest, and science self-concept. *Educational Research and Evaluation*, 19(8), 717–733.
- Tibubos, A. N., Rohrmann, S. & Ringeisen, T. (2019). How Students Learn to Moderate Group Work: The Role of Enjoyment and Boredom. *The Journal of psychology*, 1–21.
- Todt, E. & Schreiber, S. Development of Interests (L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert, Hrsg.). In: *In Interest and learning* (L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert, Hrsg.). Hrsg. von Hoffmann, L., Krapp, A., Renninger, K. A. & Baumert, J. IPN. Kiel: IPN, 1998, S. 25–40. ISBN: 3-89088-131-9.
- Trautwein, U., Lüdtke, O., Nagy, N., Lenski, A., Niggli, A. & Schnyder, I. (2015). Using individual interest and conscientiousness to predict academic effort: Additive, synergistic, or compensatory effects? *Journal of personality and social psychology*, 109(1), 142–162.
- Tröbst, S., Kleickmann, T., Lange-Schubert, K., Rothkopf, A. & Möller, K. (2016). Instruction and Students Declining Interest in Science: An Analysis of German Fourth- and Sixth-Grade Classrooms. *American Educational Research Journal*, 53(1), 162–193.

- Tsai, Y.-M., Kunter, M., Lüdtke, O., Trautwein, U. & Ryan, R. M. (2008). What makes lessons interesting? The role of situational and individual factors in three school subjects. *Journal of Educational Psychology*, 100(2), 460–472.
- Ültay, N. & Çalik, M. (2012). A Thematic Review of Studies into the Effectiveness of Context-Based Chemistry Curricula. *Journal of Science Education and Technology*, 21(6), 686–701.
- van Dijk, T. A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension* (1. Aufl.). Academic Press.
- van Duzor, A. G. (2016). Using Self-Explanations in the Laboratory To Connect Theory and Practice: The Decision/Explanation/Observation/Inference Writing Method. *Journal of Chemical Education*, 93(10), 1725–1730.
- van Vorst, H., Fechner, S. & Sumfleth, E. (2018). Unterscheidung von Kontexten für den Chemieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 98(2), 189.
- Vandenberg, R. J. & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, 3(1), 4–70.
- Viljaranta, J., Kiuru, N., Lerkkanen, M.-K., Silinskas, G., Poikkeus, A.-M. & Nurmi, J.-E. (2017). Patterns of word reading skill, interest and self-concept of ability. *Educational Psychology*, 37(6), 712–732.
- Viljaranta, J., Tolvanen, A., Aunola, K. & Nurmi, J.-E. (2014). The Developmental Dynamics between Interest, Self-concept of Ability, and Academic Performance. *Scandinavian Journal of Educational Research*, 58(6), 734–756.
- Vygotskij, L. S. & Cole, M. (Hrsg.). (1978). *Mind in society: The development of higher psychological processes*. Harvard Univ. Press.
- Walpuski, M. & Sumfleth, E. (2007). Strukturierungshilfen und Feedback zur Unterstützung experimenteller Kleingruppenarbeit im Chemieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 181–198.
- Wang, J. & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. Wiley.
- Weidinger, A. F., Spinath, B. & Steinmayr, R. (2016). Why does intrinsic motivation decline following negative feedback? The mediating role of ability self-concept and its moderation by goal orientations. *Learning and Individual Differences*, 47, 117–128.
- Weiner, B. (2010). The Development of an Attribution-Based Theory of Motivation: A History of Ideas. *Educational Psychologist*, 45(1), 28–36.
- Weinert, F. E. Concept of Competence: A Conceptual Clarification (D. S. Rychen & L. H. Salganik, Hrsg.). In: *In Defining and selecting key competencies* (D. S. Rychen & L. H. Salganik, Hrsg.). Hrsg. von Rychen, D. S. & Salganik, L. H. Seattle: Hogrefe & Huber, 2001, S. 45–65. ISBN: 978-0889372481.
- Wentorf, W., Höffler, T. N. & Parchmann, I. (2015). Schülerkonzepte über das Tätigkeitsspektrum von Naturwissenschaftlerinnen und Naturwissenschaftlern: Vorstellungen, korrespondierende Interessen und Selbstwirksamkeitserwartungen. *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 207–222.
- West, S. G., Taylor, A. B. & Wu, W. Model Fit and Model Selection in Structural Equation Modeling (R. H. Hoyle, Hrsg.). In: *Handbook of structural equation modeling* (R. H. Hoyle, Hrsg.). Hrsg. von Hoyle, R. H. New York, NY: Guilford Press, 2012, S. 209–231. ISBN: 9781606230770.
- Widaman, K. F., Ferrer, E. & Conger, R. D. (2010). Factorial Invariance within Longitudinal Structural Equation Models: Measuring the Same Construct across Time. *Child development perspectives*, 4(1), 10–18.
- Wigfield, A. & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, 30(1), 1–35.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25(1), 68–81.
- Wolff, F., Helm, F. & Möller, J. (2019). Integrating the 2I/E model into dimensional comparison theory: Towards a comprehensive comparison theory of academic self-concept formation. *Learning and Instruction*, 62, 64–75.
- Wolff, F., Helm, F., Zimmermann, F., Nagy, G. & Möller, J. (2018). On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Journal of Educational Psychology*, 110(7), 1005–1025.
- Wright, B. D., Linacre, J. M., Gustafson, J. E. & Martin-Lof, P. (1994). Reasonable mean-square fit values. *Rasch measurement transactions*, 8(3), 370.
- Wu, A. D., Li, Z. & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research and Evaluation*, 12(3), 1–26.
- Xu, J. (2018). Reciprocal effects of homework self-concept, interest, effort, and math achievement. *Contemporary Educational Psychology*, 55, 42–52.
- Yan, F. & Talanquer, V. (2015). Students' Ideas about How and Why Chemical Reactions Happen: Mapping the conceptual landscape. *International Journal of Science Education*, 37(18), 3066–3092.
- Yao, J.-X., Guo, Y.-Y. & Neumann, K. (2017). Refining a learning progression of energy. *International Journal of Science Education*, 9(2), 1–21.

Teil IV

Anhang





# Das RIASEC+N Modell: Deskriptive Befunde und Validierung



**Tab. A.1:** Deskriptive Befunde zur Beschreibung des RIASEC+N Konstrukts. Es wurden nur Schülerinnen und Schüler berücksichtigt, die zum Erhebungszeitpunkt den Chemieunterricht nicht abgewählt hatten. Klassen, aus denen  $n < 5$  Schülerinnen und Schüler partizipierten, wurden exkludiert. Zur Berechnung der psychometrischen Werte wurden alle durchgehend eingesetzten Items ( $n_{\text{Items}} = 28$ ) genutzt.

Jg.	Skala	$n_{\text{SuS}}$	$n_{\text{Klassen}}$	SuS je Klasse		M	SD	ICC(1)	ICC(2)	$\alpha$	$\omega$
				$[n_{\text{min}}, n_{\text{max}}]$							
Messzeitpunkt 1:											
5	R	464	23	[11, 27]		3.24	0.85	.06	.54	.73	.73 [.68,.78]
5	I	465	23	[11, 27]		3.03	0.87	.02	.29	.73	.73 [.69,.78]
5	A	462	23	[11, 27]		2.99	0.91	.00	.01	.71	.72 [.67,.77]
5	S	455	23	[11, 27]		2.73	0.84	.09	.67	.82	.82 [.79,.85]
5	E	459	23	[11, 27]		2.88	0.95	.07	.61	.77	.77 [.73,.81]
5	C	461	23	[11, 27]		2.61	0.88	.00	.05	.76	.75 [.71,.79]
5	N	461	23	[11, 27]		3.05	0.88	.08	.62	.79	.79 [.75,.82]
6	R	254	14	[ 7, 27]		3.10	0.89	.08	.61	.75	.75 [.70,.81]
6	I	253	14	[ 7, 27]		2.83	0.89	.08	.62	.73	.73 [.67,.79]
6	A	252	14	[ 7, 27]		2.87	0.97	.10	.66	.78	.79 [.74,.83]
6	S	252	14	[ 7, 27]		2.63	0.86	.06	.54	.81	.81 [.77,.86]
6	E	251	14	[ 7, 27]		2.76	0.90	.09	.63	.72	.72 [.66,.79]
6	C	250	14	[ 7, 27]		2.50	0.92	.11	.69	.76	.76 [.71,.82]
6	N	250	14	[ 7, 27]		2.83	0.92	.12	.70	.80	.80 [.75,.84]
7	R	249	14	[ 8, 27]		3.07	0.88	-.00	-.06	.72	.72 [.65,.78]
7	I	249	14	[ 8, 27]		2.74	0.95	-.02	-.42	.79	.79 [.74,.84]
7	A	249	14	[ 8, 27]		2.74	1.02	.02	.28	.75	.76 [.71,.81]
7	S	250	14	[ 8, 27]		2.44	0.89	.03	.32	.86	.86 [.83,.89]
7	E	247	14	[ 8, 27]		2.67	1.00	.02	.29	.76	.76 [.71,.82]
7	C	249	14	[ 8, 27]		2.33	0.91	.02	.25	.82	.82 [.78,.86]
7	N	250	14	[ 8, 27]		2.71	0.96	-.02	-.45	.78	.79 [.74,.84]
8	R	263	13	[10, 26]		2.94	0.88	.09	.68	.69	.69 [.61,.76]
8	I	264	13	[10, 26]		2.55	0.91	.07	.61	.80	.80 [.76,.84]
8	A	264	13	[10, 26]		2.66	0.97	.02	.27	.66	.67 [.61,.74]
8	S	265	13	[10, 26]		2.35	0.85	.02	.28	.84	.85 [.81,.88]
8	E	265	13	[10, 26]		2.55	0.93	.01	.19	.66	.68 [.62,.75]
8	C	266	13	[10, 26]		2.22	0.88	.01	.23	.79	.79 [.74,.84]
8	N	265	13	[10, 26]		2.60	0.95	.06	.56	.74	.75 [.70,.80]
9	R	486	25	[10, 25]		2.96	0.86	.07	.61	.68	.69 [.64,.74]
9	I	485	25	[10, 25]		2.69	0.90	.11	.70	.77	.77 [.73,.80]
9	A	486	25	[10, 25]		2.60	0.97	.09	.66	.73	.74 [.70,.77]
9	S	484	25	[10, 25]		2.49	0.88	.07	.60	.89	.89 [.88,.91]
9	E	485	25	[10, 25]		2.56	0.95	.06	.55	.76	.76 [.73,.80]
9	C	484	25	[10, 25]		2.29	0.86	.03	.39	.79	.79 [.76,.83]
9	N	485	25	[10, 25]		2.71	0.92	.08	.61	.74	.76 [.72,.79]
10	R	223	16	[ 6, 24]		2.83	0.96	.16	.73	.73	.73 [.67,.79]
10	I	223	16	[ 6, 24]		2.62	0.90	.25	.82	.76	.76 [.71,.82]
10	A	223	16	[ 6, 24]		2.48	0.95	-.01	-.14	.74	.73 [.67,.80]
10	S	222	16	[ 6, 24]		2.38	0.90	.19	.76	.89	.90 [.88,.93]
10	E	223	16	[ 6, 24]		2.51	0.92	.12	.66	.77	.78 [.73,.83]
10	C	222	16	[ 6, 24]		2.18	0.88	.09	.57	.78	.79 [.73,.84]
10	N	223	16	[ 6, 24]		2.62	0.96	.21	.79	.76	.78 [.73,.83]

wird fortgesetzt ...

...Fortsetzung

Jg.	Skala	n <sub>SuS</sub>	n <sub>Klassen</sub>	SuS je Klasse		M	SD	ICC(1)	ICC(2)	α	ω
				[n <sub>min</sub> , n <sub>max</sub> ]							
11	R	307	21	[ 6, 25]		2.75	0.94	.06	.50	.71	.71 [.65,.77]
11	I	307	21	[ 6, 25]		2.58	0.91	.26	.83	.81	.81 [.77,.85]
11	A	307	21	[ 6, 25]		2.32	0.92	.05	.41	.69	.70 [.65,.76]
11	S	307	21	[ 6, 25]		2.37	0.90	.19	.77	.91	.92 [.90,.93]
11	E	307	21	[ 6, 25]		2.50	0.95	.20	.78	.78	.80 [.76,.83]
11	C	307	21	[ 6, 25]		2.22	0.89	.10	.61	.80	.81 [.77,.85]
11	N	307	21	[ 6, 25]		2.55	0.95	.30	.86	.81	.82 [.79,.85]
12	R	73	7	[ 5, 18]		2.48	0.94	-.01	-.14	.74	.74 [.66,.82]
12	I	73	7	[ 5, 18]		2.37	0.95	-.01	-.07	.87	.86 [.81,.92]
12	A	73	7	[ 5, 18]		2.40	0.98	.02	.21	.67	.66 [.54,.79]
12	S	73	7	[ 5, 18]		2.15	0.90	.10	.53	.89	.89 [.85,.94]
12	E	73	7	[ 5, 18]		2.38	0.98	-.01	-.10	.86	.86 [.80,.92]
12	C	73	7	[ 5, 18]		2.11	0.92	.21	.74	.87	.87 [.82,.92]
12	N	73	7	[ 5, 18]		2.38	0.95	.10	.54	.84	.86 [.80,.91]
Messzeitpunkt 2:											
6	R	449	23	12, 25		3.22	0.84	.03	.40	.72	.72 [.67,.77]
6	I	450	23	12, 24		2.93	0.89	.02	.26	.79	.79 [.75,.83]
6	A	449	23	12, 24		2.92	0.95	.05	.48	.70	.72 [.67,.76]
6	S	447	23	12, 24		2.74	0.86	.01	.18	.86	.86 [.83,.89]
6	E	448	23	12, 24		2.81	0.93	.01	.21	.75	.76 [.72,.80]
6	C	446	23	12, 24		2.53	0.88	.02	.33	.80	.80 [.76,.83]
6	N	446	23	12, 24		2.94	0.93	.04	.47	.78	.79 [.75,.82]
10	R	428	24	6, 26		2.80	0.91	.10	.66	.74	.74 [.70,.78]
10	I	427	24	6, 26		2.56	0.90	.17	.79	.79	.79 [.75,.83]
10	A	428	24	6, 26		2.47	0.97	.07	.58	.72	.73 [.68,.77]
10	S	427	24	6, 26		2.43	0.87	.16	.77	.89	.89 [.88,.91]
10	E	428	24	6, 26		2.51	0.93	.14	.75	.78	.79 [.76,.83]
10	C	428	24	6, 26		2.22	0.86	.12	.71	.78	.78 [.74,.82]
10	N	428	24	6, 26		2.62	0.92	.20	.82	.78	.79 [.75,.82]
Messzeitpunkt 3:											
7	R	418	22	8, 28		3.19	0.82	.08	.62	.64	.65 [.58,.71]
7	I	418	22	8, 28		2.84	0.85	.12	.72	.75	.75 [.71,.79]
7	A	416	22	8, 28		2.81	0.93	.06	.53	.71	.72 [.68,.77]
7	S	416	22	8, 28		2.64	0.88	.06	.55	.88	.88 [.85,.90]
7	E	417	22	8, 28		2.76	0.93	.09	.66	.78	.78 [.74,.82]
7	C	415	22	8, 28		2.36	0.87	.09	.65	.80	.79 [.76,.83]
7	N	415	22	8, 28		2.91	0.90	.08	.61	.76	.77 [.73,.81]
11	R	286	19	5, 26		2.77	0.94	.12	.66	.77	.77 [.73,.82]
11	I	286	19	5, 26		2.67	0.87	.14	.72	.78	.78 [.73,.83]
11	A	286	19	5, 26		2.47	0.95	.03	.29	.71	.71 [.66,.77]
11	S	286	19	5, 26		2.48	0.90	.05	.46	.90	.91 [.89,.93]
11	E	286	19	5, 26		2.53	0.95	.12	.67	.81	.81 [.77,.85]
11	C	286	19	5, 26		2.31	0.87	.04	.40	.80	.80 [.75,.84]
11	N	286	19	5, 26		2.66	0.93	.09	.60	.83	.84 [.81,.87]
Messzeitpunkt 4:											
8	R	439	22	7, 26		3.15	0.82	.07	.61	.67	.67 [.62,.73]
8	I	439	22	7, 26		2.79	0.87	.04	.48	.76	.77 [.73,.81]
8	A	439	22	7, 26		2.75	0.95	.05	.53	.70	.72 [.68,.76]
8	S	439	22	7, 26		2.60	0.84	.04	.48	.90	.90 [.88,.92]
8	E	439	22	7, 26		2.72	0.89	.07	.60	.76	.77 [.73,.80]
8	C	439	22	7, 26		2.39	0.86	.07	.58	.78	.78 [.75,.82]
8	N	439	22	7, 26		2.84	0.87	.05	.53	.75	.76 [.72,.80]

wird fortgesetzt ...

...Fortsetzung

Jg.	Skala	n <sub>SuS</sub>	n <sub>Klassen</sub>	SuS je Klasse		M	SD	ICC(1)	ICC(2)	α	ω
				[n <sub>min</sub> , n <sub>max</sub> ]							
12	R	180	13	5, 23		2.89	0.91	.03	.31	.76	.75 [.70,.81]
12	I	180	13	5, 23		2.71	0.86	.00	.06	.73	.73 [.66,.80]
12	A	180	13	5, 23		2.52	0.98	.03	.29	.75	.78 [.73,.83]
12	S	180	13	5, 23		2.57	0.89	.06	.46	.90	.90 [.88,.93]
12	E	180	13	5, 23		2.59	0.91	.03	.32	.80	.81 [.77,.85]
12	C	180	13	5, 23		2.34	0.84	.04	.38	.80	.80 [.74,.86]
12	N	180	13	5, 23		2.71	0.89	.05	.41	.82	.83 [.78,.87]

Tab. A.2: Fit-Statistik für fünf Modelle (M1-M5) mit zunehmender Anzahl Freiheitsgraden und jeweils zunehmender Anzahl an latenten Klassen (CL).

N <sub>CL</sub>	AIC	BIC	ABIC	SSA-BIC	Entropy	p (VLMR)	p (LMR)	N Schülerinnen und Schüler je latenter Klasse									
								CL 1	CL 2	CL 3	CL 4	CL 5	CL 6	CL 7			
Modell 1: fixierte Varianzen, keine Kovarianzen																	
2	25626.04	25743.57		25673.68	0.88	0.00	0.00	650.69	893.31								
3	22625.60	22785.87	-2957.70	22690.56	0.92	0.00	0.00	278.05	504.19	761.76							
4	21098.01	21301.01	-1484.86	21180.29	0.92	0.00	0.00	193.32	585.90	447.47	317.30						
5	20350.46	20596.19	-704.81	20450.06	0.90	0.05	0.05	176.24	344.46	370.86	469.71	182.73					
6	19801.12	20089.60	-506.60	19918.06	0.91	0.03	0.03	91.56	171.35	434.91	347.45	167.29	331.45				
7	19460.00	19791.21	-298.39	19594.25	0.90	0.03	0.03	87.73	161.91	361.70	312.93	308.25	210.02	101.46			
Modell 2: fixierte Varianzen, fixierte Kovarianzen																	
2	12159.81	12389.52		12252.92	0.72	0.00	0.00	259.33	1284.67								
3	12114.09	12386.54	-2.98	12224.53	0.63	0.00	0.00	210.08	531.96	801.96							
4	11791.96	12107.15	-279.40	11919.72	0.92	0.00	0.00	278.33	510.62	618.90	136.15						
5	11746.36	12104.28	-2.87	11891.44	0.90	0.00	0.00	201.56	81.33	614.94	509.82	136.36					
6	11715.99	12116.65	12.37	11878.39	0.89	0.54	0.54	254.13	16.84	419.78	493.45	133.27	226.53				
7	11674.71	12118.10	1.45	11854.43	0.88	0.36	0.37	189.92	66.21	19.27	225.71	491.67	417.96	133.27			
Modell 3: freie Varianzen, keine Kovarianzen																	
2	12159.81	12389.52		12252.92	0.72	0.00	0.00	259.33	1284.67								
2	25563.72	25718.64		25626.51	0.89	0.00	0.00	735.83	808.17								
3	22548.29	22783.35	-2935.29	22643.57	0.92	0.01	0.01	338.35	706.16	499.49							
4	21014.35	21329.54	-1453.81	21142.11	0.92	0.00	0.00	207.73	446.26	330.51	559.50						
5	20192.37	20587.69	-741.85	20352.61	0.92	0.09	0.09	197.36	385.96	439.04	355.25	166.38					
6	19576.63	20052.08	-535.61	19769.35	0.92	0.06	0.06	44.04	435.10	190.13	369.51	345.56	159.66				
7	19174.93	19730.51	-521.57	19400.13	0.92	0.10	0.11	351.23	43.87	174.95	355.61	255.66	246.65	116.03			
Modell 4: freie Varianzen, fixierte Kovarianzen																	
2	12122.45	12389.55		12230.72	0.93	0.00	0.00	74.25	1469.75								
3	12042.07	12389.31	-0.25	12182.82	0.48	0.13	0.13	664.43	368.95	510.63							
4					0.57			26.21	478.31	522.50	516.98						
5	11723.25	12230.75		11928.96	0.88	0.10	0.11	106.17	147.96	656.62	486.44	146.81					
6	11592.06	12179.69	-51.06	11830.25	0.89	0.28	0.29	116.25	144.37	462.67	643.15	86.62	90.93				
7	11529.69	12197.46	17.76	11800.36	0.85	0.33	0.34	76.58	151.66	429.09	515.30	188.52	90.96	91.88			
Modell 5: freie Varianzen, freie Kovarianzen																	
2	11963.12	12342.41		12116.86	0.92	0.00	0.00	161.69	1382.31								
3	11717.17	12288.78	-53.63	11948.87	0.93	0.25	0.25	1357.78	39.46	146.76							
4	11585.08	12349.00	60.22	11894.73	0.66	0.72	0.72	297.99	232.12	731.65	282.23						
5	11493.58	12449.82	100.81	11881.18	0.72	0.73	0.73	53.02	825.13	457.99	75.23	132.62					
6	11466.16	12614.72	164.90	11931.71	0.72	0.26	0.26	53.26	637.53	47.75	574.45	156.23	74.78				
7	11509.33	12850.21	235.49	12052.84	0.74	0.72	0.72	52.53	139.43	750.35	85.74	102.55	244.94	168.47			

## Knowing more about things you care less about

**Table B.1:** RIASEC+N Questionnaire: dimensions and items to the question asking „How much are you interested in doing the following activities in chemistry lessons?“

Dimension	Variable	Item
Realistic	R1	To do experiments guided by an instruction
	R2	To build a machine guided by an instruction
	R3	To arrange all the equipment for an experiment
	R4	To make a chemical product
Investigative	I1	To plan experiments to investigate something
	I2	To find the solution to a chemistry problem
	I3	To investigate an object closely
	I4	To find an explanation for an observation
Artistic	A1	To draw an observation
	A2	To draw the set-up of an experiment
	A3	To create a poster or a presentation for a chemical topic
	A4	To recreate something seen in nature
Social	S1	To explain something to fellow students in chemistry
	S2	To explain chemical topics to fellow students
	S3	To help fellow students with chemistry tasks
	S4	To help fellow students doing experiments
Enterprising	E1	To organize a small chemistry project
	E2	To present the results of a small chemistry project
	E3	To organize a small chemical investigation
	E4	To lead a student group in a group work
Conventional	C1	To table the results of an experiment
	C2	To collect and organize chemical products (substances, information, data)
	C3	To make a diagram from the results of an experiment
	C4	To search for and organize chemical information
Networking	N1	To talk to fellow students about chemical topics
	N2	To work on a chemical topic together with fellow students
	N3	To investigate a chemical topic together with fellow students
	N4	To experiment with fellow students

**Table B.2:** Sample and age distribution across grades and type of course covering chemistry content.

Grade	Integrated Science or Physics and/or Biology		Chemistry		No chemistry	
	N	Age (M ± SD)	N	Age (M ± SD)	N	Age (M ± SD)
5	484	10.6 ± 0.57				
6	253	11.6 ± 0.56				
7			262	12.6 ± 0.61		
8			265	13.6 ± 0.58		
9			487	14.6 ± 0.58		
10			221	15.8 ± 0.76	62	15.8 ± 0.77
11			322	17.1 ± 0.76	154	17.1 ± 0.77

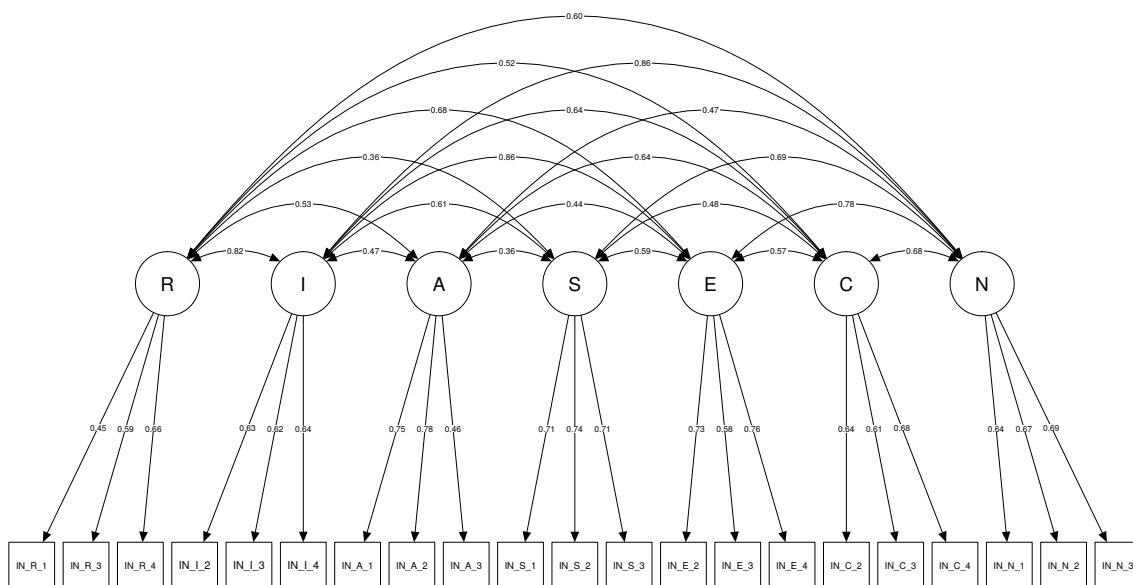
**Table B.3:** Fit indices for the multiple-group confirmatory factor analysis (MGCFA) and grade-specific CFAs

Title	Grades	N	$\chi^2$	df	CFI	TLI	RMSEA	SRMR
CFA	5 – 11	2268	1458.187	167	.938	.922	.058	.047
CFA grade-specific	5	469	340.403	167	.949	.936	.047	.041
CFA grade-specific	6	250	278.468	168	.945	.931	.051	.047
CFA grade-specific	7	255	334.561	167	.929	.911	.063	.053
CFA grade-specific	8	265	325.928	166	.928	.909	.060	.052
CFA grade-specific	9	486	487.717	167	.930	.912	.063	.056
CFA grade-specific	10	221	316.895	167	.933	.916	.064	.066
CFA grade-specific	11	322	395.013	166	.940	.924	.065	.055
MGCFA	5 – 11	2268	3050.265	1378	.919	.913	.061	.081

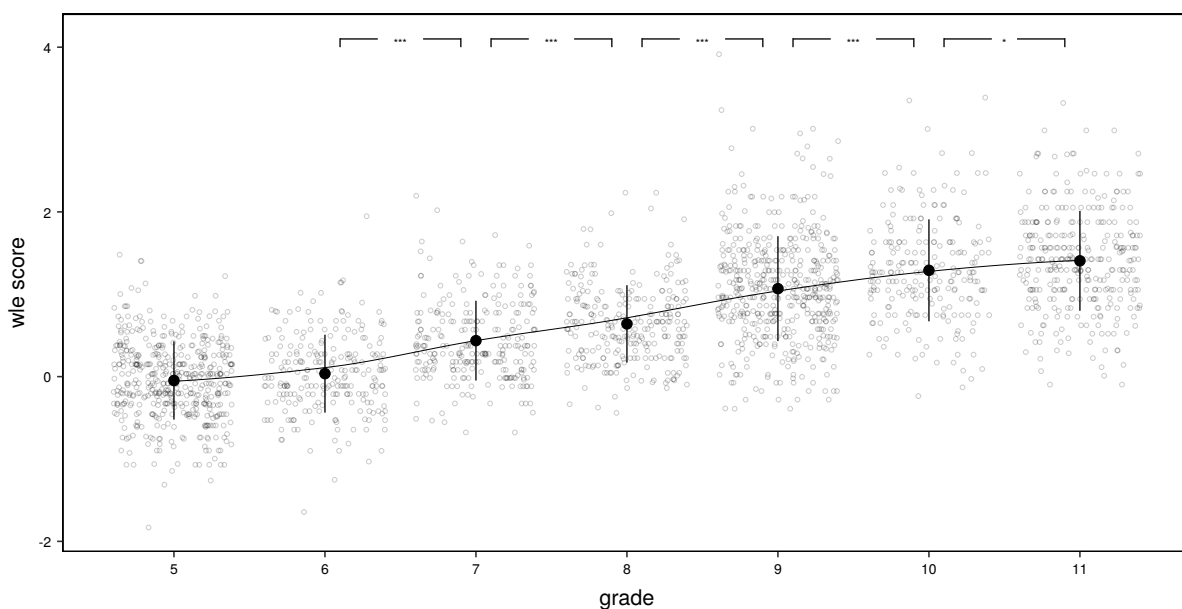
**Table B.4:** Goodness-of-Fit Indices for Measurement Invariance

Multiple-Group Model	$\chi^2$	df	CFI	TLI	RMSEA	SRMR
Configural measurement invariance	2478.985	1168	.936	.920	.059	.052
Metric/weak measurement invariance	2704.207	1252	.929	.917	.060	.062
Scalar/strong measurement invariance	2967.502	1336	.921	.913	.061	.067
Strict measurement invariance	3662.518	1524	.896	.900	.066	.095

\*\*p < .01, \*\*\*p < .001: CFI, comparative fit index; TLI, Tucker–Lewis index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.



**Figure B.1:** RIASEC+N model for interest in science activities. The CFA model was estimated for the whole sample (without grouping of grades).



**Figure B.2:** Scatter plot of students' WLE parameters across Grades 5 to 11, including mean values and standard deviations per grade and significant differences in mean values between adjacent grades (\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ ). Only students attending chemistry courses are included.





# Longitudinal couplings between interest and conceptual understanding in secondary school chemistry



## C.1 Supplemental Materials for Instruments

### RIASEC+N Questionnaire

**Table C.1:** Scales and items to the question asking „How much are you interested in doing the following activities in chemistry lessons?“ Items that were excluded in the reported analysis are crossed out.

Dimension	Variable	Item
Realistic	R1	To do experiments guided by an instruction
	<del>R2</del>	<del>To build a machine guided by an instruction</del>
	R3	To arrange all the equipment for an experiment
	R4	To make a chemical product
Investigative	I1	To plan experiments to investigate something
	<del>I2</del>	<del>To find the solution to a chemistry problem</del>
	I3	To investigate an object closely
	I4	To find an explanation for an observation
Artistic	A1	To draw an observation
	A2	To draw the set-up of an experiment
	A3	To create a poster or a presentation for a chemical topic
	A4	To recreate something seen in nature
Social	S1	To explain something to fellow students in chemistry
	S2	To explain chemical topics to fellow students
	S3	To help fellow students with chemistry tasks
	S4	To help fellow students doing experiments
Enterprising	E1	To organize a small chemistry project
	E2	To present the results of a small chemistry project
	E3	To organize a small chemical investigation
	E4	To lead a student group in a group work
Conventional	C1	To table the results of an experiment
	C2	To collect and organize chemical products (substances, information, data)
	C3	To make a diagram from the results of an experiment
	<del>C4</del>	<del>To search for and organize chemical information</del>
Networking	N1	To talk to fellow students about chemical topics
	N2	To work on a chemical topic together with fellow students
	N3	To investigate a chemical topic together with fellow students
	<del>N4</del>	<del>To experiment with fellow students</del>

## Achievement test

Exemplary items with coding scheme (in squared brackets behind each answer option) and item fit indices; Item names reflect the underlying concept area (first two letters; CR = chemical reaction, EN = energy, MA = matter) and the grade level (third and fourth letter; CO = core items, i.e. anchor items that were administered in all grade levels, 05 to 11 for the respective grade level).

**Table C.2:** Item CRCO2

---

Atoms are an important part in chemical reactions.  
What happens in chemical reactions with the atoms?

1. In chemical reactions, new atoms are formed. [0]
2. In chemical reactions, new bonds are formed. [2]
3. In chemical reactions, atoms mix in a new way. [1]
4. In chemical reactions, the shape of the atoms changes. [0]

---

Weighted MNSQ	0.93
Discrimination	0.70
Item Thresholds	-0.73, -0.13
Distribution of students' responses	a) 0.12%; b) 0.51%; c) 0.16%; d) 0.14%; missings = 0.08 (Grades 5 to 11)

---

**Table C.3:** Item ENCO4

---

In a cup tea is brewed. Then the cup of tea is left.  
What happened?

1. The tea is cooled with time, because a part of its heat energy is destroyed. [0]
2. The tea is cooled with time, because a part of its heat energy is emitted to its environment and a part disappears. [1]
3. The tea cools with time, because everything which is hot, cools down with time. [0]
4. The tea is cooled with time, because a part of its heat energy is emitted to the environment. [2]

---

Weighted MNSQ	0.91
Discrimination	0.70
Item Thresholds	-0.91 -0.33
Distribution of students' responses	a) 0.09%; b) 0.11%; c) 0.1%; d) 0.66%; missings: 0.04% (Grades 5 to 11)

---

**Table C.4:** Item MACO4

Garage doors made of metal open heavier in summer than in winter. This is because metal expands when heated.

Why expand metals when heated?

1. The metal particles move faster and thus have a greater distance from each other. [2]
2. The metal particles have reacted with the air particles. Thus, the particles become larger. [1]
3. By heating the metal particles melt a little, which is making them further apart. [0]
4. The metal particles are further apart by the heat, their movements do not change. [1]

Weighted MNSQ	0.96
Discrimination	0.61
Item Thresholds	-1.18 -0.10
Distribution of students' responses	a) 0.44%; b) 0.09%; c) 0.18%; d) 0.22%; missings: 0.06% (Grades 5 to 11)

**Table C.5:** Item CR064

Chemical reactions are sometimes intense, sometimes you barely notice them.

What happens in chemical reactions with the materials?

1. In chemical reactions, the starting materials are no longer present after the reaction. [1]
2. In chemical reactions, the colour of materials changes. [0]
3. In chemical reactions, new substances arise, but there can also the starting materials afterwards still be present. [2]
4. In chemical reactions dangerous products are formed. [0]

Weighted MNSQ	1
Discrimination	0.68
Item Thresholds	-0.94 -0.63
Distribution of students' responses	a) 0.09%; b) 0.17%; c) 0.4%; d) 0.11%; missings: missings: 0.23% (Grade 6 only)

**Table C.6:** Item MA092

In an airtight glass box 2g of water are added. The water-filled glass box has a weight of 102g. The water is allowed to evaporate and the box is weighted again.

How heavy is the box now?

1. The box weighs less than 100 grams, because the water is turned into air, and air is very light. [0]
2. The box weighs 102 g, because the water is turned into steam, but still weighs as much as before. [1]
3. The box weighs more than 102 g, because it has now become hot and so it has increased. [0]
4. The box weighs less than 102 grams, because some drops of water are gone. [0]

Weighted MNSQ	0.94
Discrimination	0.83
Item Thresholds	-0.12
Distribution of students' responses	a) 0.17%; b) 0.6%; c) 0.06%; d) 0.15%; missings: 0.02% (Grade 9 only)

**Table C.7:** Item EN113

---

Batteries allow us mobility, because we can use electricity on the go.  
What processes take place in a battery that provide the electricity?

1. The substances contained in a battery react with each other, this leads to the transformation of chemical energy to electrical energy. [1]
2. The substances contained in a battery have electrical energy stored, which they release when turning on the device. [0]
3. The substances contained in a battery are broken and released by energy. [0]
4. The substances contained in a battery react with each other, which results in the formation of electrical energy. [0]

---

Weighted MNSQ	0.95
Discrimination	0.68
Item Thresholds	0.55
Distribution of students' responses	a) 0.55%; b) 0.20%; c) 0.04%; d) 0.16%; missings: 0.04% (Grade 11 only)

---

## C.2 Supplemental Materials for Data analysis

## Missing data and method effect

As students in the present study were nested in different schools and classes, a multilevel approach would be necessary to systematically analyse the effect of this kind of nested data structure. However, this was not possible in the present study due to the limited number of classes participating in this study ( $N_{\text{classes}} = 25$ ) which is below common recommendations for multilevel approaches. To gain some intuition regarding the extent to which the nested data structure might affect the results, the intra-class correlation (ICC; i.e., the correlation among individuals within the same class; the bigger the ICC, the greater the impact of the class level) was calculated for each variable and measurement point, resulting in a mean value of  $M(\text{ICC}) = .12$ . This value indicates that only 12% of the variance in the different variables can be accounted for on the class level. B. O. Muthén and Satorra (1995) argued that it is not the size of the intra-class correlation, but the design effect (i.e., a function of the intra-class correlation and the average class size) and that a design effect greater than 2 indicates that the clustering in the data needs to be taken into account during estimation. The mean design effect across variables and measurement points in this study is 1.80, again indicating that the nested structure of the data might only marginally affect the results.

**Table C.8:** Reliabilities, descriptive statistics, and correlation coefficients for the RIASEC+N (Realistic, Investigative, Artistic, Social, Enterprising, Conventional, Networking) model and WLE Scores in Grades 9 to 11.

Grade	Scale	n	Mean	SD	$\alpha$	Correlation coefficients (Pearson)							
						R	I	A	S	E	C	N	
9	Realistic	486	2.89	0.86	0.65								
	Investigative	486	2.61	0.88	0.74	0.43							
	Artistic	489	2.58	0.96	0.72	0.40	0.41						
	Social	488	2.50	0.88	0.89	0.27	0.57	0.26					
	Enterprising	486	2.59	0.95	0.76	0.38	0.60	0.37	0.51				
	Conventional	488	2.33	0.86	0.75	0.25	0.55	0.41	0.47	0.49			
	Networking	487	2.54	0.89	0.76	0.30	0.63	0.34	0.60	0.61	0.61		
	WLE Score	503	0.27	0.60	0.74	0.11	0.27	0.01	0.20	0.12	0.11	0.19	
10	Realistic	428	2.73	0.92	0.68								
	Investigative	427	2.50	0.88	0.77	0.45							
	Artistic	426	2.42	0.96	0.73	0.44	0.36						
	Social	428	2.42	0.87	0.89	0.27	0.57	0.31					
	Enterprising	428	2.50	0.93	0.80	0.43	0.63	0.35	0.58				
	Conventional	426	2.26	0.86	0.76	0.27	0.54	0.38	0.47	0.53			
	Networking	426	2.49	0.91	0.78	0.33	0.70	0.26	0.62	0.67	0.58		
	WLE Score	456	0.33	0.54	0.75	0.07	0.23	-0.07	0.23	0.16	0.08	0.24	
11	Realistic	395	2.63	0.95	0.71								
	Investigative	395	2.54	0.88	0.74	0.49							
	Artistic	393	2.37	0.96	0.72	0.47	0.34						
	Social	395	2.39	0.89	0.9	0.28	0.53	0.26					
	Enterprising	395	2.40	0.94	0.80	0.49	0.67	0.37	0.54				
	Conventional	393	2.24	0.87	0.75	0.37	0.54	0.39	0.46	0.52			
	Networking	393	2.45	0.94	0.84	0.38	0.68	0.26	0.68	0.69	0.59		
	WLE Score	409	0.67	0.71	0.77	0.18	0.35	0.02	0.29	0.20	0.20	0.33	

**Table C.9:** Goodness-of-Fit indices for testing measurement invariance per dimension of the RIASEC+N model. Model fits were compared by CFI and RMSEA difference testing. Hereby, a change of  $< .01$  in CFI supplemented by a change of  $< .015$  in RMSEA between subsequent models is generally considered to indicate invariance (G. W. Cheung & Rensvold, 2002). As this criterion was slightly exceeded in the invariance testing of the social dimension ( $\Delta\text{CFI} = .011$ ), McDonald's NCI and gammaHat were calculated. These global fit indices were found to be below recommended thresholds (cf. Meade et al., 2008; West et al., 2012), so that strong measurement invariance was assumed for the social dimension as well.

Scale	Invariance	ChiSq (df)	CFI	RMSEA	SRMR	ACFI	ARMSEA	ASRMR	ANCI	ΔGHAT	ΔadjGHAT
Realistic	configural	12.220 (15)	1.000	0.000	0.028	0.000	0.000	0.001	0.002	0.000	0.011
	metric	14.156 (19)	1.000	0.000	0.029	0.000	0.000	0.003	0.001	0.000	0.018
	strong	21.250 (23)	1.000	0.000	0.032	0.000	0.000	0.006	0.002	0.000	0.025
Investigative	strict	23.469 (29)	1.000	0.000	0.038	0.000	0.000	0.006	0.002	0.000	0.025
	configural	18.710 (15)	0.996	0.019	0.029	0.001	0.005	0.004	0.001	0.000	0.015
	metric	21.683 (19)	0.997	0.014	0.033	0.001	0.005	0.004	0.001	0.000	0.015
Artistic	strong	24.446 (23)	0.999	0.009	0.034	0.002	0.005	0.001	0.002	0.000	0.024
	strict	30.015 (29)	0.999	0.007	0.045	0.000	0.002	0.011	0.002	0.000	0.033
	configural	19.821 (15)	0.996	0.021	0.030	0.002	0.002	0.007	0.002	0.000	0.020
Social	metric	26.328 (19)	0.994	0.023	0.037	0.002	0.002	0.002	0.001	0.000	0.032
	strong	29.693 (23)	0.994	0.020	0.039	0.000	0.003	0.002	0.001	0.000	0.032
	strict	47.360 (29)	0.985	0.030	0.052	0.009	0.010	0.013	0.009	0.002	0.043
Enterprising	configural	23.187 (15)	0.992	0.028	0.026	0.001	0.005	0.003	0.001	0.000	0.033
	metric	25.954 (19)	0.993	0.023	0.029	0.001	0.005	0.003	0.001	0.000	0.033
	strong	40.988 (23)	0.982	0.033	0.030	0.011	0.010	0.001	0.007	0.001	0.054
Conventional	strict	45.671 (29)	0.983	0.028	0.033	0.001	0.005	0.003	0.006	0.001	0.072
	configural	22.961 (15)	0.993	0.027	0.034	0.001	0.001	0.007	0.001	0.000	0.032
	metric	28.411 (19)	0.992	0.026	0.041	0.001	0.001	0.003	0.003	0.001	0.052
Networking	strong	35.875 (23)	0.988	0.028	0.044	0.004	0.002	0.003	0.003	0.001	0.052
	strict	48.408 (29)	0.983	0.031	0.045	0.005	0.003	0.001	0.008	0.002	0.070
	configural	19.404 (15)	0.994	0.020	0.029	0.005	0.004	0.011	0.002	0.000	0.018
Networking	metric	26.626 (19)	0.989	0.024	0.040	0.005	0.004	0.004	0.002	0.000	0.018
	strong	29.995 (23)	0.990	0.021	0.044	0.001	0.003	0.004	0.002	0.000	0.029
	strict	35.625 (29)	0.991	0.018	0.066	0.001	0.003	0.022	0.002	0.000	0.039
Networking	configural	60.459 (39)	0.992	0.028	0.030	0.000	0.003	0.003	0.001	0.000	0.039
	metric	64.416 (45)	0.992	0.025	0.033	0.000	0.003	0.003	0.001	0.000	0.039
	strong	81.186 (51)	0.988	0.029	0.034	0.004	0.004	0.001	0.006	0.001	0.066
Networking	strict	82.201 (59)	0.991	0.024	0.036	0.003	0.005	0.002	0.001	0.001	0.091

## C.3 Supplemental Materials for Results – Latent growth models

**Table C.10:** Fit indices of the first order latent growth curve models.

Scale	Shape	CFI	TLI	RMSEA	SRMR
Realistic	linear	0.998	0.994	0.020	0.014
Realistic	free	1.000	1.000	0.000	0.000
Investigative	linear	0.973	0.918	0.083	0.032
Investigative	free				
Artistic	linear	0.974	0.923	0.078	0.029
Artistic	free	1.000	1.000	0.000	0.000
Social	linear	1.000	1.003	0.000	0.012
Social	free	1.000	1.000	0.000	0.000
Enterprising	linear	1.000	1.011	0.000	0.010
Enterprising	free	1.000	1.000	0.000	0.001
Conventional	linear	1.000	0.999	0.006	0.014
Conventional	free	1.000	1.000	0.000	0.001
Networking	linear	1.000	1.025	0.000	0.006
Networking	free	1.000	1.000	0.000	0.001
WLE Score	linear	0.937	0.811	0.147	0.070
WLE Score	free	1.000	1.000	0.000	0.002

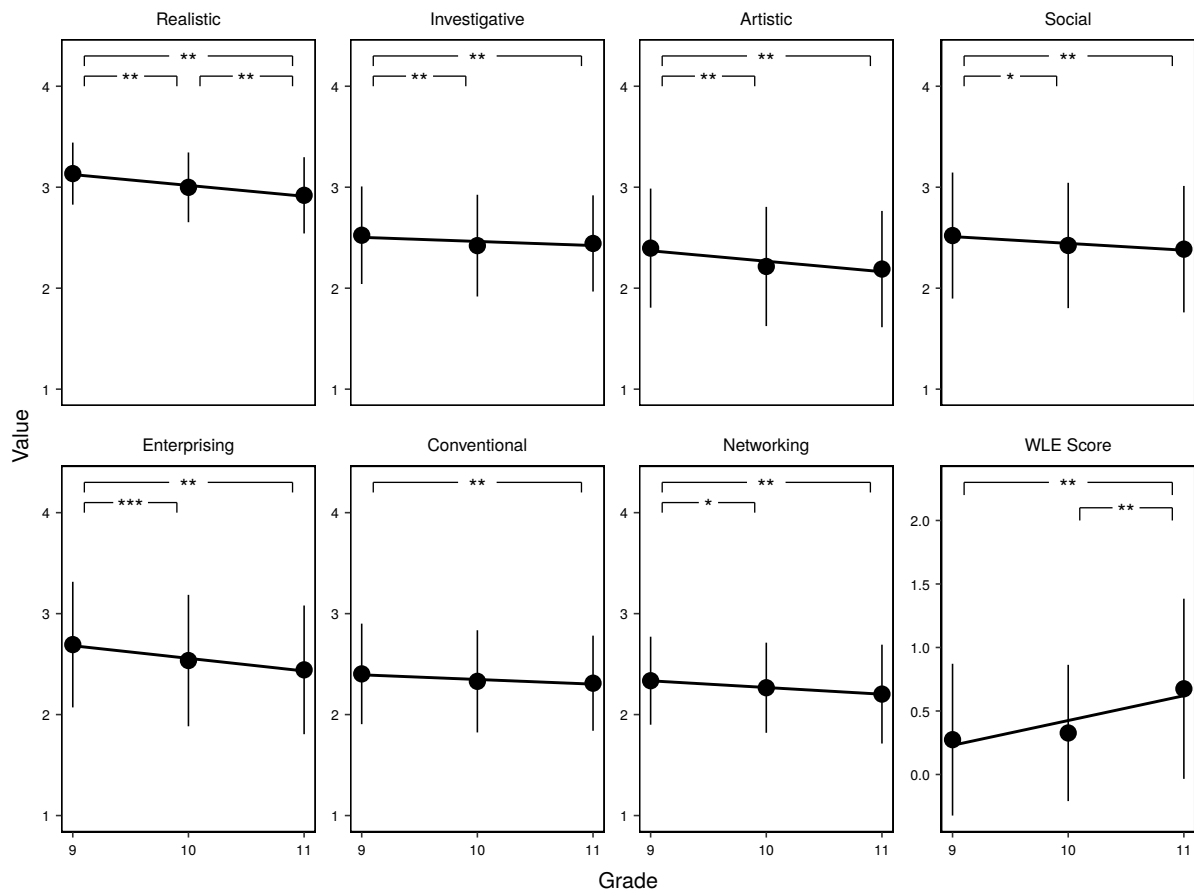
**Table C.11:** Fit indices of the True Individual Change Models (Steyer et al., 2000).

Scale	CFI	TLI	RMSEA	SRMR
Realistic	0.990	0.984	0.021	0.044
Investigative	0.982	0.972	0.033	0.030
Artistic	1.000	1.003	0.000	0.032
Social	0.988	0.985	0.029	0.034
Enterprising	0.994	0.991	0.020	0.039
Conventional	0.999	0.998	0.009	0.034
Networking	0.988	0.982	0.028	0.044

**Table C.12:** Fit indices and parameters of the first order latent growth curve models (nonlinear for Artistic and WLE Score, otherwise linear). The LGCM of the investigative dimension with a freely estimated shape didn't fit the data. The slope variance of the WLE scores was constrained to be non-negative.

	Conceptual understanding (WLE score)	Realistic	Investigative	Artistic	Social	Enterprising	Conventional	Networking
Mean initial level	0.277***	2.949***	2.696***	2.602***	2.490***	2.592***	2.315***	2.540***
Variance initial level	0.206***	0.153***	0.633	0.476	0.366***	0.331***	0.197***	0.266***
Mean slope	0.182***	-0.134***	-0.081***	-0.093***	-0.071**	-0.112*	-0.050*	-0.071**
Variance slope	0.001***	0.035	0.105	0.076	0.057	0.038	0.041	0.013
Correlation I ~ S	0.885	0.285	-0.73	-0.638	-0.317*	-0.163	-0.032	-0.199
Cohen's d (T1 → T3)	0.623	-0.533	-0.142	-0.314	-0.191	-0.338	-0.155	-0.225
CFI	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000
RMSEA	0.000	0.020	0.000	0.000	0.000	0.000	0.006	0.000
SRMR	0.004	0.014	0.001	0.000	0.012	0.010	0.014	0.006





**Figure C.1:** Mean factor scores and corresponding standard deviations of students' interest in school science activities (Realistic, Investigative, Artistic, Social, Enterprising, Conventional, Networking) and conceptual understanding (WLE Score) across Grades 9 to 11 based on the True Individual Change Model (for interest variables; Steyer et al., 2000) respectively manifest means (for the WLE score). Significant differences between grades are reported for comparisons between grades within two years of each other (\* p < .05; \*\* p < .01; \*\*\* p < .001).

## C.4 Supplemental Materials for Results – Cross-lagged models

**Table C.13:** Cross-lagged models for interest in school science activities and WLE scores: Standardized factor loadings, path coefficients, residual variances, and model fit indices.

	Realistic model		Investigative model		Artistic model		Social model		Enterprising model		Conventional model		Networking model	
	R	WLE Score	I	WLE Score	A	WLE Score	S	WLE Score	E	WLE Score	C	WLE Score	N	WLE Score
Factor loadings	.58 – .74		.57 – .86		.41 – .88		.69 – .91		.58 – .87		.53 – .89		.49 – .90	
Autoregressive effects														
T1 → T2	.63***		.62***		.61***		.60***		.62***		.48***		.56***	
T2 → T3	.70***		.63***		.67***		.61***		.64***		.60***		.49***	
Cross-lagged effects														
WLE → R	R → WLE	WLE → I	I → WLE	WLE → A	A → WLE	WLE → S	S → WLE	WLE → E	E → WLE	WLE → C	C → WLE	WLE → N	N → WLE	
T1 → T2	.04	.11*	.11**	-.02	.00	.09*	.10**	.07	.09*	.07	.06	.18***	.10**	
T2 → T3	.03	.11*	.09**	-.01	.00	.08*	.08**	.06	.07*	.07	.05	.15***	.08**	
Effects of covariates at T1														
Age	.03	.02	.02	.06	-.02	-.06	-.01	.01	-.01	.04	-.01	.05	-.01	
Gender	.23***	.10*	.12**	-.15***	.11**	.03	.11**	.07	.11**	-.07	.11**	-.02	.11**	
KFT	.10	.07	.36***	.00	.36***	.15**	.36***	.02	.36***	.07	.36***	.07*	.36***	
Residual variance														
T1	.15	.03	.15	.03	.15	.03	.15	.01	.15	.01	.15	.01	.15	
T2	.26	.43	.44	.41	.43	.39	.43	.41	.43	.26	.43	.40	.43	
T3	.49	.42	.46	.50	.41	.43	.42	.42	.42	.39	.42	.31	.42	
Model fit														
CFI	.943		.974		.983		.988		.971		.962		.971	
RMSEA	.037		.028		.023		.022		.032		.034		.032	
SRMR	.058		.038		.044		.037		.046		.049		.046	



# Probing Precursors of Students' Long-term Engagement in Chemistry



## D.1 Supplemental Materials for Method – Measurement Invariance

Model fits were compared by CFI, RMSEA and SRMR difference testing to establish measurement invariance. Hereby, for sample sizes  $N > 300$  a change of  $\geq .010$  in CFI supplemented by a change of  $\geq .015$  in RMSEA and a change of  $\geq .030$  in SRMR between the configural and the metric model as well as a change of  $\geq .010$  in CFI supplemented by a change of  $\geq .015$  in RMSEA and a change of  $\geq .030$  in SRMR between subsequent models is generally considered to indicate non invariance (scalar and strict) (Chen, 2007, p. 501). Additionally, for changes in McDonald's NCI and gammaHat (GHAT) critical values for single factors with three items are between  $\Delta\text{NCI} < .0065$  (Meade et al., 2008, p. 586) and  $\Delta\text{NCI} < .0200$  (G. W. Cheung & Rensvold, 2002, p. 251) and  $\Delta\text{GHAT} < .0010$ , respectively.

**Table D.1:** Goodness-of-Fit indices for testing measurement invariance over time of the bivariate latent change score models for lower secondary students (LS5) and upper secondary students (US9).

Scale	Invariance	CFI	RMSEA	SRMR	CFI	RMSEA	SRMR	NCI	GHAT
Lower secondary students (LS5).									
I	configural	0.959	0.047	0.037					
I	metric	0.957	0.046	0.04	0.002	0.001	0.003	0.0028	0.0003
I	scalar	0.954	0.046	0.04	0.003	0.000	0.000	0.0067	0.0007
S	configural	0.980	0.039	0.026					
S	metric	0.979	0.039	0.029	0.001	0.000	0.003	0.0032	0.0003
S	scalar	0.977	0.040	0.030	0.002	0.001	0.001	0.0076	0.0006
N	configural	0.946	0.056	0.045					
N	metric	0.944	0.056	0.048	0.002	0.000	0.003	0.0041	0.0004
N	scalar	0.942	0.055	0.049	0.002	0.001	0.001	0.0068	0.0008
Upper secondary students (US9).									
I	configural	0.920	0.076	0.049					
I	metric	0.920	0.073	0.052	0.000	0.003	0.003	0.0012	0.0003
I	scalar	0.913	0.074	0.056	0.007	0.001	0.004	0.0109	0.0014
S	configural	0.964	0.064	0.031					
S	metric	0.965	0.061	0.033	0.001	0.003	0.002	0.0018	0.0001
S	scalar	0.960	0.062	0.036	0.005	0.001	0.003	0.0096	0.0011
N	configural	0.953	0.065	0.041					
N	metric	0.952	0.063	0.044	0.001	0.002	0.003	0.0035	0.0004
N	scalar	0.950	0.062	0.046	0.002	0.001	0.002	0.0067	0.0009

Note. I = investigative, S = social, N = networking, CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual, NCI = McDonald's non-centrality index, GHAT = Gamma Hat.

## D.2 Supplemental Materials for Method – Instruments

**Table D.2:** *RIASEC+N Questionnaire*. Scales and items to the question asking „How much are you interested in doing the following activities in chemistry lessons?“ Items that were excluded in the reported analysis are crossed out (LS5 = lower secondary students, US9 = upper secondary students). Item R5 was not assessed at the first measurement occasion.

Scale	Variable		Item
	LS5	US9	
Realistic	R1	R1	To do experiments guided by an instruction
	<del>R2</del>	R2	To build a machine guided by an instruction
	R3	R3	To arrange all the equipment for an experiment
	R4	<del>R4</del>	To make a chemical product
	R5	<del>R5</del>	To work with laboratory equipment
Investigative	I1	I1	To plan experiments to investigate something
	<del>I2</del>	I2	To find the solution to a chemistry problem
	I3	<del>I3</del>	To investigate an object closely
	I4	I4	To find an explanation for an observation
Artistic	A1	A1	To draw an observation
	A2	A2	To draw the set-up of an experiment
	A3	A3	To create a poster or a presentation for a chemical topic
	A4	<del>A4</del>	To recreate something seen in nature
Social	S1	<del>S1</del>	To explain something to fellow students in chemistry
	S2	S2	To explain chemical topics to fellow students
	S3	S3	To help fellow students with chemistry tasks
	S4	S4	To help fellow students doing experiments
Enterprising	<del>E1</del>	<del>E1</del>	<del>To organize a small chemistry project</del>
	E2	E2	To present the results of a small chemistry project
	E3	E3	To organize a small chemical investigation
	E4	E4	To lead a student group in a group work
Conventional	C1	C1	To table the results of an experiment
	<del>C2</del>	C2	To collect and organize chemical products (substances, information, data)
	C3	C3	To make a diagram from the results of an experiment
	C4	<del>C4</del>	To search for and organize chemical information
Networking	N1	N1	To talk to fellow students about chemical topics
	N2	N2	To work on a chemical topic together with fellow students
	N3	N3	To investigate a chemical topic together with fellow students
	<del>N4</del>	<del>N4</del>	<del>To experiment with fellow students</del>

**Table D.3:** *Achievement test.* Exemplary item CRCO2 with coding scheme (in squared brackets behind each answer option) and item fit indices

---

Atoms are an important part in chemical reactions.  
 What happens in chemical reactions with the atoms?

1. In chemical reactions, new atoms are formed. [0]
2. In chemical reactions, new bonds are formed. [2]
3. In chemical reactions, atoms mix in a new way. [1]
4. In chemical reactions, the shape of the atoms changes. [0]

---

Weighted MNSQ	0.93
Discrimination	0.70
Item Thresholds	-0.73, -0.13
Distribution of students' responses	a) 0.12%; b) 0.51%; c) 0.16%; d) 0.14%; missings = 0.08 (Grades 5 to 11)

---

Table D.4: Reliabilities, descriptive statistics, and correlation coefficients within grades for lower secondary school students (LS5).

Grade	Scale	n	M	SD	Reliability <sup>a</sup>	Omega	R	I	A	S	E	C	N	GV	UF
6	R	401	3.26	0.84	0.65	0.66 [0.59,0.74]	.59***								
6	I	406	2.84	0.89	0.75	0.75 [0.71,0.80]	.36***	.40***							
6	A	408	2.87	0.94	0.68	0.72 [0.67,0.77]	.36***	.53***	.32***						
6	S	403	2.80	0.83	0.81	0.82 [0.78,0.86]	.47***	.59***	.44***	.53***					
6	E	404	2.87	0.91	0.77	0.77 [0.73,0.81]	.40***	.66***	.43***	.47***	.54***				
6	C	403	2.51	0.89	0.77	0.77 [0.73,0.81]	.48***	.70***	.41***	.59***	.63***	.64***			
6	N	407	2.80	0.91	0.75	0.75 [0.71,0.80]	.18***	.21***	.01	.05	.15**	.11*	.20***		
6	ACH	444	-0.38	0.42	0.75	0.68 [0.62,0.74]									
7	R	386	3.19	0.83	0.68	0.70 [0.64,0.76]	.51***	.34***							
7	I	386	2.80	0.85	0.69	0.70 [0.64,0.76]	.29***	.48***	.31***	.51***	.46***				
7	A	385	2.78	0.91	0.70	0.74 [0.70,0.79]	.41***	.53***	.44***	.42***	.59***	.53***			
7	S	386	2.71	0.86	0.85	0.85 [0.82,0.88]	.49***	.53***	.37***	.42***	.59***	.53***			
7	E	386	2.80	0.93	0.78	0.78 [0.74,0.83]	.34***	.58***	.39***	.59***	.59***	.53***			
7	C	383	2.36	0.88	0.77	0.78 [0.75,0.82]	.43***	.58***	.39***	.59***	.59***	.53***			
7	N	384	2.78	0.89	0.73	0.74 [0.69,0.79]	.11*	.16**	-.08	.04	.02	-.04	.11*		
7	ACH	446	-0.14	0.51	0.77	0.65 [0.58,0.72]									
8	R	425	3.10	0.86	0.65	0.74 [0.69,0.78]	.51***	.36***							
8	I	425	2.73	0.88	0.72	0.78 [0.74,0.82]	.44***	.42***	.24***	.43***	.47***				
8	A	426	2.71	0.94	0.72	0.76 [0.72,0.80]	.37***	.49***	.39***	.57***	.55***	.54***			
8	S	425	2.65	0.83	0.86	0.86 [0.84,0.89]	.42***	.54***	.30***	.37***	.47***				
8	E	425	2.74	0.89	0.76	0.77 [0.73,0.81]	.42***	.54***	.30***	.37***	.47***				
8	C	425	2.41	0.86	0.74	0.76 [0.72,0.80]	.43***	.63***	.35***	.57***	.55***	.54***			
8	N	426	2.70	0.86	0.75	0.75 [0.70,0.80]	.26***	.38***	.20***	.29***	.30***	.32***	.33***		
8	SU	429	3.25	1.04	0.81	0.81 [0.77,0.84]	.29***	.45***	.17***	.29***	.36***	.37***	.41***	.33***	
8	UF	429	3.14	1.10	0.84	0.84 [0.81,0.87]	.05	.22***	-.10*	.11*	.06	.08	.19***	.63***	
8	ACH	430	0.04	0.46	0.77	0.65 [0.58,0.72]									

<sup>a</sup> Reliability is equal to Cronbach's alpha (RIASEC+N, SU, UF) and WLE reliability (ACH) respectively.

Note. R = realistic, A = artistic, I = investigative, S = social, E = enterprising, N = networking, S = social, ACH = achievement, SU = social utility, UF = utility for future. \*p < .05.

\*\*p < .01, \*\*\*p < .001.



**Table D.5:** Reliabilities, descriptive statistics, and correlation coefficients within grades for lower secondary school students (US9).

Grade	Scale	n	M	SD	Reliability <sup>a</sup>	Omega	R	I	A	S	E	C	N	GV
9	R	308	3.05	0.83	0.61	0.62 [0.53,0.70]								
9	I	307	2.76	0.88	0.66	0.67 [0.60,0.74]	.60***							
9	A	307	2.50	0.94	0.71	0.74 [0.69,0.79]	.29***	.36***						
9	S	305	2.49	0.88	0.85	0.86 [0.82,0.89]	.32***	.45***	.35***					
9	E	306	2.62	0.94	0.78	0.79 [0.74,0.83]	.49***	.62***	.40***	.52***				
9	C	307	2.32	0.84	0.74	0.77 [0.72,0.81]	.28***	.46***	.45***	.44***	.49***			
9	N	306	2.59	0.89	0.76	0.77 [0.72,0.82]	.36***	.58***	.34***	.60***	.63***	.61***		
9	ACH	317	0.30	0.56	0.74	0.77 [0.72,0.82]	.17**	.21***	.02	.23***	.14*	.11	.16**	
10	R	296	2.89	0.89	0.70	0.71 [0.64,0.78]								
10	I	296	2.61	0.88	0.71	0.71 [0.65,0.77]	.61***							
10	A	296	2.39	0.96	0.71	0.74 [0.69,0.80]	.21***	.32***						
10	S	297	2.44	0.86	0.84	0.85 [0.81,0.88]	.32***	.51***	.28***					
10	E	298	2.50	0.92	0.80	0.82 [0.78,0.86]	.49***	.65***	.34***	.56***				
10	C	296	2.25	0.83	0.76	0.78 [0.73,0.82]	.19***	.46***	.40***	.44***	.47***			
10	N	296	2.52	0.90	0.79	0.79 [0.75,0.84]	.41***	.62***	.28***	.56***	.64***	.57***		
10	ACH	320	0.36	0.53	0.75	0.79 [0.75,0.84]	.13*	.19***	-.11	.17**	.14*	.09	.23***	
11	R	252	2.86	0.94	0.75	0.75 [0.69,0.81]								
11	I	252	2.62	0.88	0.71	0.72 [0.65,0.78]	.65***							
11	A	252	2.33	0.97	0.70	0.75 [0.69,0.80]	.32***	.32***						
11	S	253	2.44	0.88	0.87	0.88 [0.85,0.91]	.28***	.44***	.21***					
11	E	253	2.44	0.94	0.78	0.80 [0.75,0.85]	.53***	.70***	.42***	.51***				
11	C	252	2.29	0.88	0.76	0.77 [0.71,0.82]	.35***	.48***	.46***	.39***	.47***			
11	N	252	2.51	0.93	0.84	0.84 [0.80,0.88]	.47***	.63***	.30***	.67***	.70***	.54***		
11	ACH	263	0.73	0.71	0.77	0.84 [0.80,0.88]	.23***	.31***	-.04	.32***	.24***	.21***	.37***	
12	SU	305	3.19	1.02	0.78	0.78 [0.73,0.83]								
12	UF	305	2.80	1.14	0.85	0.85 [0.82,0.88]								.54***

<sup>a</sup> Reliability is equal to Cronbach's alpha (RIASEC+N, SU, UF) and WLE reliability (ACH) respectively.  
 Note. R = realistic, A = artistic, I = investigative, S = social, E = enterprising, N = networking, S = social, ACH = achievement, SU = social utility; UF = utility for future.  
 \*p ≤ .05. \*\*p ≤ .01. \*\*\*p ≤ .001.

### D.3 Supplemental Materials for Method – Measurement Invariance

Model fits were compared by CFI and RMSEA difference testing to establish measurement invariance. Hereby, for sample sizes  $N > 300$  a change of  $\geq .010$  in CFI supplemented by a change of  $\geq .015$  in RMSEA and a change of  $\geq .030$  in SRMR between the configural and the metric model as well as a change of  $\geq .010$  in CFI supplemented by a change of  $\geq .015$  in RMSEA and a change of  $\geq .030$  in SRMR between subsequent models is generally considered to indicate non invariance (scalar and strict) (Chen, 2007, p. 501). Additionally, for changes in McDonald's NCI and gammaHat (GHAT) critical values for single factors with three items are between  $\Delta\text{NCI} < .0065$  (Meade et al., 2008, p. 586) and  $\Delta\text{NCI} < .0200$  ((G. W. Cheung & Rensvold, 2002, p. 251) and  $\Delta\text{GHAT} < .0010$ , respectively.

**Table D.6:** Goodness-of-Fit indices for testing measurement invariance per dimension of the RIASEC+N model for lower secondary students (LS5).

Scale	Invariance	CFI	RMSEE	SRMR	CFI	RMSEA	SRMR	NCI	GHAT
R	configural	0.951	0.048	0.042					
R	metric	0.952	0.046	0.044	0.001	0.002	0.002	0.0000	0.0001
R	scalar	0.951	0.045	0.044	0.001	0.001	0.000	0.0000	0.0003
I	configural	0.959	0.047	0.037					
I	metric	0.957	0.046	0.040	0.002	0.001	0.003	0.0028	0.0003
I	scalar	0.954	0.046	0.040	0.003	0.000	0.000	0.0067	0.0007
A	configural	0.966	0.046	0.036					
A	metric	0.965	0.045	0.038	0.001	0.001	0.002	0.0020	0.0002
A	scalar	0.963	0.045	0.037	0.002	0.000	0.001	0.0056	0.0006
S	configural	0.980	0.039	0.026					
S	metric	0.979	0.039	0.029	0.001	0.000	0.003	0.0032	0.0003
S	scalar	0.977	0.040	0.030	0.002	0.001	0.001	0.0076	0.0006
E	configural	0.978	0.037	0.029					
E	metric	0.976	0.037	0.032	0.002	0.000	0.003	0.0026	0.0002
E	scalar	0.972	0.038	0.032	0.004	0.001	0.000	0.0099	0.0007
C	configural	0.974	0.038	0.031					
C	metric	0.973	0.037	0.033	0.001	0.001	0.002	0.0000	0.0001
C	scalar	0.973	0.036	0.033	0.000	0.001	0.000	0.0000	0.0002
N	configural	0.946	0.056	0.045					
N	metric	0.944	0.056	0.048	0.002	0.000	0.003	0.0041	0.0004
N	scalar	0.942	0.055	0.049	0.002	0.001	0.001	0.0068	0.0008

*Note.* I = investigative, S = social, N = networking, CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual, NCI = McDonald's non-centrality index, GHAT = Gamma Hat.

**Table D.7:** Goodness-of-Fit indices for testing measurement invariance per dimension of the RIASEC+N model for upper secondary students (US9).

Scale	Invariance	CFI	RMSEE	SRMR	CFI	RMSEA	SRMR	NCI	GHAT
R	configural	0.946	0.061	0.042					
R	metric	0.940	0.062	0.050	0.006	0.001	0.008	0.0104	0.0000
R	scalar	0.938	0.061	0.052	0.002	0.001	0.002	0.0124	0.0012
I	configural	0.920	0.076	0.049					
I	metric	0.920	0.073	0.052	0.000	0.003	0.003	0.0012	0.0000
I	scalar	0.913	0.074	0.056	0.007	0.001	0.004	0.0109	0.0014
A	configural	0.909	0.089	0.052					
A	metric	0.913	0.084	0.052	0.004	0.005	0.000	0.0077	0.0000
A	scalar	0.913	0.082	0.052	0.000	0.002	0.000	0.0065	0.0000
S	configural	0.964	0.064	0.031					
S	metric	0.965	0.061	0.033	0.001	0.003	0.002	0.0018	0.0000
S	scalar	0.960	0.062	0.036	0.005	0.001	0.003	0.0096	0.0011
E	configural	0.953	0.066	0.038					
E	metric	0.943	0.070	0.044	0.010	0.004	0.006	0.0211	0.0014
E	scalar	0.941	0.069	0.045	0.002	0.001	0.001	0.0248	0.0020
C	configural	0.958	0.058	0.040					
C	metric	0.958	0.056	0.042	0.000	0.002	0.002	0.0000	0.0000
C	scalar	0.951	0.058	0.044	0.007	0.002	0.002	0.0137	0.0012
N	configural	0.953	0.065	0.041					
N	metric	0.952	0.063	0.044	0.001	0.002	0.003	0.0035	0.0000
N	scalar	0.950	0.062	0.046	0.002	0.001	0.002	0.0067	0.0000

*Note.* I = investigative, S = social, N = networking, CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual, NCI = McDonald's non-centrality index, GHAT = Gamma Hat.

#### D.4 Supplemental Materials for Results – Latent Change Score Models

Students' initial interests in school science activities are predominantly positively correlated with their initial achievement, with the exception of lower secondary (LS5) students' interest in artistic and social activities and upper secondary (US9) students' interest in artistic and conventional school science activities. Students' individual changes of interests and achievement gains were only significantly correlated in upper secondary school. Here, changes ( $\Delta_{23}$ ) of students' interests in realistic, investigative, social, enterprising, and conventional activities and achievement were positively correlated. Moreover, the analyses yielded positive cross-lagged effects of students' achievement on subsequent changes of students' interest in investigative (LS5) and networking school science activities (US9). Otherwise, lower secondary students' achievements negatively predicted students' interest in artistic activities (e.g., drawing and designing). Moreover, students' interests in investigative (US9), social (US9), enterprising (US9), and networking (LS5, US9) school science activities predicted subsequent learning gains. There were significant change-change paths between interest in realistic school science activities and achievement gains (LS5) and achievement gains and interest in realistic school science activities (US9). All correlations and regression coefficients reported above can be interpreted as small ( $< .30$ ). Across the two cohorts analyzed in this study, the findings indicated a shift from a rather weak relations between interest and achievement in lower secondary school to interest driven achievement gains and reciprocal relations between students' interests in networking school science activities and their achievement.

With regard to lower secondary students, the findings suggest that students' prior interests and achievement positively predict utility values for future life, except for achievement when controlling interest in investigative activities, and societal utility value. With regard to upper secondary students, the findings suggest that students' interest and achievement mainly positively predict their utility values. Students' interest in realistic, artistic and conventional school science activities did not predict their utility values for future and also students' interest in artistic activities did not predict their societal utility value.

**Table D.8:** Results of the bivariate latent change score models (LCSM) for lower secondary students (LS5) including utility values. Displayed are standardized  $\beta$ -regression coefficients, correlation and in brackets standard errors (SE).

	R	I	A	S	E	C	N
<b>Lagged Effects (Change <math>\leftarrow</math> Factor, <math>\beta</math>)</b>							
${}_{12}IN \leftarrow IN_1$	-0.505 (0.074)***	-0.525 (0.060)***	-0.579 (0.044)***	-0.489 (0.047)***	-0.493 (0.051)***	-0.565 (0.045)***	-0.570 (0.048)***
${}_{23}IN \leftarrow IN_2$	-0.234 (0.089)**	-0.369 (0.069)***	-0.339 (0.065)***	-0.399 (0.065)***	-0.394 (0.068)***	-0.478 (0.058)***	-0.367 (0.063)***
${}_{12}ACH \leftarrow ACH_1$	-0.362 (0.052)***	-0.367 (0.052)***	-0.359 (0.051)***	-0.361 (0.051)***	-0.372 (0.051)***	-0.365 (0.052)***	-0.375 (0.051)***
${}_{23}ACH \leftarrow ACH_2$	-0.361 (0.064)***	-0.390 (0.063)***	-0.379 (0.063)***	-0.380 (0.062)***	-0.381 (0.062)***	-0.369 (0.062)***	-0.384 (0.062)***
<b>Cross-lagged Effects (Change <math>\leftarrow</math> Factor, <math>\gamma</math>)</b>							
${}_{12}IN \leftarrow ACH_1$	-0.074 (0.042)	0.079 (0.039)*	-0.068 (0.032)*	0.038 (0.034)	0.046 (0.033)	-0.007 (0.033)	0.012 (0.034)
${}_{23}IN \leftarrow ACH_2$	-0.104 (0.058)	0.098 (0.049)*	-0.088 (0.041)*	0.054 (0.047)	0.068 (0.050)	-0.008 (0.041)	0.017 (0.048)
${}_{12}ACH \leftarrow IN_1$	0.009 (0.044)	0.058 (0.039)	-0.042 (0.038)	0.061 (0.036)	0.052 (0.037)	0.049 (0.037)	0.108 (0.038)**
${}_{23}ACH \leftarrow IN_2$	0.008 (0.042)	0.052 (0.035)	-0.041 (0.037)	0.064 (0.037)	0.054 (0.037)	0.049 (0.037)	0.106 (0.037)**
<b>Stabilities (Change <math>\leftarrow</math> Change, <math>\phi</math>)</b>							
${}_{23}IN \leftarrow {}_{12}IN$	-0.369 (0.108)**	-0.151 (0.095)	-0.291 (0.073)***	-0.181 (0.081)*	-0.215 (0.089)*	-0.212 (0.075)**	-0.307 (0.077)***
${}_{23}ACH \leftarrow {}_{12}ACH$	-0.332 (0.060)***	-0.304 (0.060)***	-0.304 (0.061)***	-0.301 (0.061)***	-0.299 (0.060)***	-0.307 (0.060)***	-0.300 (0.060)***
<b>Cross-lagged Effects (Change <math>\leftarrow</math> Change, <math>\xi</math>)</b>							
${}_{23}IN \leftarrow {}_{12}ACH$	-0.041 (0.070)	-0.119 (0.064)	0.087 (0.060)	-0.047 (0.057)	-0.100 (0.064)	0.022 (0.054)	-0.062 (0.058)
${}_{23}ACH \leftarrow {}_{12}IN$	0.144 (0.065)*	0.061 (0.059)	0.013 (0.053)	0.008 (0.057)	-0.014 (0.056)	0.060 (0.053)	0.005 (0.053)
<b>Correlations (<math>\rho</math>)</b>							
$IN_1 \sim ACH_1$	0.129 (0.058)*	0.171 (0.053)**	0.008 (0.057)	0.024 (0.049)	0.195 (0.052)***	0.112 (0.054)*	0.139 (0.054)*
${}_{12}IN \sim {}_{12}ACH$	0.090 (0.065)	0.119 (0.071)	-0.054 (0.059)	-0.013 (0.057)	-0.058 (0.061)	-0.025 (0.059)	-0.005 (0.061)
${}_{23}IN \sim {}_{23}ACH$	-0.081 (0.079)	0.043 (0.073)	-0.055 (0.057)	0.028 (0.062)	-0.029 (0.067)	-0.010 (0.060)	0.058 (0.066)
<b>Regression Estimates (Utility Values <math>\leftarrow</math> Factor)</b>							
$SU \leftarrow IN_3$	0.344 (0.065)***	0.368 (0.060)***	0.161 (0.067)*	0.265 (0.056)***	0.371 (0.051)***	0.404 (0.060)***	0.297 (0.061)***
$SU \leftarrow ACH_3$	0.226 (0.057)***	0.210 (0.054)***	0.222 (0.053)***	0.258 (0.052)***	0.172 (0.053)**	0.185 (0.053)***	0.250 (0.052)***
$UF \leftarrow IN_3$	0.283 (0.063)***	0.470 (0.057)***	0.190 (0.068)**	0.309 (0.054)***	0.295 (0.053)***	0.335 (0.063)***	0.370 (0.060)***
$UF \leftarrow ACH_3$	0.327 (0.055)***	0.107 (0.055)	0.306 (0.050)***	0.173 (0.055)**	0.259 (0.051)***	0.270 (0.050)***	0.160 (0.053)**
<b>Intercorrelations (Utility Values)</b>							
$SU \sim UF$	0.571 (0.038)***	0.547 (0.040)***	0.594 (0.036)***	0.587 (0.037)***	0.559 (0.038)***	0.566 (0.038)***	0.568 (0.038)***
<b>Model Fit</b>							
CFI	0.948	0.949	0.950	0.973	0.974	0.972	0.938
RMSEA	0.046	0.049	0.051	0.042	0.037	0.036	0.056
SRMR	0.042	0.037	0.039	0.031	0.031	0.033	0.047

Note. R = realistic, I = investigative, A = artistic, S = social, E = enterprising, C = conventional, N = networking, ACH = achievement, UF = utility for future life, SU = societal utility. Indices represent measurement waves; Significance level: \*\*\*p < .001, \*\*p < .01, \*p < .05, .p < .1.

**Table D.9:** Results of the bivariate latent change score models (LCSM) for lower secondary students (LS5) including utility values. Displayed are standardized  $\beta$ -regression coefficients, correlation and in brackets standard errors (SE).

Lagged Effects (Change $\leftarrow$ Factor, $\beta$ )									
${}_{12}IN \leftarrow IN_1$	-0.366 (0.086)***	-0.285 (0.121)*	-0.447 (0.066)***	-0.493 (0.055)***	-0.466 (0.067)***	-0.534 (0.060)***	-0.398 (0.080)***		
${}_{23}IN \leftarrow IN_2$	-0.325 (0.109)**	-0.404 (0.095)***	-0.280 (0.088)**	-0.511 (0.062)***	-0.371 (0.080)***	-0.315 (0.085)***	-0.428 (0.070)***		
${}_{12}ACH \leftarrow ACH_1$	-0.533 (0.047)***	-0.544 (0.047)***	-0.519 (0.047)***	-0.553 (0.047)***	-0.545 (0.046)***	-0.530 (0.046)***	-0.539 (0.047)***		
${}_{23}ACH \leftarrow ACH_2$	-0.123 (0.068)	-0.147 (0.069)*	-0.119 (0.067)	-0.153 (0.067)*	-0.131 (0.068)	-0.117 (0.068)	-0.148 (0.070)*		
Cross-lagged Effects (Change $\leftarrow$ Factor, $\gamma$ )									
${}_{12}IN \leftarrow ACH_1$	0.051 (0.053)	0.025 (0.069)	-0.040 (0.050)	0.051 (0.050)	0.039 (0.046)	0.019 (0.046)	0.150 (0.055)**		
${}_{23}IN \leftarrow ACH_2$	0.044 (0.046)	0.021 (0.056)	-0.036 (0.045)	0.045 (0.044)	0.040 (0.048)	0.019 (0.046)	0.111 (0.041)**		
${}_{12}ACH \leftarrow IN_1$	0.079 (0.053)	0.106 (0.052)*	-0.015 (0.047)	0.175 (0.042)***	0.138 (0.046)**	0.077 (0.046)	0.135 (0.044)**		
${}_{23}ACH \leftarrow IN_2$	0.080 (0.055)	0.099 (0.051)	-0.014 (0.043)	0.151 (0.038)***	0.126 (0.042)**	0.066 (0.040)	0.126 (0.044)**		
Stabilities (Change $\leftarrow$ Change, $\varphi$ )									
${}_{23}IN \leftarrow {}_{12}IN$	-0.152 (0.134)	-0.163 (0.152)	-0.301 (0.101)**	-0.052 (0.081)	-0.297 (0.094)**	-0.173 (0.102)	-0.154 (0.096)		
${}_{23}ACH \leftarrow {}_{12}ACH$	-0.255 (0.072)***	-0.259 (0.074)***	-0.264 (0.071)***	-0.263 (0.071)***	-0.256 (0.070)***	-0.266 (0.071)***	-0.253 (0.071)***		
Cross-lagged Effects (Change $\leftarrow$ Change, $\xi$ )									
${}_{23}IN \leftarrow {}_{12}ACH$	0.164 (0.076)*	0.078 (0.084)	0.070 (0.072)	-0.044 (0.066)	0.001 (0.067)	0.070 (0.074)	-0.046 (0.067)		
${}_{23}ACH \leftarrow {}_{12}IN$	0.002 (0.095)	0.021 (0.115)	-0.024 (0.080)	0.048 (0.065)	0.072 (0.080)	0.014 (0.080)	-0.071 (0.093)		
Correlations ( $\rho$ )									
$IN_1 \sim ACH_1$	0.179 (0.071)*	0.245 (0.067)***	0.035 (0.062)	0.203 (0.057)***	0.170 (0.062)**	0.121 (0.062)	0.162 (0.062)**		
${}_{12}IN \sim {}_{12}ACH$	-0.051 (0.075)	0.150 (0.093)	-0.071 (0.070)	0.095 (0.070)	0.026 (0.068)	0.042 (0.071)	0.088 (0.073)		
${}_{23}IN \sim {}_{23}ACH$	0.187 (0.082)*	0.261 (0.092)**	0.027 (0.086)	0.159 (0.079)*	0.186 (0.080)*	0.191 (0.080)*	0.133 (0.082)		
Regression Estimates (Utility Values $\leftarrow$ Factor)									
$SU \leftarrow IN_2$	0.336 (0.076)***	0.205 (0.076)**	0.133 (0.070)	0.182 (0.078)*	0.412 (0.058)***	0.196 (0.069)**	0.227 (0.079)**		
$SU \leftarrow ACH_3$	0.346 (0.079)***	0.342 (0.080)***	0.420 (0.078)***	0.358 (0.080)***	0.293 (0.077)***	0.369 (0.080)***	0.343 (0.082)***		
$UF \leftarrow IN_3$	0.154 (0.085)	0.467 (0.070)***	0.045 (0.080)	0.289 (0.075)***	0.267 (0.067)***	0.079 (0.070)	0.428 (0.072)***		
$UF \leftarrow ACH_3$	0.383 (0.076)***	0.225 (0.085)**	0.431 (0.076)***	0.319 (0.080)***	0.347 (0.075)***	0.407 (0.077)***	0.261 (0.082)**		
Intercorrelations (Utility Values)									
$SU \sim UF$	0.527 (0.047)***	0.502 (0.053)***	0.534 (0.045)***	0.524 (0.046)***	0.481 (0.052)***	0.534 (0.046)***	0.502 (0.049)***		
Model Fit									
CFI	0.937	0.911	0.906	0.949	0.942	0.950	0.939		
RMSEA	0.059	0.071	0.080	0.065	0.064	0.055	0.065		
SRMR	0.050	0.052	0.052	0.038	0.043	0.044	0.047		

Note: R = realistic, I = investigative, A = artistic, S = social, E = enterprising, C = conventional, N = networking, ACH = achievement, UF = utility for future life, SU = societal utility. Indices represent measurement waves. Significance level: \*\*\*p < .001, \*\*p < .01, \*p < .05, .p < .1.

## Erklärung

Hiermit erkläre ich, dass die vorliegende Dissertation – abgesehen von der Beratung durch meine Betreuer – nach Inhalt und Form meine eigene ist. Sie wurde weder im Ganzen noch in Teilen an einer anderen Stelle im Rahmen eines Promotionsverfahrens vorgelegt. Teile dieser Arbeit wurden bereits in Form von Zeitschriftenartikeln publiziert bzw. sind zur Publikation eingereicht. Die Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden. Es wurde mir kein akademischer Grad entzogen.

Kiel, den 20. Januar 2021

Lars Höft