DiVA
http://www.diva-portal.org

This is the published version of a paper published in *Journal of Research in Science Teaching*.

Citation for the original published paper (version of record):

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
https://doi.org/10.1002/tea.21475

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-152502

Received: 3 July 2017 | Revised: 2 May 2018 | Accepted: 7 May 2018

DOI: 10.1002/tea.21475

RESEARCH ARTICLE

WILEY | JRST

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

Lars Höft[1] | **Sascha Bernholt[1]** | **Janet S. Blankenburg[1]** | **Mikael Winberg[2]**

[1] Department of Chemistry Education, Leibniz-Institute for Science and Mathematics Education (IPN), Kiel, Germany

[2] Department of Science and Mathematics Education, Umeå University, Umeå, Sweden

Correspondence
Lars Höft, Olshausenstraße 62, D-24118 Kiel, Germany.
Email: hoeft@ipn.uni-kiel.de

Abstract

The development of students' interest in school science activities, their understanding of central chemical concepts, and the interplay between both constructs across Grades 5–11 were analyzed in a cross-sectional paper-and-pencil study ($N = 2{,}510$, mean age 11–17 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, for example, school grades or general subject-related interest, and focus on science as an integrated subject instead of specific disciplines, for example, chemistry. For this article, 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, Höffler, & Parchmann, 2014) 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

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2018 The Authors. Journal of Research in Science Teaching published by Wiley Periodicals, Inc. on behalf of National Association for Research in Science Teaching

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, and 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
chemistry education, conceptual understanding, interest

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, Sinatra, Jones, Taasoobshirazi, & Lombardi, 2014) as well as study success (Bengmark, Thunberg, & Winberg, 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, Krapp, & Winteler, 1992). Furthermore, uncertainty remains on how the association between interest and knowledge develops over the time of schooling (Denissen, Zarrett, & Eccles, 2007; Rotgans & Schmidt, 2017). 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, Ortony, & Revelle, 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 this research, we examine the development of and the relationship between students' interest and their chemistry knowledge across Grades 5–11. 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 linguistical 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, Höffler, Blankenburg, Peters, & Parchmann, 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.

2 | THEORETICAL BACKGROUND

2.1 | 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 multidimensional construct (Knogler, Harackiewicz, Gegenfurtner, & Lewalter, 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, that is, 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 (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 (Ainley & Ainley, 2011; Fredricks, Hofkens, Wang, Mortenson, & Scott, 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, 2002a). 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, for example, 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, 1985) and suggests that this predisposition will only be realized if a person satisfies the need of competence, autonomy (Tsai, Kunter, Lüdtke, Trautwein, & Ryan, 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 (Deci & Ryan, 1985; Guay, Ratelle, & Chanal, 2008; Jang, Reeve, & Deci, 2010). Although interest may initially be triggered by situational features that appeal to the individual, for example, 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, that is, 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, Hoffmann, Langeheine, Rost, & Sievers, 1998), or activities (Schmidt, Rosenberg, & Beymer, 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., 2014, 2016). The adapted version of the RIASEC model takes up the multifaceted structure of the original model, that is, 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 ("doing experiments guided by an instruction"), investigative ("plan experiments to investigate something"), artistic ("draw an observation"), social ("explaining something to fellow students"), enterprising ("organizing a small chemistry project"), conventional ("tabling the results of an experiment"), and networking ("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., 2014). 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 et al., 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, (i) the quality and type of instruction (Kunter et al., 2014; Schiefele & Schaffner, 2015; Tröbst, Kleickmann, Lange-Schubert, Rothkopf, & Möller, 2016); (ii) developmental issues (students in adolescence prioritize new developmental tasks); and (iii) 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.

2.2 | 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, for example, the four-phase model of interest development (Hidi & Renninger, 2006; Renninger & Hidi, 2011), or broader frameworks such as the social cognitive career theory (Lent, Brown, & Hackett, 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). Similarly, according to Eccles, Fredricks, and Epstein (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 (Ainley et al., 2002) as well as self-regulation (Lee, Lee, & Bong, 2014). In a meta-analysis, Schiefele et al. (1992) found a mean correlation of $r = .30$ (95% CI: .04–.56) between domain-specific interest and achievement. Other studies indicate that this correlation seems to increase over time (Denissen et al., 2007; Kim, Jiang, & Song, 2015; Köller, Baumert, & Schnabel, 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, Lüdtke, & Schroeders, 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, Schroeders, Lüdtke, & Marsh, 2015; Jansen et al., 2016; Krapp & Prenzel, 2011; Wigfield & Cambria, 2010). Longitudinal studies indicate small reciprocal effects between interest and achievement (Garon-Carrier et al., 2016; Harackiewicz, Durik, Barron, Linnenbrink-Garcia, & Tauer, 2008; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005), but results regarding the causal relation are still inconclusive (Jõgi, Kikas, Lerkkanen, & Mägi, 2015; Rotgans & Schmidt, 2017; Viljaranta, Tolvanen, Aunola, & Nurmi, 2014).

2.3 | 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, Brown, & Cocking, 2000). In many countries, novel educational standards have adopted this view by incorporating disciplinary core concepts (Bernholt, Neumann, & Nentwig, 2012; NGSS Lead States, 2013). Hence, modern curricula underlie the assumption that "learning ought to be coordinated and sequenced along conceptual trajectories (Driver, Leach, Scott, & Wood-Robinson, 1994), developmental corridors (Brown, 1997), and learning progressions" (Duschl, Maeng, & Sezen, 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, Schweingruber, & Shouse, 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 (Löfgren & Helldén, 2009; Morell, Collier, Black, & Wilson, 2017; Stefani & Tsaparlis, 2009), chemical reaction (cf., Talanquer, 2008; Taskin & Bernholt, 2014), or energy (cf., Herrmann-Abell & DeBoer, 2018; Neumann, Viering, Boone, & Fischer, 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, Hill, Black, & Lipsey, 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, Feltovich, & Glaser, 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" (Pellegrino, 2012, p. 84). 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 (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, Chudowsky, & Glaser, 2001).

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), this 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 5–11) 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., 2014). 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, that is, from Grades 5–11. While the assessment of students' understanding of or interest in different science topics or contexts is also a plausible account (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, Pekrun, Dicke, & Goetz, 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 5–11?

   The instrument based on the adapted version of the RIASEC model has been used in different grade levels in different studies (Blankenburg et al., 2016; Dierks et al., 2014), 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.

4 | DESIGN AND ANALYSIS PROCEDURE

4.1 | Instruments

4.1.1 | Interest

To capture students' interest in school science activities, we used the RIASEC + N instrument developed by Dierks et al. (2014) and adapted by Blankenburg et al. (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 (Supporting Information Table S1). 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 adaptions of the published version of the questionnaire were made for this study.

In terms of validity (Joint Committee, 2014; Messick, 1994), 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., 2014). 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).

4.1.2 | Conceptual understanding

As we intended to investigate the conceptual understanding of a diverse sample, ranging from Grades 5 to 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 (Figure 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, Bernholt, Liu, Neumann, & Parchmann, 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, Alonzo, Schwab, & Wilson, 2006).

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, Neumann, Bernholt, Liu, & Parchmann, 2016; Energy: Opitz, 2017).

The test construction followed a common item anchor design with grade specific items. 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

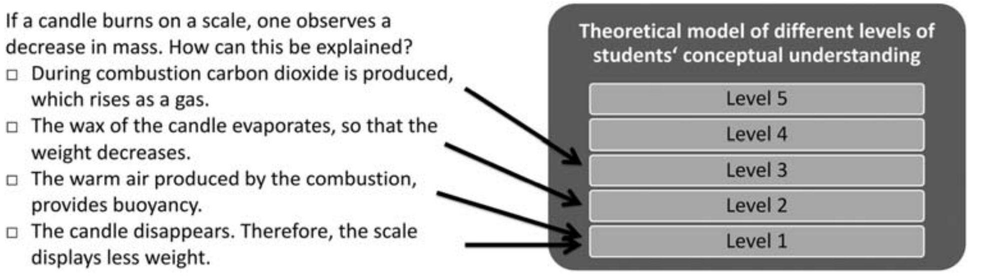

FIGURE 1 Example of an ordered multiple choice (OMC) item for the basic concept "chemical reaction." Each option of the item corresponds to a specific level of students' conceptual understanding and was scored accordingly based on a partial credit model

selected to addresses 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, & Neumann, 2014). Besides interest in school science activities and conceptual understanding, the final questionnaire included questions regarding demographic variables (gender, age, first language, and SES) and several other constructs (self-concept, motivation, and epistemic beliefs) that are not in focus of this article.

4.2 | 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 langue at home) from 136 classes, ranging from Grades 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, Kampa, & Köller, 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 (Supporting Information Table S2).

As the constitution and also the quality of school lessons depends on multiple factors, for example, 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 & 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 and 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, that is, 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 min 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.

4.3 | Data analysis

4.3.1 | Interest

To analyze the construct validity of the RIASEC + N model, confirmatory factor analysis (CFA) models with a maximum likelihood estimator were estimated (Brunner, Nagy, & Wilhelm, 2012; MacCallum & Austin, 2000) in Mplus 7.4 (Muthén & Muthén, 1998–2015). 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) >.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, Balla, & McDonald, 1998).

The questionnaire data met the normality assumption with absolute values for skewness and kurtosis of <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 (α = .67–.87) after deleting one item per dimension because of poor item-scale correlation (Table 1). Additionally, score reliabilities (ρ) defined as the ratio of true variance to observed variance were calculated for each scale

TABLE 1 Reliabilities (Cronbach's α, McDonald's ω, and reliability coefficient ρ of the latent trait reliability model (LTRM)) and descriptive statistics for interest

Scale	*n* Items[a]	α	ω	ρ	ρ 95% CI	*M*[b]	*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.
M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for the reliability coefficient ρ.

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.

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 et al., 2016; Dierks et al., 2014, 2016). This model had the following characteristics: (i) each activity measure had a nonzero loading on one first-order RIASEC + N factor but zero loadings on the others; (ii) correlations among the factors were allowed, and (iii) 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.[1]

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 (Supporting Information, Figure S1). To compare students from different grades by analyzing latent values, strong measurement invariance had to be ensured (Wu, Li, & Zumbo, 2007). First, seven CFAs were conducted for each grade separately, indicating an appropriate fit of the model in each grade (Supporting Information Table S3).

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

[1]Also a one-dimensional model ($\chi^2 = 6{,}027.755$, $p < .001$, df = 188, CFI = .720, TLI = .687, RMSEA = .117, SRMR = .078) and a two-dimensional model ($\chi^2 = 5{,}427.924$, $p < .001$, df = 187, CFI = .749, TLI = .718, RMSEA = .111, SRMR = .074), which distinguishes between students' interest in school science activities involving group work or individual work, respectively, were fitted to the data. However, both models showed poor fits, suggesting that the data are best represented by the seven-dimensional model.

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 (Supporting Information Table S4). 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 (Cheung & Rensvold, 2002). Using this criterion, we could demonstrate scalar measurement invariance, as ΔCFI was smaller than .01 and ΔRMSEA was smaller than .015 between all consecutive models (Supporting Information Table S4). 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).

4.3.2 | Conceptual understanding

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[2] multi-group generalized partial credit model (MG-GPCM, as implemented in the R package TAM; Test Analysis Modules; Kiefer, Robitzsch, & Wu, 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 .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 cutoff values (WMNSQ from 0.7 to 1.3; Wright, Linacre, Gustafson, & Martin-Lof, 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 5–11.

4.3.3 | 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 ($\chi^2 = 3{,}198.324$, $p < .001$, df = 1,476, CFI = .917, TLI = .909, RMSEA = .060, SRMR = .079); the Mplus syntax of the MIMIC approach is included in Supporting Information.

[2]As the test construction covered three basic concepts, also a three-dimensional model was fitted to the data. However by comparing the goodness-of-fit of the one- and the three-dimensional models, model fit parameters were in favor of the one-dimensional model, suggesting that the items cover a single dimension of conceptual understanding, not three distinct dimensions.

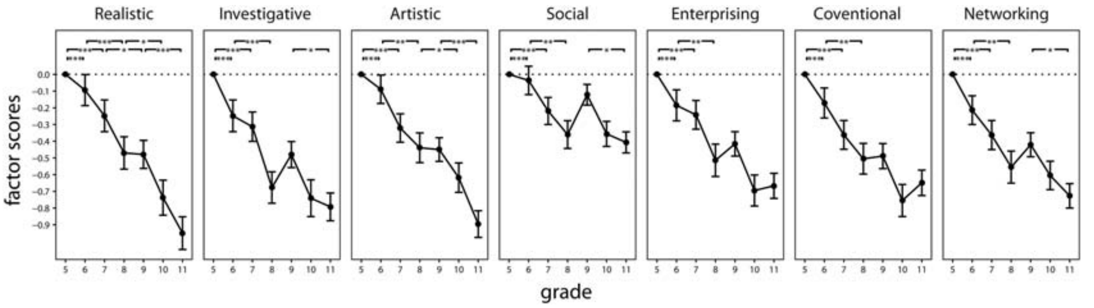

FIGURE 2 Mean factor scores and corresponding standard errors of students' interest in school science activities across Grades 5–11 based on the MG-CFA. Values in Grade 5 (reference group) are set to 0 due to identification reasons. Significant differences between grades are reported for comparisons between grades within two years of each other (* $p < .05$; ** $p < .01$; *** $p < .001$)

5 | RESULTS

5.1 | Trends over time

As the main goal of this study is to compare the interrelation of interest and conceptual understanding across Grades 5–11, 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 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 ($N = 2{,}294$), thus excluding students who decided to drop out from chemistry.

5.1.1 | 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 (Figure 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.

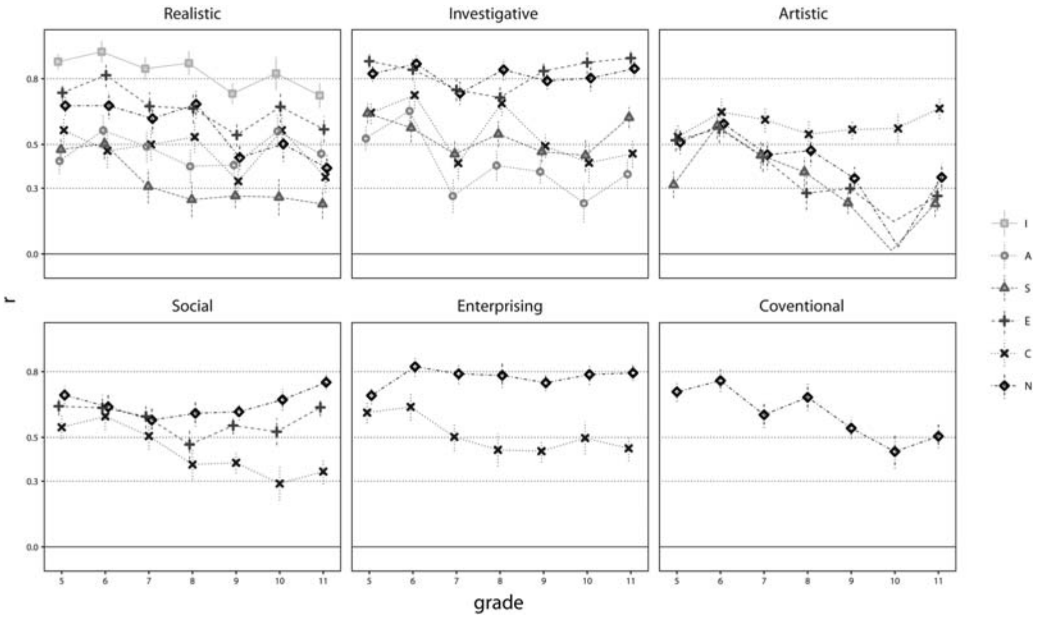

FIGURE 3 Latent correlations between the seven dimensions of students' interest in school science activities according to the RIASEC + N model (MG-CFA) across Grades 5–11 (R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; N = Networking). All correlation 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

With regard to the interrelations of students' interest between the seven dimensions of the RIASEC + N model, certain patterns emerge (Figure 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.

5.1.2 | 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 ($r(2{,}476) = .68$, $p < .001$). 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 (Supporting Information, Figure S2). We then compared the mean WLE parameter across the Grades 5–11. Based on an ANOVA, there was a significant effect of the grade level on students' WLE parameter, $F(6, 890.46) = 387.56$, $p < .001$, $\eta^2 = 0.50$. This result is supported by planned contrasts, indicating that students' understanding significantly progresses between Grades 6 and 11, while the mean difference between Grades 5 and 6 is nonsignificant (Table 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 ($d = 0.84$) and 8 to 9 ($d = 0.73$), respectively, followed by the transition from 7 to 8 with an (almost) middle-sized effect ($d = 0.43$). The gains from Grades 9 to

TABLE 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[a]
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).

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 2; Bloom et al., 2008). The calculated effect sizes in this study are substantially larger for transitions between Grades 6 and 10, while the effect size for the transition from Grades 10 to 11 is comparable. The effect size for the transition from Grades 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 Grades 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 ($0.18 < d < 0.84$) when compared to the benchmarks in Bloom et al. (2008).

5.2 | 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 (Figure 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

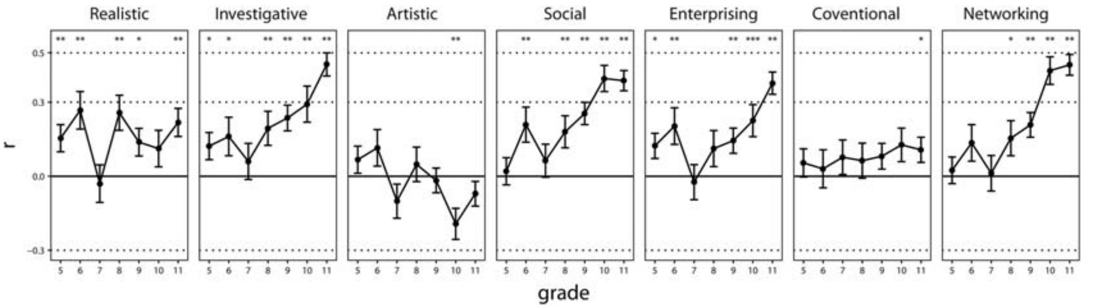

FIGURE 4 Latent correlations between students' conceptual understanding (WLE parameter of the MG-GPCM) and students' interest in school science activities according to the RIASEC + N model (MIMIC model). Correlation coefficients were tested to be significantly different from 0 (*$p < .05$; **$p < .01$; ***$p < .001$)

11 (additionally investigative and enterprising). Here, correlation coefficients fall into the range of .3–.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.

6 | DISCUSSION

The goal of this study was (i) to investigate the validity and reliability of the measures for students' interests in school science activities as well as students' conceptual understanding across Grades 5–11, (ii) to analyze the development of students' interest in school science activities and their conceptual understanding across this grade span, and (iii) 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).

6.1 | 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 5–11 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 et al., 2016) and, moreover, enhances previous findings through the conjunction with learning outcomes.

6.2 | 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, Bernholt, & Ropohl, 2016), 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, Rakoczy, Hattie, Besser, & Klieme, 2013; Keller, Neumann, & Fischer, 2016; Kunter et al., 2014).

6.3 | Development of conceptual understanding

While students' interests in school science activities decrease across the grades, their conceptual understanding increases. These learning gains are primarily achieved in Grades 6–9 and levels out from Grade 9 and onward, 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 and 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 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 2). This result that can probably be traced back to specific discrepancies between the meta-analysis and this study, for example, 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 5–11, supporting the assumption that the developed instrument provides a thorough measure of students' conceptual understanding across this grade span.

6.4 | Interplay between interest in school science activities and conceptual understanding

The main findings obtained in this study are summarized in Figure 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 onward, 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 4 also reveals differing interrelations between conceptual understanding and interest: (i) small to middle-sized, positive relations, increasing from Grade 7 to 11, between interest in investigative, social, enterprising, and networking activities and conceptual understanding; (ii) small positive, but almost stable relations between interest in realistic activities and conceptual understanding; and (iii) 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 (i), 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., 2014). 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 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 (Ainley & Ainley, 2011; Harackiewicz, Smith, & Priniski, 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 (ii), 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 nor 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, Zimmerman, Croker, & Härtig, 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, Grube, & Bögeholz, 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 (cf. Hofstein, 2004; Pickering, 1980) which has been shown to impact on conceptual understanding (Duit, Treagust, & Widodo, 2008; Hofstein & Lunetta, 2004).

With regard to (iii), 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.

6.5 | Limitations

Although all 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 and probably their interest as indicated by the questionnaire do not cover the full variance that could be observed in the student population.

6.6 | 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, for example, 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 impact 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, 1985) and of competence (Eccles, 2009; cf., 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).

ORCID

Lars Höft http://orcid.org/0000-0002-9040-8995

REFERENCES

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. https://doi.org/10.1002/tea.21036

Ainley, M., & Ainley, J. (2011). 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. https://doi.org/10.1016/j.cedpsych.2010.08.001

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. https://doi.org/10.1037//0022–0663.94.3.545

Alonzo, A. C., & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education*, *93*(3), 389–421. https://doi.org/10.1002/sce.20303

Anderhag, P., Wickman, P., Bergqvist, K., Jakobson, B., Hamza, K., & 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. https://doi.org/10.1002/sce.21231

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. https://doi.org/10.1080/10508406.2012.730082

Bengmark, S., Thunberg, H., & Winberg, T. M. (2017). Success-factors in transition to university mathematics. *International Journal of Mathematical Education in Science and Technology*, *48*(7), 988–1001. https://doi.org/10.1080/0020739X.2017.1310311

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. https://doi.org/10.1016/j.learninstruc.2010.06.002

Bernholt, S., Neumann, K., & Nentwig, P. (Eds.) (2012). *Learning outcomes in science education: Making it tangible*. Münster: Waxmann. http://www.content-select.com/index.php?id=bib_view&ean= 9783830976448

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. https://doi.org/10.18452/8207

Blankenburg, J. S., Höffler, T. N. M., & Parchmann, I. (2016). Fostering today what is needed tomorrow: Investigating students' interest in science. *Science Education*, *100*(2), 364–391. https://doi.org/10.1002/sce.21204

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. https://doi.org/10.1080/19345740802400072

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience and school*. Washington, D.C.: 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. https://doi.org/10.1207/s15326977ea1101_2

Brown, A. L. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist*, *52*(4), 399–413. https://doi.org/10.1037/0003-066X.52.4.399

Browne, M. W., & Cudeck, R. (1993). Alternative ways of testing structural equation models. In K. A. Bollen & S. J. Long (Eds.), *Sage focus editions. Testing structural equation models* (Vol. *154*, pp. 136–162). Newbury Park, CA: Sage.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, *80*(4), 796–846.https://doi.org/10.1111/j.1467-6494.2011.00749.x

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.https://doi.org/10.1207/S15328007SEM0902_5

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 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. https://doi.org/10.1016/S0160-791X(01)00051-3

Cordova, J. R., Sinatra, G. M., Jones, S. H., Taasoobshirazi, 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. https://doi.org/10.1016/j.cedpsych.2014.03.006

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Boston, MA: Springer US.

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. https://doi.org/10.1111/j.1467-8624.2007.01007.x

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. https://doi.org/10.1080/09500693.2016.1138337

Dierks, P. O., Höffler, T. N., & Parchmann, I. (2014). Profiling interest of students in science: Learning in school and beyond. *Research in Science & Technological Education*, *32*(2), 97–114. https://doi.org/10.1080/02635143.2014.895712

Ding, N., & Harskamp, E. G. (2011). Collaboration and peer tutoring in chemistry laboratory education. *International Journal of Science Education*, *33*(6), 839–863. https://doi.org/10.1080/09500693.2010.498842

DiSessa, A., & Wagner, J. (2005). What coordination has to say about transfer. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 121–154). Greenwich, CT: Information Age Publishing.

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. https://doi.org/10.1080/03057269408560040

Duit, R., Treagust, D. F., & Widodo, A. (2008). Teaching science for conceptual change: Theory and practice. In S. Vosniadou (Ed.), *Educational psychology handbook series. International handbook of research on conceptual change* (pp. 629–646). New York: Routledge.

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. https://doi.org/10.1080/03057267.2011.604476

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school*. Washington, D.C.: 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. https://doi.org/10.1080/00461520902832368

Eccles, J. S., Fredricks, J. A., & Epstein, A. (2015). Understanding well-developed interests and activity commitment. In A. K. Renninger, M. Nieswandt, & S. Hidi (Eds.), *Interest in mathematics and science learning* (pp. 315–330). Washington, D.C.: American Educational Research Association. https://doi.org/10.3102/978-0-935302-42-4_18

Field, A. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock 'n' roll*. (4th ed.). MobileStudy. Los Angeles, London, New Delhi: Sage.

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. https://doi.org/10.1002/tea.21419

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. https://doi.org/10.1037/a0026895

Garon-Carrier, G., Boivin, M., Guay, F., Kovas, Y., Dionne, G., Lemelin, J.-P., … 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. https://doi.org/10.1111/cdev.12458

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. https://doi.org/10.1037/a0012758

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. https://doi.org/10.1021/ed3006192

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. https://doi.org/10.1002/tea.21312

Hadenfeldt, J. C., Repenning, B., & Neumann, K. (2014). Die kognitive Validität von Ordered Multiple Choice Aufgaben zur Erfassung des Verständnisses von Materie [The cognitive validity of ordered multiple choice items to assess students' understanding of matter]. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *20*(1), 57–68. https://doi.org/10.1007/s40573-014-0003-7

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. https://doi.org/10.1037/0022-0663.100.1.105

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 220–227. https://doi.org/10.1177/2372732216655542

Harks, B., Rakoczy, K., Hattie, J., Besser, M., & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: The role of feedback's perceived usefulness. *Educational Psychology*, *34*(3), 269–290. https://doi.org/10.1080/01443410.2013.785384

Harris, C. J., Penuel, W. R., D'Angelo, C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., … 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. https://doi.org/10.1002/tea.21263

Häussler, P., Hoffman, 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. https://doi.org/10.1080/0950069980200207

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. https://doi.org/10.1002/tea.21411

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. https://doi.org/10.3102/00346543070002151

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hofstein, A. (2004). The laboratory in chemistry education: Thirty years of experience with developments, implementation, and research. *Chemistry Education Research and Practice*, *5*(3), 247–264. https://doi.org/10.1039/B4RP90027H

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*(1), 28–54. https://doi.org/10.1002/sce.10106

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Lutz, FL: Psychological Assessment Resources.

Holstermann, N., Grube, D., & Bögeholz, S. (2010). Hands-on activities and their influence on students' interest. *Research in Science Education*, *40*(5), 743–757. https://doi.org/10.1007/s11165-009-9142-0

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. https://doi.org/10.1080/10705519909540118

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. https://doi.org/10.1037/a0019682

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. https://doi.org/10.1016/j.cedpsych.2016.05.004

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. https://doi.org/10.1037/edu0000021

Jõgi, 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. https://doi.org/10.1016/j.lindif.2015.03.018

Joint Committee (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Keller, M. M., Neumann, K., & Fischer, H. E. (2016). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *Journal of Research in Science Teaching*. Advance online publication: https://doi.org/10.1002/tea.21378

Kiefer, T., Robitzsch, A., & Wu, M. (2017). TAM: Test analysis modules. Retrieved from https://CRAN.R-project.org/package=TAM

Kim, S.-I., Jiang, Y., & Song, J. (2015). The effects of interest and utility value on mathematics engagement and achievement. In A. K. Renninger, M. Nieswandt, & S. Hidi (Eds.), *Interest in mathematics and science learning* (pp. 63–78). American Educational Research Association. https://doi.org/10.3102/978-0-935302-42-4_4

KMK (2004). Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss: Luchterhand Verlag GmbH. Retrieved from http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Chemie.pdf

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. https://doi.org/10.1016/j.cedpsych.2015.08.004

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. https://doi.org/10.2307/749801

Krapp, A. (2002a). An education-psychological theory of interest and its relation to SDT. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 405–427). Rochester, NY: University of Rochester Press.

Krapp, A. (2002b). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, *12*(4), 383–409. https://doi.org/10.1016/S0959-4752(01)00011-1

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, *33*(1), 27–50. https://doi.org/10.1080/09500693.2010.518645

Kühn, S. M. (2016). Aufgaben in (zentralen) Abschlussprüfungen. Theoretische und empirische Perspektiven auf ein interdisziplinäres Forschungsfeld [Tasks in (centrally administered) graduation exams]. In S. Keller & C. Reintjes (Eds.), *Aufgaben als Schlüssel zur Kompetenz: Didaktische Herausforderungen, wissenschaftliche Zugänge und empirische Befunde* (1st ed., pp. 73–92). Münster, New York: Waxmann Verlag GmbH.

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2014). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*(3), 805–820. https://doi.org/10.1037/a0032583

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. https://doi.org/10.1016/j.cedpsych.2014.02.002

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. https://doi.org/10.1006/jvbe.1994.1027

Leucht, M., Kampa, N., & Köller, O. (Eds.) (2016). *Fachleistungen von Abiturienten: Vergleich allgemeinbildender und beruflicher Gymnasien in Schleswig-Holstein: LISA 6. Münster*. New York: Waxmann. Retrieved from http://www.content-select.com/index.php?id=bib_view&ean=9783830984344

Liu, X. (2012). Using learning progression to organize learning outcomes: Implications for assessment. In S. Bernholt, K. Neumann, & P. Nentwig (Eds.), *Learning outcomes in science education: Making it tangible* (pp. 309–325). Münster: Waxmann.

Liu, X., & Lesniak, K. M. (2005). Students' progression of understanding the matter concept from elementary to high school. *Science Education*, *89*(3), 433–450. https://doi.org/10.1002/sce.20056

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. https://doi.org/10.1080/09500690802154850

Luce, M. R., & Hsi, S. (2015). Science-relevant curiosity expression and interest in science: An exploratory study. *Science Education*, *99*(1), 70–97. https://doi.org/10.1002/sce.21144

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*(1), 201–226. https://doi.org/10.1146/annurev.psych.51.1.201

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1998). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, *103*(3), 391–410. https://doi.org/10.1037/0033-2909.103.3.391

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. https://doi.org/10.1037/a0019227

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. https://doi.org/10.1037/dev0000146

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. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Menthe, J., & Parchmann, I. (2015). Getting involved: Context-based learning in chemistry education. In M. Kahveci & M. Orgill (Eds.), *Affective dimensions in chemistry education*. (pp. 51–67). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45085-7_3

Messick, S. (1994). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *ETS Research Report Series*, *1994*(2), i–28. https://doi.org/10.1002/j.2333-8504.1994.tb01618.x

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. https://doi.org/10.1037/0003-066X.50.9.741

Messick, S. (1998). Test validity: a matter of consequence. *Social Indicators Research*, *45*(1/3), 35–44. https://doi.org/10.1023/A:1006964925094

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. https://doi.org/10.1007/s40573-014-0010-8

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. https://doi.org/10.1002/tea.21397

Muthén, L. K., & Muthén, B. O. (1998–2015). *Mplus user's guide: Statistical analysis with latent variables*. Muthén & Muthén.

National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, D.C.: National Academy Press.

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. https://doi.org/10.1002/tea.21061

NGSS Lead States (2013). *Next generation science standards*. Washington, D.C.: National Academies Press.

OECD (2005). *Formative assessment: Improving learning in secondary classrooms*. Paris: OECD Publishing and Centre for Educational Research and Innovation. Retrieved from https://doi.org/10.1787/9789264007413-en

Opitz, S. T. (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*, https://doi.org/10.12973/eurasia.2017.00703a

Osborne, J., & Collins, S. (2000). *Pupils' & parents' views of the school science curriculum.* London: Kings College.

Palmer, D. H. (2009). Student interest generated during an inquiry skills lesson. *Journal of Research in Science Teaching*, *46*(2), 147–165. https://doi.org/10.1002/tea.20263

Pellegrino, J. W. (2012). The design of an assessment system focused on student achievement: A learning sciences perspective on issues of competence, growth, and measurement. In S. Bernholt, K. Neumann, & P. Nentwig (Eds.), *Learning outcomes in science education: Making it tangible* (pp. 87–117). Münster: Waxmann.

Pellegrino, J. W., Chudowsky, N., & R. Glaser (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington: National Academies Press.

Pickering, M. (1980). Are lab courses a waste of time? *Chronicle of Higher Education*, 19.

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. https://doi.org/10.1177/0748175609344096

Renninger, K. A. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Educational psychology. Intrinsic and extrinsic motivation: The search for optimal motivation and performance*. (pp. 373–404). Elsevier. https://doi.org/10.1016/B978-012619070-0/50035-0

Renninger, K. A., & Bachrach, J. E. (2015). Studying triggers for interest and engagement using observational methods. *Educational Psychologist*, *50*(1), 58–69. https://doi.org/10.1080/00461520.2014.999920

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, *46*(3), 168–184. https://doi.org/10.1080/00461520.2011.587723

Risch, B. (Ed.). (2010). *Teaching chemistry around the world*. Münster: Waxmann.

Rönnebeck, S., Bernholt, S., & Ropohl, M. (2016). Searching for a common ground – A literature review of empirical research on scientific inquiry activities. *Studies in Science Education*, *52*(2), 161–197. https://doi.org/10.1080/03057267.2016.1206351

Rotgans, J. I., & Schmidt, H. G. (2017). The relation between individual interest and knowledge acquisition. *British Educational Research Journal*, *43*(2), 350–371. https://doi.org/10.1002/berj.3268

Schiefele, U. (2001). The role of interest in motivation and learning. In J. M. Collis & S. Messick (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement* (pp. 163–194). Mahwah, NJ: L. Erlbaum.

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research.

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. https://doi.org/10.1016/j.cedpsych.2015.06.005

Schiepe-Tiska, A., Schmidtner, S., Müller, K., Heine, J.-H., Neumann, K., & Lüdtke, O. (2016). Naturwissenschaftlicher Unterricht in Deutschland in PISA 2015 im internationalen Vergleich. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme, & O. Köller (Eds.), *PISA 2015* (pp. 133–176). Deutschland: Waxmann.

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. https://doi.org/10.1002/tea.21409

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. https://doi.org/10.1002/tea.21320

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. https://doi.org/10.1016/j.jvb.2009.10.015

Stefani, C., & Tsaparlis, 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. https://doi.org/10.1002/tea.20279

Steinmetz, H. (2011). Estimation and comparison of latent means across cultures. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *European association of methodology series. Cross-cultural analysis: Methods and applications* (pp. 85–116). New York, NY: Routledge.

Stiller, J. (2016). Scientific Inquiry im Chemieunterricht. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät.

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. https://doi.org/10.1002/tea.21010

Talanquer, V. (2008). Students' predictions about the sensory properties of chemical compounds: Additive versus emergent frameworks. *Science Education*, *92*(1), 96–114. https://doi.org/10.1002/sce.20235

Taskin, V., & Bernholt, S. (2014). Students' understanding of chemical formulae: A review of empirical research. *International Journal of Science Education*, *36*(1), 157–185. https://doi.org/10.1080/09500693.2012.744492

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. https://doi.org/10.1037/pspp0000034

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. https://doi.org/10.3102/0002831215618662

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. https://doi.org/10.1037/0022-0663.100.2.460

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. https://doi.org/10.1021/acs.jchemed.6b00093

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. https://doi.org/10.1177/109442810031002

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. https://doi.org/10.1080/00313831.2014.904419

Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus. Wiley series in probability and statistics*. Chichester, West Sussex: Wiley.

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Seattle: Hogrefe & Huber.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

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. https://doi.org/10.1016/j.dr.2009.12.001

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.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Höft L, Bernholt S, Blankenburg J, Winberg M. 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. *J Res Sci Teach.* 2019;56:184–210. https://doi.org/10.1002/tea.21475