



# How Do Students Understand Energy in Biology, Chemistry, and Physics? Development and validation of an assessment instrument

Sebastian Tobias Opitz

Leibniz Institute for Science and Mathematics Education (IPN), GERMANY  
CREATE for STEM at Michigan State University, UNITED STATES

Knut Neumann

Leibniz Institute for Science and Mathematics Education (IPN), GERMANY

Sascha Bernholt

Leibniz Institute for Science and Mathematics Education (IPN), GERMANY

Ute Harms

Leibniz Institute for Science and Mathematics Education (IPN), GERMANY

Received 10 August 2016 • Accepted 3 April 2017

## ABSTRACT

Science standards of different countries introduced disciplinary core ideas and crosscutting concepts—such as energy—to help students develop a more interconnected science understanding. As previous research has mostly addressed energy learning in specific disciplinary contexts, this study targets students' cross-disciplinary understanding of energy. Since no respective test instrument was available, we present the development and validation of an instrument that can be used to compare students' progressing energy understanding across contexts from biology, chemistry, and physics. In a cross-sectional study, we administered the new instrument to  $N = 752$  students at the end of grades 6, 8, and 10. In addition to a detailed discussion of the instrument's reliability and validity, the study findings compare progressing energy understanding in the three disciplinary contexts. With regard to energy as a crosscutting concept, the results are then used to discuss how students' energy understanding may be connected across disciplinary boundaries.

**Keywords:** energy, crosscutting concept, disciplinary core idea, biology, chemistry, physics

## INTRODUCTION

Large-scale assessments have identified limited connections between content elements in both students' knowledge and curricula (Bransford, Brown, & Cocking, 2000; Michaelis, Shouse, & Schweingruber, 2008). Experts, on the contrary, possess a well-connected knowledge base that is built around the core ideas of their domain (Bransford et al., 2000). These findings underline the need to particularly support low-achieving students who struggle to connect the wide-range of contents that they learn in science (Linn, Eylon, & Davis,

© **Authors.** Terms and conditions of Creative Commons Attribution 4.0 International (CC BY 4.0) apply.

**Correspondence:** Sebastian Tobias Opitz, *CREATE for STEM Institute, Michigan State University, United States.*

✉ [opitz@msu.edu](mailto:opitz@msu.edu)

### **State of the literature**

- Energy is both a core idea and a crosscutting concept in science education. As such, energy is used to bind together science topics across grade levels and school science subjects.
- Recent studies have described progressing energy understanding along four central energy aspects: energy forms, energy transfer/transformation, energy degradation, and energy conservation. These have been identified as relevant across disciplinary boundaries.
- Most studies on students' energy understanding have so far described energy understanding in physics contexts. There are very few insights into how students understand energy across different disciplinary contexts. To study this field, a respective test instrument is required.

### **Contribution of this paper to the literature**

- We provide one of the first approaches to analysing energy understanding across disciplinary contexts.
- We connect our findings to those from earlier studies by describing progressing understanding along four central energy aspects. These are used as a basis to compare between progressing energy understanding in biology, chemistry, and physics contexts.
- As a resource to future studies, we provide the developed instrument in online appendices, including detailed information about each item.

2004; Michaelis et al., 2008). Accordingly, multiple countries have revised their science standards to include disciplinary core ideas and interdisciplinary crosscutting concepts as a means to organise and connect science contents (e.g. Next Generation Science Standards [NGSS], 2013). The overarching goal is to foster cumulative learning and thereby enable students to develop a more connected science understanding. Energy is the only concept that is considered both a core idea in biology, chemistry, and physics, as well as a concept cutting across these disciplines.

With regard to energy as a core idea, learning progressions (Alonzo & Gotwals, 2012) have been proposed as a basis for designing teaching materials and assessment tools in order to help students develop a deeper understanding of core ideas and crosscutting concepts (Fortus, Sutherland, Reiser, & Krajcik, 2015). Descriptions of how students progress in understanding the energy concept have mostly focussed on individual disciplines (Neumann et al., 2013; Jin & Anderson, 2013; Lacy, Tobin, Wiser, & Crissman, 2014). To our knowledge, there is yet no clear theory about how students would learn to connect a central concept like energy across disciplinary boundaries. Likewise, little insight is available into the systematic assessment and comparison of middle school students' progressing energy understanding across different disciplinary contexts (cf. Park & Liu, 2016). Due to the discipline-specific scope of most prior studies, a test instrument for a respective study was not yet available when our study was conducted (Neumann et al., 2013; Jin & Anderson, 2012; Liu and McKeough, 2005).

To address this gap in the current research base, our article first presents the development of a 43-item test instrument on energy understanding across biology, chemistry, and physics contexts. Secondly, findings are presented from administering this instrument to

a cross-sectional study with  $N = 752$  students from grades 6, 8, and 10, thereby providing first insights to students' progressing energy understanding across disciplinary contexts.

## THEORETICAL BACKGROUND

### **Energy as a Core Idea and a Crosscutting Concept in Science Education**

The energy concept is a fundamental idea in science (e.g. Bevilacqua, 2014; Bodzin, 2012; Chen et al., 2014; Duit, 1984) and has accordingly been identified as a central concept for science education.

From a content perspective, learning about energy can facilitate the analysis of science contexts. Experts were found to be able to quickly recognise and address the key ideas relevant for solving a given problem (Bransford et al., 2000). As one such key idea, the energy concept is helpful for addressing problems in various science contexts. Students can use the energy concept as a tool with which they analyse energy changes within the boundaries of a system. For example, analysing the conversion between different energy forms can enable students to make predictions about the extent to which processes can occur in a defined system (Lerner & Trigg, 2005).

As a core idea, the energy concept can support learning by organising knowledge. Students learn about energy in a wide variety of contexts which require them to transfer energy-related knowledge along a spatial, temporal, and domain-specific continuum (Pugh & Bergin, 2006). Effective cognitive structures for knowledge transfer have been described as interconnected. Experts' knowledge structures are organised around core ideas, which were found to support knowledge transfer (Bransford et al., 2000; Pugh & Bergin, 2006). If science instruction provides adequate opportunities for knowledge transfer, for example, via instruction along core ideas, students are considered to organise newly learnt contents more cumulatively and thereby build an integrated science understanding. By 'integration of understanding', we refer to the construct of knowledge integration, i.e. the extent to which knowledge about encountered phenomena, situations, and abstractions is added, differentiated, and evaluated in relation to prior conceptions (Linn et al., 2004, p. 30). As students have been shown to rather spontaneously add knowledge instead of differentiating it, fostering students' levels of knowledge integration is a goal of ongoing research (Linn et al., 2004). As such, an integrated science understanding is considered beneficial for the critical analysis, evaluation, and decision making in future challenges (Sakschewski, Eggert, Schneider, & Bögeholz, 2014).

To support knowledge integration in science education, core ideas and crosscutting concepts have been introduced into the science standards and curricula of several countries (e.g. NGSS, 2013). At the level of individual disciplinary contexts, core ideas aim to organise and revise contents, while crosscutting concepts are intended to enable students to organise contents more coherently across disciplinary boundaries (NGSS, 2013). Energy incorporates a unique position, as it is both a core idea in different science disciplines, as well as a crosscutting

concept spanning across them. In order to be able to use energy as a crosscutting concept, aspects of understanding energy in different disciplines have to be integrated by students across disciplinary boundaries.

In conclusion, core ideas such as energy can function as a pivot to underline links between contents and highlight similarities in the complexity of science contents. Thereby, core ideas are thought to facilitate knowledge integration of students' science understanding. However, the extent to which this goal is achieved depends on the coherence (Fortus & Krajcik, 2012) with which concepts like energy are taught across different (disciplinary) contexts (Fortus et al., 2015).

### **Research on Students' Energy Understanding**

Four aspects have been identified to play a major role in understanding the energy concept: (1) Energy forms and sources, (2) energy transfer and transformation, (3) energy degradation and dissipation, and (4) energy conservation (e.g., Neumann et al., 2013; Duit, 1984; Lancor, 2015). While these four aspects are part of instruction in all disciplinary fields, the focus on them can vary substantially across different disciplinary backgrounds (Eisenkraft et al., 2014). A biology approach typically analyses energy transfer and transformation in open systems, for example, animals. The extent to which an organism transforms limited available energy into desirable energy forms (e.g. kinetic energy) instead of degrading it into thermal energy is an important factor in natural selection. Therefore, the efficiency of energy transformations and the emission of heat are often focussed on in biological approaches (e.g. in energy flow diagrams). In contrast, a physics perspective may address the question to what extent work can be conducted through an energy transformation. Here, the focus lies on work, while emitted heat is frequently seen as a by-product or even disregarded when the analysis focuses on idealised, closed systems (Stacy, Chang, Coonrod, & Claesgens, 2014). The four energy aspects have been used to analyse energy understanding (e.g. Herrmann-Abell & DeBoer, 2011; Liu & McKeough, 2005) and to structure energy teaching (e.g. Nordine et al., 2010). Further characteristics of energy understanding, for example, the relevance of systems boundaries, have been discussed by Doménech et al. (2007).

Generally, students' energy understanding has been found to be limited (e.g. Chen et al., 2014; Watts, 1983). A wider body of studies focused on (persistent) alternative conceptions and categorised these into frameworks (Boyes & Stanisstreet, 1990; Burger, 2001; Nordine et al., 2010; Trumper, 1993; Watts, 1983). Learning difficulties with respect to the energy concept are imposed, for example, by the concept's abstract nature, its wide and varied applications, connotations in everyday language (e.g. Liu & Ruiz, 2008; Trumper, 1993), as well as confusions with related concepts such as force, work, power (Kurnaz & Sağlam-Arslan, 2011) or matter (Lin & Hu, 2003).

The following paragraphs provide an overview of research on students' energy understanding. These studies span from early, less discipline-specific energy learning via

discipline-specific aspects of energy understanding, towards differentiation and integration processes in students' progressing energy understanding.

**Early energy learning.** Few studies have addressed the energy understanding of young learners (e.g. Lacy et al., 2014). In an energy learning sequence initially suggested by Driver, Squired, Rushworth, and Wood-Robinson (1994), students are expected to enter formal education with alternative conceptions from everyday experiences, such as a feeling of being 'energetic'. Next, the students progress by relating energy first to living and then to non-living entities before apprehending stored energy and, finally, energy degradation and conservation. Research suggests that early energy learning is possible (Shultz & Coddington, 1981), and can be sustainably effective until high school (Novak, 2005). If alternative conceptions are explicitly addressed (Trumper, 1993), alternative ideas such as the application of energy forms or energy as a quasi-material substance, can be helpful scaffolds for later understanding (Jin & Anderson, 2012; Lancor, 2015).

**Energy learning in different disciplinary contexts.** Since studies on energy understanding in physics have been summarized in several publications (e.g. Kurnaz & Sağlam-Arslan, 2011; Millar, 2005; Nordine et al., 2010; Tatar & Oktay, 2007), this section focuses on biology and chemistry contexts. In these two disciplines, students' alternative conceptions have been found to exhibit similar learning constraints as in physics, for example, the difficulty in identifying reference systems. Further discipline-specific ideas (e.g. concerning endo-/exothermic reactions) have been described by Tastan et al. (2008). Chemical engineering students' energy understanding in bond breaking/bond forming has been shown to rely strongly on textbook definitions. Similar to physics, students only incompletely applied normative conceptions of energy conservation and degradation and only rarely reasoned at concrete and abstract levels simultaneously (Liu, Ebenezer, & Frazer, 2002).

Energy understanding in chemical reactions forms the basis for detailed analyses of energy changes in biology (Quinn, 2014). However, the connection in energy understanding between the two contexts is not clearly understood (Hirça et al., 2008). Energy understanding in biology is frequently linked to concrete manifestations (e.g. energy sources) instead of abstract notions such as work or conservation (Forde, 2003). Learning opportunities are often more contextualised than in physics, thereby promoting confusion between the scientific energy concept and energy connotations in everyday language (Jin & Anderson, 2012; Lin & Hu, 2003). Categories of students' energy understanding in life processes have been abstracted from a larger questionnaire study (Burger, 2001). Specific alternative student ideas on energy sources for living beings have been identified, for example, in the ideas that plants get energy from soil, air or water (Boyes & Stanisstreet, 1990). While the analysis of energy flow through ecosystems is a highly relevant aspect of understanding energy in biological contexts (Eisenkraft et al., 2014), students frequently confuse system boundaries or cycle versus flow conceptualisations (Doménech et al., 2007; Jin & Anderson, 2012; Lin & Hu, 2003). The application of energy flow charts builds on an understanding of energy conservation. In contrast, the threshold-concept character of energy in biology is mostly limited to transfer and

transformation (Ross et al., 2010). A focus on the latter energy aspect also becomes apparent in learning materials and curricula, while energy degradation/dissipation and energy conservation are less understood by students and only sparsely represented in biology instruction (e.g. Forde, 2003; Opitz et al., 2015). Chabalengula et al. (2011) found that even university students exhibit difficulties in transferring energy contents from physics to biological contexts, thereby raising the question how energy understanding is related between disciplinary contexts (Lancor, 2015).

**Promoting integrated energy understanding.** Learning progressions are models that describe how core ideas are understood in successively more complex ways (Krajcik, Sutherland, Drago, & Merritt, 2012; Michaelis et al., 2008). If curricula are organised along learning progressions, students are thought to connect knowledge more meaningfully around the respective core idea and thereby achieve a more integrated science understanding (Alonzo & Gotwals, 2012; Linn et al., 2004). Learning progressions are derived and validated through empirical research findings. Principles for assessing these data have been described by Krajcik et al. (2012) and, more specifically, with regard to an energy learning progression by Neumann et al. (2013).

In the past years, several studies have presented energy learning progressions or provided data to inform their development. Predictors for energy test scores have been found—in this order of impact—in the cognitive levels of tasks, the context (everyday vs. scientific), students' age, and in the specific content (Liu & Ruiz, 2008). These results underline the impact that assessment designs can have on the interpretation of study findings and the development of learning progressions (Alonzo & Gotwals, 2012). Most approaches to an energy learning progression have so far focussed on physics contexts. In these studies, students were found to reach an understanding of energy forms at elementary level, while understanding of energy transfer was marginally achieved in lower secondary school. Energy degradation and conservation were not even fully understood in high school (Liu & McKeough, 2005). Similar results have reported, for example, by Neumann et al. (2013) or Herrmann-Abell and DeBoer (2011). Dawson-Tunik (2008) found that students' progressing energy understanding is characterised by both consolidation and transitory stages during which students achieve higher levels of abstraction. Other learning progressions have been proposed for specific contexts, for example, in relation to energy understanding in carbon-transforming processes (Jin & Anderson, 2012). These approaches can provide valuable insights for an interdisciplinary energy learning progression.

Progressing energy understanding should also be characterised by the ability to apply the energy concept across an increasing number of different contexts and disciplines (Jin & Anderson, 2012; Novak, 2005). Accordingly, recent studies have focused on integration processes (Linn et al., 2004) that occur as students learn about energy. Nordine et al. (2010) designed a teaching sequence that successfully fostered a more integrated energy understanding. While levels of energy knowledge integration have been shown to increase across middle school, they are generally low, especially among life science students (Lee &

Liu, 2009). At an interdisciplinary level, purpose-built curricula with high levels of coherence between individual units have been able to promote integrated energy understanding across disciplines and school years (Fortus et al., 2015).

These findings indicate that the aforementioned learning goal of a teaching towards a crosscutting energy concept is practically feasible. In a qualitative analysis, Lancor (2015) found that most university students employed a substance-based analogy for energy. Furthermore, students in discipline-specific courses referenced the energy aspects more prominently than students in a general science course. Hence, knowledge integration in students' energy understanding likely varies with the structure of the educational system. With respect to energy understanding in different disciplines, Park and Liu (2016) showed that college students' discipline-specific subsets of energy understanding are highly related, especially between physics and chemistry (biology–chemistry:  $r = .79$ ; biology–physics:  $r = .77$ ; chemistry–physics:  $r = .91$ ). These findings indicate that energy understanding across disciplines is either built on one energy conceptualisation that is consistently applied to varying contexts, or, that students' energy understanding is generally vague, possibly based on alternative conceptions from everyday language, and therefore highly related between disciplines.

### Research Objectives and Questions

A summary of previous research on students' energy understanding shows that (i) the majority of studies have focused on physics contexts (e.g. Liu & McKeough, 2005), whereas energy learning in other disciplinary contexts has received little attention (cf. Jin & Anderson, 2012). (ii) In addition, it has been concluded that learners are likely challenged by connecting energy knowledge across topics from different disciplines, thus potentially making it difficult for learners to use energy in the sense of a crosscutting concept (Eisenkraft et al., 2014). However, few empirical insights describe how students' progressing energy understanding in different disciplinary contexts is interconnected (Hirça et al., 2008; Lancor, 2015; Park and Liu, 2016). Specifically, research has widely overlooked cross-disciplinary energy learning in middle school and high school. A more robust empirical basis in this field is important for the development of energy learning progressions and the strengthening of coherence in energy instruction across topics and school science subjects (Fortus et al., 2015). (iii) When this study was conducted, no test instrument for the systematic comparison of middle/high school students' progressing energy understanding across different disciplinary contexts was available (for a recent and valuable addition to this field, see Park and Liu, 2016). Earlier instruments focus on single disciplinary contexts (e.g. Neumann et al., 2013; Jin & Anderson, 2012) or employ items from varied disciplinary backgrounds (e.g. Lee & Liu, 2010; Liu & McKeough, 2005), thus hampering a systematic comparison of students' energy understanding across disciplines.

To address these issues through our study, we pursued two objectives: Firstly, we aimed to develop and validate an instrument to assess progressing energy understanding

across middle and high school biology, chemistry, and physics contexts (objective 1). Secondly, this instrument was deployed in a cross-sectional study to obtain insights into cross-disciplinary energy understanding (objective 2). This study was guided by two research questions (RQs):

- (RQ 1) What progression trends of energy understanding can be identified in biology, chemistry, and physics contexts?
- (RQ 2) How is students' energy understanding related across these contexts?

## METHODS

### Procedure

With respect to the first objective, an instrument for energy understanding in biology, chemistry, and physics contexts was developed around a shared framework for energy understanding (Neumann et al., 2013). This approach was taken to ensure a parallel test design with similar items in the three disciplinary contexts, hence allowing comparisons across these. For the second objective, a cross-sectional study was employed in which students' energy understanding was assessed via the new instrument at the end of grades 6, 8, and 10. The study also assessed students' crystallised intelligence ('CI' – LPS subtests 1 and 2, Horn, 1983) as a control variable. The assessment of crystallised intelligence served two purposes: The variable (i) provided an additional measure to assess the comparability between students at different grade levels, and (ii) it was used – in extension to previously analysed covariates (Opitz et al., 2015) – as an important external variable with which to address discriminant validity. To facilitate readability, the rationale for specifically including this variable is provided in the validity discussion.

### Instrument Development (Objective 1)

**Item construction.** For this study, physics and biology items were employed from previous studies (Neumann et al., 2013; Opitz et al., 2015). These items, as well as newly developed chemistry items, were constructed around the same framework and used similar guidelines for item development. The underlying two-dimensional framework of energy understanding was operationalised for the development of all items. In this framework, the first dimension concerns the aforementioned four energy aspects, while the second dimension consists of four levels of increasing cognitive complexity (facts, mappings, relations, and concepts). The dimension of energy aspects was used to generate the respective contents for the items, while the dimension of cognitive complexity was used as a means to vary item difficulties (for details on the theoretical justification of the framework see Neumann et al., 2013). The operationalisation of items showed that the number of factors that needed consideration varied substantially across contexts and disciplines. To take this variance into account, three types of multiple choice items were employed (1-attractor, 2-attractor, and complex MC-items). The differences in item type had only a small effect on student performance (see [Appendix 2](#) for details).



All items were controlled for adequate and comprehensive language, i.e., by using feedback from teachers for chemistry items and qualitative student feedback for the physics and biology items (American Association for the Advancement of Science, 2007).

**Item selection.** The instrument was designed for students from lower middle to high school. In this regard, a total of 128 items (biology: 48; physics: 48; chemistry: 32) were field tested with  $N = 336$  grade 6, 8, 10, and 12 students. In field testing, each item was answered by an average of 78 students. A total of 48 items were selected for the instrument (16 items per disciplinary context, four items per energy aspect in each discipline). As the instrument was designed as a developmental test, grade 10 and 12 students served as a reference group with respect to the instruments' distribution of item difficulties. Item selection was conducted iteratively along the following sequence of criteria and included new parameter computation after each step: (1) Exclusion of items with negative correlation between solution probability and grade level ( $r < -.1$ )<sup>1</sup>, (2) exclusion of items with a discrimination  $< .15$ , (3) selection of discipline-specific item sets with regard to best fitting normal distribution of item difficulties in grade 10/12, (4) selection of items with the largest possible item discrimination, and (5) selection of items in accordance to the distribution of levels of cognitive complexity. All items, item parameters and attractor/distractor selection frequencies for each grade level are provided in [Appendix 1](#).

**Item discipline-specificity.** In addition to the items' construction by researchers from the respective disciplines, an expert rating was conducted to ensure the items' discipline-specificity.  $N = 12$  high school teachers and scientists from biology, chemistry and physics rated—without further instruction—all 48 items with respect to the items' disciplinary background. The findings from the rating relate to the first research objective and will be presented in the results section.

### Data Collection and Sample for the Study (Objective 2)

Importantly, all students in grades 6, 8 and 10 received the same items from all three disciplinary contexts. We chose to give all items to all students to ensure that we could get a reliable measure about how scores on the three disciplinary subtests were interrelated. Hence, we explicitly chose not to use a test booklet design to avoid using only few linking items as a basis to assess this relationship.

Data sampling was conducted during school lessons in the last four weeks of the school year. All students were able to finish the test within the 90-minute testing sessions. 47 classes from 10 metropolitan secondary schools in northern Germany (Hamburg) participated in this study, including all socio-economic strata (Hamburg Social Index, range 1–6,  $m = 4.00$ ) and both prevalent school-types—'Gymnasium' (high-performance students) and 'Stadtteilschule' (mixed-performance students). Participating students received science classes with content from mixed disciplinary backgrounds until grade 6 and separate biology, chemistry, and physics classes from grade 7 onwards. The relevant curricula and science standards for Hamburg (e.g. KMK 2005a, 2005b, 2005c) include energy as a core idea and a concept to bridge

the science subjects. The curricula list numerous energy learning opportunities from grade 2 to 12/13 (see [Appendix 3](#)).

**Table 1.** Characteristics of subsamples (grade levels). Different lower case letters denote significant differences,  $p < .05$ , n.s.= not significant

Grade level	Average crystallised intelligence, T-standardised	Percentage of male students	Percentage of students from high-performance schools	Average socio-economic index
6	46.07 – n.s.	51.8 – n.s.	56.9 – a	3.97 – a
8	45.03 – n.s.	57.4 – n.s.	84.0 – b	4.43 – b
10	45.97 – n.s.	50.2 – n.s.	64.6 – b	4.49 – c
6–10	45.69	53.13	67.6	4.30

The sample analysed below consisted of  $N = 752$  students from grade 6 ( $m_{\text{age}} = 12.3$ ;  $n = 255$ ), 8 ( $m_{\text{age}} = 14.3$ ;  $n = 268$ ), and 10 ( $m_{\text{age}} = 16.4$ ;  $n = 229$ ). [Table 1](#) shows characteristics of the subgroups (grade levels). The data indicate comparability between the groups, with significant differences only between the percentage of students from high-performance schools and the socio-economic index, the latter indicating only negligible differences. To account for these small differences, the results below were weighted with regard to the school types.

### Statistical Procedures

For the first objective, the discipline-specificity of the items was assessed through the expert rating described above. Fleiss Kappa ( $k$ ) was used as an index to describe the agreement between the raters. The discussion of validity also concerned the question how much variance the assessment could explain and to what extent the findings were generalizable from the sample to the population level. These issues were addressed by using  $r^2$  and shrinkage indices derived from (forced entry) multiple regression models (Field, 2009, p. 221). As this was an auxiliary analysis, its findings are presented in [Appendix 4](#).

For the second objective, mean energy scores for the test and the respective subtests were employed as the measure for comparison. Outliers were assessed via standard SPSS box plots and then excluded from the data presented below. Data weighting with regard to the school-type was conducted according to recommendations by Maletta (2007). Neither the exclusion of outliers nor the data weighting had substantial effects on the results. Comparisons of energy scores across grades were conducted via ANOVA, using conservative Games-Howell post-hoc tests (Field, 2009, p. 374). Effect sizes were calculated via  $\omega^2$  (Field, 2009, p. 391; Kirk, 1996).

In addition to these classical statistics procedures, the objectives of this study have also been addressed by Rasch analysis (Wu, Adams, and Wilson, 2007). The advantage of this method is, for example, that the adequacy of item difficulties can be easily assessed with regard to the ability spectrum of a given sample. However, some of the in-detail analyses for this article were not computable through Rasch estimation. The results from Rasch analysis, however, produced results similar to those from the classical statistics procedures mentioned above. Thus, in order to avoid confusion between the different statistical procedures, this paper

presents findings solely from classical statistic. Interested readers can access the methods and results from Rasch analysis in [Appendix 4](#). These results incorporate a Wright map of student abilities/item difficulties, WLE person separation reliabilities, and an analysis of differential item functioning (DIF) regarding gender.

## RESULTS

### Instrument Development (Objective 1)

**Item discipline-specificity.** The 12 raters agreed on a 'near-perfect' basis ( $k = .81$ , Kirk, 1996) on the discipline-specificity of the items. Similar values were obtained for the items of each disciplinary item set ( $k = .78-.85$ ). The rating was used to sort items into four categories: (1) high discipline-specificity (> 80 % agreement, 34/48 items), (2) substantial specificity (66–80% agreement, 6/48 items), (3) lesser specificity (50–66% agreement, 3/48 items), and, (4) specificity for another than the intended discipline (> 50% agreement on another discipline, 5/48 items). While the three items from category (3) were kept to represent more interdisciplinary contexts, the five items from group (4) were excluded from further data analyses. Therefore, the final instrument consisted of 43 items (biology: 15, chemistry: 13, physics: 15).

**Reliability.** Reliabilities across grades were high for the complete test ( $\alpha = .85$ ) and satisfactory for the three subtests ( $\alpha = .61-.71$ ). Similar values were obtained from the Rasch analysis (compare [Appendix 4](#)): Across grades 6–10, WLE person separation reliability was high (.82) for the full test (grade 6: .57, grade 8: .76, grade 10: .83) and sufficient to satisfactory for the three disciplinary subtests (.58–.62).

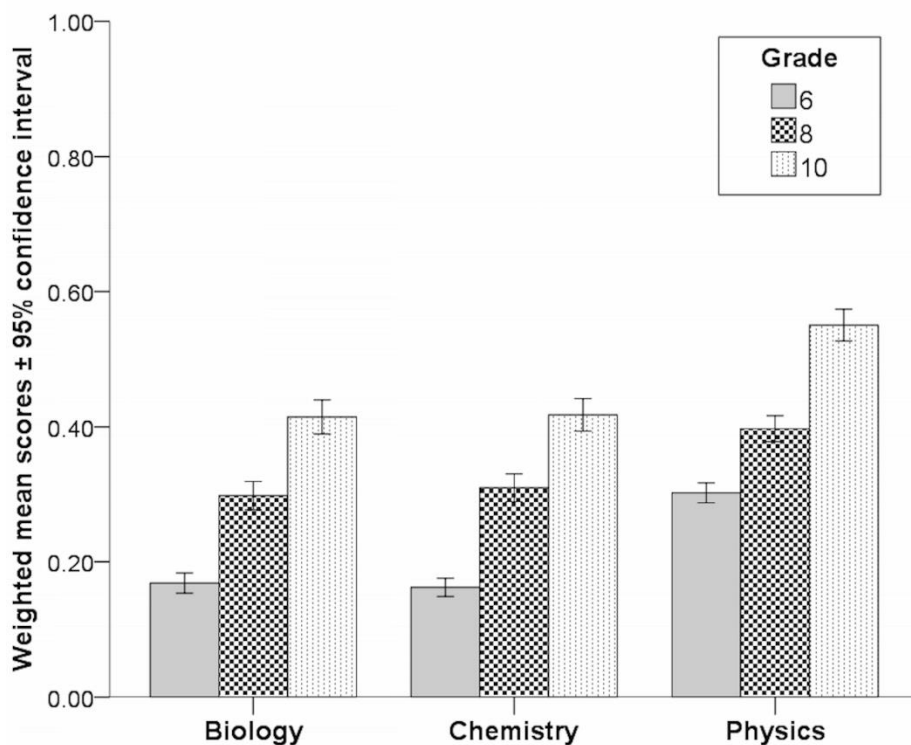
**Item difficulties and discrimination.** Item parameters for each item and grade level are provided in [Appendix 1](#). In the reference group grade 10, the distribution of item difficulties was widespread and could thereby address students of all abilities (range of item difficulties,  $p = .03-.96$ ). Due to the study's original intent to also include grade 12 students, mean item difficulties slightly exceeded student abilities (grade 10:  $m_{\text{difficulty}} = .42$ , biology: .36, chemistry: .35, physics: .53). Item discrimination increased with grade level and showed good values across grades 6 to 10 ( $m_{\text{biology}} = .32$ ,  $m_{\text{chemistry}} = .29$ ,  $m_{\text{physics}} = .34$ ).

**Control variable.** The connection between energy scores and age-standardised crystallised intelligence showed low correlations with energy understanding both across ( $r = .14$ ,  $p < .01$ ), and within grades (grade 6:  $r = .25$ ; grade 8:  $r = .17$ ; grade 10:  $r = .22$ ,  $p < .01$ ).

### Energy Understanding across the Three Disciplines (Objective 2)

Progressing understanding across grades 6–10 (RQ 1). With respect to the first research question, [Figure 1](#) shows progression trends in energy understanding in biology, chemistry, and physics contexts for grades 6 to 10. Student scores increased significantly and to a similar extent in all contexts (biology:  $F = 121.37$ ,  $df = 2$ ,  $p < .001$ ; chemistry:  $F = 132.83$ ,  $df = 2$ ,  $p < .001$ ;

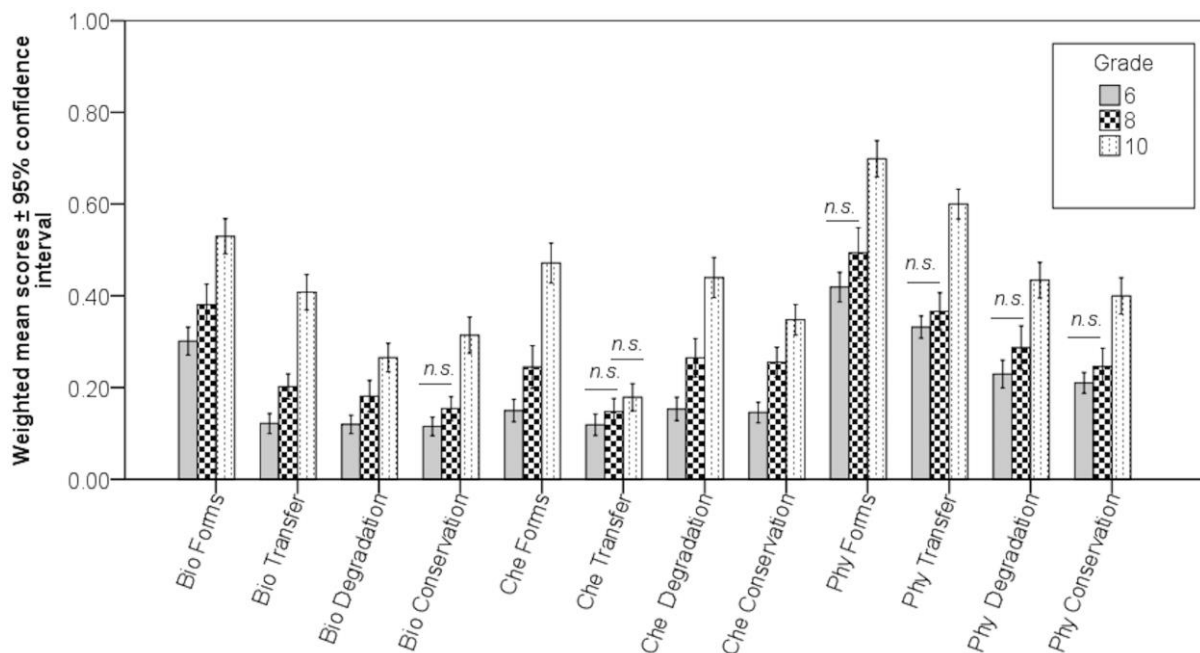
physics:  $F = 156.89$ ,  $df = 2$ ,  $p < .001$ ). These progressions relate to 'large' (Kirk, 1996) effects in biology ( $\omega^2 = .24$ ), chemistry ( $\omega^2 = .26$ ) and physics ( $\omega^2 = .29$ ).



**Figure 1.** Progressing energy understanding in biology, chemistry, and physics contexts from grades 6 to 10. All differences between grade levels are significant,  $p < .001$

**Discipline-specific foci in energy understanding (RQ1).** In the three disciplinary contexts, students showed similar patterns in their progressing understanding along the four energy aspects, especially in biology and physics (Figure 2). Here, only 3–4 items were available per energy aspect in each of the three subtests. Consequently, the results address broad comparisons of mean item difficulties, but not the comparison of individual student performances.

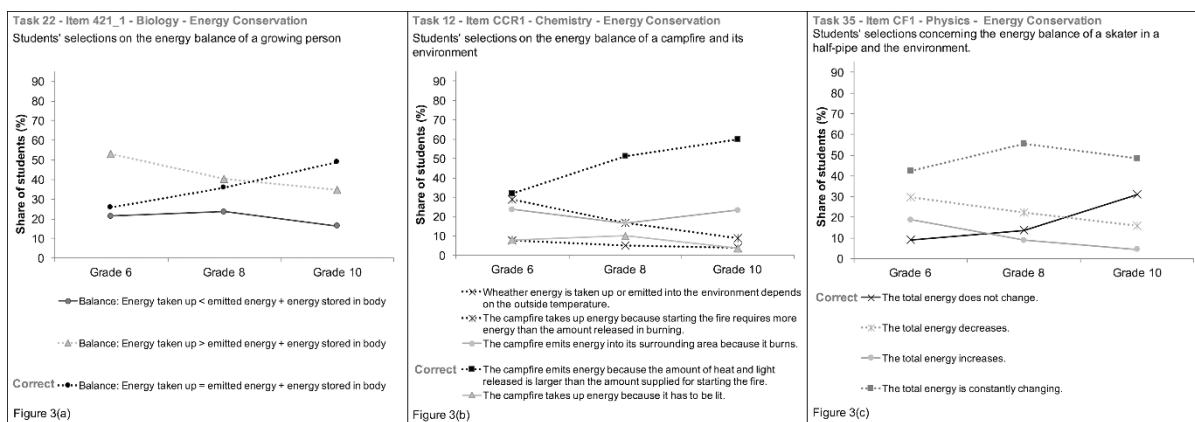
Items on energy forms were the easiest in all disciplines and across the grade levels, while conservation items were generally more difficult. In biology and physics, mean item difficulties increased along the following order of energy aspects: forms, transfer, degradation, and/or conservation. While in biology, degradation and conservation items were similarly difficult, physics items on energy conservation were more difficult and only answered correctly by older students than the items on degradation. In chemistry, the progression pattern differed from biology and physics, as items for energy transfer were comparatively difficult, while the share of students that answered these items correctly did not increase significantly across grades. When related to the distribution of item difficulties across aspects in biology and physics, students scored comparatively higher on chemistry degradation items.



**Figure 2.** Grade 6–10 students' mean scores on items for four energy aspects in biology ('Bio'), chemistry ('Che'), and physics ('Phy') contexts. Mean scores correspond to mean item difficulties ( $p$ ). Differences between columns are significant ( $p < .01$ ), if not marked as *n.s.* ( $p > .05$ )

**Relation of energy understanding between disciplines (RQ 2).** Student scores on the three discipline-specific subtests were strongly correlated with each other (across grades 6 to 10: biology–chemistry,  $r = .63$ ,  $p < .001$ ; biology–physics:  $r = .62$ ,  $p < .001$ ; chemistry–physics,  $r = .60$ ,  $p < .001$ ). Only slightly higher correlations were obtained in a random assignment of the instrument's 43 items to three subscales ( $r = .64$ – $.69$ ,  $p < .001$ ).

**Examples of progressing understanding.** To illustrate the relation of energy understanding across disciplines, **Figures 3(a)–(c)** present selection frequencies of attractors and distractors in items that concern energy conservation in one exemplary biology, chemistry and physics context, each.



**Figure 3.** Progression trends in distractor/attractor selection frequencies for one biology, chemistry, and physics item on energy conservation. As data are cross-sectional, trend lines between data points are only auxiliary

All three items focus on identifying energy balances. In the biological item (Figure 3a), the percentage of students who correctly identified energy uptake to be equal to energy storage and release increased by the factor two from grade 6 to 10 (grade 6: 26%, grade 10: 49%). Here, a relatively large percentage of students retained (53% → 35%) the notion that energy uptake was larger than energy storage and release, thereby relating to the idea of energy ‘loss’. The chemistry item on the energy balance of a campfire (Figure 3b) shows a similarly large increase in students’ ability to identify the correct energy balance (32% in grade 6, 60% in grade 10). The physics item (Figure 3c) showed a comparable developmental pattern, even though generally fewer students (grade 6: 9%, grade 10: 31%) harboured the correct idea that the total energy of the skater and his/her environment are constant.

Tendencies of students to select similar answering options across disciplines can also be found in other items. For example, students adhered similarly to the alternative idea that energy is used up in a process. In a physics context, 45% of the 6<sup>th</sup> to 10<sup>th</sup> graders selected the distractor describing that a pendulum uses up its kinetic energy when swinging back and forth (item DF1, Task 32, see Appendix 1). Similarly invariant across grade levels, roughly 30% of 6<sup>th</sup> to 10<sup>th</sup> graders selected the answering choice that energy from gasoline is used up, if a combustion engine uses fuel to propel a car (chemistry item CDC1, Task 9). Similar examples of progressing energy understanding can be found for all 43 items in Appendix 1.

## DISCUSSION

### Instrument Development (Objective 1)

**Reliability.** Similar reliability values in terms of Cronbach’s  $\alpha$  and WLE person separation reliability (see Appendix 4) indices support the findings presented in this study. The observed reliability is satisfactory for the purpose of identifying broad learning trends across disciplines.

The lower values in grade 6 indicate that the subtests for the three disciplinary contexts should not be employed as stand-alone tests for students below grade 8.

**Validity.** Evidence and arguments in relation to six aspects of validity are discussed according to suggestions and standards by the Joint Committee (2014).

**Test content.** Energy is clearly a core idea in curricula and science standards relevant for the sample tested in this study (KMK 2005a, 2005b, 2005c; see [Appendix 3](#)). Unlike in the NGSS (2014), however, the respective expectations for energy learning were, in most cases, too broadly formulated for a systematic development of test items. Thus, the developed instrument does not assess specific standards with respect to energy learning. Instead, the fit between item contents and expectations expressed in science standards was approximated by ensuring that the four energy aspects (see framework, Neumann et al., 2013) were adequately representing the learning expectations. This criterion was assessed through analyses of school textbooks and curricula, as well as through a teacher survey (Opitz et al., 2015). In other educational systems, the four energy aspects have likewise been described as central elements of energy education in different disciplinary contexts (e.g. Lancor, 2015; Liu & McKeough, 2005; Wang, Wang, & Wei, 2014).

**Cognitive processes.** Think-aloud protocols (Ericsson & Simon, 1993) were conducted to assess if students ( $N = 9$  from grades 6, 8, 10) used the intended cognitive processes for solving items of the new instrument. Detailed results are available in Wernecke (2013). Students' think-aloud protocols were analysed by differentiating content-related strategies of item solving (i.e. arguments based on energy understanding) from non-content-related strategies (i.e. guessing, deciding according to item format). Across items, students obtained the clear majority of solutions by using content-related strategies, while non-content-related strategies comprised just a small share of the observed solutions (grade 6: 13%, grade 10: 3%). This analysis suggests that grade 6–10 students can solve the items of the new instrument validly with regard to the intended cognitive processes.

**Internal structure.** This aspect of validity was addressed through two approaches. First, two expert ratings were used to structure the development of the instrument in relation to (i) the three disciplinary contexts and (ii) the four aspects of energy understanding from the employed framework. The first rating regarding item discipline-specificity (see above) resulted in the exclusion of five items, thereby decreasing the correlations between the three subtests and ensuring their discipline-specificity (before exclusion:  $r \sim .85$ ,  $N_{\text{items}} = 48$ ; after exclusion:  $r \sim .62$ ,  $N_{\text{items}} = 43$ ). A second, previously conducted rating indicated substantial agreement on the items' specificity for the four energy aspects of the framework ( $k = .61$ , Opitz et al., 2015).

The second approach used confirmatory factor analysis (CFA) to study potentially underlying dimensions of students' energy understanding across the three disciplinary contexts. As this analysis is extensive, it is presented in a separate research article (Opitz et al., 2016). The respective findings suggest that students' energy understanding is one-dimensional

across disciplinary contexts. The combination of these results with the ones from the expert rating on item discipline-specificity (see above) indicate that it may be possible to develop an instrument that assesses energy understanding in biology, chemistry, and physics contexts, but that students likely employ a single understanding of energy to answer the items from the different disciplinary contexts. <sup>2</sup>

As additional evidence for the validity of the instrument's internal structure, **Appendix 4** includes results from an analysis of differential item functioning (see Joint Committee, 2014, p. 16).

**Relation to other variables.** Different external variables (e.g. reading skills, fluid intelligence, motivation/interest) have been assessed in previous research (Opitz et al., 2015), but only had minor effects on energy test performance. To extend these findings, crystallised intelligence (CI) was included as a potential covariate of progressing energy understanding (see methods section). CI is a measure of a person's ability to apply general knowledge (Horn, 1983). In contrast to measures considered as more developmentally stable (e.g. working memory capacity/efficiency), levels of CI increase as students mature and learn more. This sub-aspect of intelligence has been found to be strongly connected to different measures of academic achievement (Marzano, 2003, p. 134). As the energy test scores assessed by our instrument concern the understanding of a wide range of everyday contexts, we assumed test performance to depend on a general knowledge measure like CI. However, students' results on the CI-test were just weakly correlated with energy scores ( $r \sim .20$ ). This finding – coupled with the previous ones described above – indicates that the employed items assess a different construct than a general performance measure.

**Relation with criteria.** This aspect of validity is concerned with the accuracy of the test scores' prediction of a criterion – in this case 'energy understanding'. Firstly, to ensure that differences in energy understanding between groups (grade levels) were not affected by construct-irrelevant sources of variance (Joint Committee, 2014, p. 18), we assessed the comparability of students from different grades along several variables (see **Table 1**).

Secondly, the relation between the scores assessed by the new instrument and the intended criterion (energy understanding) can be approximated by correlations between the derived test scores and those from similar instruments (Joint Committee, 2014, p. 17). As the items for physics contexts in the new instrument stemmed from the one presented in Neumann et al. (2013), the strong correlation between the test scores on the physics subtest and the biology/chemistry subtests indicates that these two instruments assess a highly similar construct. In extension, Park (2013, p. 98) found that energy scores on her instrument were to a large extent predictable ( $r^2 = .86$ ) through scores from the test presented by Neumann et al. (2013). We therefore conclude that the instrument presented here is substantially related to the intended criterion.

Lastly, multiple regression was used (see **Appendix 4**) to determine how well findings from the sample employed in this study can be generalised to the population level (Field, 2009,



p. 221; Joint Committee, 2014, p. 19). The regression model included grade level and crystallised intelligence as predictors of students' performance on the developed energy test. The model results showed negligible shrinkage of .1% ( $r^2 = .43$ ), thus indicating good generalisability of the findings.

**Consequences of testing.** The developed instrument is intended for research and not for assessments with diagnostic or student-placement impact. In this study, teachers and their classes participated voluntarily and they did not have to prepare for participation. All testing was conducted by the research team. All results were evaluated anonymously and individual student-, class- or school results were not released. Thus, besides potential effects on students' interest in the energy concept, we have seen no evidence that the application of the instrument had any negative consequences for participating students, teachers, or schools.

**Limitations of the instrument.** The instrument presented in this study was designed to assess discipline-specific sets of energy understanding and their relation to each other. Even though items with more interdisciplinary scope were retained in the instrument, specifically interdisciplinary dimensions of energy understanding are not sufficiently assessable through the presented items. A further limitation lies in the few items per energy aspect within each discipline. Consequently, the subtests should be extended for more fine-grained analyses with regard to these aspects, for example, when knowledge integration patterns within disciplines are of interest.

In accordance with the field testing reported in 'Instrument Development', additional data had originally been collected from a 12<sup>th</sup> grade sample, as well. This subsample had to be excluded in the analysis due to an unresolvable school-type effect that would have biased the comparability between grade levels. The developed instrument was designed to assess energy understanding up to high school and can be applied accordingly in future studies.

**Conclusion.** The developed instrument fitted students' abilities well. Reliabilities were satisfactory and relevant validity aspects were addressed. With regard to the scope of the newly developed instrument, we thus assume that our assessment tool is able to provide sufficiently reliable and valid measures of middle school students' energy understanding in biology, chemistry, and physics contexts.

### Energy Understanding across the Three Disciplines (Objective 2)

**Progression in energy understanding.** For research question 1, the results indicate similarities in progression trends in biology, chemistry, and physics contexts (see **Figures 1** and **2**), as well as in the selection of similar ideas across different disciplines (see **Figure 3**). The reason for the comparatively high performance in physics may be attributed to the fact that these contexts are typical examples of students' first formal energy instruction. Additionally, physics contexts are comparatively less complex than the ones in biology or chemistry (e.g. pendulums, Duit, 1984). While previous findings identified energy knowledge transfer between disciplines as a major learning hurdle (Chabalengula et al., 2011), learning transfer from the less complex physics contexts may also provide a starting point from which to access

more complex contexts in other disciplines. However, transfers in the opposite direction must also be considered for coherent energy instruction, as early, and often implicit learning opportunities in biological contexts can form a basis for later, more explicit energy instruction (Opitz et al., 2015).

In biology and physics, the sequential, yet overlapping progression of energy understanding along energy forms, transfer, degradation, and conservation was similar to findings in other studies (e.g. Neumann et al., 2013; DeBoer & Herrmann-Abell, 2011; Liu & McKeough, 2005). An understanding of energy forms was well established in all disciplines and can therefore function as a scaffold to more advanced energy aspects (Jin & Anderson, 2012). In contrast, energy conservation was little understood in all disciplinary contexts, suggesting that students at this stage are not yet able to use an integrated energy concept in the sense of a conceptual tool (Lacy et al., 2014). In chemistry, students' understanding of energy transfer was low and didn't progress, while understanding of degradation was surprisingly well established. However, these strengths and weaknesses in students' energy understanding in chemistry can only partly be explained by energy contents proposed in the respective learning standards (KMK, 2005b; see [Appendix 3](#)).

**Relatedness in energy understanding across disciplines.** With respect to research question 2, the findings showed clear similarities and relations between progressing energy understanding in different disciplinary contexts. These similarities are stronger than one might expect, as (i) discipline-specific foci regarding the energy aspects are generally accepted (Eisenkraft et al., 2014), (ii) empirical results of learners' performances have been found to reflect these foci (Park and Liu, 2016), and (iii) some energy aspects are rarely addressed in the instruction of individual disciplines, for example, conservation in biology (Opitz et al., 2015).

A critical argument would be that the similarities in progressing energy understanding across disciplines and the strong correlations ( $r \sim .62$ ) between the three subtest scores may be caused by dependencies in the instrument. However, the subtests were clearly identified as discipline-specific. Furthermore, correlations computed before the exclusion of items in the rating ( $r > .80$ ) are comparable to those found by Park and Liu (2016), who used a similar instrument without a rating on item discipline-specificity. Additionally, the correlations between the three scales of randomly assigned items were found to be only slightly higher ( $r \sim .67$ ) than those between the three discipline-specific subtests. These results suggest that students' energy understanding is less discipline-dependent as indicated by the aforementioned arguments.

A possible reason for the similarities in energy understanding across disciplines lies in the conceptual dependence between energy aspects, which may strengthen an inherently similar progression pattern across disciplinary contexts, for example, the dependence on energy forms when describing energy transfer. The strong relation between disciplines may be furthermore promoted by general answering characteristics, for example, by students' widespread application of problematic ideas like 'energy loss' or their inconsistent distinction

between abstract conceptualisations and concrete observations (Liu et al., 2002; compare also physics item 'Task 35' or biology item 'Task 22' in [Appendix 1](#)).

**Limitations of the study.** Further discipline-specific foci on energy understanding (e.g. energy flow in biology or macro/micro-level analyses in chemistry) were not addressed in detail by the new instrument. Similarly, students' distinction between energy in states and energy in processes was not a primary focus of this study (Papadouris & Constantinou, 2014). While these important facets of understanding can be derived from specific answering options in some items (see [Appendix 1](#)), future studies should address these characteristics of energy understanding in more detail.

This study presented findings from students who were taught in separate science classes throughout middle school. While an educational system with separate science classes may not be common in all countries, this organization is widely employed, for example, throughout the European Union (European Commission, 2011, p. 60). As the items of this instrument were not developed in regard to specific educational standards, the instrument can be applied in other educational systems, as well. For example, studies may analyse learning processes that take place as students progress from integrated science instruction in middle school to separated science classes in high school.

### **Implications and Perspective**

To our knowledge, this study proposes one of the first approaches for the analysis of students' energy understanding across disciplinary contexts. The attached items and additional information offer interested science education researchers the opportunity to further elaborate discipline-specific as well as interdisciplinary aspects of students' progressing energy understanding.

Coherent curricula that are structured around core ideas and crosscutting concepts like energy can promote an integrated science understanding (Fortus et al., 2015). Students depend on an integrated energy understanding if they are to use energy as a concept across topics and disciplinary contexts. With this in mind, the importance of a stronger focus on energy conservation and degradation can be derived as a goal for teaching in all science subjects. In contrast, students' widespread understanding of energy forms (Jin & Anderson, 2012), as well as their comparatively high performance in physics contexts could be further developed as the stepping stones to more complex energy contexts in biology and chemistry.

It is not yet well understood in what ways students draw on different disciplinary contexts when learning about energy and what hurdles they encounter in transferring knowledge across these contexts (Chabalengula et al., 2011; Hirça et al., 2008). The findings in this study suggest strong correlations in energy understanding across disciplines. However, in-detail analyses still have to determine how the relation between energy understanding in different disciplinary contexts changes as students progress in school and learn more about energy (Fortus et al., 2015; Opitz et al., 2016). Future experimental studies have to identify

effective educational settings that can enable students to use energy both as a disciplinary core idea and as a crosscutting concept (NGSS, 2013).

#### ACKNOWLEDGEMENT

The city of Hamburg kindly funded this research project and the 'Hamburger Schulversuch alles>>könnern' was a great help for this study. We are grateful for the support of all participating teachers and to the students who took part in the study. Also, we thank Michael Leucht, Jan Retelsdorf, and Olaf Köller for statistical and methodological advice, as well as Susanne Weßnigk for her help with the item development. For introducing new ideas for data analysis and item development, we acknowledge the valuable input from Ilka Parchmann.

#### NOTES

<sup>1</sup> Negative correlations between item solution probability and grade level were considered likely to represent contents specific to individual grade levels only. However, the interest of this study lay – with respect to cumulative learning – on contents addressed across grade levels.

<sup>2</sup> The results from the mentioned CFA analysis (Opitz et al., 2016) are also reflected in the supplemental Rasch Analysis (see appendix 4), where we present the developed items as a single dimension.

#### REFERENCES

- American Association for the Advancement in Science (2007). Getting Assessment Right. *2061 Today*, 17(1), 1-7.
- Alonzo, A. C., & Gotwals, A. W. (2012). *Learning Progressions in Science - Current Challenges and Future Directions*. Rotterdam: Sense Publishers.
- Bevilacqua, F. (2014). Energy: Learning from the Past. *Science & Education*, 23(6), 1231-1243. doi:10.1007/s11191-014-9690-1
- Bodzin, A. (2012). Investigating Urban Eighth-Grade Students' Knowledge of Energy Resources. *International Journal of Science Education*, 34(8), 1255-1275.
- Boyes, E., & Stanisstreet, M. (1990). Pupils' Ideas Concerning Energy Sources. *International Journal of Science Education*, 12(5), 513-529. doi:10.1080/0950069900120505
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience and School*. Washington, D.C.: National Academy Press.
- Burger, J. (2001). *Schülervorstellungen zu "Energie im biologischen Kontext" - Ermittlungen, Analysen und Schlussfolgerungen*. [Student conceptions concerning energy in biological contexts: Research, analysis and conclusions]. (Unpublished doctoral dissertation). University of Bielefeld, Bielefeld/Germany.
- Chabalengula, V., Sanders, M., & Mumba, F. (2011). Diagnosing Students' Understanding of Energy and Its Related Concepts in Biological Contexts. *International Journal of Science and Mathematics Education*, 10(2), 241-266. doi: 10.1007/s10763-011-9291-2
- Chen, B., Eisenkraft, A., Fortus, D., Krajcik, J. S., Neumann, K., Nordine, J., & Scheff, A. (2014). *Teaching and Learning of Energy in K-12 Education*. New York: Springer.

- Dawson-Tunik, T. L. (2006). Stage-Like Patterns in the Development of Conceptions of Energy. In X. F. Liu & W. Boone (Eds.), *Applications of Rasch Measurement in Science Education* (pp. 111-136). Maple Grove, MN: JAM Press.
- Doménech, J., Gil-Pérez, D., Gras-Martí, A., Guisasola, J., Martínez-Torregrosa, J., Salinas, J., . . . Vilches, A. (2007). Teaching of Energy Issues: A Debate Proposal for a Global Reorientation. *Science & Education*, 16(1), 43-64. doi:10.1007/s11191-005-5036-3
- Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making Sense of Secondary Science: Supporting Material for Secondary Teachers*. London: Routledge.
- Duit, R. (1984). Learning the Energy Concept in School - Empirical Results from the Philippines and West Germany. *Physics Education*, 19(2), 59-66. doi:10.1088/0031-9120/19/2/306
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. Cambridge: MIT Press.
- Field, A. (2009). *Discovering Statistics Using SPSS* (Third Ed. ed.). London: Sage Publications.
- Forde, T. (2003). "When I Am Watching Television I Am Not Using Any Energy" - An Empirical Study of Junior Science Students' Intuitive Concepts of Energy. *Irish Educational Studies*, 22(3), 71-89.
- Fortus, D., & Krajcik, J. (2012). Curriculum Coherence and Learning Progressions. In B. J. Fraser, K. Tobin & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 783-798). Springer: Netherlands.
- Fortus, D., Sutherland, L., Reiser, B. J., & Krajcik, J. S. (2015). Assessing the Role of Curriculum Coherence in Student Learning About Energy. *Journal of Research in Science Teaching*, 52(10), 1408-1425. doi:10.1002/tea.21261
- Herrmann-Abell, C. F., & DeBoer, G. E. (2011). Investigating Students' Understanding of Energy Transformation, Energy Transfer, and Conservation of Energy Using Standards-Based Assessment Items. *Paper presented at the NARST Annual International Conference*. Orlando, USA.
- Hirça, N., Çalik, M., & Akdeniz, F. (2008). Investigating Grade 8 Students' Conceptions of 'Energy' and Related Concepts. *Journal of Turkish Science Education*, 5(1), 75-87.
- Horn, W. (1983). *Leistungsprüfsystem - LPS (Vol. 2)*. [Intelligence test LPS] Rosdort/Göttingen: Hogrefe.
- Jin, H., & Anderson, C. W. (2012). A learning progression for energy in socio-ecological systems. *Journal of Research in Science Teaching*, 49(9), 1149-1180. doi:10.1002/tea.21051
- Joint Committee (2014). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Education Research Association (AERA).
- Kirk, R. E. (1996). Practical Significance: A Concept Whose Time Has Come. *Educational and Psychological Measurement*, 56(5), 746-759. doi:10.1177/0013164496056005002
- KMK - Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005a). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss - Beschluss vom 16.12.2004*. [German middle school learning standards for biology]. München: Luchterhand.
- KMK - Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005b). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss: Beschluss vom 16.12.2004*. [German middle school learning standards for chemistry]. München: Luchterhand.
- KMK - Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005c). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss Beschluss vom 16.12.2004*. [German middle school learning standards for physics]. München: Luchterhand.
- Krajcik, J. S., Sutherland, L. A., Drago, K., & Merritt, J. (2012). The Promise and Value of Learning Progression Research. In S. Bernholt, K. Neumann & P. Nentwig (Eds.), *Making it tangible: Learning outcomes in science education* (pp. 261-283). Münster: Waxmann.

- Kurnaz, M. A., & Sağlam-Arslan, A. (2011). A Thematic Review of Some Studies Investigating Students' Alternative Conceptions About Energy. *Eurasian Journal of Chemistry and Physics Education*, 3(1), 51-74.
- Lacy, S., Tobin, R., Wisner, M., & Crissman, S. (2014). Looking Through the Energy Lens: A Proposed Learning Progression for Energy in Grades 3-5. In B. Chen, A. Eisenkraft, D. Fortus, J. S. Krajcik, K. Neumann, J. Nordine & A. Scheff (Eds.), *Teaching and Learning of Energy in K-12 Education*. Cham Heidelberg New York Dordrecht London: Springer.
- Lancor, R. A. (2015). An Analysis of Metaphors Used by Students to Describe Energy in an Interdisciplinary General Science Course. *International Journal of Science Education*, 1-27. doi: 10.1080/09500693.2015.1025309
- Lee, H.-S., & Liu, O. L. (2009). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education*, 94(4), 665-688. doi:10.1002/sce.20382
- Lerner, R. G., & Trigg, G. L. (2005). *Encyclopedia of Physics Volume 1 (A-L)*. Weinheim: Wiley-VCH.
- Lin, C.-Y., & Hu, R. (2003). Students' understanding of energy flow and matter cycling in the context of the food chain, photosynthesis, and respiration. *International Journal of Science Education*, 25(12), 1529-1544. doi:10.1080/0950069032000052045
- Linn, M. C., Eylon, B.-S., & Davis, E. A. (2004). The Knowledge Integration Perspective on Learning and Instruction. In M. C. Linn, E. A. Davids & P. Bell (Eds.), *Internet Environments for Science Education* (pp. 29-46). Mahwah, NJ: Lawrence Erlbaum.
- Liu, X., Ebenezer, J., & Fraser, D. M. (2002). Structural characteristics of university engineering students' conceptions of energy. *Journal of Research in Science Teaching*, 39(5), 423-441. doi: 10.1002/tea.10030
- Liu, X., & McKeough, A. (2005). Developmental growth in students' concept of energy: Analysis of selected items from the TIMSS database. *Journal of Research in Science Teaching*, 42(5), 493-517. doi: 10.1002/tea.20060
- Liu, X., & Ruiz, M. E. (2008). Using data mining to predict K-12 students' performance on large-scale assessment items related to energy. *Journal of Research in Science Teaching*, 45(5), 554-573. doi:10.1002/tea.20232
- Maletta, H. (2007). *Weighting*. Retrieved March 25th, 2015, from <http://www.spsstools.net/Tutorials/WEIGHTING.pdf>
- Marzano, R. J. (2003). *What Works in Schools: Translating Research into Action*. Alexandria, VA, USA: Association for Supervision and Curriculum Development.
- 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. doi: 10.1002/j.2333-8504.1994.tb01618.x
- Michaelis, S., Shouse, A. W., & Schweingruber, H. A. (2008). *Ready, Set, Science! Putting Research to Work in K-8 Science Classrooms*. Washington, D.C.: The National Academic Press.
- Millar, R. (2005). *Teaching about Energy: Research Paper 2005/11*. York, University of York, Department of Educational Studies.
- 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.
- Nordine, J., Krajcik, J., & Fortus, D. (2010). Transforming energy instruction in middle school to support integrated understanding and future learning. *Science Education*, 95(4), 670-699. doi:10.1002/sce.20423

- NGSS (Lead States) (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.
- Novak, J. D. (2005). Results and Implications of a 12-Year Longitudinal Study of Science Concept Learning. *Research in Science Education*, 35(1), 23-40. doi:10.1007/s11165-004-3431-4
- Opitz, S., Harms, U., Neumann, K., Kowalzik, K., & Frank, A. (2015). Students' Energy Concepts at the Transition between Primary and Secondary School. *Research in Science Education*, 49(5), 691-715. doi:10.1007/s11165-014-9444-8
- Opitz, S., Neumann, K., Bernholt, S., & Harms, U. (2016). Energy – a Crosscutting Concept? The Structure of Students' Progressing Energy Understanding in Biology, Chemistry and Physics. In S.T. Opitz, Students' Progressing Understanding of the Energy Concept: An Analysis of Learning in Biological and Cross-Disciplinary Contexts (pp.66-92). Doctoral thesis, University of Kiel/Germany. Available at: [http://macau.uni-kiel.de/receive/dissertation\\_diss\\_00019005](http://macau.uni-kiel.de/receive/dissertation_diss_00019005)
- Papadouris, N. K., T & Constantinou, C.P. (2014). An Exploratory Investigation of 12-Year-Old Students' Ability to Appreciate Certain Aspects of the Nature of Science through a Specially Designed Approach in the Context of Energy. *International Journal of Science Education*, 36(5), 755-782. doi:10.1080/09500693.2013.827816
- Park., M. (2013). *Developing an Instrument for Assessing Students' Understanding of the Energy Concept Across Science Disciplines*. (Unpublished doctoral dissertation). University of Buffalo, Buffalo, New York/USA. Retrieved from <http://search.proquest.com/pao/docview/1511026690>
- Park, M., & Liu, X. (2016). Assessing Understanding of the Energy Concept in Different Science Disciplines. Advance online publication. *Science Education*. doi:10.1002/sce.21211
- Pugh, K. J., & Bergin, D. A. (2006). Motivational Influences on Transfer. *Educational Psychologist*, 41(3), 147-160. doi:10.1207/s15326985ep4103\_2
- Quinn, H. (2014). A Physicist's Musings on Teaching Energy. In B. Chen, A. Eisenkraft, D. Fortus, J. S. Krajcik, K. Neumann, J. Nordine & A. Scheff (Eds.), *Teaching and Learning of Energy in K-12*. New York: Springer.
- Ross, P. M. T., Charlotte E. ; Hughes, Chris ; Whitaker, Noel ; Lutze-Mann, Louise ; Kofod, Michelle ; Tzioumis, Vicky. (2010). Threshold Concepts in Learning Biology and Evolution. *Biology International*, 47, 47-52.
- Sakschewski, M., Eggert, S., Schneider, S., & Bögeholz, S. (2014). Students' Socioscientific Reasoning and Decision-making on Energy-related Issues—Development of a measurement instrument. *International Journal of Science Education*, 36(14), 2291-2313. doi:10.1080/09500693.2014.920550
- Shultz, T. R., & Coddington, M. (1981). Development of the concepts of energy conservation and entropy. *Journal of Experimental Child Psychology*, 31(1), 131-153.
- Stacy, A. M., Chang, K., Coonrod, J., & Claesgens, J. (2014). Launching the Space Shuttle by Making Water: The Chemist's View of Energy. In J. S. Krajcik, B. Chen, A. Eisenkraft, D. Fortus, K. Neumann, J. Nordine & A. Scheff (Eds.), *Teaching and Learning Energy in K-12 Energy Education*. New York: Springer.
- Tastan, Ö. Y., Eylem, & Boz, Y. (2008). Effectiveness of Conceptual Change Text-oriented Instruction on Students' Understanding of Energy in Chemical Reactions. *Journal of Science Education and Technology*, 17(5), 444-453. doi:10.1007/s10956-008-9113-7
- Tatar, E., & Oktay, M. (2007). The effectiveness of problem-based learning on teaching the first law of thermodynamics. *Research in Science & Technological Education*, 29(3), 315-332.
- Trumper, R. (1993). Children's energy concepts: a cross-age study. *International Journal of Science Education*, 15(2), 139-148. doi:10.1080/0950069930150203

- Wang, L., Wang, W., & Wei, R. (2014). What Knowledge and Ability Should High School Students Have for Understanding Energy in Chemical Reactions? An Analysis of Chemistry Curriculum Standards in Seven Countries and Regions In R. Chen, A. Eisenkraft, D. Fortus, J. S. Krajcik, K. Neumann, J. Nordine & A. Scheff (Eds.), *Teaching and Learning of Energy in K-12 Education*. New York: Springer.
- Watts, D. M. (1983). Some alternative views of energy. *Physics Education*, 18(5), 213. doi:10.1088/0031-9120/18/5/307
- Wernecke, U. (2013). *Erfassung des Schülerverständnisses von Energie - Validierung eines quantitativen Erhebungsinstruments durch die Methode des Lauten Denkens*. [Assessing students' energy understanding. Validating a quantitative assessment tool via think-aloud protocols]. (Unpublished master thesis). Kiel University, Kiel/Germany.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (2007). *ACER ConQuest version 2.0: generalized item response modelling software*. Camberwell, VIC: Acer Press.

## APPENDICES

### Appendix 1

#### *Items, Item Parameters and Distractor/Attractor Analyses*

Available under <http://www.ipn.uni-kiel.de/en/the-ipn/departments/biology-education/research-and-projects/completed-projects/ipn-energy-project-biology>

### Appendix 2

#### *Analysis of Differences in Item Types*

Available under <http://www.ipn.uni-kiel.de/en/the-ipn/departments/biology-education/research-and-projects/completed-projects/ipn-energy-project-biology>

### Appendix 3

#### *Energy Learning Opportunities at Schools in Hamburg/Germany*

Available under <http://www.ipn.uni-kiel.de/en/the-ipn/departments/biology-education/research-and-projects/completed-projects/ipn-energy-project-biology>

### Appendix 4

#### *Additional Statistical Procedures: Multiple Regression and Rasch Analysis*

Available under <http://www.ipn.uni-kiel.de/en/the-ipn/departments/biology-education/research-and-projects/completed-projects/ipn-energy-project-biology>

**<http://iserjournals.com/journals/eurasia>**