International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

Refining a learning progression of energy

Jian-Xin Yao, Yu-Ying Guo & Knut Neumann

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2017
https://doi.org/10.1080/09500693.2017.1381356

Routledge
Taylor & Francis Group

Refining a learning progression of energy

Jian-Xin Yao [a,b], Yu-Ying Guo [a] and Knut Neumann [c]

[a]Department of Physics, Beijing Normal University, Beijing, People's Republic of China; [b]National Center for School Curriculum and Textbook Development, Ministry of Education, Beijing, People's Republic of China; [c]Department of Physics Education, Leibniz Institute for Science and Mathematics Education at Kiel University, Kiel, Germany

ABSTRACT

This paper presents a revised learning progression for the energy concept and initial findings on diverse progressions among subgroups of sample students. The revised learning progression describes how students progress towards an understanding of the energy concept along two progress variables identified from previous studies – key ideas about energy and levels of conceptual development. To assess students understanding with respect to the revised learning progression, we created a specific instrument, the Energy Concept Progression Assessment (ECPA) based on previous work on assessing students' understanding of energy. After iteratively refining the instrument in two pilot studies, the ECPA was administered to a total of 4550 students (Grades 8–12) from schools in two districts in a major city in Mainland China. Rasch analysis was used to examine the validity of the revised learning progression and explore factors explaining different progressions. Our results confirm the validity of the four conceptual development levels. In addition, we found that although following a similar progression pattern, students' progression rate was significantly influenced by environmental factors such as school type. In the discussion of our findings, we address the non-linear and complex nature of students' progression in understanding energy. We conclude with illuminating our research's implication for curriculum design and energy teaching.

ARTICLE HISTORY
Received 29 June 2016
Accepted 14 September 2017

KEYWORDS
Learning progression; energy concept; conceptual development

Energy is one of the most fundamental scientific concepts, around which science learning can be organised (Chen et al., 2014; Nordine, Krajcik, & Fortus, 2011). Learning progressions are empirically validated descriptions of successively more sophisticated ways of understanding scientific concepts (Smith, Wiser, Anderson, & Krajcik, 2006). As such, learning progression is expected to offer a framework for organising coherent instruction, which aligns multiple facets of science education (e.g. standards, curriculum, and assessment, see Duncan & Hmelo-Silver, 2009) based on the integration of empirical research on students' understanding of the concept, cognitive theories, and pedagogical tradition (Duschl, Maeng, & Sezen, 2011; Gotwals & Alonzo, 2012; National Research

CONTACT Yu-Ying Guo yyguo@bnu.edu.cn Department of Physics, Beijing Normal University, Beijing 100875, People's Republic of China

Council, 2007; Yao & Guo, 2014). Following the paradigm and methodology of learning progression research, in this study, we propose a refined learning progression for the energy concept and present findings from a large-scale cross-sectional study to support its validity.

Theoretical framework

During the past decade, in science education, learning progressions received increasing attention as a means to describe students' progression in understanding core concepts of science (e.g. Lee & Liu, 2010; Neumann, Viering, Boone, & Fischer, 2013; Stevens, Delgado, & Krajcik, 2010) and engaging in scientific practices (e.g. Gotwals, Songer, & Bullard, 2012; Lehrer & Schauble, 2012; Schwarz et al., 2009) across grades or grade bands. Although the term 'learning progression' is relatively new (Smith, Wiser, Anderson, Krajcik, & Coppola, 2004), the idea of describing how students develop an understanding of a scientific concept across grades can be traced back to research on student's 'conceptual trajectories' or 'progression in learning' (Driver, Leach, Scott, & Wood-Robinson, 1994; Taber, 1995). Driver et al. (1994), for example, already suggested that students enter formal schooling with their very own conceptions of a science concept and progress towards a scientific conception through a series of intermediate conceptions. Still, learning progression research offers some distinctive new features compared to the research on student's 'conceptual trajectories' or 'progression in learning' from the 1990s: (1) a focus on key ideas about a concept; (2) a foundation in theories of cognition; and (3) the usage of modern measurement methods (Black, Wilson, & Yao, 2011; Duschl et al., 2011; Yao & Guo, 2014). As the research on learning progressions keeps evolving, new trends like fusing disciplinary core concepts of science with scientific practice or including instructional factors have emerged (Duschl et al., 2011; Krajcik, Sutherland, Drago, & Merritt, 2012; Lehrer & Schauble, 2015). However, as we aimed at collecting evidence for national standard revision and building a foundation for future instructional research, we adopted a more traditional approach to learning progression research: the iterative 'design-test-revise' circle based on evidence from assessment (Black et al., 2011; Lehrer & Schauble, 2015; Salinas, 2009).

Learning progressions are supposed to not simply break down curriculum standards by grades, but to build on one or more so called progress variables. Progress variables define students' initial understanding when entering the learning progression and students' expected understanding at the end of the learning progression as well as the intermediate levels of understanding in between (e.g. Duschl et al., 2011; Yao & Guo, 2014). In this sense, a progress variable delineates one aspect of students' learning about a scientific concept. Information about the progress variables, which is relevant to students learning about a scientific concept, can stem from domain analysis, prior research on students' understanding of the concept, as well as (general) theories of cognition and learning (Black et al., 2011; Wilson, 2009). In order to identify progress variables and build an energy learning progression, we reviewed research on students' conceptions of energy and existing approaches to developing an energy learning progression.

Students' alternative conceptions about energy

Research on students' conceptions about energy indicated that for most students the concept of energy is far from being well understood (e.g. Duit, 1984; Solomon, 1983a; Trumper, 1993). In fact, this research indicated that students' hold a wide range of alternative conceptions that prove to be relatively stable throughout school (Solomon, 1983a, 1983b; Watts, 1983). Watts (1983) was amongst the first to propose framework of students' alternative conceptions of energy. This framework identifies a total of seven different alternative conceptions: energy as (1) something that is mainly associated with human beings, (2) something stored in certain objects such as batteries, (3) a dormant ingredient that needs a trigger to be released, (4) an obvious activity, (5) a (by)product of a particular situation, (6) a very general kind of fuel, and (7) a fluid that can be 'put in', 'transported', or 'conducted'. Research in different cultural settings has confirmed that students' conceptions about energy, as they enter formal learning and even after formally learning about energy, can be categorised into one of these seven alternative conceptions (e.g. Duit, 1984; Finegold & Trumper, 1989; Tan, 2010; Trumper, 1993). The findings that students hold (alternative) conceptions about energy as they enter formal learning, had a significant impact on energy instruction (e.g. Trumper, 1990, 1991). However, as they found a considerable number of students to still hold alternative conceptions after formally learning about energy (e.g. Duit, 1984), scholars continued in their effort towards developing a coherent framework for the teaching and learning of energy to promote a deeper understanding of energy.

Students' progression in developing a scientific conception of energy

An initial sequence of how students may develop a scientific conception of energy was proposed by Driver, Rushworth, Squires, and Wood-Robinson (1994). Driver et al. (1994) acknowledged that students hold alternative conceptions of energy as they enter formal learning about energy, but suggested that some of these alternative conceptions are more sophisticated than others and can be developed into a scientific conception by successively introducing students to key ideas about energy – amongst them, that living and non-living things can possess energy, how events can be described in terms of energy, and why energy is conserved even though it is degraded (see Table 1). Building

Table 1. The progression sequence proposed by Driver et al., Liu et al., and Neumann et al.

Driver et al.	Liu et al. (Grade/Mean age)	Neumann et al. (Grade)
Personal energeticness		
Energeticness of other living things		
Non-living things spontaneously can do things	Activity/Work (Grade 3/9.2)	
Energeticness of some non-living things that possess energy	Source/Form (Grade 4/10.2)	Source/Form (Grade 6)
Stored energy in elastic materials		
Gravitational potential energy		Transfer & Transform (Grade 8)
Describing events in energy terms	Transfer & Transform (Grade 8/14.2)	Dissipation (Grade 8)
Energy conservation	Degradation (High school/18.0)	Conservation (part of Grade 10)
Energy degradation and efficiency	Conservation[a]	

[a]Liu and McKeough adhere students' grades and age when 50% samples show 'understand' in that developmental level.
[b]Even high-school students failed to achieve a 50% correct rate in energy conservation, so there is no grade/age presented here.

on Driver Rushworth, et al.'s (1994) work, Liu and McKeough (2005) hypothesised that students progression in understanding the energy concept is characterised by five distinct, hierarchically ordered conceptions: perceiving energy as activities or abilities to do work (Activity/Work), identifying different energy sources and forms (Form/Source), understanding the nature and processes of energy transfer and transformation (Transfer/Transformation), recognising energy degradation (Degradation), and realising energy conservation (Conservation). Using data from the TIMSS study, Liu and McKeough (2005) were the first to provide evidence that students of greater age were holding more sophisticated conceptions of the energy concept – suggesting that with the key idea of energy being related to activity or work as a point of departure, the sequence of key ideas *forms/sources*, *transfer/transformation*, *degradation*, and *conservation* represent an ideal learning sequence. These findings were subsequently confirmed by several other researchers (Dawson-Tunik, 2006; Lee & Liu, 2010; Neumann et al., 2013). However, although previous research on students' progression towards a scientific understanding of energy uniformly confirms the key ideas about energy as a progress variable, this research has not reached consensus about how students learn about the individual key ideas.

A theory of students' progression in developing a scientific conception of energy

In addition to confirming the key ideas about energy as a progress variable, recent research has identified a second progress variable relevant to delineating an energy learning progression: a theory of cognitive development (e.g. Dawson-Tunik, 2006; Lee & Liu, 2010; Neumann et al., 2013). Using the key ideas as a progress variable reflects a conceptual change perspective. As students learn about the key ideas in the above sequence, their conception of energy changes – from a more alternative conception of energy as being related to human activity or the ability of machines to do work to a more scientific conception incorporating all four key ideas up to the idea that altogether energy is conserved. However, a learning progression built solely on the key ideas as a progress variable does not provide insights into how students learn about the individual key ideas (see Neumann et al., 2013). Integrating a theory of cognitive development as a progress variable holds the potential of providing a more differentiated description of how students' progress in developing a scientific conception of energy and thus a better guidance for organising instruction that best supports students' learning about energy.

The principle idea of integrating conceptual change with a theory of cognitive development is already present in some theories of conceptual change (for an overview of different theories of conceptual change, see Vosniadou, 2013). These theories understand students' conceptions of a scientific concept not as comprehensive and cohesive understandings (in the sense of beliefs, see DiSessa & Sherin, 1998), but as complex systems of fundamental elements (so called p-prims) that can be constructed, adjusted, and connected to others (e.g. diSessa, 2002). In this sense, the development or progression of student's conceptions about a particular scientific concept corresponds to a structural change of the complex systems of fundamental elements, including the integration of new elements as well as the establishment of new connection between existing elements. The process of integrating new elements and adjusting links between existing elements is expected to be organised by key elements (in the sense of key ideas, diSessa, 1988). That is students' progression in developing a scientific conception about energy can be described as the development of an increasingly complex system of information about the energy concept – guided by key ideas about energy such as forms, transformation, degradation, or conservation.

In the description of cognitive development, much research has been built on the idea that cognitive development can be described in terms of levels of increasingly more complex cognitive operations (e.g. Biggs & Collis, 1982; Commons et al., 1998; Fischer, 1980; Gagne, Wager, Golas, Keller, & Russell, 2005; Kauertz & Fischer, 2006). Fischer (1980), for example, proposed a hierarchy of 13 levels of cognitive development from single representations, over representational systems, systems of representational systems up to abstract systems of information and single principles. For example, Commons proposed a hierarchy model with 13 levels of cognitive development (Model of Hierarchical Complexity, MHC) and found it is effective for task difficulty prediction in some Piaget tasks (Commons et al., 2008). Kauertz and Fischer (2006) in turn suggested six levels to describe the complexity of a student's knowledge about a given concept: one fact, several facts, one relation, several unconnected relations, several connected relations, and conceptual understanding. Similar hierarchies of developmental levels can also be observed for the structure of observed learning outcome (SOLO) taxonomy (Biggs & Collis, 1982) or the knowledge integration framework (Liu, Lee, Hofstetter, & Linn, 2008). A comparison of these models exhibits a particular consistency across the individual (Table 2). For instance, in the lower performance levels, the 'pre-structural level' in the SOLO taxonomy corresponds to the 'no-link' and 'partial link' in the knowledge

Table 2. The four conceptual development levels.

Levels	Level description	Connection to previous hierarchical theories
Fact level	Students describe and interpret daily life phenomena using everyday experience and piece knowledge, which is unconnected to each other.	[Tasks structures: in orders lower than '8. Concrete' in MHC][a] [Knowledge integration level: no-link][b] [SOLO taxonomy: Pre-structural][c]
Mapping level	Students can generate concept by mapping its abstract feature to observable physical quantities.	[Tasks structures: order '9. Abstract' in MHC] [Knowledge integration level: partial link] [SOLO taxonomy: Uni-structural]
Relation level	Students can articulate relationships between several concepts or specific mechanisms.	[Tasks structures: order '10. Formal' in MHC] [Knowledge integration level: full link and complex link] [SOLO taxonomy: Multi-structural]
Systematic level	Students can coordinate more than one concept in multivariate systems in a variety of contexts.	[Tasks structures: order '11. systematic' in MHC] [Knowledge integration level: systemic link] [SOLO taxonomy: Extended abstract]

[a]There are 13 orders in the MHC proposed by Commons, Trudeau, Stein, Richards, and Krause (1998). They are: 0. Calculatory; 1. Sensory and motor; 2. Circular sensory–motor; 3. Sensory–motor; 4. Nominal; 5. Sentential; 6. Preoperational; 7. Primary; 8. Concrete; 9. Abstract; 10. Formal; 11. Systematic; 12. Meta-systematic.

[b]There are five levels in the model of knowledge integration proposed by Linn (2006). They are: No-link; partial link; full link; complex link; systemic link.

[c]There are four levels in the SOLO taxonomy. They are: Pre-structural; uni-structural; multi-structural; extended abstract.

integration framework; in the higher performance levels, the 'systemic link' level in the knowledge integration framework corresponds to the 'Systematic level' in the MHC. In addition to the consistency, these models have received widespread empirical validation in the context of science education in general (e.g. Bernholt & Parchmann, 2011; Kauertz & Fischer, 2006; Liu et al., 2008) and the energy concept in particular (e.g. Dawson-Tunik, 2006; Lee & Liu, 2010). Dawson-Tunik (2006), for example, examined students' progression in understanding energy based on Fischer's (1980) skill theory, and Lee and Liu (2010) examined students' progression in understanding based on the knowledge integration framework. None of this research, however, has successfully explored how students' progress in developing understanding of the individual core ideas.

In summary, previous research suggests that in addition to the key ideas as a progress variable, a theory of cognitive development needs to be taken into account as a second progress variable. Integration of both progress variables should yield information not only about the sequence of conceptions through which students progress in developing a scientific conception of energy, but how specifically students progress in developing each of the conceptions. Based on our review of the literature we proposed four *conceptual development levels* to describe students' progression in developing conceptions aligned with each of the four key ideas about energy: *Fact*, *Mapping*, *Relation*, and *Systematic*. A detailed description of each level and how it compares to other proposed hierarchies of cognitive development levels is also provided in Table 2.

The refined learning progression of energy and research questions

From our review of previous research towards developing a learning progression of energy, we identified two different approaches. The first is to focus on key ideas about energy, the second to utilise the cognitive development theory to describe how students develop a scientific conception about energy. In our research, we combine these two approaches, in order to seek a more differentiated and thus accurate description of students' progression. In a previous study we already examined the validity of a learning progression built on the integration of the two progress variable (Neumann et al., 2013). However, in this study, the hypothesised learning progression of energy could only be partly validated. Therefore, we aimed to re-examine the validity based on a more refined description of our learning progression.

The refined learning progression of energy is a two-aspect framework (Figure 1), adopting the four key ideas of energy (*form*, *transfer and transform*, *dissipation*, and *conservation*) as a first progress variable, and the four conceptual development levels (*Fact*, *Mapping*, *Relation*, and *Systematic*) as a second progress variable. We hypothesised that students progress from an alternative conception of energy towards a scientific conception about energy along a sequence of four key ideas about energy (horizontal axis), and four conceptual development levels (vertical axis). We developed a more specific description of each of the four conceptual development levels for each of the four key ideas based on the analysis of policy documents from prior research on students' understanding of energy (see also, Yao & Guo, 2014).

In addition to validating the refined learning progression in terms of the extent to which the key ideas and conceptual development levels are suitable to describe students' progression in developing a scientific conception about energy, we also investigated factors

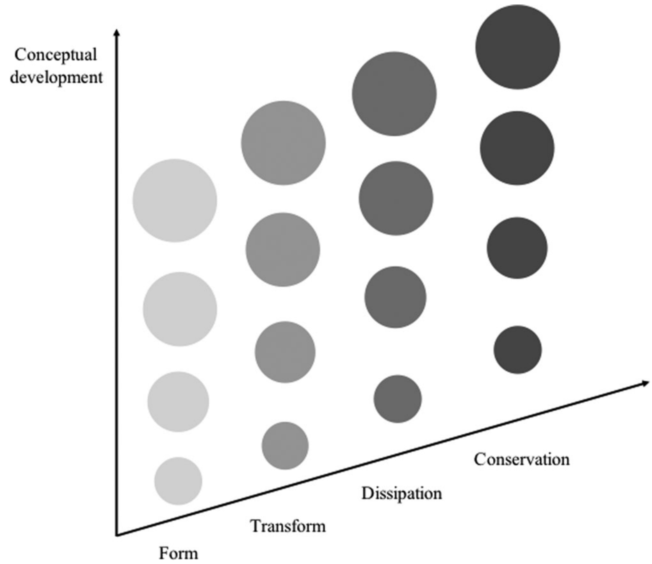

Figure 1. Hierarchical model for learning progression of energy concept.

potentially influencing students' progression. A major discussion around learning progressions focused on the question to which extent a learning progression necessarily applies to all students (e.g. Duncan & Hmelo-Silver, 2009). Research has found that different educational environments lead to different learning trajectories (e.g. Bianchini, 2017; Lee & Buxton, 2010; Lee & Liu, 2010; Plummer & Krajcik, 2010; see also National Research Council, 2012). Different educational environments sometimes are caused by social issues like inequality of regional economic development and school resources (e.g. Bianchini, 2017; Lee & Buxton, 2010), and sometimes are caused by curriculum or instruction (e.g. Lee & Liu, 2010; Plummer & Krajcik, 2010). For example, Lee and Liu found that students who took a physics science course exhibited significantly higher knowledge integration levels than students who took a life or earth science course, when assessing their energy understanding (Lee & Liu, 2010).

Based on previous research, we expected two factors – school district (urban vs. suburban) and school type (normal vs. model) to potentially influence students' progressions. More specifically we expected students form urban areas to exhibit an accelerated progression as urban areas are correlated with higher education investments compared to suburban and rural areas in Mainland China (Zhang, 2012). In the same way we expected students from model schools to exhibit an accelerated progression as model schools provide higher instructional quality in comparison to normal schools in Mainland China (Zhang, 2009).

We formulated three research questions to guide our study: To which extent (1) can students' progression in developing a scientific conception of energy be described in terms of key ideas about energy (i.e. the first progress variable), (2) can students' progression in developing a scientific conception of each of the four key ideas about energy be described in terms of conceptual development levels (i.e. the second progress variable), and (3) do students in different educational environments exhibit similar patterns in their progression towards developing a scientific conception of energy?

Method

Following previous works (e.g. Neumann et al., 2013), we developed a specific instrument, the Energy Concept Progressions Assessment (ECPA), to investigate students' progression in understanding the energy concept. We utilised Rasch analysis to examine instrument quality and obtain information about students' progression with respect to the hypothesised learning progression. Rasch analysis provides a technique by which the psychometric functioning of an instrument can be monitored, and the distribution of items and persons across a latent trait can be examined (e.g. Bond & Fox, 2006; Boone & Scantlebury, 2006). Following two pilot studies, we conducted the main study with sample of $N = 4550$ students from Grades 8 to 12 of schools from two districts in a major city in Mainland China.

Instrument development

In developing the ECPA we utilised items from Trumper (1998), Singh and Rosengrant (2003), and Neumann et al. (2013), as well as publicly available items from the TIMSS and PISA studies. All items were categorised with respect to the two progress variables. When items could not be categorised using the two progress variables, they were adapted in order to fit the progress variables if possible. We then invited experts in science education (including university researchers, professional test developers, and experienced school teachers) to assess the items' validity. In piloting stage, we conducted two studies both involving a paper and pencil test ($N_1 = 735$; $N_2 = 1033$) and an interview ($N_{i1} = 10$; $N_{i2} = 13$) with students from Grade 8 to Grade 12. We collected students' answers, their teachers' feedback, and the *dimensionality indicator*, *item fit*, and *item information curve* in Rasch analysis results. The experts were asked to review item design, item classification with respect to the learning progression, and item appropriateness for testing students at the respective grade levels – based on the wording of the items, our illustration about the learning progression, and the information from pilot-testing, as well as their own expertise. The experts identified a total of 56 quality items appropriate for inclusion in the ECPA item pool. The items, their classification with respect to the two progress variables and information such as item difficulty from piloting stage were compiled into a technical manual for future reference.

Cultural background and sample

In Mainland China, students start compulsory education at the age of 7 years. Compulsory education is composed of primary school (Grade 1–Grade 6) and middle school (Grade 7–Grade 9). Following an entrance examination, about half students continue learning in high school (Grade 10–Grade 12), others go to vocational education or work directly. Most provinces of China, including the city, in which our data were collected, adopt unified national curriculum standards. From Grade 4 to Grade 6, the general science course involves some instruction on energy. Physics courses start in Grade 8 and continue to Grade 12. Middle school physics courses are organised in terms of topics covered by grades, whereas high-school courses are organised in terms of curriculum modules. The energy, although very likely addressed in every topic, is explicitly covered in Grades 9, 10, and 12.

A total of 4550 students from Grades 8 to 12 of schools in a major city in Mainland China took part in the main study. Of these students, 49.5% ($N_f = 2254$) were girls, and 49.1% ($N_m = 2234$) were boys (62 students did not report their gender). The city, which is one of the most developed cities in China, has higher high school and college entrance rates than average. The schools were randomly chosen from two districts in the city: District F ($N_1 = 1873$), and District H ($N_2 = 2677$). District F used to be a suburban area not so far ago, District H is located in an urban area (in China, education is typically better in urban schools). Schools in both districts included normal schools and model schools. Model schools, with a reputation of providing better instructional quality, are more attractive to students. Therefore, some students switch schools, when they move from middle to high school (after Grade 9). Detailed information on the sample is provided in Table 3.

Data collection and analysis

A total of 52 items from ECPA item pool were used in the data collection. In order to obtain reliable measures for students from different grades, we compiled specific booklets for each grade. Each booklet contained more than one third of these items were identical across booklets for 2 or 3 grades. Data collection was simultaneously carried out at the beginning of the second semester of the school year. Students responses were scored as 0-wrong, 1-correct for multiple-choice items, open-ended items were scored as 0-wrong, 1-partially correct, 2- correct. Scoring was performed by two raters using a computer-based scoring system. In order to ensure a maximum of objectivity in the scoring process the raters received intensive training. Only after the raters achieved an interrater-agreement of 90% on training data, scoring of the actual data commenced.

52 items from ECPA item pool were used in final test. For comparing students in different grades, a *vertical linking* method was used when assembling questionnaires for every respondent population. This method arranged more than 1/3 same items (*linking items*) for different questionnaires that be used to test two or more grades. Then, putting all the items and respondents in one Rasch computation can make them linked together for future comparison and analysis (Bond & Fox, 2006). Objective items were coded dichotomy, and open-ended items were scored as 0-wrong, 1-partial right, 2- right, by at least two raters supporting by computer-based marking system. Before formal rating, raters must achieve an above 90% inter-rater reliability in the rater training process.

Rasch analysis was utilised in order to examine the quality of the ECPA test instrument and to locate students from different grades (i.e. students who answered different test booklets) on the same scale for further comparison and analysis (for details see Bond & Fox, 2006 or Liu & Boone, 2006). More specifically we employed the Partial Credit

Table 3. Sample information.

	District *F* (model school)	District *F* (normal school)	District *H* (model school)	District *H* (normal school)	Sum
Grade 8	231	120	335	142	828
Grade 9	244	126	426	175	972
Grade 10	283	148	401	207	1038
Grade 11	314	127	346	193	980
Grade 12	138	150	308	136	732
				In total:	4550

Rasch Model, which is an extension of the dichotomous Rasch model for polytomous data (Masters, 1982). We utilised the software Winsteps (version 3.72) to conduct the Rasch analysis (for a detailed description of Rasch analysis and its applications in science education, see Liu & Boone, 2006).

In a first iteration of Rasch analysis, we examined the extent to which the items fit the assumption of the Rasch model. A total of seven items that did not meet the relatively strict cut-off criteria we employed (i.e. infit MNSQ values between 0.80 and 1.20, see Neumann, Neumann & Nehm, 2011) were identified. These items were excluded from the data set and a second iteration of Rasch analysis was conducted. In this second iteration all of the (remaining) 47 items were found to meet the cut-off criteria. The item reliability was found to be 1.00 (item separation = 23.59), person reliability was 0.82 (person separation = 2.10). The average person ability was $M_{\text{person}} = -0.12$, which indicates that the test is slightly too difficult for the sample as the mean of the items was fixed to $M_{\text{item}} = 0$. However, as the variance of the person ability distribution was fixed to $\sigma_{\text{item}} = 1$, overall, the difficulty of the test may be considered adequate for the sample. A total of 46.6% raw variance can be explained, and the biggest unexplained variance in first contrast (potential dimension) is 2.1%. This suggests that the items form unidimensional latent scale, as was expected from previous research (e.g. Lee & Liu, 2010; Liu & McKeough, 2005; Neumann et al., 2013).

Results

The central result from a Rasch analysis is the so-called Wright Map. A Wright Map shows the distribution of the person ability estimates on the one side, and the item difficulty estimates on the other. This allows for comparing person ability, factors-related person ability to item difficulty, and factors related to item difficulty. We began our analysis with an inspection of the Wright Map (Figure 2) in regards to compare students' empirical learning progression and the hypothesised learning progression (Figure 1). The left side of the Wright Map shows the distribution of person ability estimates for three different groups: middle school students (Grade 8–Grade 9, light blue line), high-school students (Grade 10–Grade 11, dark blue line), and senior high-school students (Grade 12, red line). The right side shows the item difficulty distribution for the items grouped by key ideas about energy (*form*, *transfer and transform*, *dissipation*, and *conservation*). The item label colours indicate the conceptual development level assessed by the respective item: green for *Fact* level; blue for *Mapping* level; orange for *Relation* level; red for *Systematic* level. Interestingly, we find no sign of a clear sequence of the four key ideas about energy, while items are clearly arranged by conceptual development level. *Fact* level items are located at the bottom of the Wright map, which means these items were the easiest for sample students to solve, whereas *Systematic* level items are located at the top, indicating these items were the most difficult for students. *Mapping* and *Relation* level items are ranging in between as expected. In addition, there is correspondence between items' conceptual development levels and distributions of students (for example, it seems that students in middles school are 'progressing' from fact level to mapping level). In the following, we use statistical analyses as a lens to examine the validity of the two progress variables and patterns of students' progression as well as factors explaining differences between patterns.

Students' progression in terms of the progress variables

A valid progress variable for a learning progression should be an effective indicator of item difficulty (e.g. Neumann et al., 2013; Park & Liu, 2016). To examine the extent to which the hypothesised progress variables predict item difficulty, we conducted a two-step linear regression. The key ideas about energy and the conceptual development levels were included as independent variables and the item difficulty parameters from the Rasch analysis as the dependent variable. In the first step we used an all-possible-regression procedure, in the second step a backward elimination procedure. Results of linear regression (Table 4) show that the conceptual development levels have robust predication ability for item difficulty ($R^2 = .808$, $F_{(1,44)} = 185.43$, $p < .001$). However, adding another variable (the key ideas of energy) could not significantly improve the model ($\Delta R^2 = .002$, $p = .497$). This result preliminarily confirms the patterns we observed on the Wright map that the conceptual development levels manifested a hierarchical structure while the key ideas of energy did not.

To further examine the patterns on the Wright map, we individually analysed the effect of the two progress variables on item difficulty. As students' progression in developing a scientific conception of energy was found to be best represented as a uni-dimensional construct, we used one-way ANOVA to compare the mean difficulty of items across the four

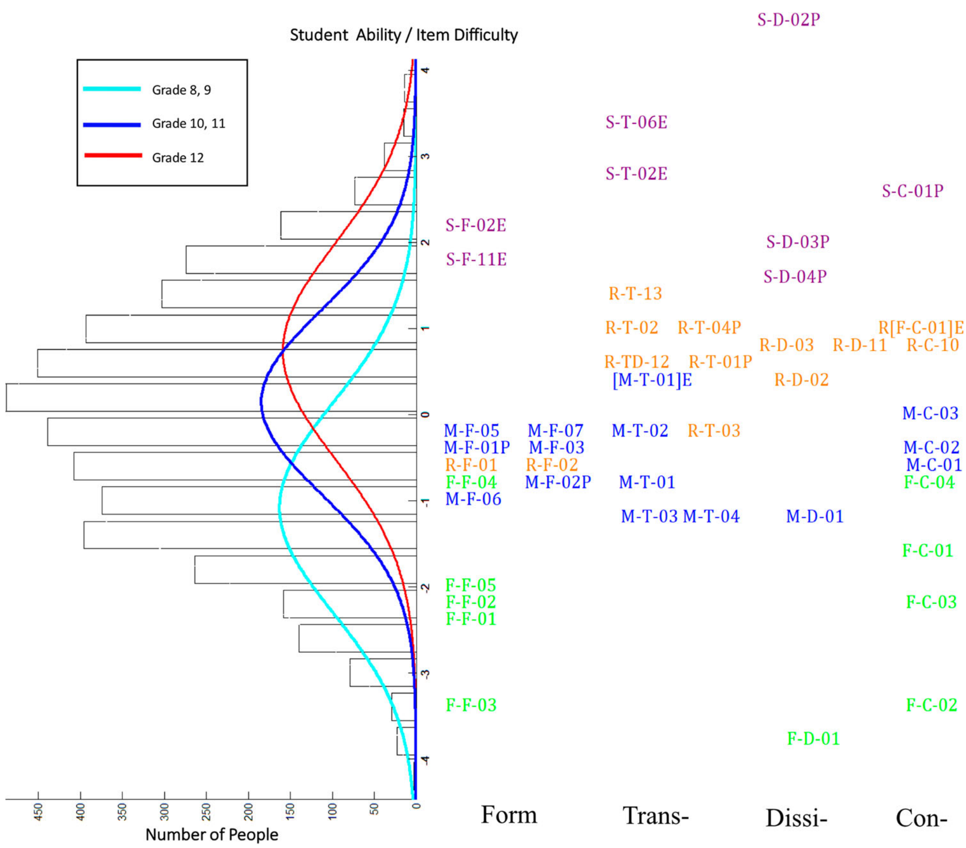

Figure 2. The Wright map of the final test.

Table 4. Results of linear regression.

	B	SE B	β
Step 1			
Constant	−3.782	0.296	
Conceptual development levels	1.543	0.113	.899***
Step 2			
Constant	−3.931	0.370	
Conceptual development levels	1.540	0.114	.860***
Key ideas of energy	0.067	0.098	.046

Note: Step 1's $R^2 = .808$, Step 2's $R^2 = .810$, $\Delta R^2 = .002$ ($p = .497$).

Table 5. Means of difficulty measurements and K–S test results of four developmental levels.

Variable	N	$M(SE)$	D	p
Fact	10	−2.25(0.32)	0.488	.971
Mapping	16	−0.56(0.12)	0.477	.977
Relation	12	0.40(0.17)	0.787	.566
Systematic	8	2.51(0.36)	0.503	.962

Table 6. Result of *post hoc* analysis on four developmental levels.

Between levels	ΔM	SE	p
Fact vs. mapping	−1.71	0.31	<.001
Mapping vs. relation	−1.07	0.29	.004
Relation vs. system	−2.12	0.34	<.001

key ideas and conceptual development levels. The results showed no significant differences in item difficulty across the four key ideas, $F(3,42) = 2.362$, $p = .085$. For the conceptual development levels, we first examined that the distribution of item difficulty among each conceptual development levels fits the normal distribution (Table 5). Then the ANOVA indicated significant differences in item difficulty across the four levels, $F(3,42) = 65.695$, $p < .001$, $\omega = .81$. A subsequent *post hoc* analysis (with Bonferroni correction) revealed significant differences between consecutive levels (Table 6).

In summary, the statistical results confirm the patterns observed in the Wright Map. Interestingly, we could not find the first progress variable (i.e. the key ideas about energy) to influence item difficulty significantly, while we found a clear effect of the second progress variable – conceptual development levels – on item difficulty estimates. In addition, we were able to confirm that the four conceptual levels as distinct levels in terms of the mean item difficulty, and that they differ across the levels and increase with higher levels of conceptual development.

Students' progression in terms of the educational environment

Using a Kruskal–Wallis one-way ANOVA, we found students' performance to increase with years of schooling (Table 7), $\chi^2(2) = 1855.57$, $p < .001$. A comparison of average item difficulty (Table 5) and average student ability (Table 7) suggests corresponding patterns between conceptual development levels and students' progression. One year into middle school students had reached the *Fact* level ($M_8 = -1.90$), after another year of study they reached the *Mapping* level ($M_9 = -0.48$). At the beginning of high school,

Table 7. Description statistics of the whole sample.

Code	Person count	Mean measure	Median	*S.D.*	*D* (K–S test)
Mid-Energy (G8)	828	−1.90	.03	0.91	1.84**
Mid-Energy (G9)	972	−.48	.04	1.19	1.61*
High-Energy (G10)	1038	.13	.04	1.16	2.19***
High-Energy (G11)	980	.69	.04	1.29	1.74**
High-Energy (G12)	732	.93	.05	1.41	1.69**
Total	4550	−.12	.02	1.54	

students range between *Mapping* and *Relation* level ($M_{10} = 0.13$). About half of the high-school students (52%) mastered the *Relation* level ($M_{11} = 0.69$) by Grade 11, but less than a quarter (17%) demonstrated a *Systematic* level performance, even after taking one additional physics modules in Grade 12 ($M_{12} = 0.93$). Note, that these are cross-grade comparison since we did not follow students across the year, but tested students from different grades at the same time. However, the findings provide initial insight into students' progression with respect to the hypothesised learning progression of energy.

In addition to investigating the students' progression as a whole, we also examined students' progression patterns for different educational environments: school area (urban vs. suburban) and school type (mode vs. normal school). Table 8 shows students' average performance for each of the four groups – broken down by grade. Figure 3 shows a graphical representation of this data – revealing different progression rates for the different groups. Note the dotted line between Grades 9 and 10 indicating that students had the opportunity to change school after *high-school entrance examination*. As a result, the differences in performance between these two grades are complicated to interpret, and their explanation is beyond the scope of this research.

At the beginning of, or more specifically the first year in middle school, model school students' abilities were found to outperform normal school students' in both District F, $t(311.88) = 2.25$; $p < 0.05$ and District H, $t(311.20) = 2.25$; $p < 0.001$. Students in suburban model schools performed better than students in urban model schools, while students in suburban normal schools performed worse than students in urban normal schools. But there was no significant difference between urban and suburban schools for the same school type. For students with one more year of education (i.e. students in Grade 9), we observed a significant difference between model schools and normal schools within each district: District F, $t(599.00) = 2.25$; $p < .001$, and District H, $t(368.00) = 2.25$; $p < .05$. In addition, although there was no significant difference between normal

Table 8. Performance of different groups.

	F (model school)		*F* (normal school)		*H* (model school)		*H* (normal school)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Grade 8	−1.7776	1.0421	−2.1681	0.7604	−1.8354	0.8928	−2.0153	0.7540
Grade 9	−0.5327	1.1367	−0.8396	1.1019	−0.1575	1.1783	−0.9487	1.0803
Changing (9–8)	1.2449	/	1.3285	/	1.6779	/	1.0666	/
Grade 10	−0.4682	1.1881	−0.4911	0.8406	0.8481	0.9321	−0.0148	0.9473
Changing* (10–9)	0.0645	/	0.3485	/	1.0056	/	0.9339	/
Grade 11	0.4456	1.0430	−0.778	1.0181	1.6159	1.0039	0.3787	1.0223
Changing (11–10)	0.9138	/	−0.2869	/	0.7678	/	0.3935	/
Grade 12	0.6015	1.1095	−0.1335	1.1613	1.9523	1.0551	0.1215	1.0541
Changing (12–11)	0.1559	/	0.6445	/	0.3364	/	−0.2572	/
Changing (12–10)	1.0697	/	0.3576	/	1.1042	/	0.1215	/

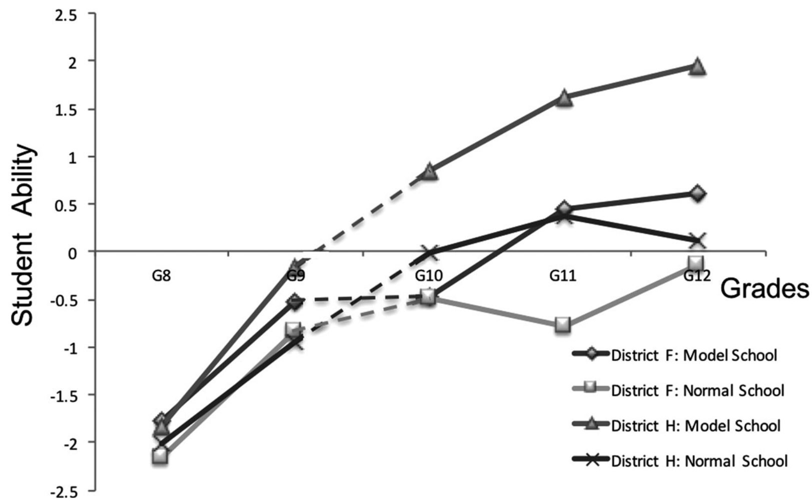

Figure 3. Performance changing pattern of different groups.

school students from the two districts, students in urban model school significantly ($p < .001$) outperformed their suburban counterparts. Model school students from Grade 9 in District F mastered the *Relation* level, while the other students were still struggling with mastering the *Mapping* level.

In high school, urban model school students outperformed other students from the beginning (i.e. in Grade 10). The interesting thing to note is the difference between urban normal school and suburban model school students. The comparison of urban normal school and suburban model school students indicated that urban normal school students showed a better performance in the starting year of high school (Grade 10), $t(488) = 4.54$, $p < .001$. However, one year later that difference was gone, and urban normal school students were eventually outperformed by suburban model school students in Grade 12, $t(272.00) = -3.67$, $p < .001$. That is, while urban normal school students start out high, their progression appears to be little, whereas suburban model school students show a significantly steeper progression. It seems that the school type significantly affects how and how much students can progress during high school. In terms of conceptual development levels, only students in urban model school were moving towards a systematic understanding. It also appears that other students were 'blocked' from reaching a respective level.

In summary, students' progression in developing a scientific conception follows one overall pattern: students progress further in higher grades (i.e. with more teaching on energy). In addition, we found a particular impact of the educational environment as students from urban model schools showed an increasingly progressed understanding of energy up to the level of a systematic understanding – compared to other schools where students (on average) did not progress beyond the *Relations* level.

Discussion and implications

The study presented in this paper aims to add to previous research by investigating a refined learning progression of energy based on the findings from previous research on students' progression in developing a scientific conception about energy. In doing so, we hypothesised two progress variables: key ideas about energy and conceptual development levels. In contrast to previous research (e.g. Liu & McKeough, 2005; Neumann et al., 2013) our results did not validate the key ideas about energy as a progress variable – putting the role of the key ideas students' progression in understanding the energy concept into question (RQ1). The results suggest, however, that students' progression can be described in terms of the second progress variable investigated in this study, the conceptual development levels (RQ2). These findings suggest that students' progression in developing a scientific conception of energy is more complex and requires further research.

In previous research on students' progression on understanding energy, researchers identified a sequence of conceptions related to four key ideas about energy – *forms*, *transfer and transformation*, *dissipation*, and *conservation* – along which they expected students to progress (e.g. Driver et al., 1994; Liu & McKeough, 2005; Neumann et al., 2013). This sequence was – as a whole – confirmed in multiple empirical studies (Dawson-Tunik, 2006; Liu & McKeough, 2005; Neumann et al., 2013). Interestingly in our study, although following the same approach as previous studies, we could not find any evidence supporting the hypothesis that students progress along this sequence in their understanding of energy. A more detailed review of previous research reveals that the key ideas did not necessarily represent distinct stages or levels of progression. For example, Liu and McKeough (2005) found no significant difference between 'Conservation' and 'Dissipation', while we observed no significant difference between 'Transfer and Transform' and 'Dissipation' in our previous study (Neumann et al., 2013). These findings, in combination with our findings in the present study, suggest that although the idea of energy forms serves as a foundation for developing a deeper understanding of energy, the other ideas may not necessarily be developed in a distinct sequence.

In addition to investigating students' progression in terms of key ideas about energy, previous research has also attempted to describe students' progression in terms of cognitive theories (e.g. Dawson-Tunik, 2006; Lee & Liu, 2010; Liu & McKeough, 2005; Neumann et al., 2013). Some positive results were achieved in their research as well as this one: the learning progression of energy can become more distinct with the lens of cognitive theories like hierarchical complexity, knowledge integration, and levels of conceptual development. These findings suggest that students' progression in developing a scientific conception should not solely be described in terms of a sequence of key ideas, but that students' progression in developing each of the key ideas needs to be taken into account.

In light of these findings, the conclusion we can draw, if we can draw one, is that students' progression in developing a scientific conception of energy is non-linear and complex. Most notably however, all the research – no matter whether investigating students' progression in terms of key ideas, conceptual development levels or both – shows a particular overlap between the hypothesised levels of development (e.g. Liu & McKeough, 2005; Neumann et al., 2013); to the extreme of finding no difference in how students progress in their understanding of the individual key ideas, suggesting that students develop an understanding of the key ideas in parallel. As a result, we propose a further refinement of our original learning progression of energy (Figure 1). The proposed refinement incorporates an integration perspective based on findings on how students develop an integrated understanding of energy (e.g. Nordine et al., 2011) and a re-interpretation of the theory of knowledge integration (e.g.

Linn, 2006). The idea is that students do not develop an understanding of each key idea independently or dependently in the sense that understanding of one key idea is the prerequisite for understanding another key idea, but interdependently (Figure 4; see Nordine, 2016). That is, we expect students to be able to acquire knowledge about all four ideas, but understanding to develop through an integration process that begins with understanding of forms, followed by understanding transfer and transformation, dissipation and conservation. There might be pieces of isolated knowledge (i.e. students may have memorised that energy is neither created nor destroyed without a deeper understanding of it), but as students progress towards a scientific conception of energy knowledge about each key idea will be increasingly linked with each other (see also Lee & Liu, 2010). For instance, reasoning that friction causing 'a moving object to stop also results in an increase in the thermal energy in both surfaces' (National Research Council, 2012, p. 125), requires at least knowledge about two forms of energy (kinetic energy and thermal energy), their transformation and transfer need to be linked. More sophisticated reasoning might even involve a discussion of how the thermal energy spreads out into the environment (i.e. leading to a degradation of the energy despite it being conserved). One next step of our efforts in developing a learning progression of energy will be to validate this further refinement in relation to previous research about energy.

In addition to obtaining important evidence for further refining a learning progression, we found that although students' progression follows the same overall pattern, students in different educational environments were progressing at different rates (RQ3). These results suggest that students progress through the identified learning progression levels in the same order – as it should be if the learning progression is valid (Duschl et al., 2011) – but do not necessarily reach the individual levels at the same time. The progression rate depends on the quality of the educational environment or instruction respectively confirms that students' progression is not (solely) driven by maturation but determined by schooling on energy. Again, this speaks for the validity of the learning progression (Krajcik et al., 2012). More importantly, however, the results echo concerns expressed by the National Research Council of the United States:

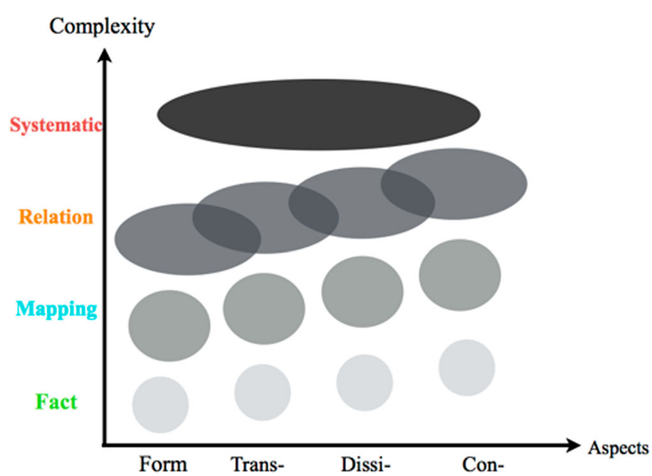

Figure 4. Integrating hierarchical model for learning progression of energy concept.

> Today there are profound differences among specific demographic groups in their educational achievements and patterns of science learning (…) The reasons for these differences are complex, and researchers and educators have advanced a variety of explanations (…) The first links differences in achievement to differences in opportunities to learn because of inequities across schools, districts, and communities. (National Research Council, 2012, p. 279)

The mere fact that there are differences is nothing new. So to some extent, the progression pattern of students from urban model schools (the schools likely to provide the highest quality of education in Mainland China) was expected. The insight, however, that although students in suburban normal schools start out higher, they are eventually outperformed by suburban model schools shows the importance of high quality instruction over contextual factors such as socio-economic status (see also Hattie, 2009). The findings also highlight the importance of ensuring high quality instruction for all students (Frankenberg, Garces, & Hopkins, 2016).

The findings from our study have – beyond informing the development of a learning progression of energy – several implications both for science instruction and science education policy. The implications for science instruction align mostly with existing approaches to energy instruction (i.e. Nordine et al., 2011; Nordine, 2016; see also Chen et al., 2014). Our findings suggest that energy instruction should not solely be organised by key ideas about energy but take into account how students learn about the key ideas and how they link the key ideas to each other. That is, students should not be taught thoroughly about energy forms, before receiving teaching about transformation, or energy dissipation before conservation, as suggested by previous research (i.e. Liu & McKeough, 2005; Neumann et al., 2013), but engage in modelling and explaining phenomena, planning and performing investigations of phenomena or discuss about different explanations of a phenomenon – all based on bringing together their knowledge about the different key ideas about energy (see also Nordine, 2016; Yao, Guo, & Neumann, 2016).

As to science education policy, in line with previous research (Liu & McKeough, 2005; Neumann et al., 2013) this study highlights the importance of aligning national standards with empirical insights into students' progressions in K-12. As previous work on students' progression in learning about energy has informed the science education policies such as the *Next Generation Science Standards* (NGSS Lead States, 2013), we think our work can inform the continued efforts to refine existing policies such as China's national science curriculum standards (e.g. Ministry of Education, P. R. China, 2017a, 2017b, 2017c). Levels of students' progression in developing an understanding of core science concepts, which are not only based on expert experience but also on empirical data about students' actual progression, can help design a 'spiral curriculum' (Bruner, 1977) to best support not only students' learning with-in one grade, but across multiple grades or even grade bands. The People's Republic of China has taken a first step by formulating national standards for primary school science based on learning progressions (Ministry of Education, P. R. China, 2017a).

Another important implication we observed comes from the inequity in students' performance across different educational settings. While such equity is an issue we are facing across a wide range of countries (see Bianchini, 2017 for a discussion), it is an issue that clearly needs to be addressed in the not-too-distant future. More specifically we think, that our findings on the effect of the educational environment highlights the importance of and need for opportunity-to-learn standards in addition to content or performance standards (see National Research Council, 2014). In addition, an 'adaptive instruction' (Corcoran, Mosher, & Rogat, 2009) has been highly recommended in suggestions for the implementation of standard (e.g. Ministry of Education, P. R. China, 2017b, 2017c).

Overall, our study emphasises on how learning progressions are a powerful lens for the systematic examination and subsequent optimisation of science education. Integrating cognitive science and science education, learning progression research can help meet the demand for and alignment of standards, curriculum (thus instruction), and assessment.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Ministry of Education of the People's Republic of China [contract grant number 13YJA880022].

ORCID

Jian-Xin Yao http://orcid.org/0000-0001-7301-8963

References

Bernholt, S., & Parchmann, I. (2011). Assessing the complexity of students' knowledge in chemistry. *Chemistry Education Research and Practice*, *12*(2), 167–173. doi:10.1039/C1RP90021H

Bianchini, J. A. (2017). Equity in science education. In K. S. Taber & B. Akpan (Eds.), *Science education: An international course companion* (pp. 455–464). Rotterdam: Sense.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York, NY: Academic Press.

Black, P., Wilson, M., & Yao, S.-Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspective*, *9*(2–3), 71–123. doi:10.1080/15366367.2011.591654

Bond, T. G., & Fox, C. M. (2006). *Applying the Rasch model: Fundamental measurement in the human sciences*. New York, NY: Routledge.

Boone, W. J., & Scantlebury, K. (2006). The role of Rasch analysis when conducting science education research utilizing multiple-choice tests. *Science Education*, *90*(2), 253–269. doi:10.1002/sce.20106

Bruner, J. S. (1977). *The process of education*. Cambridge, MA: Harvard University Press.

Chen, R. F., Eisenkraft, A., Fortus, D., Krajcik, J., Neumann, K., Nordine, J. C., & Scheff, A. (2014). *Teaching and learning of energy in K–12 education*. Heidelberg: Springer.

Commons, M. L., Goodheart, E. A., Pekker, A., Dawson, T. L., Draney, K., & Adams, K. M. (2008). Using Rasch scaled stage scores to validate orders of hierarchical complexity of balance beam task sequences. *Journal of Applied Measurement*, *9*(2), 182.

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. *Developmental Review*, *18*(3), 237–278. doi:10.1006/drev.1998.0467

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). *Learning progressions in science: An evidence based approach to reform*. Consortium for Policy Research in Education Report #RR-63. Philadelphia, PA: Consortium for Policy Research in Education.

Dawson-Tunik, T. L. (2006). Stage-like patterns in the development of conceptions of energy. In X. Liu & W. J. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 111–136). Maple Grove, MN: JAM Press.
diSessa, A. A. (2002). Why 'conceptual ecology' is a good idea. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 28–60). Dordrecht: Kluwer Academic.
diSessa, A. A. (1988). Knowledge in pieces. In G. F. P. Pufall (Ed.), *Constructivism in the computer age* (pp. 49–70). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
DiSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education*, *20*(10), 1155–1191. doi:10.1080/0950069980201002
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*, 75–100. doi:10.1080/03057269408560040
Driver, R., Rushworth, P., Squires, A., & Wood-Robinson, V. (1994). *Making sense of secondary science: Support materials for 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
Duncan, R. G., & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, *46*(6), 606–609. doi:10.1002/tea.20316
Duschl, R., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education*, *47*(2), 123–182. doi:10.1080/03057267.2011.604476
Finegold, M., & Trumper, R. (1989). Categorizing pupils' explanatory frameworks in energy as a means to the development of a teaching approach. *Research in Science Education*, *19*(1), 97–110. doi:10.1007/BF02356850
Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, *87*(6), 477–531. doi:10.1037/0033-295X.87.6.477
Frankenberg, E., Garces, L. M., & Hopkins, M. (Eds.). (2016). *School integration matters: Research-based strategies to advance equity*. New York, NY: Teachers College Press.
Gagne, R. M., Wager, W. W., Golas, K. C., Keller, J. M., & Russell, J. D. (2005). *Principles of instructional design*. Belmont, CA: Wadsworth.
Gotwals, A. W., & Alonzo, A. C. (2012). Introduction: Leaping into learning progressions in science. In A. Alonzo, & A. Gotwals (Eds.), *Learning progressions in science* (pp. 3–12). Rotterdam: Sense.
Gotwals, A. W., Songer, N. B., & Bullard, L. (2012). Assessing students' progressing abilities to construct scientific explanations. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science* (pp. 183–210). Rotterdam: Sense.
Hattie, J. (2009). *Visible learning*. London: Routledge.
Kauertz, A., & Fischer, H. E. (2006). Assessing students' level of knowledge and analysing the reasons for learning difficulties in physics by Rasch analysis. In X. Liu & W. J. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 212–246). Minneapolis, MN: JAM Press.
Krajcik, J. S., Sutherland, L. M., 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. 261–284). Münster: Waxmann.
Lee, O., & Buxton, C. A. (2010). *Diversity and equity in science education: Research, policy, and practice*. New York, NY: Teachers College.
Lee, H. S., & Liu, O. L. (2010). 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
Lehrer, R., & Schauble, L. (2012). Seeding evolutionary thinking by engaging children in modeling its foundations. *Science Education*, *96*(4), 701–724. doi:10.1002/sce.20475
Lehrer, R., & Schauble, L. (2015). Learning progressions: The whole world is NOT a stage. *Science Education*, *99*(3), 432–437. doi:10.1002/sce.21168

Linn, M. C. (2006). The knowledge integration perspective on learning and instruction. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 243–264). New York, NY: Cambridge University Press.

Liu, X., & Boone, W. J. (2006). *Applications of Rasch measurement in science education*. Maple Grove, MN: JAM Press.

Liu, O. L., Lee, H.-S., Hofstetter, C., & Linn, M. C. (2008). Assessing knowledge integration in science: Construct, measures, and evidence. *Educational Assessment*, *13*(1), 33–55. doi:10.1080/10627190801968224

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

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149–174. doi:10.1007/BF02296272

Ministry of Education, P. R. China. (2017a). *Primary school science curriculum standards for compulsory education*. Beijing: Beijing Normal University Press.

Ministry of Education, P. R. China. (2017b). *Physics curriculum standards for senior high school*. Beijing: People's Education Press.

Ministry of Education, P. R. China. (2017c). *Chemistry curriculum standards for senior high school*. Beijing: People's Education Press.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.

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

National Research Council. (2014). *Developing assessments for the next generation science standards*. Washington, DC: National Academies Press.

Neumann, I., Neumann, K., & Nehm, R. (2011). Evaluating Instrument Quality in Science Education: Rasch based analyses of a Nature of Science test. *International Journal of Science Education*, *33*(10), 1373–1405.

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. doi:10.1002/tea.21061

NGSS Lead States. (2013). *The next generation science standards*. Washington, DC: National Academies Press.

Nordine, J. (2016). What students should know about energy. In J. Nordine (Ed.), *Teaching energy across the sciences K-12* (pp. 17–38). Washington, DC: NSTA Press.

Nordine, J., Krajcik, J., & Fortus, D. (2011). Transforming energy instruction in middle school to support integrated understanding and future learning. *Science Education*, *95*(4), 670–699. doi:10.1002/sce.20423

Park, M., & Liu, X. (2016). Assessing understanding of the energy concept in different science disciplines. *Science Education*, *100*(3), 483–516. doi:10.1002/sce.21211

Plummer, J. D., & Krajcik, J. (2010). Building a learning progression for celestial motion: Elementary levels from an earth-based perspective. *Journal of Research in Science Teaching*, *47*(7), 768–787. doi:10.1002/tea.20355

Salinas, I. (2009). *Learning progressions in science education: Two approaches for development*. Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., … Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*(6), 632–654. doi:10.1002/tea.20311

Singh, C., & Rosengrant, D. (2003). Multiple-choice test of energy and momentum concepts. *American Journal of Physics*, *71*(6), 607–617. doi:10.1119/1.1571832

Smith, C. L., Wiser, M., Anderson, C. W., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic-molecular theory. *Measurement Interdisciplinary Research & Perspectives*, *4*(1-2), 1–98. doi:10.1080/15366367.2006.9678570

Smith, C., Wiser, M., Anderson, C., Krajcik, J., & Copolla, B. (2004). Implications of research on children's Learning for assessment: Matter and atomic molecular theory. Paper Commissioned by the National Academies Committee on Test Design for K–12 Science Achievement, Washington, DC: National Academies.

Solomon, J. (1983a). Learning about energy: How pupils think in two domains. *European Journal of Science Education*, *5*(1), 49–59. doi:10.1080/0140528830050105

Solomon, J. (1983b). Messy, contradictory and obstinately persistent: A study of children's out-of-school ideas about energy. *School Science Review*, *65*(231), 225–230.

Stevens, S. Y., Delgado, C., & Krajcik, J. S. (2010). Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal of Research in Science Teaching*, *47*(6), 687–715. doi:10.1002/tea.20324

Taber, K. S. (1995). Development of student understanding: A case study of stability and lability in cognitive structure. *Research in Science & Technological Education*, *13*(1), 89–99. doi:10.1080/0263514950130108

Tan, X. (2010). *An investigation on middle school students' understanding about the concept of energy in physics*. Unpublished dissertations, Beijing: Beijing Normal University.

Trumper, R. (1990). Being constructive: An alternative approach to the teaching of the energy concept – part one. *International Journal of Science Education*, *12*(4), 343–354. doi:10.1080/0950069900120402

Trumper, R. (1991). Being constructive: An alternative approach to the teaching of the energy concept part two. *International Journal of Science Education*, *13*(1), 1–10. doi:10.1080/0950069910130101

Trumper, R. (1993). Children's energy concepts: A cross-age study. *International Journal of Science Education*, *15*(2), 139–148. doi:10.1080/0950069930150203

Trumper, R. (1998). A longitudinal study of physics students' conceptions on energy in pre-service training for high school teachers. *Journal of Science Education and Technology*, *7*(4), 311–318. doi:10.1023/A:1021867108330

Vosniadou, S. (2013). *International handbook of research on conceptual change* (2nd ed.). New York, NY: Routledge.

Watts, D. M. (1983). Some alternative views of energy. *Physics Education*, *18*(5), 213–217. doi:10.1088/0031-9120/18/5/307

Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching*, *46*(6), 716–730. doi:10.1002/tea.20318

Yao, J.-X., & Guo, Y.-Y. (2014). 为学生认知发展建模: 学习进阶十年研究回顾及展望 [Model building for students' cognitive development: A review of ten-year research on learning progression]. *Journal of Educational Studies*, *5*, 35–42.

Yao, J.-X., Guo, Y.-Y., & Neumann, K. (2016). *Integrated learning progressions advancing synergetic development of energy understanding and scientific explanation*. Paper presented at the 2016 Annual Meeting of the American Educational Research Association (AERA), Washington D.C., USA.

Zhang H. (2009). 我国公立普通高中"校际差异"与"择校"关系探究 [Investigation on public high schools' school difference and school selection in China]. *Research on Fundamental Education*, (1), 3–4.

Zhang, W. (2012). 城乡教育一体化:教育公平的时代诉求 [Integration of rural and urban education: Epoch request of educational equity]. *Educational Research*, *391*(8), 13–18.