

Probing the Relation between Students' Integrated Knowledge and Knowledge-in-Use about Energy using Network Analysis

Marcus Kubsch ^{1*}, Jeffrey Nordine ¹, Knut Neumann ¹, David Fortus ², Joseph Krajcik ³

¹ IPN - Leibniz Institute for Science and Mathematics Education at Kiel University, GERMANY

² Department of Science Teaching, Weizmann Institute of Science, Rehovot, ISRAEL

³ Michigan State University, East Lansing, Michigan, USA

Received 14 January 2019 • Revised 13 February 2019 • Accepted 14 February 2019

ABSTRACT

Modern science standards emphasize knowledge-in-use, i.e., connecting scientific practices with content. For knowledge to become usable in knowledge-in-use performances, students need well organized knowledge networks that allow them to activate and connect sets of relevant ideas across contexts, i.e. students need integrated knowledge. We conducted a longitudinal interview study with 30 students in a 7th grade energy unit and used network analysis to investigate students' integrated knowledge, i.e., their knowledge networks. Linking these results with results from knowledge-in-use assessments, we found a strong connection between integrated knowledge and knowledge-in-use about energy. Further, we found evidence that well-connected ideas around the idea of energy transfer were particularly helpful for using energy ideas in the knowledge-in-use assessments. We present network analysis as a valuable extension of existing approaches to investigating students' knowledge networks and the connection between them and knowledge-in-use.

Keywords: energy, integrated knowledge, knowledge-in-use, network analysis

INTRODUCTION

Science standards such as the US *Next generation science standards* (NGSS Lead States, 2013) or the German *Bildungsstandards* (Sekretariat der ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2004) emphasize that a central goal of science education is that students are enabled to use their knowledge to make sense of the designed and natural world, i.e., students should be enabled to demonstrate knowledge-in-use by integrating disciplinary knowledge with scientific practices (Harris, Krajcik, Pellegrino, & McElhaney, 2016). For knowledge to become usable in such a way, students need to develop a well-connected knowledge network around central ideas of the domain as this facilitates retrieval and application (Bransford, 2000). The importance of knowledge networks organized around core ideas is recognized in modern educational standards, e.g., the NGSS emphasize "Disciplinary Core Ideas" and the German Science Standards introduced "Basic Concepts".

However, as Schwartz and Arena (2013) argue, science assessments often primarily focus on how much knowledge student have acquired instead of measuring to what extent students have an integrated knowledge. Students that have an integrated knowledge have well organized knowledge networks that allow them to activate and connect sets of relevant ideas across contexts. During the last decade, a number of authors have started to address this issue and developed different measures that describe to what extent students can connect ideas (Lee, Liu, & Linn, 2011; O. L. Liu, Lee, Hofstetter, & Linn, 2008; O. L. Liu, Ryoo, Linn, Sato, & Svihi, 2015; Nordine, Krajcik, & Fortus, 2011; Won, Krabbe, Ley, Treagust, & Fischer, 2017). These studies have often used the knowledge integration perspective (Linn, 2006), which views learning as a process in which students develop increasingly better organized networks of ideas, i.e., while students learn, they (re)organize their knowledge networks around core ideas in a domain. These knowledge networks are considered to be the basis that allows students to

Contribution of this paper to the literature

- This paper contributes to the existing literature on students' integrated knowledge by providing evidence that supports the premise of modern science standards of deep learning around small sets of core ideas as a key factor for knowledge-in-use.
- This paper also contributes to the literature on the teaching and learning of energy as it supports theoretical work that suggests that energy transfer can be a particularly helpful energy idea to make sense of phenomena.
- Lastly, this paper presents how network analysis can be used to provide qualitative and quantitative information about students' knowledge networks.

demonstrate knowledge-in-use, i.e., connecting scientific practices, disciplinary core ideas, and cross-cutting concepts to interpret and explain real world phenomena (Harris et al., 2016).

Existing measures of students' knowledge networks have either provided quantitative information about the extent to which students have organized knowledge networks (e.g., Lee & Liu, 2010) or qualitative information about how students organize their knowledge networks (Won et al., 2017). Neither approach has linked their results to knowledge-in-use in the sense of connecting practices, content, and cross-cutting concepts. In consequence, the relationship between students' knowledge-in-use and integrated knowledge in a domain is only little researched. We present a network analytical approach that provides qualitative and quantitative information about students' integrated knowledge, i.e., how students activate and connect ideas, and explore connections between quantitative network measures and knowledge-in-use.

This study is part of a broader project on the teaching and learning of energy in middle school. As students usually hold a wide range of ideas about energy, the energy concept provides an excellent venue to explore how network analysis can help to investigate students' knowledge networks and their relationship to students' ability to demonstrate knowledge-in-use. Because the knowledge-in-use assessments used in this study have been described in depth elsewhere (Neumann et al., 2018), we only revisit them relatively quickly.

BACKGROUND

Integrated Knowledge and Knowledge-in-Use

Knowledge, or more specifically the organization of knowledge, is often characterized as a network-like structure (e.g., Anderson, 1983; National Academies of Sciences, Engineering, and Medicine, 2018). This characterization has been recognized and successfully applied in neuro- and cognitive science (Griffiths, Kemp, & Tenenbaum, 2008; McClelland & Axel Cleeremans, 2009; Park & Friston, 2013; Thagard, 2000). Synthesizing findings from psychology, the learning sciences, and educational research, learning can be viewed as a process in which students' knowledge networks undergo a process of restructuring, i.e., students add new ideas to their knowledge networks, sort out ideas, and establish, change, refine, and strengthen the connections between ideas (diSessa & Sherin, 1998; Linn, 2006). If certain ideas are commonly activated together and across a range of instances, they form patterns of activation which represent conceptual understanding (Derry, 1996). Within these patterns of activation, some ideas have larger span (diSessa, 2013), i.e., some ideas are more strongly connected than others and serve as hubs that connect multiple ideas. Those hubs represent core ideas in a domain. Further, patterns of activation become more likely to be activated in certain contexts if they prove to be of explanatory value, i.e., they can be used productively to interpret and explain phenomena (Smith, diSessa, & Roschelle, 1994). In sum, students that have an integrated knowledge have well organized knowledge networks around core ideas which allow them to activate and connect relevant ideas across context.

This perspective on learning resonates well with major findings from the cognitive sciences and science education. The focus on the re-structuring of idea networks acknowledges that students come to the classroom with strong assumptions about how the world works (Bransford, 2000) and that this prior knowledge plays an important role in any learning process. Science education research has identified many of these conceptions or frameworks (e.g. Watts, 1983) and developed strategies to engage them by pointing to the superior explanatory value of their scientific counterparts (e.g. diSessa, 1988). Research on differences among novices and experts has revealed that novices and experts have differently organized knowledge networks (e.g. Chi, Feltovich, & Glaser, 1981). Chi et al. (1981) found that in the context of physics, experts used core ideas such as the principle of energy conservation to make sense of phenomena while novices relied on surface features of the phenomena. In essence, the knowledge of experts is well-organized around core ideas in the domain like energy conservation while the knowledge of novices typically exists as a set of loosely connected or disconnected ideas, resulting in novices and experts often

perceiving the same phenomenon differently – experts are more likely to notice deep structure while novices are commonly drawn to the surface features that experts ignore (Hmelo-Silver & Pfeffer, 2004).

Recent science standards such as the NGSS or the German science standards emphasize that a central goal of science education is to enable students to apply their knowledge in the context of scientific practices in a wide range of contexts (Harris et al., 2016), i.e., to demonstrate knowledge-in-use. To be able to connect disciplinary ideas, scientific practices and cross-cutting ideas to make sense of a range of phenomena, students need knowledge networks that are organized around core ideas of the domain, that is, integrated knowledge (Bransford, 2000; diSessa, 1988; Linn, 2006; Schneider & Stern, 2009). From an instructional perspective, the question arises which of the possible organizations of knowledge networks help to promote knowledge-in-use. More specifically, we might ask what is the relative importance of ideas in a given domain or what ideas are better suited as core ideas for an integrated knowledge? As studies usually focus either on knowledge-in-use or integrated knowledge (e.g., Lee and Liu, 2010) there has been little research into the relationship between the two. Further, existing measures of how students connect ideas largely focus either on qualitative features of students' knowledge networks but do not connect those results to students' performance on related knowledge-in-use assessments (Won et al., 2017) or distinguish broader levels of connectedness in students' knowledge networks that relate to students' performance on related assessments (not knowledge-in-use assessments) but provide little information about how students actually connect ideas (Lee & Liu, 2010).

Existing Measures of Integrated Knowledge

In order to measure integrated knowledge i.e., how students organize their knowledge networks and activate and connect ideas to make sense of phenomena, researchers have taken a number of approaches, making different assumptions and using different testing formats: Kauertz and Fischer (2006) used multiple choice (MC) tasks, Lee and Liu (2010) constructed response (CR) tasks, Nordine et al. (2011) used interviews, and Won et al. (2017) used concept mapping.

Multiple Choice Tasks. Kauertz and Fischer (2006) used a measure of six different levels of complexity that distinguished the number of ideas and the number of connections between those ideas. Low complexity was characterized by disconnected individual ideas, while high complexity was related to systemic connections between ideas organized around core ideas in a domain. Thus, the complexity measure targets different qualitative stages in students' knowledge networks. Using the complexity framework, Kauertz and Fischer designed MC tasks that match the different levels of complexity. More precisely, Kauertz and Fischer defined which ideas students had to link to answer a question correctly at the different levels of complexity, e.g., on a lower complexity level a student might have to link a form of energy with speed, on a higher complexity level a student might have to link two forms of energy with the idea of energy transformation. While the choices for linking ideas such as "energy transformation" are reasonable, they might not reflect the ideas that students actually used when answering the question as MC tasks allow for guessing or other test-wiseness related strategies (Pellegrino, Chudowsky, & Glaser, 2004). However, using IRT modeling, they found a strong and highly significant correlation between the complexity levels that they had assigned to the tasks and task difficulty. By mapping individual ability and task difficulty using IRT, Kauertz and Fischer concluded that students that score higher are able to elaborate more complex scientifically valid links among ideas relevant to the tasks, i.e., that higher scoring students have better organized knowledge networks.

Constructed Response Tasks. Liu, Lee, Hofstetter and Linn (2008) established one approach that was tested and refined over multiple studies (e.g., Lee & Liu, 2010). In this approach a combination of multiple choice (MC) and constructed response (CR) items are used (see [Figure 1](#) as an example).

The source of energy for the earth's water cycle is the

- (a) Wind
- (b) Sun's radiation
- (c) Earth's radiation
- (d) Sun's gravity

Explain your choice.

Main Ideas:

- Sun idea: Sun warms (heats) up earth.
- Evaporation idea: Water is evaporated (water changes from liquid to gas or water vapors).
- Condensation idea: Water changes from gas to liquid to form clouds.
- Precipitation idea: Different types of precipitation fall from sky.
- Energy idea: Energy is required in evaporation.
- Water cycle idea: Water cycle consists of evaporation, condensation, and precipitation.

Links:

- Sun-energy link: Sun's radiation energy is used for evaporation of water.
- Evaporation Link: Water molecules are broken apart from the surface of the liquid water.

Knowledge Integration Level	Criteria	Examples
Irrelevant: Off-task	Wrote some text unrelated to the item.	<ul style="list-style-type: none"> • Because I think so.
No-link: Nonnormative ideas	Elicited nonnormative ideas or restated the multiple-choice answer chosen.	<ul style="list-style-type: none"> • Sun's radiation gives light to the water cycle. • Wind moves the water through the air.
Partial-link: Normative ideas	Elicited any one idea listed above.	<ul style="list-style-type: none"> • Sun warms up water. • Sun evaporates water.
Full-link: Single link between two normative ideas	Used either sun-energy link or evaporation link.	<ul style="list-style-type: none"> • The water molecules are broken off from the surface of the water. • Sun heats up the water to evaporate.
Complex-link: Two or more links among three or more normative ideas	Used either sun-energy or evaporation link and added an additional idea or link listed above.	<ul style="list-style-type: none"> • The sun's radiation makes the water evaporate into the air which goes through evaporation, condensation, then precipitation.

Figure 1. Example for items and knowledge integration rubric used by Lee and Liu (2010)

Beginning with the MC part, a phenomenon is presented, and students have to choose the correct answer. In the associated CR part, students are asked to explain or justify their answer from the MC part. The extent to which students are able to connect ideas identifiable in the explanations is then coded on a scale from not linking relevant ideas ("no link") through an intermediate level of linking ("partial link") to the level of elaborating a scientifically valid link between two ideas relevant to a given context ("full link"). A further "complex link" level is identified when students use more than one link between at least three ideas, but only very few students reach that level (Lee et al., 2011). The authors use IRT models to link the scales of the MC questions to those of the CR question. Higher levels of performance on the MC questions map onto higher levels of performance on the CR questions, thus linking higher test scores to higher levels of connecting ideas. Thus, the study supports the relationship between better organized knowledge networks and performance on related assessments found by Kauertz and Fischer (2006). Further, Lee and Liu (2010) measure how students connect ideas less distal than Kauertz and Fischer (2006) as they observe the extent to which students link ideas in their answers directly. However, as Lee and Liu (2010) note, the extent to which students connect ideas may be underestimated due to students known lack of commitment to formulating scientific explanations. Students who are not motivated to explain their choice or not motivated to do so in a thorough way will link few if any ideas in their explanations.

Interviews. Nordine et al. (2011) measured which ideas students connect in the domain of energy using an interview about instances approach (Osborne and Gilbert, 1980). In this approach, the interviewed student is presented pictures that illustrate various everyday situations and asked whether the pictures illustrate their idea of

energy. If the student answers "yes", the interviewer asks for the students' reasoning. If students use ambiguous wording the interviewer can, using the language of the student, probe into the students' reasoning to clarify what is meant. To measure how students connect ideas from these interviews, the researchers used a qualitative analysis to determine whether the students used "transformation" to link at least two forms of energy (kinetic energy etc.). This level of connecting ideas appeared to be equivalent with a middle ground between the "full link" and "partial link" level identified by (Lee et al., 2011). Thereby, how students connect ideas was dichotomously assigned to represent either at least a "partial link" level or to be below the "partial link" level. The level of connecting ideas was significantly correlated with students' scores on an energy concept questionnaire adapted from Swackhamer and Hestenes (2005). This supports that better organized knowledge networks are in fact related to higher performance in interpreting and explaining phenomena using core ideas in a domain. While measuring how students connect ideas on a coarser grain size than Lee and Liu (2010), the interview approach can help to counter the motivational issues in the Lee and Liu (2010) approach and is potentially more sensitive to developmental stages in students' knowledge networks where students may not yet be able to articulate links between ideas in a written form.

Concept mapping. Won et al. (2017) asked students to draw concept maps (Novak, 1990) on the topic of energy and used the multiple choice energy concept assessment from (Neumann, Viering, Boone, & Fischer, 2013). Concept maps provide vivid representations of how students connect ideas and allow to distinguish qualitative differences in how students connect ideas. Different features related to students' knowledge networks, e.g., the number of ideas in a map or the number of connections between ideas, can be used to score the maps. However the scoring of the maps employed by Won et al. (2017) is only weakly related to how students perform on the energy concept assessment. Won et al. (2017) identify various issues such as students' inexperience with the concept mapping task that may explain the unexpectedly weak relationship between students' concept maps and their performance on the energy concept assessment. In addition, we consider concept maps problematic for measuring how students connect ideas to make sense of phenomena, as they are typically constructed by the students themselves. There is a threat to validity in this approach: a student may certainly be able to draw a concept map that connects two relevant ideas, e.g. energy transformation and different forms of energy, but not be able to apply those ideas when asked to make sense of a phenomenon and vice versa. Thus, concept maps may be suited to assess the structure of students' declarative knowledge (Ruiz-Primo, 2004) but appear less suited to assess students' ability to link ideas to make sense of phenomena, i.e., students' integrated knowledge.

In summary, the MC, the CR, and the interview approach to measuring students' integrated knowledge found a strong relation between different levels of connectedness in students' knowledge networks and performance on related assessments (although not knowledge-in-use assessments). The MC approach hinges on the assumption that one can reliably infer the ideas that a student has connected from the answer he gives on a MC test. The CR approach assigns levels of knowledge integration based on the extent to which a student uses specific a priori defined linking ideas. The same is true for the interview approach which is potentially more sensitive to less developed stages of knowledge integration than the CR based approach. Finally, the concept mapping approach does not find a strong relation between students' knowledge networks as assessed by concept mapping and students' performance on related MC assessments but provides insight into qualitative differences in how students connect ideas.

Measuring Integrated Knowledge using Network Analysis

We propose to synthesize the existing approaches to measuring how students connect ideas through the application of network analytical tools. First, consider again the definition of integrated knowledge: it describes students' ability to coordinate a set of core ideas consistently across phenomena to make sense of them. Thus, we propose, similar to the CR or interview approach, to measure which ideas students use when interpreting and explaining phenomena. However, instead of assessing whether students use a priori defined linking ideas and assigning respective knowledge integration scores, we draw on the central assumption that all ideas a student uses in response to a phenomenon, are connected. We consider this assumption, which is used with great success in automated semantic analysis (Landauer, 2014), automated coding of student responses (e.g. Zehner, Sälzer, & Goldhammer, 2016), or classification of texts (Blei & Lafferty, 2007), warranted because we aggregate across phenomena and do not interpret co-occurrences of ideas in a single phenomenon. When we aggregate across phenomena, sets of ideas that co-occur across many phenomena appear to be stronger linked than those sets of ideas that are only used in a single phenomenon or across few phenomena. Stronger linked ideas should reflect core ideas in a domain as - following the definition above - the central feature of core ideas is that they are used consistently across phenomena. Following this approach leads to relational data that can be visualized as a network. Similar to a concept map, the network should reveal major characteristics of how students link ideas, yet these maps are generated by analyzing students' use of a concept when making sense of phenomena rather than mapping the concept itself. In addition, network analytical methods allow for the quantification of aspects of the networks

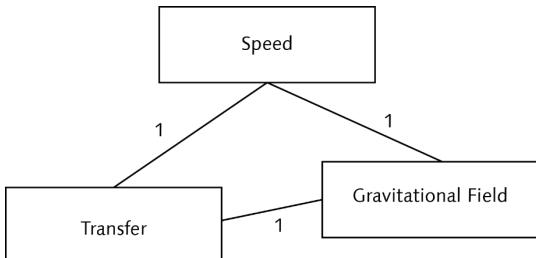


Figure 2. Example of network for one student and one phenomenon. Lines between boxes represent co-occurrence and the number next to the lines indicated frequency of co-occurrence

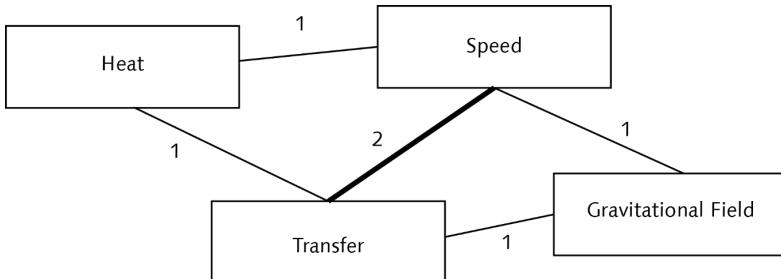


Figure 3. Example of network for one student for two combined phenomena. Lines between boxes represent co-occurrence and the number next to the lines indicated frequency of co-occurrence

which can then be related to students' performance on knowledge-in-use assessments. In the following, we will describe how knowledge networks can be constructed and quantitatively analyzed.

From Ideas to Networks. To derive networks from student explanations of a range of phenomena, we draw on the central assumption that ideas that co-occur within a student's response to a single phenomenon are connected and that repeated co-occurrence (across multiple phenomena) is an indicator for the strength of the connection. In the following, we provide an example for an individual student that has explained two phenomena A and B. Assume that the student used the ideas *speed*, *transfer of energy* and *gravitational field* in his explanation of phenomenon A. This leads to a network for phenomenon A in which all three ideas - *speed*, *transfer of energy* and *gravitational field* – are connected with each other (**Figure 2**).

Now assume that the same student uses the ideas *speed*, *heat*, and *transfer of energy* in the explanation of phenomenon B. Now, combining the networks for phenomenon A and phenomenon B results in a network (**Figure 3**) in which we add the idea *heat* to the existing network for phenomenon A and connect *heat* to *speed* and *transfer of energy*.

Further, since *speed* and *transfer of energy* co-occur in phenomenon A and phenomenon B, we increase the strength of the connection between *speed* and *transfer of energy* which is represented by the increase in width of the line connecting the ideas in **Figure 3**. In the same manner, we can combine networks across more of phenomena. Further, aggregating across individual-level networks gives collective networks, e.g., for students in a classroom. Such aggregated networks can be useful for understanding the prevalence of ideas and connections that exist in student populations (Loh & Subramaniam, 2018).

Quantifying Networks. Visual inspection of the network tells us which ideas co-occur and how often they do so. Network analysis provides the tools to construct numerical measures that help to quantify knowledge networks with respect to integrated knowledge.

The Role of Single Ideas. Not all ideas are equally important and powerful to make sense of phenomena. Central ideas of a domain that serve as hubs that connect multiple ideas are more powerful as they are applicable across a range of phenomena and contexts (National Research Council, 2012). For example, the idea that kinetic energy is related to the square or speed is correct, but not as central or broadly applicable as the idea that energy is conserved. Core ideas serve as linking ideas in knowledge integration measures (Lee and Liu, 2010) and have a larger span (diSessa, 1988; diSessa & Sherin, 1998). With respect to our networks this translates to the strength of the connections of a single idea to the remainder of the network. Network analysis offers a range of so-called centrality measures that capture this idea (Freeman, 1978). The simplest of these is degree and it captures how connected an idea is to the remainder of the network. Degree quantifies the number of other ideas a single idea is connected to, taking the strength of the connections into account, e.g., *energy transfer* in **Figure 3** has a degree of four (one connection to *heat*, one connection to *gravitational field*, and one connection of strength two to *speed*) and *heat* has a degree of two. Thus, the higher degree of *energy transfer* with respect to *heat* reflects that *energy transfer* was used

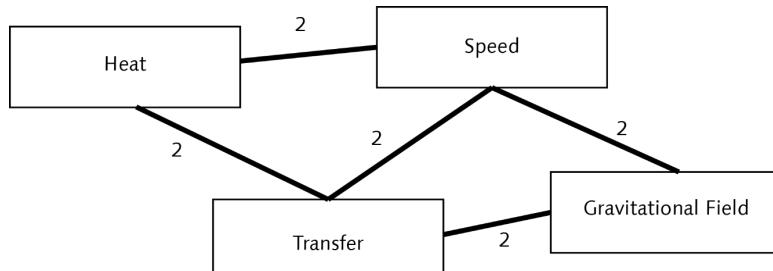


Figure 4. Example of network for one student for two combined phenomena. Lines between boxes represent co-occurrence and the number next to the lines indicated frequency of co-occurrence

across more phenomena and in coordination with a broader range of ideas than *heat*. Note that the numerical values of network measures are often only meaningful relative to other networks that were created using the same process and often have little meaning on their own.

A Network Measure of Integrated Knowledge. Higher levels of integrated knowledge are typically associated with increasingly interconnected coordination of ideas, assuming that the ideas that are normative (Anderson & Schunn, 2000; Bransford, 2000; diSessa, 1988; Koponen & Huttunen, 2013; Linn, 2006; Rumelhart & McClelland, 1986). To quantify how interconnected a network as a whole is, we draw on what Rafols and Meyer (2010) call network coherence. To calculate network coherence, we calculate the closeness (Freeman, 1978) for each idea to each other idea and average across all ideas. Closeness C is defined as the inverse of the sum of the distance d between a single idea x and all other ideas y in the idea network:

$$C(x) = \frac{1}{\sum_y d(y, x)}$$

In the formula, $d(y, x)$ is the shortest path between any two ideas. In our idea networks, we would consider *energy transfer* and *speed* strongly connected with a connection strength of 4 when a student uses those two ideas in four out of five phenomena. To interpret the strength of connection between ideas as a distance, we simply take the inverse of the strength of connection between two ideas to ensure that strong connections between ideas reflect short distances and vice versa. Let us provide an example: consider the network in **Figure 2**. To ensure that strong connections reflect short distances we simply take the inverse. Thus, the shortest path between the stronger connected ideas *energy transfer* and *speed* becomes 0.5, whereas the shortest path between the looser connected ideas *energy transfer* and *heat* becomes 1. Following the formula, the closeness of *energy transfer* is 0.4 as adding the shortest paths between *energy transfer* and *heat*, *energy transfer* and *speed*, and *energy transfer* and *gravitational field* gives 2.5 (1+0.5+1) and the inverse of 2.5 is 0.4. Following the same procedure for *gravitational field* gives 0.25 (the inverse of 1+1+2) which reflects that *gravitational field* is less strongly connected with the remaining network than *energy transfer*. Averaging across the closeness of each idea in the network gives a network coherence of 0.325. Compare the network in **Figure 3** with the network in **Figure 4**.

The network in **Figure 4** suggests that the student used all four ideas in both phenomena, demonstrating a more integrated knowledge as she coordinated a larger number of ideas consistently across phenomena than the student from the network in **Figure 3**. The measure of network coherence reflects this as it takes the larger value of 0.58. Similar to degree, network coherence has no natural scale, i.e., the numerical value can be used to compare different networks that were created using the same process but has no meaning of its own. To account for different numbers of ideas in networks, network coherence can be normalized by multiplying it with the number of ideas in a network (Freeman, 1978). In sum, network coherence is high in dense networks with strong connections and low in sparse networks with weak connections and should thusly provide a holistic measure of integrated knowledge.

Research Questions

To investigate to what extent the proposed network analytical approach illuminates how qualitative differences in students' knowledge networks are related to students knowledge-in-use, we asked the following research questions:

RQ1: How is the change in students' integrated knowledge over the course of instruction as measured by a network analytical approach related to the change in students' knowledge-in-use?

RQ2: What qualitative information about students' knowledge networks does the network analytical approach provide and how is this related to students' knowledge-in-use?

Table 1. Mean and standard deviation for average grade across science, math, and, English (11 point grading system), and learning gain (z-Score) between whole sample and interview group

	Group	M	SD
Average grade	Interview group	8.10	3.74
	Whole Sample	6.67	3.75
Learning gain	Interview group	0.47	0.70
	Whole Sample	0.44	0.54

METHODS

To address the research questions, we conducted semi-structured interviews before and after a middle school unit about energy. In the interviews, students were presented brief videos of five different phenomena and asked to explain them. We then used network analytical methods to characterize and represent how students organized their ideas within idea networks and to assess the extent to which students developed an integrated knowledge over the course of the instructional unit.

We assessed students' knowledge-in-use about energy using distal, NGSS aligned knowledge-in-use assessment (Neumann et al., 2018) before and after the energy unit and linked students' performance on those tasks to how students' idea networks changed over the course of instruction.

In the following sections, we describe the energy unit, describe our data sources, and then focus on how we analyzed the interviews to create representations of students' idea networks.

Sample and Setting

Our study was conducted as part of a larger project investigating the teaching and learning of energy in middle school. The learning environment consisted of an approximately 10-week long energy unit (Nordine et al., 2018) that emphasized energy transfers in interpreting and explaining phenomena as emphasized in the NGSS (NGSS Lead States, 2013). The energy unit was enacted by three teachers in the 7th grade of two rural middle schools located in the Midwestern United States. A sample of $N=30$ students, representing approximately 10% of the total sample of $N=294$ students, was interviewed. The teachers selected the students based on a request from the researchers to identify a sample of students who they felt would be comfortable being interviewed and who represented a range of student abilities in their classes. To test the representativeness of our interview-sample for the whole sample, we compared our interview sample to the sample as a whole based on students' average grade and gain on the unit's pre / post-test (**Table 1**).

Although the differences between interview sample and the sample as a whole on both measures (gain: $t(234) = 0.32, p = .75, d = 0.07$; average grade: $t(234) = 1.52, p = .13, d = 0.38$) are statistically not significant, they indicate that our interview sample consisted of slightly higher achieving students.

Energy Ideas in the Unit

Prior to the unit, we expected students' ideas about energy to be relatively disconnected and mostly about different forms of energy and some ideas about the transformation of energy which they may have picked up in earlier learning opportunities (Chen et al., 2014). From the literature on student conceptions of energy (for an overview see Duit, 2014) and students usage of energy ideas to solve problems (Chabalengula, Sanders, & Mumba, 2012; Driver and Warrington, 1985) we expected that prior to the unit, students would rarely use energy ideas to explain and interpret phenomena and instead use ideas such as forces and everyday interactions between objects.

This study is situated in an energy unit (Nordine et al., 2018) that emphasizes energy transfers between systems as envisioned in the *Framework for K-12 Science Education*. The so-called "systems-transfer approach" is a transfer-only approach and does not distinguish between different forms of energy but instead treats energy as a unitary entity. During phenomena, interacting systems transfer energy between them and in parallel, each system undergoes a process that changes the amount of energy the system has. Energy conservation is implied because when energy is transferred from a system, it is always transferred to another system. Therefore, we would expect that, at the end of the unit students should have developed idea networks that are organized around the central idea of energy transfer and are able make sense of phenomena using those ideas, i.e., demonstrate knowledge-in-use about energy. Other energy ideas such as forms of energy or energy transformations should play a less significant role.

Data Sources

Interviews - Protocol and Phenomena. To gain insight in how students connect ideas to make sense of phenomena, students were interviewed individually according to a semi-structured interview protocol. In this protocol, students were shown short (5-10 second) videos of five different phenomena and prompted to explain each one. After introductions and obtaining student consent, we showed students the video of the first phenomenon and asked: "How would you explain this phenomenon?". After students' initial answer, non-instructional prompts were used to clarify ambiguous statements and elicit which ideas students used, e.g. if a student referred to energy transfer to/from an object, the interviewer might ask "What do you mean by energy transfer?" After prompting, students were shown the video of the next phenomenon.

The pre and post interviews were identical and addressed five phenomena: a cup above a burning candle that starts spinning, a bouncing ball that eventually stops bouncing, melting ice¹, a person pushing a barrel up a ramp, and an electric heater heating up (all phenomena are available as online supplemental). The phenomena were selected to cover a broad range of phenomena and topics (thermodynamic, mechanics, electricity, biological systems, etc.) and to be rich in the sense that one could interpret and explain them on different levels of detail (e.g., using a macroscopic or particle perspective to explain melting ice).

Knowledge-in-Use Energy Assessments. We used NGSS-aligned knowledge-in-use assessments (Neumann et al., 2018) to investigate how students' performance at interpreting and explaining phenomena using energy ideas changes over the course of the unit. The assessments were distal in the sense that they are not aligned to what was taught in the unit but the NGSS energy performance expectations for middle school. The assessments emphasized knowledge-in-use, i.e., blending a disciplinary knowledge (in this case about energy) with science practices such as constructing an explanation and cross-cutting ideas such as systems (Harris et al., 2016). The tests included twelve open-ended tasks. We used evidence-centered design as laid out in the NRC report on *Developing assessments for the Next Generation Science Standards* (Pellegrino et al., 2014) and more specifically following the procedure by (Harris et al., 2016) to ensure that the tasks blended scientific practices, disciplinary core ideas, and cross-cutting practices, i.e., the task were designed to capture knowledge-in-use. Eight of those tasks were aligned to the NGSS energy middle school performance expectations and four to the K-5 energy performance expectations. The tasks were scored by experienced teachers (average percentage agreement across all tasks: 83%). We used a polytomous IRT model (Bond & Fox, 2015) to calculate students' knowledge-in-use ability. The infit and outfit of the items was between 0.90 and 1.20 and reliability was found to be .76.

Analysis

Coding – from Interviews to Ideas. As a first step, we transcribed all interviews. Then we used qualitative content analysis (Mayring, 2014) to determine which ideas students' used in their explanations. We followed a deductive approach in which phrases served as a coding unit and the complete explanatory account of one phenomenon as context unit, i.e., the number of times a student mentioned an idea such as speed in his or her explanation of one phenomenon was not accounted for. The deductive categories consist of numerous energy ideas and other scientific concepts (e.g., gravity, forces, or the particle model of matter) that students could use to construct scientifically valid explanations of the phenomena. We referred to literature on energy learning progressions (e.g., Neumann, Viering, Boone, & Fischer, 2013) to identify relevant energy ideas (e.g., forms, transfer, transformation). For all other concepts we considered grade band appropriate conceptions, e.g., for the concept of "force" we considered the respective literature on student conceptions (Gilbert, Watts, & Osborne, 1982) and labeled the resulting code "pushes / pulls". An example of our category system can be found in **Table 2**.

Table 2. Example part of category system used for deductive coding of ideas

Variable	Definition	Anchor Example	Borderline Cases
Transfer	S talks about energy being transferred from one system (including fields and objects) to another.	" <i>The fire is transferring energy up to the cup.</i> "	Energy has to be transferred from one to another system. Energy only being transferred "from" or "to" something is not sufficient.
Transformation	S talks about one form of energy being transformed into another.	" <i>When the person drops the ball, it [the energy] goes from potential to kinetic energy.</i> "	
Interaction	S talks about interactions between objects that are characterized by pushing / pulling.	" <i>this boy is pushing a barrel up a hill</i> "	This includes friction and changes in shape.
Gravity	S talks about gravity pulling something down.	" <i>Because gravity is trying to force it down.</i> "	

In some cases, interviewers asked the student to describe the phenomenon prior to explaining it. Ideas mentioned in a merely descriptive part were not coded. Further, in some cases the interviewers used instructional prompts (e.g. directly prompting for energy). In these cases, anything after such prompts was ignored. If instructional prompts occurred in more than one scenario, we excluded the interview because ideas potentially introduced by the interviewer would distort our results. In consequence, data from 13 students were excluded from the analysis because of instructive prompts. The remaining sample of N=17 students represents 7% of the total number of 236 students that studied the unit with these teachers and took the unit pre- and post-test.

We used a second rater to conduct an inter-rater reliability analysis in which we found very good agreement (Cohen's Kappa $\kappa = .85$, (Landis and Koch, 1977)) and after an analysis of the conflicting codes, we were able to resolve them.

Network Analysis. The codes resulting from qualitative content analysis provide us with a list of ideas that co-occur in each explanatory account of each phenomenon for each student. Our central assumption for constructing networks is that ideas that co-occur in the explanatory account of a phenomenon are connected. We combined networks across all phenomena the student explained to construct an individual-level network for each student. Further, aggregating across individual-level networks gives networks for the interview sample as a whole for pre- and post. We used the network measure of degree to investigate the role of single ideas in students' networks and the network measure of coherence to investigate students' levels of knowledge integration. For our calculations we used the *igraph* (Csardi & Nepusz, 2006) and *tneat* (Opsahl, 2009) packages in R (R Development Core Team, 2008).

Assessing Knowledge Integration. To investigate how our network analytical approach compares to the existing approaches to measure how students connect ideas based on a priori defined linking ideas as in e.g., Lee and Liu (2010), we adopted the knowledge integration scoring scheme used by Lee and Liu (2010) to the phenomena used in our study. **Figure 5** shows an example rubric for the bouncing ball phenomenon.

Knowledge Integration level	Criteria	Examples
Irrelevant: Off-task	Elicited unrelated ideas.	-
No-link: Non-normative ideas	Elicited non-normative ideas.	<i>"Well, when it falls it bounces up but it's like, it's slowly stopping."</i>
Partial-link: Normative ideas	Elicited any one idea stated.	<i>„Well, it's getting pulled down by the gravity and it's bouncing up 'cause of the material it's made of.“</i>
Full-link: Single link between two normative ideas	Used any of the links stated.	<i>„When it hits like the ground it transfers the energy to the ground.“</i>
Complex-link: Two or more links among three or more normative ideas	Used any of the links stated and added an additional idea or link.	<i>“As it drops, the gravity pulls it down. Every time the ball drops energy is transferred a little bit into the floor and the surroundings.”</i>

Bouncing ball

Ideas:

- Energy: 1) kinetic energy, 2) potential energy, 3) thermal energy
- Fields: 1) fields between the atoms, 2) gravitational field
- Force: 1) gravity pulls the ball down
- Particles: 1) the ball hits against air particles, 2) the ball is made up of particles
- Temperature: 1) the temperature of the ball / floor / air increases

Links:

- Energy transfer link: 1) energy is transferred to / from the gravitational field from / to the ball, 2) energy is transferred to / from the field between the atoms of the ball from / to the ball, 3) energy is transferred from the ball to the ground / air (surroundings)
- Energy transformation link: 1) kinetic energy is transformed into potential energy and vice versa, 2) kinetic energy is transformed into thermal energy, 3) kinetic energy is transformed into elastic energy and vice versa
- Energy degradation link: 1) energy is transferred from the ball to the environment

Figure 5. Knowledge integration rubric for bouncing ball phenomenon based on Lee and Liu (2010)

We assigned each phenomenon a knowledge integration score based on the rubric and averaged across the phenomena to come up with a knowledge integration score for each student. We used a second rater to conduct an inter-rater reliability analysis on 16 randomly chosen explanatory accounts (representing 10% of the explanatory accounts) and found substantial agreement (Cohen's Kappa $\kappa = .83$, (Landis & Koch, 1977)).

RESULTS

RQ1

During instruction, coherence (which is measured across phenomena) of students individual networks increased significantly ($t(6.82) = 30.24, p < .001$) from pre to post (Figure 6).

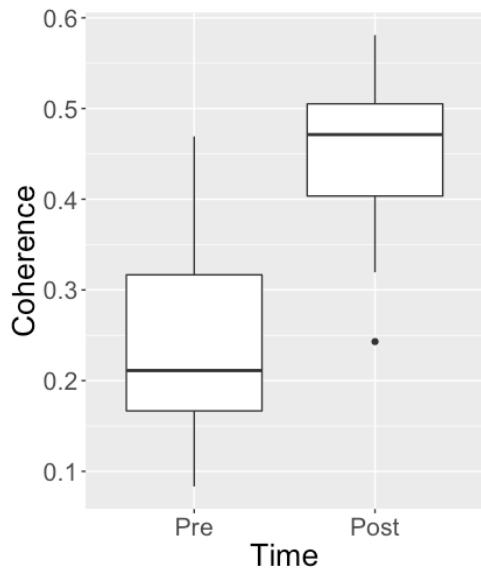


Figure 6. Network coherence for pre- and post-networks of individual students

Table 3. Mean and standard deviations of the network measure, knowledge integration measures, and knowledge-in-use assessments

Variable	M	SD
Knowledge-in-use assessments gain	0.22	0.61
Network Coherence gain	0.25	0.16
Knowledge integration measure gain	1.73	0.65

Table 4. Regression models predicting knowledge-in-use gain

Model	β^2	p	R ²
Knowledge-in-Use gain ~ network coherence gain	0.55	0.022	0.30
Knowledge-in-Use gain ~ knowledge integration gain (based on Lee and Liu (2010))	0.41	0.05	0.23

To explore to what extent a change in students integrated knowledge as measured by coherence is related to the change in students' knowledge-in-use, we used a linear model to estimate how strongly the change in coherence is related to the change in students' performance on the knowledge-in-use energy assessments. To compare the network analytical approach to existing approaches to measuring how students connect ideas, we also modelled the gain in the knowledge-in-use assessments using the Lee and Liu (2010) based knowledge integration measure. **Table 3** shows the mean and standard deviations of the coherence measure, knowledge integration measure, and the knowledge-in-use assessments. The results of the models predicting the gain in the knowledge-in-use assessments from the gain in the coherence measure and the gain in the knowledge integration measure are presented in **Table 4**.

An increase in one standard deviation in coherence gain maps onto an increase of .55 ($p=.022$) standard deviations in the gain on the knowledge-in-use energy assessment. The model explains roughly a third of the variance in students' gain from pre to post ($R^2 = .3$). Similar results hold for the regression model calculated with the measure of knowledge integration adapted from Lee and Liu (2010).

To further investigate the validity of the network analytical measure, we compared the initial answers to the four scenarios from the post interviews. **Table 5** contrasts the initial student answers that led to above median (coherence = 0.55) and below median (coherence = 0.37) coherence networks from the post interviews in **Figure 7**.

Table 5. Initial student answers to the four scenarios from a below median coherence student and above median coherence student (based on post interviews). Usage of “Transfer” ideas printed bold

Scenario	Below median coherence student answers	Above median coherence student answers
Spinning cup	Well, when someone put the sterno candle under the cup, the cup started spinning Spinning cup because the heat transferred energy to the cup and the air passes through the holes. They are in a certain direction so that's how it spins.	Well, that's a chemical reaction. Burning is always chemical reaction and then the energy transfers from the flame through the air molecules into the cup and then the air molecules go out the hole in the cup and it causes the cup to get a speed increase and spin.
Bouncing ball	The person drops the ball and it bounces back up but every time it bounces back up it loses more and more energy to the field .	When he has the ball up here the gravitational field has the energy stored and then he drops it, it gives it back but it can't go back higher because the... isn't something to do with how the gravitational field, like it, each time it goes the surroundings so it's losing just a tad bit of energy each time to the surroundings so it can't bounce back to the same height. When he drops the ball the energy goes from the field to the ball and some of it goes to like the air.
Pushing a barrel up a ramp	Well, the person is pushing it and it's spinning and going up.	Well you could say the kid gives the barrel energy by pushing on it and then the speed increase for the barrel.
Electric heater	The person turned on the knob and it got hotter because there's a cord that's plugged into an outlet and that outlet gives the machine, it gives it fuel to be able to be heated up when it turned on.	Well there's like a little box some I'm guessing energy transfers from inside the box to the each one of those wires and then the heat waves cause it to heat up because there's no such thing as warm energy or cold energy. So the energy transfers to the ... whatever those are and then they start to heat up more and more.

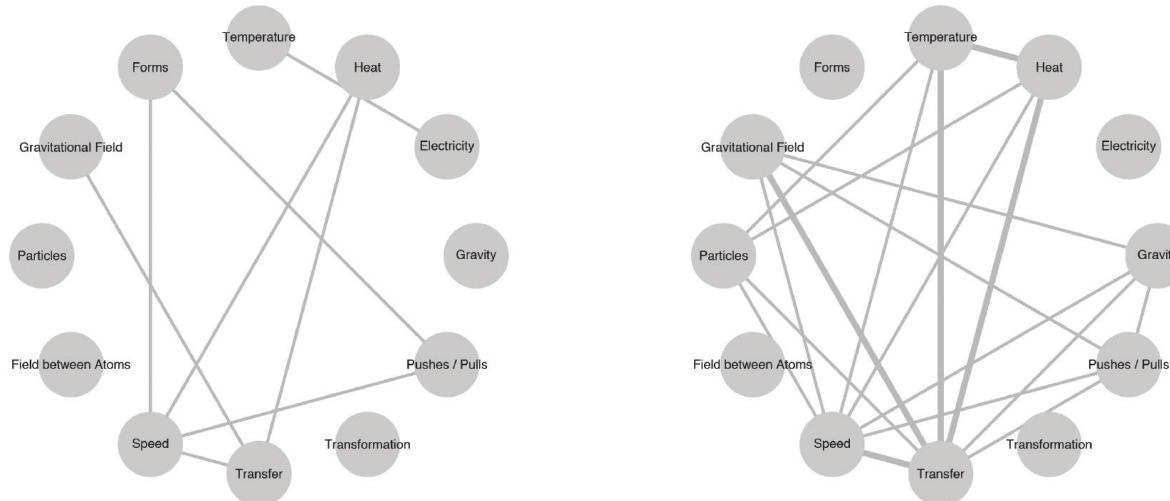


Figure 7. Post-networks of individual students, low coherence left, high coherence right

The above median coherence student uses the idea of energy transfer across all four scenarios while the below median coherence student uses the idea only once. The above median student also uses a broader range of ideas together with energy transfer, e.g., in the explanation of the bouncing ball, the high coherence student elaborates in a more detailed fashion how energy is transferred away from the bouncing ball. Thus, the coherence measure reflects differences in the extent to which students connect a set of core ideas across phenomena.

In sum, we found a strong relationship between the change in students' knowledge-in-use and the change in students' network coherence. We found a similar relationship using a measure of knowledge integration based on Lee and Liu (2010). Together with the fact that differences in network coherence between students are reflected in the extent to which they consistently use ideas across phenomena, this lends validity to the network analytical measure of coherence.

RQ2

Figure 8 shows an individual student's networks for pre and post whose gain from pre to post on the knowledge-in-use assessments is representative for the average of the interview sample.

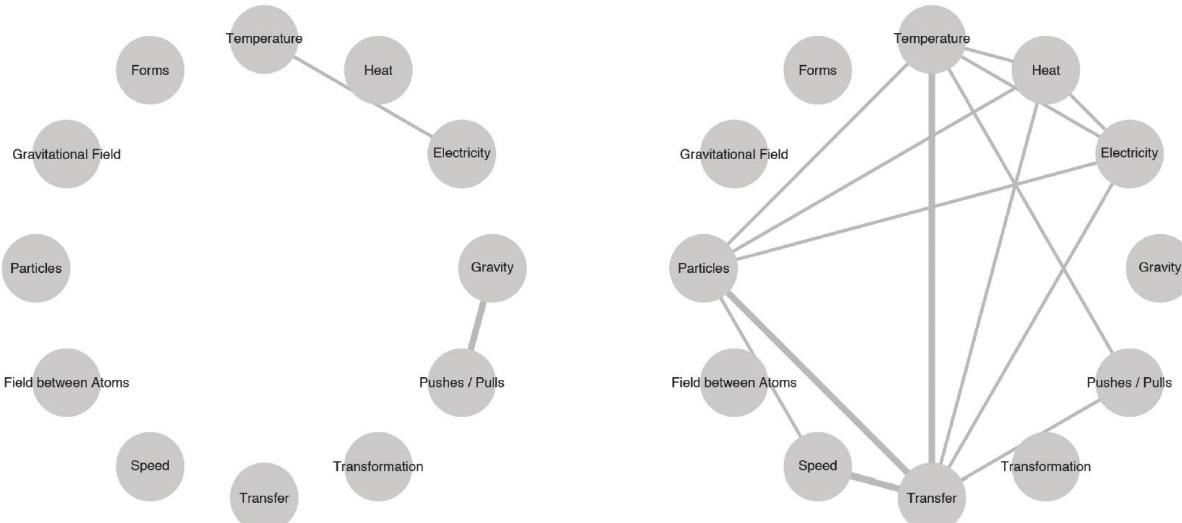


Figure 8. Pre (left) and post (right) network of an individual student

Visual inspection readily reveals differences between the two networks. Prior to instruction, the student only connects "Electricity" with "Temperature" and "Gravity" with "Pushes / Pulls". The first connection comes from the electric heater scenario, "*I see like electricity flowing through like the wires, well, I guess like the things to heat up a room. I don't ... I don't really know*". The second is made in the bouncing ball scenario, "*Uhm, I would explain it like uhm the kid's putting pressure on it to keep it up even if the gravity wants to ... to pull it back down*." The third connection comes in the person pushing a barrel up a ramp scenario, "*Eh, because the friction, I guess, is pulling it down. Yeah, and like gravity*.". The pre-network corresponds to a network coherence of 0.09. The post-network, however, shows more connections between a broader range of ideas, corresponding to a network coherence of 0.46. "Transfer" is the best-connected idea in the network with a degree of six and consistently used across scenarios to connect other ideas. In the electric heater scenario "Electricity" is now connected to the idea of "Temperature" via "Transfer": „*I saw... uh, heat like, well, no, I don't really see heat, but like the strings I guess you call it I don't know, uh they heated up once he turned the dial and mostly they do that because there's like probably a battery inside of it and once you turn that ... energy transfers to those and then those transfer like heat out, I don't know*“. Similarly, the student consistently uses "Transfer" across the person pushing a barrel up a ramp and the bouncing ball scenario and connects it to a range of other ideas such as "Speed" or "Temperature": "*Uh the boy is transferring energy to the barrel to make it go up. I saw the guy drop the basketball onto the ground and then it not bounce back up not to its original height and that happens because energy is being transferred to the ground. As it falls the ball hits that and it transfers energy to the ground and the surroundings and then it makes the temperature increase*." The change in how the student only connects few ideas before the unit and connects various ideas and consistently uses "Transfer" across scenarios after the unit is captured in the notable increase in integration as measured by network coherence (0.09 to 0.46) and degree of the "Transfer" idea (0 to 6).

Aggregated Development. The change in students' knowledge structure over the course of the unit is most vividly apparent in **Figure 9** which shows aggregated networks for pre-and post.

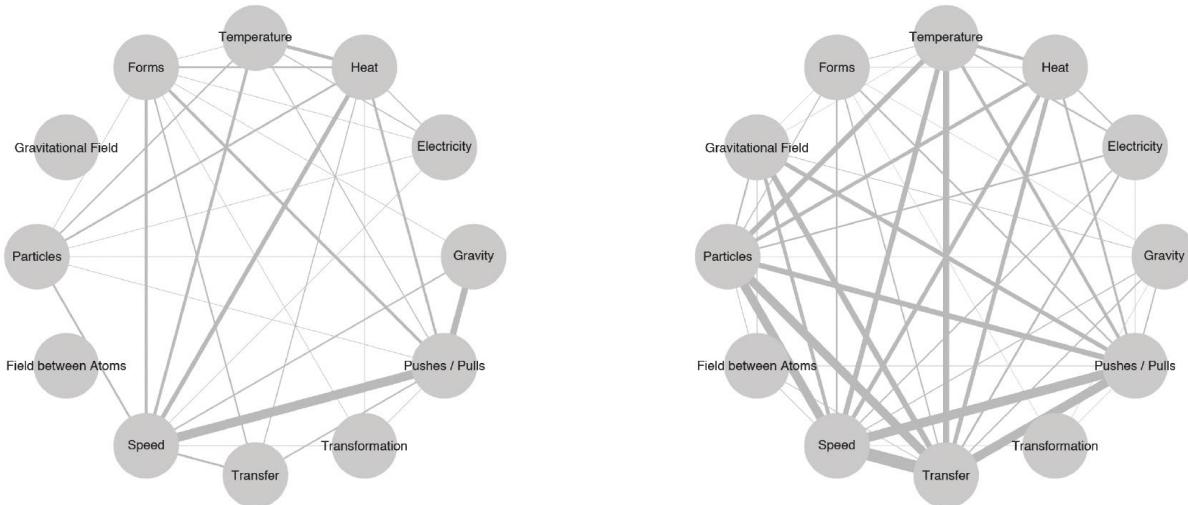


Figure 9. Aggregated network representation of co-occurrences of student ideas, pre-(left) and post-(right). Line thickness represents number of co-occurrences

Table 6. Standardized regression coefficient β^2 , p -value, variance explained R^2 , F-test, and information criterion AIC for five regression models in which students' learning gain on the energy pre/post-test was predicted by the gain in degree of the respective idea listed in the table

Idea	β^2	p	R^2	AIC
Transfer	0.49	0.046	0.24	48.57
Pushes / Pulls	0.37	0.150	0.13	50.76
Speed	0.44	0.080	0.19	49.63
Forms	0.39	0.125	0.15	50.46
Particles	0.19	0.476	0.03	52.62

Although no individual student network is identical to these aggregated networks, they reflect the central tendencies: we see a strong increase in degree (the number of other ideas an idea is connected to) for "Transfer", "Gravitational Field", and "Field between Atoms" – all ideas emphasized in the unit. Further, "Transfer" has the strongest degree in the post-network. Lastly, "Forms" and "Transformation" - energy ideas that are not emphasized in the unit – do not change in degree and have relatively low degree. The diagrams in **Figure 9** suggest that during instruction, students developed more well-connected ideas about energy and that the concept of transfer gained a more prominent role as an idea that students activated.

Connecting Quantitative Features of Students' Knowledge Networks to Students' Knowledge-in-Use. In the background section, we have argued that the degree measure allows to determine the relative importance of ideas in students' knowledge structures, i.e., the ability to identify linking-ideas empirically. To do so, we considered the top five ideas with highest degree in students aggregated post-instruction networks, i.e., the best-connected ideas, and investigated to what extent an increase in connections of these ideas from pre to post (as it should occur during learning) predicted student learning. The results of the five regression models are presented in **Table 6**.

We found the standardized regression coefficient of degree gain for "Transfer" to be the largest (0.49). Only the effect of "Speed" comes close which, however, is only approaching statistical significance, explains less variance, and has a higher AIC. "Pushes / Pulls" and "Forms" which together with "Transfer" would be natural candidates for linking-ideas from a theoretical perspective, explain about 10% less variance than "Transfer" and have higher AIC. Further, their effects are about 0.1 smaller and not statistical significant. In general, evaluating the models by their ability to explain the data R^2 and accuracy AIC, "Transfer" is the best predictor as it has the highest R^2 and lowest AIC of the five. This means that the increase in connections between energy transfer and other ideas in students' idea networks is more strongly related to students' performance on a distal knowledge-in-use assessment than the increase between other high-degree ideas and other ideas.

In sum, visualizing the networks gives qualitative insights into students' knowledge networks. In addition, derived quantitative measures allow to identify core ideas in students' knowledge networks.

DISCUSSION

The premise of the many current standards documents emphasis on a small set of powerful science ideas, such as the US *Framework for K-12 Science Education* (National Research Council, 2012) or the German *Bildungsstandards* (Sekretariat der ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2004), is that having a well-developed knowledge base organized around major ideas of a domain is the basis of scientific literacy (Bransford, 2000). Further, emphasizing these major ideas (e.g., "disciplinary core ideas" in the NGSS and "Basiskonzepte" (basic concepts) in the *Bildungsstandards*) is supposed to help students develop such organized knowledge. In the domain of energy for example, increasingly connected knowledge networks around central energy ideas should be positively associated with students' ability to make sense of phenomena using energy ideas (Bransford, 2000; Chi et al., 1981; Linn, 2006). We have developed a ground-up approach that provides qualitative and quantitative information about students' ability to activate and connect sets of relevant ideas as they make sense of phenomena across contexts, i.e., students' integrated knowledge. We found that students linked ideas around the central idea of energy transfer (which serves as a disciplinary core-idea in middle school and was emphasized in the energy unit in which the students participated) and that students that did so to a larger extent, performed better on a knowledge-in-use assessment in which students had to connect scientific practices and energy ideas. This supports the premise of the *Framework for K-12 Science Education* or the *Bildungsstandards* of deep learning around a small set of core ideas. In addition, our results support theoretical work that suggests that energy transfer can be a particularly helpful energy idea to make sense of phenomena (Brewe, 2011; Ellse, 1988; Nordine et al., 2018; Quinn, 2014; Swackhamer, 2005).

Current approaches to measuring how students connect ideas often (e.g., Kauertz and Fischer (2006) or Lee and Liu (2010)) define linking-ideas a priori. In general, as our approach is more open in the sense that it does not rely on a priori established linking-ideas, it enables a more full and flexible description of students' ideas without filtering through a pre-defined rubric (Steedle & Shavelson, 2009). This appears especially valuable when it is not really clear or at least contested which ideas should be considered central and how students' ideas develop. For example, in domains with multiple central ideas such as energy it is not necessarily clear which possible linking ideas (energy transfer, energy transformation) are most productive (Papadouris & Constantinou, 2016). The degree measure can help to resolve such issues.

Further, our results extend work that has linked how students connect ideas with their performance on related MC assessments (Kauertz & Fischer, 2006; Lee & Liu, 2010; Nordine et al., 2011) as we used knowledge-in-use assessments in which students had to connect disciplinary ideas with scientific practices and cross-cutting concepts. In general, the relationship of scientific practices, disciplinary core ideas, and cross-cutting concepts is not well defined when those components are integrated in knowledge-in-use performances. Our results provide a first step towards better understanding the role of well-organized disciplinary knowledge for knowledge-in-use, as the coherence measure that is insensitive to elements of a scientific practice, explained a large share of the variance in students' knowledge-in-use.

Limitations

The network analysis method we present is built on the assumption that we consider ideas connected if they co-occur. Thus, we do not differentiate between students that meaningfully connect e.g., "force" with "friction" and those that use both ideas within the same explanatory account in isolation. However, the premise of the present approach is probabilistic as it assumes that given a sufficient number of phenomena, ideas that co-occur repeatedly, do so because they are connected meaningfully. Network coherence as a measure of integrated knowledge does not describe a single students' explanation of a phenomenon but describes a student's idea network given multiple explanations of phenomena from the student. Further, we ignore the order in which ideas are presented. Thus, we cannot differentiate between students that start their explanatory account with deep conceptual ideas such "transfer" and then use surface level ideas such as "speed" and those that start with surface level ideas and then go on to more conceptual ones. While future network approaches could at least in principle consider the order in which ideas are presented, it is also questionable to what extent order is a relevant feature of idea networks and not feature of language or other cognitive processes, given the prominent role of parallel processing theories in cognitive science (Rumelhart & McClelland, 1986).

In sum, we consider that interpreting co-occurrence as connection and ignoring order are helpful in balancing the simplicity of assumptions and their effectiveness in practice which these two assumptions have already demonstrated in the contexts of automated semantic analysis (Blei, 2012; Landauer, 2014; Zehner, Sälzer, & Goldhammer, 2016).

Implications and Future Directions

Our network analytical approach provides qualitative and quantitative information about students' integrated knowledge which can be linked to students' knowledge-in-use while not relying on a priori defined linking ideas. Thus, it allows researchers fuller access to the complexity of conceptual development and growth in understanding over time and addresses the problem, that developmental pathways may remain hidden due to the filtering through pre-defined learning progressions (Duncan & Rivet, 2013). Given the promising results we presented, we consider future research that uses the approach and helps to further develop and refine it warranted.

Our network analytical approach is build ground-up from the ideas that students activate. This should not only help to avoid filtering through pre-defined learning progressions but also help to reveal intermediary stages in the developments of students' knowledge networks where students might only be able to activate ideas, but not to connect them. This finer grain size could for example inform the discussion in the energy literature about the influence of developmental stages for energy learning progressions (X. Liu & McKeough, 2005).

We focused on the normative part of students' knowledge networks. Future research could extend the approach to include non-normative ideas to better understand the role that they play in students' knowledge networks and how they develop over time. In the domain of energy for example, there is a rich literature on student conceptions of energy (Duit, 2014; Watts, 1983) but the extent to which they impair or may even enhance (as productive intermediary stages) how students develop well-organized knowledge networks is little researched.

Further, what we have sketched here for the domain of energy is in principle applicable to other domains and core-ideas like force or evolution in biology.

Apart from such research oriented outcomes, our network analytical approach has the potential to provide rich diagnostic information for teachers and curriculum developers that current assessments do seldomly provide (Duncan & Hmelo-Silver, 2009). However, to be applicable in wide spread practice, one would have to find a way to scale the approach, possibly using automation techniques inspired by the field of learning analytics.

ACKNOWLEDGEMENTS

This work was supported by the NSF (Grant Number DUE-1431725).

ENDNOTES

During the analysis of the interviews we found that students were often confused by the melting ice phenomenon. Interviewers reacted by starting to explain the experimental setup in the video. However, during this they often included leading prompts. In consequence, the melting ice phenomenon was excluded from further analysis.

REFERENCES

- Anderson, J. R. (1983). Cognitive science series. The architecture of cognition. Hillsdale, NJ, US.
- Anderson, J. R., & Schunn, C. (2000). Implications of the ACT-R learning theory: No magic bullets. *Advances in Instructional Psychology, Educational Design and Cognitive Science*, 1-33.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77. <https://doi.org/10.1145/2133806.2133826>
- Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of Science. *The Annals of Applied Statistics*, 1(1), 17-35. <https://doi.org/10.1214/07-AOAS114>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: fundamental measurement in the human sciences* (Third edition). New York ; London: Routledge, Taylor and Francis Group.
- Bransford, J. (2000). *How people learn: brain, mind, experience, and school*. (National Research Council (U.S.), Ed.) (Expanded ed). Washington, D.C: National Academy Press.
- Brewe, E. (2011). Energy as a substancelike quantity that flows: Theoretical considerations and pedagogical consequences. *Physical Review Special Topics - Physics Education Research*, 7(2), 020106. <https://doi.org/10.1103/PhysRevSTPER.7.020106>
- Chabalengula, V. M., Sanders, M., & Mumba, F. (2012). Diagnosing Students' Understanding Of Energy And Its Related Concepts In Biological Contexts. *International Journal of Science and Mathematics Education*, 10(2), 241-266. <https://doi.org/10.1007/s10763-011-9291-2>

- Chen, R. F., Eisenkraft, A., Fortus, D., Krajcik, J., Neumann, K., Nordine, J., & Scheff, A. (2014). *Teaching and learning of energy in K-12 education*. Cham: Springer. Retrieved from <http://gso.gbv.de/DB=2.1/PPNSET?PPN=783976232>
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices*. *Cognitive Science*, 5(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695(5), 1-9. <http://igraph.sf.net>
- Derry, S. J. (1996). Cognitive schema theory in the constructivist debate. *Educational Psychologist*, 31(3–4), 163–174. <https://doi.org/10.1080/00461520.1996.9653264>
- diSessa, A. A. (1988). Knowledge in Pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the Computer Age* (pp.49-70). Hillsdale, NJ: Lawrence Erlbaum.
- diSessa, A. A. (2013). A bird's-eye view of the "pieces" vs "coherence" controversy (from the "pieces" side of the fence). In Stella Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 31–48). New York, NY: Routledge.
- diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education*, 20(10), 1155–1191. <https://doi.org/10.1080/0950069980201002>
- Driver, R., & Warrington, L. (1985). Students' Use of the Principle of Energy Conservation in Problem Situations. *Physics Education*, 20(4), 171–176.
- Duit, R. (2014). Teaching and Learning the Physics Energy Concept. In Chen, R.F., Eisenkraft, A., Fortus, D., Krajcik, J., Neumann, K., Nordine, J., and Scheff, A. (Eds.), *Teaching and Learning of Energy in K-12 Education* (pp. 67–85). Cham: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-05017-1_5
- 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. <https://doi.org/10.1002/tea.20316>
- Duncan, R. G., & Rivet, A. E. (2013). Science Learning Progressions. *Science*, 339(6118), 396–397. <https://doi.org/10.1126/science.1228692>
- Ellse, M. (1988). Transferring Not Transforming Energy. *School Science Review*, 69(248), 427–437.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239.
- Gilbert, J. K., Watts, D. M., & Osborne, R. J. (1982). Students' conceptions of ideas in mechanics. *Physics Education*, 17(2), 62–66. <https://doi.org/10.1088/0031-9120/17/2/309>
- Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition.
- Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & McElhaney, K. W. (2016). *Constructing assessment tasks that blend disciplinary core Ideas, crosscutting concepts, and science practices for classroom formative applications*. Menlo Park, CA.
- Hmelo-Silver, C., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, 28(1), 127–138. [https://doi.org/10.1016/S0364-0213\(03\)00065-X](https://doi.org/10.1016/S0364-0213(03)00065-X)
- Kauertz, A., & Fischer, H. E. (2006). Assessing students' level of knowledge and analysing the reasons for learning difficulties in physics by Rasch analysis. *Applications of Rasch Measurement in Science Education*, 212–246.
- Koponen, I. T., & Huttunen, L. (2013). Concept Development in Learning Physics: The Case of Electric Current and Voltage Revisited. *Science and Education*, 22(9), 2227–2254. <https://doi.org/10.1007/s11191-012-9508-y>
- Landauer, T. K. (2014). *Handbook of latent semantic analysis*. New York: Routledge.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Lee, H.-S., & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades. *Science Education*, 94(4), 665–688. <https://doi.org/10.1002/sce.20382>
- Lee, H.-S., Liu, O. L., & Linn, M. C. (2011). Validating Measurement of Knowledge Integration in Science Using Multiple-Choice and Explanation Items. *Applied Measurement in Education*, 24(2), 115–136. <https://doi.org/10.1080/08957347.2011.554604>
- Linn, M. C. (2006). The Knowledge Integration Perspective on Learning and Instruction. In *The Cambridge handbook of: The learning sciences*. New York, NY: Cambridge University Press.
- Liu, O. L., Lee, H.-S., Hofstetter, C., & Linn, M. (2008). Assessing Knowledge Integration in Science: Construct, Measures, and Evidence. *Educational Assessment*, 13(1), 33–55. <https://doi.org/10.1080/10627190801968224>

- Liu, O. L., Ryoo, K., Linn, M. C., Sato, E., & Sviha, V. (2015). Measuring Knowledge Integration Learning of Energy Topics: A two-year longitudinal study. *International Journal of Science Education*, 37(7), 1044-1066. <https://doi.org/10.1080/09500693.2015.1016470>
- Liu, X., & McKeough, A. (2005). Developmental growth in students' concept of energy. *Journal of Research in Science Teaching*, 42(5), 493-517. <https://doi.org/10.1002/tea.20060>
- Loh, A. S. L., & Subramaniam, R. (2018). Mapping the knowledge structure exhibited by a cohort of students based on their understanding of how a galvanic cell produces energy. *Journal of Research in Science Teaching*. <https://doi.org/10.1002/tea.21439>
- Mayring, P. (2014). *Qualitative Content Analysis*. Beltz.
- McClelland, J. L., & Cleeremans, A. (2009). Connectionist Models. In T. Byrne, Axel Cleeremans, and P. Wilken (Eds.), *Oxford Companion to Consciousness*. New York: Oxford University Press.
- National Academies of Sciences, Engineering, and Medicine. (2018). *How People Learn II: Learners, Contexts, and Cultures*. Washington, D.C.: National Academies Press. <https://doi.org/10.17226/24783>
- National Research Council. (2012). *A framework for K-12 science education*. Washington, D.C.: The National Academies Press. Retrieved from <http://www.worldcat.org/oclc/794415367>
- Neumann, K., Kubisch, M., Nordine, J., Fortus, D., & Krajcik, J. (2018). Assessing students' progression in developing a deeper understanding of energy. Paper presented at NARST 2018 national conference. Atlanta.
- Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, 50(2), 162-188. <https://doi.org/10.1002/tea.21061>
- NGSS Lead States. (2013). *Next generation science standards*. Washington DC: National Acad. Press.
- Nordine, J., Fortus, D., Krajcik, J., Neumann, K., & Lehavi, Y. (2018). *Modelling Energy Transfers between Systems to Support Energy Knowledge in Use*. Manuscript submitted for publication.
- 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. <https://doi.org/10.1002/sce.20423>
- Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, 27(10), 937-949. <https://doi.org/10.1002/tea.3660271003>
- Opsahl, T. (2009). *Structure and evolution of weighted networks* (PhD Thesis). Queen Mary, University of London.
- Osborne, R. J., & Gilbert, J. K. (1980). A technique for exploring students' views of the world. *Physics Education*, 15(6), 376.
- Papadouris, N., & Constantinou, C. P. (2016). Investigating middle school students' ability to develop energy as a framework for analyzing simple physical phenomena. *Journal of Research in Science Teaching*, 53(1), 119-145. <https://doi.org/10.1002/tea.21248>
- Park, H.-J., & Friston, K. (2013). Structural and Functional Brain Networks: From Connections to Cognition. *Science*, 342(6158), 1238411-1238411. <https://doi.org/10.1126/science.1238411>
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2004). *Knowing what Students Know* (3. print). Washington, DC: National Acad. Press. Retrieved from <http://gso.gbv.de/DB=2.1/PPNSET?PPN=487618513>
- Pellegrino, J. W., Wilson, M. R., Koenig, J. A., Beatty, A. S., National Research Council (U.S.) (Eds.). (2014). *Developing assessments for the Next Generation Science Standards*. Washington, D.C: The National Academies Press.
- Quinn, H. R. (2014). A Physicist's Musings on Teaching About Energy, In Chen, R.F., Eisenkraft, A., Fortus, D., Krajcik, J., Neumann, K., Nordine, J., and Scheff, A. (Eds.), *Teaching and Learning of Energy in K-12 Education*. Cham: Springer. https://doi.org/10.1007/978-3-319-05017-1_2
- R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. *Scientometrics*, 82(2), 263-287. <https://doi.org/10.1007/s11192-009-0041-y>
- Ruiz-Primo, M. A. (2004). Examining concept maps as an assessment tool. Proceedings of the First International Conference on Concept Mapping. Pamplona, Spain.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: explorations in the microstructure of cognition*. Cambridge, Mass: MIT Press.
- Schneider, M., & Stern, E. (2009). The Inverse Relation of Addition and Subtraction: A Knowledge Integration Perspective. *Mathematical Thinking and Learning*, 11(1-2), 92-101. <https://doi.org/10.1080/10986060802584012>

- Schwartz, D. L., & Arena, D. (2013). Measuring What Matters Most, 192.
- Sekretariat der ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). *Bildungsstandards Physik-Mittlerer Schulabschluss*.
- Smith III, J. P., diSessa, A., & Roschelle, J. (1994). Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *Journal of the Learning Sciences*, 3(2), 115–163. https://doi.org/10.1207/s15327809jls0302_1
- Steedle, J. T., & Shavelson, R. J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching*, 46(6), 699–715. <https://doi.org/10.1002/tea.20308>
- Swackhamer, G. (2005). *Cognitive Resources for Understanding Energy*.
- Swackhamer, G., & Hestenes, D. (2005). An energy concept inventory. *Arizona State University*.
- Thagard, P. (2000). *Coherence in thought and action*. Cambridge, Mass.: MIT Press.
- Watts, D. M. (1983). Some alternative views of energy. *Physics Education*, 18(5), 213. <https://doi.org/10.1088/0031-9120/18/5/307>
- Won, M., Krabbe, H., Ley, S. L., Treagust, D. F., & Fischer, H. E. (2017). Science Teachers' Use of a Concept Map Marking Guide as a Formative Assessment Tool for the Concept of Energy. *Educational Assessment*, 22(2), 95–110. <https://doi.org/10.1080/10627197.2017.1309277>
- Zehner, F., Sälzer, C., & Goldhammer, F. (2016). Automatic Coding of Short Text Responses via Clustering in Educational Assessment. *Educational and Psychological Measurement*, 76(2), 280–303. <https://doi.org/10.1177/0013164415590022>

<http://www.ejmste.com>