Chapter 4
Intended Learning Outcomes and Assessment of Computer-Based Scaffolding

Abstract In this chapter, I describe the intended learning outcomes of scaffolding—content knowledge and higher-order thinking abilities—and link these to the goals advanced by the Next Generation Science Standards and related documents from recent curricular revisions in STEM education. Furthermore, I address different ways in which scaffolding's effect can be measured (assessment level), and explore whether there are differences in the magnitude of scaffolding's effect according to assessment level. Meta-analysis results show that there is no difference in effect size magnitude on the basis of intended learning outcome (i.e., content knowledge or higher-order thinking abilities). Scaffolding's effect was greater when measured at the principles level than when measured at the concept level. But scaffolding's effect was statistically greater than 0 and substantial for all three assessment levels (i.e., concept, principles, and application). These results are then discussed.

Keywords Application-level assessment · Argumentation · Assessment levels · Common Core · Concept-level assessment · Epistemology · Intended learning outcomes · Next Generation Science Standards · Principles-level assessment · STEM education

4.1 Rationale for this Chapter

In science, technology, engineering, and mathematics (STEM) education, computer-based scaffolding has been deployed to help enhance students' higher-order thinking skills (Belland, 2010; Cho & Jonassen, 2002; Eck & Dempsey, 2002; M. Kim & Hannafin, 2011) and deep content learning (Chang & Linn, 2013; Davis, 2003; Hwang, Shi, & Chu, 2011). These diverse learning outcomes may be seen by some as evidence of two categorically different interventions that cannot be considered alongside each other. But these dual emphases of scaffolding can be seen as congruent with the emphases on learning the process of STEM, as well as learning cross-cutting concepts and disciplinary core ideas in the Next Generation Science Standards (NGSS; Achieve, 2013; National Science Board, 2010). Needless to say, scaffolding's emphases did not emerge in direct response to the writing of the NGSS, as such emphases were formed well before the NGSS existed. Rather, scaffolding's intended learning outcomes arose within and alongside the currents of

B. R. Belland, *Instructional Scaffolding in STEM Education,*
DOI 10.1007/978-3-319-02565-0_4

the transformation of education from a didactic process of information transfer to one of construction of knowledge.

In this chapter, to provide context and to help the reader understand the seeming dichotomy of learning goals of scaffolding, I first situate scaffolding relative to the calls for the enhancement of content knowledge and higher-order thinking skills in the NGSS (Achieve, 2013; Krajcik, Codere, Dahsah, Bayer, & Mun, 2014; National Science Board, 2010) and the Common Core State Standards (McLaughlin & Overturf, 2012; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

Second, I expand on the intended learning outcomes of scaffolding. Variation in intended learning outcomes of scaffolding largely aligns with differences in the theoretical underpinnings of scaffolding, which were discussed in Chap. 2: "Instructional Scaffolding: Foundations and Evolving Definition." I also explore if the effectiveness of scaffolding varies according to intended learning outcome, as informed by the meta-analysis results.

Just as it is important to consider intended learning outcomes, it is also important to consider how learning is assessed (Belland, 2012; Belland, French, & Ertmer, 2009; Furtak & Ruiz-Primo, 2008; Messick, 1989). Indeed, one is often advised to consider assessment before even designing objectives and instructional materials/strategies (Gagné, 1965; Wiggins & McTighe, 2005). By considering how scaffolding's influence on cognitive outcomes varies according to how it is assessed—at the concept, principles, or application level (Sugrue, 1995)—one can see if scaffolding as a whole delivers stronger impacts on content learning or various types of higher-order thinking skills. It is important to consider this alongside the intended learning outcome, as (a) just because an intervention is designed to increase content learning or higher-order skills does not necessarily mean that it does, and (b) just because scholars claim that scaffolding is intended to help students enhance their skill in a particular area does not always mean that the learning is being assessed at that level.

In this chapter, I discuss these ideas, and present meta-analysis results comparing scaffolding's impact according to intended learning outcome and assessment levels.

4.2 Targeted Learning Outcomes of Scaffolding

Scaffolding has been designed to promote higher-order skills such as ill-structured problem-solving ability (Ge & Land, 2004; Liu & Bera, 2005) and argumentation ability (Belland, Gu, Armbrust, & Cook, 2015; McNeill & Krajcik, 2009), and enhanced/deep content knowledge (Davis & Linn, 2000; Koedinger & Corbett, 2006). It is important to note that in the intelligent tutoring systems literature, authors posit a focus on enhancing procedural knowledge (production rules) by which individuals can apply declarative knowledge. Some may argue that this is a form of problem-solving skill. But I argue that it is a form of content learning, as each production rule is concerned with how to apply one highly specific domain knowledge element (Anderson et al., 2004).

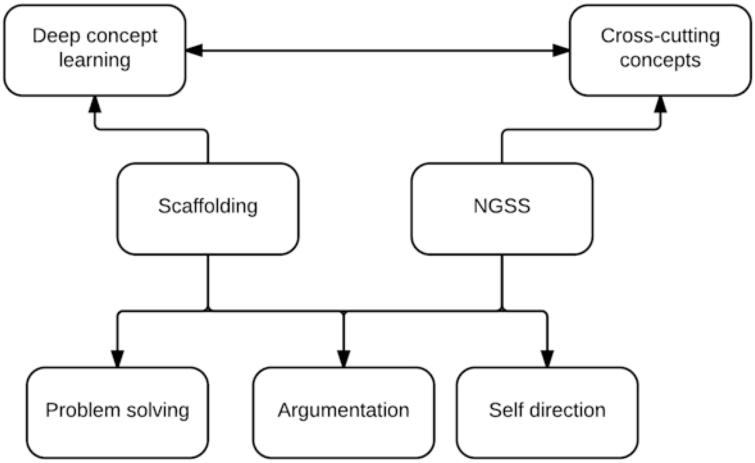

Fig. 4.1 The relationship between the intended learning outcomes of instructional scaffolding and of the Next Generation Science Standards (Achieve, 2013)

The interrelationship between the intended learning outcomes of scaffolding and of the NGSS are illustrated in Fig. 4.1 and expanded upon in the sections that follow.

4.2.1 Higher-Order Thinking Skills

4.2.1.1 Ill-Structured Problem-Solving Ability

Scaffolding to promote problem-solving ability is closest to the original instructional scaffolding definition (Wood, Bruner, & Ross, 1976). Problem-solving ability in this case refers to the ability to solve ill-structured problems—problems with many possible valid solutions and many valid solution paths (Jonassen, 2000, 2011).

To be successful solving ill-structured problems, learners need to qualitatively model such problems so that they can determine what entities interact in the problem, how they interact, and what such interaction means to each entity (Chi, Feltovich, & Glaser, 1981; Jonassen, 2003; Klahr & Simon, 1999; Lesh & Harel, 2003; Nersessian, 2008). But then they need to characterize the disparity between the goal state and the current state and determine an appropriate way to bridge the gap (Jonassen, 2000). However, this process is different from the means-ends analysis that describes how people often solve well-structured problems. Rather, solving ill-structured problems is an iterative process of defining the problem and identifying and weighing potential goal states and different methods of arriving at those goal states (Chi et al., 1981; Giere, 1990; Jonassen, 2000, 2003; Nersessian, 2008). By definition, ill-structured problems often have many solutions that are equally valid (Jonassen, 2011). In this way, the suitability of solutions to ill-structured problems needs to be judged on the basis of evidential support (Belland, Glazewski, & Richardson, 2008; Ford, 2012; Jonassen & Kim, 2010). Students thus need to have the opportunity to build and evaluate evidence-based arguments to be able to engage in ill-structured problem-solving, and to prepare for the modern workforce (Ford, 2012; Gu & Belland, 2015; Jonassen, 2011; Osborne, 2010; Perelman & Olbrechts-Tyteca, 1958).

The ability to solve ill-structured problems is qualitatively different from solving well-structured problems such as the story problems found in many mathematics textbooks (Jonassen, 2000; Lesh & Harel, 2003; Nersessian, 2008). One can solve well-structured problems with only the information given in the problem description, whereas solving ill-structured problems requires the acquisition, evaluation, and use of much data beyond that given in the problem description. Well-structured problems have only one correct answer, and often only one solution path, whereas ill-structured problems have multiple potentially correct solutions, and many ways of arriving at them. Given these differences, the strategies by which one addresses well-structured problems and ill-structured problems differ (Jonassen, 2000). As such, one cannot promote the enhancement of ill-structured problem-solving ability by engaging students in well-structured problem-solving; rather, one should engage students in ill-structured problem-solving along with instructional support such as scaffolding (Abd-El-Khalick et al., 2004; Jonassen, 2011).

4.2.1.2 Argumentation Ability

Argumentation ability refers to the ability to back claims with evidence by way of premises, and evaluate and respond to the extent to which claims presented by others are well supported by evidence (D. Kuhn, 1991; Perelman & Olbrechts-Tyteca, 1958; van Eemeren, Grootendorst, & Snoeck Henkemans, 2002). Two prominent models of argumentation are those of persuasive argumentation and dialectical argumentation. According to the former, there is no such thing as a universally valid argument; rather, arguments are successful to the extent to which the audience agrees with its central claim. As such, the goal of argumentation is to lead the audience to adhere to the validity of one's claim (Perelman & Olbrechts-Tyteca, 1958; Walton, 1989). In persuasive argumentation, novice arguers often focus on strengthening one's own position (D. Kuhn, 1991; Vellom & Anderson, 1999).

Dialectical argumentation starts off with individuals creating evidence-based arguments, but from there it diverges. Specifically, rather than simply supporting one's own claims, in dialectical argumentation, one also engages with claims of others (Asterhan & Schwarz, 2009; Jonassen & Kim, 2010; Keefer, Zeitz, & Resnick, 2000). This can include attempting to weaken the position of others (Asterhan & Schwarz, 2009; D. Kuhn, 1991) or negotiating with opposing parties in pursuit of an ultimate truth (Jonassen & Kim, 2010; Keefer et al., 2000; van Eemeren & Houtlosser, 2001). In the latter case, the opposing parties make concessions in their arguments in the service of improving their claims and ultimately moving toward an ultimate truth that is not directly knowable, but which can be approached through negotiation of arguments.

Argumentation can be considered a subset of problem-solving ability (Jonassen & Kim, 2010; D. Kuhn, 1991), and is the process by which scientific knowledge advances (Ford, 2012; Osborne, 2010). As discussed earlier, argumentation is core to how the quality of solutions to ill-structured problems is judged. Having arrived at initial solutions to such problems, argumentation is also how such solutions are iteratively improved, as well as the evidential support for the solutions (Ford, 2012; Osborne, 2010). K-12 (Belland et al., 2008; Driver, Newton, & Osborne, 2000; Glassner, Weinstock, & Neuman, 2005; McNeill & Pimentel, 2010) and college students (Abi-El-Mona & Abd-El-Khalick, 2011; Cho & Jonassen, 2002; Uskola, Maguregi, & Jiménez-Aleixandre, 2010) often struggle with argumentation, and thus it is important to help them learn this skill. But rather than teaching such didactically, it is important to put them in a situation about which to argue (Aufschnaiter, Erduran, Osborne, & Simon, 2008; Belland et al., 2008; Driver et al., 2000; Jonassen & Kim, 2010) and support them with such tools as scaffolding (Belland et al., 2008; Cho & Jonassen, 2002; Clark & Sampson, 2007; Nussbaum, 2002).

4.2.1.3 Self-Directed Learning Ability

Self-directed learning refers to the ability to identify learning issues, plan and execute a strategy to address the learning issues, and evaluate the quality with which the learning issues were addressed; in other words, it is the ability to identify and regulate one's pursuit of learning issues (Bolhuis, 2003; Loyens, Magda, & Rikers, 2008). Being able to do so is central to addressing ill-structured problems (Giere, 1990; Jonassen, 2011; Nersessian, 2008), and thus is an important skill to support to facilitate student success in problem-centered approaches to instruction (Lohman & Finkelstein, 2000; Loyens et al., 2008; Merriënboer & Sluijsmans, 2008).

Identifying learning issues to be addressed requires that learners assess what information is needed to address the problem, and what among the needed knowledge is a knowledge deficiency—either not present in the problem presentation or part of their preexisting knowledge (Hmelo-Silver, 2004; Loyens et al., 2008). This allows for a good deal of autonomy on the part of students in that they can define the content to be learned, which in turn can enhance student motivation (Deci & Ryan, 2000; Wijnia, Loyens, & Derous, 2011). This clearly goes beyond the traditional practice in teacher-centered classrooms in which the teacher determines what is to be learned.

Planning and executing a strategy to address learning issues requires that learners select appropriate learning resources (Hmelo-Silver, 2004; Loyens et al., 2008). The effective evaluation of the quality of sources is considered key to information literacy and solving problems, as without it, one can be lost in the vast amount of information on the web, and not be able to distinguish between credible information and non-credible information (Berzonsky & Richardson, 2008; Van de Vord, 2010). Yet, college (Berzonsky & Richardson, 2008; Van de Vord, 2010) and K-12 (Kuiper, Volman, & Terwel, 2005; Nicolaidou, Kyza, Terzian, Hadjichambis, & Kafouris, 2011; Williams, 2005) students often experience much difficulty searching for and effectively evaluating the quality of online information. For example, K-12 students often search for information in an unsystematic manner and rapidly decide if a page is usable; they then quickly search for an answer to a specific question (Kuiper et al., 2005). Furthermore, K-12 students often see all evidence as equally valid (Nicolaidou et al., 2011). Unaided college science students are often unable to distinguish between peer-reviewed sources and non-peer-reviewed sources (Berzonsky & Richardson, 2008). Students' poor ability to evaluate and use sources effectively can stem from such phenomena as conflicting information across sources, complexity of the target information and the way in which it is portrayed, and the structure that the text follows (Britt, Richter, & Rouet, 2014). Unsophisticated epistemic beliefs can cause students to struggle to distill important messages from sources and fail to question the credibility of sources (Bråten, Britt, Strømsø, & Rouet, 2011). Furthermore, students' evaluation of sources is often short-circuited by a desire for quick learning (Berzonsky & Richardson, 2008; Zimmerman, 1995), which is often experienced by students with unsophisticated epistemic beliefs (Chinn, Buckland, & Samarapungavan, 2011; Hofer & Pintrich, 1997; Qian & Alvermann, 1995). Clearly, students' struggles identifying appropriate learning issues and determining promising ways to address such present a prime opportunity to use computer-based scaffolding (Kuiper et al., 2005).

The last part of self-directed learning ability is the ability to evaluate the quality of one's own learning and learning processes, also known as metacognition (Loyens et al., 2008; Quintana, Zhang, & Krajcik, 2005; Schraw, Crippen, & Hartley, 2006). Metacognition is desirable in part to enable the smooth operation and success of a student-centered learning environment. This is because if students define and pursue their own learning issues, and different student groups in the same classroom pursue a wide variety of learning issues in a wide variety of manners, it is difficult for one teacher to provide sufficient feedback to ensure that all students are on the right track. Metacognition can work in concert with teacher feedback to provide a consistent corpus of feedback to inform the revision of learning processes as needed. Metacognition has been an important process that scaffolding seeks to support (Cuevas, Fiore, & Oser, 2002; Quintana et al., 2005).

4.2.1.4 Alignment with NGSS

The intended learning outcome of promoting higher-order thinking skills aligns with NGSS's emphasis on students learning STEM processes and engaging with the culture of STEM and with authentic STEM issues (Achieve, 2013; National Science Board, 2010), as detailed in the following sections.

4.2.1.4.1 STEM Processes

The goal of helping students learn to apply STEM processes includes helping students learn to (a) identify important problem characteristics to investigate further, (b) design strategies to investigate those problem aspects, (c) interpret appropriately data and other information collected, (d) arrive at reasonable conclusions, and (e) engage in a variety of valued scientific discourse patterns (Achieve, 2013; Duschl, 2008; National Science Board, 2010). This does not mean that all citizens need to know and be able to apply such processes at the same level as a professional chemist or engineer, but they should be able to converse with STEM processes and issues to the extent that they can make informed decisions about scientific issues that impact their local communities and nation (Duschl, 2008; Kolstø, 2001; Sadler, Barab, & Scott, 2007). Each of these subpoints is addressed in the following pages.

4.2.1.4.1.1 Identify Important Problem Characteristics to Investigate Further

One of the key processes in STEM is asking cogent questions and identifying key aspects of problems (Carr, Bennett, & Strobel, 2012; Giere, 1990; Klahr & Simon, 1999; National Research Council, 2012; Nersessian, 2008). Going into a problem with a vague goal of figuring it out is unlikely to lead to a meaningful solution (Jonassen, 2011). Rather, one needs to determine the involved variables, how they interact, and what about how they interact is problematic (Belland et al., 2008; Jonassen, 2011). This is a key scientific process, and one that does not require the asker to be a professional scientist. But it is a skill that individuals do not naturally have; rather, it needs to be developed through instruction (Jonassen, 2003). By habitually asking questions about scientific phenomena, citizens will identify key issues facing their communities, and be prepared when others present arguments and explanations about STEM-related issues in their community (Kolstø, 2001; Sadler et al., 2007; Zeidler, Sadler, Simmons, & Howes, 2005).

4.2.1.4.1.2 Design Strategies to Investigate Problem Aspects

Students need to think of scientific problems from different perspectives (Jonassen, 2011). They also need to recognize and apply the key role of iteration in addressing scientific questions (Klahr & Simon, 1999; Nersessian, 2008). Specifically, they need to understand that one cannot effectively address a scientific question with just one piece of scientific evidence. Rather, they need to collect data/reason scientifically in one way, consider the limitations of such, and design and carry out additional investigations accordingly (Carr et al., 2012; Giere, 1990; Klahr & Simon, 1999). In other words, they need to understand STEM from an epistemological standpoint—for example, that one cannot arrive at definitive answers to STEM questions by consulting just one source or conducting just one investigation (Chinn et al., 2011; Duschl, 2008; Hogan & Maglienti, 2001; Mason, Boldrin, & Ariasi, 2010; Sandoval, 2005) and that most knowledge is not certain (Bråten et al., 2011; Giere, 1990; Hofer & Pintrich, 1997). But it is not enough to simply understand this; citizens need to also be able to and be willing to apply this understanding to real STEM problems (Chinn et al., 2011; Mason & Scirica, 2006).

In designing investigations, students need to be able to apply the tools of mathematics and computation, and recognize the influence of such tools and specifically the ways in which the tools are used in the problem solution process (Lesh & Harel, 2003; National Research Council, 2012; Schoenfeld, 1985). It is important to note that applying the tools of mathematics does not simply mean setting up equations. Rather, it is important to think, at a conceptual level, about what type of data should be collected and how it will be analyzed to address the research questions (Kerlinger & Lee, 2000; Schoenfeld, 1985). This is important so that the right type of data is collected. At the same time, students need to understand that not all problem-solving strategies need to involve the use of mathematics. Rather, attempting to see where the presented problem and an idealized, qualitative model depart from each other is a viable problem-solving strategy (Nersessian, 2008).

4.2.1.4.1.3 Interpret Data and Other Information Appropriately

Students need to be able to analyze data in a systematic manner, but also realize that the job is not done until such analysis is interpreted in light of a theoretical framework (Giere, 1990; National Research Council, 2012). This is important because many individuals have the mistaken impression that scientific investigations always take place in a theoretical vacuum. To the contrary, theoretical frameworks always drive the design, conduction of, and interpretation of the results of research (Abi-El-Mona & Abd-El-Khalick, 2011; Ford, 2012; Giere, 1990; D. Kuhn, 2010). For example, theoretical frameworks can influence the choice of problems to investigate and the selection of variables on which to focus in an investigation (Lather, 2012; Miles & Huberman, 1984). Furthermore, knowing that differences in property A are statistically different between two objects means little without interpreting the finding in light of a theoretical framework. This is important both as something to do when investigating scientific phenomena, but also to remember that other scientists themselves do this when investigating scientific phenomena (Abi-El-Mona & Abd-El-Khalick, 2011; Giere, 1990).

4.2.1.4.1.4 Arrive at Reasonable Conclusions

Much of arriving at reasonable conclusions involves interpreting findings in light of a theoretical framework (Abi-El-Mona & Abd-El-Khalick, 2011). But it also involves actively searching for conflicting findings in the literature. For K-12 students, the literature includes books, interviews with experts, and Internet resources. K-12 students need to be able to reconcile conflicting findings to arrive at reasonable conclusions. This can involve looking for what the preponderance of studies show, privileging findings from more reputable sources, considering limitations and delimitations of studies, and synthesizing different elements of findings to create a cohesive whole (Britt et al., 2014). This is a challenging activity for such students (Bråten et al., 2011), who often are blinded by my-side bias (Britt et al., 2014; D. Kuhn, 1991; Stanovich & West, 2008).

4.2.1.4.1.5 Engage in Scientific Discourse Patterns

Students also need to know and be able to apply and interpret patterns of STEM discourse, including explanations (Britt et al., 2014; Sandoval, 2003) and persuasive and dialectical argumentation (Bricker & Bell, 2008; Ford, 2012; Osborne, 2010; Perelman & Olbrechts-Tyteca, 1958). Behind all scientific explanations are theories, data, and/or biases. Students need to be able to recognize such, both as they create scientific explanations, but also as they interpret those produced by others. For example, if a proposal is advanced to dam a river to produce power, citizens need to be able to weigh the proposed benefits and drawbacks. Furthermore, they need to be able to judge the extent to which an arguer's stakeholder position influences his/her biases, and by consequence, his/her claims and evidence advanced in support of his/her position. As part of this process, they need to be able to evaluate the credibility of evidence, something with which K-12 and college students often struggle (Britt et al., 2014; Nicolaidou et al., 2011).

4.2.1.4.2 Engaging in the Culture of STEM

Key to helping students engage in the culture of STEM is helping them learn the iterative nature of STEM, as well as the importance of modeling, argumentation, and epistemology.

4.2.1.4.2.1 Iterative Nature of STEM

Engaging students in the culture of STEM does not mean getting students to engage in the "scientific method," as the latter is in fact heavily simplified (Abd-El-Khalick, 2012; Lawson, 2010; Tang, Coffey, Elby, & Levin, 2010). STEM professionals do not always start an investigation with a hypothesis, but often engage in an exploratory investigation to identify pertinent variables or to simply observe and describe a system (Franklin, 2005; Klahr & Simon, 1999; Lawson, 2010). For example, exploratory investigations helped scientists uncover the phenomenon of gene expression (Franklin, 2005). Such exploratory studies often do not involve a control condition, and yet they can lead to very important scientific discoveries, and guide further inquiry (Klahr & Simon, 1999). That is, they can indicate and lead to descriptions of important phenomena. As such observations accumulate, STEM professionals can begin to build theory to explain the phenomena. Further investigations can explore whether the new theory explains and predicts other instances of similar phenomena (Klahr & Simon, 1999; Lawson, 2010).

This accumulation of studies along a line of inquiry does not proceed in a linear manner. Rather, it proceeds in fits and starts—in a very iterative manner. Students should have the opportunity to experience the iterative nature of STEM (T. S. Kuhn, 1996; Lammi & Becker, 2013; Nersessian, 2008). The iterative nature holds at its core theory; theory drives the creation of problem representations (modeling; described below), the design and conduct of investigations to understand problems further, the creation of claims, and backing claims with evidence (argumentation; described below) (Giere, 1990; Klahr & Simon, 1999; Nersessian, 2008). The initial model of a problem situation will necessarily be idealized; it can be improved through such processes as establishing limiting cases (Nersessian, 2008), reacting to phenomena that cannot be sufficiently explained through existing theory (Klahr & Simon, 1999) and engaging with other STEM professionals who often apply different perspectives to problems (Giere, 1990). Not all citizens will engage in the entire process of model-building, but they need to understand the process such that they can engage in authentic scientific discourses centered on locally relevant scientific problems (Kolstø, 2001; Sadler et al., 2007).

Scientists need to revisit theory at multiple stages within the problem-solving process, as it can provide a lens through which to view and interpret data, and suggest new directions to go in an investigation (Giere, 1990; Nersessian, 2008). For example, the discovery of the double helix structure of DNA did not occur all at once, but rather happened through iteration of ideas and interaction with arguments from other scientists (Crick, 1974). Needing to iterate toward an ever-improving solution to a scientific problem can be frustrating to students (Belland, Kim, & Hannafin, 2013). Furthermore, students can often see authentic science as consisting of only collecting data, and not analyzing such (Gu, Belland, Weiss, Kim, & Piland, 2015). Thus, it is important to help students control negative emotions and promote positive emotions throughout this process (Belland et al., 2013; Kim & Hodges, 2012; Kim & Pekrun, 2014; Turner & Husman, 2008). But it is also important to help students perceive that they can be successful in this endeavor (Bandura, 1977; Belland et al., 2013; Britner & Pajares, 2006) and that it is of value (Belland et al., 2013; Wigfield & Eccles, 2000).

4.2.1.4.2.2 Modeling

To be conversant in STEM, individuals also need to be able to use the tools of science, engineering, and mathematics to model natural phenomena, and use those models in reasoning and argumentation (Anzai & Yokoyama, 1984; Lesh & Harel, 2003; Pluta, Chinn, & Duncan, 2011; Sensevy, Tiberghien, Santini, Laubé, & Griggs, 2008; Stratford, Krajcik, & Soloway, 1998). This means representing the constituent parts of the system and how they interact. This is key to the first part of problem-solving—representing the problem (Chi et al., 1981; Jonassen, 2003). It is important to be able to model phenomena both qualitatively and also with the language of mathematics (Chi et al., 1981; Giere, 1990; Jonassen, 2011; Larkin, McDermott, Simon, & Simon, 1980). Modeling phenomena qualitatively means thinking widely about the involved entities, using words rather than numbers to describe how such entities interact and connecting the problem elements to existing domain knowledge (Anzai & Yokoyama, 1984; Jonassen, 2003; Lesh & Harel, 2003). However, students often suffer from limited understanding of complex causality, which can limit their ability to model a problem appropriately (Hmelo-Silver & Pfeffer, 2004; Perkins & Grotzer, 2005). That is, one cannot identify a factor A that directly causes factor B in all systems; students who think that they should always find such a relationship will likely often create an incorrect model (Perkins & Grotzer, 2005).

Students often also suffer from a poor understanding of the words with which to precisely describe a scientific relationship; this can lead them to construct representations of scientific phenomena that do not reflect reality (Leont'ev, 1974; Sensevy et al., 2008). Furthermore, they often perceive that they need to enter values from the problem description into an equation, rather than attempt to construct a qualitative representation (Van Heuvelen & Zou, 2001). When developing a qualitative model, a representation is conducted at first in a learner's mind, and then can be externalized in such forms as a concept map, a textual representation, and/or a diagram (Chi et al., 1981; Jonassen, 2003). The process of articulation can lead to improvement of the model (Belland et al., 2008; Land & Zembal-Saul, 2003; Quintana et al., 2004). Qualitative representations can then be iteratively improved.

Modeling phenomena with mathematics includes setting up an equation that describes the phenomena. It is important to note that effective problem solvers do not solely model problems qualitatively or quantitatively; rather, they use both sorts of representation, as each informs the other and together can lead to a more effective solution and solution process (Chi et al., 1981; Jonassen, 2003; Van Heuvelen & Zou, 2001). For example, after creating a qualitative model, one may proceed to create a quantitative model. The finished qualitative model will influence how the quantitative model is set up. One should then see where the models are consistent, and where they contradict each other; in this way, the models can be progressively improved. By spending adequate time modeling, one can engage in more effective problem-solving, as it guides subsequent investigations, can activate solution schemas, and can provide the framework by which one can simulate what would happen when a variable is manipulated (Anzai & Yokoyama, 1984; Chi et al., 1981; Jonassen, 2003; Sins, Savelsbergh, & van Joolingen, 2005).

Just as it is important to learn to create models, it is also important to be able to interpret the models created by others, especially in terms of what these diverse models say differently about the underlying problems (diSessa, 1988; Seufert, 2003; Wu, Krajcik, & Soloway, 2001). Doing so can lead to enhanced understanding of the problem (Seufert, 2003). This is particularly challenging for K-12 students (Bråten et al., 2011; Seufert, 2003). Indeed, learners often simply adhere to the model that is closest to their own early experiences, or the simplest explanation of the underlying phenomenon, even when presented with a more accurate model (diSessa, 1988; Perkins & Grotzer, 2005). This may be explained in part by most K-12 students' lack of familiarity with complex causal models, such as those that explain changes in a factor through indirect action from a combination of factors A and B (Perkins & Grotzer, 2005). While some evidence indicates that reluctance to consider an alternative model is widespread among learners of differing levels of prior knowledge and skill, other evidence indicates that it may be more prevalent among lower-achieving students (Seufert, 2003). Thus, it is especially important to endeavor to increase modeling skills from a social justice vantage point and to broaden participation in STEM (Lynch, 2001).

4.2.1.4.2.3 Argumentation

Science is very much a social endeavor, as no scientist works in a vacuum (Ford, 2012). Rather, scientists work in a large community of practice in which they share and defend findings to one another, and build off of others' work. At the core of this is argumentation, defined as both backing claims with evidence and models, but also effectively evaluating claims on the basis of evidence and models (Ford, 2012; Osborne, 2010). The argumentation process allows scientific models and theories to be iteratively improved (Ford, 2012). To be able to engage in STEM effectively as citizens, individuals also need to be able to engage in clear argumentation (Aufschnaiter et al., 2008; Jonassen, 2011; Osborne, 2010; Perelman & Olbrechts-Tyteca, 1958). For example, when scientific issues are discussed, citizens need to be able to sort out well-founded claims from less-well-founded claims. K-12 students (Hogan & Maglienti, 2001; Weinstock, Neuman, & Tabak, 2004) and adults (D. Kuhn, 1991) often struggle to evaluate arguments, in part due to poor ability to evaluate the credibility of evidence (Bråten et al., 2011; Nicolaidou et al., 2011).

There are several key areas that need to be addressed in the course of learning to argue. First, there is the conceptual level—helping students understand what a well-founded argument is and is not, and by extension recognize strong and weak arguments. After all, before one can hope to help students learn a skill, they need to be familiar at a conceptual level with the skill that is being learned (Wood et al., 1976). Specifically, students need to understand that an argument is linking a claim to evidence by way of premises to which the claimer and the audience adhere, in the pursuit of leading the audience to adhere to the claim (Perelman & Olbrechts-Tyteca, 1958). A well-founded argument is one that performs this function well, within the framework of generally accepted rhetorical principles. Being able to distinguish between strong and weak arguments relies in part on sophisticated epistemological understanding (Hogan & Maglienti, 2001; Weinstock et al., 2004), which refers to how one thinks that knowledge is established and justified (Mason & Scirica, 2006). This is described in more detail in the next section.

Next, individuals need to learn about the process of argumentation. This involves first making a claim. But before one can establish a claim, one needs to thoroughly understand the underlying problem, including the involved entities and how they interrelate. To do so, one needs to define the problem, determine needed information, and find and organize the information (Belland et al., 2008). Next, one needs to connect evidence to the claim. In so doing, one needs to appeal to premises by which the evidence connects to the claims. Ideally, one employs premises with which the audience already agrees (Perelman & Olbrechts-Tyteca, 1958). Premises that are widely held by the majority of the audience can be left unsaid, while premises that are not held as given by the majority of the audience need to be stated (Perelman & Olbrechts-Tyteca, 1958). For example, if one wanted to claim that Brazilians are unhappy that the Brazilian team was knocked out of the World Cup, one could provide evidence that the Brazilian team in fact was knocked out of the World Cup and that many Brazilians are unhappy. One would also rely on a premise that people tend to be unhappy when their national team in their most popular sport loses.

4.2.1.4.2.4 Epistemology

Closely connected to learning argumentation is a need to develop sophisticated epistemic beliefs, defined as beliefs about the sources, certainty, justification, and simplicity of knowledge that align with that of most STEM professionals (Bendixen & Rule, 2004; Hofer & Pintrich, 1997). With sophisticated epistemic beliefs, an individual knows that claims need to be supported with well-justified, converging evidence, such as evidence collected through tests of a refutable question (Chinn et al., 2011; Hogan & Maglienti, 2001; Mason & Scirica, 2006; Weinstock et al., 2004). Next, with sophisticated epistemic beliefs, one understands that justification for knowledge claims should come from rational arguments or empirical evidence, rather than an appeal to authority (Hogan & Maglienti, 2001;

Jiménez-Aleixandre, 2014). Furthermore, with sophisticated epistemic beliefs, one understands that arriving at correct information/conclusions will often not happen instantaneously (Chinn et al., 2011; Greene, Azevedo, & Torney-Purta, 2008). Someone with sophisticated epistemic beliefs will also understand that most knowledge is not certain, and rather is subject to verification through further research (Hofer & Pintrich, 1997). Without sophisticated epistemic beliefs, individuals often jump to erroneous conclusions (Hofer, 2001; Weinstock et al., 2004). Epistemic beliefs influence individuals' ability to interpret conflicting information from multiple scientific texts (Bråten et al., 2011). The sophistication of middle school students' epistemic beliefs significantly predicted their ability to produce arguments, counter-arguments, and rebuttals (Mason & Scirica, 2006). Epistemic beliefs have also been associated with conceptual change: the more sophisticated the epistemic beliefs, the easier it is to achieve conceptual change given the proper instruction, and vice versa (Hofer, 2001).

4.2.1.4.3 Engaging with Authentic STEM Issues

To be clear, the idea of helping all citizens learn some cross-cutting concepts does not mean reestablishing a rhetoric of conclusions approach to science education (Chinn & Malhotra, 2002; Duschl, 2008)—one focused on transmitting an unchanging body of scientific knowledge. Rather, it means to teach core concepts in science for which evidence is overwhelming, such as the role of DNA and genetic expression in determining such characteristics as the size, shape, and function of organisms. One can do this by engaging students with authentic STEM problems. *Authentic problems* are characterized by the following factors: they (a) are locally relevant, (b) have multiple valid solutions and solution paths, and (c) relate to one or more aspects of STEM, and addressing them requires the use of the tools of the discipline (Barab, Squire, & Dueber, 2000; Chinn & Malhotra, 2002; Hung & Chen, 2007; Jonassen, 2011).

Authentic problems suitable for use in STEM education include (a) dilemmas, a problem type represented by many socioscientific issues, and (b) design problems, which may be centered in or at least involve engineering education (Jonassen, 2000). A socioscientific dilemma can address whether a factory should be built that would cause pollution and degrade habitat, but would increase jobs (Tal & Kedmi, 2006). To address this problem, students need to consider such scientific concepts as what contributes to the health or lack thereof of coastal habitats. But they also need to consider social equity issues related to the right to work in an appropriate job. Many such problems can involve multiple areas within STEM, as interdisciplinary work can lead to more robust problem solutions (Belland & Fee, 2012; Porter & Rafols, 2009) and is becoming more common in STEM research (Murray, Atkinson, Gilbert, & Kruchten, 2014; Porter & Rafols, 2009).

A design problem could involve how to use design to prevent erosion while supporting local habitat on barrier islands (Kolodner et al., 2003). To address this problem, middle school students need to employ engineering design principles and processes, draw on scientific knowledge, identify and research needed knowledge, and engage in extensive iteration. This engages students in the culture of STEM, but also helps them gain important STEM skills and knowledge. Another design problem could involve the design of an alarm to respond to specific needs (Silk, Schunn, & Cary, 2009). Addressing this problem again requires the use of engineering approaches and scientific knowledge.

Requiring the use of the tools of the discipline means that students should need to engage in similar processes and use similar tools as professionals in the target field (Chinn & Malhotra, 2002; Hung & Chen, 2007). It is clear that no students except the most advanced graduate students will use exactly the same processes and tools as professional scientists and engineers, but they should use similar epistemic processes, defined as approaches to designing and conducting investigations, as well as interpreting data and making conclusions (Chinn & Malhotra, 2002).

4.2.2 Learning Content Deeply

Learning content deeply goes beyond simple declarative learning; rather, it refers to the ability to describe knowledge in one's own words and apply it to new situations, as well as recognize the connections between the knowledge and related knowledge (Belland et al., 2009; Bloom, Englehart, Furst, Hill, & Krathwohl, 1956). This outcome has been the focus on much work in scaffolding. One line of such research is that of knowledge integration (Clark & Linn, 2013; Linn, 2000). According to this framework, the knowledge learners bring to school does not need to be replaced by more accurate models, but rather can be used as a base on which to build greater understanding. This is because students' existing knowledge base about science consists of mini theories developed through experience that may be at least partially correct (diSessa, 1988). One can help students build upon their existing knowledge base by encouraging them to engage in authentic problem-solving scenarios supported by scaffolds. However, the goal is not directly to improve problem-solving ability. Rather, it is to help students (a) build enhanced mental models of such things as natural phenomena, and (b) realize that what they are learning applies equally well at home and out in the world as in school (Clark & Linn, 2013; Linn, 2000). However, there is the thought that this in turn could lead to more effective problem-solving (Linn, 2000).

Another line of research on scaffolding that focuses on deep content learning is that of intelligent tutoring systems. In this context, learning content deeply has a different meaning than in scaffolding to support knowledge integration. Namely, intelligent tutoring systems seek to develop students' procedural (production rules) and declarative knowledge related to a particular skill (Anderson, Matessa, & Lebiere, 1997; Self, 1998; VanLehn, 2011). Scaffolding embedded in intelligent tutoring systems helps students apply declarative knowledge to problems. In this way, students develop production rules by which the declarative knowledge can be applied without conscious control to similar problems in the future (Koedinger & Aleven, 2007). But Adaptive Character of Thought-Rational (ACT-R) also endeavors to help students learn declarative knowledge deeply, which means that it can be deployed independently in the future.

4.2.2.1 Alignment with STEM Education Goals

The NGSS and Common Core posit learning content deeply as an important goal (Achieve, 2013; McLaughlin & Overturf, 2012; National Science Board, 2010). For example, one part of the NGSS calls for students to learn cross-cutting concepts. Cross-cutting concepts takes at its core the idea that certain concepts—"patterns; cause and effect: mechanism and explanation; scale, proportion, and quantity; systems and system models; energy and matter; flows, cycles, and conservation; structure and function; and stability and change"—are applicable across a range of STEM disciplines (National Research Council, 2012, p. 3). For example, cause and effect applies equally in science and engineering, and indeed among the many subdisciplines in science and engineering. It is important to note that one cannot always find a single cause that by itself leads to a given effect; often there are multiple causal factors that either together lead to the given effect, or which moderate each other's effect (Hmelo-Silver, Marathe, & Liu, 2007; Perkins & Grotzer, 2005). Seeking to find causal factors for phenomena is a core activity in science (Achieve, 2013) and engineering (Brophy, Klein, Portsmore, & Rogers, 2008; Carr et al., 2012). Furthermore, in engineering, one most often aims to design a product, tool, or strategy that causes a desired outcome (National Research Council, 2012). Considering scale, quantity and proportion is just as important in physics as it is in chemistry, and indeed is important in mechanical and other forms of engineering.

Such cross-cutting concepts are key to the participation of common citizens in discourses about STEM problems. For example, without knowing about flows and cycles as well as systems, one would not be able to intelligently discuss issues related to water quality and access. It is unreasonable to expect everyone to take environmental science classes to learn about such concepts within the context of water quality, and chemistry classes to learn about the application of such concepts in chemistry, and so on. Rather, the hope is that students can learn the concept as a cross-cutting concept in one context, and add depth to their knowledge when learning the same cross-cutting concept in another context, as in a spiral curriculum (Achieve, 2013; Bruner, 2009). Or, at the very least, they would have the base knowledge so that when an authentic socioscientific issue arises, they would be able to converse with it intelligently (Reiser, Krajcik, Gouvea, & Pellegrino, 2014).

Cross-cutting concepts may be best learned in the context of problem-centered instructional models (National Research Council, 2007, 2012). However, abstracting a generalizable cross-cutting concept from such a problem is not easy (Perkins & Grotzer, 2005). First, the target concept may be experienced as context-specific by the student (Perkins & Salomon, 1989). Next, it is not an easy feat to both encode such a concept and include the necessary information to be able to retrieve it later in a new situation in which the concept could be applied (Perkins & Grotzer, 2005). Thus, one may need to be explicit about the cross-cutting nature of concepts, as well as situations in which they can be applied in the future, though this does not need to be done in a didactic manner.

The NGSS also call for students to learn disciplinary core ideas, defined as a few key ideas in each STEM discipline around which one can build STEM curricula (Achieve, 2013; National Research Council, 2012). For example, a core idea in physical sciences revolves around the structure and properties of matter (National Research Council, 2012). A core idea in life sciences relates to the growth and development of organisms (National Research Council, 2012). This approach reflects in many ways the idea of science from a few ideas—the idea that it is more important to know very well a few core ideas in a scientific field, rather than know less well a wide breadth of topics in the given science discipline (Clark, 2000; Pritchard, Barrantes, & Belland, 2009; Schmidt, Wang, & McKnight, 2005). The six countries that performed the best in the Third International Mathematics and Science Study (TIMSS) focused on a much narrower range of key science concepts than most states/districts in the USA (Schmidt et al., 2005). Understanding core ideas does not mean simply being able to describe the idea, but rather to use the idea to describe natural phenomena (Bloom et al., 1956; Reiser et al., 2014). This aligns with the focus on deep content learning of much scaffolding (Clark & Linn, 2013; Linn, Clark, & Slotta, 2003).

4.2.3 Results from Meta-Analysis

In the meta-analysis, outcomes were coded according to whether scaffolding in the studies aimed to increase higher-order thinking skills ($n_{outcomes}=237$), content learning ($n_{outcomes}=95$), or motivation ($n_{outcomes}=1$; See Table 4.1; Belland et al., In Press). This means that 71.2% of included outcomes aimed at enhancing higher-order skills, 28.5% aimed at enhancing content knowledge, and 0.3% aimed to enhance motivation. Results indicated that there was no statistically significant difference between average effect sizes when scaffolding intended to increase higher-order thinking skills ($g=0.45$) versus deep content learning ($g=0.50$). This suggests that scaffolding is a robust instructional approach that can be used to promote diverse learning goals. This is interesting, in that educational interventions tend to not have equally positive influences on content learning and higher-order skills. For example, lecture is well known to be efficient and effective at influencing content learning, but to be ineffective at influencing higher-order thinking abilities (Albanese & Mitchell, 1993; Bland, Saunders, & Frisch, 2007). Problem-based learning tends to lead to strong impacts on higher-order thinking skills, and not on immediate recall of content (Gijbels, Dochy, Van den Bossche, & Segers, 2005; Walker & Leary, 2009). Thus, scaffolding appears to remedy one of the weaknesses of problem-based learning, by helping students learn content knowledge effectively.

Table 4.1 Table of results of moderator analyses on the effect of intended learning outcome on cognitive outcomes

Level	*n* outcomes	Effect size estimate	95% confidence interval	
			Lower limit	Upper limit
Content learning	95	0.50	0.41	0.58
Enhance motivation	1	0.86	0.2	1.52
Higher-order thinking	237	0.45	0.39	0.51

4.3 Assessment

Scaffold designers can set out to design scaffolds with the intention of enhancing students' higher-order thinking abilities or content knowledge. But to be able to verify if the scaffolding that is produced actually enhances such knowledge and skills, it is necessary to consider how the learning is assessed (Cronbach, 1949; Messick, 1989). After all, an assessment that is on the topic of problem-solving does not necessarily assess problem-solving ability. To assess problem-solving, one would need to assess students' abilities to define the problem, determine needed information, and find and synthesize the needed information to arrive at a solution (Belland et al., 2009; Sugrue, 1995).

To assess learning appropriately, it is important to consider the constructs of interest, defined as a characteristic of an individual or group (e.g., intelligence, fluency, and argumentation ability) that cannot be directly measured, and for which one can only measure certain related behaviors (e.g., ability to construct an argument given a scenario and argument construction parameters) (Belland et al., 2009; Kerlinger & Lee, 2000; Messick, 1989). It is necessary to carefully define the constructs to be assessed, and craft a set of activities that can reliably and validly assess the extent to which the test takers evidence a grasp of the target construct (Anastasi & Urbina, 1997; Belland, 2012; Belland et al., 2009; Cronbach, 1949; Messick, 1989). To be reliable, test scores need to be consistent when taken multiple times in close temporal proximity by the same person and also display similar response patterns among people of similar abilities (Kerlinger & Lee, 2000; Messick, 1989). To be valid, a variety of evidence needs to support the conclusion that the set of test scores issuing from the administration of a test are a fair reflection of the amount of the underlying construct the test taker has (Kerlinger & Lee, 2000; Messick, 1989). To be valid, a set of test scores needs to also be reliable (Messick, 1989).

When examining assessment of learning results from the use of computer-based scaffolding, it is useful to consider the assessment framework of Sugrue (1995), who classified assessments in terms of whether they measure at the concept, principles, or application level (see Fig. 4.2). When doing so, it is important to avoid the temptation to label all multiple choice assessments as concept-level assessments, and all open-response assessments as principles or application-level assessments (Hancock, 1994). Measuring at the concept level means that the assessment measures how well students can define or recognize examples of a given concept. This could include assessments ranging from multiple choice tests in which students need to choose a definition, to sorting tasks, and short answer assessments. Measuring at the principles level means that students are provided scenarios involving relationships among several variables and need to predict what would happen if one of the variables were manipulated in a particular way. This again could take many different forms, ranging from multiple choice to writing essays. Measuring at the application level means that students need to design and conduct an investigation using the newly learned material. This is often a performance-based assessment, but can take other forms, such as multiple choice (Hancock, 1994). In many ways, the concept, principles, and application levels parallel the intended learning outcomes

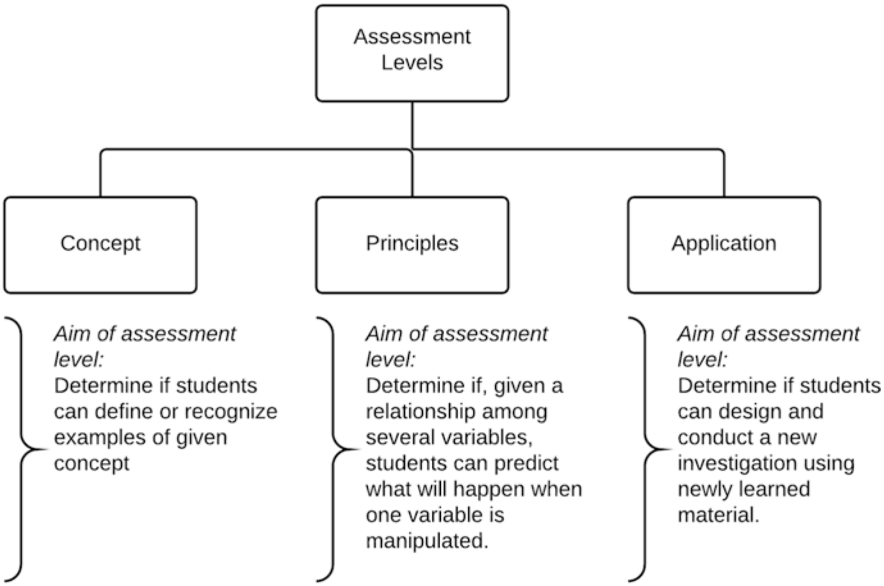

Fig. 4.2 Aims of assessments at the concept, principles, and application levels, as proposed by Sugrue (1995)

of scaffolding. But it is important to make the distinction between intended learning outcomes and assessment levels, as the former are goals towards which designers work when designing scaffolds, and the latter are the ways in which student learning is assessed. These are not always one and the same (Boud & Falchikov, 2006).

4.3.1 Results from Meta-Analysis

It is natural to question whether there are any differences in effect sizes of computer-based scaffolding according to the different assessment levels. For example, if scaffolding is designed to promote problem-solving ability, one would imagine that assessment at the principles or application levels would be more sensitive to the effect of said scaffolding. And if scaffolding is intended to influence content learning, then one would expect that concept-level assessment would be most sensitive to the effect of the scaffolding. Outcomes at the concept level ($n_{outcomes}=125$), principles level ($n_{outcomes}=167$), and application level ($n_{outcomes}=41$) were included (See Table 4.2; Belland, Walker, Kim and Lefler, In Press). Scaffolding's impact on cognitive outcomes was statistically greater when measured at the principles level ($g=0.51$) than when measured at the concept level ($g=0.40$). The effect size for scaffolding at the application level was $g=0.44$. Thus, the effect size point estimate for scaffolding ranged from 0.40 to 0.51 for the three assessment levels. The

Table 4.2 Table of results of moderator analyses on the effect of assessment level on cognitive outcomes

Level	*n* outcomes	Effect size estimate	95 % confidence interval	
			Lower limit	Upper limit
Concept	125	0.40	0.33	0.47
Principles	167	0.51	0.44	0.59
Application	41	0.44	0.32	0.57

95 % confidence intervals—(0.33–0.47), (0.44–0.59), and (0.32–0.57) for concept, principles, and application level assessment, respectively—indicate that one can have great confidence that scaffolding leads to substantial effects across all three assessment levels. This is intriguing, in that it is rare for educational interventions to have such a consistent effect across assessment levels. For example, the underlying instructional models with which scaffolding is used often produce strong effects in one or two of the assessment levels, but not all three. Problem-based learning (PBL) meta-analyses have indicated the PBL leads to effects that are statistically greater than zero at the principles (Gijbels et al., 2005) or the principles and application levels (Walker & Leary, 2009), but not at the remainder of the assessment levels.

There are several possible explanations of the robust effect of scaffolding across assessment levels. First, scaffolding designed to impact higher-order thinking abilities may only be assessed at the principles and application levels, and be mostly successful at influencing student learning as measured by the given assessments; likewise, scaffolding designed to influence content learning may be assessed largely at the concept level, and be mostly successful in influencing learning at that level. Next, it may be possible that scaffolding designed to enhance content learning is also assessed at the principles and application levels, and it also has a positive influence at those levels. It is possible also that scaffolding designed to enhance higher-order thinking abilities is assessed at the concept, principles, and application levels, and leads to strong learning outcomes at all three levels. After all, one of the arguments for promoting content learning in the context of problem-solving is that this will increase students' abilities to solve problems through the enhancement of students' mental models (Anderson, 1983; Clark & Linn, 2013; Johnson-Laird, 2001).

It is especially interesting that scaffolding leads to such a strong effect at the application level. The lower limit of its confidence interval was 0.32, which is an effect of a substantial magnitude—one that is higher than one often finds in educational technology applications for mathematics learning ($ES = 0.15$; Cheung & Slavin, 2013). To perform well on an application level assessment, one must understand the target strategy to a sufficient extent to be able to apply it to a new situation (Sugrue, 1995). This is a very difficult bar to clear, as it requires abstraction of the underlying strategy, and application of said strategy in a new situation that likely differs in key aspects. In short, it is essentially far transfer that is being targeted, which is very difficult to promote (Barnett & Ceci, 2002; Salomon & Perkins, 1989).

References

Abd-El-Khalick, F. (2012). Examining the sources for our understandings about science: Enduring conflations and critical issues in research on nature of science in science education. *International Journal of Science Education, 34*(3), 353–374. http://doi.org/10.1080/09500693.2011.629013.

Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D., & Tuan, H. (2004). Inquiry in science education: International perspectives. *Science Education, 88*(3), 397–419. http://doi.org/10.1002/sce.10118.

Abi-El-Mona, I., & Abd-El-Khalick, F. (2011). Perceptions of the nature and "goodness" of argument among college students, science teachers, and scientists. *International Journal of Science Education, 33*(4), 573–605. http://doi.org/10.1080/09500691003677889.

Achieve. (2013). Next generation science standards. Retrieved on August 8, 2013 from http://www.nextgenscience.org/next-generation-science-standards.

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning—A review of literature on its outcomes and implementation issues. *Academic Medicine, 68*(1), 52–81.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th edn.). Upper Saddle River, NJ, USA: Prentice Hall.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA, USA: Harvard University Press.

Anderson, J. R., Matessa, M., & Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. *Human-Computer Interaction, 12*(4), 439–462. http://doi.org/10.1207/s15327051hci1204_5.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036–1060. http://doi.org/10.1037/0033-295X.111.4.1036.

Anzai, Y., & Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction, 1*(4), 397–450. http://doi.org/10.1207/s1532690xci0104_2.

Asterhan, C. S. C., & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialog. *Cognitive Science, 33*(3), 374–400. http://doi.org/10.1111/j.1551-6709.2009.01017.x.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. http://doi.org/10.1037/0033-295X.84.2.191.

Barab, S. A., Squire, K. D., & Dueber, W. (2000). A co-evolutionary model for supporting the emergence of authenticity. *Educational Technology Research & Development, 48*(2), 37–62. http://doi.org/10.1007/BF02313400.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637. http://doi.org/10.1037//0033-2909.128.4.612.

Belland, B. R. (2010). Portraits of middle school students constructing evidence-based arguments during problem-based learning: The impact of computer-based scaffolds. *Educational Technology Research and Development, 58*(3), 285–309. http://doi.org/10.1007/s11423-009-9139-4.

Belland, B. R. (2012). The role of construct definition in the creation of formative assessments in game-based learning. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 29–42). New York: Springer.

Belland, B. R., & Fee, S. B. (2012). Conclusion: Building on the strengths of interdisciplinarity. In S. B. Fee & B. R. Belland (Eds.), *The role of criticism in understanding problem solving: Honoring the work of John C. Belland* (pp. 245–248). New York: Springer.

Belland, B. R., Glazewski, K. D., & Richardson, J. C. (2008). A scaffolding framework to support the construction of evidence-based arguments among middle school students. *Educational Technology Research and Development, 56*(4), 401–422. http://doi.org/10.1007/s11423-007-9074-1.

Belland, B. R., French, B. F., & Ertmer, P. A. (2009). Validity and problem-based learning research: A review of instruments used to assess intended learning outcomes. *Interdisciplinary Journal of Problem-Based Learning, 3*(1), 59–89. http://doi.org/10.7771/1541-5015.1059.

Belland, B. R., Kim, C., & Hannafin, M. (2013). A framework for designing scaffolds that improve motivation and cognition. *Educational Psychologist, 48*(4), 243–270. http://doi.org/10.1080/00461520.2013.838920.

Belland, B. R., Gu, J., Armbrust, S., & Cook, B. (2015). Scaffolding argumentation about water quality: A mixed method study in a rural middle school. *Educational Technology Research & Development, 63*(3), 325–353. http://doi.org/10.1007/s11423-015-9373-x.

Belland, B. R., Walker, A. E., Kim, N., & Lefler, M. (In Press). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis.

Bendixen, L. D., & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist, 39*(1), 69–80. http://doi.org/10.1207/s15326985ep3901_7.

Berzonsky, W. A., & Richardson, K. D. (2008). Referencing science: Teaching undergraduates to identify, validate, and utilize peer-reviewed online literature. *Journal of Natural Resources and Life Sciences Education, 37*, 8–13.

Bland, M., Saunders, G., & Frisch, J. K. (2007). In defense of the lecture. *Journal of College Science Teaching, 37*(2), 10–13.

Bloom, B. S., Englehart, M. B., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals (Vol. Handbook 1: Cognitive Domain)*. New York: Longman.

Bolhuis, S. (2003). Towards process-oriented teaching for self-directed lifelong learning: A multidimensional perspective. *Learning and Instruction, 13*(3), 327–347. http://doi.org/10.1016/S0959-4752(02)00008–7.

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31*(4), 399–413. http://doi.org/10.1080/02602930600679050.

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48–70. http://doi.org/10.1080/00461520.2011.538647.

Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education, 92*(3), 473–498. http://doi.org/10.1002/sce.20278.

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching, 43*(5), 485–499. http://doi.org/10.1002/tea.20131.

Britt, M. A., Richter, T., & Rouet, J.-F. (2014). Scientific literacy: The role of goal-directed reading and evaluation in understanding scientific information. *Educational Psychologist, 49*(2), 104–122. http://doi.org/10.1080/00461520.2014.916217.

Brophy, S., Klein, S., Portsmore, M., & Rogers, C. (2008). Advancing engineering education in P-12 classrooms. *Journal of Engineering Education, 97*(3), 369–387. http://doi.org/10.1002/j.2168-9830.2008.tb00985.x.

Bruner, J. S. (2009). *The process of education*. Cambridge, MA, USA: Harvard University Press.
Carr, R. L., Bennett, L. D., & Strobel, J. (2012). Engineering in the K-12 STEM standards of the 50 U.S. states: An analysis of presence and extent. *Journal of Engineering Education, 101*(3), 539–564. http://doi.org/10.1002/j.2168-9830.2012.tb00061.x.
Chang, H.-Y., & Linn, M. C. (2013). Scaffolding learning from molecular visualizations. *Journal of Research in Science Teaching, 50*(7), 858–886. http://doi.org/10.1002/tea.21089.
Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88–113. http://doi.org/10.1016/j.edurev.2013.01.001.
Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science: A Multidisciplinary Journal, 5*(2), 121–152. http://doi.org/10.1207/s15516709cog0502_2.
Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*(2), 175–218. http://doi.org/10.1002/sce.10001.
Chinn, C. A., Buckland, L. A., & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist, 46*(3), 141–167. http://doi.org/10.1080/00461520.2011.587722.
Cho, K., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem-solving. *Educational Technology Research and Development, 50*(3), 5–22. http://doi.org/10.1007/BF02505022.
Clark, D. B. (2000). Scaffolding knowledge integration through curricular depth (PhD Dissertation). University of California, Berkeley, United States—California. ProQuest Dissertations & Theses Full Text. (Publication Number AAI3001793).
Clark, D. B., & Linn, M. C. (2013). The knowledge integration perspective: Connections across research and education. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 520–538). New York: Routledge.
Clark, D. B., & Sampson, V. D. (2007). Personally-seeded discussions to scaffold online argumentation. *International Journal of Science Education, 29*(3), 253–277. http://doi.org/10.1080/09500690600560944.
Crick, F. (1974). The double helix: A personal view. *Nature, 248*(5451), 766–769. http://doi.org/10.1038/248766a0.
Cronbach, L. J. (1949). *Essentials of psychological testing*. Oxford, UK: Harper.
Cuevas, H. M., Fiore, S. M., & Oser, R. L. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in computer-based training environments. *Instructional Science, 30*(6), 433–464. http://doi.org/10.1023/A:1020516301541.
Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences, 12*(1), 91–142. http://doi.org/10.1207/S15327809JLS1201_4.
Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education, 22*, 819–837. http://doi.org/10.1080/095006900412293.
Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. http://doi.org/10.1207/S15327965PLI1104_01.
diSessa, A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 49–70). Mahwah, NJ, USA: Lawrence Erlbaum Associates.
Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*(3), 287–312. http://doi.org/10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A.
Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education, 32*(1), 268–291. http://doi.org/10.3102/0091732X07309371.

Eck, R. V., & Dempsey, J. (2002). The effect of competition and contextualized advisement on the transfer of mathematics skills a computer-based instructional simulation game. *Educational Technology Research and Development, 50*(3), 23–41. http://doi.org/10.1007/BF02505023.

Ford, M. J. (2012). A dialogic account of sense-making in scientific argumentation and reasoning. *Cognition and Instruction, 30*(3), 207–245. http://doi.org/10.1080/07370008.2012.689383.

Franklin, L. R. (2005). Exploratory experiments. *Philosophy of Science, 72*(5), 888–899. http://doi.org/10.1086/508117.

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education, 92*(5), 799–824. http://doi.org/10.1002/sce.20270.

Gagné, R. M. (1965). *The conditions of learning and theory of instruction*. New York: Holt, Rinehart, & Winston.

Ge, X., & Land, S. M. (2004). A conceptual framework for scaffolding ill-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development, 52*(2), 5–22. http://doi.org/10.1007/BF02504836.

Giere, R. N. (1990). *Explaining science: A cognitive approach*. Chicago: University of Chicago Press.

Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research, 75*(1), 27–61. http://doi.org/10.3102/00346543075001027.

Glassner, A., Weinstock, M., & Neuman, Y. (2005). Pupils' evaluation and generation of evidence and explanation in argumentation. *British Journal of Educational Psychology, 75*, 105–118. http://doi.org/10.1348/000709904X22278.

Greene, J., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist, 43*(3), 142–160. http://doi.org/10.1080/00461520802178458.

Gu, J., & Belland, B. R. (2015). Preparing students with 21st century skills: Integrating scientific knowledge, skills, and epistemic beliefs in middle school science. In X. Ge, D. Ifenthaler, & J. M. Spector (Eds.), *Full STEAM ahead—Emerging technologies for STEAM* (Vol. 2). New York: Springer.

Gu, J., Belland, B. R., Weiss, D. M., Kim, N., & Piland, J. (2015). Middle school students' science interest and epistemic beliefs in a technology-enhanced, problem-based, scientific inquiry unit. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL, USA.

Hancock, G. R. (1994). Cognitive complexity and the comparability of multiple-choice and constructed-response test formats. *The Journal of Experimental Education, 62*(2), 143–157.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235–266. http://doi.org/10.1023/B:EDPR.0000034022.16470.f3.

Hmelo-Silver, C. E., & 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. http://doi.org/10.1016/S0364-0213(03)00065-X.

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *Journal of the Learning Sciences, 16*(3), 307–331. http://doi.org/10.1080/10508400701413401.

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review, 13*(4), 353–383. http://doi.org/10.1023/A:1011965830686.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88–140. http://doi.org/10.3102/00346543067001088.

Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions. *Journal of Research in Science Teaching, 38*(6), 663–687. http://doi.org/10.1002/tea.1025.

Hung, D., & Chen, D.-T. V. (2007). Context-process authenticity in learning: Implications for identity enculturation and boundary crossing. *Educational Technology Research & Development, 55*(2), 147–167. http://doi.org/10.1007/s11423-006-9008-3.
Hwang, G.-J., Shi, Y.-R., & Chu, H.-C. (2011). A concept map approach to developing collaborative mindtools for context-aware ubiquitous learning. *British Journal of Educational Technology, 42*(5), 778–789. http://doi.org/10.1111/j.1467–8535.2010.01102.x.
Jiménez-Aleixandre, M. P. (2014). Determinism and underdetermination in genetics: Implications for students' engagement in argumentation and epistemic practices. *Science & Education, 23*, 1–20. http://doi.org/10.1007/s11191-012-9561-6.
Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Sciences, 5*(10), 434–442. http://dx.doi.org/10.1016/S1364-6613(00)01751-4.
Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development, 48*(4), 63–85. http://doi.org/10.1007/BF02300500.
Jonassen, D. H. (2003). Using cognitive tools to represent problems. *Journal of Research on Technology in Education, 35*(3), 362–381. http://doi.org/10.1080/15391523.2003.10782391.
Jonassen, D. H. (2011). *Learning to solve problems: A handbook for designing problem-solving learning environments*. New York: Routledge.
Jonassen, D. H., & Kim, B. (2010). Arguing to learn and learning to argue: Design justifications and guidelines. *Educational Technology Research and Development, 58*(4), 439–457. http://doi.org/10.1007/s11423-009-9143-8.
Keefer, M. W., Zeitz, C. M., & Resnick, L. B. (2000). Judging the quality of peer-led student dialogues. *Cognition and Instruction, 18*(1), 53–81. http://doi.org/10.1207/S1532690XCI1801_03.
Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research* (4th ed.). South Melbourne, AU: Wadsworth.
Kim, M., & Hannafin, M. (2011). Scaffolding 6th graders' problem solving in technology-enhanced science classrooms: A qualitative case study. *Instructional Science, 39*(3), 255–282. http://doi.org/10.1007/s11251-010-9127-4.
Kim, C., & Hodges, C. (2012). Effects of an emotion control treatment on academic emotions, motivation and achievement in an online mathematics course. *Instructional Science, 40*(1), 173–192. http://doi.org/10.1007/s11251-011-9165-6.
Kim, C., & Pekrun, R. (2014). Emotions and motivation in learning and performance. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 65–75). New York: Springer.
Klahr, D., & Simon, H. A. (1999). Studies of scientific discovery: Complementary approaches and convergent findings. *Psychological Bulletin, 125*(5), 524–543. http://doi.org/10.1037/0033-2909.125.5.524.
Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*(3), 239–264. http://doi.org/10.1007/s10648-007-9049-0.
Koedinger, K. R., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning sciences to the classroom. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–78). Cambridge, UK: Cambridge University Press.
Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design(tm) into practice. *Journal of the Learning Sciences, 12*(4), 495–547. http://doi.org/10.1207/S15327809JLS1204_2.
Kolstø, S. D. (2001). Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. *Science Education, 85*(3), 291–310. http://doi.org/10.1002/sce.1011.
Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the next generation science standards. *Journal of Science Teacher Education, 25*(2), 157–175. http://doi.org/10.1007/s10972-014-9383-2.
Kuhn, D. (1991). *The skills of argument*. Cambridge, UK: Cambridge University Press.

Kuhn, D. (2010). Teaching and learning science as argument. *Science Education, 94*(5), 810–824. http://doi.org/10.1002/sce.20395.

Kuhn, T. S. (1996). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Kuiper, E., Volman, M., & Terwel, J. (2005). The web as an information resource in K–12 education: Strategies for supporting students in searching and processing information. *Review of Educational Research, 75*(3), 285–328. http://doi.org/10.3102/00346543075003285.

Lammi, M., & Becker, K. (2013). Engineering design thinking. *Journal of Technology Education, 24*(2), 55–77.

Land, S. M., & Zembal-Saul, C. (2003). Scaffolding reflection and articulation of scientific explanations in a data-rich, project-based learning environment: An investigation of progress portfolio. *Educational Technology Research and Development, 51*(4), 65–84. http://doi.org/10.1007/BF02504544.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, New Series, 208*(4450), 1335–1442. http://doi.org/10.1126/science.208.4450.1335.

Lather, P. (2012). The ruins of neo-liberalism and the construction of a new (scientific) subjectivity. *Cultural Studies of Science Education, 7*(4), 1021–1025. http://doi.org/10.1007/s11422-012-9465-4.

Lawson, A. E. (2010). Basic inference of scientific reasoning, argumentation, and discovery. *Science Education, 94*(2), 336–364. http://doi.org/10.1002/sce.20357.

Leont'ev, A. N. (1974). The problem of activity in psychology. *Soviet Psychology, 13*(2), 4–33. http://doi.org/10.2753/RPO1061-040513024.

Lesh, R., & Harel, G. (2003). Problem solving, modeling, and local conceptual development. *Mathematical Thinking and Learning, 5*(2–3), 157–189. http://doi.org/10.1080/10986065.2003.9679998.

Linn, M. C. (2000). Designing the knowledge integration environment. *International Journal of Science Education, 22*(8), 781–796. http://doi.org/10.1080/095006900412275.

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education, 87*(4), 517–538. http://doi.org/10.1002/sce.10086.

Liu, M., & Bera, S. (2005). An analysis of cognitive tool use patterns in a hypermedia learning environment. *Educational Technology Research and Development, 53*(1), 5–21. http://doi.org/10.1007/BF02504854.

Lohman, M. C., & Finkelstein, M. (2000). Designing groups in problem-based learning to promote problem-solving skill and self-directedness. *Instructional Science, 28*, 291–307. http://doi.org/10.1023/A:1003927228005.

Loyens, S. M. M., Magda, J., & Rikers, R. M. J. P. (2008). Self-directed learning in problem-based learning and its relationships with self-regulated learning. *Educational Psychology Review, 20*(4), 411–427. http://doi.org/10.1007/s10648-008-9082-7.

Lynch, S. (2001). "Science for all" is not equal to "one size fits all": Linguistic and cultural diversity and science education reform. *Journal of Research in Science Teaching, 38*(5), 622–627. http://doi.org/10.1002/tea.1021.

Mason, L., & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction, 16*(5), 492–509. http://doi.org/10.1016/j.learninstruc.2006.09.007.

Mason, L., Boldrin, A., & Ariasi, N. (2010). Epistemic metacognition in context: Evaluating and learning online information. *Metacognition & Learning, 5*(1), 67–90. http://doi.org/10.1007/s11409-009-9048-2.

McLaughlin, M., & Overturf, B. J. (2012). The common core: Insights Into the K–5 Standards. *The Reading Teacher, 66*(2), 153–164. http://doi.org/10.1002/TRTR.01115.

McNeill, K. L., & Krajcik, J. (2009). Synergy between teacher practices and curricular scaffolds to support students in using domain-specific and domain-general knowledge in writing arguments to explain phenomena. *Journal of the Learning Sciences, 18*(3), 416–460. http://doi.org/10.1080/10508400903013488.

McNeill, K. L., & Pimentel, D. S. (2010). Science discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203–229. http://doi.org/10.1002/sce.20364.

Merriënboer, J. J. G., & Sluijsmans, D. M. A. (2008). Toward a synthesis of cognitive load theory, four-component instructional design, and self-directed learning. *Educational Psychology Review, 21*(1), 55–66. http://doi.org/10.1007/s10648-008-9092-5.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd edn., pp. 13–103). New York: American Council on Education.

Miles, M. B., & Huberman, A. M. (1984). Drawing valid meaning from qualitative data: Toward a shared craft. *Educational Researcher, 13*(5), 20–30. http://doi.org/10.3102/0013189X013005020.

Murray, J. L., Atkinson, E. J. O., Gilbert, B. D., & Kruchten, A. E. (2014). A novel interdisciplinary science experience for undergraduates across introductory biology, chemistry, and physics courses. *Journal of College Science Teaching, 43*(6), 46–51.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common core state standards*. http://www.corestandards.org/the-standards.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press. http://www.nap.edu/catalog/11625/taking-science-to-school-learning-and-teaching-science-in-grades.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press. http://www.nap.edu/catalog/13165/a-framework-for-k-12-science-education-practices-crosscutting-concepts.

National Science Board. (2010). *Preparing the next generation of STEM innovators: Identifying and developing our nation's human capital (National Science Board Publication No. NSB-10-33)*. Washington, DC: National Science Foundation.

Nersessian, N. J. (2008). *Creating scientific concepts*. Cambridge, MA, USA: MIT Press.

Nicolaidou, I., Kyza, E. A., Terzian, F., Hadjichambis, A., & Kafouris, D. (2011). A framework for scaffolding students' assessment of the credibility of evidence. *Journal of Research in Science Teaching, 48*(7), 711–744. http://doi.org/10.1002/tea.20420.

Nussbaum, E. M. (2002). Scaffolding argumentation in the social studies classroom. *The Social Studies, 93*(2), 79–83. http://doi.org/10.1080/00377990209599887.

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science, 328*(5977), 463–466. http://doi.org/10.1126/science.1183944.

Perelman, C., & Olbrechts-Tyteca, L. (1958). *La nouvelle rhétorique: Traité de l'argumentation [The new rhetoric: Treatise on argumentation]*. Paris: Presses Universitaires de France.

Perkins, D. N., & Grotzer, T. A. (2005). Dimensions of causal understanding: The role of complex causal models in students' understanding of science. *Studies in Science Education, 41*(1), 117–165. http://doi.org/10.1080/03057260508560216.

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 18*(1), 16–25. http://doi.org/10.3102/0013189X018001016.

Pluta, W. J., Chinn, C. A., & Duncan, R. G. (2011). Learners' epistemic criteria for good scientific models. *Journal of Research in Science Teaching, 48*(5), 486–511. http://doi.org/10.1002/tea.20415.

Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics, 81*(3), 719–745. http://doi.org/10.1007/s11192-008-2197-2.

Pritchard, D. E., Barrantes, A., & Belland, B. R. (2009). What else (besides the syllabus) should students learn in introductory physics? In M. Sabella, C. Henderson, & C. Singh (Eds.), *Proceedings of the annual conference of the American Institute of Physics* (pp. 43–46). Ann Arbor, MI, USA: American Institute of Physics. http://doi.org/10.1063/1.3266749.

Qian, G., & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology, 87*(2), 282–292. http://doi.org/10.1037/0022-0663.87.2.282.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences, 13*(3), 337–386. http://doi.org/10.1207/s15327809jls1303_4.

Quintana, C., Zhang, M., & Krajcik, J. (2005). A framework for supporting metacognitive aspects of online inquiry through software-based scaffolding. *Educational Psychologist, 40*(4), 235–244. http://doi.org/10.1207/s15326985ep4004_5.

Reiser, B. J., Krajcik, J., Gouvea, J., & Pellegrino, J. W. (2014). Building theory while supporting implementation of the NGSS. Paper presented at the Discovery Research K-12 PI meeting, Washington, DC, USA.

Sadler, T. D., Barab, S. A., & Scott, B. (2007). What do students gain by engaging in socioscientific inquiry? *Research in Science Education, 37*(4), 371–391. http://doi.org/10.1007/s11165-006-9030-9.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist, 24*(2), 113–142. http://doi.org/10.1207/s15326985ep2402_1.

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences, 12*(1), 5–51. http://doi.org/10.1207/S15327809JLS1201_2.

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education, 89*(4), 634–656. http://doi.org/10.1002/sce.20065.

Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies, 37*(5), 525–559. http://doi.org/10.1080/0022027042000294682.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL, USA: Academic.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education, 36*(1–2), 111–139. http://doi.org/10.1007/s11165-005-3917-8.

Self, J. (1998). The defining characteristics of intelligent tutoring systems research: ITSs care, precisely. *International Journal of Artificial Intelligence in Education (IJAIED), 10*, 350–364.

Sensevy, G., Tiberghien, A., Santini, J., Laubé, S., & Griggs, P. (2008). An epistemological approach to modeling: Cases studies and implications for science teaching. *Science Education, 92*(3), 424–446. http://doi.org/10.1002/sce.20268.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*(2), 227–237. http://doi.org/10.1016/S0959-4752(02)00022–1.

Silk, E. M., Schunn, C. D., & Cary, M. S. (2009). The impact of an engineering design curriculum on science reasoning in an urban setting. *Journal of Science Education and Technology, 18*(3), 209–223. http://doi.org/10.1007/s10956-009-9144-8.

Sins, P. H. M., Savelsbergh, E. R., & van Joolingen, W. R. (2005). The difficult process of scientific modelling: An analysis of novices' reasoning during computer-based modelling. *International Journal of Science Education, 27*(14), 1695–1721. http://doi.org/10.1080/09500690500206408.

Stanovich, K. E., & West, R. F. (2008). On the failure of cognitive ability to predict myside and one-sided thinking biases. *Thinking & Reasoning, 14*(2), 129–167. http://doi.org/10.1080/13546780701679764.

Stratford, S. J., Krajcik, J., & Soloway, E. (1998). Secondary students' dynamic modeling processes: Analyzing, reasoning about, synthesizing, and testing models of stream ecosystems. *Journal of Science Education and Technology, 7*(3), 215–234. http://doi.org/10.1023/A:1021840407112.

Sugrue, B. (1995). A theory-based framework for assessing domain-specific problem-solving ability. *Educational Measurement: Issues and Practice, 14*(3), 29–35. http://doi.org/10.1111/j.1745-3992.1995.tb00865.x.

Tal, T., & Kedmi, Y. (2006). Teaching socioscientific issues: Classroom culture and students' performances. *Cultural Studies of Science Education, 1*(4), 615–644. http://doi.org/10.1007/s11422-006-9026-9.

Tang, X., Coffey, J. E., Elby, A., & Levin, D. M. (2010). The scientific method and scientific inquiry: Tensions in teaching and learning. *Science Education, 94*(1), 29–47. http://doi.org/10.1002/sce.20366.

Turner, J. E., & Husman, J. (2008). Emotional and cognitive self-regulation following academic shame. *Journal of Advanced Academics, 20*(1), 138–173. http://doi.org/10.4219/jaa-2008-864.

Uskola, A., Maguregi, G., & Jiménez-Aleixandre, M. (2010). The use of criteria in argumentation and the construction of environmental concepts: A university case study. *International Journal of Science Education, 32*(17), 2311–2333. http://doi.org/10.1080/09500690903501736.

Van Eemeren, F. H., & Houtlosser, P. (2001). Managing disagreement: Rhetorical analysis within a dialectical framework. *Argumentation & Advocacy, 37*(3), 150–157.

Van Heuvelen, A., & Zou, X. (2001). Multiple representations of work–energy processes. *American Journal of Physics, 69*(2), 184–194. http://doi.org/10.1119/1.1286662.

Van Eemeren, F. H., Grootendorst, R., & Snoeck Henkemans, A. F. (2002). *Argumentation: Analysis, evaluation, presentation*. Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Van de Vord, R. (2010). Distance students and online research: Promoting information literacy through media literacy. *The Internet and Higher Education, 13*(3), 170–175. http://doi.org/10.1016/j.iheduc.2010.03.001.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221. http://doi.org/10.1080/00461520.2011.611369.

Vellom, R. P., & Anderson, C. W. (1999). Reasoning about data in middle school science. *Journal of Research in Science Teaching, 36*(2), 179–199. http://doi.org/10.1002/(SICI)1098-2736(199902)36:2<179::AID-TEA5>3.0.CO;2-T.

von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching, 45*(1), 101–131. http://doi.org/10.1002/tea.20213.

Walker, A., & Leary, H. (2009). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-Based Learning, 3*(1), 12–43. http://doi.org/10.7771/1541-5015.1061.

Walton, D. N. (1989). Dialogue theory for critical thinking. *Argumentation, 3*(2), 169–184. http://doi.org/10.1007/BF00128147.

Weinstock, M., Neuman, Y., & Tabak, I. (2004). Missing the point or missing the norms? Epistemological norms as predictors of students' ability to identify fallacious arguments. *Contemporary Educational Psychology, 29*(1), 77–94. http://doi.org/10.1016/S0361-476X(03)00024-9.

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

Wiggins, G., & McTighe, J. (2005). *Understanding by design*. Alexandria, VA, USA: Association for Supervision and Curriculum Development.

Wijnia, L., Loyens, S. M. M., & Derous, E. (2011). Investigating effects of problem-based versus lecture-based learning environments on student motivation. *Contemporary Educational Psychology, 36*(2), 101–113. http://doi.org/10.1016/j.cedpsych.2010.11.003.

Williams, R. L. (2005). Targeting critical thinking within teacher education: The potential impact on society. *Teacher Educator, 40*(3), 163–187. http://doi.org/10.1080/08878730509555359.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. http://doi.org/10.1111/j.1469-7610.1976.tb00381.x.

Wu, H.-K., Krajcik, J. S., & Soloway, E. (2001). Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching, 38*(7), 821–842. http://doi.org/10.1002/tea.1033.

Zeidler, D. L., Sadler, T. D., Simmons, M. L., & Howes, E. V. (2005). Beyond STS: A research-based framework for socioscientific issues education. *Science Education, 89*(3), 357–377. http://doi.org/10.1002/sce.20048.

Zimmerman, B. J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist, 30*(4), 217–221. http://doi.org/10.1207/s15326985ep3004_8.