ZDM (2020) 52:207–218
https://doi.org/10.1007/s11858-020-01135-z

ORIGINAL ARTICLE

Exploring the affordances of Bayesian networks for modeling usable knowledge and knowledge use in teaching

Nicole B. Kersting[1] · James E. Smith[1] · Beau Vezino[1] · Mei-Kuang Chen[1] · Marcy B. Wood[1] · James W. Stigler[2]

Accepted: 21 January 2020 / Published online: 19 February 2020

Abstract
In this article we propose the use of Bayesian networks as a potentially promising way to model usable knowledge. Using the Classroom Video Analysis (CVA and CVA-M) assessments as a lab model for studying teachers' usable knowledge, we first explored whether we can identify the knowledge (pieces) underlying teachers' written responses. In the CVA approach we ask teachers to respond to short video clips of authentic classroom instruction based on different prompts that are designed to simulate common teaching tasks. We then explored the affordances of Bayesian networks to functionally model usable knowledge as an interconnected dynamic knowledge system consisting of different knowledge pieces and connected pathways weighted by situation-specific relevance and applicability. We explore the implications of these models for studying the development and growth of usable knowledge and propose the use of Bayesian networks as a novel and potentially promising way to model usable knowledge and for understanding how knowledge is used in teaching.

Keywords Usable knowledge · Knowledge use · Teacher knowledge · Teaching practice · Knowledge system · Bayesian networks

1 Introduction

Although there is widespread agreement that teaching requires a great deal of knowledge, we still know little about how knowledge becomes usable and how teachers' use their knowledge for instructional decision making (Ball & Bass, 2000; Ball, Thames, & Phelps, 2008; Blömeke, Gustaffson, & Shavelson, 2015). In this article, we explore the Classroom Video Analysis instrument (CVA and CVA-M), which asks teachers to respond to short video clips of authentic classroom instruction as a lab model for how teachers use their knowledge in real classroom situations. By lab model we mean that teachers engage in the same basic cognitive processes when responding to the teaching situations shown in the video clips, although in more limited ways, that they engage in in their classrooms, and that, therefore, teachers' answers to the CVA or CVA-M instruments can be considered good approximations of their situation-specific usable knowledge. We explore implications of this view for understanding usable knowledge and knowledge use in teaching. Specifically, we will address two research questions:

1. Can we identify the knowledge contained in teachers' responses to the CVA or CVA-M?
2. What are the affordances of Bayesian networks for modeling the knowledge contained in teachers' responses to the CVA or CVA-M to conceptualize usable knowledge and knowledge use in teaching?

Since Shulman's theoretical analysis of the knowledge base required for teaching (Shulman, 1986, 1987), much of the research on teacher knowledge has focused on identifying knowledge domains specific to teaching (Ball 2000, 2003) and on examining their impact on teaching and student learning (Hill, Schilling & Ball, 2004; Baumert, Kunter, Blum, Brunner et al., 2010). Building on and extending Shulman's work in the area of mathematics, the Mathematics Knowledge for Teaching (MKT) construct with its six subdomains and associated items represent the most well-known example of this broader effort within

✉ Nicole B. Kersting
nickik@email.arizona.edu

1 College of Education, University of Arizona, 1430 2nd Street, Tucson, AZ 85721, USA

2 Department of Psychology, University of California Los Angeles, Box 951563, Los Angeles, CA 90095-1563, USA

the US (Ball, Thames, & Phelps, 2008; Hill, Schilling, & Ball, 2004).

Classifying knowledge into different types and measuring them separately—e.g., content knowledge, pedagogical content knowledge, pedagogical knowledge, etc.—holds obvious benefits for research and measurement purposes, but it may divert attention away from important questions related to knowledge use. Such classifications also raise theoretical and practical questions about distinguishing between content and pedagogical content knowledge in a teaching context because content knowledge that is applied in pedagogically purposeful ways, might be considered pedagogical content knowledge (Baumert et al., 2010). Furthermore, items on traditional assessments of teacher knowledge measure single pieces of knowledge in isolation rather than multiple knowledge pieces in connection as they are used in the process of teaching. In the larger context, this inadvertently reinforces a modular view of knowledge in which different kinds of knowledge and knowledge pieces are seen as separate entities. It implicitly presumes that if teachers have acquired a sufficient amount of knowledge in key knowledge domains, they can flexibly apply it while teaching their students.

This approach, however, does little to advance our understanding of how teachers access and connect different pieces of knowledge as they carry out teaching tasks. As Ball, Thames, and Phelps noted (2008, p. 403): "How such knowledge is actually used and what features of pedagogical thinking shape its use, remains tacit and unexamined.'' There is a notable difference between being able to recall a specific common student mistake on a multiple-choice exam, for example, and effectively addressing that same student mistake in an actual classroom situation. In a classroom, teachers may draw upon combinations of knowledge: knowledge of the common student mistakes, foundational or supportive content knowledge, and other knowledge to decide on the best teaching strategy for a particular learner and context.

To date we have not been able to answer the question of how teachers activate and flexibly combine knowledge in the process of teaching from a functional point of view. Earlier models of cognition (Anderson, 1983) have proposed that declarative knowledge (knowing what) needs to be converted into procedural knowledge (knowing how to) initially through basic interpretative processes and through iterative refinement. In Anderson's model, procedural knowledge consists of condition-action (if–then) pairs, called productions, activated according to rules relating to a goal structure, for example, assessing student thinking. Multiple productions can be combined to carry out complex tasks (Renkl, Mandl, & Gruber. 1996) thereby connecting knowledge to behavior. Anderson's model appears to suggest that the more declarative knowledge teachers have converted into procedural knowledge, the more efficiently they can enact their knowledge in the process of teaching.

A more recent approach proposed by Blömeke and colleagues (2015) hypothesized an intervening construct to link knowledge to performance, which they refer to as situation-specific skills of perceiving, interpreting, and decision-making. The model suggests that these situation- specific skills may mediate or affect whether teachers are able to apply their knowledge in a relevant context. Thus, strengthening teachers perceptive and interpretative abilities presumably enables them to apply their situation-relevant knowledge more efficiently leading to better decision making and more effective teaching performance. The model reconceptualizes the process of knowledge conversion from declarative to procedural found in the Anderson model as a set of situation-specific skills that identify basic cognitive processes, shown to shape actions and behaviors, emphasizing the context-bound nature of knowledge use. Both models resolve why teachers may be able to produce specific knowledge in an assessment context but not in a real classroom situation. For Anderson, declarative knowledge has not yet been "productionized"; for Blömeke and colleagues this is due to insufficient or low situation-specific skills. Neither model details how knowledge activation can be envisioned functionally or how it affects decision making.

In this article we explore questions about knowledge activation and application as they relate to usable knowledge in mathematics teaching. Section 1 addresses research question 1. Different from past studies where we assigned scores to teachers' CVA or CVA-M responses to obtain measures of their usable knowledge, here we examine teachers' responses to the CVA and CVA-M to identify the pieces of knowledge that underlie teachers' responses to the teaching situations shown in the video clips. In Sect. 2, we explore the affordances of Bayesian networks to computationally model the knowledge pieces contained in teachers' responses as a dynamic knowledge network that consists of interconnected pieces of knowledge weighted by situation-specific relevance. We describe how the process of activating and connecting different pieces of knowledge could be envisioned functionally within such a network and describe how such models could be used for predicting instructional decision-making, thereby linking knowledge to teaching practice. In Sect. 3, we discuss implications of using Bayesian networks for research on usable knowledge and knowledge use in teaching.

2 Identifying usable knowledge in teacher responses to the classroom video analysis instruments (CVA and CVA-M)

The classroom video-analysis instruments (the original CVA and the Common Core-aligned CVA-M) use video clips of authentic classroom instruction as stimuli to elicit teachers' knowledge in concrete teaching situations. Teachers' view a set of video clips of mathematics instruction to approximate a real teaching situation and provide written responses to the video clips based on different prompts, which simulate common teaching tasks (Kersting et al., 2016). To obtain measures of teachers' usable knowledge, the written responses are scored according to reliable rubrics.

In the CVA, the same prompt is associated with each of the video clips. The prompt is fairly open-ended, asking teachers to comment on the mathematics, the teacher, and the student(s) and interactions between them. The video clips reflect key mathematical ideas in a given content domain although not a specific content framework. Each response is scored along four reliable rubrics (0–2) indicating whether the response analyzed the mathematics shown in the clip (MC), analyzed student thinking and understanding (ST), contained suggestions for improvement (SI), and provided an in-depth interpretation (DI). The DI rubric indicates whether the response remained entirely descriptive (0), was interpretative in some aspect (1), or whether different interpretative ideas formed a coherent argument (2). For the three remaining rubrics a score of "0" indicates that a response did not address a particular rubric. A score of 1 indicates that a response addressed a scoring rubric and used knowledge descriptively (MC and ST) or pedagogically (SI), whereas a score of 2 indicates that a response provided an in-depth analysis of the mathematics (MC), the mathematical understanding of the student (ST), or provided a mathematically-based suggestion for improvement (SI). Given the intentional open-endedness of the prompt allowing teachers to focus on different aspects of the teaching episodes, the CVA captures a wide range of usable knowledge.

In contrast, the CVA-M is more narrowly focused on the mathematics. The video clips and scoring rubrics are aligned with the Common Core content and practice standards. The prompts of the CVA-M are explicitly focused on the mathematics and simulate common teaching tasks, such as generating targeted mathematical questions to help the student(s) improve their understanding, diagnosing students' mathematical thinking, or providing suggestions for improving the teaching episode that are specific to the mathematics. Thus, different from the CVA scores, which reflect teachers usable mathematical and general pedagogical knowledge, teachers' scores on the CVA-M reflect teachers' usable mathematics knowledge for teaching as it connects to the Common Core.

We hypothesized that by using authentic video clips and teaching-focused prompts, teachers' written responses would not only reflect their knowledge as it pertains to the video clips but also their ability to apply that knowledge in a real classroom situation (Kersting et al., 2010). In prior work, we have shown that quantifying qualitative differences in teachers' responses to the CVA or CVA-M predicted teachers' own teaching measured as instructional quality and their students learning (Kersting et al., 2010, 2012). We have interpreted the empirical evidence to suggest that the CVA and CVA-M capture usable knowledge.

To examine what knowledge teachers draw upon and how teachers combine different knowledge (pieces) when responding to the video clips, we analyze four responses to a fractions video clip, each produced by an experienced 4th or 5th grade classroom teacher, two from the CVA and two from the CVA-M. We include example responses from both measures in our analysis to understand the feasibility of knowledge identification and classification in both kinds of responses and to illustrate how different prompts produce variation in knowledge activation. We purposefully selected higher quality responses to show variation in the activated knowledge. To identify the mathematics knowledge contained in the responses we use the Common Core standards (National Governors Association Center for Best Practices, Council of Chief State School Officers 2010). We apply the classification of general pedagogical knowledge, which includes pedagogical and psychological components, suggested by Guerriero (2013).

To provide some context for understanding the teacher responses, we first describe the teaching episode. We recognize that the following description is no substitute for the actual visual and auditory input experienced by the teacher participants in our study.

In a lesson about fractions, the teacher in the video clip posed the following word problem to her students: "If 132 cupcakes are distributed evenly among 6 containers, how many cupcakes would end up in each container?" Students in the clip appear confused about how to solve the problem, so the teacher brings the class together and invites students to give a clue or a hint rather than the answer. One student suggests dividing 132 by 6. The teacher asks the class whether the hint is helpful to them, and when there is no response, the teacher shifts gears and asks what 1/6 of 132 means. When students remain confused, the teacher instructs students to talk at their tables and come up with hints on how to solve the problem.

After about 1 min during which the teacher circulates the classroom and listens into some of the groups, she calls the class together and invites one student, Elizabeth, to the overhead projector to give a clue. Elizabeth repeats the earlier clue of dividing 132 by 6 and shows her solution on the overhead, which shows a tape diagram with six empty boxes. Elizabeth explains that she divided 132 by 6, which resulted in 22 cupcakes per box, writing 22 in each of the six boxes. She then explains that one of the six boxes represents 1/6 of the 132 cupcakes because there are six boxes in total. It is not clear whether Elizabeth understands that the 132 represents the whole, which is divided into six equal parts reflected in the denominator of "6" in the fraction 1/6. Although the answer the student provides is correct, it is apparent that most students in the class are still unsure of what 1/6 of 132 means. The teaching episode concludes with the teacher asking the class if anyone has a question of Elizabeth at the overhead.

Different pieces of knowledge, both mathematical and general pedagogical,[1] could be relevant and activated when responding to this teaching episode. Based on our understanding, the primary mathematical ideas are interpreting the product $(a/b) \times q$ as a parts of a partition of q into b equal parts (or $1/6 \times 132 = 132 \div 6$) and understanding fractions as division ($132/6 = 132 \div 6$). Supporting mathematical knowledge that teachers might draw on includes division and multiplication of three-digit numbers, understanding a fraction *1/b* as the quantity formed by dividing one whole into b equal parts, understanding 132/6 as equivalent to 22/1 or 22/132 as equivalent to 1/6, or more generally multiplication and division as inverse operations such that $132 \div 6 = 22$ and $22 \times 6 = 132$.

General pedagogical knowledge that might be activated could include strategies of classroom management and assessment, instructional strategies used by the teacher in the video that are identified and interpreted from a pedagogical or psychological learning-process perspective. In addition, comments about choices of representations (e.g., tape diagram) and/or manipulatives would reflect knowledge of diverse learners.

Finally, the teaching episode might also activate knowledge of teaching strategies specific to mathematics (mathematical practices), such as making sense of problems, persevering in solving them, constructing viable arguments and critiquing the reasoning of others (National Governors Association Center for Best Practices, Council of Chief State School Officers 2010).

In what follows, we present two example responses to the CVA prompt that asks teachers to make sense of the observed teaching episode: "View the clip and explain how the teacher and the students interact with the mathematics and with each other." This general prompt allows teachers to choose what to attend to in the teaching episode—a choice we hypothesize will be determined by their usable knowledge.

[1] Instead of differentiating between content and pedagogical content knowledge, which raises both theoretical and practical issues of demarcation (Baumert et al., 2010), we identify both as mathematical knowledge.

2.1 CVA and CVA-M example responses

2.1.1 Example response 1

In the first response, several knowledge pieces can be identified that reflect the teacher's general pedagogical and mathematical knowledge. The response starts with comments focused on pedagogical strategies leading to an assessment of Elizabeth's mathematical understanding and concludes with a suggestion for how to solidify students' understanding of the mathematics:

> I love the ending of "…how many agree with what Elizabeth said." It doesn't quite affirm that Elizabeth is correct, but it doesn't contradict it. Just by eliciting Elizabeth's help to come to the overhead to explain, there is affirmation of a good answer. I love the way the teacher wants the students to help without giving away an answer. I think Elizabeth did an excellent job of explaining that 22 is one-sixth of 132 by putting the 22 in each box. It would have helped if she (or the teacher) followed it up by multiplying 22 by 6 to show that it equals 132.

The response starts off with three pedagogical observations, all three focused on strategies the teacher in the video uses to create a student-centered classroom that engages all students in mathematical work. Specifically, the response indicates knowledge of assessment strategies (*"…how many agree with what Elizabeth said.")*, which is interpreted from a pedagogical and motivational learning process perspective. The other two comments identify teaching strategies, i.e., having Elizabeth explain at the overhead, and giving hints rather than answers, and demonstrate knowledge of their pedagogical value.

The assessment of Elizabeth's understanding—the responding teacher assumes that Elizabeth understood the meaning of 1/6 of 132 because she explained why 22 cupcakes are in each box—shows relevant mathematical knowledge. We cannot be certain which specific mathematical ideas the responding teacher draws on because the response summarizes Elizabeth's explanation. Given the wording in the response, it is likely that the mathematical knowledge that is activated relates to the primary mathematics in the clip, interpreting the 22 cupcakes as the product of 1/6 and 132 or one part of the whole (132) when it is divided into 6 equal parts. It is also possible, however, that the responding teacher applied knowledge of fractions as division ($132/6 = 132 \div 6 = 22$) and on the foundational concept of fractions as partitions of a whole into equal parts, i.e., understanding a fraction $1/b$ (1/6) as the quantity formed by 1 part when a whole (132) is partitioned into b equal parts (6 parts) in order to interpret 22 as 1/6 of 132.

The responding teacher's concluding suggestion of multiplying 6 by 22 to show that it equals 132 is akin to making an instructional decision and aims at solidifying students' understanding of the mathematics. The suggestion uses knowledge about the inverse relationship between division and multiplication, but by wording it within a whole number operations context ($22 \times 6 = 132$) rather than a fractions context, it is not clear whether the responding teacher is aware that this suggestion may lead some students in the class to continue to interpret the problem as a division problem rather than a fraction problem.

The response shows that the responding teacher activated and connected five or six identifiable pieces of knowledge: Three pieces of pedagogical knowledge (knowledge of assessment strategies and knowledge of two teaching strategies) followed by one or two pieces of mathematical knowledge reflected in the assessment of Elizabeth's understanding, and finally, one additional piece of mathematical knowledge underlying the suggestion.

2.1.2 Example response 2

The second example response shows a similar structure and analytic focus. It also shows the use of multiple pieces of general pedagogical and mathematical knowledge. Different from the first response, this second responding teacher, however, makes explicit the question of whether students see the relation between division and fractions in the context of the mathematics problem:

> I liked the way that the teacher encouraged the students to "give a clue" rather than to give the entire answer. I also appreciated how she wanted the students to share something with the entire class, and then when they were reluctant, suggested then discussing with their small table groups. The students may be less intimidated to share a concern with the smaller group, then with the entire class. When one student suggests "dividing 132 by six" for a clue, that seems to demonstrate that they are seeing this as a division problem. In this clip, it was not evident if the students are then led to see how the division problem is also a fraction problem. Perhaps further instruction to show that 22/132 is equivalent to 1/6.

This response identifies three teaching strategies "giving a clue", "sharing with the class", and "discussing in groups", which indicates pedagogical knowledge of their pedagogical and motivational value on the learning process and learning environment. Different from response 1, this responding teacher's attention is firmly placed on students mathematical thinking and understanding. Perhaps this is why this responding teacher perceives the relevance of the first clue "dividing 132 by 6", given earlier by one of the students in the class, but not followed up on by the teacher in the video. It is clear the responding teacher interprets the student comment as evidence that students might be viewing the problem as a division and not as a fraction problem.

The responding teacher's understanding of the students' primary struggle motivates the concluding suggestions, which can be considered the "lab equivalent" to instructional decision-making. The responding teacher draws on two fraction concepts, interpreting fractions as division of the numerator by the denominator ($a/b = a \div b$), and knowledge of equivalence, *22/132 = 1/6,* to support students in recognizing that the division problem is also a fraction problem and to help them make the connection between dividing by 6 and 1/6, both identifiable mathematical knowledge.

2.1.3 Example response 3

Next, we present two example responses to the same video clip, only this time teachers answered the teacher question (TQ) prompt of the CVA-M, which is more targeted and explicitly focused on the mathematics. The prompt asks teachers to generate a mathematical question ("If you were a teacher in this situation what mathematical question might you pose to the students and how would your question help improve the students' mathematical understanding?"). Again, we observe differences in the knowledge underlying teachers' responses.

In this example, we can identify two mathematical knowledge pieces in the response, one underlying the question, the other underlying the rationale motivating the question:

> I would ask them, "What does the denominator represent in the fraction?" When the students understand the concept of how the denominator represents the amount of pieces, then we can discuss how fractions are related to division, and how the whole is the 132 cupcakes and how the sixths are how many pieces or in this case how many boxes they want to split the cupcakes into.

The mathematical knowledge reflected in the teacher question represents basic meaning of fractions knowledge, specifically, the meaning of the denominator as indicating the number of fractional parts a whole has been partitioned into (i.e., interpreting a fraction $1/b$ as the quantity formed when a whole is partitioned into b equal parts). The teacher question is intended to help students make the connection to fractions by understanding fractions as division in order to interpret the 132 as the whole that is divided into six pieces or sixths, which can be identified as mathematical knowledge. By linking meaning of fractions to fraction as division knowledge, the response addresses a key mathematical idea, that is, that fractions can be interpreted as division of the numerator by the denominator ($132/6 = 132 \div 6 = 22$) and that the denominator indicates the number of fractional parts the whole was divided into, thereby allowing students to interpret each sixth of the whole as 22 for this problem.

2.1.4 Example response 4

The teacher in the final example fully articulates the fraction concept that is the learning goal for this math problem: Interpreting the product of a unit fraction and whole number as the size of one part, which can be found when dividing the whole into equal parts. In total, two pieces of mathematical knowledge can be identified in the response.

> I think the students are not looking at the problem as part of a whole. I would ask the students how many containers the 132 were divided into (6), then ask them if they see that number represented in the fraction 1/6 (the denominator, 6). I would then ask, "So the whole has been divided into how many parts?" (6). Then I would circle one box of 22 and ask, "This box represents how many parts of the 6 boxes?" (1). I would then ask them if they see that number represented in the fraction 1/6 (the numerator). This would help students see that taking a part of a whole is the same as dividing the whole into parts.

Different from the prior examples, this response provides a sequence of questions and expected answers, which makes for straightforward identification of the knowledge used in the response. The response starts by identifying students struggling to recognize that the problem is a fraction problem. To address this struggle, the response uses supporting knowledge of fractions, specifically understanding a fraction *1/b* as the quantity formed by dividing one whole into *b* equal parts to generate a series of three questions to help students understand the meaning of 1/6. The response concludes that the sequence of proposed questions would help students understand the key fraction concept.

2.2 Discussion of knowledge identification

In each response, teachers drew upon multiple pieces of knowledge and connected this knowledge in pedagogically meaningful yet different ways to make sense of the teaching situation. These differences in the activated mathematical and general pedagogical knowledge influenced the teachers' assessment of student understanding and led to somewhat differing interpretations of the teaching situation. Differing interpretations in turn led to different suggestions akin to instructional decisions these teachers might make if they found themselves in the same situation.

We found it to be fairly straightforward to identify and categorize the underlying pedagogical and mathematical knowledge pieces in the responses. Moving forward, it would be desirable to use knowledge classifications that can be widely agreed upon so that results from such investigations can be compared across studies. Only example 1 required us to make some inference about the mathematical knowledge underlying the response, which might lead to some inaccuracies if frequent. Nevertheless, based on our analysis of these four example responses, the process of knowledge identification and classification in teachers' responses to the CVA or CVA-M appears to be feasible and could be formalized for future studies. In the next section, we explore the affordances of Bayesian networks as a novel analytic approach to model usable knowledge.

3 Exploring the affordances of Bayesian network to functionally model usable knowledge

Advances in cognitive science have successfully modeled complex decision making from a probabilistic perspective using Bayesian networks (Chater, Tenenbaum, & Yuille, 2006; Gopnik & Tenenbaum, 2007; Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010; Jacobs & Kruschke, 2011; Pouget et al., 2013). Such networks consist of nodes reflecting different variables, and directed edges expressing direct and conditional relationships between variables as numerical probabilities. The structure and resulting conditional probabilities of such networks are determined by comparing all possible combinations (including directionality of dependencies) between variables and identifying those through a network that most accurately model the data. In this way, such models (structure and probabilities) represent approximations of the observed data, not the actual data.

In the context of modeling usable knowledge, these networks then estimate the probabilities of knowledge pieces being activated given the activation state (activate or inactivate) of other knowledge pieces. Thus, the variables or nodes in a Bayesian network are the knowledge pieces identified in teachers' responses to a specific teaching situation, and the directed edges represent causal relationships between them. Specifically, the directed edges represent the probabilistic relationships between the knowledge pieces, i.e., situation-specific weights. In this way, the situation-specific nature of the knowledge being modeled is an integral part of the model itself.

3.1 Basics relationships of Bayesian networks

There are four basic relationships in Bayesian networks (Nagarajan, Scutari, & Lebre, 2013). We describe them in the context of usable knowledge as shown in Fig. 1. (a) Two knowledge pieces A and B that are marginally independent—meaning activation of knowledge piece A has no effect on activation of knowledge piece B and vice versa—are not

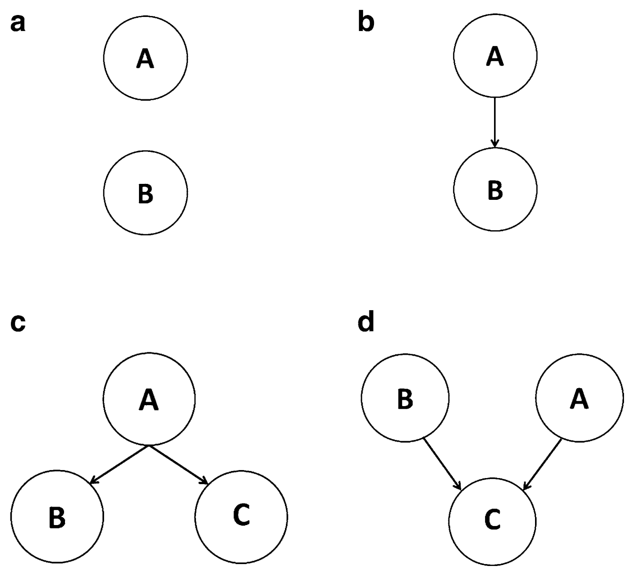

Fig. 1 **a–d** Four types of probabilistic relationships in a Bayesian network

connected through directed edges. (b) Two knowledge pieces A and B, A → B, are directly dependent or causally related. This means that the probability of activation of knowledge piece B is altered if knowledge piece A is activated. (c) Two direct relationships connecting three knowledge pieces, A → B and A → C, causes B and C to become conditionally independent if knowledge piece A is activated (divergent relationship). Intuitively, this stems from the fact that activation of knowledge piece A contains all the information that determines activation of knowledge pieces B and C. (d) Two direct relationships, A → C and B → C, form a convergent relationship. If knowledge piece C is activated, A and B become conditionally dependent because they are coupled by C. The probability of activation of A and B is affected by the activation of knowledge piece C but not by each other. In other words, if C is activated, we can make inferences about activation of A and B, but information about B is not needed to make inferences about A and vice versa. A key requirement in these networks is that there are no directed circles (acyclic).

A strength of Bayesian networks is that once the full probability distribution for the causal relationships has been approximated through the network, conditional probabilities for any combination of connected nodes can be determined. Another strength of these models is that they can be used in descriptive ways, showing the causal dependencies within a knowledge network, in predictive ways, such as predicting instructional decisions or student learning from specific patterns of activated knowledge, and explanatory, where the usefulness of a theoretically derived usable knowledge network (e.g., an ideal answer) can be evaluated against empirical data. All three modes, descriptive predictive, and explanatory are of great value to study usable knowledge.

3.2 Simulated Bayesian network model of knowledge activation

To illustrate the affordances of Bayesian networks for analyzing knowledge activation in teacher responses, we randomly sampled with replacement from the original four responses to generate a sample of size 500 using the *sample* function in R and the hill-climbing, score-based algorithm in the *bnlearn* package in R (Scutari, 2010) to create the Bayesian network shown in Fig. 2. Random sampling with replacement produced the following probabilities for each of the activated knowledge pieces in our sample: Knowledge of *teaching methods* (0.52), *learning processes* (0.52), knowledge of *fractions as division* (0.52) and *basic fraction understanding* (0.48) are contained in about half of the responses, knowledge of *equivalence* (0.27), *multiplication of integers* (0.25), and *interpreting taking part of a whole as dividing the whole into parts* (*ABQ*; 0.23) in about one-fourth of the responses, and knowledge of *division of integers* is present in three quarters of the responses (0.75). Not surprisingly, the reported frequencies closely approximate the observed knowledge piece frequencies in the original four responses. Resampling responses preserved the knowledge relationships we described in Sect. 2, which eases interpretation of the network.

In our simulated example, the network consists of eight knowledge pieces, the total number of knowledge pieces identified across all four responses as shown in Fig. 2. Interpreting these graphs, however, is not straight forward (Butz et al. 2009).

To interpret the network, we focus on three clusters of coactivation dependencies within the network. The coactivation dependencies consists of the basic knowledge activation relationships described in Sect. 3.1. Figure 3a shows a direct dependence of knowledge of *learning processes* and *teaching methods*, indicating that the activation of *teaching methods* knowledge has a direct effect on the activation of *learning processes*.

In this network, the probability of knowledge of *learning processes* (0.52) increases to 0.92 if knowledge of *teaching methods* is activated, which reflects the fact that these two knowledge pieces cooccurred in the two responses that contained knowledge of teaching methods. We further learn that activation of *learning processes* does not affect the activation of any other knowledge. The example illustrates that Bayesian networks can help identify knowledge that may have little impact on other knowledge relationships modeled in the graph. Substantively, such branches in the graph may capture knowledge that may be applicable in a given teaching situation but could be considered supplemental. It would

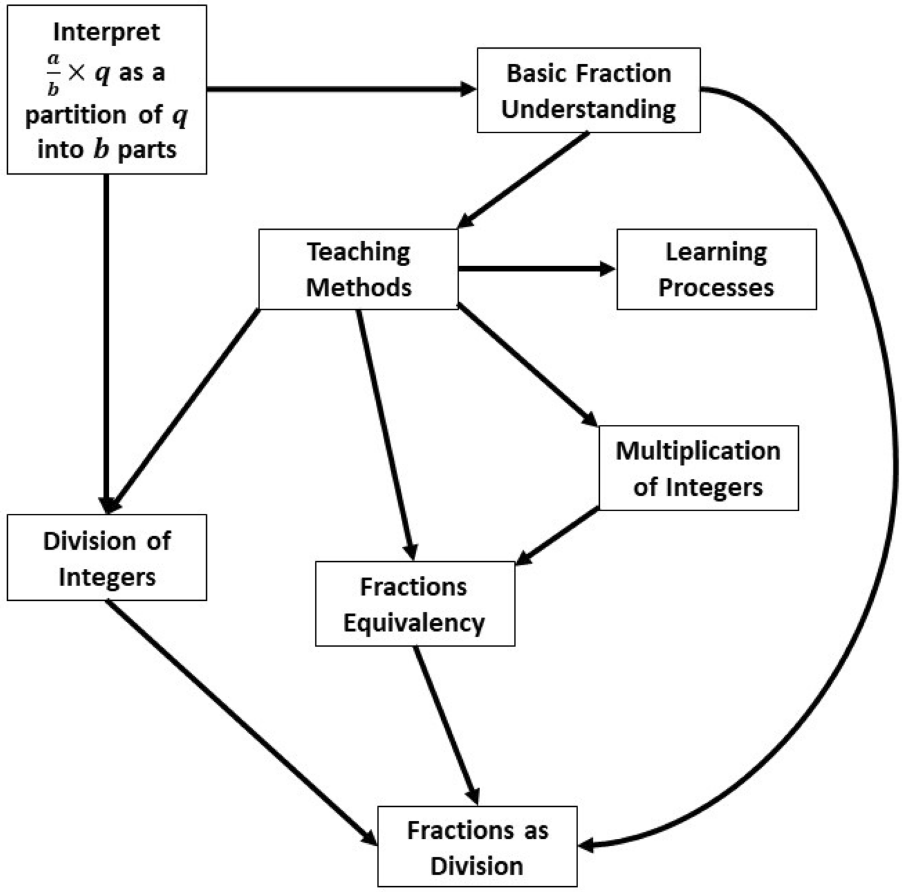

Fig. 2 Simulated Bayesian network

be impossible to identify such branches in an actual empirical dataset with the variability in knowledge that arises in authentic responses.

Figure 3b highlights three direct relationships, which form a convergent relationship. For example, if knowledge of *division of integers* is activated, the probability of activating *fractions as division* knowledge (originally 0.52) reduces to 0.37. Similarly, activated knowledge of *fraction equivalence* increases that probability to .83. Further, if *fractions as div*ision knowledge is activated, knowledge of *division of integers*, *equivalence*, and *basic fraction understanding* become conditionally dependent. That is, the activation of *fractions as division* knowledge alters the probability of activation of any of the three connected knowledge pieces.

For example, if *fractions as division* knowledge is activated but *division of integers* and *fractions equivalency* are not, the probability of activating basic fraction understanding increases from 0.52 to 0.91. This reflects the fact that in the original four responses, *fractions as division* either cooccurred with both *division of integers* and *fraction equivalency* or *basic fraction understanding*, but not all together or all separately. Alternatively, we can determine that if fractions as division is activated along with division of integers and equivalence, a knowledge combination not observed in any of the four original responses, the probability of activating knowledge of division of integers decreases to 0.07. One affordance of these networks is that probabilities for any combination of knowledge pieces can be obtained once the full probability distribution has been determined, regardless of whether the relationship was observed in the data.

Finally, Fig. 3c illustrates a divergent relationship. This relationship indicates that activation of *ABQ* makes supporting knowledge of *division of integers* and *basic fraction understanding* conditionally independent. This means that activation of *ABQ* provides all the information needed to make inferences about *either knowledge*. For instance, if *ABQ* is activated, the probability of *basic fraction understanding* increases from .52 to .85 regardless of activation state of *division of integers*. Conversely, that probability increases to .83 for *division of integers, which in turn decreases the probability of teaching methods knowledge being activated from .52 to .21*. This demonstrates that probabilities in these models not only propagate forward but also backwards, which is another affordance of Bayesian networks.

It is important to remind the reader that the directed edges in the knowledge network represent the directionality of the conditional relationships, not the order in which

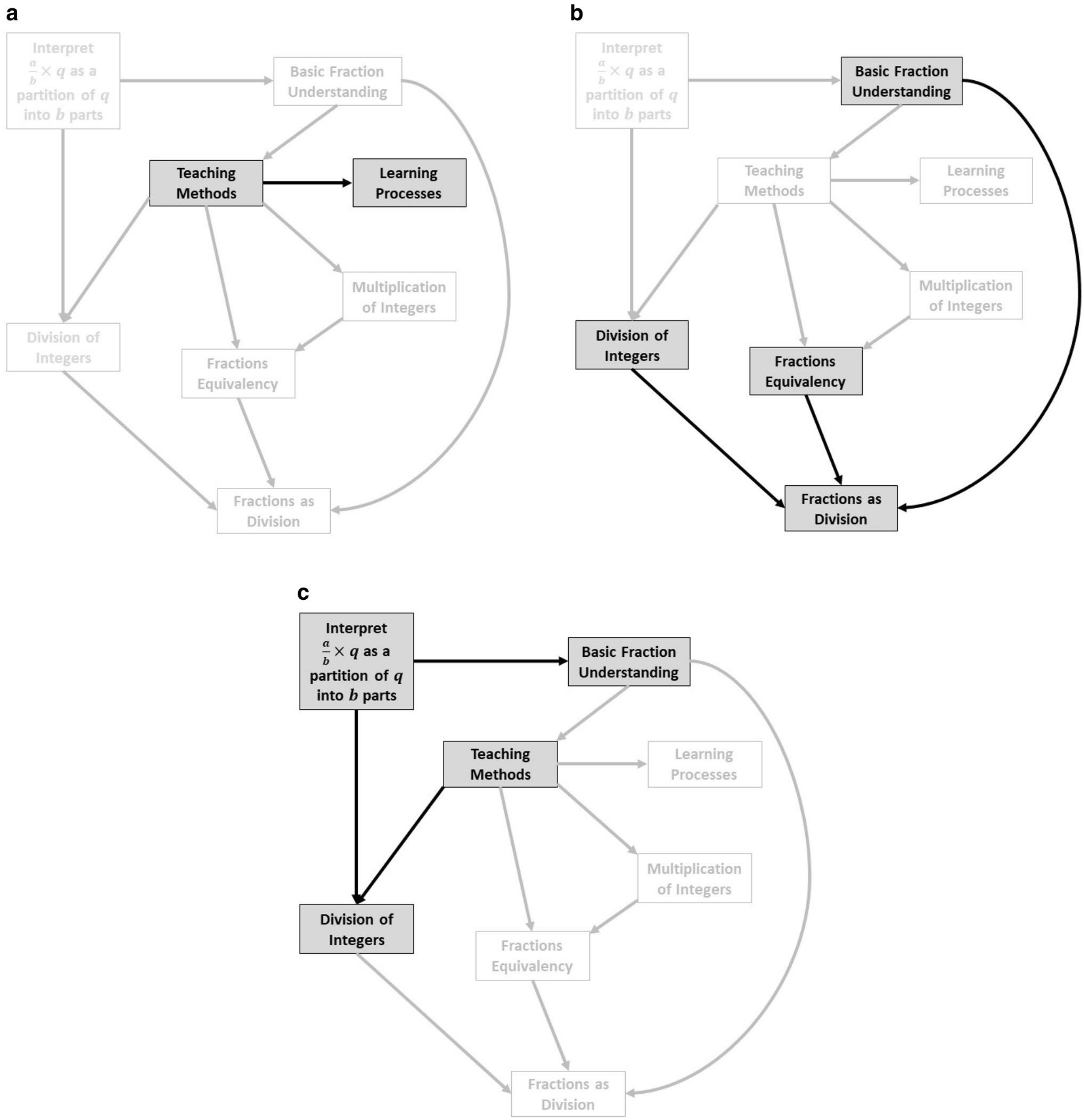

Fig. 3 **a**–**c** Coactivation dependencies

knowledge pieces occurred within responses. Therefore, it would be incorrect to interpret any top to bottom knowledge sequence to mean that the top knowledge leads to activation of the next lower connected knowledge, which in turn leads to the next lower connected knowledge and so on.

3.3 Extending the idea of dynamic knowledge networks to envision knowledge activation in individual teachers

To be clear, the weights in the network we discussed so far represent knowledge activation relationships based on the sample, not individual teachers. Nevertheless, considering this functional mechanism at the individual teacher level can provide useful insights into the process of knowledge activation. If applied to the individual teacher, the weights could be envisioned as capturing seemingly automatic activation of knowledge and subsequent instructional decisions. The processes of knowledge activation might become conscious and deliberate when weights to different connected knowledge are highly similar, rendering all knowledge equally applicable, or when the situation suggests that the most relevant knowledge, based on weights, is actually not applicable, creating dissonance. As teachers become more efficient and experienced the process no longer requires conscious activation.

Some situation-specific weights connecting different knowledge may initially be learned as part of teacher preparation courses. More often, however, building on research demonstrating the effectiveness of practices such as lesson study (Lewis & Perry, 2017) and rehearsal (Kazemi, et al., 2016; Freese, 1999), we hypothesize situation-specific weights are initially created and continuously modified based on experience and reflective practice (Santagata & Yeh, 2016), which highlights the dynamic nature of usable knowledge networks. Like a feedback loop, based on the perceived effectiveness of a knowledge piece used in a specific teaching situation, the weight will be adjusted upward or downward. From this perspective, increasing usable knowledge implies increasing connections to other knowledge pieces and greater differentiation of and more optimal situation-specific weights, in addition to increasing the number of individual knowledge pieces (Russ, Sherin, & Sherin, 2016). In the next section, we will discuss implications of Bayesian networks for research on usable knowledge and knowledge use.

4 Implications of Bayesian networks for research on usable knowledge and knowledge use

In this article, we explored the affordances of Bayesian networks for modeling usable knowledge and for functional conceptualizations of usable knowledge and knowledge use in teaching. We considered the CVA and CVA-M measures as a lab model for how teachers use their knowledge in the process of teaching and identified the knowledge (pieces) underlying teachers' CVA and CVA-M responses. We described why Bayesian networks could be useful to model knowledge activation and instructional decision making. We also showed how such networks support conceptualizing usable knowledge from a functional perspective as a dynamic knowledge network, where connections between different pieces of knowledge represent situation-specific weights that inform knowledge activation for instructional decision-making. The potential of Bayesian networks we have presented in this article has a number of implications for research on usable knowledge and knowledge use in teaching. In this section, we will discuss four implications, we consider theoretically and empirically relevant.

(a) Bayesian network models learn from data. To apply these models, we need data that are good approximations of teachers' usable knowledge. We used our CVA and CVA-M responses, but other data sources containing multiple interconnected knowledge pieces situated in a specific teaching context could also be suitable. Another, and related consideration is, the quality of the data used in these models. Although identification and classification of the individual knowledge pieces in the four example responses we analyzed, appeared to be straightforward, issues related to accuracy and reliability need to be investigated and addressed if this approach were scaled up. In this article, we used the Common Core standards because they encapsulate a view of the mathematics as interconnected ideas to identify the mathematical knowledge in teachers' responses. However, other content frameworks might work equally well and could be broadly applicable across time. For identifying general pedagogical knowledge, we relied on the classification provided by Guerriero (2013), but other classifications can be considered. It would be most effective if there was agreement and validation of knowledge taxonomies within the research community to facilitate comparability of study results.

(b) Connecting usable knowledge and knowledge use. We used the CVA and CVA-M responses as a lab model to approximate how teachers use their knowledge in the process of teaching, which conflates usable knowledge and knowledge use. There is, however, a difference between usable knowledge and knowledge use. Usable knowledge represents a theoretical construct while knowledge use represents the enactment of knowledge in a classroom. Bayesian networks, if used to predict instructional decisions (outputs) from usable knowledge (inputs), provide a simple way with a single model to connect usable knowledge to knowledge use in teaching via instructional decisions. One caveat of these models is that they cannot account for idiosyncratic knowledge, such as knowledge of individual students, which in an actual classroom situation might overwrite the knowledge relationships predicted by the model. This trade-off is outweighed, however, by the benefits of generalizability. It is, for example, conceivable to identify a set of representative video clips that stimulate knowledge activation and use the resulting usable knowledge networks to measure beginning teachers' knowledge in ways that can more accurately capture the complexity and demands of classroom teaching.

(c) Bayesian network models can offer a novel way to study knowledge growth because they efficiently represent highly qualitative information in quantified ways. Networks can provide detailed information on changes in knowledge network structures and weights that cannot be captured by scores on traditional teacher knowledge assessments. Therefore, these models might offer important complementary information to assessment scores. Given that usable knowledge in these models is functionally conceptualized as an interconnected knowledge network, weighted by situation-specific relevance, understandings of knowledge growth can move away from conceptions of simply acquiring more relevant knowledge pieces to an understanding that emphasizes a more differentiated (interconnected and optimally weighted) knowledge network.

From this functional perspective, pre-service or novice teachers' networks have fewer connections or less differentiation in situation-specific weights (making all knowledge somewhat equally likely to apply), which requires more effort from this group of teachers to connect different knowledge in purposeful ways to carry out teaching tasks and to decide which knowledge might be most relevant for a given situation. Expert teachers, on the other hand, can be thought of as having developed a highly sophisticated and differentiated knowledge network reflecting optimally weighted connections, which allows expert teachers to automatically apply the most relevant knowledge in a given teaching situation. This understanding of knowledge growth in teaching also moves beyond efforts to classify knowledge dimensions and to quantify their impact on teaching and learning.

(d) Investigations into knowledge growth. The functional modeling of usable knowledge allows for investigations into how usable knowledge grows over time, answering questions such as why some teachers develop expertise over time while others' teaching practices change little, or why teaching practices are difficult to affect at scale. If teachers actively reflect on the knowledge they draw on when teaching and intentionally activate different, more relevant knowledge to be more effective in a specific teaching situation, slowly with repeated activation of the new knowledge pieces new connections get created, situation-specific weights get modified and new knowledge combinations eventually become the teachers' go-to routine. As Bjork (1975) wrote: Retrieval modifies memory. Teachers who engage in this process over time, build more optimal connections, create more optimal weights, and develop expertise (Fadde, & Klein, 2010).

To be sure, the ideas about functionally modeling usable knowledge and knowledge use we advanced in this article are based on teachers' responses to video clips of authentic classroom instruction, not on knowledge use captured during the actual process of teaching. Nevertheless, they may represent a useful approximation of knowledge activation and application in a real teaching situation. These models add another mode of investigation to already existing frameworks, which might motivate further research (Blömeke et al., 2015). Moving forward, these models could potentially be applied to video recordings coded for enacted knowledge pieces.

Bayesian networks are a promising approach that can help explain a variety of issues related to teacher knowledge and its relationships to teaching that have a history of research but no definite answers. Among them are how to understand knowledge growth in teaching, how to understand knowledge use in teaching, and how to better understand the relationship between knowledge and changes in teaching practice. Much work remains to be done, but at the very least, the ideas we presented offer a new vantage point from which to examine the relationship between teacher knowledge and teaching practice.

Acknowledgements This material is based upon work supported by the National Science Foundation under Grant Nos. 1720866, 0949241, and 1445431. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Anderson, J. R. (1983). *The architecture of cognition*. New York, NY: Psychology Press.

Ball, D. L. (2000). Bridging practices: Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, *51*, 241–247. https://doi.org/10.1177/0022487100051003013.

Ball, D. L. (2003). Using content knowledge in teaching: What do teachers have to do, and therefore have to learn? In *Keynote address at the third annual conference on sustainability of systemic reform*, Cambridge, MA.

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics* (pp. 83–104). Westport, CT: Ablex Publishing.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: what makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180.

Blömeke, S., Gustafsson, J. E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift Fur Psychologie/Journal of Psychology, 223*(1), 3–13.

Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition. *Trends in Cognitive Sciences, 10*(7), 287–291.

Council of Chief State School Officers & National Governors Assocation Center for Best Practices, A. (2010). Common core state standards for mathematics. Commmon Core State Standards Initiative. Retrieved from http://www.corestandards.org/Math/Content/.

Fadde, P. J., & Klein, G. A. (2010). Deliberate performance: Accelerating expertise in natural settings. *Performance Improvement, 49*(9), 5–14.

Freese, A. R. (1999). The role of reflection on preservice teachers' development in the context of a professional development school. *Teaching and Teacher Education, 15*(8), 895–909.

Gopnik, A., & Tenenbaum, J. B. (2007). Bayesian networks, Bayesian learning and cognitive development. *Developmental Science, 10*(3), 281–287.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends in Cognitive Sciences, 14*(8), 357–364.

Guerriero, S. (2013). Teachers' pedagogical knowledge and the teaching profession: Background report and project objectives. Retrieved October 9, 2019, from http://www.oecd.org/edu/ceri/Background_document_%0Ato_Symposium_ITEL-FINAL.pdf.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal, 105*(1), 11–30.

Jacobs, R. A., & Kruschke, J. K. (2011). Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(1), 8–21.

Kazemi, E., Ghousseini, H., Cunard, A., & Turrou, A. C. (2016). Getting inside rehearsals: Insights from teacher educators to support work on complex practice. *Journal of Teacher Education, 67*(1), 18–31.

Kersting, N., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Teachers' analyses of classroom video predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education, 61*(1–2), 172–181.

Kersting, N., Sutton, T., Kalinec-Craig, C., Stoehr, K. J., Heshmati, S., Lozano, G., et al. (2016). Further exploration of the classroom video analysis (CVA) instrument as a measure of usable knowledge for teaching mathematics: taking a knowledge system perspective. *ZDM - Mathematics Education, 48*(1–2), 97–109.

Kersting, N. B., Givvin, K. B., Thompson, B., Santagata, R., & Stigler, J. (2012). Developing measures of usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, *49*(3), 568–590. https://doi.org/10.3102/0002831212437853.

Lewis, C., & Perry, R. (2017). Lesson study to scale up research-based knowledge: A randomized, controlled trial of fractions learning. *Journal for Research in Mathematics Education, 48*(3), 261–299.

Nagarajan, R., Scutari, M., & Lèbre, S. (2013). *Bayesian networks in R with applications in systems biology*. New York, NY: Springer, New York.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: Knowns and unknowns. *Nature Neuroscience, 16*(9), 1170–1178.

Renkl, A., Mandl, H., & Gruber, H. (1996). Inert knowledge: Analyses and remedies. *Educational Psychologist, 31*(2), 115–121.

Russ, R., Sherin, B. L., & Sherin, M. G. (2016). What constitutes teacher learning? In D. Gitomer & C. Bell (Eds.), *Handbook of Research on Teaching* (pp. 391–438). Washington, DC: American Educational Research Association.

Santagata, R., & Yeh, C. (2016). The role of perception, interpretation, and decision making in the development of beginning teachers' competence. *ZDM - Mathematics Education, 48*(1–2), 153–165.

Scutari, M. (2010). Learning Bayesian Networks with the bnlearn R Package. *Journal of Statistical Software, 35*(3), 1–22.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–23.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.