International Journal of Science Education

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

Elementary teachers' use of content knowledge to evaluate students' thinking in the life sciences

Jaime L. Sabel, Cory T. Forbes & Leslie Flynn

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
http://dx.doi.org/10.1080/09500693.2016.1183179

Routledge
Taylor & Francis Group

Elementary teachers' use of content knowledge to evaluate students' thinking in the life sciences

Jaime L. Sabel[a], Cory T. Forbes[a,b] and Leslie Flynn[c]

[a]Department of Teaching, Learning, and Teacher Education, College of Education and Human Sciences, University of Nebraska-Lincoln, Lincoln, NE, USA; [b]School of Natural Resources, University of Nebraska-Lincoln, Lincoln, NE, USA; [c]Department of Teaching and Learning, College of Education, University of Iowa, Iowa City, IA, USA

ABSTRACT

Science learning environments should provide opportunities for students to make sense of and enhance their understanding of disciplinary concepts. Teachers can support students' sense-making by engaging and responding to their ideas through high-leverage instructional practices such as formative assessment (FA). However, past research has shown that teachers may not understand FA, how to implement it, or have sufficient content knowledge to use it effectively. Few studies have investigated how teachers gather information to evaluate students' ideas or how content knowledge factors into those decisions, particularly within the life science discipline. We designed a study embedded in a multi-year professional development program that supported elementary teachers' development of disciplinary knowledge and FA practices within science instruction. Study findings illustrate how elementary teachers' life science content knowledge influences their evaluation of students' ideas. Teachers with higher levels of life science content knowledge more effectively evaluated students' ideas than teachers with lower levels of content knowledge. Teachers with higher content exam scores discussed both content and student understanding to a greater extent, and their analyses of students' ideas were more scientifically accurate compared to teachers with lower scores. These findings contribute to theory and practice around science teacher education, professional development, and curriculum development.

ARTICLE HISTORY
Received 13 October 2015
Accepted 23 April 2016

KEYWORDS
Elementary science; life science; formative assessment; elementary teachers

The *Next Generation Science Standards* (NGSS Lead States, 2013) emphasize conceptual understanding of disciplinary core ideas within the life sciences which influence how students are able to consider concepts that cut across all domains. Past research has shown that elementary students often hold a variety of alternative ideas about core life science concepts (e.g. Anderson, Ellis, & Jones, 2014; Barman, Stein, McNair, & Barman, 2006). Teachers can support students' scientific sense-making about these and other natural phenomena by engaging and responding to students' ideas through formative assessment

CONTACT Jaime Sabel jaime.sabel@huskers.unl.edu

(FA), a high-impact instructional practice that allows teachers to cultivate student-centered learning environments.

Previous studies have examined elementary teachers' use of FA for science and mathematics (e.g. Aschbacher & Alonzo, 2006; Buck & Trauth-Nare, 2009; Forbes, Sabel, & Biggers, 2015a; Hammer, Goldberg, & Fargason, 2012; Heritage, Kim, Vendlinski, & Herman, 2009; Morrison, 2013; Otero & Nathan, 2008). However, few studies have investigated relationships between elementary teachers' knowledge of disciplinary concepts and their FA practices. Little research has been conducted to explain how teachers leverage their disciplinary content knowledge when they make decisions about how to elicit students' ideas and shift instruction to target observed gaps in student understanding. In particular, more work is needed to better understand how they use their content knowledge to analyze evidence of students' thinking, a core element of FA, particularly in the life sciences. Results from such work are crucial to inform efforts that will support teachers to implement FA effectively in science classrooms.

To address these needs, we conducted a study embedded within a multi-year professional development program intended to support elementary teachers to integrate FA practices into their science instruction. Here, we investigate teachers' analysis of student artifacts to diagnose their disciplinary thinking about biological concepts. This study also contributes to two bodies of research: work focused on teaching and learning in the life sciences (e.g. Anderson et al., 2014; Barman et al., 2006; Forbes, Sabel, & Zangori, 2015b; Friedrichsen, 2001; Haefner, Friedrichsen, & Zembal-Saul, 2006; Kikas, 2004; Rice, 2005; Sabel, Forbes, & Zangori, 2015) and work focused on FA in science (e.g. Buck, Trauth-Nare, & Kaftan, 2010; Forbes et al., 2015a, 2015b; Levin, Hammer, & Coffey, 2009; Otero & Nathan, 2008; Sabel et al., 2015; Talanquer, Bolger, & Tomanek, 2015; Talanquer, Tomanek, & Novodvorsky, 2013). Although significant work has occurred in each of these two fields of study, the study presented here contributes specifically to the intersection of the two fields which has not had the benefit of much focus. In particular, this study will help inform instructional practices for elementary life science instruction, research on the factors important for teachers' use of FA, and the design of professional development for elementary teachers and science curriculum materials for elementary classrooms. The following research questions guide this study:

(1) To what extent does elementary teachers' content knowledge influence their ability to implement FA practices within life science instruction?
(2) In what ways do elementary science teachers use life science content to evaluate evidence of students' thinking?

Background and theoretical framework

FA is an instructional practice that allows teachers to engage with students' ideas, identify trends, and create responsive instruction to enhance learning, thus creating a student-centered learning environment (Bell & Cowie, 2001; Coffey, Hammer, Levin, & Grant, 2011). FA can take many forms, but it typically involves (a) *anticipating* students' ideas, (b) *evaluating* individual student progress, and (c) implementing *follow-up* instruction that is responsive to those ideas. A fundamental component of FA is the ability to effectively evaluate evidence of students' thinking. Increasing the use of effective FA has been shown to lead to significant student learning gains in science (Ruiz-Primo & Furtak, 2006).

Theoretical framework

Elementary students arrive in science classrooms with a variety of preexisting ideas about the natural world that may lack scientific accuracy (Donovan & Bransford, 2005). Science learning environments should provide students with opportunities to engage with relevant scientific phenomena in order to expand and refine these preexisting ideas to reach new or revised ideas and greater understanding of scientific concepts. To achieve this, teachers must elicit and respond to students' ideas in order to help them develop scientific understanding and engage in scientific practices (Levin et al., 2009; NGSS Lead States, 2013). In this way, students' ideas and teachers' pedagogical decisions both influence the classroom learning experience. In turn, the classroom learning experience influences students' scientific reasoning and idea development, as well as teachers' responsive pedagogical reasoning. The conceptual framework we have developed to demonstrate this relationship is shown in Figure 1 (Sabel et al., 2015).

A fundamental piece of this, and the primary focus of this study, is teachers' ability to effectively interpret the evidence they gather regarding students' ideas so that they can diagnose gaps in students' thinking and foster learning experiences that are responsive to those ideas.

Teachers' use of FA to evaluate students' ideas

Although FA is a proven instructional practice shown to lead to student learning gains, it is not a widespread practice in elementary classrooms (Coffey et al., 2011; Morrison, 2013). This may be, in part, due to teachers' lack of understanding of what FA is and how they can employ it in elementary science learning environments (Coffey et al., 2011; Hammer

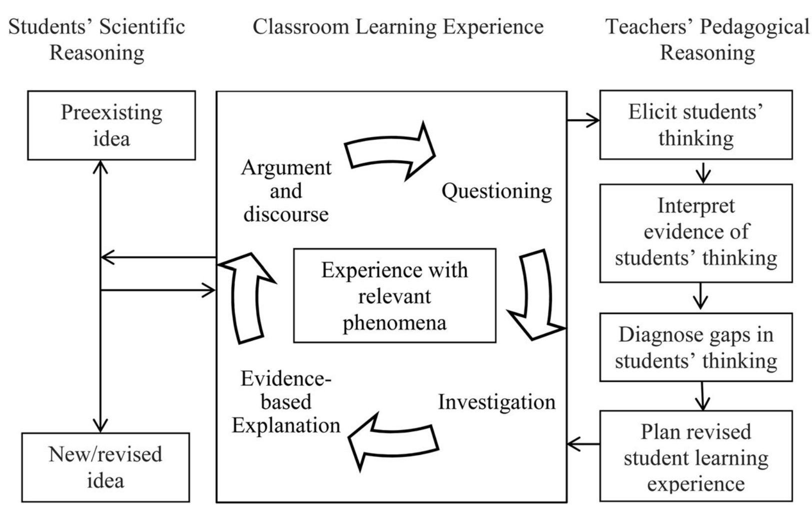


Figure 1. Relationship between responsive instruction and students' science learning (Sabel et al., 2015).

et al., 2012; Otero & Nathan, 2008). Prior research has shown that teachers often do not focus on interpreting what students know or incorporating ideas of how students learn to help advance their understanding. Instead, they tend to rely on ideas of whether students ‘get it or don’t’ , rely on the presence of particular vocabulary words or curriculum-specific language as markers for understanding, or discuss general ideas for how they would address students’ conceptions rather than specific content they might utilize to challenge students’ ideas (Forbes et al., 2015a; Gottheiner & Siegel, 2012; Otero & Nathan, 2008). However, through targeted support in both preservice and inservice contexts, teachers can develop their abilities to analyze evidence of students’ thinking and more effectively engage in the practices of FA (Forbes et al., 2015b; Sabel et al., 2015). Examining students’ work, eliciting students’ understanding, and analyzing reasoning can allow teachers to learn both about and from their students’ thinking and also lead to increased development of their own understanding of the concepts and how assessment can be used as a way to understand student learning (Kazemi & Franke, 2004).

Research on what teachers notice has shown that, in order to respond to student understanding, teachers need to attend to strategies students use to consider the task and the content, interpret their understanding, and decide how to respond to those ideas (Jacobs, Lamb, & Philipp, 2010). In order to make sense of evidence they gather about students’ reasoning, teachers need robust knowledge of disciplinary content, how students think about the content, and the context within the classroom (Van Es, 2011). Prospective teachers, however, tend to focus their interpretation on how well they see students engaging in general science process skills needed to carry out investigations rather than on the scientific plausibility of their students’ ideas (Talanquer et al., 2013). Teachers are unlikely to respond to students’ ideas in ways that consider their understanding unless they purposefully intend to do so and are provided support to learn how to do so (Jacobs et al., 2010; Levin et al., 2009). For teachers to foster classroom learning environments that place emphasis on and are responsive to their students’ ideas, they must have an understanding of how learning develops, how they might integrate students’ ideas into their instruction, and how students’ ideas align with the content within the discipline (Ball, Thames, & Phelps, 2008; Heritage et al., 2009).

Teachers’ disciplinary content knowledge and FA

The relationship between teachers’ content knowledge and their engagement in FA practices remains unclear. Coffey et al. (2011) proposed lack of sufficient science content knowledge as one reason why teachers may not implement FA in their classrooms. However, the literature includes reports of both positive relationships (e.g. Heritage et al., 2009) and nonexistent relationships (e.g. Forbes et al., 2015b) between content knowledge and FA practices, and very few studies have examined this relationship within the life sciences discipline. Therefore, additional exploration of this relationship is needed.

Robust knowledge of disciplinary concepts is widely viewed as a key dimension of teachers’ expertise and ability to engage in effective instruction, including in science (Kennedy, 1998). Prior empirical work provides evidence that teachers with more developed understanding of the content they teach implement more effective science and mathematics instruction in the classroom (Falk, 2011; Hill, Rowan, & Ball, 2005). Results from a study conducted by Hill et al. (2005), for example, show that elementary teachers' disciplinary knowledge of mathematics is a significant factor in student achievement gains (Hill et al., 2005). Yet, prior research has firmly established limitations of elementary teachers' subject matter knowledge for science, particularly the life sciences. In particular, past work has shown that teachers have difficulty understanding biological phenomena targeted in elementary science standards and curricula (Kikas, 2004; NGSS Lead States, 2013; Rice, 2005) and rely heavily on the curriculum to help them make decisions about how well students understand disciplinary concepts (Forbes et al., 2015b; Sabel et al., 2015). Furthermore, teachers have been shown to misinterpret students' scientific explanations as correct simply because they contain the correct scientific terms (Forbes et al., 2015a; Kikas, 2004).

Limitations in content knowledge notwithstanding, it is intuitive to assume that teachers' FA practices are similarly influenced by their grasp of disciplinary concepts. It stands to reason that teachers must possess reasonably robust knowledge of the ideas students are expected to learn in order to engage in effective FA practices, particularly the evaluation of student artifacts as evidence of students' scientific thinking. There is even limited evidence to suggest that engaging in FA itself can support teachers to self-identify as disciplinary thinkers by focusing on multiple and varied reasoning strategies through which students explain scientific phenomena (Ash & Levitt, 2003; Kazemi & Franke, 2004). However, Coffey et al. (2011) have argued that the FA strategies and approaches too often tend to be discipline nonspecific and disconnected from the 'disciplinary substance' upon which they are grounded. This has implications for all constituent formative assessment practices, including opportunities to elicit students' reasoning about disciplinary concepts, the ways teachers interpret evidence of students' thinking about those disciplinary concepts, and discipline-specific instructional strategies through which teachers can provide students with additional learning opportunities.

However, as previously mentioned, the relationship between teachers' science content knowledge and their FA practices, particularly their ability to evaluate and diagnose evidence of students' thinking, remains unclear. Some prior research suggests a positive relationship. For example, Heritage et al. (2009) found that teachers with greater mathematics content knowledge were more effective at diagnosing student thinking and proposing responsive instruction. However, with the exception of a few studies that have tentatively addressed some aspect of content knowledge with practice (Aschbacher & Alonzo, 2006; Gottheiner & Siegel, 2012), little comparative research has been similarly conducted in science. Furthermore, in a previous study (Forbes et al., 2015a), we found no observable relationship between teachers' knowledge of Earth science concepts and the effectiveness of their FA practices. This finding, considered in light of other research showing limited relationships between teachers' content knowledge and FA practices, suggests perhaps more nuanced and complex interactions between teachers' content knowledge and FA practices for science, particularly the evaluation of evidence of students' thinking. More work is therefore needed to determine how teachers draw upon their knowledge of science concepts through instructional practice to make sense of students' disciplinary reasoning. This study builds upon and extends our previous research by exploring this relationship within the life sciences disciplinary domain.

Study design and methods

We used a sequential explanatory mixed-methods research design (Cresswell & Plano Clark, 2011) to investigate how teachers' content knowledge influenced the ways in which they engaged in FA practices. We used quantitative methods to address the first research question and then used those findings to inform the qualitative analysis conducted in response to the second research question. The objective of the qualitative analysis was to more thoroughly illustrate and explain trends established through quantitative data analysis.

Contexts and participants

This study was part of a three-year sustained professional development program designed to support elementary (K-5) teachers in their implementation of FA for elementary science (see Forbes et al., 2015a). This study involved 32 teachers from grades 3–6 classrooms in 12 schools from 4 school districts in a single Midwestern state in the United States. The teachers were selected from a larger pool of potential participants based on their use of kit-based instructional materials for science (e.g. FOSS, Insights, STC) from commercial publishers provided by a regional science curriculum center. This study was conducted during the second year of the program.

Teachers engaged in professional development for 7 days for each of 2 summers prior to this study and met in small, collaborative learning teams 12 times during each school year. The professional development opportunities focused on engaging teachers in using FA in science, examining student work, developing FA prompts to use in their own classrooms, implementing instruction to respond to students' ideas, and increasing their science content knowledge through science investigations as learners and engaging in Curriculum topic study (CTS; Keeley, 2005). The focus of the project's second year was on life science, so the professional development, including science investigations and CTS, emphasized life science topics, instructional strategies, and learning outcomes for students.

Each teacher taught at least one 8-week life science unit using commercially published, kit-based elementary science curriculum materials and a few taught two life science units throughout the school year (see Table 1). These included five different units about plants, the human body, biological structure and foundation, and ecosystems. The timing of teachers' implementation of these life science units was dependent upon their own normal district and building curriculum and schedules and was not influenced by their participation in the program.

Data collection

As a part of the project, teachers completed project-developed FA instructional logs in which they described individual science lessons from their science units, the student artifacts they examined in those lessons, trends in students' ideas they observed, and details and rationales for follow-up instruction they used. Instructional logs have previously been used in similar classroom contexts (Camburn & Barnes, 2004; Forbes et al., 2015a; Rowan & Correnti, 2009) to capture classroom-specific data at scale. The FA logs were accessed in an online format and included open-ended and forced response items which provided

Table 1. School distribution of study participants.

District	School	% Eligible for free or reduced lunch	Grades represented	Number of teachers in the study	Life science topic
1	A	[a]	4th	1	The human body
			4th/5th/6th	1	The environment
	B	53.2	3rd	5	Plants
			4th	3	The environment
	C	44	6th	1	Plants
2	A	40.9	3rd/4th	3	Characteristics of living organisms
			5th/6th	3	Plants
	B	54.3	3rd	1	Characteristics of living organisms
			4th	1	The human body + Characteristics of living organisms
	C	23.8	3rd/4th	2	Characteristics of living organisms (1), Plants + The human body (1)
	D	26.6	5th/6th	1	Plants
	E	4.3	3rd	1	Characteristics of living organisms
			4th/5th	1	Plants
	F	20.6	3rd/4th	1	Plants
	G	26.1	4th	1	The human body
			6th	1	Plants
3	A	24.5	3rd	2	Plants + Characteristics of living organisms
			4th	2	The human body
4	A	28.5	6th	1	The environment

[a]Private school, data not available.

teachers with space to reflect upon and document their planning, evaluation, and follow-up instructional choices each time they used FA in teaching their science units. Teachers received training on how to complete the logs and feedback on how to provide complete responses during professional development workshops in the summer preceding data collection. The teachers were asked to complete FA logs for each science unit they taught throughout the school year and submit them electronically to the project team. For the life science units that are the focus of this study, teachers submitted an average of 3.6 FA logs with a range from 1 to 12 for a total of 137 life science FA logs. In addition, teachers submitted copies of student artifacts from the science lesson upon which the FA log responses were based. Teachers' individual FA logs and student artifacts were combined into a single file and stored electronically.

All teachers in the project also completed a multiple-choice exam to assess their knowledge of science content in the units they taught. Items were selected from the Misconceptions-Oriented Standards-based Assessment Resources for Teachers (MOSART) assessments which align with National Science Education Standards and research on students' misconceptions (Sadler et al., 2010). We used items designed for grades 5–8 to assess teachers' content knowledge at a slightly higher level than they teach. The exam consisted of 40 items, 13 of which were selected to align with life science concepts in teachers' life science units, thereby providing a measure of their knowledge of life science content they actually taught.

Data analysis

All life science FA logs were scored using a rubric we previously developed (Forbes et al., 2015a) to determine the extent to which teachers were engaging in FA practices. The rubric has been validated and shown to be reliable for scoring FA logs collected as part of this project, with an intraclass correlation coefficient of 0.808 as a measure of inter-scorer reliability (Forbes et al., 2015a). The constructs and criteria emphasized in the scoring reflect those used in other studies of teachers' practices around FA and interpreting evidence of students' thinking (e.g. Talanquer et al., 2015; Van Es, 2011). Here, each FA log and associated student work samples constitute the unit of analysis for characterizing teachers' evaluation of trends in student understanding. Each FA log question that focused on how teachers evaluated student work, along with the accompanying student work samples, was scored on 5-point scale that ranged from 0 to 4. On this scale, 4 indicated a response with an accurate and complete evaluation of students' ideas, 3 indicated a response with partially accurate/complete evaluation, 2 indicated a response with partially inaccurate or insufficient detail, 1 indicated a response that did not address the point or was completely inaccurate, and 0 was no response. The average evaluating score for each log for each teacher was calculated and used for subsequent analysis. An example of the rubric scale for one of the evaluating portions of the life science FA log is included in Appendix 1.

We used a multi-level mixed-model analysis of variance (ANOVA) (Littell, Milliken, Stroup, Wolfinger, & Schabengerger, 2006) in SPSS to determine the relationship between teachers' content knowledge, as represented by the content exam, and FA practices, as represented by the life science FA log subscores for each teacher. Since each teacher submitted multiple life science FA logs over the course of the study, we designed the statistical model to nest the individual FA log evaluating score per teacher. We then ran the mixed-model ANOVA with the independent variable as teachers' content exam score and the dependent variable as the FA log evaluating subscores for each teacher. The multi-level mixed-model ANOVA formula is $Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk}$ where i is time, j is the FA log score, k is the teacher, and ε is the error in the dependent and independent variables.

We then used qualitative analysis to examine how the teachers described how they evaluated students' ideas within the life science FA logs. The evaluating responses from all 137 life science logs were compiled and organized based on the FA log score for the Evaluating section. That is, all of the FA log responses that achieved a score of four on the Evaluating section were grouped together, and so on. First, all responses were qualitatively analyzed for the a priori code of life science content. Each response was coded as either content present or content absent (Table 2). We used this analysis to determine the percentage of life science FA logs that included content within each of the individual sections. The particular life science content within each response was highlighted for further analysis. Next, to answer our second research question, we used open-coding to identify emergent codes (Merriam, 2009) within the Evaluating subsection of the life science FA logs (Appendix 2). Third, to examine how teachers' inclusion of content varied based on

Table 2. Descriptive statistics of scoring groups.

Group	Number of teachers	% of participants	Average FA log Evaluating score
High scores	9	28	3.15
Average scores	11	34	2.96
Low scores	12	38	2.66

their content knowledge, we split the teachers into three categories based on sampling from the distribution of content exam scores. The first group were those teachers who were among the highest scorers on the content exam as indicated by a score more than a standard deviation above the mean, the second group consisted of those who scored within a standard deviation above or below the mean content exam score, and the third group consisted of teachers who were among the lowest scorers on the content exam as indicated by a score more than a standard deviation below the mean (Table 2).

The emergent codes were used in conjunction with the highlighted life science content codes to find common themes in how the teachers' discussed students' ideas both in the presence and in the absence of their discussion of content knowledge within each of these groups. Finally, we examined the student work associated with each of the FA logs to determine student understanding of the key concept. We compared our analysis of students' understanding based on the student artifacts to the teachers' summary of their evaluation of students' ideas to find themes among each of the three content knowledge groups of teachers.

Results

Content knowledge and engagement in FA

In research question one, we asked, 'To what extent does teachers' content knowledge influence their ability to engage in formative assessment within elementary life science units?' We used the content exam scores and scores from the evaluating portion of the FA log scores to examine relationships between teachers' life science content knowledge and FA practice. Descriptive statistics for the content assessment showed that the mean score on the content assessment was 74%, with a range from 50% to 95%. Descriptive statistics for the evaluating portion of the FA Logs showed a mean score of 2.90 out of a maximum possible score of 4; the teachers scored in a range from 1 to 4.

We observed a positive relationship between the teachers' life science content exam score and evaluating score (Figure 2). When the teachers scored high on the content exam score, they also scored higher for the FA practice of evaluating, as measured by the FA log scores. This relationship was analyzed with a multi-level mixed-model ANOVA and is statistically significant, $F(2, 11) = 269.7$, $p = .004$. These data were further supported by our analysis of the presence or absence of life science content within the Evaluating sections of the FA logs. We found that 95.6% (131 out of 137) of the evaluating responses included life science content. Together, this suggests that teachers used their content knowledge when they examined students' ideas and that those with greater content knowledge more effectively evaluated what students understood and where they still needed help.

Evaluating student ideas

In research question two, we asked, 'In what ways do elementary science teachers use life science content to evaluate evidence of students' thinking?' Given the statistically significant relationship observed between the content score and teachers' evaluation of student artifacts, we turned to qualitative analysis of teachers' FA logs to examine how the teachers

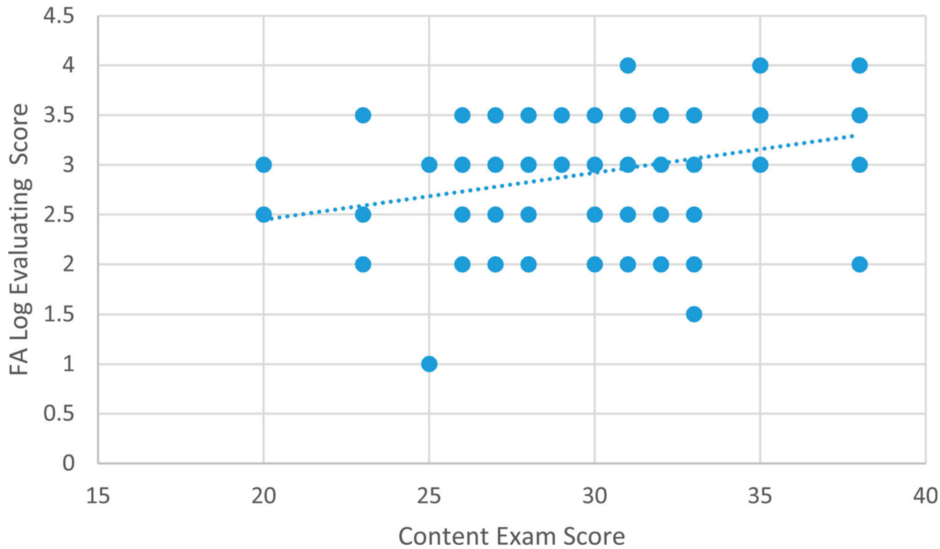


Figure 2. Alignment of *Evaluating* subscore with content exam score.

utilized their content knowledge to examine evidence of students' thinking. To illustrate trends in the qualitative findings, we examined the FA log responses within three groups based on their content exam score (i.e. those with high scores, average scores, and low scores). We present results from two representative teachers from each of the three groups. Summaries of these teachers are presented in Table 3.

High-scoring group

Teachers in the highest scoring group discussed content in depth and described both how students understood the concept and how they did not. Their evaluation aligned closely with the actual trends in the student work. Marie and Jill are both examples of teachers who achieved among the top scores on the content exam and also achieved high scores on all of their life science FA logs. Both included connection to content in their evaluation of students' ideas and accurately identified trends in student understanding. In one of Marie's lessons, students were asked to list three structures found on crayfish and to provide the function of each structure they listed. This prompt was part of a lesson in which students observed and examined the structures of crayfish and determined the function of each of the structures they observed. In her evaluation of her students' ideas, Marie said,

Table 3. Representative teachers from each of the three groups.

Group	Teacher	Life science unit	Content exam score	Average FA Log score	Average *evaluating* subscore
High scores	Marie	Structures of life	35	3.50	4
	Jill	Structures of life	35	3.46	3.6
Average scores	Heather	Plant growth & development	30	3.25	3.1
	Kelly	Plant growth & development	30	2.85	2.9
Low scores	Sandra	Plant growth & development	25	2.58	2
	Diane	Human body	23	2.74	2.5

> Most students were able to identify three structures. Some were able to identify more than one function for the structures. I did notice that some students identified incorrect structures such as fins, ears, bumps. One student listed structures but was unable to identify and functions leading me to believe that he did not understand that term. (Marie, FA Log 2)

In this example, Marie indicated that the focus of the lesson was to get student to identify three structures. Marie discussed both what students understood (they could identify at least three structures) and what they did not (some identified incorrect structures or did not identify functions of the structures). While Marie said that 'bumps' were one of the 'incorrect structures' that students identified, those structures were included as actual structures in the teacher guide for this unit. However, only one student listed bumps and did not list a correct function for them. The student said the function of bumps is 'on [its] back' (Student 1; Figure 3). Two students also listed 'ears' which Marie correctly identified as a structure not present on crayfish.

Jill evaluated student work that asked students to write about whether or not a fictional student had correctly evaluated what she saw occurring as a seed grew. The fictional student said that the first structure coming out of the seed was the stem; Jill's students were asked if they agreed or disagreed and why. They were then asked to provide an explanation for the function of roots and finally to identify the evidence the fictional student has

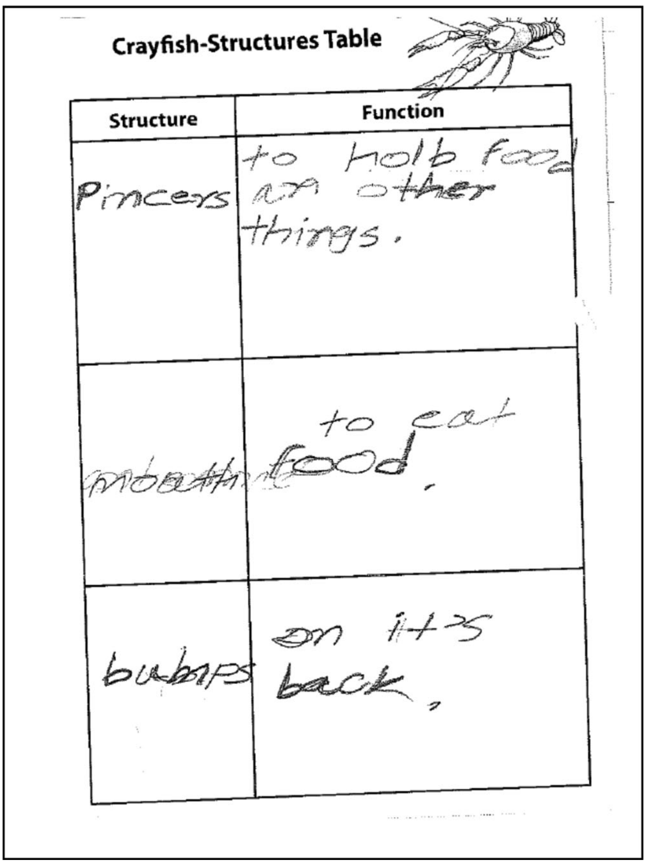


Figure 3. Student 1 example from Marie's classroom.

that the seed is living. Students completed this prompt after a lesson in which they observed germinating seeds, determined the structures seedlings have to help it grow and survive, and learned about the functions of those structures. In her evaluation of students' ideas, Jill said,

> Of the three students who did not correctly identify the first structure as the root, one agreed that it was the stem, and the other two mistakenly called it the embryo, which shows they have not grasped the idea that the embryo becomes the stem and leaves. Only one of those three students could not list a function of the roots. The children who identified more functions tended to give deeper explanations that showed they understood why the root has to develop first. Seven students could identify two or three functions of the root. (Jill, FA Log 2)

As with Marie, Jill indicated the focus of the lesson by pointing out that students were expected to identify the first structure as the root. She explained what students understood, how she knew they understood (they provided deeper explanations), and also explained specific problems that students who did not understand displayed, such as thinking the root was the stem or the embryo or not being able to list a function for the roots.

The evaluation that Jill provided aligned exactly with the student work – both what they understood and what they did not. For example, she said that 'the children who identified more functions tended to give deeper explanation that showed they understood why the root has to develop first' (Jill, FA Log 2). One student said, 'The root's job is to get water and [nutrients]' (Student 7), while a student who provided an explanation for the function said, 'the roots get water to other places like the Embryo so it can grow and make [leaves]' (Student 18). In this way, Jill correctly identified the various degrees of understanding her students had.

Average scoring group

Heather and Kelly are both examples of teachers who scored around the average score on the content exam as well as on the FA logs. As was typical of teachers in the average scoring group, both Heather and Kelly discussed life science content, but typically went into less detail than the teachers with high content exam and FA log scores.

In one of Heather's lessons, students completed a task that asked, 'From your experience of making the bee and flower models, explain how a bee pollinates the Brassica plant.' This prompt was part of a lesson on how bees and flowers are interdependent and how bees pollinate flowers. In her evaluation of students' ideas, Heather said,

> I noticed that many students did understand that the bee has to travel from one flower to another for pollination. Many students also seemed to understand the process, but forgot to use the science vocabulary terms for the parts of the flower. (They forgot to say ANTHER and STIGMA.) I didn't have many students who included that the bee goes to the flower looking for nectar. I didn't end up requiring this for a 'got it,' but I will address this in my next step strategy. (Heather, FA Log 4)

In this example, Heather mentioned the content focus of bees traveling from flower to flower for pollination. However, though she went into detail about what students struggled with, her analysis of their understanding was limited to a statement about how they 'seemed to understand the process but forgot to use the science vocabulary terms for the parts of the flower' (Heather, FA Log 4). She did not explain in any further detail how she knew that students understood the process or how she knew they forgot to use the science vocabulary terms as opposed to not understanding them.

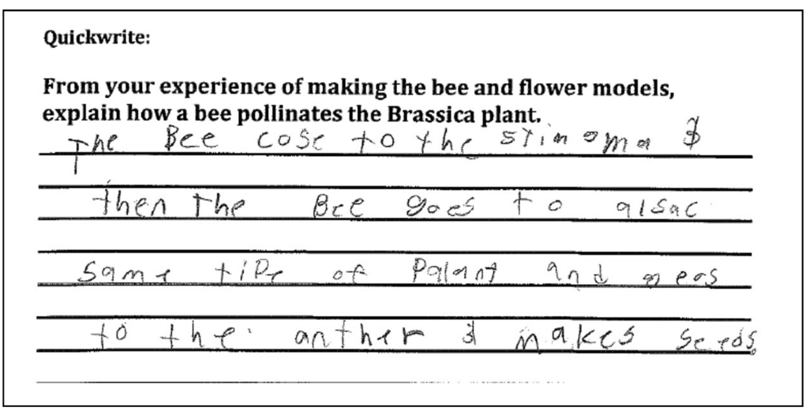
Many of the students provided vague answers and only articulated brief descriptions of the order in which pollination occurs; however, a few included how and why these phenomena take place and used appropriate lesson vocabulary. For example, a student who did use vocabulary accurately and clearly explained the process said, '1. Bee has to see the flower [petals]. 2. The Bee get pollen from the anthers. 3. The Bee [goes] to another flower. 4. Then pollen comes off with the help of the sticky stigma' (Student 4). However, a student who only focused on process steps, was vague in his answer, and did not incorporate the lesson vocabulary said, 'The Bee [goes] to the stigma and then the bee goes to also same [type] of [plant] and goes to the anther and makes seeds' (Student 9; Figure 4). In this second example, it is not clear that the student 'forgot to use the science vocabulary terms' as Heather suggested or if he did not fully understand how or why the process occurs.

Kelly's student work consisted of six steps of plant growth and asked students to number the steps 'to show the order in which things happen when a plant grows.' Students completed this prompt after a lesson on the life cycle of seeds. Regarding her evaluation of her students' ideas, she said

> All but one student got the first two steps correct with 1. being that a seed is placed in the ground and watered, and 2. The water soaks into the seed. From there on we had random orders. 14 of the 18 students got the last step correct. Some students were confused whether the plant or the root comes out of the seed first. (Kelly, FA Log 1)

Kelly indirectly mentioned the key concept by discussing the order that a plant comes out of a seed. She talked about the water soaking into the seed and whether the plant or root comes out of the seed first. However, precisely what students did understand is not clear and by mentioning the others had 'random orders,' she shows that she did not identify any further trends in student understanding.

Kelly's evaluation of student ideas was similar to that of Heather's. She mentioned that most students got the first two steps of seed growth correct and they struggled with the rest. She did not mention that six out of the 18 students had the entire sequence


Quickwrite:

From your experience of making the bee and flower models, explain how a bee pollinates the Brassica plant.

The Bee cose to the stimoma & then the Bee goes to alsac same tipe of Palant and gees to the anther & makes seeds

Figure 4. Student 9 example from Heather's classroom.

completely correct, as our analysis indicated. She did mention correctly that 'some of the students were confused whether the plant or the root comes out of the seed first' but did not mention that other students seemed confused about the role of water in the developing plant. For example, many students placed 'The swollen plant bursts out of the seed' before 'The water makes the plant and food inside the seed swell up.' Instead of recognizing this trend, Kelly suggested that the students had a 'random order' to their responses. This indicates that, rather than explore some of the specific problems students had with understanding the correct order of events in seed development, such as the importance of water, she superficially categorized students as randomly ordering events.

Low-scoring group

Sandra and Diane are examples of teachers who had both low content exam scores and low FA log scores. Similarly to other teachers in the low-scoring group, both Sandra and Diane focused on aspects of students' responses that were not based on content understanding. Sandra gave her students work that consisted of space to draw the lima bean they observed and then asked them to respond to the following prompt: 'Thinking like a scientist, observe a lima bean and describe its properties (color, shape, texture, odor, and size).' This prompt was part of a lesson in which students examined lima beans and recorded their observations. However, in her evaluation of student work, she did not include any connection to this life science topic. She said,

> We had really focused on complete sentences with punctuation as well. I was very glad to see how many kids remembered this! The class also had a long discussion about not using the word 'it' in any descriptions. I appreciated how many students remembered this also. There are 5 kids who used one descriptor that was an 'opinion'. I need to re-teach that scientific observations need to use only comparisons that are based on things that can be proven. (Sandra, FA Log 1)

Rather than on science content, Sandra's analysis consisted of focusing on her students' writing skills, such as sentence structure and punctuation, and whether or not their responses were based on their opinions or could be proven. Although these are important aspects to consider in student work, they do not provide any analysis of whether or not the students had understanding of the life science concept. Her evaluation did align with the student work, in that most students used complete sentences with punctuation. However, she mentioned that some of the students used 'one descriptor that was an "opinion"' (Sandra, FA Log 1). While this is true, it would have been difficult to the students to have had proof for their observations of something as subjective as the odor of the seed. For example, students described the odor of the lima bean as like 'a rug' (Student 1), 'salt' (Students 2 and 7; Figure 5), 'a flower stem' (Student 3), 'smelly hands' (Student 5), and 'beans' (Student 6), among other responses. A few said that it had 'no smell' at all (Students 4 and 8). However, when asked about the size of the seed, many of the students listed a precise measurement of 2 cm (Figure 5). As a result, it is unclear in this case how Sandra planned to 're-teach that scientific observations need to use only comparisons that are based on things that can be proven' (FA Log 1).

Diane's students completed student work that asked them to (a) list the three main functions of the skeleton and then to (b) choose which of those is the primary function of the skull. Students engaged in this prompt during a lesson in which they counted the

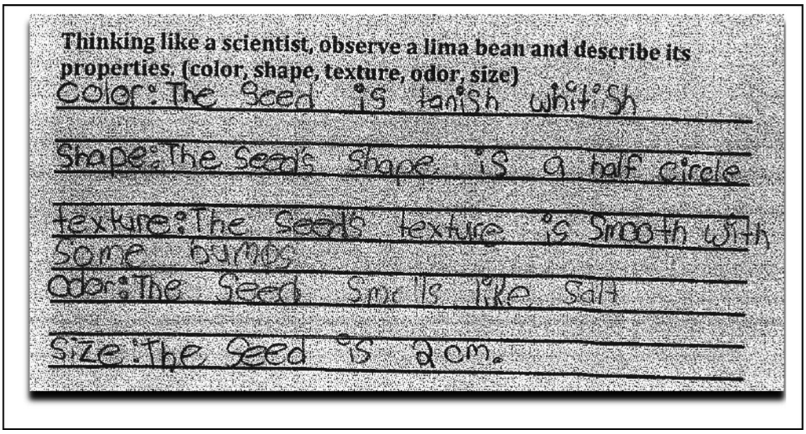

Thinking like a scientist, observe a lima bean and describe its properties. (color, shape, texture, odor, size)
Color: The Seed is tanish whit'sh
Shape: The Seed's Shape is a half circle
texture: The Seeds texture is Smooth with Some bumps
Odor: The Seed Smells like Salt
Size: The Seed is 2cm.

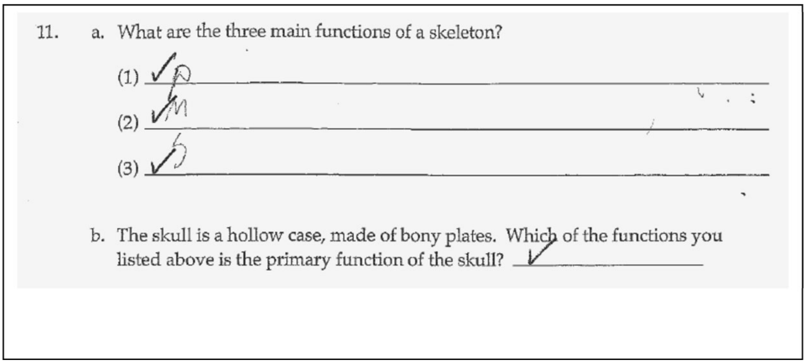
Figure 5. Student 2 example from Sandra's classroom.

bones in their body and learned about the different functions their bones served while they jumped rope. In her evaluation of their ideas, Diane said,

> Motion was used as a substitute for movement. A lot of kids skipped part b of the question even though it was pointed out to them. For the most part, students knew I was looking for them to complete the 'PMS' acronym, but some couldn't come up with the correct functions. (Diane, FA Log 2)

Here, Diane alluded to part of the key concept by mentioning movement, but did not explicitly state the entire focus that the three functions of the skeletal system are protection, movement, and support. She only mentioned that 'motion was used as a substitute for movement' though students displayed many other misunderstandings with the material. Finally, she indicated that she was looking for students to complete an acronym (PMS for protection, movement, and support) rather than show that they understood the three functions of the skeletal system.

Diane pointed out in her evaluation that 'students knew I was looking for them to complete the "PMS acronym"' (Diane, FA Log 2) and the student work did show that students chose words that began with those three letters. One student simply wrote the letters p, m, and s at the start of each line but did not write any more (Student 7; Figure 6). Others wrote three words that began with those three letters, but were not the correct functions of the skeletal system such as '[purpose], move, skeleton' (Student 8). The focus on that acronym rather than on the concept seemed to cause students to think primarily about the acronym rather than the key concept of the three functions of the skeletal system. Furthermore, though Diane was correct that many students did not finish the PMS acronym, she did not mention some of the other alternative conceptions that students had. For example, one student wrote parts of skeleton ('rib cage, skull, spine') rather than functions (Student 14). Overall, very few of Diane's students seemed to understand the concept, but this overall lack of understanding was not well represented in Diane's evaluation. In this way, Diane's focus on the acronym, rather than on the content, also seemed to cause her to miss important aspects of students' misunderstanding of the concept.


11. a. What are the three main functions of a skeleton?

(1) P

(2) M

(3) S

b. The skull is a hollow case, made of bony plates. Which of the functions you listed above is the primary function of the skull?

Figure 6. Student 7 example from Diana's classroom (*checkmarks are from Diana's analysis).

Summary of findings

Study findings show that teachers' content knowledge, as exhibited by their content exam score, is related to their effectiveness in evaluating evidence of their students' thinking. Results from quantitative analysis were supported by results from the qualitative analysis of the life science FA Log responses and the student artifacts. Overall, teachers with higher levels of life science content knowledge were able to more effectively evaluate students' ideas than teachers with lower levels of content knowledge. The teachers with higher scores on the content exam discussed both content and student understanding of the concept to a greater extent than teachers in the lower scoring groups. Furthermore, their analysis of students' ideas tended to be more scientifically accurate than those of the other teachers. Teachers in the average scoring group typically did not include as much content or identify student trends in understanding to the same extent as those in the high-scoring group. The teachers in the lower scoring group often did not include content at all in their analysis or focused on only part of the life science content focus. Their evaluation summaries tended to focus on factors other than content understanding or on tricks for remembering the content.

Synthesis and discussion

Though FA is an important instructional practice, it is still rarely a part of elementary science instruction. Empirical evidence from prior studies suggest that this may be because teachers do not have sufficient understanding of FA or the content knowledge they need to effectively implement the practice (Coffey et al., 2011; Hammer et al., 2012; Morrison, 2013; Otero & Nathan, 2008). However, limited research has focused on how teachers consider information about students' ideas they have obtained through FA and how their own understanding of disciplinary ideas contributes to this analysis of students' ideas (e.g. Bell & Cowie, 2001; Heritage et al., 2009); this is particularly true for the life sciences. Furthermore, prior research has provided mixed results on the relationship between teachers' content knowledge and effectiveness of their FA practices (e.g. Aschbacher & Alonzo, 2006; Forbes et al., 2015a). The primary contribution of this study is establishing and helping to explain the relationship between teachers' knowledge of life science content and their interpretation of students' ideas when they engage in FA.

First, study results provide empirical evidence that suggests that teachers' effectiveness in evaluating their students' ideas depends, at least in part, on their content knowledge. The teachers who had higher average scores on the content exam included more specific disciplinary content in their analysis of students' ideas than did the teachers who had lower average exam scores. This observable relationship between teachers' content knowledge and how they evaluate students' ideas shows that content knowledge is needed for instruction that is responsive to students' ideas. Past work has shown that teachers need content knowledge to effectively elicit and evaluate what students know (Aschbacher & Alonzo, 2006; Gottheiner & Siegel, 2012; Heritage et al., 2009). The results from the present study not only reinforce these prior findings but also extend them by providing evidence that the amount of content knowledge teachers have influences how well they are able to engage in these practices. Specifically, this study shows that more robust disciplinary knowledge enables teachers to engage in these practices more effectively. This not only provides empirical evidence that content knowledge is needed for teachers to successfully engage in FA, but also suggests that supporting teachers to continue to gain content knowledge throughout their teaching careers may have important impacts within science classrooms by leading to continued improvement and greater effectiveness with how teachers implement instructional practices.

While the *Next Generation Science Standards* (NGSS Lead States, 2013) emphasize particular life science core ideas, past work has shown that elementary students have a variety of alternative conceptions about life science concepts (e.g. Anderson et al., 2014; Barman et al., 2006) and teachers also have difficulty with understanding the basic biological concepts found in elementary science curricula (Kikas, 2004; Rice, 2005). As a result of teachers not fully understanding the content or how students understand that content, FA strategies often tend to focus on generic strategies teachers can use to address student thinking rather than on addressing the disciplinary concepts or the reasoning behind students' answers (Coffey et al., 2011). The findings from this study provide clear evidence of the link between content knowledge and FA practice. Furthermore, findings from this study show that improving teachers' content knowledge is an important step in improving how teachers engage with students' ideas and, as such, can help teachers begin to more effectively consider their students' understanding and scientific reasoning.

Engaging in FA and evaluating students' ideas can also help teachers begin to see themselves as disciplinary thinkers by considering their students' reasoning (Ash & Levitt, 2003; Kazemi & Franke, 2004). However, this connection of content knowledge to ability to evaluate student's ideas may be context-dependent. Although in this study we found that content knowledge of life science concepts was an important aspect of teachers engaging effectively in evaluating students' ideas, a previous study found that content knowledge was not a significant factor in teachers' evaluation of students' understanding of earth science concepts (Forbes et al., 2015a). This difference may point to other important factors in teachers' FA practices, such as their familiarity and experience with particular curriculum materials or how well the chosen student work aligned with teachers' preexisting ideas of how to evaluate students' ideas independently of their own content knowledge. Teachers also had more experience using FA practices in the present study than in the previous study, which tentatively suggests that teachers may need to reach a threshold of understanding and using FA practices before subject matter becomes a significant factor in implementation of those practices. Examination of the particular differences between how teachers engaged in these two disciplines is beyond the scope of this study, but does suggest that further research is needed to explore the contributing factors to integration of teachers' disciplinary knowledge and evaluation of students' ideas.

Second, study results also illustrate *how* teachers leverage content knowledge to evaluate students' ideas. Teachers with higher scores on the content exam evaluated and discussed both content and student understanding in a way that those with lower scores did not. This aligns with past work that has shown that teachers who have more robust disciplinary knowledge tend to enact teaching practices that are more supportive of student learning, such as FA (Aschbacher & Alonzo, 2006; Falk, 2011; Heritage et al., 2009; Hill et al., 2005). However, in this study, the teachers with higher content scores did more than just incorporate more content into their evaluation. Study findings show that teachers also more specifically considered both what students did and did not understand. Thus, other dynamics, such as the particular student work teachers choose to evaluate or how teachers choose what to look for in the student work, are likely also factors in how well teachers are able to engage in FA and will be important for future work to further explore. Understanding these dynamics and how teachers differentially engage in these practices is important because the use of effective FA has been previously tied to significant student learning gains in sciences classes (Ruiz-Primo & Furtak, 2006).

In stark contrast to how the higher scoring teachers incorporated content, the lower scoring teachers either did not include content in their evaluation of students' ideas, focused on evaluating students' performance of tasks not related to life science concepts, or relied on the presence of tricks for remembering the content rather than on evidence of robust disciplinary understanding. How these teachers evaluated student's ideas is in line with previous work that has shown that teachers may rely on 'get it or don't' conceptions of student understanding or on the presence of particular vocabulary words rather than evidence of understanding of the contexts in which those vocabulary words are used (Forbes et al., 2015a; Otero & Nathan, 2008). Furthermore, the disconnect for lower scoring teachers between content knowledge and evaluating ideas aligns with work by Heritage et al. (2009) that has shown that teachers' lack of deep understanding of disciplinary ideas contributes to the difficulties teachers have with determining what students know and deciding what to do with that information. To that end, teachers with less content knowledge may rely on the curriculum to help them understand what disciplinary content to evaluate (Forbes et al., 2015a; Sabel et al., 2015), but that same curriculum may not help them decide what next instructional steps to take. Therefore, teachers with less content knowledge are at a significant disadvantage when it comes to supporting their students in moving beyond their current understanding and onto more robust understanding of science concepts. In order to move beyond relying on general instructional strategies (Gottheiner & Siegel, 2012; Jacobs et al., 2010; Levin et al., 2009) and toward teaching strategies that specifically connect instruction to the particular ideas and alternative conceptions students have, teachers will need to learn both content and pedagogical knowledge and will need support to incorporate both into their classroom environments.

Implications and conclusion

Past work has focused on the two separate, but connected bodies of research: elementary teachers' use of FA for science (e.g. Aschbacher & Alonzo, 2006; Buck & Trauth-Nare, 2009; Forbes et al., 2015a; Hammer et al., 2012; Morrison, 2013; Otero & Nathan, 2008) and teaching and learning in the life sciences (e.g. Anderson et al., 2014; Barman et al., 2006; Forbes et al., 2015b; Kikas, 2004; Rice, 2005; Sabel et al., 2015); however, little work has focused on the intersection between teaching and learning life science content and FA practices. The work presented here contributes to this intersection of research by providing evidence of how teachers use disciplinary knowledge when they engage in FA for life science instruction and how their knowledge of life science concepts contributes to their ability to evaluate students' ideas. As such, this study has the potential to inform elementary life science instructional practice and the design of both professional development and curriculum materials for elementary teachers and elementary classrooms.

The findings here suggest that more support is needed in both preservice and inservice education for teachers to develop more discipline-specific content knowledge and that this support should be sustained throughout teachers' careers as more robust content knowledge leads to more effective instructional practice. Teacher education and professional development experiences should include meaningful connections to disciplinary content, opportunities for teachers to develop content knowledge, and support for both preservice and inservice teachers to integrate that knowledge into effective teaching strategies that respond to students' ideas. Some examples of preservice teacher education courses that integrate life science content with elementary teaching methods currently exist (Forbes et al., 2015b; Friedrichsen, 2001; Haefner et al., 2006; Sabel et al., 2015), and the professional development program in which this study was based provides a similar model for professional development. However, such programs for teacher development are the exception rather than the norm. More work is needed not only to develop and implement these types of courses and professional development experiences, but also to study the impact of their design on intended outcomes.

Teachers will also need additional support to develop pedagogical skills related to both content and evaluating students' ideas in preservice and inservice education. Experiences that integrate instruction on implementing FA strategies have been shown to be successful in the past and can serve as models for future interventions (e.g. Buck et al., 2010; Forbes et al., 2015b; Levin et al., 2009; Otero & Nathan, 2008; Sabel et al., 2015; Talanquer et al., 2013). Teachers must be supported to learn the importance of FA practices and how to implement those practices in their classrooms, as well as the content and pedagogical knowledge to effectively engage in considering the scientific accuracy of their students' ideas (Talanquer et al., 2013, 2015; Van Es, 2011). This will also include support for understanding how students learn, how students' ideas align with the content of the lessons, and how they can incorporate students' ideas into their instruction (Ball et al., 2008; Heritage et al., 2009).

In addition to courses or professional development programs, educative curriculum materials can provide another source of support for teachers in the particular disciplinary domains they teach. In order to most effectively consider students' ideas and understanding, curriculum materials should also include indications of where students may have problems with the content and ideas for how teachers can extend the curriculum to address students' ideas in follow-up instruction. This is particularly true for teachers with less robust content knowledge who may not have access to resources to support them in determining next instructional steps to best address the issues they find in their evaluation of students' ideas. Teachers should also be provided with direction to other resources where they can extend their own content knowledge about a particular topic, what research has shown on students' ideas about that topic, and effective instructional strategies for addressing common alternative conceptions students may have.

Combined, this support for both life science content knowledge and FA practices can help teachers begin to move past the 'get it or don't' assessment of students' ideas or overreliance on the presence of particular vocabulary words or curriculum-specific language (Forbes et al., 2015a; Otero, 2006) and move to a more robust evaluation of students' disciplinary thinking. Integration of these pieces is needed to prepare teachers to create effective science learning environments that are responsive to students' ideas. Engaging in FA can also help teachers in their own understanding as they consider and respond to students' ideas about the content (Ash & Levitt, 2003; Jacobs et al., 2010).

While others have previously advocated for the importance of discipline-specific knowledge for effective implementation of pedagogical practices (e.g. Falk, 2011; Hill et al., 2005), little past research has explored relationships between teachers' content knowledge and their evaluation of students' understanding. Our study leverages similar theoretical and analytical perspectives on FA and diagnosis of students' ideas (Talanquer et al., 2013, 2015; Van Es, 2011) to provide empirical evidence of the importance of this relationship in the life science domain. This study is limited in that it investigates teachers using a particular subset of curriculum materials. Future work should extend this examination of teachers' engagement with life science topics to include additional sources of life science curriculum materials. Furthermore, while previous work has shown a connection between teachers' content knowledge and students' learning gains, this particular study did not examine potential effects of teachers' content knowledge on students' learning of life science concepts that were taught. As such, future work should examine the potential relationship between teachers' content knowledge, their use of FA strategies, and students' learning of life science concepts.

Acknowledgements

We thank Laura Zangori, Tina Vo, and anonymous reviewers for their help in thinking about these issues and their thoughtful comments on earlier versions of this paper. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is funded by the Spencer Foundation, the Iowa Board of Regents, the State of Iowa, and the Title IIA Improving Teacher Quality: State Grants Program [grant number 201400084].

Notes on contributor

Jaime Sabel completed doctoral work at the University of Nebraska-Lincoln and will be an Assistant Professor of Biological Sciences at the University of Memphis beginning in the fall of 2016.

Cory Forbes is an Associate Professor of Science Education and Coordinator of the Science Literacy Initiative at the Institute for Agriculture and Natural Resources at the University of Nebraska-Lincoln.

Leslie Flynn is a Clinical Assistant Professor in the Department of Teaching and Learning at the University of Iowa.

References

Anderson, J. L., Ellis, J. P., & Jones, A. M. (2014). Understanding early elementary children's conceptual knowledge of plant structure and function through drawings. *CBE – Life Sciences Education*, *13*(3), 375–386. doi:10.1187/cbe.13-12-0230

Aschbacher, P., & Alonzo, A. (2006). Examining the utility of elementary science notebooks for formative assessment purposes. *Educational Assessment*, *11*, 179–203.

Ash, D., & Levitt, K. (2003). Working within the zone of proximal development: Formative assessment as professional development. *Journal of Science Teacher Education*, *14*(1), 23–48.

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.

Barman, C. R., Stein, M., McNair, S., & Barman, N. S. (2006). Students' ideas about plants and plant growth. *The American Biology Teacher*, *68*(2), 73–79.

Bell, B., & Cowie, B. (2001). The characteristic of formative assessment in science education. *Science Education*, *85*(5), 536–553.

Buck, G.A., & Trauth-Nare, A.E. 2009. Preparing teachers to make the formative assessment process integral to science teaching and learning. *Journal of Science Teacher Education*, *20*, 475–494.

Buck, G. A., Trauth-Nare, A., & Kaftan, J. (2010). Making formative assessment discernable to pre-service teachers of science. *Journal of Research in Science Teaching*, *47*(4), 402–421.

Camburn, E., & Barnes, C. A. (2004). Assessing the validity of a language arts instruction log through triangulation. *Elementary School Journal*, *105*(1), 49–73.

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching*, *48*, 1109–1136.

Cresswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Los Angeles, CA: Sage.

Donovan, M. S., & Bransford, J. D. (Eds.). (2005). *How students learn: Science in the classroom*. Washington, DC: National Academies Press.

Falk, A. (2011). Teachers learning from professional development in elementary science: Reciprocal relations between formative assessment and pedagogical content knowledge. *Science Education*, *96*, 265–290.

Forbes, C.T., Sabel, J.L., & Biggers, M. (2015a). Elementary teachers' use of formative assessment to support students' learning about water-earth materials interactions. *Journal of Geoscience Education*, *63*(3), 210–221. doi: 10.5408/14-063.1.

Forbes, C.T., Sabel, J.L., & Zangori, L. (2015b). Integrating life science content and instructional methods in elementary teacher education. *American Biology Teacher*, *77*(9), 5–11. doi: 10.1525/abt2015.77.9.2.

Friedrichsen, P. M. (2001). Moving from hands-on to inquiry-based: A biology course for prospective elementary teachers. *The American Biology Teacher*, *63*(8), 562–568.

Gottheiner, D.M., & Siegel, M. A. (2012). Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education*, *23*, 531–557.

Haefner, L. A., Friedrichsen, P. M., & Zembal-Saul, C. (2006). Teaching with insects: An applied life science course for supporting prospective elementary teachers' scientific inquiry. *The American Biology Teacher*, *86*(4), 206–212.
Hammer, D., Goldberg, F., & Fargason, S. (2012). Responsive teaching and the beginnings of energy in a third grade classroom. *Review of Science, Mathematics, and ICT Education*, *6*, 51–72.
Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues Practice*, *28*(3), 24–31.
Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*, 371–406.
Jacobs, V. R., Lamb, L. L. C., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, *41*(2), 169–202.
Kazemi, E., & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education*, *7*, 203–235.
Keeley, P. (2005). *Science curriculum topic study*. Thousand Oaks, CA: Corwin Press.
Kennedy, M. M. (1998). Education reform and subject matter knowledge. *Journal of Research in Science Teaching*, *35*(3), 249–263.
Kikas, E. (2004). Teachers' conceptions and misconceptions concerning three natural phenomena. *Journal of Research in Science Teaching*, *41*(5), 432–448.
Levin, D. M., Hammer, D., & Coffey, J. E. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education*, *60*(2), 142–154.
Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). SAS for mixed models (2nd ed.). Cary, NC: SAS Institute, Inc.
Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Josey-Bass.
Morrison, J. A. (2013). Exploring exemplary elementary teachers' conceptions and implementation of inquiry science. *Journal of Science Teacher Education*, *24*, 573–588.
NGSS Lead States. (2013). *Next generation science standards. For states, by states*. Washington, DC: The National Academies Press.
Otero, V. K. (2006). Moving beyond the "get it or don't" concept of formative assessment. *Journal of Teacher Education*, *57*(3), 247–255.
Otero, V. K., & Nathan, M. J. (2008). Preservice elementary teachers' views of their students' prior knowledge of science. *Journal of Research in Science Teaching*, *45*(4), 497–523.
Rice, D. C. (2005). I didn't know oxygen could boil! What preservice and inservice elementary teachers' answers to 'simple' science questions reveals about their subject matter knowledge. *International Journal of Science Education*, *27*(9), 1059–1082.
Rowan, B. and Correnti, R. 2009. Studying reading instruction with teacher logs: Lessons from the Study of Instructional Improvement. *Educational Researcher* , *38*, 120–131.
Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment*, *11*, 205–235.
Sabel, J.L., Forbes, C.T., & Zangori, L. (2015). Promoting prospective elementary teachers' learning to use formative assessment for life science instruction. *Journal of Science Teacher Education*, *26*(4), 419–445. doi: 10.1007/s10972-015-9431-6.
Sadler, P. M., Coyle, H., Miller, J. L., Cook-Smith, N., Dussault, M., & Gould, R. R. (2010). The astronomy and space science concept inventory: Development and validation of assessment instruments aligns with the K-12 National science standards. *Astronomy Education Review*, *8*(1), 1–26.
Talanquer, V., Bolger, M., & Tomanek, D. (2015). Exploring prospective teachers' assessment practices: Noticing and interpreting student understanding in the assessment of written work. *Journal of Research in Science Teaching*, *52*(5), 585–609.
Talanquer, V., Tomanek, D., & Novodvorsky, I. (2013). Assessing students' understanding of inquiry: What do prospective science teachers notice?. *Journal of Research in Science Teaching*, *50*(2), 189–208.
Van Es, E. A. (2011). A framework for learning to notice student thinking. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 134–151). New York: Taylor and Francis.

Appendix 1

Table A1. Example from Rubric for Lesson Log Scoring.

Q3: What did you notice in the student work you reviewed? Did you observe any misconceptions? What trends in students' understanding of the key concept did you see?

Score	Description	Example
4	Response includes an accurate description of students' understanding	'Most students correctly identified the mistakes. A few identified the mistakes, but did not know how to correct it. For example, they know there was a mistake in the arm bones, but they did not know that it was that the humerus and the radius/ulna were switched. Also, quite a few students said that the radius and ulna should NOT be twisted like it Shows in the picture if the thumb if the left arm has the thumb pointed in.' (FA Log 10.8.3)
3	Response includes a mostly accurate description of students' understanding	'Many students focused on the parts of the plant rather than including needs. Most students included mushrooms and many included mold in the list of plants. Many didn't include sun but most listed water and soil as needs.'(FA Log 5.7.3)
2	Response includes a mostly inaccurate description of students' understanding	'Misconceptions: Any liquid would soften the seed coat, plants need the light not warmth from the sun, seeds need food to germinate (not realizing that the cotyledon serves as the food at first.' (FA Log 35.12.3)
1	Response does not address or inaccurately describes students' understanding	'We had really focused on complete sentences with punctuation as well. I was very glad to see how many kids remembered this! The class also had a long discussion about not using the word "it" in any descriptions. I appreciated how many students remembered this also. There are 5 kids who used one descriptor that was an "opinion." I need to re-teach that scientific observations need to use only comparisons that are based on things that can be proven.' (FA Log 67.10.3)
0	No Response	

Appendix 2

Table A2. Codes Used in Qualitative Analysis.

Code Category	Code	Example
A priori	Life science content present	'A few students noted that not only does the fruit contain the seeds, but that it has a covering (shield) to protect the seed.' (FA Log 9.13.3)
	Life science content absent	'We had really focused on complete sentences with punctuation as well. I was very glad to see how many kids remembered this! The class also had a long discussion about not using the word "it" in any descriptions. I appreciated how many students remembered this also. There are 5 kids who used one descriptor that was an "opinion." I need to re-teach that scientific observations need to use only comparisons that are based on things that can be proven.' (FA Log 64.10.3)
Emergent	Specificity/completeness in student's answers	'Overall, most kids who still need help, just need to be more specific in their answers.' (FA Log 43.2.3)
	Correctness in students' answers	'The ones that did not understand the concept were generally the ones that did not answer the question correctly.' (FA Log 45.4.3)
	Teachers' focus on terms	'Students knew the process but did not use key vocabulary words.' (FA Log 84.12.3)
	Teachers' need for modifications to lesson	'Upon reviewing these answers, I think I need to clarify that perhaps the plants may not die, but instead just be unhealthy.' (114.10.3)