



Western Michigan University ScholarWorks at WMU

Scientific Literacy and Cultural Studies Project

Mallinson Institute for Science Education

2014

Pedagogy of Science Teaching Tests: Formative assessments of science teaching orientations

William W. Cobern

Western Michigan University, bill.cobern@wmich.edu

David Schuster

Western Michigan University, david.schuster@wmich.edu

Betty Adams

Western Michigan University, b.adams@wmich.edu

Brandy Ann Skjold

Western Michigan University, b.skjold@wmich.edu

Ebru Zeynep Mugaloglu

Bogazici University

See next page for additional authors

Follow this and additional works at: http://scholarworks.wmich.edu/science_slcsp

 Part of the [Science and Mathematics Education Commons](#)

WMU ScholarWorks Citation

Cobern, William W.; Schuster, David; Adams, Betty; Skjold, Brandy Ann; Mugaloglu, Ebru Zeynep; Bentz, Amy; and Sparks, Kelly, "Pedagogy of Science Teaching Tests: Formative assessments of science teaching orientations" (2014). *Scientific Literacy and Cultural Studies Project*. 51.

http://scholarworks.wmich.edu/science_slcsp/51

This Article is brought to you for free and open access by the Mallinson Institute for Science Education at ScholarWorks at WMU. It has been accepted for inclusion in Scientific Literacy and Cultural Studies Project by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.



Authors

William W. Cobern, David Schuster, Betty Adams, Brandy Ann Skjold, Ebru Zeynep Mugaloglu, Amy Bentz, and Kelly Sparks

This article was downloaded by: [Western Michigan University]

On: 17 June 2014, At: 11:10

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Science Education

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tsed20>

Pedagogy of Science Teaching Tests: Formative assessments of science teaching orientations

William W. Cobern^a, David Schuster^a, Betty Adams^a, Brandy Ann Skjold^a, Ebru Zeynep Muğaloğlu^b, Amy Bentz^a & Kelly Sparks^c

^a The Mallinson Institute for Science Education, Western Michigan University, Kalamazoo, MI, USA

^b Science Education, Boğaziçi University, Turkey

^c Teacher Education Department, University of Southern Indiana, Evansville, IN, USA

Published online: 16 May 2014.

To cite this article: William W. Cobern, David Schuster, Betty Adams, Brandy Ann Skjold, Ebru Zeynep Muğaloğlu, Amy Bentz & Kelly Sparks (2014): Pedagogy of Science Teaching Tests: Formative assessments of science teaching orientations, International Journal of Science Education, DOI: [10.1080/09500693.2014.918672](https://doi.org/10.1080/09500693.2014.918672)

To link to this article: <http://dx.doi.org/10.1080/09500693.2014.918672>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Versions of published Taylor & Francis and Routledge Open articles and Taylor & Francis and Routledge Open Select articles posted to institutional or subject repositories or any other third-party website are without warranty from Taylor & Francis of any kind, either expressed or implied, including, but not limited to, warranties of merchantability, fitness for a particular purpose, or non-infringement. Any opinions and views expressed in this article are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor & Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages,

and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

It is essential that you check the license status of any given Open and Open Select article to confirm conditions of access and use.

Pedagogy of Science Teaching Tests: Formative assessments of science teaching orientations

William W. Cobern^{a*}, David Schuster^a, Betty Adams^a,
Brandy Ann Skjold^a, Ebru Zeynep Muğaloğlu^b, Amy Bentz^a
and Kelly Sparks^c

^a*The Mallinson Institute for Science Education, Western Michigan University, Kalamazoo, MI, USA;* ^b*Science Education, Boğaziçi University, Turkey;* ^c*Teacher Education Department, University of Southern Indiana, Evansville, IN, USA*

A critical aspect of teacher education is gaining pedagogical content knowledge of how to teach science for conceptual understanding. Given the time limitations of college methods courses, it is difficult to touch on more than a fraction of the science topics potentially taught across grades K-8, particularly in the context of relevant pedagogies. This research and development work centers on constructing a formative assessment resource to help expose pre-service teachers to a greater number of science topics within teaching episodes using various modes of instruction. To this end, 100 problem-based, science pedagogy assessment items were developed via expert group discussions and pilot testing. Each item contains a classroom vignette followed by response choices carefully crafted to include four basic pedagogies (didactic direct, active direct, guided inquiry, and open inquiry). The brief but numerous items allow a substantial increase in the number of science topics that pre-service students may consider. The intention is that students and teachers will be able to share and discuss particular responses to individual items, or else record their responses to collections of items and thereby create a snapshot profile of their teaching orientations. Subsets of items were piloted with students in pre-service science methods courses, and the quantitative results of student responses were spread sufficiently to suggest that the items can be effective for their intended purpose.

Keywords: *Formative assessment; Conceptual development; Teacher development*

*Corresponding author. The Mallinson Institute for Science Education, Western Michigan University, Kalamazoo, MI, USA. Email: bill.cobern@wmich.edu

Introduction

One of the most important goals for science teacher education is that prospective teachers acquire knowledge of science teaching pedagogy, and toward this end, teacher education programs include a science teaching methods course. Such courses commonly include readings, observations of science teachers (live or by film), micro-teaching, and the practice writing of science lesson plans. In the course, the assessment of pre-service teachers' acquisition of pedagogical knowledge of science teaching is commonly done by evaluating the science lessons they have created as well as their micro-teaching.

The range of science topics typically found in K-8 curricula (see, for example, the *National science education standards* (National Research Council, 1996) and *A framework for K-12 science education*) is very broad. Given the time and scope limitations of a college course in science teaching methods, observing practicing teachers or constructing lesson plans can only touch on a fraction of the topics taught and the possible pedagogical strategies used across the K-8 grades. Consider the pedagogical example of teaching science by inquiry. Abd-El-Khalick et al. (2004) acknowledge that 'research has consistently indicated that what is enacted in classrooms is mostly incommensurate with visions of inquiry put forth in reform documents, past . . . and present' (p. 398). Despite this, in any science methods course, at best students will see inquiry instruction applied to a few science topics, and apply it themselves to even fewer. Based on such limited exposure, feedback, and experience, one has to wonder how far these teachers will be able to operationalize and extend their general knowledge of inquiry instruction to multiple specific topics. We know from cognitive studies that the transfer of knowledge best occurs when the learner sees the knowledge applied in various situations (Donovan & Bransford, 2005). Would this not apply as well to the acquisition of science teaching pedagogy? Thus, one would expect the development of knowledge about inquiry instruction and the ability to implement it to be related to the number of different teaching situations in which inquiry has been encountered. Obviously quantity of exposure has to be balanced by the quality and appropriateness of the teaching situations that the student sees. Multiple exposures to poor teaching are not helpful. Moreover, exposure to even several good teaching situations does not reduce the essential need for application and practice; combining quality and quantity is optimal.

Our research and development work began with the recognition that time constraints within a typical methods course or professional development program severely limit exposure to contextualized science teaching pedagogies. Only a few films of teachers' actual practice can be viewed. Only a few lessons can be constructed or enacted by the student. We therefore sought a supplemental strategy for the teaching and learning of science pedagogy that would cover a much broader range of implementation. An efficient and effective method for doing this could have significant instructional impact as well as research utility. A new kind of assessment for use in a methods course, comprising a broad range of realistic, problem-based items of suitable nature, has the potential to achieve these goals.

The Nature of the Assessment

The assessment draws on ideas from worked examples in the physical sciences and mathematics (Heller & Heller, 2010; Johnson, 2001), from Assessment for Learning (Black & Wiliam, 2006), and from problem-based learning (PBL) (Capon & Kuhn, 2004). There is a long history of assigning problems and questions to students in the physical sciences and mathematics. The idea is that to gain expertise, one must practice during learning by applying one's knowledge to a diverse range of examples. As noted by Capon and Kuhn (2004, p. 61), the 'integration of new information with existing knowledge structures can be activated' by experience with multiple examples. Recent work on the efficacy of retrieval tasks for enhancing learning suggests why such practice is effective (Karpicke & Blunt, 2011). Of particular interest to us is the formative use of worked examples as Assessment for Learning (Black, Harrison, Lee, Marshall, & Wiliam, 2003). We suggest that the preparation of science teachers can be substantially enhanced by giving students the experience of working through science pedagogy problems, that is, by presenting students with classroom situations where they have either to make an instructional decision or evaluate a particular instructional approach adopted by a teacher. Assessment items for this purpose are essentially problems or questions involving alternative pedagogical approaches to a given teaching situation. Working through such items with students operates as a scaffold for novices' current lack of schemas and promotes effective instruction based on active engagement with sample cases. Certain studies suggest that students learn effectively from suitable worked teaching examples, rather than just attempting many problems on their own (Cooper & Sweller, 1987; Maloney, 1994; Sweller & Cooper, 1985; Trafton & Reiser, 1993; Ward & Sweller, 1990).

We thus undertook the creation of problem sets for use in methods courses by drawing upon ideas from PBL, among others. PBL is an approach widely used in medical education (Albanese & Mitchell, 1993; Peterson, 1997) and more recently adopted in science teacher education (Dean, 1999; Ngeow & Kong, 2001; Wang, Thompson, Shuler, & Harvey, 1999). In our model, PBL presents pre- and in-service teachers with a practical teaching problem in the form of a realistic scenario, vignette, or case. Our model uses realistic K-8 science teaching situations where each item/problem begins with a brief classroom vignette followed by a question and set of response choices representing alternative pedagogies.

Basing a pedagogy-of-science assessment on PBL has advantages. First, given that item vignettes depict actual classroom teaching situations, the assessment is situated and more authentic. Our vignettes are like mini-cases. Regarding case methods, Bencze, Hewitt, and Pendretti (2001, p. 196) note that

formal teacher education cannot prepare student teachers for all possible teaching and learning situations. Case methods offer an excellent opportunity along these lines [of becoming a reflective practitioner] because of their potential for challenging student teachers to analyse authentic instructional scenarios.

Shulman (1992) further makes the case that ‘cases may reduce the problems of transfer because they simulate the way in which the most effective forms of learning are situated in specific contexts and circumstances’ (p. 24). A second advantage of problem-based assessments is that they do not lapse into measurement of rote memory or generalities about pedagogy. Each item specifically requires either application or evaluation, in terms of Bloom’s taxonomy (Anderson & Krathwohl, 2001) involving particular instructional situations. Successful application and evaluation require that one understands science pedagogy and its use in particular science content areas at the appropriate grades.

Once a set of problem-based items on science pedagogy for specific teaching examples is available, the items can be used to help students develop and reinforce a usable understanding of the general principles they are learning. This article reports on the design, development, and testing of a large collection of selected-response items involving alternative pedagogies that provide a practical means for the presentation of multiple science teaching events. They are freely available in various formats at: <http://www.wmich.edu/science/inquiry-items/index.html>. The items can be used for supplemental, instructional, formative, and research purposes.

Science Teaching Approaches

As noted by the *A framework for K-12 science education* (National Research Council [NRC], 2011) and *Ready, set, science!* (Michaels, Shouse, & Schweingruber, 2007), there is more than one type of learning objective in science education to be addressed by suitable science pedagogies. *Ready, set, science!* suggests four major interconnected objectives: understanding scientific explanations, generating scientific evidence, reflecting on scientific knowledge, and participating productively in science. Our focus is mainly (though not entirely) directed toward the objective of ‘understanding scientific explanations’. This emphasis is based on the perspective that understanding scientific concepts and explanations is fundamental and is needed to achieve the other objectives. Furthermore, it has been our experience that both pre- and in-service teachers are more inclined toward non-inquiry instruction for the teaching of basic scientific concepts than, for example, for teaching to the objective of ‘generating scientific evidence’. Data reported by Banilower and Smith (2013) show the persistent use of non-inquiry forms of instruction in the USA. Interestingly, news reports in the USA on the Next Generation Science Standards (NGSS, 2013) often cite NGSS as a response to science teaching for passive learning (Leone-Cross, 2013; Mervis, 2013). Hence, we see a substantial need for an effective formative assessment strategy with respect to teaching and learning important science content. Nevertheless, within this primary focus, it will be seen that our assessments bring in aspects of the other objectives in a natural way as appropriate.

Even within a focus on conceptual understanding of science content, one finds in practice a great variety of science teaching strategies. Nevertheless, at a basic level, most of these strategies are variants of two fundamental epistemic modes of instruction: either some form of inquiry instruction or some form of direct instruction.

Students either develop science content knowledge in an inquiry-based fashion, via suitable guided explorations of phenomena, or the science content is presented and explained directly to them, commonly followed by practical confirmation. In the National science education standards for an increasing number of countries and in the science education literature, it is inquiry that is typically advocated (Cavanagh, 2005; European Commission, 2011). In the USA, the 2011 National Research Council document, *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*, adopts the terminology of scientific practices, but retains a primary focus on inquiry instruction. Note also that inquiry instruction for the understanding of scientific content is not of a single form. According to the National Research Council (2000b), there are various degrees of inquiry instruction, which can vary from instruction strongly guided by the teacher to a very student-centered open inquiry.

The historic counterpoint to inquiry-based instruction is direct instruction in its various forms. Direct instruction tends to be typically but misleadingly portrayed as teaching-by-telling with passive reception (Cavanagh, 2004; Schroeder, Scott, Tolson, Huang, & Lee, 2007; Thomson & Gregory, 2013; Wise, 1996). However, we have eschewed any simplistic or absolutist contrasts between direct and inquiry instruction in favor of the more flexible concept of teaching orientations. In the science education literature, teaching orientation can refer to a number of different ideas. The orientation dimension we focus upon in our work is somewhat less complex than the full ‘sets of beliefs that teachers hold’ (Friedrichsen, Van Driel, & Abell, 2011, p. 372), and at the same time reflects aspects of both the National Research Council descriptions and the dimensions of influential teacher conceptions that Lotter, Harwood, and Bonner (2007) describe. It provides a spectrum of common orientations along a direct–inquiry continuum, described in greater detail in the next section. Within this set of orientations one usually finds that a more direct orientation is more teacher-directed, while an inquiry orientation is a more learner-directed approach (NRC, 2000b, p. 29). A more direct orientation aligns more with information transmission than independent thinking (Lotter et al., 2007). Our orientation spectrum also encompasses many of the orientation types proposed by Magnusson, Krajcik, and Borke (1999), including process, didactic, activity-driven, discovery, inquiry, and guided inquiry, in a form aligned closely with our project goals.

Science Teaching Orientation Spectrum

For our purposes, science teaching expertise translates to knowledge of content, practices, scientific inquiry, science pedagogy, and inquiry pedagogy. On a practical level, teaching expertise includes being able to recognize, create, and follow teaching plans that are in accord with best practices in science education. This demanding combination constitutes pedagogical content knowledge for science teaching. One of the first choices that a teacher will make, either implicitly or explicitly, is whether to present and explain scientific concepts and principles directly to the students, or

Fundamental epistemic mode	Variant for each mode	Operationalized description (abbrev.)
Science presented as factual knowledge... "Ready-Made-Science"	1 Didactic Direct	Teacher presents and explains science content directly... illustrates with example or demo. No student activities.
	2 Active Direct	Teacher presents and explains science content directly... students actively engage in verification/confirmation.
Science as developed by process of scientific inquiry... "Science-in-the-Making"	3 Guided Inquiry	Students actively explore phenomenon or idea with teacher guidance toward desired science content.
	4 Open Inquiry	Students actively explore phenomenon or idea as they choose... teacher facilitates process but does not prescribe.

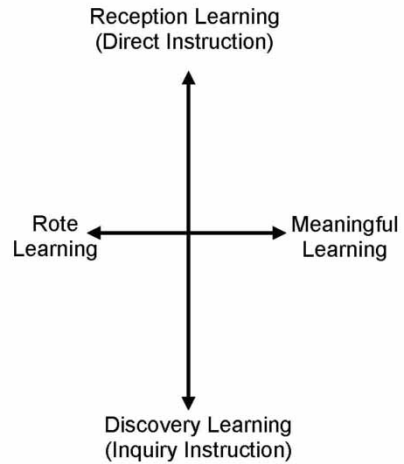


Figure 1. The pedagogical foci for item responses and Ausubel's axes

have the students play some role in exploring and finding out the scientific explanations themselves. This most basic distinction (between what are commonly classed as *direct* and *inquiry* modes of instruction) is represented in the left column of the table in Figure 1. The middle column breaks this dimension down further into two variations on each fundamental mode, thus providing four common teaching orientations. These are briefly described in the right column. These instructional mode options represent the epistemic decisions a teacher will need to make (consciously or subconsciously) in designing and implementing science instruction to teach scientific content for any given topic (the concepts, principles, relationships, and explanations).

This spectrum of epistemic approaches is extremely valuable for constructing assessment items based on a consistent set of instructional approaches. The four categories along the instructional spectrum are labeled as didactic direct, active direct, guided inquiry, and open inquiry. These are not to be seen as rigid compartments, but as a useful way of broadly characterizing instructional approaches found in practice. It is likely that a variation exists in exactly how people feel each instructional type should be defined, but the brief descriptions give the basic nature of each and make the distinctions between them clear. Beyond this, instructional method details will depend on the particular aspect of instruction involved in each case, and hence on the item at hand. Using items based on this set of basic approaches, science teaching orientations could be identified, responses could potentially be quantified, and teaching orientation profiles obtained.

Theoretical Framework: Types of instruction and types of learning

Our approach to instructional strategies is framed by David Ausubel's theory of learning and instruction (Ausubel, Novak, & Hanesian, 1986). His ideas form a theoretical/conceptual framework for our work because they elegantly integrate foundational educational issues while focusing upon arguably the most important desired outcome, meaningful learning. This remains especially cogent today amid ongoing debates about science instructional approaches. Ausubel's two-dimensional diagram is depicted on the right in Figure 1, representing the nature of learning along the horizontal axis and the type of instruction along the orthogonal vertical axis. Type of instruction aligns with the spectrum of basic science teaching approaches described above. Note that Ausubel's original terms reception and discovery are best reflected today by direct and inquiry instruction. Meaningful learning is defined as the result of a learner's cognitive engagement such that new knowledge becomes integrated within the learner's conceptual schemata. Ausubel made the point that learning can actually range from rote to meaningful independently from instructional type, which may range from reception to discovery, hence the orthogonal axes. Ausubel (1961, 1963), Ausubel et al. (1986) and Novak (1976, 1979) clarified that the important learning goal was *meaningful learning* as opposed to rote learning, *whatever the type of instruction*. They believed that reception learning could be meaningful with appropriate instructional design; Novak (1976) referred to this as direct facilitation of concept learning, and various tools such as advance organizers and concept mapping for fostering meaningful reception learning have been developed (Mayer, 1979; Stone, 1983; Trowbridge & Wandersee, 2005). Research on conceptual change (Duit & Treagust, 2003; Posner, Strike, Hewson, & Gertzog, 1982; Thorley & Stofflett, 1996), the use of explanatory analogies (Dagher, 1995, 2005) and bridging analogies (Clement, 1982, 1998; Clement, Brown, & Zietsman, 1989), research on combining verbal learning with visual learning (Clark & Paivio, 1991; Culatta, 2012), etc. all involve forms of instruction that try to directly facilitate meaningful conceptual learning. Strategies from this research have shown instructional success at shifting learning from rote/fragmented to meaningful. Clearly one aims at meaningful learning whether by direct or inquiry routes, implying a need to focus on the two right-hand quadrants of Ausubel's diagram for both instruction and research.

Ausubel and Novak recognized the value of hands-on activities for learning science and viewed this partly as a method for promoting cognitive activity (i.e. minds-on). The idea of active learning is quite Ausubelian; teaching for active learning means using instructional techniques that encourage cognitive engagement with the subject matter rather than passive listening. Ausubel further cautions that hands-on activity without cognitive engagement would not lead to meaningful learning. What he called *discovery* learning, as advocated by Bruner (1961, 1971) and others (Guthrie, 1967), subsequently developed into today's commonly known inquiry instruction (National Research Council, 2000b). This approach advocates that learners engage with the practices of science during concept learning, that is, in activities that reflect the investigative nature of science (Guthrie, 1967). Unfortunately,

Ausubel's two-dimensional framework of orthogonal constructs often tended to be collapsed to one dimension, with direct instruction implicitly equated with rote learning, and inquiry instruction with meaningful learning. Although reception learning research in various forms continues today (Clark, Eyler, Rivas, & Wagner, 2011; Cobern et al., 2010; Klahr, 2002; Sweller, 2009), by the late 1980s the rote/meaningful learning dimension tended to be forgotten as the inquiry/direct instructional dichotomy became the focus. In 2000, the widely referenced book *How people learn* (National Research Council, 2000a) specifically advocates active learning and inquiry instruction without mention of Ausubel, Novak, reception learning, or even meaningful learning. The idea of meaningful learning does appear in *A framework for K-12 science education* (2011), and a nod is given toward the possible use of direct instruction ('instruction may include teacher talk', pp. 10–19), but beyond that the text is oriented only toward inquiry instruction. The *Framework* document (NRC, 2011) notes that 'Current research in K-12 science classrooms reveals that earlier debates about such dichotomies as "direct instruction" and "inquiry" are simplistic, even mistaken, as a characterization of science pedagogy'. As a characterization of all science pedagogy, we would agree. However, as argued earlier in this paper, the distinction between inquiry and direct instruction as fundamental epistemic approaches definitely has application to instructional design decisions for the teaching of science concepts (Cobern et al., 2010).

As we developed items for assessing science pedagogy orientations, we were mindful to distinguish instructional mode from type of learning. For this reason, the construction of items was guided by Ausubel's idea that instruction for meaningful learning can potentially range from reception to discovery modes and that we should reflect this in the item options.

Structure of Assessment Items

All items were cast in a standard multiple-choice question (MCQ) format as shown in Figure 2. Each item begins with a short, titled vignette representing a real instructional situation for a particular topic. The vignettes also specify instructional aim and approximate grade level. Although topics are not tightly tied to grade level, stating a level within a grade band appropriate to the topic adds a sense of realism, as does giving a name to the teacher in the vignette. Providing a small picture/icon adds some context and interest. The scenarios are kept fairly brief in order to narrow the focus to a particular aspect of pedagogy, and this also helps avoid many of the possible disadvantages of case methods (Shulman, 1992, p. 26), particularly in terms of time and distracting complexity. Reasonably concise items also maximize the number and variety of examples students can encounter. After the vignette comes a lead-in sentence to the four responses, posed in terms of an instructional method choice for the reader. Lead-in sentences begin with, 'Thinking about how you would teach this lesson, of the following ...' The purpose is to encourage the responder to envision himself or herself in the particular teaching situation, play the role of decision-maker, and respond accordingly. The phrase 'of the following' is intended to

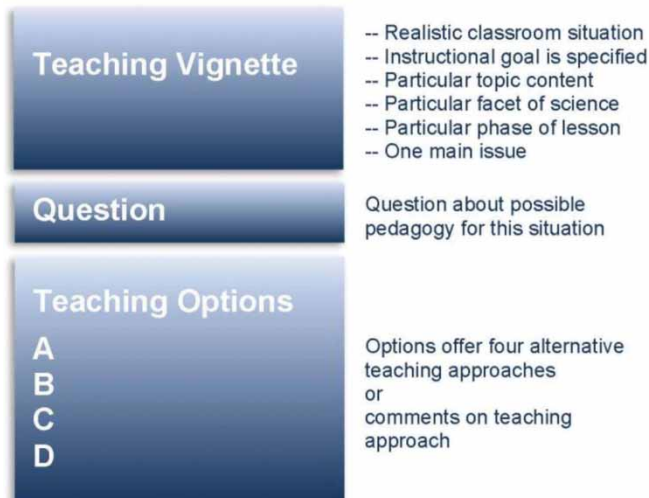


Figure 2. Standard MCQ format of assessment items

keep attention on the four responses. The lead-in sentence concludes with a question such as ‘which one is most similar to what you would do?’ or ‘how would you evaluate Mr. Goodchild’s lesson?’ or ‘how would you advise Ms. Katinka to structure her lesson?’ The question is then followed by four response options representing didactic direct, active direct, guided inquiry, and open inquiry instructional approaches. This overall item format is consistent across items, although the order of the response options varies from item to item. By constructing items using a fixed set of instructional types, it becomes possible to compare responses over a series of items and thus build up teaching orientation profiles for individuals or groups. Nevertheless, our intent is not to be ‘labeling or pigeonholing teachers, using a predetermined list of [orientation] categories’ (Friedrichsen et al., 2011, p. 372). Rather we seek to elicit and examine tendencies along this central epistemic dimension in teaching approach.

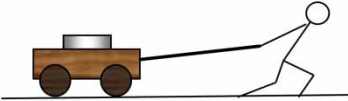
The nature and structure of such an assessment item is best appreciated via an example. The example in Figure 3 involves alternative pedagogy options for overall lesson design for the topic of force and motion. The four possible responses are randomly ordered.

Item Creation and Refinement

A team of four science education faculty, two doctoral research associates with school teaching experience, and five experienced K-8 teachers initially worked together to collect a large number of ideas for vignettes and alternative responses. It was important to have teachers and teacher educators involved in item creation because they have direct knowledge of situations that teachers encounter and how they are dealt with. The team created a wide range of draft items, working both individually and

Lesson on force and motion

Ms. Brandt is preparing a lesson to introduce her 5th grade students to the relationship between force and motion, namely that a net force will cause an object to speed up or slow down (Newton's 2nd Law). The classroom has available a loaded wagon to which a pulling force can be applied. Ms. Brandt is considering four different approaches to the lesson.



Thinking about how you would want to teach this lesson, of the following, which one is most similar to what you would do?

- Write a clear statement of Newton's 2nd Law on the board and explain it carefully to my students. Then I would demonstrate the law by pulling on a loaded wagon with a constant force in front of the class as they observe the motion.
- Raise the question of what kind of motion results from a constant force. I would then guide my students to explore the question themselves by pulling on a loaded wagon and observing what happens. From the evidence they would then propose a possible law.
- Write a clear statement of Newton's 2nd Law on the board and explain it carefully to my students. I would then have the students verify the law by pulling on a loaded wagon themselves and confirming what type of motion results.
- Raise the question of whether there is any relationship between force and motion. My students would then be free to explore this safely in the lab. Afterward we would have a class discussion of their findings.

Figure 3. Example of a pedagogy assessment item

together, and then refined them in light of extensive discussion. Ideas for the vignettes came from the teaching experiences of those on the item writing team or were created with the instrument in mind for topics they were familiar with teaching. Others arose from thoughts and insights about various pedagogies in context, as well as from the literature (Tippins, Koballa, & Payne, 2002), local and state science objectives, *The National Standards on Science Education*, the *Harvard-Smithsonian Case Studies in Science Education* (Annenberg Foundation, 2013), and the National Research Council (2011). Ideas for vignettes were then drafted into our item format, i.e. vignette/question/options. Due to limitations on teachers' time, from this point on the revising–testing–revising of items was done by the science education faculty and two doctoral research associates, hereafter referred to as the developers.

The final revision work was to align items with the basic orientations in Figure 1, revising options until there was a consensus within the developers that each option reflected the intended pedagogical approach. Samples of items were subsequently dissected and discussed in detail by focus groups composed of from 3 to 15 pre- or in-service teachers meeting with a project investigator. The pre-service teachers were volunteers from a science methods course (none of whom knew any of the project investigators). The teachers were volunteers from area schools. Each focus group was asked to respond to a small set of draft items and then to discuss their thoughts on the items. We were particularly interested in whether the focus groups found the item vignettes and response options realistic, and whether the items made sense to them. We also asked several science education professors at other universities to comment on a sample of items. Both groups found the items

Table 1. Distribution of items

Grades	Sciences				
	Earth	Life	Physical	General	All
K-2	5	8	14	2	29
3-5	8	7	14	1	30
6-8	14	10	15	2	41
All	27	25	43	5	100

understandable and appropriate, and the vignettes and responses consistent with their experiences or expectations with respect to the breadth of instructional practice.

As a result of our item development work, we have 100 selected-response items distributed over three science areas (Earth Science, Life Science, and Physical Science) and three grade band groupings (K-2, 3-5, 6-8). Table 1 shows the current item distribution. One of our goals was to have a reasonable spread across science areas and grade bands to represent the broad range of instructional situations that teachers might encounter. The cells in Table 1 are all reasonably populated, with some imbalance due to the nature of topics at various grade levels. A few items represent generic rather than topic-specific aspects of science and of science instruction, and are categorized as 'general'.

Piloting of Selected Items

The items were constructed to avoid response options that would have little appeal to pre-service or in-service teachers (what might be called straw man responses), or that represented poor teaching generally. We wanted choices that would be seen as reasonable and reflective of common practices, thus engendering reflection and discussion. We did not want choices that would be avoided by all but the rarest person. Focus group findings on individual items were encouraging, but until a larger number of the items were piloted, we could not know if most persons reading an item would focus mainly on a single response or if we would observe wide ranges of responses.

Ultimately, the effectiveness of items for formative purposes can only be assessed by classroom discussions of individual items and responses. However, a quicker assessment of potential item effectiveness can be achieved by a quantitative approach, looking at distribution profiles obtained along the pedagogical spectrum we have outlined. Items that produce little response variation are less likely to illuminate the potential range of teacher thinking about pedagogy, and less likely to promote energetic classroom discussion of teaching approaches.

For this type of quantitative evaluation of items, two forms or instruments of 16 items each were compiled. Such compiled sets are called Pedagogy of Science Teaching Tests (POSTT), and the two forms are called POSTT-1 and POSTT-2. Copies of these first test instruments may be downloaded from <http://www.wmich.edu/science/>

[inquiry-items/index.html](#). The choice of 16 items per instrument was based on an estimate of three minutes for the minimum time it would take a person to respond thoughtfully to each item; thus, 16 items in a POSTT form should take less than 60 minutes (trials typically took 30–40 minutes). A 16-item instrument achieves this and can still contain a range of three science subjects and three grade bands. We began with two instrument versions to increase the number of items piloted, with four items common to both versions to serve as an item reliability check (response patterns to the four common items did not vary statistically between POSTT-1 and POSTT-2). Hence, both tests together utilize 28 unique items out of the bank of 100 items. Both versions were administered mid-semester in three sections of an upper level, pre-service elementary science methods course at a Midwestern university. By the time of taking the POSTT, all of the students had been introduced to science teaching strategies including inquiry and had taken at least three science content courses taught using a predominantly inquiry format. Twenty-eight subjects took POSTT-1 and 32 took POSTT-2.

Findings

Overall Results

Because our primary aim in developing items was for formative assessment purposes during teacher preparation or workshops, we are especially interested in the response spread in the pilot data. As noted above, we wanted to know whether or not items would precipitate a breadth of response choices across the spectrum of four teaching options. Table 2 shows the distributions obtained across students and across item responses, for the item sets administered. No students chose the same instructional type across all items; all students always chose at least two different types among the four response categories. Similarly, every item precipitated a range of responses; no item precipitated only a single response type. Almost all students (except four) selected three or more possible strategies at least once (32 of 60 students used all four responses at least once). Likewise, 24 items out of the 28 total across POSTT-1 and POSTT-2 had all four responses selected at least once.

Table 2. Student and item response variation

Student response variation		Item response variation	
No. of different choices (1–4)	No. of students (%)	No. of different choices (1–4)	No. of items (%)
1	0 (0.0%)	1	0 (0.0%)
2	4/60 (6.7%)	2	2/32 (6.3%)
3	24/60 (40.0%)	3	6/32 (18.7%)
4	32/60 (53.3%)	4	24/32 (75.0%)

Detailed Results for Four Items

We present four POSTT assessment items as examples, giving the complete item with vignette, question and response options, along with the corresponding histogram of cumulative student responses per item (as opposed to individuals' orientation profiles), and a histogram showing how the item developers responded to the item. Including the developers' data is for qualitative comparative purposes. The comparison suggests how the views of science teacher educators and pre-service teachers can vary, which in turn strengthens the validity of the items for formative assessment purposes. Note that while the bar charts below show profiles per item across subjects, results can also be analyzed and profiles depicted per subject (or per group) over all the items. Some of the project activity to date (especially international) has focused on these kinds of analyses.

When administered to participants, item responses were in random order and labeled A, B, C, and D (as in [Figure 3](#)). For purposes of illustration and discussion in this paper, the responses below are placed in order of instructional mode and labeled DD, AD, GI, and OI (Didactic Direct, Active Direct, Guided Inquiry, and Open Inquiry). Item response options are specifically designed so that an increase along the POSTT scale (ordered from 1 to 4, i.e., DD to OI) reflects increased degree of inquiry teaching mode coincident with decreased degree of direct teaching mode. We recognize that this modal dimension is neither precisely quantifiable, nor precisely divisible into four equal intervals. Moreover, different contexts between items are an essential aspect of the primarily qualitative use of these items and instruments. However, for quantitative analyses, reporting the median (a single integer on this scale) would not adequately convey the variance in results across pedagogical choices. Therefore, the POSTT scale can be construed as an 'ordinally interval hybrid scale' (Hair, Bush, & Ortinau, 2003, p. 390), in which resulting means and standard deviations can be considered meaningful as simple descriptions of central tendency and dispersion within items, instruments, individuals, or groups.

The four items offered below are those that produced the strongest inquiry orientation overall, the narrowest response distribution, the most even response distribution, and the strongest direct instruction orientation overall, respectively. The item shown in [Figure 4](#) clearly drew a strong open inquiry response from students with only 3 students of 28 opting for either form of direct instruction. What is arresting about this item is that the developers were visibly less inquiry-oriented than the pre-service teachers. For formative use, an item like this affords an opportunity to probe and discuss differing views among instructors as well as between instructors and students. The visible difference between students' and developers' response profiles is also reflected in the basic statistical results. Again, descriptive comparisons can be made by taking the numerical scale along the X -axis to represent the spectrum of pedagogies (assigning DD = 1, AD = 2, GI = 3, and OI = 4). The mean of the students' responses is at 3.50 (between Guided Inquiry and Open Inquiry), while the mean of the developers' responses is at 2.83 (slightly more direct than the Guided Inquiry category). Standard deviations (representing diversities of opinion) for both

Magnetic attraction

Mr. Golden is beginning a unit on magnetism with his 2nd grade students, and his objective is for them to learn about magnetic attraction. He gives each student group a bar magnet and a tray that contains a paper clip, a coin, an iron nail, school scissors, a pencil, some keys, a marble, a crayon, aluminum foil, some sand, and students can add a few objects of their own. Mr. Golden introduces the term ‘magnetic attraction,’ and demonstrates how to test a couple of objects with a magnet. Student groups are then asked to sort the objects in their trays according to whether they are attracted by the magnet or not.



Thinking about how you would teach, of the following, how would you evaluate Mr. Golden’s lesson?

- DD) This is a good lesson because Mr. Golden introduces the important terminology right at the start. However, having demonstrated how to test an object using a magnet, he might as well have demonstrated what happens with all the objects, sorting as he goes.
- AD) This is a good lesson because Mr. Golden introduces the important terminology right at the start, and follows up with the students doing a hands-on activity, testing and sorting the objects themselves.
- GI) Instead of beginning with terminology, Mr. Golden should have had the students first test the various objects themselves and discuss their ideas about it. In wrapping up the session, Mr. Golden could introduce the term magnetic attraction, and how it applies to what they observed.
- OI) Mr. Golden should have allowed the students to explore freely with magnet and objects, without bringing up terminology. He could then let them discuss any ideas they might have about it and share these with the class. The only contribution he needs to make is to present the term magnetic attraction at the end.

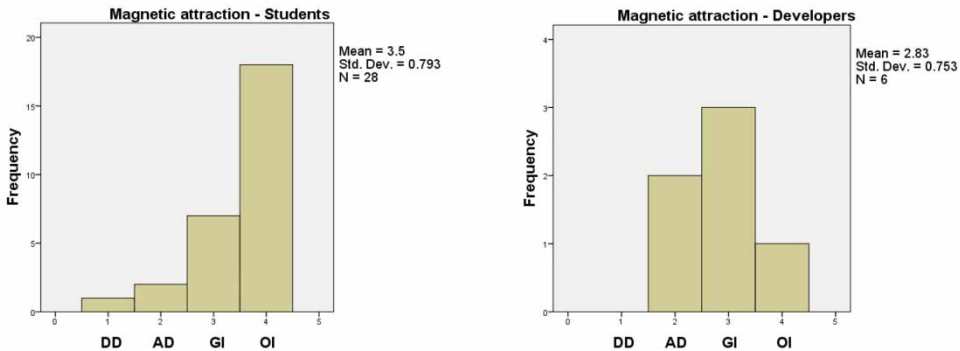


Figure 4. Strongest inquiry instruction response orientation overall—magnetic attraction

students and developers are comparable, only slightly higher for the students than the developers.

The ‘Earth materials’ item (Figure 5) was 1 of only 2 items (out of 28) that elicited only 2 kinds of responses from 29 participating students. Students preferred guided inquiry with only a few opting for active direct. The developers also chose guided inquiry, with one opting for open inquiry. A quick comparison of means confirms that developers leaned farther toward the inquiry side of the spectrum from Guided Inquiry (3.17), and students leaned just slightly more direct than Guided Inquiry (2.83). One would want to know the reasons given by both groups for their choices. Especially for formative use, a course instructor would want to know what reasons the students had for their choices.

Twenty-nine students responded to the ‘Sink or float’ item (Figure 6). The results and comparisons are very interesting. All six developers took the guided inquiry

Earth materials

Mr. Sanchez wants his 3rd grade students to be able to recognize and describe different types of earth materials, namely rock, mineral, clay, gravel, sand, and soil samples, which he has available for use in the lesson. Mr. Sanchez is considering four different approaches to the lesson. Thinking about how you would teach this lesson, of the following, which one is most similar to what you would do?



DD) I would write the different types of earth materials on the board and define them for my students. Then I would individually describe the unique characteristics of each type of material to the students, and pass the samples around.

AD) I would write the different types of earth materials on the board and define them for my students. Based on the descriptions on the board, I would then ask the students to sort the earth materials, and describe why they sorted the materials the way they did.

GI) I would have the students sort and describe the various earth materials displayed on their tables, according to their unique characteristics. I would then guide a class discussion about these different types of earth materials.

OI) I would ask the students to think about what types of materials the earth is made up of. The students would be free to explore this question with different earth materials in the classroom, and then report back on their conclusions.

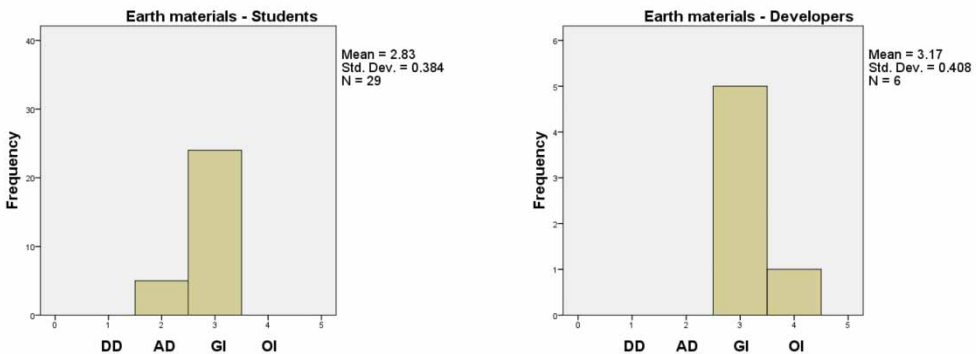


Figure 5. Narrowest response distribution—earth materials

position. While the mode of the distribution for the students was also guided inquiry, different students, in marked contrast to the developers, were attracted by *all* the options including Didactic Direct. Moreover, the students taking the POSTT were upper-division teacher education students who had already taken several inquiry-oriented science content courses and were more than halfway through their science methods course, making this an unexpected response distribution (compared to the unanimity of the developers' responses). We suggest, therefore, that the value of the POSTT as an instructive, formative assessment tool is particularly evident in the responses to this item.

The 'Frog dissection' item (Figure 7) was one of the four items that appeared in both POSTT-1 and POSTT-2, hence the participant number of 60. The student responses on items ranged from didactic direct to open inquiry, with the mode of the distribution clearly at active direct. The mode for the six developers was also active direct, but none of the developers opted for open inquiry. Given the response spread among both subjects and developers, this item can be expected

Sink or float

Ms. Hoo has her kindergarten students gather around a small pool of water. She has a set of objects of different sizes and different materials; some will sink and some will float. Ms. Hoo’s goal is for her students to first distinguish the objects by whether they sink or float, and then realize that this does not depend on the size of the object but on what it is made of (e.g., the stones will all sink no matter how big or small they are, and the wooden blocks will all float).



Thinking of how you would teach this lesson, of the following, what would you most likely do?

- DD) Drop objects one by one into the water, and have the children notice that some sink and some float. Point out that all the stones sank, no matter how big or small, and all the wooden blocks floated, etc. Conclude by stating the lesson objective, that it is not size that matters but the material the object is made of.
- AD) Have students come one by one and drop an object into the water, with everyone calling out whether it sank or floated. Point out that all the stones sank, no matter how big or small, and all the wooden blocks floated, etc. Conclude with the lesson objective, that it is not size that matters but the material the object is made of.
- GI) Have students come by one by one and drop an object into the water, with everyone calling out whether it sank or floated. Ask them to suggest what this depended on; when some suggest size and others what it is made of, have them test these ideas by dropping more objects. Then have them agree on a conclusion.
- OI) Have all the students dropping various objects in the water and seeing what happens. Then have them talk among themselves about this and ask volunteers to give their ideas about it, with others saying if they agreed or not.

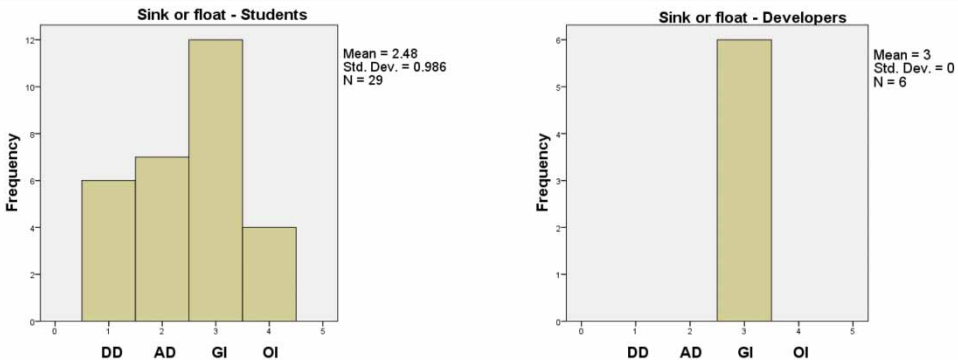


Figure 6. Most even response distribution—sink or float

to prompt lively and diverse discussion of alternative teaching approaches for particular topics.

The frog dissection item also illustrates the ongoing process of discussing and refining items for both formative and summative use. For example, this science topic can serve the goal of helping middle school students understand the body as ‘a system of multiple interacting subsystems’ and know that ‘special structures are responsible for particular functions in organisms’ (NGSS, MS-LS1-3). It is a likely activity, but it is perhaps more suitable for eighth grade than for sixth, and it probably needs to make clear that in all four modes the actual dissection activity is outlined for students in advance of giving them dissection tools. The essential difference is in whether they are first given direct instructions on what to look for and why, or are directed to

Frog dissection

Mr. Goodchild is doing a frog dissection with his 6th graders to help teach them about anatomy.

Thinking about how you would teach a lesson, of the following, which is most similar to what you believe is the best way to incorporate a dissection into a lesson?

- DD) It should be used as a demonstration with Mr. Goodchild explicitly pointing out what students need to know.
- AD) It should be used as a follow-up student activity after Mr. Goodchild explains what students need to know.
- GI) Students should do the dissection first, followed by teacher-led discussion.
- OI) It should be used as a stand-alone activity for students to explore and discuss on their own.

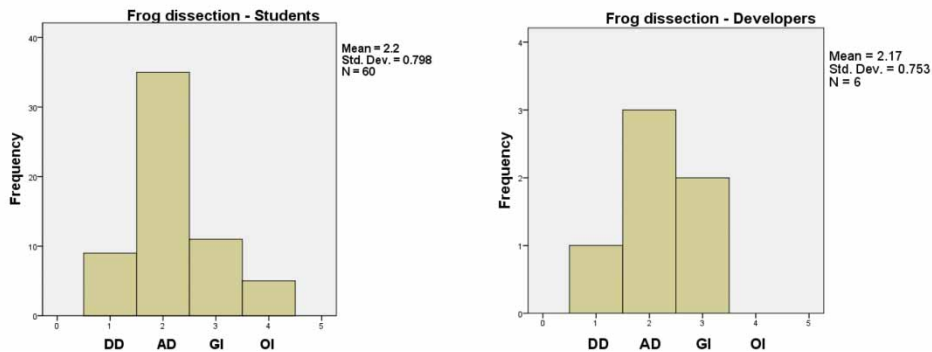


Figure 7. Strongest direct instruction response orientation overall—frog dissection

first observe and carefully consider possibilities themselves. The process of trying to create perfect items is incremental at best, and as the usage of these POSTT items and instruments moves ahead, we look forward to making additional refinements in light of constructive comments. Another interesting feature of this item is that, while dissections are common events in biology labs, the usual unqualified recommendation for teaching science by inquiry may result in instruction being too undirected, in particular cases, for particular purposes, for some students, and by some teachers. This all makes for good discussion between methods teachers and their students who will be teachers. As a strategy of formative assessment, the discussion affords the instructor the opportunity to assess employment of pedagogical content knowledge.

Discussion

The responses to the 4 items described above are typical with respect to the balance of the 28 items used for POSTT-1 and POSTT-2, and to the entire item bank. All piloted items precipitated a range of responses from students, and each student chose some range of response types for different items, suggesting that classroom discussions based on these items could usefully indicate how the students understand and value different approaches to science pedagogies, and under what circumstances they

believe they would employ them. Such discussions, coupled with the ability to analyze quantitative class results, should help science methods educators gauge progress toward course goals on understanding science pedagogies. The response spread for the various items raises other interesting questions. Does the area of science or particular topic make a difference in responses to an item? Does the function or location of a particular pedagogy within the vignette's lesson make a difference? Does grade level/band make a difference? These questions are relevant to instruction and to research. Of great interest are the reasons students give for their choices. A potential use for a POSTT instrument might be to look for shifts in means between pre- and post-tests, with pedagogical methods instruction occurring in between, while another would be to look for differences between the means from different groups of teachers in different circumstances.

Another pertinent question is whether degree of science content knowledge affects ability to choose a preferred pedagogy from the four options. Feedback from both respondents and developers indicates it is not an issue for choosing or evaluating pedagogy, as long as vignette situations are couched in descriptive rather than technical terms. None of the respondent groups raised it as problematic during discussions, and developers whose expertise was in one subject felt comfortable evaluating a pedagogy item whose context was another; in fact this was sometimes seen as an advantageous clarity check. However, it remains an avenue for further research. The POSTT items can be expected to have applications in teacher education programs and research in many countries. One of our members has made cultural adaptations to POSTT-1 for use at an English language Turkish university and has begun to collect data. Many of the response patterns are similar to ours, but there are also intriguing differences to be studied. Recently in South Africa an earlier set of POSTT items, specifically for physical science, were used to assess and compare the pedagogical orientations of in-service physical science teachers practicing in township (disadvantaged) schools and suburban (advantaged) schools, and results so far indicate marked differences between the preferred teaching practices of the two groups of teachers in their particular circumstances (Ramnarain & Schuster, 2014). For information regarding the Turkish language POSTT and the Turkish English language POSTT, contact Ebru Muğaloğlu (akturkeb@boun.edu.tr), and for a Korean language POSTT, contact Young-Shin Park (parkyoungshin1968@gmail.com). For use in South African, contact Umesh Ramnarain (uramnarain@uj.ac.za).

Potential for summative assessment use: There is likely to be interest in using POSTT items in an instrument format for summative assessment purposes. Here we point out that the assessment identifies science teaching orientations and produces orientation profiles. Items reflect a consistent spectrum of possible instructional approaches and are not assumed to have one correct instructional option with the others being wrong distracters as with conventional content MCQs. This, along with the (desirable) response spread shown in our pilot study, suggests caution in summative implementation, interpretation, or instrument testing. Nevertheless, cognizant of this, we conducted preliminary studies of the test characteristics of the items. We deliberately

placed four common items in POSTT-1 and POSTT-2 so we could estimate reliability. Sixty students responded to these four items across three sections of a science methods course. There were no significant differences for any of the four items between groups (*at the 5% level*). The rest of the items were taken by 28 or 29 students; and, of those 24 items, 21 items showed no significant differences between sections responding to the POSTT items (*at the 5% level*). The participant sizes are not large, but the findings nonetheless suggest that 25 items perform rather reliably (the correlation matrix is available at <http://www.wmich.edu/science/inquiry-items/index.html>), and eventually we will have data on all 100 items. At that point, the POSTT website will have all the items along with the response histograms. (As noted, all 100 items are available at the <http://www.wmich.edu/science/inquiry-items/index.html> website. As we gather data on each item, we are posting the student response histogram and the developers' histogram with each item.)

One might ask whether or not the items represent some single construct, as one might ordinarily expect of summative assessment. However, one must bear in mind that the aim here is to obtain science teaching orientation profiles, and that different teaching situations may evoke different pedagogical preferences. Thus, the very characteristic that makes the items useful for formative assessment, response spread, is problematic for conventional summative assessment. The response variation shows itself in the weak inter-item correlations that we calculated. An exploratory factor analysis showed small clusters of items loading on separate factors. While it is possible that subsets of highly correlated items may be located among the 100-item inventory which could then be used for summative assessment, we suggest two alternative approaches. We are posting the developers' histogram for each item. Anyone who teaches K-8 science methods or works in science teacher development will have their own perspective on how the various item responses correspond with their instructional goals for pre- and in-service teachers. Hence, one approach to a summative assessment is for items to be selected and scored in a way which is consistent with particular instructional goals, making it a criterion-referenced assessment. For example, if one were teaching guided inquiry, then one could select a set of topics and items that best fit that model, and likewise for the other modes.

Note that the two POSTT forms compiled so far are general instruments in that they include items from all three subject areas and three grade bands. Additional POSTT versions can be compiled by selecting items with particular science content, classroom grade level, or aspect of instruction. An instrument could easily be used both qualitatively and quantitatively. Formative assessment discussions or even an added comment section below each item could readily serve to draw out other critical ideas or issues not present in the response options themselves. Participants can be asked to state or write their reasons for their instructional choices and say why they did not choose the other options. In this case, the assessment would be both about teaching orientations and the merits of rationales given for specific decisions.

Conclusion

As noted, our work is motivated by a concern that pre-service K-8 teachers typically are not able to see or engage with very many examples of science teaching pedagogies for science content learning across various topics. Moreover, with the amount of time given to science instruction falling at the elementary level (Petrinjak, 2011), the prospects are not good that student teachers will be exposed to a wide variety of science instruction examples. Yet cognitive science findings and studies of the use of worked examples both suggest that people need multiple exposures to instances of concepts and practices over a wide range of situations in order to develop competence and adaptable expertise (Donovan & Bransford, 2005; Heller & Heller, 2010; Johnson, 2001).

Our project contributes a new type of pedagogical assessment which can be used for formative, summative, and research purposes. Rather than being cast in terms of declarative generalizations about teaching, items are problem based in nature, each situated in a realistic classroom situation for teaching a specific science topic, to best assess whether students are able to apply and integrate their knowledge and understanding of science pedagogies. Individual items are available from a bank of 100 items, cross-referenced according to topic, grade band, and instructional aspect categories. Several compiled POSTT instruments are currently available and further instruments can be purpose built with specific focuses. Such items and instruments can be used to provide novice science teachers with multiple exposures to science teaching pedagogies for a variety of topics and teaching situations in relatively short time frames. The response spread obtained in our pilot studies is a promising outcome, indicating that the items are likely to prompt lively discussion during teacher preparation, or professional development situations about ways to teach science and deal with common learning and teaching scenarios. Instructors who have already administered a POSTT have commented that they are very interested in using items formatively to spark targeted classroom discussions, including consideration of the possible relevance of the science topic and grade level to pedagogical choices. Having such discussions encourages teachers to think through instructional options and the reasons for making one choice over another, in various cases, and the contextual factors and constraints that might come into play. POSTT instruments provide teacher educators and professional development leaders with an assessment of student/teacher understanding of science pedagogical decision-making, thus providing formative feedback to them about how to shape their own teacher education courses and programs. POSTT items are particularly well suited for use with clicker technology for in-class formative assessment with immediate feedback and discussion. Judiciously selected individual items can be displayed for students to view and consider at various stages during a methods lesson or in professional development settings. Item discussion leads to further learning as students reflect on their ideas; while the instructor is able to use the assessment responses to adjust a current lesson on the spot as well as to shape future lessons.

The bank of 100 items publically available, along with several POSTT instruments, constitutes a valuable resource for the improvement of science teacher education. We hope to be able to invite constructive comments and continue the item creation and refinement processes. We would also like to create an open data bank for volunteered items offered in the same vein and format. We would process and refine the most promising items and add them to the formal collection of accessible, searchable, and downloadable items, as well as including them in further POSTT instruments.

Acknowledgements

This project was based upon work supported by the National Science Foundation Award #0512596. Any opinion, 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

- Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N. G., Naaman, R. M., Hofstein, A., . . . Tuan, H. L. (2004). Inquiry in science education: International perspectives. *Science Education*, 88(3), 397–419.
- Albanese, M., & Mitchell, S. (1993). Problem based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, 68(1), 52–81.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.
- Annenberg Foundation. (2013). *Learning science through inquiry: An eight part K-8 professional development workshop series*. Retrieved August 21, 2013, from <http://www.learner.org/workshops/inquiry/>
- Ausubel, D. P. (1961). In defense of verbal learning. *Educational Theory*, 11(1), 15–25.
- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York, NY: Grune & Stratton.
- Ausubel, D. P., Novak, J. D., & Hanesian, H. (1986). *Educational psychology: A cognitive view* (2nd ed.). New York, NY: Werbel & Peck.
- Banilower, E. R., & Smith, P. S. (2013). *Highlights from the 2012 National Survey of Science and Mathematics Education*. Retrieved February 4, 2013, from <http://www.horizon-research.com/2012nssme/wp-content/uploads/2013/01/NSSME-bag-lunch-at-NSF.pdf>
- Benceze, L., Hewitt, J., & Pendretti, E. (2001). Multi-media case methods in pre-service science education: Enabling an apprenticeship for praxis. *Research in Science Education*, 31(2), 191–209.
- Black, P., & Wiliam, D. (2006). *Inside the black box: Raising standards through classroom assessment*. Retrieved February 17, 2006, from <http://www.pdkintl.org/kappan/kbla9810.htm>
- Black, P. J., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning*. Berkshire: Open University Press.
- Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, 31(1), 21–32.
- Bruner, J. S. (1971). Some elements of discovery. Chapter 4 of J. S. Bruner *The Relevance of Education* (pp. 68–81). New York, NY: W.W. Norton.
- Capon, N., & Kuhn, D. (2004). What's so good about problem-based learning? *Cognition and Instruction*, 22(1), 61–79.
- Cavanagh, S. (2004). NCLB could alter science teaching. *Education Week*, 24(11), 12–13.

- Cavanagh, S. (2005). *Asian equation*. Retrieved January 27, 2008, from <http://www.edweek.org/ew/articles/2007/06/06/39china.h26.html?tmp=1543524430>
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–170.
- Clark, N., Eyler, G., Rivas, A., & Wagner, T. (2011). *Direct instruction vs. scientific inquiry: Evaluating student outcomes*. Retrieved June 25, 2012, from emurillo.org/Classes/Class2/DirectvsScientific.docx
- Clement, J. J. (1982). Student preconceptions in introductory mechanics. *American Journal of Physics*, 50(1), 66–71.
- Clement, J. J. (1998). Expert novice similarities and instruction using analogies. *International Journal of Science Education*, 20(10), 1271–1286.
- Clement, J. J., Brown, D. E., & Zietsman, A. (1989). Not all preconceptions are misconceptions: Finding ‘anchoring conceptions’ for grounding instruction on students’ intuitions. *International Journal of Science Education*, 11(5), 554–565.
- Cobern, W. W., Schuster, D. G., Adams, B., Applegate, B., Skjold, B., Undreiu, A., . . . Gobert, J. D. (2010). Experimental comparisons of inquiry and direct instruction in science. *Research in Science & Technological Education*, 28(1), 81–96.
- Cooper, G. A., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem solving transfer. *Journal of Educational Psychology*, 79, 347–362.
- Culatta, R. (2012). *Dual coding theory (A. Paivio)*. Retrieved June 22, 2012, from <http://www.instructionaldesign.org/theories/dual-coding.html>
- Dagher, Z. R. (1995). Analysis of analogies used by science teachers. *Research in Science Teaching*, 32(3), 259–270.
- Dagher, Z. R. (2005). The case for analogies in teaching science for understanding. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), *Teaching science for understanding: A human constructivist view* (pp. 195–211). San Diego, CA: Academic Press.
- Dean, C. D. (1999, April). *Problem-based learning in teacher education*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.
- Donovan, M. S., & Bransford, J. D. (2005). *How students learn: Science in the classroom*. Washington, DC: Committee on How People Learn: A Targeted Report for Teachers, National Research Council, The National Academies Press.
- Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25(6), 671–688.
- European Commission. (2011). *Science education in Europe: National policies, practices and research*. Brussels: Education, Audiovisual and Culture Executive Agency.
- Friedrichsen, P., Van Driel, J. H., & Abell, S. K. (2011). Taking a closer look at science teaching orientations. *Science Education*, 95(2), 358–376.
- Guthrie, J. T. (1967). Expository instruction versus a discovery method. *Journal of Educational Psychology*, 68(1), 45–40.
- Hair, J. F., Bush, R. P., & Ortinau, D. J. (2003). *Marketing research*. New York, NY: McGraw-Hill/Irwin.
- Heller, K., & Heller, P. (2010). *Cooperative problem solving in physics: A user’s manual*. Retrieved August 14, 2013, from <http://www.aapt.org/Conferences/newfaculty/upload/Coop-Problem-Solving-Guide.pdf>
- Johnson, M. (2001). Facilitating high quality student practice in introductory physics. *American Journal of Physics*, 69(S1), S2–S11.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331, 772–775.
- Klahr, D. (2002). *Paths of learning and their consequences: Discovery learning versus direct instruction in elementary school science teaching*. Retrieved April 10, 2009, from <http://www.lrdc.pitt.edu/supergroup/>

- Leone-Cross, L. (2013). *State seeks new science education standards*. Retrieved July 29, 2013, from http://www.sj-r.com/top-stories/x369942208/State-seeks-new-science-education-standards?zc_p=2
- Lotter, C., Harwood, W. S., & Bonner, J. J. (2007). The influence of core teaching conceptions on teachers' use of inquiry teaching practices. *Journal of Research in Science Teaching*, 44(9), 1318–1347.
- Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 95–132). Dordrecht: Kluwer Academic.
- Maloney, D. P. (1994). Research on problem solving: Physics. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 327–354). New York, NY: MacMillan.
- Mayer, R. E. (1979). Twenty years of research on advance organizers: Assimilation theory is still the best predictor of results. *Instructional Science*, 8(2), 133–167.
- Mervis, J. (2013). Science standards begin long, hard road to classroom. *Science*, 340(6139), 1391.
- Michaels, S., Shouse, A. W., & Schweingruber, H. A. (2007). *Ready, set, science! Putting research to work in K-8 science classrooms*. Washington, DC: National Research Council.
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
- National Research Council. (2000a). *How people learn*. Washington, DC: National Academy Press.
- National Research Council. (2000b). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academy Press.
- National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.
- Next Generation Science Standards. (2013). *The Next Generation Science Standards (NGSS)*. Retrieved May 18, 2013, from <http://www.nextgenscience.org/>
- Ngeow, K., & Kong, Y.-S. (2001). Learning to learn: Preparing teachers and students for problem-based learning. Retrieved, May 9, 2014, from <http://files.eric.ed.gov/fulltext/ED457524.pdf>.
- Novak, J. D. (1976). Understanding the learning process and effectiveness of teaching methods in the classroom, laboratory, and field. *Science Education*, 60(4), 493–512.
- Novak, J. D. (1979). *The reception learning paradigm*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.
- Peterson, M. (1997). Skills to enhance problem-based learning. *Medical Education Online* [Serial Online], 2(3). Retrieved July 20, 2011, from <http://www.med-ed-online.org/f0000009.htm>
- Petrinjak, L. (2011). *Elementary teachers getting less time for science*. Retrieved July 21, 2011, from <http://www.nsta.org/publications/news/story.aspx?id=58727>
- Posner, G., Strike, K. A., Hewson, P., & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66, 211–227.
- Ramnarain, U. & Schuster, D. (2014). The Pedagogical orientations of south african physical sciences teachers toward inquiry or direct instructional approaches. *Research in Science Education*. (ISI). Advance online publication. doi: 10.1007/s11165-013-9395-5.
- Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T.-Y., & Lee, Y.-H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching*, 44(10), 1436–1460.
- Shulman, L. S. (1992). Toward a pedagogy of cases. In J. H. Shulman (Ed.), *Case methods in teacher education* (pp. 1–30). New York, NY: Teachers College Press.
- Stone, C. L. (1983). A meta-analysis of advance organizer studies. *Journal of Experimental Education*, 51(4), 194–199.
- Sweller, J. (2009). What human cognitive architecture tells us about constructivism. In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* (pp. 127–143). New York, NY: Routledge.

- Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2, 59–89.
- Thomson, M. M., & Gregory, B. (2013). Elementary teachers' classroom practices and beliefs in relation to US science education reform: Reflections from within. *International Journal of Science Education*, 35(11), 1800–1823.
- Thorley, N. R., & Stofflett, R. T. (1996). Representation of the conceptual change model in science teacher education. *Science Education*, 80(3), 317–339.
- Tippins, D. J., Koballa, T. R., & Payne, B. D. (2002). *Learning from cases: Unraveling the complexities of elementary science teaching*. Boston, MA: Allyn and Bacon.
- Trafton, J. G., & Reiser, B. J. 1993. *The contributions of studying examples and solving problems to skill acquisition*. The proceedings of the fifteenth annual conference of the cognitive science society, Boulder, CO, pp. 1017–1022.
- Trowbridge, J. E., & Wandersee, J. H. (2005). Theory-driven graphic organizers. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), *Teaching science for understanding: A human constructivist view* (pp. 95–131). San Diego, CA: Academic Press.
- Wang, H. C. A., Thompson, P., Shuler, C., & Harvey, L. (1999, January). *Problem-based learning approach for science teachers' professional development*. Paper presented at the annual meeting of the Association for the Education of Teachers in Science, Austin, TX.
- Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, 7(1), 1–39.
- Wise, K. C. (1996). Strategies for teaching science: What works? *Clearing House*, 69(6), 337–338.