Orchestrating productive whole class discussions: The role of designed student responses

Sheila Evans
University of Nottingham

Clare Dawson
University of Nottingham

Received: 1 January 2017 Accepted: 8 February 2017
© Mathematics Education Research Group of Australasia, Inc.

The value of students publicly sharing and discussing their solutions to unstructured problems is widely recognized. This can, however, be pedagogically challenging. The solutions may be partial, unclear and unpredictable. For many teachers, particularly those new to working with such problems with their students, the improvisation needed to orchestrate productive discussions can be unmanageably high. In this paper we present a pedagogical tool to help teachers. Specifically, teachers orchestrate discussions of *designed, worked-out solutions* to unstructured problems. The worked-out solutions have been carefully pre-designed for teachers to use in lessons. Knowing the range of distinct solutions students are to work with supports teachers' planning. The reduced need for improvisation means they are better placed to learn and practice new ways of probing students' reasoning. These acquired practices may then be applied to discussions of students' own responses to a problem.

In the study we explore the question: for a teacher new to working with unstructured problems, how do discussions of worked-out solutions (called in this paper designed student responses) differ from discussions of students' own solutions? Our findings indicate that discussions of authentic student solutions tend to focus on procedural descriptions, whereas the discussions of designed solutions stimulated student explanation and evaluation. The work reported here represents the initial part of an on-going study.

Keywords · whole class discussion · teacher tools · problem solving · designed student response · worked-out example · mathematics discourse

Introduction

Whole class discussions held after students have attempted to solve unstructured problems can make a substantial difference to learning. By expressing ideas publicly, students make their mathematical reasoning visible for scrutiny. As well as informing the teacher about what students know and what they need to learn, whole class discussions provide opportunities to develop students' individual ideas into collective understandings (Staples, 2007). Misconceptions may surface, different solution strategies may be compared, connected and contrasted, and generalizations may be sought (Brousseau, 1997). However, to productively orchestrate these discussions is demanding. In this paper we present a pedagogical tool to help teachers.

The study forms part of a large design research project in which over one hundred Formative Assessment Lessons were designed and developed to support US Middle and High Schools in implementing the new Common Core State Standards for Mathematics. These resources can be found online at: http://map.mathshell.org.uk/materials/lessons.php. About a third of the lessons developed involve the tackling of unstructured problems. A novel feature is the inclusion of several worked-out solutions to each problem. These are carefully selected and modified versions of authentic responses taken from trials of the problem with other classes of students (Swan & Burkhardt, 2014). The intended learning goals of the lesson determine the content and structure of the worked-out solutions. The worked-out solutions provide opportunities for students to make sense of and critique responses without the social pressures associated with scrutinizing a peer's ideas, which can be constraining. Because the solutions are included in the materials, teachers can plan how to orchestrate the discussion of these responses before the lesson in a detailed way that is not possible when using authentic student responses that are generated in the course of the lesson. We conjecture that the reduced need for improvisation may impact on the nature of the discussion. Through questioning, a novice teacher may focus more on students' mathematical reasoning rather than just the steps involved in the construction of a solution. The experience of these discussions may then prompt teachers to practice similar discourse moves within discussions of students' own responses to a problem. Thus we regard designed student responses as a potentially useful pedagogical tool to help teachers shift the topic of conversation from requests for elaboration of the process involved in constructing a solution to requests for elaboration of the meaning of the constructed solution. We see this study as the initial step in what could be an important avenue of research.

The research described here contributes to a growing body of work concerning tools to support teachers in their orchestration of whole class discussions (e.g. Michaels & O'Connor, 2015; Stein, Eagle, Smith, & Hughes, 2008). With a particular focus on mathematical reasoning, we explore the question: for a teacher new to working with unstructured problems, how do discussions of worked-out solutions (called in this paper designed student responses) *differ* from discussions of students' own solutions? To this end we compare how the whole class discusses authentic and designed responses to the same unstructured problem. We begin the paper by considering the research perspective that frames the work. We then outline the potential learning benefits of whole class discussions together with the pedagogical challenges of productively orchestrating them. This is followed by consideration of how designed responses may address some of these challenges. After this we analyze video data of whole-class discussions of both authentic and designed responses to a problem. We then proceed to examine the notable differences between these two types of discussion. We draw on these differences to speculate how practices evident in discussions of designed responses might be transferred to discussions of authentic student work. We finally reflect on how the materials could be refined to provide further opportunities for collective learning.

Definition of terms

To help minimize the risk of misinterpretation we will define some of the terms used in this paper. Firstly, the unstructured problems students attempt to solve are set within a real-world context. To understand the context there is no need for specialist knowledge beyond a teenager's everyday knowledge. We describe these problems as unstructured because, although the solution goal is made explicit, there is little guidance on how to achieve this goal. Thus, each problem may be tackled in different ways depending on a student's current mathematical understanding. The term "novice teacher" is used to refer to any teacher who is relatively new to working with unstructured problems, but not necessarily new to teaching. Later in the paper we will exemplify, in terms of working with unstructured problems, the notion of an "expert" and "novice" teacher. Finally, we use the term "teacher practice" to mean recurrent patterns in communication that draw on a particular system of knowledge (Cole & Scribner, 1975). Thus development of practice indicates change in behavior and extension of knowledge.

Background

This research supports the view that learning takes place within a social framework, distributed among co-participants within a community (Lave & Wenger, 1991; Vygotsky, 1978). Teacher and students work together within the constraints of the social and socio-mathematical norms of the class (Yackel & Cobb, 1996) to accomplish culturally valued goals, involving the use of cultural mediating tools (Lave & Wenger, 1991) including, in our case, designed student responses to problems. Within this setting the teacher's role is not to judge ideas, but to help structure and nurture student talk by carefully organizing turn-taking, sensitively supporting and clarifying student contributions, and prompting students to critically reflect on and make connections between mathematical ideas (O'Connor, 1998). By encouraging students to exchange ideas, a shared meaning may be constructed that individuals would not have attained alone (Sfard & Kieran, 2001).

Benefits of whole class discussions

Within whole class discussions, important mathematical ideas may emerge and be identified with and related to other ideas. For instance, two students may have used distinctly different approaches to solve a problem (e.g. a scale drawing and a numeric solution), but by drawing students' attention to relationships between variables connections may be made. Networks of related ideas may be strengthened (Silver, Ghousseini, Gosen, Charalambous, & Font Strawhum, 2005), enabling students to achieve "a coherent, comprehensive, flexible and more abstract knowledge structure" (Seufert, Janen, & Brunken, 2007).

Empirical studies suggest that students can develop metacognitive behaviors, characteristic of expert problem solvers, when they are encouraged to make judgments about the work of others. By commenting on the work of peers, students' criteria for success are made visible for scrutiny (Black & Wiliam, 1998), differences surface and opportunities arise for students to reflect on and adapt their success criteria to accommodate new values. Critically reviewing solutions in this way may enable them to develop more self-regulated approaches to solving problems (Brousseau, 1997), without being solely dependent on the memorization of the various procedures embedded in each approach (Brousseau, 1997; Chazan & Ball, 1999; Lampert, 2001; Stein et al., 2008). Moreover, encouraging students not only to make sense of a solution but also to make judgments about its quality may shift their perspective from viewing solutions as a process to viewing them as objects to be interpreted and evaluated. This shift can promote deeper understanding of the mathematics (e.g. Sfard, 1991). These benefits can help both students actively contributing to the discussion and those who make no contributions, but are actively listening (O'Connor, Michaels, Chaping, & Harbaugh, 2016).

Challenges of orchestrating whole class discussions

The research literature demonstrates that there are significant pedagogical demands involved in productively orchestrating whole class discussions (Carlson, 1999; Lampert, 2001; Sherin, 2002; Swan & Burkhardt, 2014). Students use varied and unanticipated approaches, and unforeseen difficulties may arise. Teachers can find it challenging, in-the-moment, to select appropriate samples for class discussion, follow students' chains of reasoning, and then synthesize these ideas in order to tease out significant conceptions (e.g. Stein et al., 2008). This is particularly the case when students' thinking is a work in progress, incomplete, includes errors, and is not clearly communicated. The teacher may experience a tension between maintaining a student's social status, avoiding insincere praise, and ensuring emerging ideas are understandable to peers. In the midst of a discussion novice teachers, unable to draw on knowledge available to experts, may be uncertain how best to respond to what students say or do. Their self-efficacy as teachers may be diminished (Smith, Hardman, Wall, & Mroz, 2004; Stein et al., 2008).

Teachers, whether novice or expert, may begin a whole class discussion by asking a student an open-ended question such as "How did you solve the problem?" As this type of prompt is open to a range of explanatory or descriptive interpretations it allows most to participate. A disadvantage is that student responses may not include any mathematics or their ideas may be difficult for other members of the class to follow. Asking effective follow-up questions requires an extensive network of teacher knowledge. Required is knowledge of mathematical content, pedagogy, and students as students (Ball, Thames, & Phelps, 2008). Only through skillful questioning by the teacher can individual understandings be used to develop the collective understanding of the class. This involves achieving the appropriate balance between giving students authority over their emerging ideas, ensuring that these ideas are held accountable (Chazan & Ball, 1999), and upholding students' social status. Students need to provide enough detail to enable teacher and peers to comment and build on their ideas (Sfard & Kieran, 2001), and make connections (Franke et al., 2009) between them. At the same time, perceived personal risk of participating in the discussion needs to be minimized. It is perhaps understandable that many teachers prefer a "show and tell" model in which students take turns to present their ideas; these are publicly acknowledged and applauded, but there is little critical engagement with them. The absence of questions probing students' mathematical reasoning limits the discussion to what was done, rather than why it was done. Students' mathematical authority may be fostered but without accountability the development of the collective mathematical reasoning of the class is constrained (Mercer, 1996).

For teachers to flexibly and productively engage with students' ideas a form of *knowing* in the moment of instruction is required. Ball and colleagues (2008) termed this kind of knowing *knowledge-in-action*. Even when novice teachers understand *how to* interact with students, productively improvising their way through a discussion can be challenging.

Method

This leads us to the question: how can novice teachers develop their knowledge-in-action? We speculate the designed student responses provide a setting for novice teachers to practice discourse moves that would not normally be available, particularly when teaching students new to problem solving. They open the door to discussions of distinctly different, often powerful, responses to unstructured problems, a door not guaranteed to be open when discussing authentic student responses.

Student resources: designed student responses

Using worked-out examples (or designed student responses) is a key pedagogical tool in many instructional settings (Renkl, Gruber, Weber, Lerche, & Schweizer, 2003). In this study, they were used to prompt further reflective analysis after students tackled the problem for themselves. The aim was to encourage important mathematical ideas to surface and misconceptions to be tackled. Their design was guided by both the learning goals of the lessons and the responses of students to the problem under consideration. For example, if the learning goal was to compare and critique the value of alternative representations to a problem, the designed student responses illustrate a range of such representations (e.g. a graphical solution, a tabular solution and an algebraic solution). Alternatively, if the intention was that a particular misconception should be highlighted, then this was included in a response (Booth, Lange, Koedinger, & Newton, 2013). We sometimes left designed responses incomplete so that students may be set the challenge of interpreting and continuing the methods illustrated (see Stark, 1999).

We typically included a response that students would produce unprompted along with two more responses that many students may not have yet considered. The intention was that including a familiar response would aid students' ability to make sense of unfamiliar ones. To support this process and so facilitate "connection making," we carefully constructed the solutions. For example, we included the same values for variables in two solutions (e.g. the coordinates of a point of a graph were explicitly specified in one response and the same values were then repeated in a response consisting of a table of numbers).

The anonymity of the designed student responses means that teachers need not be concerned about issues around exposing students' incomplete or incorrect thinking. Also, the emotional aspect of students publicly describing their own work for others to review is removed. Instead, they describe the solution of another, unknown to the class. Similarly, teachers can ask direct questions such as, "What are the advantages and disadvantages of this approach?" without worrying about maintaining the author's social status or credibility as a mathematician (Evans & Swan, 2014). This may prompt students to voice negative as well as positive comments about a solution. Encouraging students to scrupulously think about other students' thinking may help them to think about their own thinking when solving a problem. And as previously stated, this in turn may improve their metacognitive skills. Equally, explaining and judging another, anonymous student's solution may motivate more students to participate in the discussion.

Hypotheses

With respect to our research questions, we considered the following hypotheses.

Compared to discussions of authentic solutions, when discussing designed student responses:

1. The teacher will ask more follow-up questions, explanatory questions, and evaluative questions.
2. Students' responses will include more mathematical reasoning.
3. There will be more connections made between responses.

The intervention

The study consisted of one teacher, teaching six problems to the same heterogeneous class of thirty Year 9 (13 and 14-year-olds) mixed gender students in a UK secondary school. For the study, we selected a teacher who, although they had three years teaching experience, had little prior experience of working with unstructured problems, or of using designed student responses. Before the intervention the teacher confirmed her "novice" status. This was verified when observing the first intervention lesson. As we outline later in the paper, she did indeed use the practices of a novice. This was particularly evidenced by how she probed student thinking through questioning.

Before and during the intervention the teacher had access to all the teacher guides accompanying each problem. They explicitly provide meaning and coherence to the resources, and so help restrict the range of possible interpretations of the resources. Each guide outlines the structure of the lesson, includes both the problem and designed student responses and offers detailed pedagogical guidance for the teacher. For example, a guide clearly states the lesson goals, provides suggestions for formative assessment, and includes examples of issues students may encounter. In the study the guides provided the only direct teacher support. The teacher reported that she had read the guides.

The typical structure of a pair of intervention lessons is outlined in Figure 1. Figure 2 shows one of the problems used in the intervention, together with three designed student responses.

Prior to lessons (20 minutes)

- Students individually tackle a problem
- Then the teacher reviews the student work and provides written suggestions in the form of questions to help students develop their thinking.

Lesson A: In pairs students tackle the problem (60 minutes)

- Students use the teacher's suggestions to individually review their own work.
- Then, in pairs, students share their initial ideas and construct a joint solution to the problem. Students are encouraged to write their thinking on a poster.
- Finally, in a whole class discussion students share their approaches.

Lesson B: In pairs students analyse designed student responses to the problem (60 minutes)

- Students discuss the designed student responses to the same problem they had worked on in Lesson A. There are usually three responses.
- Then in a whole class discussion students discuss the designed student responses.

Figure 1. Typical structure of a Formative Assessment Lesson

Students participating in the second whole class discussion would no doubt have more understanding of the problem. Similarly, the teacher would know more about student thinking. This knowledge will influence the nature of the two discussions. This we acknowledge as an inherent limitation of the design of the study.

Lucky Dip

Dominic has devised a simple game.
Inside a bag he places 3 black and 3 white balls. He then shakes the bag.

He asks Amy to take two balls from the bag without looking.

If the two balls are the same color then you win.
If they are different colors then I win.

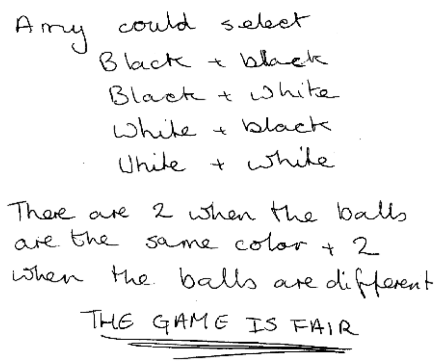
OK.
That sounds fair to me.

Is Amy right? Is the game fair?

If Amy is wrong, then who is most likely to win?

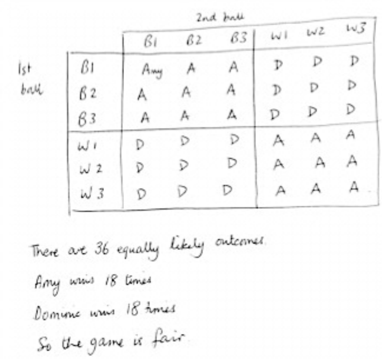
Show all your reasoning clearly.

Designed student response: Anna



Amy has provided a descriptive list of the four possible outcomes. She has not considered that some outcomes are more probable than others.

Designed student response: Ella



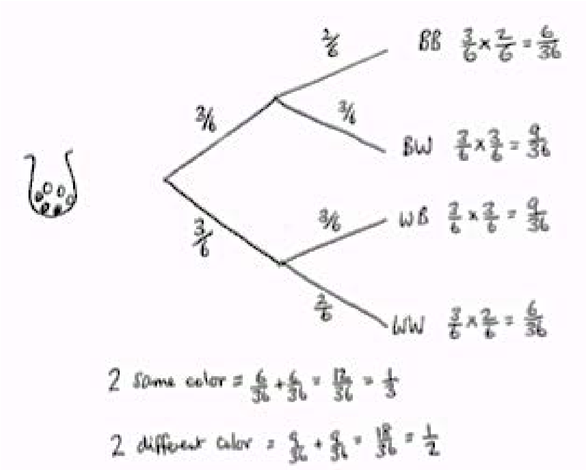
2nd ball

1st ball	B1	B2	B3	W1	W2	W3
B1	Amy	A	A	D	D	D
B2	A	A	A	D	D	D
B3	A	A	A	D	D	D
W1	D	D	D	A	A	A
W2	D	D	D	A	A	A
W3	D	D	D	A	A	A

There are 36 equally likely outcomes
Amy wins 18 times
Dominic wins 18 times
So the game is fair

Ella has drawn a sample space diagram of the situation. She has not realized that the same ball cannot be taken from the bag simultaneously, twice.

Designed student response: Jordan



Jordan has drawn a tree diagram of the situation. He has not accounted for only five balls remaining in the bag after the first ball is removed.

Figure 2. Example of an unstructured problem together with three designed student responses.

Note: The materials in Figure 2 are from the resource Representing Conditional Probabilities (http://map.mathshell.org.uk/materials/lessons.php?taskid=409&subpage=problem).

The resources were designed for US classrooms, however for practical reasons the case study described here was undertaken in a UK classroom. To account for the change in country we undertook minor revisions to the resources (e.g. the language used).

We will now explain the significant design intentions of the resources shown in Figure 2. The solution goal to the Lucky Dip problem is clear: to decide if the game is fair. However, because the route to this goal is not specified, students need to decide for themselves a strategy and then apply it. Thus we consider the problem to be unstructured.

The three designed student responses can potentially reveal different information about the problem situation. For example, students may focus on the meaning of the symbols rather than how the sample space diagram was constructed. In contrast, the tree diagram may encourage an operational understanding of the steps involved in its construction. Students may understand the solution by sequentially working through the procedures within each branch of the tree diagram. Providing responses that may be conceived differently can support students' understanding of the problem-situation.

For the Lucky Dip problem, all designed responses deliberately included common mistakes for students to correct. The intention was to signal to students that mistakes can be a source of learning. Discussing them as a class can lead to productive results for many. Recognizing the benefits may reduce the stigma attached to making errors (Staples, 2007). The errors are conceptual rather than technical. They can be regarded as symptoms of incorrect reasoning. By noticing them, students may be prompted to check their own understanding of the problem-situation. For instance, the sample space solution may encourage students to question whether the first ball selected is returned to the bag before a second ball is selected.

Analyzing whole class discussions

To minimize possible contextual effects, the lessons videoed were all with the same students and at the same time each week. We analyzed video clips and transcripts of three pairs of lessons; the fourth, fifth and sixth pairs. Although we videoed all lessons we decided to restrict our analysis to these lessons for two reasons: firstly, we wanted to allow time for the teacher and students to become familiar with the routines and structure of the lessons, and secondly we wanted the teacher and students to become more comfortable with the process of being filmed as they discussed their reasoning.

To address the tension between interpretive explicitness, pragmatic reduction of data and adherence to our perspective on learning, we undertook a multi-method approach. We began with a quantitative analysis of all three pairs of whole class discussions. This provided us with an overview of what was happening in the discussions. In an attempt to make the quantitative findings explicable within the context of the discussion we then took a narrative approach to select and analyse video clips (Derry et al., 2010).

Based on the research of Swan (2006) and Mercer, Dawes, Wegerif, and Sams (2004), we first constructed a framework focusing on the types of questions teachers ask students during a whole class discussion and the types of replies students give to these questions. This construct provides a useful means of comparing teacher practices when conducting whole class discussions of authentic responses and designed student responses. The framework for this analysis is shown in Table 1.

Table 1

Framework for analysing the video data of whole class discussions

	Teacher pedagogical strategies	Student contributions
Ensure ideas are made accessible to peers	• Ask a range of questions including descriptive and explanatory ones. • Probe student thinking by asking follow-up questions.	Use of reasoning to describe and explain their or their peer's ideas.
Ensure ideas are evaluated	• Ask evaluative questions. • Probe student thinking by asking follow-up questions.	Use reasoning to critique their own and a peer's ideas.
Ensure ideas are mathematically connected	• Ask questions about how responses are connected. • Probe student thinking by asking follow-up questions.	Use reasoning to connect differing solutions.

Teacher questions were categorized as (Swan, 2006):

Descriptive: Questions intended to elicit descriptions or clarifications of methods. For example, "Can you tell me about what you have done?"

Explanatory: Questions intended to elicit mathematical reasons or justifications. For example, "How does this assumption influence the answer?"

Factual: Questions intended to elicit the recall of mathematical facts, numerical answers, rules etc. For example, "How can I write that in symbols?"

Generative: Questions intended to promote generalisations or hypotheses or that will promote further thinking about the situation. For example, "Has anyone got a conjecture? Can anyone suggest any further examples of this? Which way do you think would be the best way of showing this information?"

As we were also concerned with how students evaluated each other's solution, we included an additional teacher question:

Evaluative: Questions intended to elicit the critique of a method (including their own). For example, "What mistakes has she made? How can she improve her work?"

None of these question types explicitly address whether the teacher asked students about making connections between solutions. We decided this construct was implicitly embedded within all the existing list of teacher questions. As such, the question-type "Making Connections" (MC) formed a subcategory of the five questions. This is consistent with Swan's (2006) original work. For example, he provided the question "Why do these two expressions always give the same answer?" as an example of an *Explanatory* teacher question. Thus, this question can be categorized as an *Explanatory* question that *Makes Connections*.

Whole class discussions are an opportunity for the class to develop their collective learning. To do this productively students need to participate by listening or talking. Thus, we decided to also identify teacher questions that required students to have listened to previous talk (denoted as "listening" questions in the Results section). For example, when the teacher asked the class: "So when they weren't sure, what did they do?" she was referring to a previously explained solution. Furthermore, to encourage students to describe their ideas in sufficient detail for others to understand and comment on, the teacher needs to ask probing questions that build on a student's previous response. Thus, we decided to identify teacher questions that demonstrated the teacher was probing students' initial contributions. These we referred to as "follow-up" questions. An example of a follow-up question is: "How did you work out that that [assumption] made a difference [to the answer]?" In an attempt to simplify the results these two types of teacher questions ("listening" and "follow-up") were not to be treated as subcategories of the existing five teacher question types. Instead we categorized them separately, in a second cycle of coding of the video transcripts of the discussions.

To investigate whether students included reasoning in their responses we drew on the work of researchers Mercer et al. (2004). Although their study focused on how students reason in science lessons, we considered that there was sufficient overlap in the aims of the two studies to justify reproducing their method. For example, the researchers investigated the presence in student talk of skills such as reasoning about outcomes, formulating hypotheses and critically examining explanations. As we also sought to gather evidence of these discursive skills, we decided to examine student contributions for the "indicator words" Mercer and colleagues associated with reasoning. These words are "because," "if," "I think," "would" and "could" contributions. For example, "It is not an even chance *because* there are more blacks than whites [balls]" (one indicator word), or "*If* you kept it out [the ball] the student [sample student] *would* be right" (two indicator words). We decided to count *all* indicator words apart from those within sentences that indisputably do not indicate mathematical reasoning (e.g. "after school I'm going to the shop *because* I want to buy some sweets"). We conjectured that if the teacher asked a preponderance of explanatory questions then correspondingly student responses would include a multitude of indicator words.

We also decided to identify those student comments that focused on critiquing solutions and those that focused on making connections. This would help ascertain whether the teacher asking evaluative questions, or questions that encouraged students to make connections between solutions, influenced student responses.

We recognize that a purely quantitative analysis does not allow us to represent, in all its complexity, how the discussions unfolded over time, and could potentially mask important contextual factors (Franke et al., 2009). In an attempt to address this issue, the quantitative results guided our selection of episodes to qualitatively examine. With the specific aim of situating the quantitative results within the narrative of the lesson, we drew on the three constructs outlined in Table 1 to select clips. The aim was to illustrate what the constructs *sound* like in the whole class discussions of authentic and designed student responses. For example, we explore episodes in which students evaluated ideas or the teacher asked a series of follow-up questions. In pursuit of coherence and because of limitation of space only one pair of lessons was analysed, the fifth. We chose this lesson because the whole-class discussions were of substantial and similar length: about 12 minutes duration for Lesson A and 15 minutes for Lesson B.

Analysis of the Data

Quantitative results

We first present the quantitative analysis of the three pairs of whole class discussions. We then focus on the qualitative analysis of two whole-class discussions of the Lucky Dip problem.
Two researchers independently coded each transcript and overall there was above 90% consistency. The few remaining inconsistencies were then reviewed jointly.

Tables 2 and 3 below summarize the data gathered from the whole class discussions in the three A lessons and the three B lesson.

Table 2

Summary of number of teacher questions in the two types of discussions

		Lesson A Discussion of authentic student work	Lesson B Discussion of designed student work
Coding of teacher questions using Swan's (2006) categories	Descriptive	23 (67%)	27 (39%)
	Explanatory	3 (9%)	15 (29.5%)
	Evaluative	1 (3%)	15 (29.5%)
	Factual	6 (18%)	1 (2%)
	Generative	1 (3%)	0 (0%)
	Total	**34 (100%)**	**58 (100%)**
Coding of teacher questions using our own categories	Listening	7 (21%)	18 (31%)
	Follow-up	16 (47%)	25 (43%)
	Other	11 (32%)	15 (26%)

Note: There were no attempts by the teacher to make connections between solutions. The percentages represent the proportion of questions asked within Lesson A or within Lesson B whole class discussions. *Other* refers to the teacher questions we could not categorize as *Listening* or *Follow-up*.

Table 3

Summary of student contributions to the two types of discussions

		Lesson A Discussion of authentic student work	Lesson B Discussion of designed student work
Coding of student contributions: *Number of indicator words used*		17	72
Coding of student contributions: *Number of comments*	Making connections	0	0
	Critiquing solutions:	2	25

In a post-intervention interview, the teacher stated she regarded whole class discussions of authentic student solutions as an opportunity to publicly acknowledge and celebrate student work, with the expectation that peers would learn through listening. This appears to indicate she regarded the discussions mainly as a "show and tell" activity (see Stein et al, 2008). Moreover, data in Table 3 confirms this emphasis on the value of listening. There was notable variation in the proportion of teacher questions that could only have been answered if students had listened to previous talk (21% in Lesson A compared to 31% in Lesson B).

The teacher opened most discussions with broad, highly inclusive questions such as "Who wants to share ideas?" (Lesson 4-A), or "Does anyone want to tell me about the [designed student response] work?" (Lesson 6–B). Students in both types of discussion often responded with a procedural description, with little elaboration as to the purpose of the procedure. For example, a student refers to creating a tally, "I did a tally of their favorites. I then added them up…" (Lesson 6–A). Or, "Basically she multiplied 3 by 0" (Lesson 6–B). Probing these responses occurred more in the discussions of designed student responses than in the discussions of students' own work. Furthermore, of the teacher's questions, there were proportionally more of an *explanatory* nature (29.5% compared to 9%). These questions included the request for clarification of ambiguous explanations, or the reasoning underlying a procedure, or the exploration of an error the "sample student" made. For example, in Lesson 6–B, the teacher uses a series of follow-up questions to encourage the students to move beyond their initial vague descriptive response. She pressed for a more detailed explanation, which in turn could facilitate the collective understanding of the class. Correspondingly, there were many more student responses containing "indicator words" in the discussions of designed student responses (72 compared to 17). A student commented, for example, in a Lesson 5-B discussion, "… that's the point where you pick out one ball so it should only be 5 because they've just picked out a ball, so there's only 5 left." This quantitative data suggests students tended to explain and justify their thinking (or the thinking of the "sample student") more in the Lesson B discussions.

Another notable difference between the discussions was the number of evaluative teacher questions (1 for the A lessons and 15 for the B lessons). When discussing designed student responses, students not only attempted to explain the work of others, but were expected to critique it. Correspondingly, there were more instances of students critiquing a designed sample solution than authentic solutions (25 compared to 2). We conjecture the teacher did not consider the class ready to have their own work critiqued by their peers. This is supported by a comment she made in a post-intervention interview: "I don't want students' only participation in the discussion to be about something that they have done wrong. When working with designed student responses, students do not need to worry about whether they've got the problem right or wrong, they can simply talk about their ideas." However, there were no attempts by the teacher in any of the discussions make connections between the solutions, nor did students offer any connections.

Qualitative Analysis of Video Clips from the Lucky Dip Whole-Class Discussions

As previously mentioned, the video episodes framed by the constructs outlined in Table 1, are selected to help illustrate of the discussions.

Whole class discussion of authentic student work (Lesson A)

In the collaborative activity preceding the whole class discussion, most students tackled the problem descriptively (using just text, and no explicit mathematics), although two pairs constructed diagrams and one pair used an empirical approach. Within the whole class discussion eight different students spoke. The discussion focused on three distinct student approaches to the problem. We consider all three here. The first strategy consisted of verbal descriptions relating to why students considered the game fair or not. The second approach used a diagram in which balls had been labeled, then the balls were pictorially combined in order to figure out probabilities. By repeatedly selecting balls from a bag, the third strategy used an empirical approach. These three methods were not explicitly compared or connected and only one method was briefly critiqued. A prominent feature of the discussion was that all substantial student contributions (e.g. more than a sentence long) were difficult to interpret.

First student method

The discussion (no accompanying written work) proceeds as follows:

S1	Okay, if you pick a black, then you pick a white ball out of the bag, then if you want to pick another out there's still three black balls in the bag, but there's only two white balls left so there's a three to, three in five chance of getting a black ball instead of a white ball, so if you think about it, there's more chance of getting a black ball and no matter how many more balls you add, like if you had one black one and another and add another white ball, then no matter which one it'd still be neither
T	Okay, so what was your conclusion then? Was the game fair or not?
S1	No
T	No? Okay. So you decided that
S1	That Dominic is a bit of a cheat
T	(laughs) that Dominic is a bit of a cheat (class laughs)

The solely verbal approach allowed the participation by a student who was still developing his understanding of the problem-situation. However, without the use of a visual representation, expressing his ideas in a way meaningful to peers required skill and precision and/or a preparedness on the part of the teacher to pursue ambiguities. Both were lacking. The student colloquially described the situation without reference to the specificity of formal mathematical language. Rather than probe what, in-the-moment, may seem to be muddled thinking, the teacher distracted the class by encouraging laughter. This strategy may be due to her commitment to maintaining, if not boosting, students' social status whilst publicly discussing work. We suspect however, that although S1 has poorly expressed his thinking, there is sense-making within it. Furthermore, generalizations were emerging: provided the numbers of black and white balls are equal, the game would remain unfair even if more balls were added to the bag. This was not followed up through teacher questioning.

This episode in the discussion provides an example of both the challenges of understanding students' incomplete thinking and possibly reluctance on the part of the teacher to risk damaging the student's social status by probing their potentially fragile understandings of the problem.

Second student method

The next three substantial student contributions to the discussion were predominantly descriptive with little explicit reference to mathematics. They depicted incomplete thinking that could again be challenging, in-the-moment, to interpret and respond to constructively. Although the teacher did ask one follow-up question and did attempt to draw others into the discussion, she did not ask students to clarify their explanations. We suggest that this again was driven by a reluctance to risk publicly exposing the student's lack of understanding.

Possibly because of the difficulties understanding student's remarks, the teacher chose to explain the second student's method. The student had labeled the three black balls A, B, or C and the three white balls 1, 2 or 3. They then listed correctly the thirty possible permutations; eighteen when different colored balls were selected (e.g. A1) and 12 the same colored balls were selected (e.g. AB).

The episode begins with a couple of introductory questions:

T	Okay, what did you find out then, or what did you do?
S4	Umm [pause] we found all the, like all the things you can get
S5	Combinations
S4	Yeah
T	Oh excellent word S5. Excellent, good. Okay, all the different combinations
S4	Yeah and we found that there's 12 in 30 chance that you can get, like, two numbers or two letters
T	By numbers and letters what do you mean?
S4	Like all the black numbers or letters and all the white numbers and numbers
T	Okay, so can you hold your board up, [to the class] everybody look at (S4) because I think it needs to be, right what (S4) and (S5) did was they actually all the black balls they gave a letter to, okay and the white balls they gave a number to and then they wrote out all the different possible possibilities that could be picked out, okay, so S4 has written A1, so you could have a black one and the first white one, you could have A2, could have the black one and second white one, okay and then she said 'right, what are all the possibilities where you'd get two letters or two blacks? And how many possibilities are there where you'd get two whites?', [to S4] and you counted those up didn't you? and used a fraction to explain them, okay, how did you work out the 18 over 30, what's that for?
S4	That's for ones where they are like not the same
T	Where they're not the same?
S4	Yeah
T	So when Dominic wins? Okay, so what did you decide overall then?

S4 Well it's not fair on Amy because Dominic has an advantage
T Ah okay, excellent, good, well done, well explained as well guys

The teacher then draws the students' attention to the technical term 'combination.' In doing so, she appears to reinforce the notion that they are part of a community of school mathematics in which exists a unique language to be used when engaging in mathematical talk. The teacher then explains to the class the student's approach. This ensures a clarity that may have been missing from the student's own explanation, however she limits her description to a procedural one. In an attempt not to assume ownership, she stresses throughout that she is describing students' (S4 and S5) work. This authorship is reiterated when, after her explanation, she asks a question to draw S4 and S5 back into the discussion. The final remark ("excellent, good, well done") seems inappropriate as the students had explained little and the teacher's judgment cuts off the possibility of further discussion.

Third student method

The episode was followed by student descriptions of an empirical approach. Again student contributions seemed difficult to follow, however during this episode the teacher chose not to reiterate them or encourage peers to comment on the approach. She simply stated their results: "when you picked them [the balls] out you normally picked out a white one and a black one? [student nods]." We speculate this indicates that although the teacher sought students' participation and their work to be publicly recognized, her priority was not that others understood this particular approach.

Whole class discussion of designed student responses (Lesson B)

In the collaborative work all three designed student responses were considered, however each pair of students worked on just one. Within the whole class discussion groups of students took turns to stand at the front of the class and explain a response. Ten different students spoke. Six of these spoke in the previous discussion of authentic student responses. Most student contributions were easier to understand than the previous discussion about students' own work. Throughout, students referred to the responses displayed on the large whiteboard positioned at the front of the class. Without this resource, students would have had difficulties understanding and following the conversation. Within the discussion students explained methods and critiqued them, but did not attempt to make any mathematical connections between them.

The discussion opened with the teacher asking "Explain what Anna has done, who wants to have a go at that?" This is followed by an episode that included eight follow-up teacher questions. For example: "She hasn't explained it very well, what do you mean?" Furthermore, on two occasions, by linking two students' contributions, the teacher emphasizes the collaborative aspect of the discussion. For example:

S5 They haven't really gone into detail about why it's fair, they've just said 'it's fair' without actually explaining it to anyone why it's fair
T Right [to S7] is that what you meant S7?
S7 Yeah
T Yeah, good, excellent, okay, [to S6] did you want to add anything?
S6 I was gonna say the reason, what [S7] was trying to say was she hasn't explained how it is fair, all she's done is written a load of combinations down, put two of the same colour and the game is fair, she hasn't explained how it is fair

Earlier in the discussion student S7 had tentatively stated "we just put: she hasn't really explained why it's fair.'' By linking S5's comment to S7's, the teacher validates their contribution and attempts to draw the student back into the conversation. S6 also acknowledges S7's contribution and then expands on it. Here is a clear example of, through teacher guidance, three students publicly recognising and then building on each other's ideas. Another feature of the discussion was the focus on understanding the responses. For example, when considering Ella's sample space diagram, S9 states:

> "We didn't understand it [Ella's response], we didn't know what any of the letters meant, like, yeah, we only knew what the A and like B mean but we didn't know what the B1 means."

To which the teacher responds with an evaluative question:

> "But that's important S9's group [referring the small group S9 had worked with] because what perhaps should Ella have done? Or might have been better for Ella to have done?"

The teacher's response (above) is to the whole class. She is legitimizing lack of understanding of a response; it should not be regarded as a deficiency on the part of the student, rather it is a deficiency on the part of the anonymous student (Ella). Following this comment, another student admitted he'd guessed what Ella's table represents and when a third student confessed his lack of understanding, a peer, unprompted, goes to the front of the class to explain the work.

In a later episode, after the teacher had through repeated questioning pursued the notion that Ella may have made a mistake, a section of dialogue occurred in which there was little input from the teacher, apart from asking an evaluative question:

S8	I know where she's gone wrong [design student, Ella]
T	She made a mistake?
S8	Yeah
T	Go on then, what's she done?
S8	Because right, I'm going to point at it, because if I pick out black 2 I can't pick out black 2 again, so she's not going to win, nobody wins because you can't pick it out again like, so white 1 you can't pick out white 1 again, so she can't win [the numbers refer to the numbering in the sample space diagram in Figure 3]
S5	But surely you could pick out white 1 white 2, white 1 white 3
S8	Yeah you can 2 and 3 but you can't pick out white 1 again if you've already picked out white 1
S5	But I put it back in
T	This is a really good discussion, well done
S8	But if you put it back in you can't, you're not going to have two whites are you?
S9	No, but then you chart it, like oh I got white 1 and then you like put it back in and then
S8	It doesn't say you put it back in again
S9	Otherwise, otherwise if you leave it out then it's not an even chance
S6	No but if you put
S6 & S8	[students talking at same time] (inaudible)
T	So what do we need to know to help us with this discussion?
S9	Whether they put it back in or leave it out

S8 seems to be questioning the validity of the solution; winning by selecting the same ball twice makes no sense to him. On the other hand S5 and S9 assume the table is accurate and correctly interpret it to mean the first ball selected is returned to the bag. Students appear to be committed to the problem, they listen and respond to each other's comments, but it is unclear whether S8 has shifted his thinking. However, when discussing Jordan's response (Figure 3) it appeared that S8 did now understand the context:

S8	Oh right, so what we found out was that if you go downwards first then its white because the first letter starts with Y and if we go up its when you pick a black ball first, [to S7] this is if, if you do pick it out the bag and you keep it out of the bag Alicia (laughs) right so if you go and pick another black one, this if you pick a white one, black one, white one, but what we found out with that is it's still a sixth chance of getting the same ball if you know what I mean, so it should be 5 not 6 because you've already picked one out of the bag
S6	Because you've picked a ball out and you're keeping it out, so there's only 5 balls left in the bag, so as soon as you get to this point, you've already, that's the point where you pick out one ball so it should only be 5 because you've just picked out a ball, so there's only 5 left
S8	So that's like one mistake he's made and the other mistake he's made is obviously following on from there he's done it here as well
S6	He's done all the multiplications as though there's 6 balls still in the bag.

S8 may have needed time to modify his understanding of the context, or he may have been unwilling to publicly admit he had changed his mind. This is a second episode within this discussion in which there was little teacher input. Being prepared to "go with the kids" (Staples, 2007) may indicate the teacher's growing confidence in her students' commitment to meaningfully discuss the mathematics.

S6's role in this discourse appears to be to re-voice S8's explanations. However, he has not simply replicated the explanations, he has emphasized (presumably) aspects of the solution he or his partner found difficult or important. This indicates a desire not only to participate in the discussion, but a commitment to developing a shared understanding of Ella's response. This sophisticated conversational device also slows down the talk, ensuring most have heard the challenging parts of the explanation at least once.

Within the exploration of Ella's response the teacher asks a series of questions concerning students' assumption that a ball is taken out of the bag but returned to it:

> "So you're assuming that he doesn't put the ball back in?"
>
> "So what about if he did put the ball back in?"
>
> "So you're assuming that his mistake that the number 6 on the bottom?"
>
> "What if his mistake was the number 2 on the top?"

Students' initial responses indicated they were reluctant to recognize they were making an assumption, with one student stating:

> "The reason that I know he's took it out and kept it out is because he wouldn't put a 2 on the black that he has took out if you hadn't already of kept it out, so he's still only got that 2 chance of getting that black ball because he's got one in his hand and it's the same for the white one."

However, after the fourth teacher question a student linked their response to a student's earlier contribution:

> "Yeah its just the same, so if, if you were putting it back in the bag like (S7) said the 6 would be right, but if it was not, if you kept it out, the 2 would be right."

The teacher appeared satisfied with this response. She re-voiced it.

Flexibly and productively engaging with students' ideas within the public situation of a whole class discussion is challenging for teachers. The teacher appeared to find discussing designed student responses that are coherently presented and authored by others less demanding. She probed students' understandings more, including asking questions that promoted reasoning and critical thinking. Correspondingly, the students responded with elaborations of their ideas in a way that was not evidenced when discussing their own work.

Discussion

There was one notable similarity between the two types of discussions. In both, the teacher makes no attempt to encourage students to link solutions and students did not spontaneously notice connections. This concurs with the literature that suggest students typically do not discern connections between different representations on their own. They need support from the teacher. (Van Someren, Reimann, Boshuisen, & de Jong, 1998). In the discussion of designed student responses student ideas were connected and built on but there was no explicit connections made between responses. We are uncertain why the teacher did not encourage students to make connections between solutions. It may be because the teacher guide only provides generic suggestions about the importance of making mathematical connections between responses. Or it may be that it is a challenging pedagogical strategy that requires more professional development than a guide is able to provide. This challenge was accentuated because the teacher chose to give each student pair just one designed student response to work with, rather than all three as the guide suggests. Therefore during the discussions, many students will have lacked the experience of generating their own ideas about a response to support the understanding of other students' expositions. Under these circumstances, making connections between solutions may have been simply too challenging.

The analysis of the data did indeed indicate there is a distinct difference between the discussions of authentic and designed student responses. Student explanation of their own work is predominantly procedural and often difficult to understand. The teacher often fails to probe students' burgeoning ideas. As most substantial student contributions were difficult to interpret, lack of clarification means it would have been difficult for other students to learn from, comment about, or use to build on their thinking. However this was not the situation when discussing *known* designed student responses. The teacher presses students to move beyond simply describing procedures to justifying their thinking. Students are encouraged to evaluate the responses, a pedagogical strategy missing from discussions of students' own work. And in turn, student responses include more mathematical reasoning. Students attempt to make sense of the mathematics within the responses and assess the validity of the responses.

The teacher's apparent concern about publicly exposing students' lack of understanding of their own work, together with uncertainty about how best to respond, in-the-moment, to students' emerging ideas, may account for lack of probing teacher questions. It may also be that because most students solved the problem descriptively, there was simply less opportunity for the teacher to ask probing questions. On the other hand, designed student responses guarantee exposure to key mathematical ideas. Factors such as the restricted number of solutions, their anonymity and the focus on understanding solutions rather than performance all contribute to reducing the complexity of improvisation. Designed student responses may provide teachers new to this work a safe way of limiting the socio-emotional component of publicly explaining and critiquing solutions.

The reassuring predictability that can be planned for when working with designed student responses was confirmed during an end of project interview in which the teacher commented: "I find it easier to structure the discussions about designed student work as I know where the talk is likely to go, whereas with their own work, students can take it in all directions." We hypothesize that when holding a planned, whole class discussion on designed student responses, a novice teacher can develop practices and gain knowledge that may be generalized for the more demanding, but clearly analogous, discussion of students' own work. Ultimately, it may help teachers, particularly those new to problem solving, to recognize that whole class discussions can be more than an opportunity to celebrate work, they can be an opportunity for students to learn together.

We recognize there are limitations within the study. The study focused on just whole class discussions in six lessons, taught by one teacher. Discussions are historically situated in each lesson. What happens earlier in a lesson will influence the content and structure of the whole class discussion. Research, for example, indicates how the task is introduced is related to the quality of the discussion (Jackson, Garrison, J, Gibbons, & Shahan, 2013). What had occurred earlier in the lesson was not accounted for in this study. Furthermore, it could be argued that the differences between the two discussions may simply be a product of discrepancies in times. The length of the discussion of designed student responses was longer than the discussion of authentic responses. Similarly, because the discussion of designed responses was after the discussion of authentic responses, some may say it is to be expected students would reason more in the second discussion. After all, students then know more. We argue that the length of the discussion may be partially explained by students' willingness to talk, which was perhaps greater in the designed responses lesson. However, the findings are inevitably securely bound to the local setting, from which the hypothesis regarding the inferred causations rather than generalities has emerged. The hypothesis may then be tested in future studies.

Acknowledgements

We would like to acknowledge the support for the study, the Bill and Melinda Gates Foundation, our co-researchers at the University of Berkeley, California and the observer team.

References

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

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.

Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K. J. (2013). Using example problems to improve student learning in algebra: Differentiating between correct and incorrect examples. *Learning and Instruction, 25*, 24-34.

Brousseau, G. (1997). *Theory of didactical situations in mathematics* (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Trans.). Dordrecht: Kluwer.

Carlson, M. P. (1999). The mathematical behaviour of six successful mathematics graduate students: Influences leading to mathematical success. *Education Studies Mathematics, 40*, 237-258.

Chazan, D., & Ball, D. L. (1999). Beyond being told not to tell. *For the Learning of Mathematics, 19*(2), 2-10.

Cole, M., & Scribner, S. (1975). Theorizing about the socialisation of cognition *Ethos, 3*, 249-268.

Derry, S. J., Pea, R., Barron, B., Engle, R. A., Erickson, F., & Goldman, R. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Learning Sciences, 19*(1), 3-53.

Evans, S., & Swan, M. (2014). Developing students' strategies for problem solving in mathematics: The role of pre-designed "Sample Student Work". *Education Designer, 7*.

Franke, M. L., Webb, N. M., Chan, A. G., Ing, M., Freund, D., & Battey, D. (2009). Teacher questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education, 60*(4), 380-392.

Jackson, K., Garrison, A. J. W., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education, 44*(4), 646-682.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction, 6*(4), 359-377.

Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal, 30*(3), 359-377.

Michaels, S., & O'Connor, C. (2015). Conceptualizing talk moves as tools: Professional development approaches for academically productive discussion. In AERA (Ed.), *Socializing Intelligence Through Academic Talk and Dialogue* (pp. 347-362).

O'Connor, C. (1998). Language socialization in the classroom. In M. Lampert & M. L. Blunk (Eds.), *Talking Mathematics in Schools* (pp. 17-56). UK: Cambridge University Press.

O'Connor, C., Michaels, S., Chapin, S., & Harbaugh, A. G. (2017). The silent and the vocal: Participation and learning in whole-class discussion. *Learning and Instruction, 48,* 5-13.

Renkl, A., Gruber, H., Weber, S., Lerche, T., & Schweizer, K. (2003). Cognitive load during learning from worked-out examples. *Zeitschrift Fur Padagogische Psychologie, 17*(2), 93-101.

Seufert, T., Janen, I., & Brunken, R. (2007). The impact of intrinsic cognitive load on the effectiveness of graphical help for coherence formation. *Computers in Human Behavior, 23*(3), 1055-1071.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.

Sfard, A., & Kieran, C. (2001). Cognition as communication: Rethinging learning-by-talking through multi-faceted analysis of students' mathematical interactions. *Mind, Culture, and Activity, 8*(1), 42-76.

Sherin, M. G. (2002). When teaching becomes learning. *Cognition and Instruction, 20*(2), 119-150.

Silver, E. A., Ghousseini, H., Gosen, D., Charalambous, C., & Font Strawhum, B. T. (2005). Moving from rhetoric to praxis: Issues faced by teachers in having students consider multiple solutions for problems in the mathematics classroom. *Journal of Mathematical Behavior, 24*, 287-301.

Smith, F., Hardman, F., Wall, K., & Mroz, M. (2004). Interactive whole class teaching in the National Literacy and Numercy Strategies. *British Educational Research Journal, 30*(3), 395-411.

Staples, M. (2007). Supporting whole-class collaborative inquiry in a secondary mathematics classroom. *Cognition and Instruction, 25*(2-3), 161-217.

Stark, R. (1999). *Learning by worked-out examples: The impact of incomplete solution steps on example elaboration, motivation, and learning outcomes.* Bern: Huber.

Stein, M. K., Eagle, R. A., Smith, M. A., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*, 313-340.

Swan, M. (2006). *Collaborative learning in mathematics: A challenge to our beliefs and practices.* London: National Institute for Advanced and Continuing Education (NIACE) for the National Research and Development Centre for Adult Literacy and Numeracy (NRDC).

Swan, M., & Burkhardt, H. (2014). Lesson design for formative assessment. *Education Designer, 7*.

Van Someren, M. W., Reimann, P., Boshuisen, H. P. A., & de Jong, T. (1998). *Learning with multiple representations*. Amsterdam: Pergamon.

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458-477.

Authors

Sheila Evans
University of Nottingham, Dearing Building, Jubilee Campus, Wollaton, Nottingham NG8 1BB
Sheila.Evans@Nottingham.ac.uk

Clare Dawson
University of Nottingham, Dearing Building, Jubilee Campus, Wollaton, Nottingham NG8 1BB
Clare.Dawson@Nottingham.ac.uk