DESIGNING A STAGE-SENSITIVE WRITTEN ASSESSMENT OF ELEMENTARY STUDENTS' SCHEME FOR MULTIPLICATIVE REASONING

Nicola M. Hodkowski
University of Colorado, Denver
nicola.hodkowski@ucdenver.edu

Peter Hornbein
University of Colorado, Denver
peter.hornbein@ucdenver.edu

Amber Gardner
University of Colorado, Denver
amber.gardner@ucdenver.edu

Heather L. Johnson
University of Colorado, Denver
heather.johnson@ucdenver.edu

Cody Jorgensen
University of Colorado, Denver
cody.jorgensen@ucdenver.edu

Ron Tzur
University of Colorado, Denver
ron.tzur@ucdenver.edu

In this paper we examine the application of Tzur's (2007) fine-grained assessment to the design of an assessment measure of a particular multiplicative scheme so that non-interview, good enough data can be obtained (on a large scale) to infer into elementary students' reasoning. We outline three design principles that surfaced through our recent effort to devise a sequence of items to assess at which stage—participatory or anticipatory—a child might have constructed the multiplicative double counting (mDC) scheme (Tzur et al., 2013). These principles include the nature and sequencing of prompts, number choice, and authenticity of responses.

Keywords: Number Concepts and Operations, Assessment and Evaluation

Introduction

In this paper we articulate a design process, focusing on assessing upper-elementary students' mathematical reasoning through a written instrument to develop an assessment tool sensitive to students' multiplicative reasoning. Through our project[1], we intended to assess over a 1,000 students, with future follow-ups likely to scale-up. Therefore, an immediate challenge presented itself: how to gain a trustworthy inference into a child's reasoning without conducting interviews with each participating student.

Following the approach of Norton and colleagues (e.g. Norton & Wilkins, 2009), we decided to try producing a written, pencil-and-paper assessment tool[2]. The written format would have to reflect the theoretical distinctions—schemes and stages in students' multiplicative reasoning—that we intend to measure. Our project team decided to assess two of the six schemes articulated by Tzur et al. (2013). In this paper, we focus on the design process, and design principles gleaned from the production of assessment items for the first of those six schemes (further explained in the next section), which marks a child's conceptual leap from additive to multiplicative reasoning. As this is not an empirical study, sections of the paper do not follow a canonical format of a research report.

We contribute an important elaboration on the work of Norton and colleagues (e.g. Norton & Wilkins 2009)—the need to distinguish a stage at which a learner's scheme might be constructed. Their work demonstrated how written assessments could provide a good enough substitute to inferences made from data collected in time-consuming, effort-intensive interview settings. A written assessment constrains and/or complicates the possibility for a researcher (or teacher) to make inferences about students' reasoning, because data obtained from written items hardly indicate language and actions the child might be using (e.g., moving fingers, counting to oneself). To overcome these challenges, assessment items Norton & Wilkins (2009) created for children's fractional schemes were highly correlated with interview-based inferences. Norton & Wilkins (2009) tested a correlation against the underlying schemes a child might be using while successfully solving assessment items—first written and then through interviewing. The premise of their correlation seems to be that a child's ability or inability to solve certain items reflects having or not having constructed the underlying scheme, respectively.

In our project, we take a compatible, yet more nuanced approach. We premise our project on the recognition of gradation in successful solution to assessment items that is contingent upon the stage at which a learner has constructed the measured (inferred) scheme. Specifically, we articulate our design process and posit design principles that attempt to distinguish children who constructed a new scheme only at the participatory stage, which limits a learner's use of an evolving scheme to items only if somehow being prompted (Simon, Placa, & Avitzur, 2016; Tzur & Simon, 2004).

Conceptual Framework

We ground our conceptual framework in constructivist scheme theory (Piaget, 1985; von Glasersfeld, 1995), meaning that we define learning as a conceptual change brought forth through a learner's own goal directed activity (Simon, Tzur, Heinz, & Kinzel, 2004). We use "scheme" to indicate a three-part conceptual 'building block': a situation ("assimilatory template") that sets the learner's goal, an activity triggered to accomplish that goal, and a result that can turn into a prediction. We distinguish two stages in a learner's development of a new scheme: participatory and anticipatory (Tzur & Simon, 2004). Broadly, we classify a learner as "participatory" if she needs a prompt to elicit a new scheme, and "anticipatory" if she does not need a prompt to elicit a new scheme. In the context of assessment, we use "prompt" to refer to a written statement or diagram used to elicit the use of a new scheme that, otherwise, would not be brought forth.

Tzur and Simon (2004) proposed the participatory/anticipatory stage distinction to explain the "next day" phenomenon, which entails the inconsistent accessibility learners seem to have to an evolving, new way of reasoning. Initially, a learner might connect the last two parts of a scheme, as she might notice novel effects that the activity brings about. While the activity and its effect(s) are enacted, the learner can anticipate future 'mental runs' of the activity to yield the same effects. However, at the participatory stage the newly noticed activity-effect is yet to be linked with the situation/goal. Thus, at a later time ("next day"), if/when the activity has not been triggered and 'ran' to bring the newly noticed effects—the learner is likely to 'revert back' to using previously established, anticipatory schemes (Tzur & Lambert, 2011).

To distinguish between participatory and anticipatory stages, our assessment of a learner's way of reasoning needs to be sensitive to a learner's independence (or lack thereof) in using a newly constructed scheme (Tzur, 2007). To make participatory/anticipatory stage distinctions, we follow his recommendation to use fine grain assessment that proceeds from prompt-less items indicative of the anticipatory stage to prompt-dependent items indicative of the participatory stage of a new scheme. We provide a twofold reason for our choice, First, when/if prompted, a learner at the participatory stage can solve assessment tasks in ways that provide 'false positive' data—as if the learner is properly using the new scheme. Therefore, providing prompts too early may prohibit distinguishing a learner at a participatory stage from a learner at an anticipatory stage of the same scheme. Second, if assessment items avoid prompting for an evolving scheme altogether, then the learner's prompt-independent schemes (anticipatory stage) are expected to trump the use of prompt-dependent schemes (participatory stage) (Tzur & Lambert, 2011). In such a case, the learner's responses to assessment items will provide 'false negative' data—as if unable to use the newly constructed scheme, whereas she can actually reactivate and properly use the new scheme once prompted. This possibility for a 'false negative is a key aspect of our elaboration on Norton & Wilkins (2009) work.

Our assessment draws on findings about students' constructions of six schemes for multiplicative reasoning (Tzur et al., 2013). A core goal of our assessment is to identify stages in learners'

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

transition from additive to multiplicative reasoning. We use learners' coordination of composite units—single "things" a learner can conceive of as being made up of sub-parts (Steffe & von Glasersfeld, 1985) —to distinguish between additive and multiplicative reasoning (Clark & Kamii, 1996; Steffe & Cobb, 1998). Consider the problem: *Maria is making treat bags for her 3 friends; each treat bag will have 8 candies; how many candies does Maria need in all?* A child using additive reasoning preserves the type of composite unit (e.g., 8 candies + 8 candies + 8 candies = 24 candies). In contrast, a child using multiplicative reasoning transforms the type of composite unit, and for the candy bag problem there would be two types of composite units (8 candies per bag; 3 bags). A child using multiplicative reasoning coordinates two different types of composite units by distributing items of one composite unit over the items of another composite unit; the result of this distribution is a third, different kind of unit (e.g., 8 *candies per bag* distributed into each of 3 *bags* = 24 *candies*). Tzur et al. (2013) termed this scheme of multiplicative reasoning *multiplicative double counting* (mDC), to reflect the way a child determines the total number of 1s in the coordinated quantity by operating simultaneously on two counting sequences (e.g., 1-is-8, 2-is-16, 3-is-24).

Design Process and Principles

In this section, we articulate our process for designing a written assessment distinguishing between participatory and anticipatory stages of the mDC scheme. We chose the mDC scheme because it reflects a child's crucial conceptual leap from additive to multiplicative reasoning. A minimum of 3 items on a written assessment is considered necessary to provide a valid and reliable score of a theoretically grounded construct (Hinkin, Tracey, & Enz, 1997; Raubenheimer, 2004). To play it safe, our team decided to begin by developing five items – up to 2 of which could later be eliminated. This paper focuses on the creation of one item—a sequence of questions—the design of which proved most challenging. This item, initially asking (with drawing like Fig. 1) "How many small, gray boxes are needed to fill the large box?", required most of our effort and recurrent revisions, which assisted in our explication of design principles.

The goal we set for each assessment item was to provide data to distinguish the stage, participatory or anticipatory (Simon et al., 2016; Tzur & Simon, 2004), at which a child had constructed the mDC scheme. Following the fine-grained assessment rationale (Tzur, 2007), each item had to progress from prompt independent ("hard") to prompt dependent ("easy") questions, so that a child at the anticipatory stage can demonstrate reasoning without any prompts at the start of each item. The assessment item on which we focus here was created as a task of figuring out, in a context of two-dimensional shapes (here rectangles – see Figure 1), how many small units fit into an entire larger unit[3].

Figure 1. A small and a large rectangle providing a context for mDC.

For this item, we took caution to counteract giving a lower-level solution to a child already at the commencing, prompt independent version. For example, a child may simply draw all small rectangles and then count them one by one, which gives no information about a child's mDC scheme. We identified the need for this caution through informal interviews that one of our team members had conducted with her 7- and 9-year old daughters. The girls' use of this "easy-way-out" strategy made us realize that such a solution can yield the correct answer even for a child who is still operating on units of one (1s) and is yet to construct numbers as composite units—let alone

Wood, M. B., Turner, E. E., Civil, M., & Eli, J. A. (Eds.). (2016). *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Tucson, AZ: The University of Arizona.

coordinate such units multiplicatively.

During an interview, an interviewer could present and monitor constraints to make “easy way out” strategies less likely. Furthermore, an interviewer could also probe (or avoid probing) to figure out if the child would spontaneously supply, say, columns (here, a composite unit of 7) and rows (here, a composite unit of 4) to organize and then execute a simultaneous count of the items that constitute each of those units (e.g., 1-has-4, 2-have-8, etc.). However, how might a written format steer students away from the “easy-way-out” strategy, and instead elicit data indicating an anticipatory, simultaneous operation on (coordination of) composite units?

Our tentative solution, which at the time of writing this paper is yet to be tested with more children, brought us face-to-face with the reading and writing capabilities, as well as known or unknown multiples of particular numbers (e.g., 5s vs. 7s), of 3rd and 4th grade students whose reasoning we intend to assess. Thus, on one hand, we strive to reduce the number of words in each assessment item as much as possible, and use active-voice phrases with simple-yet-precise mathematical terminology (e.g., the drawing is of rectangles, not of a cube or a box – which we initially used). We also want to allow a student to present her or his solution (answer + reasoning) while having to write as little as possible. On the other hand, we must provide clear directions about constraints we might have introduced in an interview (e.g., use numbers for which a child does not have a memorized sequence of multiples, such as 7). Several rounds of comments and suggested revisions from our project team members, led to changing the item’s first question into: “How many small, gray rectangles fill the big rectangle? To find the answer, do NOT draw small rectangles.” A drawing like in Figure 1 is presented below this revised question, and below that drawing three more probes are provided in large font:

A: Do you understand the question? (circle one) Yes No

B: Your Answer (fill in the blank): It takes _______ small rectangles to fill the big rectangle.

C: Show how you got your answer. Do NOT draw small rectangles.

With this first question to assess if a child has an anticipatory stage of the mDC scheme, we turned to creating subsequent questions that provide prompts for solving the problem by identifying and coordinating units a child may not yet be able to bring forth spontaneously. For each of those questions, we presented a drawing of the two original rectangles with some changes (see Fig. 2a-b-c), each figure followed by three probes like the ones above.

As our team discussed how to sequence the follow-up questions/drawings, we took notice of three design principles:

1. Sequencing gradually more explicit prompting for the mathematical reasoning at hand (here – units coordination);
2. Selecting “non easy” numbers to mitigate students’ use of memorized facts and to engineer opportunities for students to simultaneously use composite units;
3. Eliciting authentic student responses under the given constraints.

To address the first design principle, we created the first of three prompts (Fig. 2a) for this assessment item to possibly elicit a child’s organization of the rectangular space into two kinds of composite units. To this end, we drew (not fully!) 7 columns and just one row, using the same size as the small rectangle’s dimensions. We considered this as the least explicit prompt of all three, because the child has not only to notice the dimension compatibility but also supply the organization of a unit of four 1s, and then coordinate the accrual of 7 units of 4 units of 1.

The second of three prompts further explicated 4 small units of 1 that would constitute a single, 4-unit ‘column’ (Fig. 2b), to possibly trigger the child’s constitution of an anticipated, similar composite unit for each of the 6 other columns. Because adding those 4 small rectangles could confuse children about the reference in the written question, we ended up extending it as follows: "Maria drew four (4) small gray rectangles in the big rectangle. How many small, gray rectangles fill the big rectangle? To find the answer, you can draw up to 6 more small rectangles."

As seen in Fig. 2c, the third of three prompts explicated almost all small units of four 1s, while still leaving two columns for the child to supply such units and successfully complete the "filling" task. Here, we wanted to ensure a majority of the students would be successful in solving the problem while still having to anticipate the interjection of a composite unit of 4 into the two empty columns, hence showing at least a rudimentary form of an assimilatory scheme for such a unit. For students who may still be unable to solve the problem, this last prompt of the assessment item could provide an indication that, perhaps, they are yet to construct number (e.g., 4) as a composite unit.

The three prompts shown in Figures 2a-c illustrate our second design principle. For each assessment item we produce, we strategically choose numbers to mitigate students' use of memorized facts and engineer opportunities for students to use composite units. First, we chose "non-easy numbers" (4 and 7) for which many students may not have a memorized sequence of multiples that can mitigate the need to keep track of the simultaneous accrual of composite units and 1s that constitute them. Second, we chose to introduce the first 4 rectangles and allow drawing up to 6 more, so the child is still constrained to not drawing each and every unit of 1 (small rectangles) while having a possibility of producing a "mental marker" for the two sets of composite units (4 along the 'width' and 7 along the 'length').

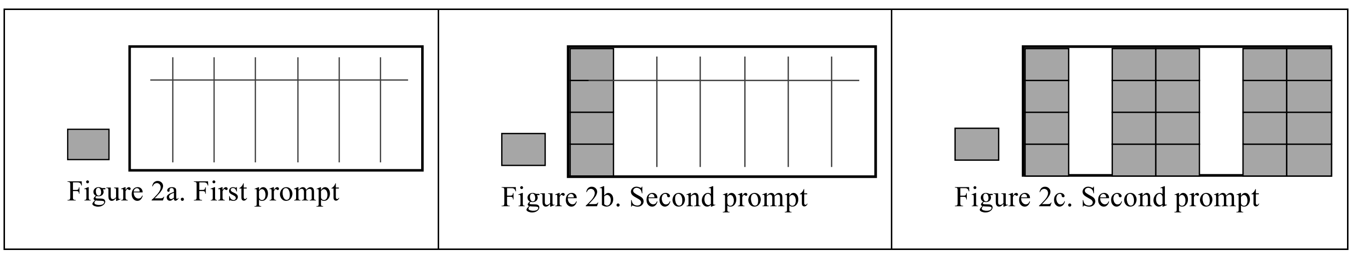

Figure 2a. First prompt

Figure 2b. Second prompt

Figure 2c. Second prompt

Figure 2. Three, diagram-based prompts for simultaneously coordinating 7 units of 4 (1s).

To address the third design principle, we faced a challenge of how to encourage children to express to others, with only a pencil and paper available, what they did in their mind to solve the item. In an interview, the child's bodily motions (e.g., fingers, head, eyes) provide an interviewer with hints as to follow-up probing questions to ask about the reasoning that took place. In a paper-and-pencil format, while trying to eliminate the need for the child to produce written, linguistic explanations, we instead guided the child as openly as possible ("Show how you got your answer") while maintaining the initial constraint ("Do NOT draw small rectangles"). We presumed that a child's drawing of some sort, likely using the figure provided as a basis, could provide a window onto the mental units and operations she or he used. It should be noted that administration of the assessment would include an explicit request to use drawing as a way to help us understand the student's thinking, as if she or he "shows it." We also decided to administer each assessment item, with its gradual-prompting versions, one question at a time. This way enables to invite the child to show how she solved a problem by approaching the person who administers the assessment. That person would document, in an abbreviated form, what the child said (e.g., "Used fingers; counted 4 and 4 = 8; then 8+8 = 16; + 8 = 24; +4 = 28). As soon as the assessment is over, that person would then select and record a code developed for conceptually distinct solutions.

Discussion

Guided by Tzur's (2007) constructivist notion of fine grain assessment, we identified three design principles that can guide the design of a written assessment for distinguishing, without conducting an interview, between a participatory and an anticipatory stage in students' reasoning (Simon et al., 2016). The principles we identified included (a) gradual sequencing of prompt-less to prompt-heavy assessment items, (b) selecting "non easy" numbers, and (c) eliciting authentic student responses. We illustrated how each of these principles was applied to the creation and refinement of items for assessing the stage at which a child might have constructed a particular mathematical idea—the multiplicative double counting (mDC) scheme (Tzur et al., 2013). In this section, we discuss two main contributions the paper can make.

First, we assert that providing guidance for assessing if a student is at the participatory or anticipatory stage in the construction of a new way of reasoning (scheme) seems complementary to the design of instruction so that a classroom zone of proximal development (ZPD) can be promoted (Murata & Fuson, 2006). We root our assertion in Jin & Tzur (2011) and Tzur & Lambert's (2011) postulation of a theoretical linkage between the participatory/ anticipatory stage distinction and Vygotsky's (1986) core notions of ZPD/ZAD (zone of actual development), respectively. Written assessments that allow such a distinction seem key to effectively selecting (a) goals for what each student should learn next based on what they already know and (b) instructional activities that may facilitate such learning. Specifically, instructional activities would attempt to foster two types of reflective processes proposed by Tzur (2011) as suitable for promoting the construction of each stage. Simply put, the design process and principles depicted in this paper explicate important considerations for assessing cognitive correlates of ZPD.

Second, this paper also contributes to the ongoing efforts in the field to figure out ways to both infer into and promote key developmental understandings (KDUs, Simon, 2006) in students. Specifically, our paper focused on the work devoted to creating a stage-sensitive assessment of the first scheme (mDC) a child may construct when embarking on the conceptual leap required to transform one's additive into multiplicative reasoning about whole numbers. As noted in the Conceptual Framework, such a shift in reasoning is considered central to children's mathematical progression (e.g., to fractions and later to algebra). Illustrating how a scalable assessment can help identify students whose initial way of reasoning multiplicatively is still prompt-dependent, and thus prone to being left behind if instruction moves along prematurely, may provide a model for creating such an assessment for other KDUs.

Endnotes

[1]This research was supported by the US National Science Foundation under grant DRL 1503206. The opinions expressed do not necessarily reflect the views of the Foundation.

[2]Testing the tool for validity and reliability would involve a sequence of steps that includes administering the assessment items in a written format and a follow-up interview format to a sample of ~25 students, and figuring out the correlation between each student's work on each item in both format.

[3]Andy Norton suggested this item to us, based on his use of a similar item to assess, without the stage distinction, a fraction scheme.

References

Clark, F. B., & Kamii, C. (1996). Identification of multiplicative thinking in children in grades 1-5. *Journal for Research in Mathematics Education*, *27*(1), 41-51.

Hinkin, T. R., Tracey, B. J., & Enz, C. A. (1997). Scale construction: Developing reliable and valid measurement instruments. *Journal of Hospitality & Tourism Research*, *21*(1), 100-120.

Jin, X., & Tzur, R. (2011). '*Bridging': An assimilation-and ZPD-enhancing practice in Chinese pedagogy*. Paper presented at The 91st Annual Meeting of the National Council of Teachers of Mathematics Research Pre-Session, St. Louis, MO.

Murata, A., & Fuson, K. (2006). Teaching as assisting individual constructive paths within an interdependent class learning zone: Japanese first graders learning to add using 10. *Journal for Research in Mathematics Education*, *37*(5), 421-456.

Norton, A., & Wilkins, J. L. M. (2009). A quantitative analysis of children's splitting operations and fraction schemes. *The Journal of Mathematical Behavior*, *28*(2-3), 150-161. doi:http://dx.doi.org/10.1016/j.jmathb.2009.06.002

Piaget, J. (1985). The equilibration of cognitive structures: The central problem of intellectual development (T. Brown & K. J. Thampy, Trans.). Chicago: The University of Chicago.

Raubenheimer, J. (2004). An item selection procedure to maximize scale reliability and validity. *SA Journal of Industrial Psychology*, *30*(4), 59.64. doi:10.4102/sajip.v30i4.168

Simon, M. A. (2006). Key developmental understandings in mathematics: A direction for investigating and establishing learning goals. *Mathematical Thinking and Learning*, *8*(4), 359-371. doi:10.1207/s15327833mt10804_1

Simon, M. A., Placa, N., & Avitzur, A. (2016). Participatory and anticipatory stages of mathematical concept learning: Further empirical and theoretical development. *Journal for Research in Mathematics Education*, *47*(1), 63-93.

Simon, M. A., Tzur, R., Heinz, K., & Kinzel, M. (2004). Explicating a mechanism for conceptual learning: Elaborating the construct of reflective abstraction. *Journal for Research in Mathematics Education, 35*(5), 305-329.

Steffe, L. P., & Cobb, P. (1998). Multiplicative and divisional schemes. *Focus on Learning Problems in Mathematics*, *20*(1), 45-62.

Steffe, L. P., & von Glasersfeld, E. (1985). Helping children to conceive of number. *Recherches en Didactique des Mathematiques*, *6*(2/3), 269-303.

Tzur, R. (2007). Fine grain assessment of students' mathematical understanding: Participatory and anticipatory stages in learning a new mathematical conception. *Educational Studies in Mathematics*, *66*(3), 273-291. doi:10.1007/s10649-007-9082-4

Tzur, R. (2011). Can dual processing theories of thinking inform conceptual learning in mathematics?. *The Mathematics Enthusiast*, *8*(3), 597-636.

Tzur, R., Johnson, H. L., McClintock, E., Kenney, R. H., Xin, Y. P., Si, L., …Jin, X. (2013). Distinguishing schemes and tasks in children's development of multiplicative reasoning. *PNA*, *7*(3), 85-101.

Tzur, R., & Lambert, M. A. (2011). Intermediate participatory stages as zone of proximal development correlate in constructing counting-on: A plausible conceptual source for children's transitory 'regress' to counting-all. *Journal for Research in Mathematics Education*, *42*(5), 418-450.

Tzur, R., & Simon, M. A. (2004). Distinguishing two stages of mathematics conceptual learning. *International Journal of Science and Mathematics Education*, *2*(2), 287-304. doi:10.1007/s10763-004-7479-4

von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. Washington, D. C.: Falmer.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT.