

Article

Generative Unit Assessment: Authenticity in Mathematics Classroom Assessment Practices

P. Janelle McFeetors ^{1,*}, Richelle Marynowski ² and Alexandra Candler ³

¹ Department of Elementary Education, Faculty of Education, University of Alberta, Edmonton, AB T6G 2G5, Canada

² Faculty of Education, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada; richelle.marynowski@uleth.ca

³ Elk Island Public Schools, Sherwood Park, AB T8B 1N2, Canada; alexandra.candler@eips.ca

* Correspondence: janelle.mcfeetors@ualberta.ca

Abstract: In our pursuit to broaden and deepen understandings of what it means to engage in an assessment activity, we explored the designing and implementing of a geometry performance task as an instantiation of authentic assessment to assess elementary school students' mathematics learning. Using participatory action research, we incorporated a performance task as an end-of-unit assessment with grade 4/5 students. We found that the authenticity within what we are calling a generative unit assessment, is understood as a process that is dynamic in contrast to conventional unit tests. We established an innovative assessment practice that emerged from the student and teacher data and is illustrated through four features applicable to any content area. Through collaborative discussions and the ensuing creation of a generative unit assessment, we found spaces to authentically understand ontological growth and continual learning through assessment.

Keywords: mathematics education; geometry; elementary school; performance assessment; summative assessment; authenticity; ontological growth



Citation: McFeetors, P.J.; Marynowski, R.; Candler, A. Generative Unit Assessment: Authenticity in Mathematics Classroom Assessment Practices. *Educ. Sci.* **2021**, *11*, 366. <https://doi.org/10.3390/educsci11070366>

Academic Editor: Azita Manouchehri

Received: 14 June 2021

Accepted: 16 July 2021

Published: 20 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

For over 20 years, calls to shift classroom assessment practices have been made by researchers and practitioners. Shepard [1] pointed to this shift, identifying that assessment practices that follow a measurement and social efficiency model stand in the way of instructional reform. Delandshere [2] called for a reconceptualization of assessment to align with current theories of learning, moving from a focus on educational measurement traditions to a consideration of the interpretive nature of what constitutes learning and knowing. Assessment can be viewed relationally as a moment of interaction and engagement with others participating together in an activity [2]. Andrade [3] recognized that classroom assessment practices have long been influenced by traditions, such as program evaluation and psychometrics, and that a change to a focus on learning needs to occur. Assessment has been viewed as transactional, where students give while teachers take and then teachers “pay” through giving marks while students take on the labels the marks dictate. This conception of assessment perpetuates an “authoritative relationship between the assessor (the knower) and the assessee (the learner)” [2] (p. 1479), where the assessor is considered an expert and the assessee is a novice.

A shift in clarifying and broadening classroom assessment practices has led to the categorization of assessment practices to support teachers in their intentions to develop a range of assessment strategies. These categorizations have been offered in an attempt to support teachers in being intentional in their assessment decision making and to expand the sources of assessment data beyond paper-and-pencil tests [4]. The different ways of categorizing assessment that have been prevalent in assessment literature are: assessment for/of/as learning, formative/summative, and the type of task assigned [5]. Assessment

for/of/as learning refers to the way assessment tasks are framed (or the purpose) [6]. The labelling of an assessment as formative/summative refers to the temporal aspect of assessment and how the temporality influences how the assessment is carried out [7]. The type of task assigned refers to what students produce as an assessment artefact [8]. What none of these framings does is attend to how a student knows or understands the concepts being assessed. Our conceptualization of assessment encompasses observing and understanding students' learning in classroom contexts, where assessment is designed to provide opportunities for students to learn through engaging in the assessment task.

Through working as a research team, as researchers and a teacher collaborating, we had the opportunity to investigate how an assessment in elementary school geometry could question previous categorizations of assessment. In our pursuit to broaden and deepen understandings of what it means to engage in an assessment activity in a mathematics classroom, we asked: what is the impact of designing and implementing a geometry performance task as an instantiation of authentic assessment to assess elementary school students' mathematics learning?

2. Literature Review

According to the National Council of Teachers of Mathematics (NCTM) [9], assessment "is a process whose primary purpose is to gather data that support the teaching and learning of mathematics" (p. 89). Additionally, NCTM [8] articulates eight effective mathematics teaching practices, one of which is "implement tasks that promote reasoning and problem solving" (p. 10). Performance assessments offer a way to gather assessment data on student reasoning and how they engage in problem solving. This information can be used by the teacher to identify strengths and areas for growth in student thinking [10]. In response, we explored the use of a performance assessment as an end-of-unit assessment, which remains underutilized compared to traditional tests [11].

2.1. Performance Assessment

Performance assessments, or performance-based assessments, have been identified as being an alternative form of assessment that can elicit a deeper understanding of student understanding than traditional tests [12]. The use of tasks as performance assessments pre-dates the use of objectively scored items in educational and non-educational settings by hundreds of years [13]. Performance assessments are used in many different contexts outside of education to ensure one's ability to perform tasks to a certain level of competency (i.e., driving tests, music performance exams). Performance assessments have been used in classrooms in test-like situations or as multiple phase projects completed individually or in groups [8]. Lane [14] describes performance assessments as "ideal tools for formative, interim, and summative assessments that can be naturally embedded in instruction" (p. 313). Additionally, Fuchs et al. [15] states that performance assessments "pose authentic problem-solving dilemmas and require students to develop solutions involving the application of multiple skills and strategies" (p. 611). Characteristics of performance assessments include contextualized tasks, incorporation of higher-level thinking skills, and creation of a product [16]. Having students complete performance tasks in collaboration with others "that is purposeful, used to support accomplishment of productive tasks, and generative of both stronger relationships and more insightful problem solving" [17] (p. 197) is another component that can be incorporated into performance assessments. Performance assessments offer students ways to demonstrate their thinking and their processes for coming to a solution to the posed problem or task in collaboration with others.

Rubrics assessing the quality of the final product are generally used to determine a student's grade on the performance assessment [8]. Lane [14] advocates for the use of performance assessments, stating that "the use of performance assessments is related to deeper student understanding of content" (p. 321). The design, implementation, and scoring of performance assessments for classroom assessment use is heavily influenced by psychometric and objectivist theory to ensure that the results of the performance assessment

can be considered reliable and valid [18]. Lane [14] calls for further research on the design and implementation of performance assessments, “and their impact on advancing student learning and informing instruction” (p. 326). Factors to be considered when developing a rubric for a performance assessment include the type of rubric (analytic or holistic), what specifically is being assessed (process, performance, or outcome), and if the scoring criteria are based on qualitative or quantitative measure [19].

Additionally, performance assessments have been equated to authentic assessments in that the tasks that students complete are more in line with performing a set of skills in a relevant context. Frey [13] problematizes the notion that a performance assessment is automatically authentic by countering, “one could easily design a performance-based assessment that is not authentic” (p. 201). Zane [18] offers that performance assessment is a way to determine how well a student “can handle the messy (ill-structured) real world” (p. 83), acknowledging that engaging in performance assessment is not necessarily about the product, rather the engagement within the process. As such, rather than grounding our investigation in principles of performance assessment, we choose to describe assessment that engages students in the process of thinking and creating by using the notion of authentic assessment instead.

2.2. Authentic Assessment

Gulikers et al. [20] described factors that one could use to determine if an assessment could be labelled as authentic. They identified “that in order to meet the goals of education, a constructive alignment between instruction, learning and assessment (ILA) is necessary” (p. 67). They continue saying that “learning and assessment are two sides of the same coin, and that they strongly influence each other” (p. 68). In order for assessments to be considered authentic, they need to be designed so that there is a similarity between the learning task and the assessment task. If those two are not aligned, there is a lack of authenticity in the assessment. Additionally, Gulikers et al. [20] articulated that “giving students ownership of the task and the process to develop a solution is crucial for engaging students in authentic learning and problem solving” (p. 71). In order for a task to be authentic, they theorized that the task itself, the physical context, the social context, the assessment result or form, and the criteria and standards must be perceived as being “representative, relevant, and meaningful” [20] (p. 71) by the users of the task. The fidelity of the task to the context of learning is what makes an assessment task authentic. A performance task is not necessarily authentic if it does not align with the learning processes of the course.

Vu and Dall’Alba [21] take a different stance on authenticity of assessment. They stated, “Authentic assessment is not an end in itself; rather, it is an opportunity for students to learn to become who they endeavour to be” (p. 779)—ontological growth. Rather than focusing on the characteristics of a task that could or could not be perceived as being authentic, Vu and Dall’Alba [21] focus on the ontological underpinnings of authenticity and draw on Heidegger [22] to describe authenticity as one’s being in the world and how we draw on our understanding of our involvement in the world. In relation to assessment, they articulated that “planning an assessment task requires understanding of who we are as teachers and who are our students, as well as what type of learning is to be assessed and how that learning can be demonstrated through the assessment” (p. 784). Vu and Dall’Alba [21] continue to say that authentic assessments allow students to fully engage in the task and to explore who they are as human beings in the world; “Authenticity is, then, not an attribute of tasks but, rather, a quality of educational processes that engage students in becoming more fully human” (p. 787). The conceptualization of authentic assessment as providing students opportunities to explore their own ways of knowing and being provides a way to think about assessment that is more about the students engaging in the task rather than the curricular goals, objectives of the task or the grading criteria. Vu and Dall’Alba [21] also stated, similarly to Gulikers et al. [20], that the physical context, social context, criteria, and form are important components of an authentic assessment. However, Vu and Dall’Alba [21] focus on the interaction of the student with the task, of

students with each other, and with the teacher, rather than the specific construction of the task. Here, we respond to their call for “the need to explore and evaluate the use of authentic assessment practices” (p. 789) to contribute to a theoretical understanding and practical implementation of assessment practices for ontological growth.

3. Materials and Methods

We used participatory action research [23–26] as a methodological framing, grounded in a nondualistic ontology. In participatory action research, understanding arises out of action through democratic practices of collaboration and inquiry as a social construction of reality toward transformation that becomes the basis for developing and sharing new knowledge as it is reified through use in the research context. Participatory action research emphasizes implementing a novel approach to improve learning, situating the inquiry within a specific context, inquiring in a systematic manner, which includes cycles of refining and reflection, initiated by teachers and contributing a finely nuanced understanding of the phenomenon to impact professional, personal, and political realms. The aim, methodologically, is transformative practice of teaching and learning by cyclically clarifying the meaning of (inter)actions. Understandings developed through collaborative, interpretive actions within the research process are shared as a coherent narrative in which the words of participants are amplified as evidence for the generation of new knowledge.

3.1. Participants and Setting

The research reported took place at Maple Springs Elementary School. Consistent with the cyclical nature of action research, Maple Springs was the location of the third macro-cycle in a larger research project investigating students’ mathematical learning through commercial game play. The first macro-cycle was exploratory, and students’ unstructured play enabled growth in mathematical processes like communicating, visualizing, problem solving, and reasoning. The first macro-cycle established the effective use of commercial games for learning mathematics and the need for more frequent student reflections [27]. Therefore, the second macro-cycle included students’ recording mathematical thinking on activity sheets every session, crafting prompts teachers posed to students during game play, and focused solely on students’ development of logical reasoning [28]. Students’ emerging understanding of geometric ideas through spatial reasoning in the second macro-cycle sparked a new avenue of investigation for the next macro-cycle.

The third macro-cycle of the research at Maple Springs occurred during the latter half of the school year. The classroom teacher and co-author, Alexandra, invited the researchers into her classroom after hearing about students’ improved learning in the first two macro-cycles of the project. Within the topic of geometry, she identified a desire to transform her mathematics teaching practices and her students’ learning experiences by shifting toward students engaging in mathematical thinking [29] and participating in cognitively demanding tasks [30,31]. The novel approach was implemented within Alexandra’s grade 4 and 5 combined class, where all students participated in all aspects of the weekly sessions and the assessment intervention. Additionally, 25 students consented to data collection processes for the research project. Aligned with participatory action research, we investigated the learning occurring for the teacher and each student as idiosyncratic to the being and becoming of each individual, rather than the collective.

The ensuing assessment intervention was situated within the broader instructional setting, where Alexandra selected two commercial games, *Equilibrio* and *Quartex*, to enhance geometry learning. The students played *Equilibrio* and *Quartex* in a one-hour mathematics class, weekly, over four months. Each week constituted a micro-cycle, as data collected was reflectively considered among the research team (teacher and researchers) in planning for learning for the following week. Our visits began before Alexandra’s geometry unit commenced and continued throughout its entirety. We introduced *Equilibrio* and *Quartex* by watching short videos of an overview to the game rules and materials. Students then alternated weekly between playing *Equilibrio* and *Quartex* and moved

from exploration toward strategic thinking [32] by using the game pieces as tools to model conjectures of good piece placement to justify their play. As students played, they were encouraged to communicate their mathematical thinking as Alexandra and the researchers posed verbal prompts and as students recorded through words and drawings their ideas on activity sheets. From one micro-cycle to the next, we refined instructional materials. Notably, the questions posed on student activity sheets were constantly refined to incorporate strategies or constructed game boards from the previous week that generated mathematical thinking, to increase cognitive demand by providing a game variation as an extension, and to integrate analysis of pieces and game boards based on new mathematical topics introduced between micro-cycles. We also adapted the verbal prompts to elicit deeper mathematical thinking in subsequent micro-cycles. Alexandra also conducted approximately three weeks of formal teaching for the geometry unit where she introduced 90° angles, shape attributes and properties, symmetry, congruency, and transformation of shapes.

3.2. Assessment Intervention

In this article, we focus on the activity of the final micro-cycle for the participatory action research at Maple Springs. The final micro-cycle was a culmination of the macro-cycle that was marked as an end-of-unit assessment just as Alexandra and her class concluded their formal geometry unit. As Alexandra reflected on her students' mathematical thinking while engaged in game play, she recognized that the depth of mathematical thinking demonstrated and expressed was deeper than what she regularly saw on paper-and-pencil evaluations, and students' persistence in seeing themselves as capable of mathematical activity was sustained. She requested that we work collaboratively to create an end-of-unit assessment in the context of game play.

We determined that Equilibrio provided an opportunity to assess a wide variety of geometry learning outcomes and to elicit more creativity from the students. Equilibrio includes a blueprint book depicting towers in two-dimensions with increasing difficulty of puzzles. The goal of the game is to solve the puzzle by constructing a corresponding three-dimensional vertical tower using a finite set of blocks. See Figure 1 for an image of students playing Equilibrio and the website www.learnmathwithgames.com (accessed on 19 July 2021) for further game description. Although it is described as a one player game, the students insisted that having two players was necessary for tricky balancing and increased enjoyment. As well, Alexandra valued the mathematical discourse [33,34] among students as they were problem solving to construct a tower.



Figure 1. Students playing Equilibrio by building a tower from a blueprint.

Beginning with the familiar context of Equilibrio game play, we collaboratively shaped the performance assessment and accompanying rubric. The research team progressed through many revisions, characterized as a “messy turn” in action research that leads

to practical and theoretical insights [35], as we had conversations about the aims of the assessment and how to elicit students' geometric thinking. See Appendix A for the complete performance task with rubric. The performance assessment invited students to create a new puzzle for the blueprint book to showcase the geometric thinking and mathematical processes they had developed over the unit. The framing of the task posed high cognitive demand [31] as it invited students to break from habituated ways of engaging in the game of moving from a 2D representation to a 3D tower, to enact creativity in first designing a 3D tower and then representing it in a 2D puzzle. In using a novel approach with familiar materials, our intention was to observe students' problem solving within geometry. Consistent with design requirements for performance assessments as presented above, this task was contextualized within a familiar game of Equilibrio, incorporated higher level thinking skills in its cognitive demand, and asked students to create a new product in the form of a new puzzle.

Some of the Equilibrio performance assessment was collaborative. Part I included the design phase, where student pairs were to record their brainstorming of the features they wanted to include in their tower. Next, the students were directed to use the Equilibrio blocks to design a 3D tower and take a photograph on an iPad, as towers were prone to collapsing and they would be able to reference the photograph as they proceeded. In Part II, students were to use their 3D design to create a new Equilibrio puzzle in three stages: (1) drawing an initial 2D representational sketch, (2) conferencing with another pair to improve their puzzle, and (3) finalizing the puzzle through a share-quality sketch. Part III introduced the mathematical defense of the puzzle creation. Because students' verbal expressions often contained more sophisticated understanding than what students produced on paper throughout the research, we encouraged students to create a screencast with the app "Explain Everything", annotating and recording a voice-over explanation of the mathematical features of their tower and puzzle. Consistent with design requirements for performance assessments as presented above, this task invited students to collaborate with other classmates in the puzzle creation and provided many ways to demonstrate their mathematical thinking and processes such as reasoning, visualizing, problem solving, and communicating as they arrived at the solution of a new puzzle.

In order to observe individual students' mathematical thinking, Part IV of the performance assessment was completed individually. We invited students to select the "four best math ideas" among the mathematical ideas they had been discussing with their partner in the collaborative component. To assess students' complete performance assessment, the authors collaboratively created a four-point rubric that focused attention on critical features of geometric thinking and mathematical processes. Alexandra also provided students the curricular learning outcomes framed as "I can . . ." statements, along with a word bank.

A week before the assessment intervention took place, students were introduced to the performance task context of creating a new puzzle for the Equilibrio blueprint book. The day before the performance assessment, Alexandra also introduced the rubric to the students to clarify expectations. Returning the following week, the task was implemented over two class periods (approximately 1.5 h). Students were provided with a blank copy of the performance task, an Equilibrio game, and an iPad. Students were introduced to the application "Explain Everything" and given instructions on how to proceed through the performance task. During the first period, the students worked in self-selected pairs on Parts I to III. During the second period, each student received a copy of the picture of their tower and then completed Part IV.

3.3. Data

Multiple forms of qualitative data were used to capture students' engagement in the performance assessment, including field notes, student artefacts, and interviews. We wrote field notes during and after the classes to record observations of students' thinking and interactions while completing the performance assessment. The artefacts produced by the students throughout the assessment intervention were simultaneously sources of data.

Students' written records of the performance assessment were collected and digitized for storage and analysis. Students' screencasts were downloaded and stored, as we found these to be critical forms of data to augment written records.

Several weeks after the performance task, we conducted interviews with all of the students who participated in the study and with Alexandra. The 20 min student interviews, conducted in pairs, encouraged students to look back retrospectively at the value of learning mathematics with games and inquired into the novel evaluation process of the performance assessment. A one-hour interview with Alexandra invited her to reflect on all the micro-cycles within the project, including the performance assessment. All interviews were audio recorded and transcribed. The interviews enriched the data analysis as multiple perspectives on the performance assessment were included to gain insight into the impact of shifting from traditional to novel assessment practices.

3.4. Data Analysis

We inductively analyzed the transcribed data and students' performance assessments using thematic analysis [36], a flexible approach that enabled us to explore openly through multiple engagements with the data to recognize, analyze, and explain patterns within the data. Following the phases suggested for conducting thematic analysis, our analytic aim was to interpret the participants' experiences to develop explanatory themes that is a "*patterned response or meaning . . . [that] captures something important*" [36] (p. 82, emphasis in original) related to the impact of an authentic assessment for all the participants. We became familiar with the data through the immersive process of reading the data corpus from the macro-cycle and analyzed it within the broader project for students' geometry learning through game play. Through this initial analysis, the importance of the novel assessment at the end of the unit for all the participants became apparent, and so we progressed by creating a data set of all the data related only to the performance assessment and reread the refined data set. In discussing our initial observations and prototypical exemplars [37], we noticed shifts in how Alexandra and her students viewed themselves and shifts we saw in their capabilities through their engagement.

With our research question foregrounded, we proceeded by coding moments of shifts by indicating the focus of the participant's shift through labelling with short phrases such as "problem solving", "perseverance" or "understand symmetry", and memo-writing brief interpretive comments to explain the meaning of the code [38]. We individually coded the interview data first as it was a retrospective look at the performance assessment. Because the prompts were quite open, we found that Alexandra and the students' responses alerted us to what they viewed as educative in their experiences. After, we discussed emerging codes with specific examples and found ideas common across the coding that began to hint at emerging themes to consider. For instance, the theme of "becoming mathematical through mathematical processes" was formed when we noticed that students invoked many processes such as problem solving, representing, and communicating while completing the task. Conversely, conceptual understanding and productive disposition are both strands of mathematical proficiency, yet we chose to keep them as two separate themes as we did not see evidence of growth in the other three strands to warrant "developing mathematical proficiency" as a theme. We then proceeded individually again by coding the students' artefacts, occasionally returning to apply a refined code to the interview data, moving back and forth among data moments as a "more recursive process" [36] (p. 86).

In discussing the codes for the data set and considering our collective experiences in mathematics education scholarship, we noted four themes that organized the codes to highlight areas of growth as the significant impact of a novel authentic assessment. Alexandra's intimate understanding of her students and the classroom context was instrumental in developing the themes. After the themes were confirmed, all the data was reread as it was sorted into an electronic file for each theme to verify membership of codes and data excerpts within each theme, and to assist in reporting the themes. It is worth noting that the only set of codes that did not fit were specific reflections on the mechanisms of the

performance assessment that were effective, which fall outside the scope of this article. The themes and their unifying concept are described in the following results section.

4. Results

Arriving at a unifying concept required us to return to problematize two aspects of assessment: the labels that are used to point to particular kinds of assessment tasks and the rhythm of assessment driven by discourse that sequences formative and summative evaluations. In looking for a type of task that would embody authenticity in assessment [21], we found that a performance assessment [8,13,14] was the closest fit. It enabled a process-oriented task where students were actively demonstrating their mathematical thinking and was multifaceted, so Alexandra could gain multiple perspectives into students' mathematical thinking.

However, Alexandra explained in her interview that her students "learned a lot from [the performance assessment] as opposed to something you're performing," leading us to question the metaphor of performance that invokes a highly polished product to many students who participate in extracurricular activities such as dance or music, where the pinnacle accomplishment is a perfected display of competency. Additionally, etymologically perform means "to form thoroughly" [39] (p. 485) or "complete, finish" [40] (p. 346). This is contrary to the nature of students' mathematical understanding and skills that are never thoroughly formed or finished but are undergoing constant (re)forming as they enrich their relational understanding. Performing also positions students as performers, etymologically derived from "one who makes a public exhibition of his skill" [41] (p. 870), rather than as learners who are valiantly trying to make external to themselves and their teacher what has developed, both internally and ephemerally, in interaction with others and the environment.

We suggest that *generative* is a stronger interpretation of the participants' experiences. This descriptor encapsulates the continued learning and growth occurring in each person as they engaged in the process of assessment—the propelling forward of what the assessment continued from the rest of the unit. Etymologically, generate means "to bring into life; to originate" [41] (p. 200), and in this instance, not only was what was learned by students revealed, but it enlivened the learning of the teacher and students. While the task was generative, the teacher and students were agentic, in the etymological sense of "having the power of generating or producing" [41] (p. 200) as they were making themselves, becoming.

Similarly, many teachers would categorize the Equilibrio performance task as summative because of its placement at the end of the unit and assessment intentions to evaluate the extent to which students learned and report publicly [12,13,42,43]. The participants' experiences showed that learning was ongoing throughout the performance task, bringing into question the finality of students' geometry learning. Students were accustomed to summative unit tests, not only in mathematics class but throughout all of their courses. The summative unit assessments were weighted heavily toward their overall grade and often led to student anxiety. While we advocate for grade creation throughout a unit and a variety of sources, it remains important to provide students an opportunity to draw together their learning at the end of a unit as a unifying whole, while still leaving the possibility of future related learning that occurs during the school year. We suggest that *unit assessment* is a more representative term. And so, we offer an original and innovative assessment approach: *generative unit assessment*.

Using a performance assessment was an effective test replacement because it was a generative unit assessment sponsoring teacher and student growth. Dewey [44] saw the aim of education as growth, and that growth occurs through engagement in educative experiences that "modifies the one who acts and undergoes" (p. 35) in the process of becoming. In responding to the research question, the authenticity of the generative unit assessment (GUA) is understood as a process that is dynamic in its impact on those involved, in contrast to conventional unit tests being seen as a static event. Four themes

that emerged from the student and teacher data illustrated four features of generative unit assessment:

1. Teachers deepening their awareness of students' mathematical thinking;
2. Students consolidating their understanding;
3. Students becoming mathematical through mathematical processes;
4. Students' emerging productive disposition.

We will elaborate each of the themes below by explaining participants' growth as demonstrating the impact of authentic assessment, illustrating each with teacher and student data. Specific examples from the data were selected among a multitude of examples sorted under each of the themes to incorporate a broad range of students, to illustrate the core ideas related to the component parts of each theme through individual instances, to develop a coherent narrative of growth through the GUA, and to showcase the variety of forms of data incorporating multiple perspectives.

4.1. Teachers Deepening Their Awareness of Students' Mathematical Thinking

The teacher's awareness of students' mathematical thinking deepened as she observed students' mathematics-in-action and examined recorded products. Talanquer et al. [45] stated, "What teachers notice or listen to enables and constrains their inferences about student knowledge, thinking, and understanding" (p. 587). The GUA provided opportunities for Alexandra to become aware of instances of students demonstrating mathematical thinking. Talanquer et al. [45] explained that observations of student's mathematics-in-action through engagement in an assessment task can allow for a "productive interpretation" of student thinking and "development of expertise in this area are marked by a shift from building general evaluations of student learning to creating specific accounts of student understanding" (p. 588). Through her deepening awareness of student mathematical thinking, Alexandra was able to identify understandings that specific students were developing. She commented that having the students engage in the GUA "helped me understand the kids more" and "when I did my report cards, it was really easy to write about the kids because I could understand their thinking." Not only was she becoming more aware of student thinking, she saw herself as a more effective communicator in making visible her assessment understanding to others.

Alexandra became more aware of how the students understood the mathematical concepts rather than simply what they knew. She stated:

So when you just memorize it, I think they don't understand it and they weren't having to apply what that meant. So I think when they had to use the word symmetry with each other and start applying that into the tower, it helped them understand it better.

This comment illustrates that the teacher was becoming aware of the difference between students developing an instrumental and a relational understanding [46] of symmetry while engaging in the GUA. Instrumental understanding is when students understand how to answer a question using a rule their teacher told them, for instance, correctly identifying a line of symmetry on a 2D shape on a test. Relational understanding encompasses interconnected associations that enables students to express why mathematical properties and procedures exist. In the case of symmetry, students can explain why shape is symmetrical because of invariance, can justify that symmetry can be used in geometric analysis, and can apply the symmetry in novel contexts. Through the GUA, Alexandra was able to observe students' explanations and application of symmetry that demonstrated they had a relational understanding of symmetry and will be illustrated by students' thinking in the next section.

Through observation of students working and engaging in conversation with them during the generative unit assessment, Alexandra commented:

Lots of students really shocked me at how much they knew and helped me understand their thinking. I did more pencil-and-paper work before this and

they weren't showing me the same level of thinking as they did when we did the performance assessment.

Alexandra was becoming aware that the work that she had students previously complete was not providing a rich picture of student mathematical thinking.

As she reflected on how she had been choosing to engage students in assessment tasks previously, she noted that the focus of assessment was not necessarily on where it should be to support students in developing mathematical processes. She stated:

I think these processes are something I also thought a lot about because the processes are what's going to help them as they become older and continue on their grades and become lifelong learners . . . Can they communicate and can they reason, is way higher level skills than can they memorize what the word symmetry means?

Kilpatrick et al. [47] commented that teachers "learn by listening to their students and by analyzing their teaching practices" (p. 384). Alexandra showed us that she grew in her thinking about herself as a teacher and her role in the classroom. She noted, "I have totally changed my approach to teaching math . . . to focus on thinking." This evolution in focusing on students' mathematical thinking-in-action in Alexandra's teaching came about because she noticed it was possible during the GUA. Her excitement for the opportunity the GUA offered to elicit students' mathematical thinking-in-action and making it external and apparent to others caused her to design future lessons to incorporate these opportunities. Key characteristics she used in future lessons were open questions and the use of problems that had multiple solution pathways, along with small group and class discussions where students could share their mathematical thinking-in-action. In this case, not only were the students learning but the teacher was, too. "Teachers whose learning becomes generative see themselves as lifelong learners who can learn from studying curriculum materials and from analyzing their practice and their interactions with students" [47] (p. 384). The implementation of the GUA by Alexandra in her class provided her the opportunity to deepen her awareness of student mathematical thinking and an opportunity for her to grow as a teacher, expanding her identity as a teacher who noticed student learning.

4.2. Students Consolidating Their Understanding

As students were engaging in the GUA, they were consolidating their understanding of mathematical concepts within the unit. Relational understanding is consolidated through multiple experiences over time by increasing and strengthening connections among mathematical ideas and by making connections to experiences and ideas outside of mathematics. Further, teachers can assess developing understanding by collecting evidence "located in the discourse, actions, and transactions of individuals in participation" [2] (p. 1478). As the students were engaged in the GUA, they were increasing and strengthening relationships with the mathematical ideas present in the task and consequently creating evidence of consolidating their understanding. Among many of the concepts that students invoked, we use symmetry and balance as illustrative examples.

During the GUA, students were consolidating their understanding of symmetry. Alexandra had introduced the idea of symmetry to students during the formal geometry unit, and they could identify and draw lines of symmetry on individual 2D shapes. Unlike a unit test where students are directed to identify a line of symmetry to demonstrate procedural fluency [47], the mathematical analysis within the GUA was open-ended for students to choose which mathematical ideas to apply from the geometry unit. This makes students' intentional choice to rely on symmetry significant as they had to notice the connection between symmetry and a shape, rather than being told by their teacher. We observed emerging understanding when Arjun (all students are referred to with pseudonyms; we have maintained original spelling in all excerpts of student work.) wrote, "It is tricky because you need to mack a line of simetry and count the lines of semetry all in one. The line of semetry is something that divides or macks something split" (see Figure 2 for

an accompanying drawing). Arjun moved beyond following a prescribed procedure to convince Alexandra of its importance. His justification included writing a definition of a line of symmetry in his own words, which indicates his growing understand of symmetry and how it applies to a legitimate context.

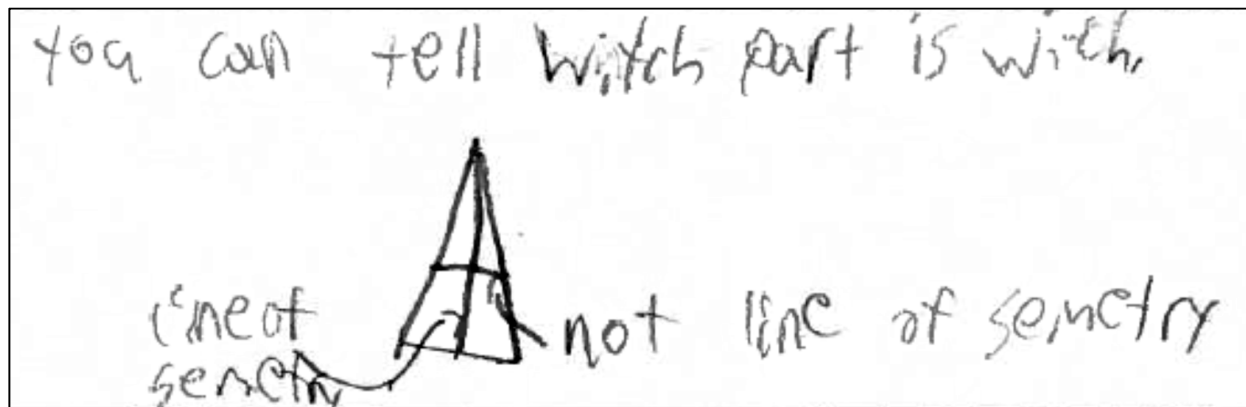


Figure 2. Arjun's pictorial representation of symmetry for a 2D shape.

Students were also applying the idea of symmetry through the GUA to their whole tower as a 3D object, consolidating understanding by connecting to mathematically more complex objects compared to the unit content. Kamri articulated her understanding of symmetry and how it applied to her tower:

It's how you gotta think of symmetry, is a line down it would have to be the same on both sides. If you were to fold it in half, would it be the same? Like, if I was to fold it in half this way, I don't know if it would be the same, but like if I was to fold it in half this way, it would be the same and if I was to fold it in half this way it would also be the same.

Kamri used the language of folding which she learned in the context of a 2D shape, and applied that concept to her 3D tower. She could visualize the lines of symmetry for her tower, which shows a consolidation of understanding of the concept of symmetry beyond what she was introduced to in class. Riya described her pair's tower saying, "Our tower is not symmetrical and it has different things on each side." This shows a recognition that symmetry means that *things* are the same on each side. Bailey and Dillon demonstrated consolidating their understanding by being able to identify how to adjust one of the pieces on their tower to create symmetry:

Dillon: Symmetry, except for the cube—top part right here.

Bailey: Are not symmetry. . . . The circled part is not symmetry. If you flip this piece, yes it would be symmetry.

These students showed that they used the GUA as an opportunity to consolidate their understanding of symmetry by explaining examples, non-examples, and how to modify their design to create symmetry. Hazel reflected on consolidating understanding through the GUA: "help[ed] us with symmetry because I bet that we're going to be doing harder symmetry next year . . . and these would help us because they—you have to think about what you're going to do and why you're going to do it." Hazel recognized learning through the GUA would contribute to her future learning when mathematical ideas would become more complex.

Students also made connections across mathematical ideas. For example, Ronin related symmetry to a reflection where, "One thing that I noticed, is like line of reflection, is same as the line of symmetry. 'Cause if you just cut the shape in half it'd be the exact same so it would be a reflection and symmetrical." Reflections were a concept that were not specifically addressed with 3D objects nor was the idea of reflection being similar to symmetry. Ronin was consolidating his understanding through making connections to his

other mathematical knowledge. This consolidation strengthens his understanding of both symmetry and reflections.

Students emphasized the notion of balance during the construction of their towers. In their final written component of the GUA, students were asked to describe ways in which their tower illustrated mathematical ideas. Of the 22 students in the class, 18 of them mentioned how pieces balance together as an element of the mathematics that they incorporated into their tower. Some examples include:

Aditi: Making the cube stand on the rectangle because of the balance point.

Nyla: You have to get the wate lavled out so you can get it to balens.

Kamri: Use a line of symmetry so it has balance. (See Figure 3a.)

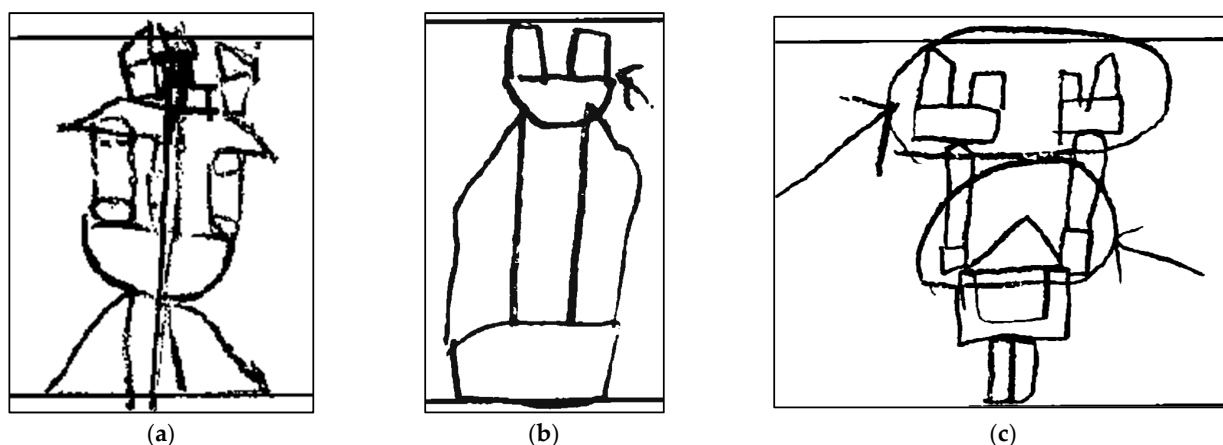


Figure 3. Students' tower drawings include: (a) Kamri's drawing, (b) Brody's drawing, and (c) Wyatt's drawing.

Sage: What makes it triky is where the half moon is right on top of the triagles is it is all balanced on that.

Brody: It's tricky to balace the cemy circle on the two ramps because it's not as stable. (See Figure 3b.)

Wyatt: Theres many things balancing on one piece. (See Figure 3c.)

Bailey: I think my puzzle is tricky because you have to balance most of the tower on the two triangles.

In Science, these students had been introduced to balance in Grade 2 and to building structures in Grade 3 without explicit connection between the grades or to mathematics class. The comments that students made about balance show that they were independently integrating their learning across two content areas, through the experiences of building and constructing their towers into the mathematical understandings they are developing. Students were also incorporating characteristics of the 3D objects into their descriptions, identifying what specifically makes balancing more or less tricky. The GUA offered students opportunities to experience balance and the characteristics of shapes that allowed them to build connections between those developing understandings.

4.3. Students Becoming Mathematical through Enacting Mathematical Procces

During the GUA, students were becoming mathematical through enacting mathematical processes. Being mathematical means "to enter and engage in the discipline of mathematics" [48] (p. 48), into ways of thinking and acting that are characteristic to mathematics. NCTM [49] emphasized five primary mathematical processes for being mathematical: problem solving, reasoning, communication, connections, and representations. Alexandra explained that the assessment allowed her to observe "how kids think about math and use it as a skill to problem solve, to make connections between ideas, to communicate and reason." We present two cases of pairs of students to demonstrate that the students were in the process of being and becoming mathematical as they chose to invoke mathematical processes to make sense of the situation and act mathematically.

In the week leading up to the GUA, Kamri and Sage arrived at school each morning with 2D sketches of possible tower designs that were created without the 3D blocks available. However, it was fascinating to observe the difficulty they had building the 3D tower from their initial plan during the GUA. They later explained:

Kamri: No, we had the half circle with the rect—with the cylinders on top of it. Trying to balance there I think . . . it just wasn't working for us, it would—it just kept falling over. And both of us tried, I tried and it fell over, she tried and it fell over, and so . . .

Sage: So we switched it up, and I built a tower on a piece of paper. And then she built a tower, and then with our math thinking, we took another piece of paper and transformed our towers together, and this is what it looked like . . . we thought that we could make a bit of changes. And, it's pretty similar but it's also different.

They engaged in problem solving as they identified the design element that was problematic—cylinders on top of the semi-circular solid (half circle)—and proceeded to explain the process they used to resolve the design issue. Figure 4a shows their final tower design where the semi-circular solid is located near the bottom of the tower with the bridge stacked on top, rather than two cylinders. On the screencast, Sage clearly communicated the procedure to build the bottom part of the tower:

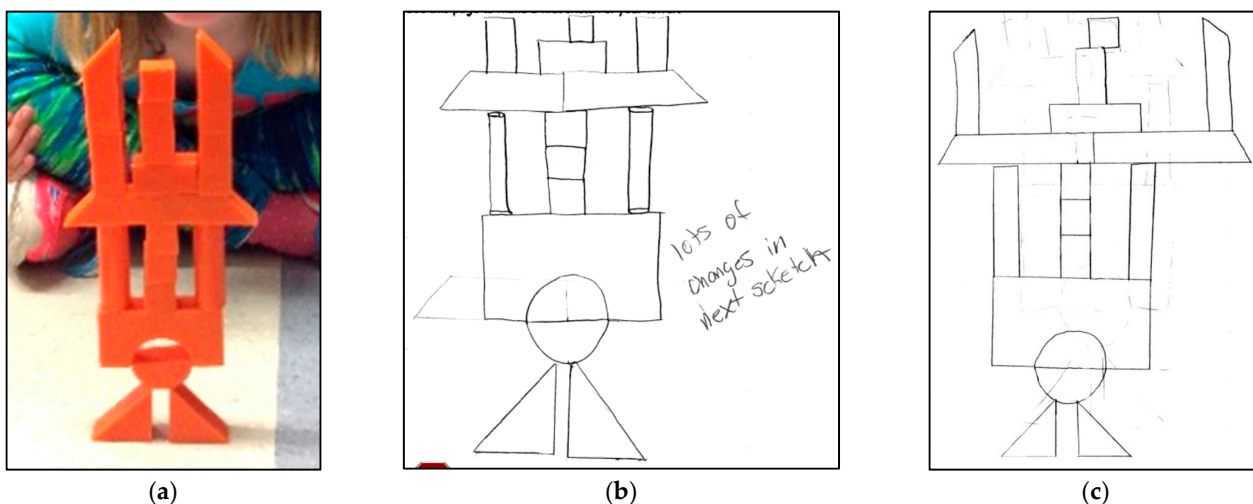


Figure 4. Kamri and Sage's tower representations: (a) tower photograph, (b) draft drawing, and, (c) share-quality drawing.

First, we will take two triangles, we will place them side by side. Then we will take the half moon and place it on top. Then we will take the bridge and place it on top of the half moon. On top of that we will take two cylinders and place them on each side of the bridge.

We observed communication of emergent spatial reasoning as Sage attempted to explain the position of each solid in space relative to the previous piece.

Kamri and Sage also modified the top of their tower, depicted in the differences between the draft drawing (Figure 4b) and share-quality drawing (Figure 4c), allowing them to incorporate all the 3D pieces. In defending their new design at the top, they also used reasoning as Sage justified, "But now it doesn't have a line of symmetry because of this block" in reference to the top cube. Applying the mathematical idea of symmetry to a game demonstrates making connections to a relevant context. The difference between the two drawings also demonstrates improvement in representations. In order to assess students' translation between 3D (tower) and 2D (drawing) representations, we recognized the significant challenge in drawing a puzzle for the first time as part of an assessment. While some details could be improved, we see progress in representing a cylinder in two-dimensions where the top and base have been eliminated in the share-quality drawing. In each of these moments, we see Kamri and Sage becoming mathematical as they relied on ways of acting mathematically to solve the problem presented in the GUA.

Harlow and Brody were confident from the beginning in their tower design, perhaps because their approach differed by first building with the 3D blocks provided. As a result, they focused on communicating their ideas, inclusive of reasoning, and representing their tower by connecting words and images. Early in the unit, the students had realized the importance of naming 3D solids to build towers collaboratively. An exchange was recorded in their screencast:

Harlow: Where is that round thingy?

Brody: What round thing?

Harlow: That, I mean, what is it called again?

Brody: Cylinder.

Harlow went on to confidently communicate by applying the geometric name when he wrote, “The cylinder that’s balancing on the triangle may slide right into the skinny part of the tower” (see Figure 5a for constructed tower). When students learn mathematical vocabulary, such as cylinder to name an object, in an authentic context, they construct meaning of the term that they can apply flexibly when the need arises [50,51]. Harlow and Brody moved beyond description in their communication to justifying mathematical thinking with sophisticated forms of reasoning. For the tower difficulty they wrote, “If you put the long [cylinder] on the triangle, it might slide down.” Their use of the if-then structure signals emergent deductive reasoning. Early and multiple opportunities to construct mathematical arguments supports students’ later engagement in creating proofs, an essential element of advanced levels of being mathematical [52].

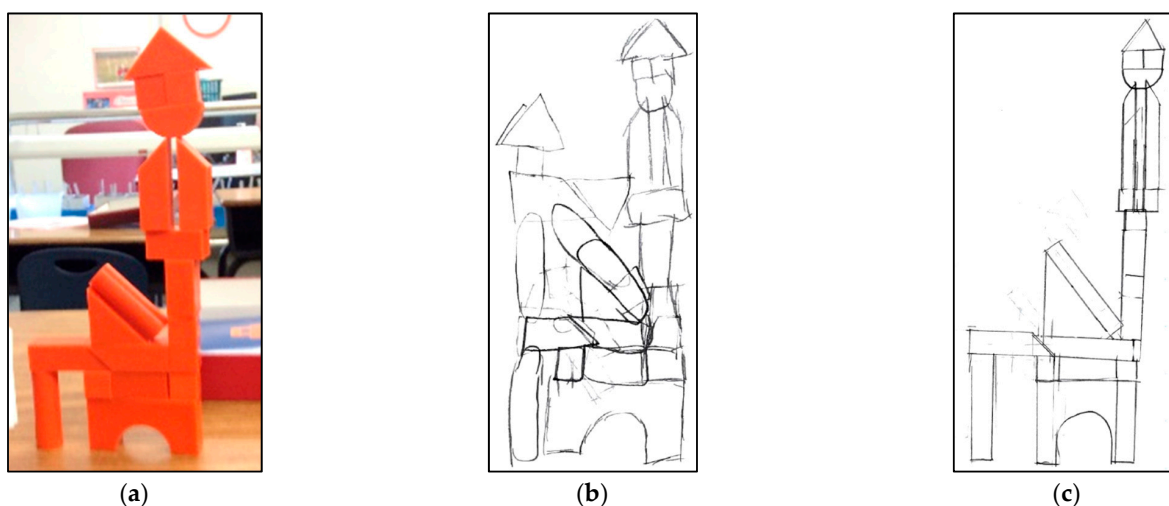


Figure 5. Harlow and Brody’s tower representations: (a) tower photograph, (b) draft drawing, and (c) share-quality drawing.

Similar to Kamri and Sage, Harlow and Brody exhibited becoming mathematical in their improvement in 2D representations of their tower between the two drawings (see Figure 5b,c). The second drawing shows improved proportions among the pieces, as well as reduced hesitations in figuring out the representational aspect. Harlow and Brody reflected on how they improved their representation of the cylinder in the 2D drawings:

Brody: We drew them curved first because they’re cylinders. But in the book we looked, and the cylinders look like rectangles . . .

Harlow: Because you can’t really draw them 3D. It’s hard. Because you’re trying to draw the picture in 2D and it’d be hard to make it look like a cylinder, so then just make it look like a rectangle.

The students took on a different perspective in creating a new pictorial representation. Additionally, we saw earlier how Brody adeptly connected verbal and pictorial representations to communicate the element of balance. When students are invited to act mathematically through assessment tasks, it results in their participation in the disciplinary

practices—here, ways of thinking and acting using mathematical processes—and each moment contributes to their growth of being mathematical.

4.4. Students' Emerging Productive Disposition

Students' growth through the GUA was particularly evident in their emerging productive disposition. Kilpatrick et al. [47] define productive disposition through students' beliefs, where they see mathematics as a subject that can be meaningful and interesting, sustained engagement contributes to learning mathematics, and they are capable of learning and doing mathematics. Thus, productive disposition extends far beyond a positive attitude toward (re)forming their identities as successful mathematics learners, so much so that "students' disposition toward mathematics is a major factor in determining their educational successes" [47] (p. 131). While many of the students in the study echoed Fiona's sentiment, "When I think of math, I don't really like math," we began to see the emergence of characteristics of productive disposition during the GUA. As Alexandra explained, in her class generally, "There's an 'I can't do math' feeling with a lot of kids, or 'math is too hard'." Yet, she noticed a shift in that the students "were really excited to do math and that level of anxiety was removed" and, in particular, through the GUA "it really felt like they wanted to show their learning versus just get [a mark] at the end." In fact, Kamri's highlight of the unit was expressed as, "I like how we did our test this time!" We believe that this unit assessment was generative because it played a role in students' emerging beliefs.

The GUA occasioned the opportunity for students to perceive mathematics as sensible and interesting. Mathematical ideas are sensible when students can make connections to their personal experiences, such as Bailey pointing out that she suggested to her partner, "How about we do Jenga? Because I've been playing Jenga a lot at home." Students also reflected on and expressed how they were active in the process of making sense through the GUA. Hazel explained, "You have to do it on your own, and you need to figure out your own strategies . . . it taught me a different way to do math. . . . you could play this; make it your own." Rather than being told to remember information a teacher would give in a lesson, Hazel shifts her understanding of mathematics as a discipline where she can make sense of content herself, even during an assessment task.

Throughout the implementation of the GUA, students' interest was palpable in the room as they focused intently, constantly communicated mathematically in debating designs and representations, and laughed in delight of the mathematical trickiness of their tower design. Students relied on the word *fun* to indicate growth in becoming interested in thinking mathematically within the task. Some examples include:

Grant: I like that it's fun and you can also learn things from it.

Fiona: We got to build our own tower and stuff, and we had to make up ideas for it to be hard and challenging . . . It's probably what were, what we would be doing in math but in a funner way.

Integrating their own creativity, as Fiona intimates in *make up ideas*, also captured students' interest. Aditi pointed out, "You get to play—build towers. And it's better than just writing out on a piece of paper your answers, because you get to use your own imagination to create something." Her partner, Theo, agreed, and enthusiastically exclaimed, "We got to build a robot!" Beyond the relevant context, Aditi and Theo labelled each shape in their 2D representation with a mathematical label and a robot part. The GUA contributed to students' enhanced views on the nature of mathematics.

Rather than the ethos of a "normal test kind of puts a little pressure on you", as Dillon explained, the GUA opened up space for students' sustained engagement in the challenge posed in creating a new Equilibrio tower and creating the drawing. Observing Kamri and Sage's initial tower creation revealed their tenacious engagement in a challenge, one they did not anticipate in the earlier sketch they brought to class until it was resolved. Riya defended the difficulty of the tower she created with Nyla: "Not to make it symmetrical so it can be harder. Sometimes the symmetrical . . . would be easier because if you look

up the side it will help you with it." She was increasing her productive disposition as she chose to apply a mathematical idea to justify a difficult tower to build. At the end of her screencast, Dillon exclaimed, "It has to be pretty challenging, because people want something like a challenge, not something easy!" Her enthusiasm for meeting the demands of a challenging task and desire to share the challenge with others demonstrates her emerging productive disposition.

And while many of the students identified the task as quite challenging, they were also eager to persevere. Nyla described the persistence she developed through the task as "you just keep on trying" in a proactive manner, with a belief that her efforts would pay off, contrasted with "normal math you have to keep on doing it, . . . and doing it, and doing it until you're dead!" Caitlin described how she persevered in the initial tower design as:

I'm really proud how I created this half of the tower because it was a very, very complicated task. I had to stack tons of blocks together until I could make it correct and then (laughter) we kept on knocking it over, until finally I finally got it to come. . . . That's how I got better at it.

On behalf of their classmates, these students demonstrate that they reformed their sense of challenge in mathematics class as worthy of persistent engagement.

Students were (re)forming their identities [53] as mathematical doers and learners through expressions of ownership. Students' ownership represents an emerging productive disposition in asserting agency and was first apparent on their written records. When Caitlin wrote, "As you see, our tower . . ." and Bailey wrote, "My puzzle is mathematical because it has symmetry . . .," their use of *our* and *my* demonstrated ownership of the product of the creative process, as they see themselves as those who can create a tower with mathematical elements. Dillon demonstrated ownership of her mathematical thinking when she wrote, "The hole tower has simatry but the top. I know becaus if you where to fold the tower except for the Top it will mach but the top wont mach." Dillon inserts herself in taking ownership of her mathematical knowing and communicates that to her teacher through asserting *I know*.

Students were also (re)forming their identities as capable of doing and learning mathematics. Wyatt, a struggling student, began to see himself as having mathematical ideas to share because he "didn't have to write stuff . . . you could build a tower instead." We see this as a nascent emergence of seeing himself as capable in his use of the pronoun *you* instead of first-person *I*, recognizing the individual growth for Wyatt in this moment. Alexandra confirmed our interpretation by explaining, "He would have really struggled to show me that level of thinking on a summative, paper-and-pencil [test]. But he was able to through this game, as he was able to talk about it and he didn't have to write sentences." Not only did students begin to see themselves as capable of doing mathematics like Wyatt, but they also saw themselves as mathematical learners when they described their learning within the assessment task as:

Riya: My idea . . . when we were building, it wasn't really in a test, and then you just included it inside a test that makes people more stronger, confident and stronger with the test because they didn't know they were doing the test anyway when they were building.

Theo: I've learned mathematical thinking . . . I'm getting really good at it, 'cause I know . . . the fractions and vertices.

Caitlin: [The task] makes a difference because it strengthens my brain into mathematics, which encourages me to do more of it, and helps me learn more about shapes and sizes.

Dillon: It's like a fun project and also you can learn lots from it . . . the videos that we made we were doing lines of symmetry.

The students saw themselves as effective in not only demonstrating their learning through the GUA, but also simultaneously as learning during the task. Kilpatrick et al. [47] noted that "children who attribute success to a relatively fixed ability are likely to approach new tasks with a performance rather than a learning orientation" (p. 171). What is apparent from the students' emerging productive disposition is that they viewed this assessment

task from a learning orientation and further demonstrated they were in the process of (re)forming their identity as capable mathematics learners. The GUA enabled the students to be mathematically who they could be, rather than limiting their possibilities of being (mathematical) through a traditional test.

5. Discussion

Through our ongoing collaborative discussions about the particular performance assessment and our ensuing creation of generative unit assessment as a novel approach, we marvel at the spaces that can be found to authentically understand students' mathematical thinking and encourage continual learning within a constrained system. Alexandra was innovative in questioning the efficacy of a traditional paper-and-pencil unit test as a routine, reliable approach to generating students' grades in summative evaluations, and acted with authenticity when she "glimpsed[d] a better and more effective way of carrying out" [21] (p. 784) assessment. In this way, she was "more integrated and coherent, or in other words, more *fully human*" [21] (p. 783, emphasis in original). Critically examining and reflecting on the impact of designing and implementing a generative unit assessment, we emphasize the need to return to authenticity to support continued growth of all involved.

We believe that the four features of generative unit assessment can be applied across all content areas and are not limited by the development of an example in mathematics class. While Vu and Dall'Alba [21] recognized the importance of authentic assessment for students' being and becoming, we extended this ontological orientation to demonstrate that the teacher is fully implicated in the growth through generative unit assessments. A teacher can grow in their awareness of students' disciplinary thinking and shift their pedagogical approaches in seeing themselves as a teacher who sponsors growth. As teachers "take a stand on our being and accept responsibility for the way we live our life" [21] (p. 781) in living authentically, they inspire students to also live with authenticity. Through students' developing richer understanding of disciplinary ideas, whether it is scientific concepts or literary devices, they can integrate these understandings as they "learn to develop their ways of being and inhabiting the world" [21] (p. 779) by participating in disciplinary activities, engaging in scientific inquiry or authoring texts, and in their growth of being mathematicians, scientists, authors, etc. Students' reflections on their learning within the generative unit assessment promote their productive sense of the discipline and themselves in (re)forming their identities as capable.

We anticipate that our initial explication of generative unit assessment will open a fruitful avenue of future exploration for educational researchers. We invite researchers to apply the characteristics of generative unit assessment to other content areas and a broader range of grades to augment the evidence of teacher and student ontological growth through generative unit assessment. Further, we suggest that, while a performance assessment task was effective in our context, researchers continue to explore and adapt other forms of assessment tasks to investigate how they can be generative unit assessments. Multiple examples could be synthesized to define characteristics of tasks that promote generative assessment. Additionally, we suspect that continuity of learning [44] can and needs to be applied more broadly, that students and teachers need multiple experiences of generative unit assessments to continue to grow and benefit from the assessment process. Longitudinal inquiry is needed over a school year, or across years, to understand the possible enduring impact of assessment practices sponsoring ontological growth.

Finally, we have come to understand that the impact of the generative unit assessment as a "model of assessment [that] is inclusive" [54] (p. 157) was not a credit to the performance task as the assessment event at the end of the unit. Rather, in augmenting epistemological matters in assessment that focus on what knowledge and skills students have attained, we follow Vu and Dall'Alba [20] to shape an assessment focus on the "integration of knowing, acting and being" in order to "empower students to form and establish themselves in the world" (p. 786). In this way, learning and assessment become ontological endeavors that attend to students' being and becoming [55]. Generative unit assessment,

as a form of authenticity in a classroom, was about an orientation on the part of the teacher and the students to imbue each moment in the classroom as an educative experience that enabled them to engage in becoming.

Author Contributions: Conceptualization, P.J.M., R.M. and A.C.; methodology, P.J.M. and A.C.; formal analysis, P.J.M., R.M. and A.C.; investigation, P.J.M. and A.C.; resources, P.J.M. and A.C.; writing—original draft preparation, P.J.M. and R.M.; writing—review and editing, P.J.M., R.M. and A.C.; project administration, P.J.M.; funding acquisition, P.J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the University of Alberta under the Support for the Advancement of Scholarship Grant Program and the Roger S. Smith Undergraduate Student Research Award Program. The APC was funded by the University of Alberta and the University of Lethbridge.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Research Ethics Board 1 of the University of Alberta (protocol Pro00051895_AME2 on 12 November 2015) and by the Cooperative Activities Program at Elk Island Public Schools (24 November 2015).

Informed Consent Statement: Informed consent was obtained from all participants and/or legal guardians for involvement in the study, including assent from students.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

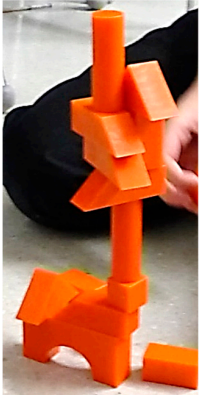
Designing an Equilibrio Tower

You've been hired by FoxMind to design a new tower for the Equilibrio puzzle book!! The game is so popular that players are demanding more puzzles for different levels of difficulty in the book. FoxMind has two design criteria:

- (1) make a tower and matching puzzle tricky so that it's a challenge;
- (2) it has to have math/geometric features to convince more teachers to use Equilibrio in math class!

In this task you will:

- Design an Equilibrio tower: Use as many pieces as you want from the game, and arrange in a tower that you build.
- Create a puzzle: Draw a matching puzzle to be included in the Equilibrio book that can be used by players to build your tower.
- Defend your puzzle: Explain what makes your tower mathematical and tricky to solve.



Part I: Design an Equilibrio Tower

Names _____ Date _____

Before you start building, make a list of at least 4 features you want to include. Use drawings and words.

Build your tower!!

Make sure you've thought about these questions:

- Is our tower stable for someone else to build?
- Will we be able to draw the puzzle picture?
- What features worked and what did we have to change?
- What math ideas are in our tower?

Take a photo with an iPad.

Part II: Create a Puzzle (Sketch)

Use this page to make a first sketch of your tower.



Use the Peer Coaching Tool and interview another group. Use their suggestions on your drawing to make a share-quality drawing.

Peer Coaching Tool: Designing an Equilibrio Tower

Names _____ Date _____

Coaches _____

Instructions for the Coaches: Using the questions below, interview a group who is defending their tower design. Here are some questions to ask the group:

?	How did you decide on the number and types of 3D solids to use to make it tricky enough?
?	How did you design your tower so you can show your great geometry ideas?
?	Why do you think your sketch shows your tower well enough for someone else to build?
!	One suggestion I have for you is...

Instructions for the Student:

Now that you have explained your tower and puzzle, listened to the feedback, and thought about your work, go back and improve your puzzle in at least one way.
--

Part II: Create a Puzzle (Share-Quality)

Use this page to make a share-quality drawing of your tower.

Part III: Finding Math Ideas with Your Partner

Use the photo on your iPad and the app "Explain Everything". Draw on the photo any math ideas you find, and record your voices explaining!

To put your puzzle in the Equilibrio puzzle book, circle the level of difficulty for your puzzle:

yellow
orange
green
blue
purple
red

Explain why your puzzle is that level of difficulty.

Part IV: Defend Your Puzzle (Explain)

Here are best four math ideas that make your tower mathematical and tricky:

1. _____
2. _____
3. _____
4. _____

Part IV: Defend Your Puzzle (Draw)

Here are some drawings to defend your best four math ideas:

1.

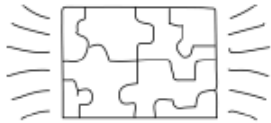
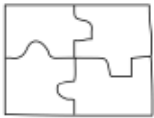


--
2.

--
3.

--
4.

--

Rubric

	Connections	Problem-Solving	Communication	Reasoning
<p>Expert I have an excellent understanding and could thoroughly teach it to someone else.</p> 	<ul style="list-style-type: none"> <input type="checkbox"/> I can make in depth connections and extend my tower/puzzle to things I have learned, or know (in and out of school). <input type="checkbox"/> I can make in depth connections between math ideas and my tower/puzzle. <input type="checkbox"/> I can make in depth connections with my built tower and my drawn puzzle. 	<ul style="list-style-type: none"> <input type="checkbox"/> I can use efficient strategies that make my tower tricky. <input type="checkbox"/> I can analyze the shapes I used and their features in depth. <input type="checkbox"/> I am able to correctly adjust my first tower and sketch (if needed) to represent my tower in my share-quality sketch. 	<ul style="list-style-type: none"> <input type="checkbox"/> I can provide an in-depth explanation of my best four math ideas using precise mathematical language. <input type="checkbox"/> I can draw an accurate puzzle to match my tower. 	<ul style="list-style-type: none"> <input type="checkbox"/> I can make an efficient and advanced concluding defense to why my tower is tricky to solve. <input type="checkbox"/> I have a comprehensive mathematical basis for explaining why my tower is mathematical.
<p>Practitioner I have a strong understanding.</p> 	<ul style="list-style-type: none"> <input type="checkbox"/> I can make thorough connections with my tower/puzzle to things I have learned, or know (in and out of school). <input type="checkbox"/> I can make thorough connections between math ideas and my tower/puzzle. <input type="checkbox"/> I can make thorough connections with my built tower and my drawn puzzle. 	<ul style="list-style-type: none"> <input type="checkbox"/> I can use effective strategies that make my tower tricky. <input type="checkbox"/> I can analyze the shapes I used and their features thoroughly. <input type="checkbox"/> I am able to adjust my first sketch (if needed) to represent my tower in my share-quality sketch. 	<ul style="list-style-type: none"> <input type="checkbox"/> I can provide a thorough explanation of my best four math ideas using specific mathematical language. <input type="checkbox"/> I can draw a somewhat accurate puzzle to match my tower. (two errors or less) 	<ul style="list-style-type: none"> <input type="checkbox"/> I can make a correct concluding defense to why my tower is tricky to solve. <input type="checkbox"/> I have a substantial mathematical basis for explaining why my tower is mathematical.
<p>Apprentice I almost got it, but still need help.</p> 	<ul style="list-style-type: none"> <input type="checkbox"/> I can make mostly complete connections with my tower/puzzle. <input type="checkbox"/> I can make mostly complete connections between math ideas and my tower/puzzle. <input type="checkbox"/> I can make mostly complete connections with my built tower and my drawn puzzle. 	<ul style="list-style-type: none"> <input type="checkbox"/> I can use workable strategies that make my tower tricky. <input type="checkbox"/> I can analyze the shapes I used and their features sufficiently. <input type="checkbox"/> I am able to adjust some parts of my first sketch (if needed) to represent my tower in my share-quality sketch. 	<ul style="list-style-type: none"> <input type="checkbox"/> I can provide a mostly complete explanation of my best four math ideas using basic mathematical language. <input type="checkbox"/> I can draw a partially accurate puzzle to match my tower. (three or more errors) 	<ul style="list-style-type: none"> <input type="checkbox"/> I can defend why my tower is tricky to solve, but I have gaps in explaining why. <input type="checkbox"/> I have a partial mathematical basis for explaining why my tower is mathematical.
<p>Novice I need help in understanding.</p> 	<ul style="list-style-type: none"> <input type="checkbox"/> I need help in making connections related to my built tower that are relevant without help. <input type="checkbox"/> I need help in making connections between math ideas and my tower/puzzle. <input type="checkbox"/> I need help in making connections with my built tower and my drawn puzzle. 	<ul style="list-style-type: none"> <input type="checkbox"/> I need help to understand the problem. <input type="checkbox"/> I need help in analyzing the shapes I used and their features. <input type="checkbox"/> I did not make any adjustments from my first sketch to my share-quality sketch, even though it was needed. 	<ul style="list-style-type: none"> <input type="checkbox"/> I can provide a sketchy explanation of my best four math ideas using inaccurate mathematical language. <input type="checkbox"/> I need help in drawing a puzzle to match my tower. 	<ul style="list-style-type: none"> <input type="checkbox"/> I need help with defending why my tower is tricky to solve. <input type="checkbox"/> I have an underdeveloped mathematical basis for explaining why my tower is mathematical.

References

1. Shepard, L.A. The role of assessment in a learning culture. *Educ. Res.* **2000**, *29*, 4–14. [CrossRef]
2. Delandshere, G. Assessment as inquiry. *Teach. Coll. Rec.* **2002**, *104*, 1461–1484. [CrossRef]
3. Andrade, H. Classroom assessment in the context of learning theory and research. In *SAGE Handbook of Research on Classroom Assessment*; McMillan, J.H., Ed.; Sage: Thousand Oaks, CA, USA, 2013; pp. 17–34.
4. Clark, I. Formative assessment: A systematic and artistic process of instruction for supporting school and lifelong learning. *Cdn. J. Educ.* **2012**, *35*, 24–40.
5. Western and Northern Canadian Protocol for Collaboration in Education. (WNCP). *Rethinking Classroom Assessment with Purpose in Mind: Assessment for Learning, Assessment as Learning, Assessment of Learning*; Manitoba Education, Citizen and Youth: Winnipeg, MB, Canada, 2006. Available online: <https://www.wncp.ca/media/40539/rethink.pdf> (accessed on 10 June 2021).
6. Dann, R. Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assess. Educ. Princ. Pol. Prac.* **2014**, *21*, 149–166.
7. Newton, P.E. Clarifying the purposes of educational assessment. *Assess. Educ.* **2007**, *14*, 149–170. [CrossRef]
8. Burke, K. *How to Assess. Authentic Learning*, 5th ed.; Corwin: Thousand Oaks, CA, USA, 2009.
9. National Council of Teachers of Mathematics. (NCTM). *Principles to Action: Ensuring Mathematical Success for All*; NCTM: Reston, VA, USA, 2014.
10. Suurtamm, C.; Thompson, D.R.; Kim, R.Y.; Moreno, L.D.; Saya, N.; Schukajlow, S.; Silver, E.; Ufer, S.; Vox, P. *Assessment in Mathematics Education: Large-scale Assessment and Classroom Assessment*; Springer Open: New York, NY, USA, 2016.
11. McMillan, J.H. Secondary teachers' classroom assessment and grading practices. *Educ. Meas. Issues Prac.* **2001**, *20*, 20–32. [CrossRef]
12. Popham, W.J. *Classroom Assessment: What Teachers Need to Know*, 7th ed.; Pearson Education: Boston, MA, USA, 2014.
13. Frey, B. *Modern Classroom Assessment*; Sage: Thousand Oaks, CA, USA, 2014.
14. Lane, S. Performance assessment. In *SAGE Handbook of Research on Classroom Assessment*; McMillan, J.H., Ed.; Sage: Thousand Oaks, CA, USA, 2013; pp. 313–330.
15. Fuchs, L.S.; Fuchs, D.; Karns, K.; Hamlett, C.L.; Katzaroff, M. Mathematics performance assessment in the classroom: Effects on teacher planning and student problem solving. *Am. Educ. Res. J.* **1999**, *36*, 609–646. [CrossRef]
16. Palm, T. Performance assessment and authentic assessment: A conceptual analysis of the literature. *Prac. Assess. Res. Eval.* **2008**, *13*. [CrossRef]
17. Darling-Hammond, L.; Barron, B.; Pearson, P.D.; Schoenfeld, A.H.; Stage, E.K.; Zimmerman, T.D.; Cervetti, G.N.; Tilson, J.L. *Powerful Learning: What We Know about Teaching for Understanding*; Jossey-Bass: San Francisco, CA, USA, 2008.
18. Zane, T.W. Performance assessment design principles gleaned from constructivist learning theory (part 1). *TechTrends* **2009**, *53*, 81–90.
19. Zane, T.W. Performance assessment design principles gleaned from constructivist learning theory (part 2). *TechTrends* **2009**, *53*, 86–94.
20. Gulikers, J.T.M.; Bastiaens, T.J.; Kirschner, P.A. A five-dimensional framework for authentic assessment. *Educ. Technol. Res. Dev.* **2004**, *52*, 67–86. [CrossRef]
21. Vu, T.T.; Dall'Alba, G. Authentic assessment for student learning: An ontological conceptualisation. *Educ. Phil. Theory* **2014**, *46*, 778–791. [CrossRef]
22. Heidegger, M. *Being and Time*; Macquarrie, J.; Robinson, E., Translators; Harper & Row: San Francisco, CA, USA, 1962.
23. Carr, W.; Kemmis, S. *Becoming Critical: Education, Knowledge, and Action Research*; Falmer Press: London, UK, 1986.
24. Kemmis, S.; McTaggart, R.; Nixon, R. *The Action Research Planner: Doing Critical Participatory Action Research*; Springer: Singapore, 2014.
25. McIntyre, A. *Participatory Action Research*; Sage Publications: Thousand Oaks, CA, USA, 2008.
26. Whyte, W.F. (Ed.) *Participatory Action Research*; Sage: Thousand Oaks, CA, USA, 1991.
27. McFeetors, P.J.; Ireland, K. Ready, SET, play. *Delta-K* **2016**, *53*, 38–47.
28. McFeetors, P.J.; Palfy, K. Educative experiences in a games context: Supporting emerging reasoning in elementary school mathematics. *J. Math. Behav.* **2018**, *50*, 103–125. [CrossRef]
29. Mason, J.; Burton, L.; Stacey, K. *Thinking Mathematically*, 2nd ed.; Prentice Hall: Harlow, UK, 2010.
30. Stein, M.K.; Grover, B.W.; Henningsen, M. Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *Am. Educ. Res. J.* **1996**, *33*, 455–488. [CrossRef]
31. Stein, M.K.; Smith, M.S.; Henningsen, M.A.; Silver, E.A. *Implementing Standards-based Mathematics Instruction: A Casebook for Professional Development*, 2nd ed.; Teachers College & National Council of Teachers of Mathematics: Reston, VA, USA, 2009.
32. McFeetors, P.J.; Mason, R.T. Learning deductive reasoning through games of logic. *Math. Teach.* **2009**, *103*, 284–290. [CrossRef]
33. Herbel-Eisenmann, B.A.; Cirillo, M. (Eds.) *Promoting Purposeful Discourse: Teacher Research in Mathematics Classrooms*; National Council of Teachers of Mathematics: Reston, VA, USA, 2009.
34. Smith, M.S.; Stein, M.K. *5 Practices for Orchestrating Productive Mathematics Discussions*; National Council of Teachers of Mathematics: Reston, VA, USA, 2011.

35. Cook, T. The purpose of mess in action research: Building rigour through a messy turn. *Educ. Action Res.* **2009**, *17*, 277–291. [[CrossRef](#)]
36. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psych.* **2006**, *3*, 77–101. [[CrossRef](#)]
37. Dey, I. Grounding categories. In *The SAGE Handbook of Grounded Theory*; Bryant, A., Charmaz, K., Eds.; Sage: Thousand Oaks, CA, USA, 2010; pp. 167–190.
38. Charmaz, K. *Constructing Grounded Theory*, 2nd ed.; Sage: Thousand Oaks, CA, USA, 2014.
39. Partridge, E. *Origins: A Short Etymological Dictionary of Modern English*; Routledge: London, UK, 2009.
40. Hoad, T.F. *Oxford Concise Dictionary of English Etymology*; Oxford University Press: New York, NY, USA, 1996.
41. Chambers. *Chambers's Etymological Dictionary of the English Language*; G. N. Morang: Toronto, ON, Canada, 1903.
42. Harlen, W. Teachers' summative practices and assessment for learning—Tensions and synergies. *Curr. J.* **2005**, *16*, 207–223. [[CrossRef](#)]
43. McMillan, J.H. (Ed.) *SAGE Handbook of Research on Classroom Assessment*; Sage: Thousand Oaks, CA, USA, 2013.
44. Dewey, J. *Experience and Education*; Touchstone: New York, NY, USA, 1997.
45. Talanquer, V.; Bolger, M.; Tomanek, D. Exploring prospective teachers' assessment practices: Noticing and interpreting student understanding in the assessment of written work. *J. Res. Sci. Teach.* **2015**, *52*, 585–609. [[CrossRef](#)]
46. Skemp, R.R. Relational understanding and instrumental understanding. *Math. Teach. Mid. Sch.* **2006**, *12*, 88–95. [[CrossRef](#)]
47. Kilpatrick, J.; Swafford, J.; Findell, B. *Adding It Up: Helping Children Learn. Mathematics*; National Academy of Sciences: Washington, DC, USA, 2001.
48. Mason, J. Being mathematical with and in front of learners: Attention, awareness, and attitude as sources of differences between teacher educators, teachers and learners. In *International Handbook of Mathematics Teacher Education: The Mathematics Teacher Educator as Developing Professionals*; Jaworski, B., Woods, T., Eds.; Sense: Rotterdam, The Netherlands, 2008; Volume 4, pp. 31–55.
49. National Council of Teachers of Mathematics. (NCTM). *Principles and Standards for School Mathematics*; NCTM: Reston, VA, USA, 2000.
50. Lampert, M. Introduction. In *Talking Mathematics in School: Studies of Teaching and Learning*; Lampert, M., Blunk, M.L., Eds.; Cambridge University Press: Cambridge, UK, 1998; pp. 1–14.
51. Lee, C. *Language for Learning Mathematics: Assessment for Learning in Practice*; Open University Press: Maidenhead, Berkshire, UK, 2006.
52. Glass, B. Comparing representations and reasoning in children and adolescents with two-year college students. *J. Math. Behav.* **2004**, *23*, 429–442. [[CrossRef](#)]
53. McFeetors, P.J.; Mason, R.T. Voice and success in non-academic mathematics courses: (Re)forming identity. *For Learn. Math.* **2005**, *25*, 16–23.
54. Brookhart, S. Successful students' formative and summative uses of assessment information. *Assess. Educ. Princ. Pol. Prac.* **2001**, *8*, 153–169. [[CrossRef](#)]
55. Packer, M.J.; Goicoechea, J. Sociocultural and constructivist theories of learning: Ontology, not just epistemology. *Educ. Psychol.* **2000**, *35*, 227–241. [[CrossRef](#)]