International Journal of Game-Based Learning
Volume 14 • Issue 1

Open-Ended Mathematics Learning:
Implications From the Design of a Sandbox Game

Matheus A. Cezarotto, New Mexico State University, USA*
https://orcid.org/0000-0001-6657-753X

Pamela N. Martinez, New Mexico State University, USA
https://orcid.org/0000-0003-4718-310X

Ruth Constansa Torres Castillo, New Mexico State University, USA

Theodore Stanford, New Mexico State University, USA

Christopher Engledowl, Lastinger Center for Learning, University of Florida, USA

Germain Degardin, University of New Mexico, USA
https://orcid.org/0009-0007-9583-1120

Barbara Chamberlin, New Mexico State University, USA

ABSTRACT

Mathematical learning has an important role and is often prioritized in education. In K-16 education, algebra is one of the most vital mathematical content domains: it represents one of the top barriers for students pursuing a postsecondary education. Game-based learning has been effective in fostering classroom math learning environments that are collaborative and focused on conceptual understanding. Sandbox games provide open-ended learning environments where players can set their own goals and level of effort. As part of the project "Math Snacks," the team designed Agrinautica, a sandbox game to enable constructivist-informed early algebra learning. This article identifies design recommendations for creating meaningful sandbox games for learning, considering students' and teachers' needs. Researchers discuss the decisions to create a sandbox game and describe challenges inherent in math learning through sandbox-type gameplay. This study provides impact results from a large-scale study of users of the game, and shares recommendations for developing future sandbox learning games.

KEYWORDS

Agrinautica, Constructivism, Early Algebra, Educational, Gameplay, Math Snacks, Sandbox, Transformational Games

DOI: 10.4018/IJGBL.337795

*Corresponding Author

INTRODUCTION

Mathematics maintains an important and prioritized role in K–16 education, and algebra is perhaps one of its most vital content domains. As such, algebra has risen to the top in maintaining a gatekeeper role that has kept many students from basic levels of access to post-secondary education, thus further disadvantaging the most vulnerable students (Moses & Cobb, 2001). To support students in overcoming some of these challenges, early algebra can provide an important gateway through which students may excel in the domain of algebra.

In addition to recognizing the importance of learning early algebra concepts within a constructivist-oriented learning environment (e.g., National Council of Teachers of Mathematics [NCTM], 2000; Blanton et al., 2015), educational technology has become ubiquitous not only in schools (Gray & Lewis, 2021) but also in homes and among the general public (Rideout, 2014). The use of digital games to learn mathematics has been increasing in popularity (Byun & Joung, 2018). However, there is much to learn, in particular when it comes to understanding the impact of such games on students' mathematics achievement and conceptual understanding (Byun & Joung, 2018). For example, all games are not necessarily built to reflect or enable a constructivist approach to learning. Some games may focus on memorization or practice, and may use a more behaviorist approach in motivating students to do routine tasks with the aim to get a reward.

As educators face the need for effective educational early algebra games and strive to understand the impact that some game genres have on players, research is emerging on how to connect constructivist approaches to math-based games, including ways to build games that foster classroom learning environments that are collaborative, dialogic, and focused on conceptual understanding of important and rigorous mathematics, such as early algebra.

In an attempt to address these complex mathematics teaching and learning needs, an experienced team of education researchers and developers created a highly successful suite of mathematics animations and games, called Math Snacks (Wiburg et al., 2016). Building on these past successes, the team moved on to the Math Snacks project, which sought to prepare learners for algebra by exposing them to foundational concepts earlier in their mathematics learning. All the games in the Math Snacks suite were designed from a constructivist perspective—attempting to enable the user to build their own understandings of how algebra concepts work. As part of that work, the Math Snacks Early Algebra project team wanted to create a sandbox game through which learners could experiment with different approaches to writing, manipulating and interpreting expressions; understand the role operators and parentheses had in those expressions; and become experienced in creating many different kinds of expressions. The game was eventually named Agrinautica.

In this article, we share the theoretical underpinnings of designing the game Agrinautica and provide results from multiple studies of students' and teachers' experiences using the game in the classroom. Sandbox games are more open ended than traditional goal- or score-oriented games and tend to give players greater autonomy in deciding how to play the game and in setting their own goals. Thus, a collection of students in a given classroom could all have different experiences playing the same game. As such, it was important for the team to review the different ways players approached Agrinautica gameplay, identify the impact of those choices on the player, identify the impact of the gameplay on teachers offering the game, and determine best practices for developing sandbox-style games for learning.

This study addresses three main questions: What is the impact of gameplay in Agrinautica on student cognitive and affective outcomes? What is the impact of the intervention on teacher affect and pedagogical practice? In what ways does the sandbox nature of the game influence these impacts? A mixed-methods approach was used across the project to analyze the formative and summative data to investigate these questions. We share what has been learned through these findings in the form of recommendations for the development of future sandbox learning games.

BACKGROUND

Game-Based Learning for Mathematics

Effective *digital game-based learning* experiences allow learners to learn specific content in a meaningful way by exploring a game's new world. According to the *integrated design framework for playful learning*, games intending to engage players in game-based learning should be first grounded in particular content, and then they should integrate all other game elements focused around that content: an engaging narrative, aesthetic elements, responsive feedback, progressive goals, and a reward system (Mayer, 2014; Plass et al., 2020a). Using the full range of game features, game-based learning interventions have clear educational goals, instructional content, and a pedagogical approach informed by discipline-specific theory and research (Plass et al., 2020b).

Games as part of classroom-based lessons have been used in differing pedagogical capacities—for example, as a "warm-up" exercise, to introduce new mathematical concepts, to consolidate skills and knowledge, and for fluency practice (Squire, 2005; Egenfeldt-Nielsen, 2010; Russo et al., 2021). Despite the popularity of using games for teaching mathematics content, there is a paucity of research on the effect of digital game-based learning on students' mathematical knowledge (Byun & Joung, 2018), which is even less expressive in early algebra knowledge.

Although some view the field of early algebra research as having matured (Blanton et al., 2015), the most recent *Compendium for Research in Mathematics Education* presents literature on early algebra that does not mention a single study utilizing a digital game and mentions only two other studies that incorporated gamified tasks (Stephens et al., 2017). Few studies of digital games have focused on early algebra. Kolovou et al. (2013) and van den Huevel-Panhuizen et al. (2013) examined game play and pre/post data from 489 students with the game Hit the Target, which addresses covarying quantities—an early algebra concept (Blanton et al., 2015; Stephens et al., 2017). Kolovou et al. (2013) found that students who played the game outperformed those who didn't on an early algebra assessment, with an effect size of $d = 0.31$. Huevel-Panhuizen et al. (2013) found a substantial variability in play time, and they found significant effects on gaining knowledge for students who exhibited some intention in their play—such as exploring relationships, as opposed to free play. In another study, ter Vrugte et al. (2017) investigated student learning outcomes after they had played the game Zeldenrust, in which players completed challenges at a hotel to make as much money as they could. Although instructional support was provided in addition to gameplay, there were struggles to motivate all students to independently engage in the tasks, underscoring the need to research how to more effectively pair game play with instruction in a classroom environment. Despite these studies' contributions, however, there is great need for research in game design for early algebra applications and into how these games impact students' knowledge gain. Although assessing the impact of a given game as an intervention is valuable, to make a progress in the game-based learning field, studies need to offer guidance about the types of activities that yield the greatest impact, with implications for ways in which the design of a game aligns with different theories and approaches for math education.

Constructivism and Sociocultural Theory in a Sandbox Experience

In developing a game for the intervention in this study, we grounded learning in constructivist and sociocultural frameworks (e.g., Ernest, 1994; 1996; Mercer & Howe, 2012) for teaching mathematics. In this vein, early algebra concepts (e.g., Blanton et al., 2015; Stephens et al., 2017) were considered alongside the types of mathematical processes students should engage in as they interact with the content. From a constructivist perspective (Cennamo, 2003), math learners should be allowed to explore different concepts, approaches, and experiences so that they can build their own understanding of how early algebra works. From this lens, the early algebra activities need to take into consideration learners' individualities (understandings, values, and beliefs) and previous experiences. One way to foster this constructivist instruction in a game is articulating activities that allow players to explore, define their goals, think critically, and then take actions. Thus, in this project scope, the game environment and supporting activities need to provide students with extensive opportunities to explore and build expressions with intent while also being able to explore how to use the four operations and grouping symbols.

The sandbox game environment provides a promising space to foster constructivist activities that can lead to meaningful learning. Sandbox games offer players an open-world environment to freely explore based on their imagination and creativity (Breslin, 2009). Researchers have been exploring the sandbox gameplay style in educational applications, including studies with the popular Minecraft Education Edition (Microsoft, 2022; Ellison & Drew, 2020; and Hébert & Jenson, 2020). According to Breslin (2009), Adams (2010), and Dormans (2012), a Sandbox gameplay offers the following features:

- *Open and alternative goals:* An open-ended world presents multiple goals that players can pursue. Different from other game genres in which the game sets the goals for a player to meet, a sandbox game offers an environment for exploration, and it also offers several different types of activities for the player to use to meet various different goals. For example, in Agrinautica, players can create 100 different types of plants in thousands of different ways. The players choose which plants to create, based on their own goals for their garden
- *Fewer gameplay restrictions:* Players in sandboxes can explore the game world and mechanics with less restricted guidance on what they should do, compared with other game genres. In Agrinautica, they are given some simple parameters, "build an expression with *these numbers,"* but are given no score or punishment based on their choices.
- *No victory condition:* A sandbox game does not present the player with a victory condition; rather, the game system lets players determine their own achievement goals for the play experience. With Agrinautica, players were often seen sharing their screens with other players, demonstrating pride in their own selected achievements.
- *Offer a free world in a framework:* Players will be introduced to this free world of possibilities with fewer restrictions; their autonomy will be guided in a set of possibilities determined by the game framework. In Agrinautica, players can explore and build different expressions following an expression structure that the game offers.

DEVELOPMENT OF THE GAME: AGRINAUTICA

For designing Agrinautica, game developers and researchers worked directly with content experts, a math professor, a math educator who taught teachers, math teachers, and graduate students from education, math, and computer science. Using the Transformational Design Process (Chamberlin & Schell, 2018), the team articulated the educational objectives to define the *change* needed in the learners, identified what learners have to *do* to create that change, and explored how the game could engage learners in the needed actions. The team started the design process by identifying the *problem* that 4th and 5th graders face when learning early algebra and establishing outcomes. With this in mind, the team's desired mathematical outcomes for students playing Agrinautica were as follows:

- Create lots of numeric expressions and become comfortable and familiar with them, ultimately understanding how to create numeric expressions that match their intent.
- Identify patterns where changing an expression in some way changes (or doesn't change) the output value and use these patterns strategically.
- Gain experience using parentheses correctly and effectively to achieve a desired output value.

- Experiment with standard "order of operations" conventions—for example, that multiplication is performed before addition in the absence of parentheses—and receive meaningful feedback to discover the ways in which the rules of operations work.

The next step was to articulate the changes intended for students in the sense of how different they would be after playing the game. Agrinautica aims for students to be able to build accurate expressions in different ways to meet their intent. After playing Agrinautica, students will be able to understand the order of syntax, understand why and when it doesn't align with their intent, create expressions, and edit existing expressions to get different results.

With the intended content and approach outlined, the team started designing activities that could lead to that transformation. The goal was to understand what was available, know what teaching approaches were working, and to expand ideas for gameplay and game mechanics.

Agrinautica allows players to create a unique planet with beautiful creations while using a rich playground for exploring math expressions. Players are encouraged to freely create expressions using four randomized numbers between 1 and 9, four operators, and sets of parentheses. The players then collect the creations (e.g,. plants, crystals, rocks), while choosing what they want to be in the game (Figure 1). They earn specific creations based on the structure of their expressions. For example, if a player creates an expression using one operation and one set of parentheses that results in "5," it creates a "fungi puddingcupca" (Figure 2), where an expression using three types operators and one set of parentheses resulting in "2," creates a "Spikucus chompia" (Figure 2). Additionally, an in-game guide book offers reference help to direct learners in choosing creations (Figure 3).

Design Challenges and Decisions

During the design process, the team addressed many design challenges, including keeping players engaged during the gameplay experience and fostering content learning. In this section we describe two of these challenges and how the team approached them.

Decisions to Improve Gameplay Engagement

While balancing gameplay and learning goals to create an engaging, meaningful learning experience, the team faced several challenges, including the following:

How can the game guide players to write meaningful expressions? We did not want to tell players which expressions to write, but we also didn't want them to be completely lost, sitting and typing random combinations of symbols (invalid expressions) that did not mean anything. We addressed this concern by giving players values to target with their expressions (0–9). Expressions that produce target numbers bring beautiful and visually appealing creations (e.g., fungus, plants, crystals). However, any valid expression that does not evaluate to an integer from 0 to 9 creates a "weed" (Figure 4), which does not tell them they are wrong, but conveys the message that they did not target the asked-for numbers.

Figure 1. Agrinautica number building bar; players click and drag parentheses, four operators, and four randomized numbers (from 1 to 9) to build expressions

Figure 2. Examples of creatures based on the expression's structure

Figure 3. In-Game guide book informing players about the creation's possibilities

Figure 4. Feedback screen when players get a weed

How do we redirect "wrong" answers? Because the number of plants is finite (92), all expressions had to offer a whole number ranging between 0 and 9. It is possible to make valid expressions in the tool that don't compute to that range. For example, a player could make the following expression: (10-5)/2. This evaluates to 2.5, which is not a whole number. The answer was to create weeds as a way of saying, "You made a valid expression! But it's not on the target list."

Thus, creating a weed is easy. For some players, that had a clear effect, and they deleted weeds immediately and tried to avoid them. Others seemed to make a lot of weeds, either because they were confused or because they chose to expend little effort. To address this challenge, the team worked on the game feedback. Agrinautica provides *feedback* on the building process for each expression: The visual procedural steps describe which operations took place and in which order (Figure 5, left). Additionally, if players are stumped, *the game offers help,* providing access to a help guide with information about expressions (Figure 5, right). With the right feedback, players who did not want weeds were better able to avoid them.

Figure 5. Visual procedural steps (left) and help guide (right)

How do we encourage diversity in expression building? Nothing in the game stops players from repeating the same expression over and over; however, the team goal was to give learners an opportunity to create many different kinds of expressions. The game provided players with an *achievement*, a visual, collectible reward for making something new (Figure 6). Players can also access the *field guide* (Figure 6), which informs players about all the species they can get in the game based on values and expression building. Serving as a resource and a motivation, the field guide and achievements offered different ways to encourage players to explore, regardless of their own goals for the game.

Decisions to Improve Content Learning

Agrinautica encourages players to write numeric expressions. Expressions are building blocks used in equations, functions, formulas, and other places. An algebraic expression may have numbers, variables, operations, and parentheses. For example, the expression $5(x+1) - y$ has three operations, one set of parentheses, two variables, and two numbers (Figure 7). A numeric expression may contain numbers, operations, and parentheses, but it does not contain variables.

A student may be motivated to create a plant called *Spiky limnoria*, which requires an expression with at least two types of operations and two sets of parentheses and which evaluates to the number 4 (Figure 8). To create such an expression, the student would need to *look for and make use of structure* (National Governors Association Center for Best Practices [NGACBP]

Figure 6. Achievements screen (left) and field guide screen (right)

Figure 7. Although there are no variables in the expressions created in Agrinautica, the study of numeric expressions lays the foundation for the later study of algebraic expressions in several important ways

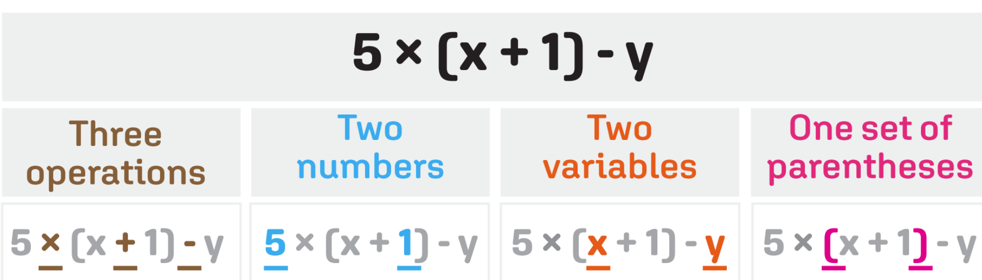


Figure 8. Screenshot from Agrinautica field guide showing how to create spiky Limnoria

& Council of Chief State School Officers [CCSSO], 2010) by recognizing how the sets of parentheses may alter what the expression evaluates to and also how the placement of operations may change how the parentheses are used. By engaging in this *reasoning* (NCTM, 2000), if the student had played the game for a more extended amount of time, there would also be a reasonable expectation that the student would engage in *expressing regularity in repeated reasoning* by recognizing a more generalizable characteristic of how operations and parentheses interact with one another when evaluating an expression. Over time, and through a learning process defined from a constructivist lens, the student would experience enough different expressions to begin to refine their *attention to precision* (NGACBP & CCSSO, 2010) in crafting intentional expressions, developing conceptual understanding, and working toward procedural fluency (Kilpatrick et al., 2001) within a risk-free, exploratory environment.

Smith and Stein (2018) and NCTM (2014) offers guidance for facilitating meaningful discussions—pedagogical moves that are research based and grounded in the same kinds of constructivist and sociocultural theory frameworks on which the Math Snacks project is based. Collaboration and mathematical discourse are highly encouraged during gameplay, and they are also highly encouraged as part of the associated inquiry-oriented classroom activities. In the classroom supporting activity that was designed for Agrinautica, students are placed into groups, and there is a specific heading in the written lesson plan (available at www.mathsnacks.org) titled “Facilitating student learning” that encourages teachers to pay close attention to what students are doing and to select a couple of expressions for larger discussion. The sections that follow in the lesson plan continue these types of suggestions by providing examples of how to engage students in having discussions with one another, supporting students in such practices as *construct viable arguments and critique the reasoning of others* (NGACBP & CCSSO, 2010) and *communication* (NCTM, 2000), while supporting them in developing *conceptual understanding* (Kilpatrick et al., 2001).

Formative Research and Evaluation

We used formative evaluation during the game development to inform design decisions. Formative research involved a mixed-methods approach by incorporating observations, interviews, and user testing with students and teachers using early versions and prototypes of the game. There were a total of 20 testing sessions throughout 2017–2019 that involved continual user testing of Agrinautica as it evolved and was refined.

Observational data of students playing early versions of the game informed the team on various aspects of the gameplay interaction and instructional design—for example, the complexity of the expressions built by students and students' reuse of expressions when playing the game. Interviews allowed the team to investigate with users specific things about the game; for example, students were asked to draw possible characters for the game and to think about titles for the game. In user testing sessions, the development team used early versions of the game to understand players' gameplay interaction, including usability and accessibility issues.

ASSESSMENT

Researchers used *summative assessment* to evaluate the impact of Agrinautica with students and teachers, which includes students' cognitive and affective outcomes and teachers' perception. The research was approved by the NMSU Institutional Review Board (22576). As part of the larger Math Snacks Early Algebra project, Agrinautica research data are summarized in previous articles (Torres Castillo & Morales, 2019; Engledowl, 2020; Engledowl et al., 2021; Morales et al., 2019; Torres et al., 2016), and the data serve as valuable reference points in summarizing the design and development process as well as the use and impact of this sandbox game. For example, a detailed analysis of all the research presents all the methods in depth for the Math Snacks project and describes the findings from a qualitative analysis undertaken after the intervention (Torres Castillo, 2020); and other statistical analysis reports on the levels of sophistication that students exhibited on the early algebra assessment (Engledowl et al., 2021).

Although this article focuses on Agrinautica's impacts on students and teachers, it is important to outline the full intervention used in the larger study, which encompasses the following elements:

- Additional games: Three games covering early algebra key content were developed: *Agrinautica* (expressions), *Curse Reverse* (expressions and variables) and *Creature Caverns* (graphing relationships). All these were informed by a constructivist lens, but Agrinautica is the only sandbox.
- Teacher guides: For each game, a teacher guide was created that included supplemental activities and lesson plans guiding teachers on how to use the game as part of a classroom activity.
- Professional learning workshops: Teachers received two training sessions on how to implement the games and a lesson plan into their classroom teaching.

Researchers implemented the intervention across 10 elementary schools during classroom time between January and February of 2019. Different instruments and measurements were used to collect data from teachers and students before, during, and after the intervention. Table 1 provides an overview of the intervention and data collection.

Subjects

Participation came from 544 students who initially signed consent forms, from 28 4th and 5th grade elementary teachers, across 10 elementary schools within a single midsized public school district in the Southwestern region of the United States. Of these 544 students, 457 completed both the pre-

Table 1. Math snacks study overview

	Sep	Oct-Dec	Jan-Feb (Intervention)	Mar-Apr	May
Students	**Pre-Survey** Background & Perceptions	**Pre-Assessment** Knowledge	**Played Math Snacks games** (classroom) **Embedded data** collection	**Post-Assessment** Knowledge	**Post-Survey** Background & Perceptions
Teachers	**Pre-Survey** Background & Perceptions	1st Classroom **Observation** 1st Professional **Workshop**	**Taught using resources** from the project Fidelity **Observations** **Reflection logs** 2nd Professional **Workshop**	2nd Classroom **Observation**	**Post-Survey** Background & Perceptions Teachers **Interviews**

2018 Sep Oct-Dec Jan-Feb Mar-Apr May 2019

and post-assessment of early algebra knowledge. Of these 457 students, we were able to uniquely identify 410 who spent time playing Agrinautica. Although data were articulated for all games in the study, data shown in this article relate specifically to the play of Agrinautica. Additionally, 28 teachers participated in the study.

Instruments Measurements

Students

Across the larger project, for students, there were pre- and post-surveys designed to collect information about student demographic information, students' experiences with computer games and computer games for learning mathematics, their mathematics anxiety, mathematics self-concept, and more broadly their persistence in learning. Moreover, an early algebra assessment was constructed that has shown promising validity and reliability evidence (Engledowl, 2020); it was administered as a pre- and post-test. Finally, embedded telemetry data were collected on student game play for Agrinautica.

Teachers

To understand the larger context, and to examine teachers' experiences in using and teaching with the Math Snacks Early Algebra products, pre- and post-surveys were used to collect information about teachers' experiences teaching math with technology and their levels of experience teaching at the elementary level; their experiences playing games themselves and using games in the classroom; their attitudes toward using games for learning mathematics; and the level of access to technology they have in their classrooms. In addition, pre- and post-observations of the classroom learning environment were conducted using an observation protocol (Valdez, 2012), one randomly selected Math Snacks Early Algebra lesson was observed using a fidelity observation protocol, and teachers submitted online self-reflection logs about their experiences teaching the lesson associated with the game.

Findings

Students

To understand the impact of the intervention on students' early algebra knowledge, two separate analyses have been conducted. For the first one, Engledowl et al. (2021) drew on Rasch person-ability scores and item difficulty scores (Bond & Fox, 2015) to describe four different levels of sophistication in early algebra knowledge and to track changes in these levels from pre- to post-test. For the second one, Engledowl et al. (in development) are using hierarchical linear modeling (Raudenbush & Byrk, 2002) with students (Level 1) nested within teachers (Level 2) to examine the impact of the intervention and to describe any differential effects.

Drawing on Rasch person-ability scores on both the pre- and post-test of early algebra knowledge, Engledowl et al. (2021) found that in addition to statistically significant improvements in students' mean early algebra knowledge, a little more than half of the students improved to a higher level of sophistication in the early algebra knowledge. This improvement is quite a large impact, given the small scale of the intervention. Moreover, analysis drawing on the use of hierarchical linear modeling (analysis in progress by Chris Engledowl, the fifth author of this paper) has found that (a) a substantial proportion of the differences in student post-test Rasch person-ability scores can be explained by differences across teachers (Intra-Class Correlation = 0.352) and (b) after accounting for differences across teachers, students' self-reported gender and race/ethnicity, and the amount of time students spent playing the game, there was a statistically significant improvement in students' early algebra knowledge with a moderate effect size ($d = 0.433$, standardized coefficient = 0.6178, $p < .001$). Time spent playing Agrinautica was found to be statistically significant, but not practically significant ($p < .01$, standardized coefficient = 0.1269), associated with post-test scores. This finding, although perhaps initially shocking, is not all that unexpected. As has been reported elsewhere (e.g., Kolovou et al., 2013), putting in more time playing a game is not necessarily expected to be highly associated with learning outcomes because it is more about engagement in the intervention itself. Moreover, the quality of that engagement tends to have more explanatory power.

Embedded telemetry data were collected on student game play for Agrinautica. This data allowed us to better understand the expressions made by 410 students (Table 2). Students who played very little, or for whom we had incomplete data, are not included. The total group is broken down in two ways: (a) By time of play, measured in total recorded actions taken (an action includes any task completed in the game, such as building a plant, looking at the reference guide, or randomizing the numbers to build from). Some students played for only 20 minutes or so, and some played for hours. (b) By the

Table 2. Expressions made by students

	Percent of students' expressions that were invalid	Percent of students' expressions that were weeds	Percent of students' plant expressions that were repeated	Percent of students' plant expressions that were distinct
All 410 students for whom we have sufficient data	5	19	47	29
204 students who took < 417 total actions	7	22	37	34
206 students who took ≥ 417 total actions	5	18	49	28
181 students who did not gain a level on the post test	6	21	45	28
229 students who gained one or more levels on the post test	5	17	48	30

results of the pre-test and the post-test. The details of the levels, and the overall effectiveness of the intervention, are given in Engledowl et al., 2021.

We take this data as supporting evidence that Agrinautica was successful. Very little time was spent by students entering meaningless expressions. Some time was spent entering weeds, or repetitions, but not an inordinate amount. The percentages do not vary that much over the subgroups, which suggests that it is normal in a sandbox game for students to spend a certain amount of time doing things that may appear unproductive (weeds and repeated plants), while at the same time spending a good amount of time on productive tasks (distinct plants).

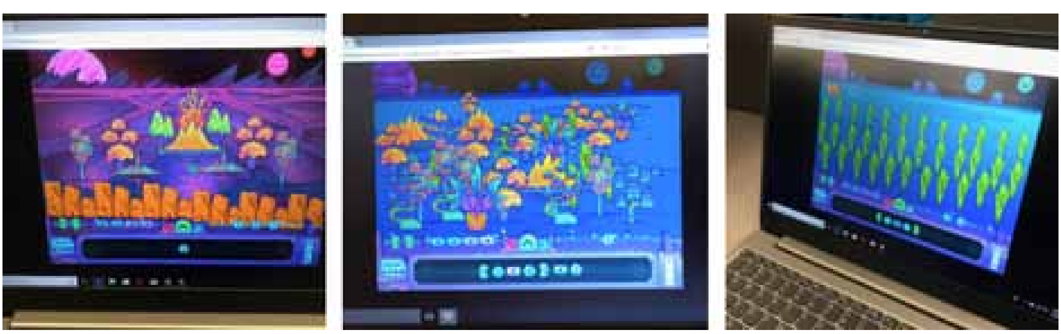
In a more traditional approach to evaluating expressions, students are often given fewer problems because they need to work through each one by hand. In a sandbox environment, by contrast, the students can do many more examples because the computational details can be handled by the game itself. The trade-off is that some of the students' examples are less productive than others. The average student in our data created 180 expressions, which is much larger than the number of expressions they could have done in a traditional "evaluate these expressions" assignment in the same amount of time. However, only roughly 30% of these on average were new expressions that met the target criteria. The research team believes that even the invalid expressions and weeds are valuable as errors to learn from, just as long as they do not dominate the activity.

Teachers

Teachers' interviews, pre-post attitude surveys, and reflection logs shed light on important findings. First, teachers' abilities to facilitate the game are critical to the success of game integration. After teachers completed the intervention and played the games themselves, their perceptions of their own technological and computer-games knowledge greatly improved. Indeed, 93% and 86% of the 28 teachers agreed or strongly agreed to knowing how to use computers and/or tablets in daily instruction and to differentiating instruction.

In addition to playing the games themselves, teachers also let *students explore the games and modeled supporting activitie*s. One of the main takeaways from teachers' reflection logs is that teachers showed great willingness to let go of control and give students some liberty to explore the games on their own or in small groups. Students play Agrinautica in different ways (Figure 9); they explore the game based on their goals, expectations and previous experiences. For example, some students' goals were to "collect" various and new creations; some students aimed to "decorate" and arrange creations in a specific pattern; and some students focused on "creating a garden of weeds." Because teachers were open to the different ways of play, students had the opportunity to explore and make sense of the content. Throughout teachers' reflection logs, expressions such as "they realized," "they found out," "they discovered," and "they figured out" were recurrent. Those expressions reflect teachers' surprise about students' ability to independently explore the games and complex mathematical concepts.

Figure 9. Different ways students play and explore Agrinautica



Math Snacks games also encouraged group work and student-to-student collaboration. In their reflection logs, teachers highlighted that students were willing to help one another by sharing strategies either about the games or about mathematics. Some teachers encouraged students to share and discuss aspects with the whole class, whereas others preferred having students practice mathematical discourse in small groups or in pairs. According to teachers' reflection logs, some students were able to make sense of mathematical complexes with the help of their classmates.

RECOMMENDATIONS

After reflecting on Agrinautica's design process and data assessments, we articulated a set of design recommendations for other sandbox games, particularly in math. The goal was to inform and support discussion by development teams when designing sandbox games for learning.

Design for Discovery

If informed by a constructivist mindset, the game needs to offer learners opportunities to discover the content on their own, being lighter on instructions and heavier on resources or things players can learn for themselves at their pace—for example, providing guides that players can access based on their own desires or needs. Achievements and other motivators can help guide learners to learning opportunities during gameplay, but they should be used to move them toward discovery, not reward them for completion. Feedback on what is expected to be learned is still important, but gameplay should encourage learners to create their personal set of achievements and ways to explore the content through gameplay.

Be Comfortable With Different Outcomes

Gameplay needs to enable different outcomes, encouraging a diverse way of play based on players' variabilities. The team needs to be comfortable with the idea that some players will not try, and others will. This point is best evidenced through the use of weeds by some students. Some players were content to exert little effort and gain a weed, then another, then another. In observations, students would often try this approach, particularly after they saw the gardens of other students. Often, it would take one student exclaiming after looking at another one's screen, "Wow! How'd you get *that* plant?" Students would gather, get the structure of the expression, and then explore ways to build a similar successful expression with the numbers they had received. These types of interactions often motivated the lower-effort weed builders to start exploring. However, we had to be comfortable with students building a lot of weeds to give them time to establish their own goals. Similarly, we had to accept that some students would be motivated by achievements, and others wouldn't. One positive aspect of this approach is that students building in a group setting served as their own motivators: The diversity of students' screens prompted players to make their own customized gardens.

Support the Game-Based Learning With Out-Of-Game Experiences

A sandbox game needs to provide meaningful learning opportunities. A good sandbox game represents an opportunity to enrich classroom activities and provide experiences that can be profitably built on. Although some games by themselves can foster good learning experiences, this experience can always be enriched with reflection, application to real-life situations, and expansion to other areas through guided classroom activities. If the game is built to be used in a classroom setting, guidance needs to be given to teachers or instructors regarding how to play the game, how to incorporate the game in the classroom, how to use the game to enrich the learning experience, and additionally, how to understand the different outcomes and different ways of playing a sandbox game. For example, effective support would guide teachers in how they can find ways to help players share and collaborate, either in the game or in class time around the game. This approach is especially true of sandboxes, where each player can have a slightly different experience, and then everyone can learn by comparing those experiences.

Consider What Learning Goals Work Best for Sandbox Environments

During the design phase, the team needs to articulate ways the content and learning goals intended for the project fit with the sandbox gameplay (an open world to explore with few gameplay instructions). Some content and learning goals can benefit extensively from the sandbox features; for example, in Agrinautica, the learning goal was to allow players to explore and practice with many different kinds of expressions and operators, making the sandbox gameplay an excellent choice. On the other hand, content or activities that demand more guidance and structure are less likely to benefit from the openness of a sandbox game. A game that simply demands memorization or practice may be best achieved with other types of motivators.

Design Learning Activities That Support Constructivist Learning

Design game activities to support students' thinking and teachers' reflection. Learning goals can be more than just about whether students learn some pre-chosen skill or concept; they can support students to better understand their thinking, and they also can inform teachers about students' thinking, background, and interests. For example, if an activity instructs students to evaluate some expressions, such as $5 \times (2+1) - 6$ and get the correct answer of 9, then there is not much to explore; the only thing teachers are going to learn from that is whether students can comply with the activity instruction correctly or not. From a constructivist learning standpoint, this is not very helpful or engaging for the teacher or the student. On the other hand, if the activity asks for a more open-ended response (e.g., "what kinds of expressions can you make?"), then students will be able to creatively think about and find solutions for the problem, while teachers will potentially learn all kinds of interesting things about how students think, what they understand, what misconceptions they have, and what motivates them.

CONCLUSION

Game-based learning experiences including mathematics content have significantly increased over the years. The use of these games is followed by a global effort to research the impact of these digital media on students in formal and informal learning settings. In this study, we discussed formative and summative data from the design and development of the early algebra Math Snacks game Agrinautica.

Data from students showed that the Agrinautica intervention was effective in improving students' early algebra knowledge. Additionally, they enjoyed the gameplay experience. Additional analysis is currently underway to examine embedded data from Agrinautica to understand (a) patterns in the different kinds of expressions that students wrote, and how that may reveal important differences in the quality of their engagement in the early algebra concepts, and (b) patterns in students' style of play that may be associated with affect and cognitive learning outcomes.

Data from teachers using Agrinautica showed that the pedagogical strategy of using games to practice mathematics concepts has been widely used. However, the implementation of Math Snacks introduced different ways of using computer games, as a way to explore concepts and to launch a lesson for a new mathematics concept. Teachers started to see computer-based games as an integral part of their future teaching practice, rather than something they "allowed" students to do outside the structure of the lesson. Moreover, teachers experienced how letting students engage in self-exploration gameplay time led to a rich inquiry, discussion-based learning environment where all participants learned from each other. Implications for research in game-based learning are that the teachers and other gatekeepers are valuable sources of data on the impact of games.

Open-world gameplay, where players can establish their own goals and explore with just enough guidance, can support early algebra learning. The constructivist approach used to foster learning by discovery, giving learners several opportunities to explore on their own, was also beneficial to players' learning. Grounded in practical and theoretical knowledge, articulations of how games work for students and teachers can help inform the design of future sandbox games for learning.

This study is part of a robust effort, but it still has limitations. Game-based learning interventions are powerful, but complex artifacts that involve various stakeholders to design and evaluate impact. This study brought together a wide range of data sources to support the impact of the game, but future studies still need to investigate the findings further. Future studies might involve a larger sample size and additional schools, and could include parents' perspectives and role in games' math-learning impact.

ACKNOWLEDGMENT

We wish to thank members of all the design teams of the Math Snacks project who developed engaging and effective learning games, including deep appreciation for project lead investigators, Karin Wiburg, Karen Trujillo, and Wanda Bulger-Tamez. We also acknowledge the hundreds of students who played the games and offered feedback to improve it, the research participants who consented to their data being used to assess their learning, and advisory board members for constant recommendations and support. We also thank Amy Smith Muise for improving the written quality of this paper. Research approved and consent received through IRB number 22576.

CONFLICT OF INTEREST

We declare there is no conflict of interest.

FUNDING AGENCY

Math Snacks materials were developed with support from the National Science Foundation (0918794 and 1503507). Any opinions, findings, and conclusions or recommendations expressed in this material are our own and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Adams, E. (2010). *Fundamentals of game design* (2nd ed.). New Riders.

Blanton, M., Stephens, A., Knuth, E., Gardiner, A. M., Isler, I., & Kim, J.-S. (2015). The development of children's algebraic thinking: The impact of a comprehensive early algebra intervention in third grade. *Journal for Research in Mathematics Education*, *46*(1), 39–87. doi:10.5951/jresematheduc.46.1.0039

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge., doi:10.4324/9781315814698

Breslin, S. (2009, July 16). *The history and theory of sandbox gameplay [web log].* Gama Sutra. http://www.gamasutra.com/view/feature/4081/the_history_and_theory_of_sandbox_. Php

Byun, J., & Joung, E. (2018). Digital game-based learning for K-12 mathematics education: A meta-analysis. *School Science and Mathematics, 118*(3–4), 113–126. https://doi.org/10.1111/ssm.12271

Cennamo, K. S. (2003). Design as knowledge construction: Constructing knowledge of design. *Computers in the Schools*, *20*(4), 13–35. doi:10.1300/J025v20n04_03

Chamberlin, B., & Schell, J. (2018). *The secret process for making games that matter* [Paper presentation]. Connected Learning Summit. MIT Media Lab.

Dormans, J. (2012). *Engineering emergence: Applied theory for game design* [Doctoral dissertation, University of Amsterdam].

Egenfeldt-Nielsen, S. (2010). The challenges to diffusion of educational computer games. *Leading Issues in Games Based Learning*, *141*, 145–158.

Ellison, M., & Drew, C. (2020). Using digital sandbox gaming to improve creativity within boys' writing. *Journal of Research in Childhood Education*, *34*(2), 277–287. doi:10.1080/02568543.2019.1675823

Engledowl, C. (2020). Constructing and validating an early algebra assessment. In *Proceedings of the 47th Annual Meeting of the Research Council on Mathematics Learning* (pp. 51–58). https://www.rcml-math.org/assets/Proceedings/RCML%202020%20Proceedings%202-19-20.pdf#page=57

Engledowl, C., Al-younes, M., Chamberlin, B., & Stanford, T. (2021). Learning acquired from a computer game-based early algebra intervention. *School Science and Mathematics*, *121*(8), 495–508. doi:10.1111/ssm.12500

Ernest, P. (1994). Social constructivism and the psychology of mathematics education. In P. Ernest (Ed.), *Constructing mathematical knowledge: Epistemology and mathematics education* (pp. 62–72). Routledge.

Ernest, P. (1996). Varieties of constructivism: A framework for comparison. In L. P. Steffe, P. Nesher, P. Cobb, G. A. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 335–350). Lawrence Erlbaum Associates, Inc. https://books.google.com/books?id=x2XRiIm-3vAC&pg=RA5-PT200&dq=theories+of+mathematical+learning+steffe&hl=en&sa=X&ei=8yg1VcrtLse4ggSQnoHYBA&ved=0CB4Q6AEwAA#v=onepage&q=theories%20of%20mathematical%20learning%20steffe&f=false

Gray, L., & Lewis, L. (2021). Use of educational technology for instruction in public schools: 2019–20 (NCES 2021-017). *National Center for Education Statistics. Institute of Education Sciences*. U.S. Department of Education. https://nces.ed.gov/pubs2021/2021017.pdf

Hébert, C., & Jenson, J. (2020). Teaching with sandbox games: Minecraft, game-based learning, and 21st century competencies. *Canadian Journal of Learning and Technology /La revue canadienne de l'apprentissage et de la technologie, 46*(3). 10.21432/cjlt27990

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.

Kolovou, A., van den Heuvel-Panhuizen, M., & Köller, O. (2013). An intervention including an online game to improve grade 6 students' performance in early algebra. *Journal for Research in Mathematics Education*, *44*(3), 510–549. doi:10.5951/jresematheduc.44.3.0510

Mayer, R. E. (2014). *Computer games for learning: An evidence-based approach*. The MIT Press., doi:10.7551/mitpress/9427.001.0001

Mercer, N., & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning, Culture and Social Interaction*, *1*(1), 12–21. doi:10.1016/j.lcsi.2012.03.001

Microsoft. (2022). *Minecraft Education Edition*. Microsoft. https://education.minecraft.net/en-us

Morales, S., Trujillo, K., Mucundanyi, G., & Torres Castillo, R. (2019). Math Snacks early algebra: Successes and challenges of a research design for computer-based games in mathematics classrooms. In K. Graziano (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference* (pp. 2185–2191). Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/primary/p/207951/

Moses, R. P., & Cobb, C. E. (2001). Algebra and civil rights? In *Radical equations: Math literacy and civil rights* (pp. 3–22). Beacon Press.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. NCTM.

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions: Ensuring mathematical success for all*. NCTM.

National Governors Association Center for Best Practices (NGACBP) and Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. CCSSO. https://CCSSO.org

Plass, J. L., Homer, B. D., Mayer, R. E., & Kinzer, C. K. (2020a). Theoretical foundations of game-based and playful learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 3–24). The MIT Press.

Plass, J. L., Mayer, R. E., & Homer, B. D. (Eds.). (2020b). *Handbook of game-based learning*. The MIT Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage Publications.

Rideout, V. (2014). *Learning at home: Families' educational media use in America*. Joan Ganz Cooney Center at Sesame Workshop. https://eric.ed.gov/?id=ED555586

Russo, J., Bragg, L. A., & Russo, T. (2021). How primary teachers use games to support their teaching of mathematics. *International Electronic Journal of Elementary Education*, *13*(4), 407–419. doi:10.26822/iejee.2021.200

Smith, M. S., & Stein, M. K. (2018). *5 practices for orchestrating productive mathematics discussions* (2nd ed.). Corwin.

Squire, K. (2005). Changing the game: What happens when video games enter the classroom? *Innovate: Journal of Online Education, 1*(6). https://nsuworks.nova.edu/innovate/vol1/iss6/5

Stephens, A. C., Ellis, A. B., Blanton, M., & Brizuela, B. M. (2017). Algebraic thinking in the elementary and middle grades. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 386–420). National Council of Teachers of Mathematics, Inc.

ter Vrugte, J., de Jong, T., Vandercruysse, S., Wouters, P., van Oostendorp, H., & Elen, J. (2017). Computer game-based mathematics education: Embedded faded worked examples facilitate knowledge acquisition. *Learning and Instruction*, *50*, 44–53. doi:10.1016/j.learninstruc.2016.11.007

Torres, R., Toups Dugas, P. O., Wiburg, K., Chamberlin, B., Gomez, C., & Ozer, M. A. (2016). Initial design implications for early algebra games. In *CHI PLAY Companion '16: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (pp. 325–333). Association for Computing Machinery., https://dl.acm.org/doi/abs/10.1145/2968120.2987748 doi:10.1145/2968120.2987748

Torres Castillo, R., & Morales, S. (2019). A protocol for reviewing off-the-shelf games to inform the development of new educational games. In K. Blashki & P. Isaías (Eds.), *Handbook of research on human-computer interfaces and new modes of interactivity* (pp. 40–58). IGI Global. doi:10.4018/978-1-5225-9069-9.ch003

Torres Castillo, R. C. (2020). *Teaching mathematics through game-play: Elementary teachers' pedagogical practices and perceptions of integrating computer games in the classroom (27955684)* [Doctoral dissertation, New Mexico State University]. https://www.proquest.com/openview/b50534257b2ebf2fb2ac480642f0b0b6/1?pq-origsite=gscholar&cbl=18750&diss=y

Valdez, A. (2012). Assessing mathematics classroom instruction: Observation of learning environments (OLE). *Psychology Journal*, *9*(1), 35–43.

van den Heuvel-Panhuizen, M., Kolovou, A., & Robitzsch, A. (2013). Primary school students' strategies in early algebra problem solving supported by an online game. *Educational Studies in Mathematics*, *84*(3), 281–307. doi:10.1007/s10649-013-9483-5

Wiburg, K., Chamberlin, B., Valdez, A., Trujillo, K., & Stanford, T. B. (2016). Impact of Math Snacks games on students' conceptual understanding. *Journal of Computers in Mathematics and Science Teaching*, *35*(2), 173–193.

Matheus Cezarotto is an assistant professor and extension educational technology specialist at New Mexico State University's Innovative Media Research and Extension Department. Dr. Cezarotto coordinates the Learning Games Lab, providing inclusive design expertise for the team's educational technology products, including learning games, animations, and virtual labs. He also researches how these affordances support learning. His research focuses on educational media design, investigating and developing effective research-based learning technology tools.

Pamela N. Martinez is an assistant professor and learning technologies specialist at New Mexico State University's Innovative Media Research and Extension Department and its Learning Games Lab. Dr. Martinez oversees product development and assessment for grant-funded research projects. She has more than 25 years' experience creating educational media and games and research strengths in diversity, equity, and inclusion, making sure all audiences are represented in developed media and have access to appropriate learning tools.

Ruth Torres Castillo earned her Ph.D. with a concentration in computer science and curriculum and instruction at New Mexico State University (NMSU). Dr. Torres Castillo does research in science, technology, engineering, and mathematics (STEM) and computer science education and diversity in computing, computational thinking (CT) practices, computer games as part of the learning process in the science, technology, engineering, arts, and mathematics (STEAM) field, human-computer interaction, mathematics, and educational technology. She was recently working as a post-doctoral researcher investigating the implementation of CT practices into performance and creative writing courses throughout an NSF-funded project called DEPICT (Discovering Computational Thinking Through Creative Writing). She also worked on the NSF-funded Math Snacks (mathsnacks.org) series of games supporting teachers as the main facilitators for games in the classroom.

Theodore Stanford is an Associate Professor (now retired) in the Department of Mathematical Sciences at New Mexico State University. His research interests include low-dimensional topology, mathematical learning games, and professional development for K-12 teachers.

Christopher Engledowl is a research scientist at the University of Florida Lastinger Center for Learning. His research focuses on the impact and effectiveness of innovative digital products for teacher and student learning broadly, but with a special interest in mathematics, statistics, and data science education. He draws on his expertise in quantitative research methods, user experience research, and data science to iteratively improve the center's products in ways that have potential to transform public education.

Germain Degardin first moved to New Mexico in 2011 to start a bachelor's in economics at NMSU. After earning his bachelor's degree, Degardin remained at NMSU to receive a master's in business, a master's in education, and a teaching license in secondary education. He developed his expertise as a research evaluator in education through his six years at the SOAR Evaluation and Policy Center, where he evaluated the implementation, quality, and impact of statewide and local outreach education programs. He currently works for Project ECHO as a science research manager, and his primary focus is on analyzing and reporting programs feedback as well as creating data visualization for program managers.

Barbara Chamberlin has been developing educational games, media, and interactive programs with New Mexico State University Extension for almost 30 years. She leads instructional design of digital tools in the department, working with the amazing programmers, artists, and designers on the team and collaborating with content experts at universities throughout the United States. Her areas of research include user testing and the processes behind developing educational media. She oversees research and development in the Learning Games Lab and developed the game models used in design at the lab.