STUDENT ENGAGEMENT, UNDERSTANDING, AND STEM INTEREST IN A GAME BASED SUPPLEMENTAL FRACTION CURRICULUM

Jessica H. Hunt
North Carolina State University
Jhunt5@ncsu.edu

Michelle Taub
University of Central Florida
Michelle.taub@ucf.edu

Matthew Marino
University of Central Florida
Matthew.marino@ucf.edu

Kenneth Holman
University of Central Florida
Kenneth.homan@ucf.edu

Alejandra Duarte
North Carolina State University
Aduarte2@ncsu.edu

Brianna Bentley
North Carolina State University
blbentle@ncsu.edu

We analyzed the effects of a game-based, supplemental fraction curriculum on fourth and fifth grade students' fraction knowledge, engagement, and STEM interest. Students with and without disabilities with intersecting identities (e.g., race, disability status, gender) comprised the sample. Results indicate significant differences in fraction concept knowledge as a result of the curriculum for all students, but not STEM interest. Furthermore, engagement was a significant predictor of STEM post test scores, but not fraction concept post test scores. Implications of the results in the context of previous research on game-based mathematics curriculums are shared.

Keywords: engagement, fractions, disability, curriculum

The study reported here addresses a game-based supplemental fraction curriculum and aligns with two central questions of the PMENA conference: (a) What design features for tools and curricula consider supporting engagement, interest, and learning for all students? (b) How might learning environments that take all students into account impact engagement, interest, and learning?

Literature Review

Engagement is often described as "participation in activity with some cognitive or affective investment" with an "inseparability of learning from the engagement through which learning takes place" (Middleton et al., 2017, p. 668). Most historical and contemporary research quantitatively defines engagement as *traits* that students evidence at set points in time or discusses engagement as sets of behaviors teachers can model for students (Klem & Connell, 2009; McLeskey et al., 2017). Yet, students' engagement can also be observed or documented during the processes of learning, or as *states*. We posit that research that examines engagement as *states* can yield important contributions to the literature and can be of particular benefit to diverse student populations for which research on engagement is sorely missing, such as students with disabilities (McKlesky et al., 2017). For example, researchers can investigate if these students' engagement differs across contexts and how different levels of engagement coincide with outcomes, such as learning and STEM interest.

Digital game-based mathematics curricula have gained increased prevalence as a means to improve students' engagement, STEM interest, and learning outcomes over the past two decades (Sies, 2018). For example, Lin et al. (2013) found digital games improved students' learning via problem-solving tasks within a game. Additionally, summaries of gaming research have identified the potential of games to enhance STEM content accessibility and interest, increase opportunities to learn via solving problems, and allow students to explore mathematics in ways that were previously inconceivable, especially for students with disabilities and other marginalized populations (Marino et al, 2013).

Despite the tremendous potential of games, there is a pressing need for rigorous empirical research in terms of the effects of game-based curriculums, how games affect student engagement and STEM interest, and the extent to which engagement, understanding of math content, and interest are related. For example, Byun and Joung (2018), in a meta-analysis of 296 articles on gaming in mathematics education, found only 17 studies with sufficient statistical data to support effect size calculations. They noted many of the articles included drill and practice games and reported performance in the absence of student engagement and interest. Byun and Joung (2018) also identified a low percentage of the study authors (i.e., 7%) with a background in mathematics education.

Enhancing Engagement, Learning, and Interest with Universal Design for Learning

Universal Design for Learning is a framework for the design and implementation of efficacious instructional materials. The framework is organized around nine guidelines and 31 checkpoints, organized vertically to proactively designing for learner variability through: (a) multiple means of engagement (i.e., considering how to engage students in multiple ways), (b) multiple means of representation (i.e., providing content in multiple formats), and (c) multiple means of action and expression (i.e., providing opportunities for students to demonstrate their understanding in multiple ways). King-Sears (2020) pointed out UDL-based interventions must include a flexible, purposeful design in order to engage a maximum number of learners. Marino and Basham (2013) reported that the proactive identification of barriers across physical, social/emotional, cultural, and cognitive aspects of the lesson are critical during STEM lessons.

Instruction should be intentionally planned so that it is personally challenging for all learners. When planning for learner variability, curriculum materials should consider specific considerations such as individual and group strengths, abilities, background knowledge, and motivation for participating in the learning activity. The implementation of UDL within the game-based program focused on integrating the three principles above across four instructional domains: 1) clear goals, 2) intentional planning for learner variability, 3) flexible methods and materials, and 4) timely progress monitoring. The goal of the game is to support student engagement and understanding around the idea that fractions are quantities with magnitudes determined by the multiplicative coordination of the numerator and the denominator.

We also designed the Dream2B interface intentionally for variability so that players can access each challenge in multiple ways (see Hunt et al., 2020). For example, an interactive learning environment motivates players by allowing them to customize the game based on their preferences. The player has a choice of flexible methods, materials, and analytical tools that they can use to employ individual strategies and ways of reasoning. Sandbox play supports players to create fractional quantities by partitioning, repeating, distributing, and coordinating units (Wilkins & Norton, 2018) without high stakes repercussions. These and other UDL features have been empirically shown to increase engagement and eliminate differences in performance between students with and without disabilities in middle school science classrooms (Marino et al., 2013). However, there is a pressing need to determine if the same results will occur in mathematics as well as how engagement, understanding, and interest are related within the context of the game-based program.

Research Questions

This study addresses the following research questions: (1) What is the change from pre- and post-test scores of conceptual understanding of fractions before and after engaging in the game-based curriculum? Does this relationship differ between student demographics? (2) What is the change from pre- and post-test scores of STEM interest fractions before and after playing Dream2B? Does this relationship differ between student demographics? (3) What is the relationship between pre-test and post-test conceptual understanding of fractions, pre-test and post-test STEM interest and student self-reported engagement across all game worlds, and (4) What is the association between pre-test scores, post-test scores, and student self-reported engagement?

Methods

The current study examines a five-unit, 36 lesson fraction curriculum with a video game embedded within it. Developed using the UDL framework, the program is designed to maximize accessibility and engagement by providing fraction conceptual understanding challenges rooted in authentic STEM careers. Each curriculum lesson has three parts: Before game previews, video game, and after-gameplay discussion activities. All components strategically link to mathematics curriculum standards (NCTM).

Participants and Setting

Program testing occurred with six 4^{th} and 5^{th} grade teachers and their students (n = 132) in two schools in the southeastern United States. Both schools included students with intersecting identities in terms of race, language, and neurodiversity. Twenty-one students were identified as having learning disabilities. The program was delivered in the mathematics classroom, which commonly includes 15-25 students and one teacher. Teachers and students engaged with the program over a period of nine weeks, which is considered best practice for technology-based programs (Gersten & Edyburn, 2007). Demographic information for students is given in Table 1.

Table 1: Student Demographics

Gender		Race				Disability*	
Female	Male	Hispanic	African American	White	Two or More	Yes	No
41%	59%	28%	21%	32%	19%	16%	84%

Teacher Development

Teachers attended four one-half day training sessions on implementation of the curriculum. Day 1 of training opened with the purpose of the study, the logic model, and the target population. Over the second and third day of training, teachers used the game and sample student gameplay to deepen their understanding of how the core components are used to bolster student learning. On the final day, a curriculum guide was given to teachers to drive delivery of the intervention, and teachers practiced using the resource through role playing in small groups, rotating between teaching roles and student roles. Teachers then engaged in the after-game tasks, discourse, and talk moves to facilitate a sample student conversation.

Measures

Three forms of data were gathered to assess the research questions. First, the *Engagement in Science Learning Activities* survey (Chung et al., 2016) was used to measure student cognitive, behavioral, and affective engagement across the curriculum. Participants respond on a Likert- type scale ranging from 1 (YES!) to 4 (NO!). Both Cronbach's α and the polychoric coefficients yielded acceptable reliability when using all eight scale items (0.80 and 0.85, respectively).

Second, fraction knowledge was measured before and after students engaged in the program using the *Test of Fraction Schemes* (Wilkins et al., 2013). Internal consistency reliability for the test was reported as 0.70; criterion-related validity was reported as 0.58 ($p < 0.01$).

Third, we used the *Upper Elementary School* (4-5) *Student Attitudes Toward STEM (S- STEM)* Survey (Friday Institute for Educational Innovation, 2012) to measure changes in students' self-reported STEM interests. The S-STEM was developed as part of a National Science Foundation (NSF) funded research program and measures students' confidence and self- efficacy in STEM subjects, 21st century learning skills, and interests in STEM careers. It contains 56 items across six constructs: math attitudes (8 items), science attitudes (9 items), engineering and technology attitudes (9 items), 21st century learning attitudes (11 items), interest in STEM career areas (12 items), and 7 "About You" items that measure short-term expectations for course success and exposure to STEM careers. Responses are supported by a five-point Likert scale, with response options ranging from "strongly disagree" (1) to "strongly agree" (5). Higher scores reflect the greater perceived value of participants. Cronbach's α of the S-STEM ranged from 0.84 to 0.86 for the grade 4-5 subscales and 0.89 to 0.91 for the middle high school subscales, respectively).

Data coding and scoring. We coded and scored two types of test scores to address our research questions. First, pre- and post-tests of conceptual understanding of fractions were scored out of the total number of questions to create a ratio score (i.e., ¿ *correct responses* ÷ 12[*total* ¿ *of questions*]).

We also coded and scored other variables that would serve as predictors and dependent variables for our models. This includes variables related to STEM interest, demographics, and engagement. STEM interest was scored by summing all components of the *S-STEM* questionnaire. Demographic information included: gender (dichotomized score of 1: Male or 2: Female), race (scores of 1: Hispanic, 2: White, 3: African American, or 4: Two or More Races), and disability status (dichotomized score of 1: Disability or 2: No disability). To code engagement, we created a composite score from the *Engagement in Science Learning Activities* survey combining the score from all 5 worlds.

Data Analysis

To determine the change from pre- and post-test scores of conceptual understanding of fractions and STEM interest before and after engaging in the game-based curriculum, we ran two ANOVAs with repeated measures using time as the repeated, within-subjects factor (i.e., fraction concept understanding score; STEM interest score) and student demographics as the between-subjects factors. Significance was set at 0.05 for both ANOVA procedures.

To determine association between pre-test scores, post-test scores, and student self-reported engagement, we first ran a bivariate correlation including all pre-test and post-test items to determine the relationship between these variables. Next, we ran multiple linear regressions, with post-test fraction and STEM interest scores as dependent variables and pre and post-test STEM interest, post-test fraction score, and self-reported engagement as our predictor variables. Significance was set at 0.05 for all regressions.

Results

To address our research questions, we ran ANOVAs with repeated measures and multiple linear regressions to examine the relationships between pre-test scores, post-test scores, student reported engagement, and student reported demographics.

Research Question 1: What is the change from pre- and post-test scores of conceptual understandings of fractions before and after playing Dream2B? Does this relationship differ between student demographics? For this question, we ran an ANOVA with repeated measures using time as the repeated, within-subjects factor and student demographics as the between-subjects factors.

Results revealed a significant, medium to large within-subjects effect for test time (eta squared = .086). Moreover, there were no significant interaction effects between time with gender, race, or disability status ($p > .05$). This indicates that regardless of gender, ethnicity, or disability status, all students demonstrated a significant increase in score from pre-test to post- test. Levene's Test was significant for both pre- and post-test (ps < .05), so we do not report the F statistic. However, results of a non-parametric Friedman test of differences among repeated measures rendered a Chi-square value of 24.43, which was also significant ($p < .001$). Results are demonstrated in Figure 1.

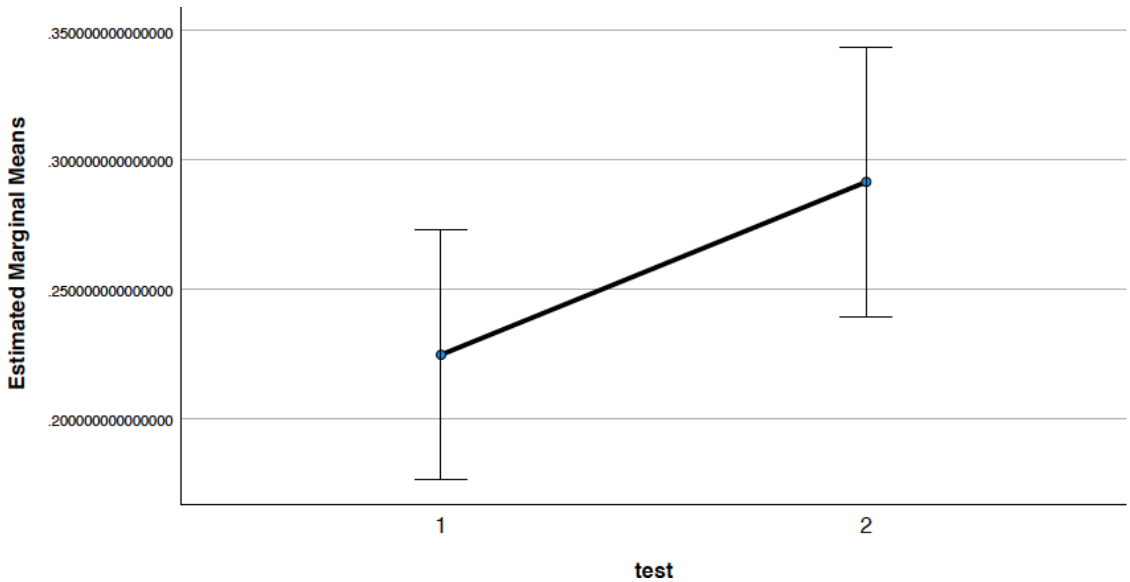

Figure 1: Main effect of time for fraction test score

Research Question 2: What is the change from pre- and post-test scores of STEM interest fractions before and after playing Dream2B? Does this relationship differ between student demographics? For this question, we ran an ANOVA with repeated measures using time as the repeated, within-subjects factor and student demographics as the between-subjects factors.

Results did not reveal a small significant within-subjects effect for test time, with eta squared = .025. However, there were significant interaction effects between time*gender, time*disability status, time*gender*ethnicity, and time*ethnicity*disability status. This indicates the relationship between pre-test and post-test STEM scores were not significantly different while ignoring the effects of gender, ethnicity, and disability status. It is important to note that Levene's Test was significant for STEM pre-test ($p < .05$), so we do not report the F statistics for the overall tests or their interactions. In addition, results of a non-parametric Friedman test of differences among repeated measures rendered a Chi-square value of .398, which was not significant ($p > .05$).

Research Question 3: What is the relationship between pre-test and post-test conceptual understanding of fractions, pre-test and post-test STEM interest, and student self-reported engagement across all game world levels? To address this research question, we ran a bivariate correlation including all pre-test and post-test items to determine the relationship between these variables. We report Pearson coefficients in Table 2, below.

Table 2: Pearson Correlations between study pre and post-test items

	1	2	3	4	5
1-Pre-test fraction	-	.739***	.204*	.197*	-.095
2-Post-test fraction		-	.208*	.261**	-.131
3-Pre-test STEM			-	.360***	.111
4-Post-test STEM				-	.163
5-Engagement					-

****p*<.001, ***p*<.01. **p*<.05.

Research Question 4: What is the association between pre-test scores, post-test scores, and student self-reported engagement? To address this research question, we broke the question into several parts, as described below.

Do pre-test fraction conceptual understanding and student self-reported engagement predict post-test STEM interest? We ran a multiple linear regression with post-test STEM interest as our dependent variable and pre-test fractions score and self-reported engagement as our predictor variables. Results revealed a significant model; $R^2 = .07$, $F(2,121) = 4.54$, $p < .05$. It was found that pre-test fraction score (□ = .21, $p < .05$) and reported engagement (□ = .18, $p < .05$) significantly predicted post-test STEM interest. Therefore, as pre-test fraction score and self- reported engagement score increased, so did student post-test STEM interest.

Do pre-test STEM interest and student self-reported engagement predict post-test fraction conceptual understanding? We ran a multiple linear regression with post-test fraction score as our dependent variable and pre-test STEM interest and self-reported engagement as our predictor variables. Results revealed a significant model; $R^2 = .067$, $F(2,129) = 4.64$, $p < .05$. It was found that pre-test STEM interest significantly predicted post-test fraction score (□ = .23, $p < .05$), however reported engagement (□ = -.16, $p > .05$) did not significantly post-test fraction score. Therefore, as pre-test STEM interest increased, so did student post-test fraction score.

Do post-test fraction conceptual understanding and student self-reported engagement predict post-test STEM interest? We ran a multiple linear regression with post-test STEM interest as our dependent variable and post-test fractions score and self-reported engagement as our predictor variables. Results revealed a significant model; $R^2 = .11$, $F(2,121) = 7.13$, $p < .01$. It was found that post-test fraction score (□ = .28, $p < .01$) and reported engagement (□ = .20, $p < .05$) significantly predicted post-test STEM interest. Therefore, as post-test fraction score and self- reported engagement score increased, so did student post-test STEM interest.

Do post-test STEM interest and student self-reported engagement predict post-test fraction conceptual understanding? We ran a multiple linear regression with post-test fraction score as our dependent variable and post-test STEM interest and self-reported engagement as our predictor variables. Results revealed a significant model; $R^2 = .092$, $F(2,121) = 6.15$, $p < .01$. It was found that pre-test STEM interest significantly predicted post-test fraction score (□ = .29, $p < .01$), however reported engagement (□ = -.16, $p > .05$) did not significantly post-test fraction score. Therefore, as post-test STEM interest score increased, so did student post-test fraction score.

Discussion and Conclusion

Overall, results reveal a significant effect of the supplemental, game-based fraction program on students' conceptual knowledge. On the other hand, the effects of the program on students' STEM interest did not produce significant effects. Results also reveal that there is a relationship between all variables, demonstrating the relationship between student learning and performance, STEM interest, and engagement in the mathematics classroom. In addition, all significant predictors were positive relationships, such that an increase in the predictors (fraction test scores, STEM interest scores, and reported student engagement) were associated with an increase in the dependent variable. However, it is worth noting that engagement scores significantly predicted STEM interest posttest scores yet not fraction concepts post-test scores, indicating an increase in engagement was associated with higher STEM interest yet not necessarily higher fraction test- scores at post-test.

The study has limitations that need to be acknowledged. First, the engagement data (and the data on students' STEM interest) were gained via students' self-report. For engagement, the measure we used breaks down cognitive, emotional, or behavioral forms, yet we did separate out these differing forms of engagement in our analyses. Furthermore, because we view engagement as a state (as opposed to a trait), gameplay data could yield valuable information as to how students engaged with the game, however we did not have access to these data for this study.

Another point can be made about the STEM interest measure. Quantitatively, our results did not show significant changes, yet qualitative data (reported in Hunt et al., under review) did. Finally, students' conceptual knowledge, while found significant in this study, was measured via a distal measure. Future work should address these limitations. For example, future work should explore students' gameplay patterns as evidence of both engagement and conceptual growth and/or other measures that are better reflective of differencing forms of engagement and of the curriculum (e.g., curriculum-based measure). Future work might also report multiple forms of data through mixed-methods approaches that can incorporate a nested students-within-teachers structure and compare and/or merge different forms of data that address the same research questions together to gain more robust accounts of program effects on student outcomes.

Finally, because the results reported here were obtained during a feasibility study of the curriculum that did not use a control group, more work needs to be done with larger samples that encompasses a more robust design that can address the limitations of the current work.

Acknowledgments

This work is supported by the National Science Foundation, grant #1949122. Any Opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation.

References

Byun, J. & Joung, E. (2018). Digital game learning for K-12 mathematics education: A meta-analysis. School Science and Mathematics, 118, 113-126.

Chung, J., Cannady, M. A., Schunn, C., Dorph, R., & Bathgate, M., (2016) Measures Technical Brief: Engagement in Science Learning Activities. Retrieved from: http://www.activationlab.org/wpcontent/uploads/2016/02/Engagement-Report-3.1-20160331.pdf.

Friday Institute for Educational Innovation (2012). Student Attitudes toward STEM Survey-Upper Elementary School Students, Raleigh, NC: Author.

Hunt, J. H., Martin, K., Khounmeuang, A., Silva, J., Patterson, B., & Welch-Ptak, J. (2020). Design, development, and initial testing of asset-based intervention grounded in trajectories of student fraction learning. Learning Disability Quarterly, 0731948720963589.

Hunt, J.H. Taub, M., Marino, M., Duarte, A., Bentley, B., Holman, K.., & Kuhlman, A. (under review). Effects of game-handed supplemental fraction curriculum on student engagement, fraction knowledge, and STEM interest. Education Sciences.

Gersten, R. & Edyburn, D. (2007). Defining quality indicators in special education technology research. Journal of Special Education Technology, Retrieved from https://doi.org/10.1177/016264340702200302

Ke, F., & Abras, T. (2013). Games for engaged learning of middle school children with special learning needs. British Journal of Educational Technology, 44(2), 225–242.
King-Sears, M. E. (2020). Introduction to special series on universal design for learning. Remedial and Special Education, 41(4), 191-193.
Lin, C. H., Liu, E. Z., Chen, Y. L., Liou, P. Y., Chang, M., & Wu, C. H. et al. (2013). Game-based remedial instruction in mastery learning for upper-primary school students. Educational Technology and Society, 16(2), 271-281.
Marino, M. T., Israel, M., Beecher, C. C., & Basham, J. D. (2013). Students and teachers perceptions of using videogames to enhance science instruction. Journal of Science Education and Technology. 22, 667-680.
Marino, M. T., Becht, K., Vasquez III, E., Gallup, J., Basham, J. D., & Gallegos, B. (2014). Enhancing secondary science content accessibility with video games. Teaching Exceptional Children, 47(1), 27-34. doi: 10.1177/0040059914542762.
McLeskey, J., Barringer, M-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., Lewis, T., Maheady, L., Rodriguez, J., Scheeler, M. C., Winn, J., & Ziegler, D. (2017, January). High-leverage practices in special education. Arlington, VA: Council for Exceptional Children & CEEDAR Center.
Middleton, J. A., Jansen, A., & Goldin, G. A. (2017). The complexities of mathematical engagement: Motivation, affect, and social interactions. Compendium for Research in Mathematics Education, 667-699.
Wilkins, J. L. M., & Norton, A. (2018). Learning progression toward a measurement concept of fractions. International Journal of STEM Education, 5(27). https://doi.org/10.1186/s40594-018-0119-2
Wilkins, J. L. M., Norton, A. & Boyce, S. J. (2013). Validating a written instrument for assessing students fractions schemes and operations. The Mathematics Educator, 22(2), 31-54.