Using a New Version of Minecraft to Promote Computational Thinking Learning Opportunities: Investigating Middle Grades Student Outcomes

Elizabeth L. Adams[1]

Paul David Foster[2]

Lawrence J. Klinkert[2]

Emma Goff[2]

Ching-Yu Tseng[2]

Yanjun Pan[2]

Leanne R. Ketterlin-Geller[2]

Eric C. Larson[2]

Corey Clark[2]

Author Note

Paper presented at the 2023 American Educational Research Association Annual Conference, Chicago, IL

Correspondences concerning this article should be addressed to Elizabeth Adams eadams@air.org or Leanne R. Ketterlin-Geller lkgeller@smu.edu

This study was supported by the National Science Foundation under NSF Award Number 1933848. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation

---

[1] American Institutes for Research

[2] Southern Methodist University, Dallas, Texas

Abstract

An interdisciplinary team piloted a prototype computer-game-based learning environment to teach computational thinking to 31 middle school students attending a summer camp. The learning environment included an extension to Minecraft, a custom tool called Minecraft Factory Planner, and the Canvas learning management system. The content focused on three Computer Science Teachers Association data analysis standards. Students played a tutorial and three lessons at the rate of one lesson per day for 40 minutes per day. Changes in students' self-efficacy and perceived usability were statistically significant. Additional lessons are being developed with more studies planned.

Using a New Version of Minecraft to Promote Computational Thinking Learning Opportunities: Investigating Middle Grades Student Outcomes

Since summer 2020, our interdisciplinary team of educators, computer scientists, and game developers has been working with an Educator Advisory Panel to develop an extension to Minecraft focused on developing computational thinking skills. In summer 2022, we tested the first prototype of our version of Minecraft with a group of middle school students. The first prototype included one partial unit of gameplay, starting with a Minecraft gameplay tutorial. The unit also included three lessons that required students to engage with Minecraft, the learning management platform Canvas, and a custom web-based tool called Minecraft Factory Planner (MFP). The first unit focuses on Computer Science Teachers Association (CSTA) data analysis learning standards: using encoding schemes (2-DA-07), collecting data using computational tools (2-DA-08), and building and refining computational models (2-DA-09). During the first lesson, students build items to help a community within Minecraft that was damaged by a natural disaster. During the second lesson, students build models using the MFP and explore ways to build items more efficiently with minimal waste to address community needs. During the third lesson, students sort collected items using an encoding scheme to make it easier to find items in the future. We were interested in assessing students' computational thinking, self-efficacy, and content knowledge before and after playing these lessons. We were also interested in assessing students' engagement and satisfaction across lessons. The research questions for this study include:

- To what extent did students' computational thinking self-efficacy scores change following the gameplay experience?
- To what extent did students' computational thinking content knowledge change following the gameplay experience?
- To what extent were students engaged in the lessons? To what extent did students' engagement change across the four lessons that students played?

- Did scores vary depending on students' grade, gender, or previous Minecraft experience?

Computational Thinking and Gaming

Computational thinking is a critical skill in today's workforce. We define computational thinking to not focus specifically on coding, but broadly focused on data analysis, abstraction, generalization, and pattern recognition; concepts that lend themselves to the Minecraft gaming experience. Gaming has shown promise for promoting computational thinking (Ketelhut & Schifter, 2011; Schifter, 2013) and user engagement (Cairns, 2016), especially when players enter a state of flow (Leroy, 2021), promoting task persistence. Educational gaming, however, risks taking the form of "chocolate covered broccoli," where educational concepts are not intentionally woven into the gaming narrative. Our project builds on an inherently motivational game and the educational aspects that exist within the game.

We employed a game design process starting with decomposition of the game mechanics and learning standards separately, followed by integration of the mechanics, learning standards, and lesson development (Adams et al., 2021). We promote equity and access for students who may not otherwise have learning experiences focused on computational thinking. The game narratives build on existing research highlighting ways that science contributes to communities and society (Allen et al., 2015).

Data

In summer 2022, we piloted a gameplay experience with 31 students participating in a summer camp at a nonprofit community center in Texas. Students played the tutorial and three lessons, using Canvas, Minecraft, and the MFP. Students participated in three groups of approximately ten students for about 40 minutes of gameplay each day for four days (i.e., one lesson per day). Most students (61%) did not miss any days. In addition to the four days of gameplay, students completed an assessment and survey on the first and last days of the pilot. Two or three university team members facilitated the gameplay experience for students with the support of camp counselors and volunteers. Most students identified as Hispanic (94%) and male (52%). Students were evenly distributed across grades 6, 7, 8.

Students completed the following measures twice, once before and once after the pilot:

- Two measures related to students' computational thinking self-efficacy. We included two measures because we wanted to test both sets of items.
    - The *Computational Thinking Scale for Computer Literacy Education* (Tsai et al., 2020) measures students' computational thinking thought processes for five scales including abstraction, decomposition, algorithmic thinking, evaluation, and generalization. Scores on this measure represent the average across three or four items assigned to each scale. Students responded to each item on a five-point scale ranging from "Strongly Disagree" to "Strongly Agree".
    - We used a portion of the *Computational Thinking Self-Efficacy Survey* (Weese & Feldhausen, 2017) focused on general problem solving, including nine items addressing self-efficacy in problem solving related to algorithms, problem decomposition, parallelization, data, and control flow. Scores on this measure represent the average across nine items. Students responded to each item on a five-point scale ranging from "Strongly Disagree" to "Strongly Agree". Coding is not a focus of our game, so coding centered items were excluded.
- A measure of computational thinking content knowledge consisting of six items from the NAEP Question Tool (https://nces.ed.gov/NationsReportCard/nqt/). We selected items related most closely to the CSTA content standards targeted with gameplay, items focused on technology and society, information and communication technology, data analysis, and design and systems. Items ranged in difficulty from easy to hard. Four items had multiple sub-questions embedded within the overall question, resulting in 27 items and sub-items in total. Scores on this measure represent the sum of questions answered correctly with a maximum score of 27. Scores on this measure ranged from 5 to 23 with a standard deviation of 4.2.

In addition to the self-efficacy and content knowledge measures, students completed the *User Engagement Scale* (Wiebe et al., 2014) each day of gameplay. The survey includes 28 items with six subscales: focused attention, perceived usability, aesthetics, and satisfaction. Focused attention is based in flow theory including concentration, absorption, and temporal dissociation (Wiebe et al., 2014). Perceived usability focuses on affective (frustration) and cognitive (effortful) aspects of the game. Aesthetics focuses on the game's visual appearance. Satisfaction measures the extent to which the experience was fun and interesting. Scores on this measure represent the average across the items assigned to the scale. Students responded to each item using a five-point scale ranging from "Strongly Disagree" to "Strongly Agree".

Method

We tested differences between the baseline and outcome scores for each scale on the surveys and on the content knowledge test using a paired sample t-test with Bonferroni-adjusted post hoc tests. As part of exploratory work given the small sample (n=31), we compared scores for students by grade, gender, and whether they had previous Minecraft experience, which was measured as part of a background survey at baseline.

We examined means across lessons and tested for change across lessons in engagement based on the scales included on the engagement survey using a repeated measures analysis of variance (ANOVA). We tested for interaction effects based on grade, gender, and whether the students had previous Minecraft experience.

We anticipated statistically significant increases in student self-efficacy and content knowledge scores. However, given the small number of lessons that students experienced, we anticipated that not all differences would be statistically significant, in part due to limited dosage (i.e., shorter durations spent playing the game). Similarly, we anticipated that engagement would remain consistently high across lessons and across students overall, including when grouped by grade, gender, and previous Minecraft experience.

Results

Table 1 shows the t-test results overall and for subgroups.

Self-efficacy. Students' self-efficacy for general problem solving increased after the gameplay experience (*p*<01). General problem solving was measured using a scale with nine items developed by Weese and Feldhausen (2017). Looking at differences based on students' grade, gender, or previous Minecraft experience, we observed that students who identified as female increased in general problem-solving self-efficacy (*p*<.05); whereas students who identified as male did not increase significantly. We also observed an increase in general problem-solving self-efficacy in students with limited Minecraft experience (*p*<.05) and students entering Grade 8 (*p*<.01).

Scores increased on the other scales measuring students' computational thinking self-efficacy developed by Tsai and colleagues (2020); however, the increases were not statistically significant. We did not observe statistically significant mean differences for student subgroups based on this self-efficacy measure.

Content knowledge. The overall difference in students' computational thinking content knowledge was not statistically significant. The results for subgroups indicate that male students scored higher on content knowledge after gameplay (*p*<.05). Increases were observed for students across the other subgroups; however, the increases were only statistically significant for male students.

User Engagement. Students tended to report the aesthetics and their satisfaction with the game higher than the other two scales (Figure 1). On average students tended to agree, as indicated by a score of 4, with statements related to aesthetics and satisfaction. Students tended to neither agree nor disagree, as indicated by a score of 3, with statements related to focused attention. Students tended to report perceived usability of the game lower with scores ranging from disagree, as indicated by a score of 2, to neither agree nor disagree depending on the lesson.

Table 2 shows the results of the ANOVA overall and for subgroups. Students reported statistically significant change in perceived usability scores across lessons. Post hoc tests indicated that the differences were marginally statistically significant between the tutorial and lesson 2 (p=.049), with perceived usability decreasing and then remaining consistent across lessons 2 and 3. We did not observe differences in scores on the engagement survey across lessons for the other scales including focused attention, aesthetics, and satisfaction.

Using interaction terms in the ANOVA, we tested for differences in scores based on students' grade, gender, or previous Minecraft experience. We observed a statistically significant interaction for gender with male students reporting consistently decreasing perceived usability across lessons; whereas female students reported decreasing perceived usability across the tutorial, lesson 1, and lesson 2, but perceived usability increased for lesson 3.

Scholarly Significance

This paper reflects evidence related to students' experiences playing an extension of Minecraft designed to provide students with learning opportunities related to computational thinking. We observed evidence that the game provides students with learning opportunities related to the learning standards, although differences before and after gameplay were not statistically significant for most students. We also observed that students generally liked the aesthetic of the game and enjoyed playing the game. This preliminary evidence supports that mapping content standards to existing game mechanics can support building on inherently educational games in ways that explicitly map the learning standards.

We observed increases in self-efficacy on one measure of general problem solving in computational thinking but not on the other measure of computational thinking self-efficacy. This disconnect may reflect the broad interdisciplinary focus of the game that is not specifically aligned to computer science and coding. It is possible that students did not have enough exposure to the game to realize changes on specific measures of self-efficacy related to computational thinking concepts.

Perceived usability decreased across lessons, with statistically significant differences between males and females on lesson 3. The tutorial focused on game mechanics (e.g., how to jump) and did not require students to engage in computational thinking as it is described in the content standards, which may explain why students reported higher usability for the tutorial. It is interesting that perceived usability decreased across lessons 1 through 3 for males and decreased across lessons 1 and 2, increasing in lesson 3 for females. The lessons generally increased in difficulty, building on concepts learned in previous lessons and incorporating new content. The gender difference for lesson 3, which focused on sorting items using an encoding scheme, warrants further investigation.

This work builds on equity-oriented educational gaming focused on computational thinking (e.g., Leonard et al., 2016) by focusing on concepts related to computational thinking that are not specific to coding. This game positions students as engineers in the Minecraft world, which may inspire students who would not otherwise consider computer science (e.g., girls with limited Minecraft experience) to feel efficacious in and build computer science knowledge and computational thinking.

References

Allen, J. M., Muragishi, G. A., Smith, J. L., Thoman, D. B., & Brown, E. R. (2015). To grab and to hold: Cultivating communal goals to overcome cultural and structural barriers in first-generation college students' science interest. *Translational Issues in Psychological Science, 4*(4), 331-341.

Adams, E. L., Tseng, C., Foster, P., Luo, V., Ketterlin-Geller, L. R., Larson, E. C., & Clark, C. (2021, January). *A standard decomposition process to inform the development of game-based learning environments focused on computational thinking*. Paper presented to the International Conference on Computational Thinking and STEM Education (Virtual).

Cairns, P. (2016). Engagement in Digital Games. In H. O'Brien & P. Cairns (Eds.), Why Engagement Matters: Cross-Disciplinary Perspectives of User Engagement in Digital Media (pp. 81–104). Springer International Publishing. https://doi.org/10.1007/978-3-319-27446-1_4

Feldhausen, R., Weese, J. L., & Bean, N. H. (2018). Increasing students' self-efficacy in computational thinking via STEM outreach programs. Paper presented to SIGCSE'18 in Baltimore, MD.

Ketelhut, D. J., & Schifter, C. C. (2011). Teachers and game-based learning: Improving understanding of how to increase efficacy of adoption. *Computers and Education, 56*(2), 539–546.

Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O. S., Hubert, T., & Almughyirah, S. (2016). Using robotics and game design to enhance children's self-efficacy, STEM attitudes, and computational thinking skills. *Journal of Science Education and Technology, 25*, 860-876.

Leroy, R. (2021). Immersion, Flow and Usability in video games. Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 1–7. https://doi.org/10.1145/3411763.3451514

Schifter, C. C. (2013). Games in learning, design, and motivation. In M. Murphy, S. Redding, & J. Twyman (Eds.), *Handbook on innovations in learning* (pp.149-164). Information Age Publishing.

Tsai, M., Liang, J., & Hsu, C. (2021). The computational thinking scale for computer literacy education. *Journal of Educational Computing Research, 49*(4), 579-602.

Wiebe, E., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior, 32*, 123-132.

Table 1

T-Test Results on Baseline and Outcome Measures Overall and for Subgroups

Scale	n	Baseline Mean (SD)	Outcome Mean (SD)	T-Statistic
Self-Efficacy – Abstraction[3] Overall	**26**	**3.61 (.93)**	**3.77 (.70)**	**.88**
• Male Only	14	3.57 (1.10)	3.93 (.70)	1.51
• Female Only	12	3.65 (.71)	3.58 (.68)	.22
• Limited Minecraft Experience Only	11	3.27 (.97)	3.52 (.44)	.93
• Minecraft Experience Only	15	3.85 (.84)	3.95 (.81)	.38
• Grade 6 Only	10	3.78 (1.06)	3.95 (.83)	.54
• Grade 7 Only	9	3.78 (.57)	3.56 (.60)	.83
• Grade 8 Only	7	3.14 (1.05)	3.79 (.62)	1.92
Self-Efficacy – Decomposition[1] Overall	**26**	**3.63 (.64)**	**3.68 (.79)**	**.32**
• Male Only	14	3.45 (.62)	3.71 (.86)	1.56
• Female Only	12	3.83 (.63)	3.64 (.73)	.70
• Limited Minecraft Experience Only	11	3.52 (.74)	3.55 (.62)	.14
• Minecraft Experience Only	15	3.71 (.58)	3.78 (.90)	.29
• Grade 6 Only	10	3.47 (.67)	3.60 (.99)	.55
• Grade 7 Only	9	3.93 (.68)	3.70 (.56)	.71
• Grade 8 Only	7	3.48 (.47)	3.76 (.81)	.89
Self-Efficacy – Algorithmic Thinking[1] Overall	**26**	**3.98 (.67)**	**4.01 (.71)**	**.17**
• Male Only	14	3.82 (.66)	4.07 (.72)	1.39
• Female Only	12	4.17 (.66)	3.94 (.72)	.78
• Limited Minecraft Experience Only	11	3.75 (.68)	3.89 (.68)	.69
• Minecraft Experience Only	15	4.15 (.63)	4.10 (.74)	.19
• Grade 6 Only	10	3.95 (.78)	3.93 (.83)	.09
• Grade 7 Only	9	4.11 (.36)	4.02 (.76)	.27
• Grade 8 Only	7	3.86 (.86)	4.10 (.54)	.79
Self-Efficacy – Evaluation[1] Overall	**26**	**3.93 (.72)**	**4.00 (.74)**	**.45**
• Male Only	14	3.88 (.76)	4.13 (.81)	1.41
• Female Only	12	4.00 (.70)	3.85 (.64)	.59
• Limited Minecraft Experience Only	11	3.89 (.66)	4.04 (.75)	1.05
• Minecraft Experience Only	15	3.97 (.78)	3.97 (.76)	.00
• Grade 6 Only	10	3.78 (.83)	4.03 (.74)	.89
• Grade 7 Only	9	4.36 (.59)	4.19 (.77)	.63
• Grade 8 Only	7	3.61 (.48)	3.71 (.71)	.51

---

[3] Tsai et al., 2020

Scale	n	Baseline Mean (SD)	Outcome Mean (SD)	T-Statistic
Self-Efficacy – Generalization[1] Overall	**26**	**3.92 (.72)**	**3.95 (.78)**	**.16**
• Male Only	14	3.95 (.68)	4.13 (.86)	.85
• Female Only	12	3.90 (.80)	3.75 (.65)	.46
• Limited Minecraft Experience Only	11	3.86 (.78)	3.82 (.71)	.18
• Minecraft Experience Only	15	3.97 (.71)	4.05 (.84)	.31
• Grade 6 Only	10	4.03 (.67)	3.83 (.94)	.62
• Grade 7 Only	9	4.08 (.79)	4.11 (.74)	.09
• Grade 8 Only	7	3.57 (.69)	3.93 (.62)	1.08
Self-Efficacy – General Problem Solving[4] Overall	**26**	**3.50 (.76)**	**3.92 (.61)**	**3.19****
• Male Only	14	3.37 (.92)	3.79 (.68)	2.05
• Female Only	12	3.65 (.49)	4.07 (.50)	2.53*
• Limited Minecraft Experience Only	11	3.25 (.91)	3.84 (.70)	2.38*
• Minecraft Experience Only	15	3.67 (.58)	3.99 (.55)	2.11
• Grade 6 Only	10	3.64 (.68)	3.87 (.68)	1.05
• Grade 7 Only	9	3.59 (.48)	3.96 (.58)	.20
• Grade 8 Only	7	3.16 (1.09)	3.95 (.63)	3.77**
Content Knowledge Overall	**29**	**16.17 (4.07)**	**17.17 (4.25)**	**1.60**
• Male Only	15	16.27 (4.79)	17.87 (4.22)	2.30*
• Female Only	13	15.77 (3.24)	16.15 (4.36)	.33
• Limited Minecraft Experience Only	13	15.69 (4.31)	15.62 (5.08)	.11
• Minecraft Experience Only	16	16.56 (3.97)	18.44 (3.05)	1.96
• Grade 6 Only	11	16.72 (3.22)	18.72 (3.29)	1.90
• Grade 7 Only	9	16.89 (3.48)	17.44 (4.28)	.51
• Grade 8 Only	9	14.78 (5.45)	15 (4.77)	.20

*** p<.001; ** p<.01; * p<.05

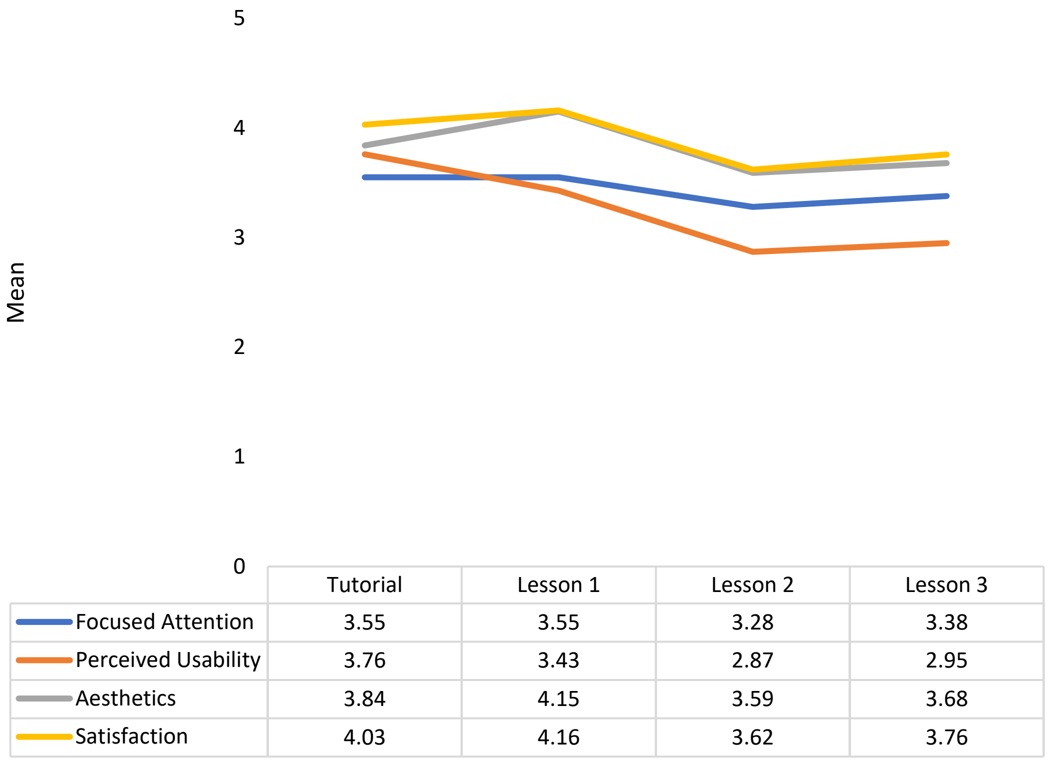
---

[4] Weese & Feldhausen, 2017

Table 2

Repeated Measures ANOVA Results on Baseline and Outcome Measures Overall and for Subgroups

Engagement Scale	n	Tutorial Mean	Lesson 1 Mean	Lesson 2 Mean	Lesson 3 Mean	F-Statistic
Focused Attention Overall	**19**	**3.55**	**3.56**	**3.29**	**3.38**	**.67**
• Male	9	3.60	3.46	3.07	3.42	.56
• Female	10	3.51	3.65	3.49	3.34	
• Limited Minecraft Experience	10	3.21	3.66	3.19	3.09	.84
• Minecraft Experience	9	3.93	3.44	3.40	3.69	
• Grade 6	6	3.48	3.29	2.85	3.29	.71
• Grade 7	6	3.63	3.31	3.67	3.38	
• Grade 8	7	3.55	4.00	3.33	3.44	
Perceived Usability Overall	**19**	**3.76**	**3.43**	**2.87**	**2.95**	**3.34***
• Male	9	4.03	3.99	3.21	2.36	5.47**
• Female	10	3.51	2.94	2.56	3.49	
• Limited Minecraft Experience	10	3.51	3.24	2.74	2.90	.11
• Minecraft Experience	9	4.02	3.65	3.01	3.01	
• Grade 6	6	4.41	4.10	3.44	2.98	.47
• Grade 7	6	3.31	2.98	2.65	3.29	
• Grade 8	7	3.57	3.25	2.57	2.64	
Aesthetics Overall	**19**	**3.84**	**4.15**	**3.59**	**3.68**	**1.66**
• Male	9	3.91	4.20	3.71	3.84	.20
• Female	10	3.78	4.10	3.48	3.54	
• Limited Minecraft Experience	10	3.64	3.92	3.30	3.78	.95
• Minecraft Experience	9	4.07	4.40	3.91	3.58	
• Grade 6	6	3.80	4.07	3.40	3.47	.51
• Grade 7	6	3.80	4.20	4.07	3.60	
• Grade 8	7	3.91	4.17	3.34	3.94	
Satisfaction Overall	**19**	**4.04**	**4.16**	**3.62**	**3.76**	**1.84**
• Male	9	4.30	4.33	3.83	3.84	.14
• Female	10	3.80	4.00	3.44	3.69	
• Limited Minecraft Experience	10	4.04	4.09	3.39	3.84	.89
• Minecraft Experience	9	4.03	4.24	3.89	3.67	
• Grade 6	6	4.26	4.38	3.52	4.14	.45
• Grade 7	6	4.00	4.02	4.14	3.52	
• Grade 8	7	3.88	4.08	3.27	3.63	



	Tutorial	Lesson 1	Lesson 2	Lesson 3
Focused Attention	3.55	3.55	3.28	3.38
Perceived Usability	3.76	3.43	2.87	2.95
Aesthetics	3.84	4.15	3.59	3.68
Satisfaction	4.03	4.16	3.62	3.76

Figure 1

Means Across Lessons on the User Engagement Survey

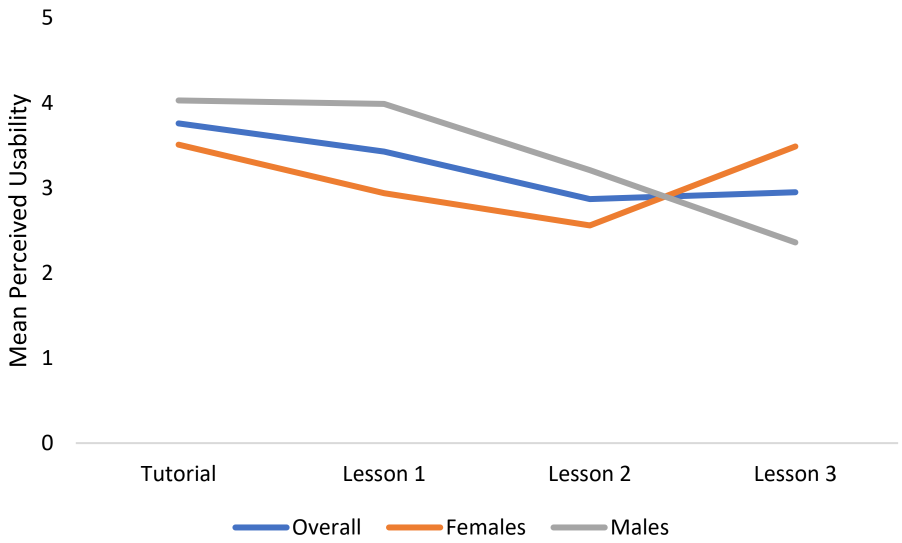


Figure 2

Mean Perceived Usability Overall and for Males and Female Students