What do Students Learn About Experimental Research by Designing Interactive Fiction Games?

Anna Amato, New York University, ada437@nyu.edu
Camillia Matuk, New York University, cmatuk@nyu.edu
Dylan Schouten, Northeastern University, dylan.schouten@gmail.com
Steven Sutherland, University of Houston-Clear Lake, Sutherland@UHCL.edu
Gillian M. Smith, Worcester Polytechnic Institute, gmsmith@wpi.edu
Casper Harteveld, Northeastern University, c.harteveld@northeastern.edu

Abstract. Research design is challenging to learn, and students have few hands-on experiences to practice it. We explore how StudyCrafter, a platform for creating and running research studies in the form of interactive fiction games, promotes students' perceived and measured abilities in certain key research skills. Fourteen graduate students in a game design course used StudyCrafter during a 5-week-long unit on designing and conducting experimental research studies. Analyses of pre- and post-tests, and of students' written post-unit reflections, showed that students' confidence and abilities in designing research studies increased, but that their abilities to critique studies did not. This study contributes an understanding of the opportunities and challenges in using interactive fiction games to support students in developing their experimental research skills.

Keywords: game design, narrative, research, university students

Major issues addressed and potential significance

Learning to design experimental research studies is challenging for university students. They have difficulty designing researchable hypotheses, identifying and operationalizing variables, controlling for variability, aligning methods with research goals, and grounding conclusions in evidence (Woolley et al., 2018). Moreover, instructors have little time and few materials necessary to support their students in authentic ways. We explore how StudyCrafter, a platform for creating interactive fiction games, can support students' confidence and abilities in key experimental research design skills. More broadly, this project seeks to explore the role of narrative as a scaffold for students in learning to design and critique experimental research. The current analysis shows our first steps to identifying the impacts of this approach on specific research competencies.

Theoretical background

Competence in experimental research involves abilities to both design and critique studies in terms of researchable hypotheses, operationalized variables, and the alignment of these with study designs that account for confounding variables. Self-perceived competence in research is also important, as greater self-perceived competence implies greater interest, persistence, and effort, which in turn contribute to learning and achievement (Zimmerman & Bandura, 1994). Games and narrative can make experimental research design more approachable, effective, and feasible for learners. In particular, game making can promote learners' domain understanding, disciplinary identification, and metacognition (Kafai & Burke, 2015); and narrative offers an intuitive framework for making sense of unfamiliar concepts (Bruner, 1991). We moreover propose that narrative offers a useful framework for designing robust research studies (Matuk et al., 2019). For example, just as a good interactive fiction game gives players a compelling narrative context that facilitates the suspension of disbelief and the perception of agency, robust research designs invite participants' buy-in to an experimental manipulation. This buy-in can ensure that participants' behaviors better reflect how they would normally respond to these events had they occurred in the real world. Such research designs thus have greater internal and external validity.

Building on these ideas, we created StudyCrafter, a platform that allows users to create research experiments in the form of interactive fiction games. It also allows them to run these experiments with real research participants, which is otherwise challenging within the constraints of a typical semester-long course (Harteveld et al., 2017). With StudyCrafter, creators select and arrange characters, objects, and backgrounds from an asset library, organize these into scenes, and program them to respond to players' actions (See Figure 1). Creators then distribute a link to potential participants. In a typical StudyCrafter project, a player takes on the role of a character, and makes choices within a story context. These choices become the data that creators analyze to answer their research question. StudyCrafter was designed to be incorporated into university-level courses for students across domains to gain hands-on practice with designing and conducting research studies.

Research questions

This study explores how creating and conducting experimental research studies in the form of interactive fiction games impacts: (1) students' self-perceived competencies in key skills related to designing and critiquing experimental research studies; (2) students' measured competence in designing and critiquing studies; and (3) how students' reflections on research learning experiences resonate with their perceived and measured competencies in research study design.

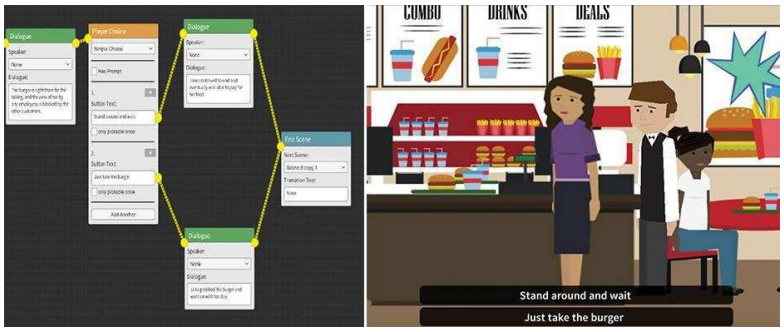

Figure 1. Left: Scriptor interface for programming game assets; Right: Scene from a StudyCrafter project that tests whether the ethics of a player's character impacts decision making.

Participants and context

Participants were 14 students (9 males, 4 females) enrolled in a graduate-level introductory game design course at a university in the eastern United States. Students had varying disciplinary backgrounds and prior experiences with research study design. In a 5-week long unit during the last part of the semester, students used StudyCrafter to individually design and conduct research studies that explored questions of their choice relevant to game design or social psychology (e.g., loot boxes, role of player-character, ethical decision making, course selection, anxiety, altruism). Students recruited participants, analyzed data collected, and produced written reports of their findings, which they exchanged for peer critique and feedback.

Study design, data, and analysis

To address (RQ1), we administered an 18-item pre- and post-test at the start and end of the 5-week unit, in which students rated their self-perceived research competencies (e.g., "I am confident I can formulate researchable hypotheses."). We created overall scores by adding the scores on each dimension (7 points total), and used a Wilcox signed rank test to find significant differences between pre and post-test scores.

To measure change in students' design and critique abilities (RQ2), we administered a pre/post-test, in which students designed and critiqued a study (See Table 1) by responding to a series of prompts (e.g., "State the study's hypothesis", "Describe the study's dependent and independent variables").

Two researchers developed rubrics based on the activity's intended learning outcomes, and to capture the range of quality of students' responses. These rubrics identified 13 design dimensions and 8 critique dimensions (0-2 points each). For example, on the ability to *identify independent variables (IV)*, scores included 0 (IV is misidentified), 1 (does not extend logically from the study design or it is not properly defined), and 2 (IV extends logically from the study design and is properly defined). We added scores of each dimension to create overall scores for design (26 points total) and critique (16 points total). We used a Wilcox signed rank test to find significant differences between pre- and post-test scores.

Table 1: The Design and Critique pre-test items, which were each followed by guiding prompts (not shown). Post-test items focused on the theory of Delayed Gratification

Design a study: *The Forbidden Fruit Theory* refers to a person's tendency to react in the opposite manner to what someone wants, as a way of resisting the perceived constraint on their freedom of choice. Design an experiment to investigate whether this phenomenon exists in a game environment.
Critique a study: Sam designed an experiment to explore the *Forbidden Fruit Theory* in a game setting. She created an online adventure game in which players collect coins along the path of a precarious terrain. At certain points in the game, players can choose to exchange their coins either for resources necessary for survival, or for objects with no value.

Sam randomly assigned 6 players to each of two conditions. In *Condition A*, players were free to choose which objects to purchase and when. In *Condition B*, an in-game character appeared periodically to remind players that the road was precarious, and that they would need certain supplies if they wanted to succeed in winning the game. Sam found that players in Condition B were significantly more likely to purchase non-necessary supplies than participants in Condition A.

To explore how students' perceived competencies (from RQ1) and measured competencies (from RQ2) resonated in reflections on their learning experiences (RQ3), we prompted students to write open-ended reflections on their experiences at the end of the unit. Two researchers independently read the reflections and identified emergent themes through inductive coding (Strauss & Corbin, 1998). Through discussion, we came to consensus on a rubric with three themes, each with up to 5 sub-themes: user experience (e.g., adequacy of instructional support), learning experience (e.g., what was learned about research), and self-reflection (e.g., self-critique of their process). We then noted how many of the *Learning Experiences* statements resonated with the key research competencies by coding these in terms of the competencies identified in RQ1 and RQ2.

Major findings

Regarding RQ1, we found a significant increase in participants' mean competency scores from pre-test (M=76.9, SD=14.3) to post-test (M=88.1, SD=10.7). Students showed significant improvement in overall confidence in designing research studies, and confidence in their abilities to: (1) formulate researchable hypotheses, (2) devise appropriate levels of measurement for each variable, (3) design a study that tests a specific hypothesis, (4) make good use of the available experiment-creation tools to address research questions, (5) design experimental tasks and settings that participants will take seriously, (6) be aware of how stereotypes are reinforced through research design, (7) design effective studies, and (8) design an experiment in a reasonable amount of time. Effect sizes for these competencies ranged from r=.56 to r=.77. These findings suggest that StudyCrafter was successful in helping students develop their sense of confidence in key research skills.

Regarding RQ2, we found a significant increase in participants' mean scores for designing studies from pre-test (M=12.6, SD=5.5) to post-test (M=17.4, SD=4.9). Students showed significant increases in: (1) formulating researchable hypotheses, (2) identifying independent variables, (3) identifying dependent variables, and (4) identifying confirming evidence. Effect sizes for these design competencies ranged from r=.53 to r=.72, with competency 1 (formulating researchable hypotheses) improving the most. Notably, our observation of students' improved abilities on this latter competency are consistent with students' perceptions of improvement. Meanwhile, students showed no significant improvement on critique items between pre-test (M=8.2, SD=2.9) and post-test (M=8.9, SD=3.4), which suggests a need for more instructional support for critiquing studies.

Regarding RQ3, we found that of the 437 statements in students' reflections, 47% were about learning experiences, 47% about self-reflection, and 6% about user experiences. Out of 11 emergent themes in *Learning Experiences*, 8 corresponded to research competencies (76% of reflections coded as learning experience). The top three were study design (29%), study design alignment (13%), and data analysis (17%). For example, one student wrote, "You have to correctly identify your variables first and choose the method that fits your research goal. Also picking the right scope is crucial for your study." Another reflects on the challenge of aligning a hypothesis with data analysis: "Another difficulty I personally ran into was what analysis to do for the data. It supposedly tied with the hypothesis and what we should do." These examples show that students recognize the importance of alignment, data analysis, and study design even though they did not show improvement on post-tests. These results point to key opportunities for instructional support.

Four to six students (up to 42.9% of participants) reflected on identifying and operationalizing variables, identifying confounding variables, and identifying confirming evidence. For example, "differentiating between independent and dependent variables became much easier as I worked on my project. It became much clearer which variable was affecting the others." Another wrote, "The thing is all your conclusions should be drawn from your experimental results. You cannot show the result of your own inference without proofs." Findings from RQ2 shows that while students improved on two of these competencies, they struggled to identify confounding variables. Yet, their reflections indicate their metacognitive awareness of the importance and challenge of these competencies, which is an important step in learning (Panadero, 2017).

Conclusions and implications

These findings inform new StudyCrafter features, including prompts to guide students' hypothesis construction and variable identification. They also inform refinements to the surrounding instruction, such as including a rubric to guide students' critique of their own and others' studies. In ongoing analyses of students' StudyCrafter projects and reports, we are exploring relationships between the quality of students' narrative and research designs, and

investigating students' abilities to interpret findings from their studies. This work is an example of how interactive fiction design can develop students' experimental research competencies. It also contributes an understanding of the role of narrative and game making in promoting students' disciplinary skills, self-competence, and learning.

References

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*(1), 1-21.

Ford, M. J. (2005). The game, the pieces, and the players: Generative resources from two instructional portrayals of experimentation. *The Journal of the Learning Sciences*, *14*(4), 449-487.

Harteveld, C., Manning, N., Abu-Arja, F., Menasce, R., Thurston, D., Smith, G., & Sutherland, S. C. (2017, March). Design of playful authoring tools for social and behavioral science. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion* (pp. 157-160).

Kafai, Y. B., & Burke, Q. (2015). Constructionist gaming: Understanding the benefits of making games for learning. *Educational Psychologist*, *50*(4), 313-334. DOI: 10.1080/00461520.2015.1124022

Matuk, C., Sutherland, S., Althoff, W., Snodgrass, S. *Partlan, N. Smith, G., Seif El-Nasr, M. & Harteveld, C. (2019, Apr 5-9). Synergies between research and game design: Reflections on interactive narrative experiments by student game designers. Poster presented at the *2019 Annual Meeting of the American Educational Research Association*, Toronto, ON.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422.

Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research Techniques and Procedures for Developing Grounded Theory* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Woolley, J. S., Deal, A. M., Green, J., Hathenbruck, F., Kurtz, S. A., Park, T. K., ... & Jensen, J. L. (2018). Undergraduate students demonstrate common false scientific reasoning strategies. *Thinking Skills and Creativity*, *27*, 101-113.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, *29*(3), 663-676.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants No. 1736065, 1736056 and 1736185.