Getting Academical: A Choice-Based Interactive Storytelling Game for Teaching Responsible Conduct of Research

Edward F. Melcer
University of California, Santa Cruz
Santa Cruz, CA
eddie.melcer@ucsc.edu

Katelyn M. Grasse
University of California, Santa Cruz
Santa Cruz, CA
katy@ucsc.edu

James Ryan
Carleton College
Northfield, MN
joryan@ucsc.edu

Nick Junius
University of California, Santa Cruz
Santa Cruz, CA
njunius@ucsc.edu

Max Kreminski
University of California, Santa Cruz
Santa Cruz, CA
mkremins@ucsc.edu

Dietrich Squinkifer
Independent Artist
Montreal, QC, Canada
hey@squinky.me

Brent Hill
University of Utah
Salt Lake City, UT
brent.hill@hsc.utah.edu

Noah Wardrip-Fruin
University of California, Santa Cruz
Santa Cruz, CA
nwardrip@ucsc.edu

ABSTRACT

Concepts utilizing applied ethics, such as responsible conduct of research (RCR), can prove difficult to teach due to the complexity of problems faced by researchers and the many underlying perspectives involved in such dilemmas. To address this issue, we created *Academical*, a choice-based interactive storytelling game for RCR education that enables players to experience a story from multiple perspectives. In this paper, we describe the design rationale of *Academical*, and present results from an initial study comparing it with traditional web-based educational materials from an existing university RCR course. The results highlight that utilizing a choice-based interactive story game is more effective for RCR education, with learners developing significantly higher engagement, stronger overall moral reasoning skills, and better knowledge scores for certain RCR topics.

CCS CONCEPTS

• **Human-centered computing**;

KEYWORDS

choice-based; role-playing; interactive storytelling; narrative game; educational game; responsible conduct of research; ethics

1 INTRODUCTION

Topics such as the responsible conduct of research (RCR) are difficult to teach due to the complexity of applied ethics and ethical decision-making [3], the need for moral reasoning [58], and the lack of existing educational tools that are motivating and foster critical thinking [19]. While past work has attempted to address these issues through alternative learning approaches such as group mentoring [72] and role-playing [5, 59], these issues have still remained largely unaddressed—resulting in ill-defined content, format, and goals, as well as minimal evidence for effectiveness [18]. Conversely, in the context of educational games, choice-based interactive storytelling is a popular format for narrative videogames [12, 40, 56]. There have even been educational interactive narratives designed specifically to teach issues related to ethics [17], although they have yet to be evaluated for effectiveness. Interactive storytelling (and educational games in general [23, 34, 35]) have also been shown to increase engagement/motivation and learning for more rote topics with clearly defined answers and educational outcomes, such as in the areas of STEM [53, 70, 73]. However, past work has not fully examined the capabilities of choice-based interactive storytelling games in teaching more ambiguous concepts such as moral reasoning and ethical decision-making.

RCR in particular is an important concept that warrants study of and improvement to existing training tools. This is because it comprises fundamental ethical topics that inform all aspects of the research process, which can also be further complicated by many factors such as power dynamics and marginalized identities. As a result, RCR requires understanding a variety of perspectives and dilemmas that impact underlying research ethics [21, 60]. Additionally, current educational RCR tools suffer from a notable lack of user engagement and motivation when learning the material [19].

Interactive storytelling games may be particularly effective for addressing the above issues with RCR education. Specifically, we hypothesized that the choice-based, role-playing nature of interactive storytelling games could also be employed to improve student engagement, learning outcomes, and moral reasoning within

ethically complex topics such as RCR education—which requires learners to understand a variety of perspectives and perform ethical decision-making. As a result, we created *Academical*, a choice-based interactive storytelling game for RCR education that allows players to experience a story from multiple perspectives. In this paper, we discuss the design of *Academical*, and provide results from an initial study comparing engagement and learning outcomes of our web-based game with traditional web-based educational materials from an existing RCR course at the University of Utah. We conclude with a discussion of the results and their implications for the usage of choice-base interactive storytelling games for teaching ethics knowledge, moral reasoning skills, RCR, and improving the overall experience of educational role-playing.

2 BACKGROUND

In this section, we provide background information on our project, with an emphasis on choice-based interactive storytelling and its use in learning materials. We also discuss RCR, the subject area for which *Academical* serves as an educational resource, and past research exploring RCR education.

2.1 Choice-based Interactive Storytelling

Though it is attested as far back as the sixteenth century [38, 54], choice-based interactive storytelling was made famous by the *Choose Your Own Adventure* book series [51, 55] and is now most prominent as a popular format for narrative videogames [12, 40, 56]. For instance, the various titles developed by Telltale Games, e.g., [64, 65]. In this format, players navigate a *plot graph* [71] by making decisions (typically on behalf of a character) at branching points in the narrative (see Figure 1 for an excerpt from the plot graph for *Academical*). Research in this area has typically concerned the history [12, 38, 54, 56], analysis [31, 32, 40], or procedural generation [15, 33, 45] of works in the choice-based format. Of particular relevance to our study here is prior work that has argued for the format's power in terms of evoking empathy [4, 56, 57],[1] providing therapeutic benefits [9, 63], and enabling learning experiences, the latter of which we discuss next in a dedicated section.

2.2 Interactive Storytelling and Learning

Interactive storytelling has substantial potential for education and games [6, 8, 36, 41, 69]. Specifically, narrative/storytelling is an important element that can be incorporated into educational games in order to maintain and increase students' motivation [7, 10, 44, 53], with some suggesting that integration of a good story into an educational game will determine its success or failure [13]. Interactive storytelling has been incorporated into a number of educational games focusing on topics such as history [7, 61], STEM [8, 70, 73], and bullying [2, 67]. However, the majority of research on educational interactive storytelling games has focused on adaptivity [14, 24], interactivity [61, 73], emergent narrative [2], player and knowledge modeling [29, 52], narrative planning and generation [16, 50, 66, 74], and the game creation process itself [7, 62]. As a result, there is surprisingly little work evaluating the impact of an interactive storytelling approach on learning outcomes (exceptions

---

[1]Though see [49] for a critique of this notion.
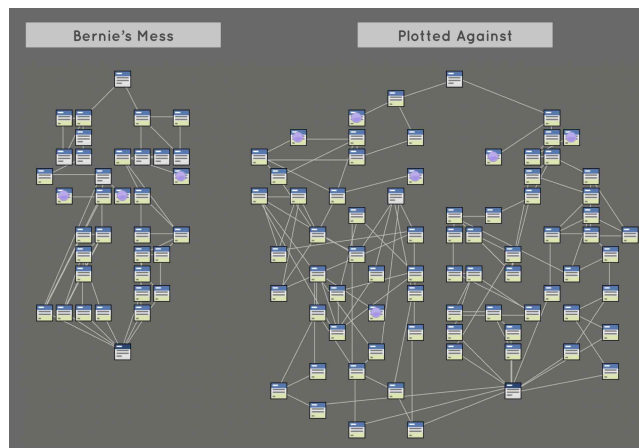
Figure 1: Plot graphs for two of *Academical*'s playable scenarios, visualized in the Twine authoring environment. Each node in these graphs is a Twine "passage" (story unit), some of which are player choice points that link to other passages. As the game progresses, the scenarios become more complex—of the two scenarios shown here, the one on the right comes later in the game.

being [37, 53, 67, 70, 73]), especially for topics such as RCR with ethically complex concepts that require a variety of perspectives.

2.3 Responsible Conduct of Research

Although students generally know that they should report data honestly and cite sources accurately, they might not know specific standards or obligations of RCR—such as criteria for co-authorship and maintaining the confidentiality of manuscripts reviewed for publication [48, 59]. The importance of RCR is such that many major funding agencies, such as the National Institutes of Health (NIH) and National Science Foundation (NSF), explicitly require researchers supported by their grants to receive RCR training [43, 47]. Currently, the NIH provides a guideline of nine core RCR topics [20]: 1) *conflict of interest*, 2) *human and animal subjects*, 3) *mentoring*, 4) *collaboration*, 5) *peer review*, 6) *data management*, 7) *research misconduct*, 8) *authorship and publication*, and 9) *scientists and society*. Past research on RCR education has ranged from issues teaching ethical theories underlying RCR [3] and identifying metacognitive reasoning strategies that facilitate ethical decision-making [25, 39] to the use of group mentoring [72] and role-playing [5, 59] for improved training efficacy. However, there is still a notable engagement issue within current RCR education, and a critical need for a variety of tools to improve discussion, engagement, and critical thinking [19]. As a result, an interactive storytelling approach may prove effective for increasing motivation and fostering deeper critical thinking.

3 ACADEMICAL

Academical is a work of choice-based interactive storytelling [26, 31, 32] that was created using the Twine authoring framework [12, 56]. The game comprises nine playable scenarios, each pertaining to a specific topic in RCR [20]. These scenarios are adapted (with
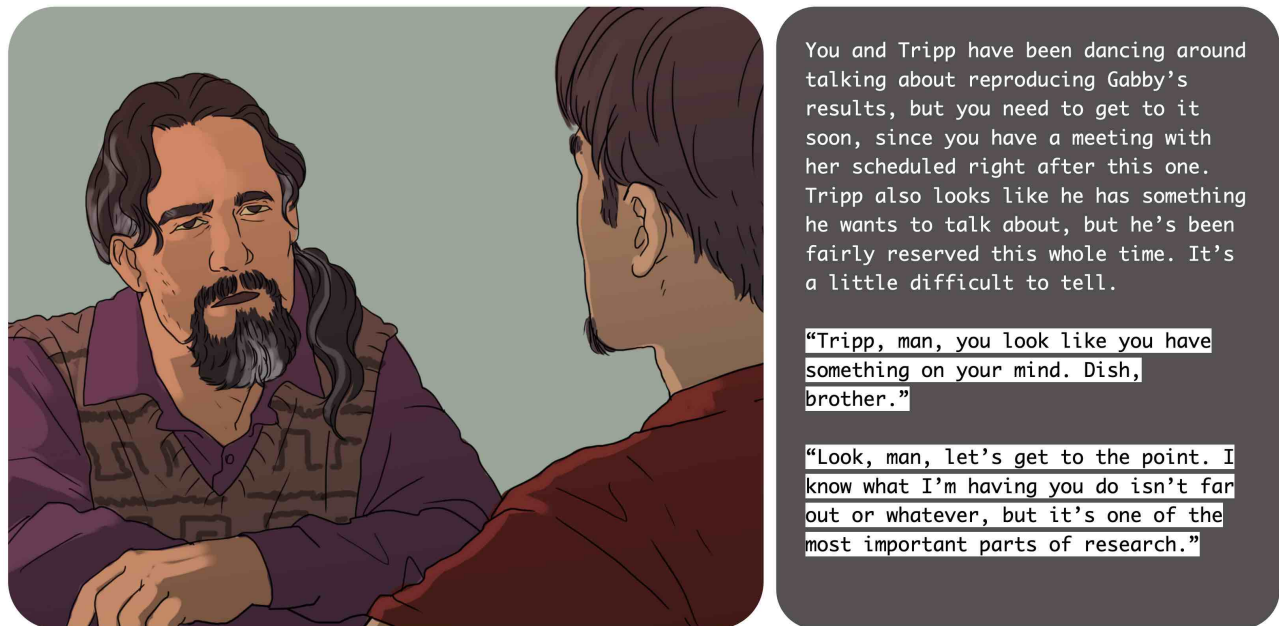
Figure 2: A choice point from *Academical*'s final scenario, "Fallen Angel Y2K." In this scene, the player controls a busy professor whose graduate student suspects that a postdoc in the lab has fabricated research results. The two highlighted text blocks represent dialogue options between which the player must select. To complete the scenario, the player must also navigate the situation responsibly while acting as the graduate student.

permission) from a series of existing educational RCR role-playing prompts [5, 59]. Figure 2 shows a screenshot taken during gameplay, which occurs in a web browser.

Each playable scenario in *Academical* centers on a conversation between two stakeholders in the RCR issue at hand, one of whom is controlled by the player—in the sense that they select dialogue options for that character. By virtue of these choices, the player will ultimately reach one of several possible endings, a subset of which represent successful navigation of the situation. Upon reaching a good ending for the first character, the player then unlocks the other interlocutor and replays the scenario from that person's viewpoint. In turn, reaching a good ending for the second character in a given scenario unlocks the next scenario/RCR topic. The game concludes upon completion of the final scenario. Generally, the scenarios become more complex (and difficult to navigate) as the game proceeds, as Figure 1 illustrates.

At the outset of the project, we decided that the format of choice-based interactive storytelling—which allows a player to experience a story from multiple perspectives and replay scenes to see how different actions play out—would demonstrate the complicated nature of RCR to students in a compelling way. In adapting the role-playing prompts, we sought to show how seemingly obvious answers around questions of research ethics can be complicated by factors such as power dynamics and marginalized identities and experiences. Instead of cleanly delineating right and wrong answers, *Academical* showcases complexity and uncertainty to

provoke questions around how courses of action could have unexpected consequences. In turn, while all successful paths through the game's scenarios represent the player character acting responsibly, not all of the situations reach clear resolutions. Specifically, many scenarios feature paths that appear to represent obvious solutions, but ultimately lead to bad outcomes. Through replaying and selecting new options, the player explores the social concerns encompassed in a given RCR scenario, which will lead to a richer understanding of the ethical complications that one can encounter while conducting research as well as aid future moral reasoning.

4 METHODOLOGY

RCR is a complicated topic to teach that requires understanding a variety of perspectives and dilemmas that impact research ethics [21, 60]. As a result, we wanted to evaluate whether a choice-based interactive storytelling design, such as the one employed in *Academical*, could prove more effective than traditional approaches for teaching ethically complex topics. We hypothesized that the choice-based, role-playing nature of *Academical*—which is specifically designed to highlight how research ethics can be complicated by many factors such as power dynamics and marginalized identities—would be 1) more engaging, 2) as effective as traditional RCR educational materials at developing knowledge of RCR concepts, and 3) result in stronger moral reasoning skills. In order to explore these hypotheses, we conducted a between-subjects study comparing our choice-based interactive storytelling game approach with web-based educational materials from an existing RCR course

Unit 1 The Protection of Human Subjects

Claude Bernard

Claude Bernard (July 12, 1813 — February 10, 1878), a French physiologist, is often called the father of modern physiology. In 1865, he articulated his own standard of ethical conduct in the following quote:

> "The principle of medical and surgical morality consists in never performing on man an experiment which might be harmful to him to any extent, even though the result might be highly advantageous to science, i.e., the health of others."

This is not a standard that we meet today. There are risks involved with participating in modern research, which range from unintended disclosure of sensitive or personal information to side effects of experimental drugs. It is important to protect yourself, as well as other individuals that are involved in research studies. Bernard's central idea of minimizing risk to humans is the basis on which modern ethical conduct of research is based.
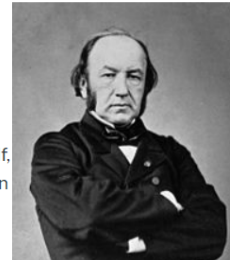
Figure 3: An excerpt from the traditional web-based educational materials used in this study. As is common with current educational RCR tools, the material is more heavily focused on historical context and case studies than *Academical*. These materials were borrowed from an existing university RCR course.

(see Figure 3). The study consisted of two conditions: 1) a group that read through two modules of the web-based educational RCR materials covering peer review and authorship; and 2) a group that played two chapters of *Academical* covering peer review and authorship content.

4.1 Procedure

Participants were told that the study was to explore different approaches to RCR education, and that they would either play a game or read materials teaching selected RCR concepts. They then completed an online survey collecting demographic information (age, prior gaming experience, prior RCR experience, and so forth). Upon completing the survey, participants were randomly assigned to one of the two conditions (web materials or *Academical*). After completing the RCR training for peer review and authorship, participants then completed a post-test that assessed their 1) engagement with the training material, 2) quantitative knowledge of peer review and authorship RCR concepts and 3) qualitative moral reasoning skills for these same concepts. All participants completed the same topics in the same order for both the training and testing phases.

4.2 Participants

A convenience sample of 28 university graduate and undergraduate students—the standard target populations for RCR training—were recruited for the study (age: μ=24.8, σ=7.6). There were 10 female, 14 male, and 3 non-binary participants, with 1 declining to disclose gender. During the study, participants were randomly assigned to one of the two conditions: web materials (**14 total**; 3 female, 2 non-binary, 8 male, 1 decline to answer) and *Academical* game (**14 total**; 7 female, 1 non-binary, 6 male). None of the participants reported prior RCR training within the past 2 years.

4.3 Measures

4.3.1 Temple Presence Inventory, Engagement Subscale. Engagement is an critical aspect of the learning process [22], drastically

influencing a learner's motivation to continue interacting with a system and the educational content [42]. In order to assess participant engagement with the two educational RCR tools employed, we utilized the Engagement subscale of the Temple Presence Inventory (TPI) [27]. The TPI is an instrument that has been validated for use with games [28] and measuring game engagement [30].

4.3.2 Peer Review and Authorship RCR Quizzes. To assess and compare how effective the two RCR tools were for teaching knowledge of peer review and authorship concepts, we utilized two quizzes from the existing online RCR course at the University of Utah. Each quiz consists of three questions around a respective topic, and each question is either true/false, yes/no, or multiple choice (see Appendix A).

4.3.3 Qualitative Assessment of Moral Reasoning. To assess and compare how effective the two RCR tools were for teaching moral reasoning skills, we utilized qualitative test materials from a previous study that evaluated the effect of role-play on RCR learning outcomes [59]. These test materials included two RCR-themed short stories obtained from the Online Ethics Center for Engineering and Research (OEC; https://www.onlineethics.org, Appendix B) and three short answer questions that the previous study designed to characterize a student's ability to 1) analyze a moral problem, 2) consider the viewpoints of all individuals involved, and 3) propose solutions and anticipate their possible short- and long-term consequences. Participants first read and wrote responses to the short story about peer review, then answered the same three questions for the other scenario involving authorship. After completion of the study, two of the authors scored these answers using the behaviorally anchored rating scale (BARS) method (see Figure 4). The coders initially used the same rubric described in the previous study to separately evaluate all answers, then compared results to assess score distributions and inter-rater reliability. Similar to the previous study, it was necessary to relax some grading criteria for questions that rarely received "ideal" answers. Using these updated rubrics,

Identify Issues	Representative Response
1: Indicates that there is no problem, or states that there is a simple disagreement amongst the parties.	"Mike and Lisa are not clear about the partnership."
3: Misses some of the moral issues present in the case. Primarily restates the issues as presented in the case without naming the issue or mentioning specific standards.	"The main issue in this scenario is that not everyone who worked on the experiment is getting the credit they deserve. Mike was convinced by his adviser that he should take the credit because it would further his career."
5: Accurately identifies and names most or all of the moral issues present in the case. If applicable, mentions relevant standards.	"Mike failed to make a more meaningful impact with his paper because he decided to submit the paper as sole author. Although one might argue that being both the designer and experimenter of a paper is more prestigious, it is not worth sacrificing your vision and purpose by removing the experiment which gives it validation among the scientific community. In addition, he tried to take credit for the work that Lisa did for their project, which is definitely unacceptable."

Describe Viewpoints	Representative Response
1: Primarily restates the behaviors of the parties involved as they are given in the case; states that there is no excuse for the behavior of one or more of the parties.	"Slater was asked to review a manuscript from competitor's lab, he thinks he could be objective and shared the manuscript with his student."
3: Explains at least two viewpoints. However, the focus is either primarily on the interest of only one of the parties involved, or the student indicates that the parties involved are entitled to their opinions but that one perspective is "more correct" than other perspectives without providing justification.	"Slater stands to benefit by sabotaging the competitor's work, but both Slater and Parker could possibly tarnish their reputations if this is exposed in the science world. The authors who submitted the manuscript that was rejected are just being completely screwed over."
5: Presents a balanced view from the perspective of several involved parties. States the different attitudes, values, and possible motives of the parties without making unfounded assumptions about intent.	"The first viewpoint is from the professor's perspective; he thinks that he can review the paper objectively despite the circumstances. The second viewpoint is from the grad student, whose professor put them in a compromising position. The third viewpoint is from the authors of the paper who received a reject review from a competing lab that also took a tip from their paper. The fourth is from the Journal of Cool Results that thought they were getting an objective review from the professor, but really received a biased reject."

Propose Solutions	Representative Response
1: Solution is to ignore the problem, to interfere or "go behind someone's back", or act immediately without considering whether this is the best course of action. Student does not mention, or devalues, the undesirable consequences of the chosen solution.	"Unfortunately, this is unavoidable. Slater and Parker were aware of the rules of conduct for peer reviewing, and they chose to subvert them. Any sense of competition will incite this kind of behavior. However, given that peer reviews often summon multiple people to provide feedback, I think that the quality of a work will be recognized by the majority."
3: Solution is practical, but incomplete or vaguely formulated. Student understands some of the consequences of the proposed solution but does not propose strategies for minimizing these consequences.	"I believe Slater and Parker should withdraw their statement of the manuscript since it is biased, and either credit or not use the solution found by the competition's research. Not using the solution may not be that simple, but if they do then they need to credit where they found the idea from. Additionally, they should refrain from responding to research that is bias on their end in the future. There was clear conflict of interest, and it should be addressed instead of agreeing to do the research."
5: Solution is practical and directly addresses the issues at hand. Solution aims to optimize the outcomes of all parties involved and to maintain relationships and reputations. Solution adopts standard best practices and does not violate ethical standards. Student understands the consequences of the solution and mentions strategies for minimizing negative consequences.	"Prof. Slater should write back to the Journal of Cool Results with his feedback, along with a description of his situation regarding his current work and the conflict of interest. Prof. Slater and Ms. Parker might want to contact the author directly for permission to use the original author's work and discuss credit in their paper. When Prof. Slater and Ms. Parker publish their results, they should mention the original author as the person who came up with the technique. The Journal of Cool Results might find Prof. Slater to be unprofessional/unethical, leading to a stain on his image. If he mentioned "sharing of the paper with Ms. Parker" with the Journal, he might be barred from reviewing papers any further, and increased scrutiny in their current work. The original author might want more credit than what Prof. Slater and Ms. Parker want to share, according to original author's perception of the contribution of his technique in their work."

Figure 4: Initial BARS rubric for scoring qualitative answers and representative responses. The left column is taken directly from [59] while the right column provides representative responses from our study participants. The final rubric was applied similarly to both of the RCR topics.

Table 1: Post-test results for the TPI Engagement subscale, Peer Review test, and Authorship test. The table contains mean scores, standard deviations, t-test and Wilcoxon rank sum scores for significance, and effect size—which is medium to large for significant differences.

| | Quantitative Test Results | | | | | | |
| | Web | | Game | | Sig | ES | |
Measures	μ	σ	μ	σ	*p*	*d*	*r*
TPI Engagement	23.4	9	30.1	6.1	**.029**	**.87**	**.4**
Peer Review Test	2.14	0.77	2.93	0.27	**.002**	**1.4**	**.56**
Authorship Test	2.36	0.75	2	0.79	.23	-.47	-.23

Table 2: Post-test results for the qualitative assessment of moral reasoning. The table contains mean scores, standard deviations, Wilcoxon rank sum test scores for significance, and effect size—which is medium to large for significant differences.

| | Qualitative Test Results | | | | | | |
| | Web | | Game | | Sig | ES | |
Measures	μ	σ	μ	σ	*p*	*d*	*r*
Identify Issues	6.93	1.9	8.57	1.6	**.023**	**.92**	**.42**
Describe Viewpoints	4.71	2.8	7.36	2.5	**.016**	**.99**	**.44**
Propose Solutions	4.71	2.3	7.14	2.3	**.015**	**1.1**	**.47**
Total Score	16.4	5.7	23.1	4.7	**.004**	**1.3**	**.54**

the coders again separately scored all answers and then met to discuss rationale for any discrepancies. In the end, the scores for each of the six questions had good inter-rater reliability, with acceptable levels of percent agreement (ranging .82-1) and Cohen's kappa values (ranging .72-1). Final scores for the few unresolved ratings were calculated as the average of the two coders' scores.

5 RESULTS

In this section, we provide the results of our study in terms of participant prior knowledge and experience, as well as differences between the two conditions with regard to engagement with the materials and learning outcomes.

5.1 Prior Knowledge and Experience

According to a series of independent samples t-tests, participants in the two conditions did not differ with respect to age, prior game experience, or prior interactive story experience (all p values >= .12). Similarly, no participants reported prior RCR training in the past 2 years. Therefore, we can assume that participants in both groups had similar prior RCR, game, and interactive story experience.

5.2 Engagement with RCR Training Tools

We first examine participant engagement between the different RCR educational tools. In order to analyze differences between the web materials and *Academical* game conditions, we used an independent samples t-test. The first row of Table 1 shows descriptive statistics for scores on the TPI Engagement subscale, as well as significant differences and effect sizes. Results found a significant difference in favor of *Academical* increasing participant engagement (p = .029, r = .4), suggesting that a choice-based interactive story game is a more engaging experience for RCR training than traditional web reading materials.

5.3 RCR Learning Outcomes

5.3.1 Peer Review and Authorship RCR Quizzes. To better understand participants' knowledge of RCR concepts, we analyzed post-test scores on the RCR peer review and authorship quizzes. Descriptive statistics, statistical significance, and effect sizes for the two measures are shown in the bottom two rows of Table 1. A series of Wilcoxon rank sum tests showed that participants in the *Academical* condition scored significantly higher on the peer review test (p =

.002, r = .56) and neither significantly better or worse than the web materials for the authorship test (n.s., p = .23). This suggests that, in terms of short-term learning, a choice-based interactive story approach is more effective than traditional educational materials for developing knowledge of certain RCR topics.

5.3.2 Qualitative Assessment of Moral Reasoning. To better understand participants' moral reasoning skills, we analyzed a series of qualitative responses they wrote evaluating multiple aspects of two scenarios addressing either peer review or authorship concepts (see Appendix B). Descriptive statistics, statistical significance, and effect sizes for these measures are shown in Table 2. A series of Wilcoxon rank sum tests showed that participants in the *Academical* group scored significantly higher overall on the qualitative tests of moral reasoning (total score: p = .004, r = .54). Combining the scores across the two scenarios revealed that these participants had similarly significant improvements for all three aspects of moral reasoning (Issues: p = .023, r = .42; Viewpoints: p = .016, r = .44; Solutions: p = .015, r = .47). The *Academical* group also demonstrated better moral reasoning skills all together for both scenarios (Peer Review: p = .019, r = .44; Authorship: p = .005, r = .53; see Figure 5). These results indicate that, in terms of short-term learning, a choice-based interactive story approach is more effective than traditional educational RCR materials for developing moral reasoning skills necessary to properly employ RCR.

6 DISCUSSION

The results from this study suggest that a choice-based interactive story game design is effective as an RCR education tool, with learners developing significantly higher engagement, stronger overall moral reasoning skills, and better knowledge scores for certain RCR topics with neither significantly better or worse scores for others. Results from our study highlight the potential of choice-based interactive storytelling games for improving student engagement and learning outcomes within RCR education as a whole. We discuss our results in more detail below.

6.1 Engagement with RCR Training

An independent samples t-test for the TPI Engagement subscale showed that *Academical* was significantly more engaging than traditional web-based RCR educational materials. This confirmed our
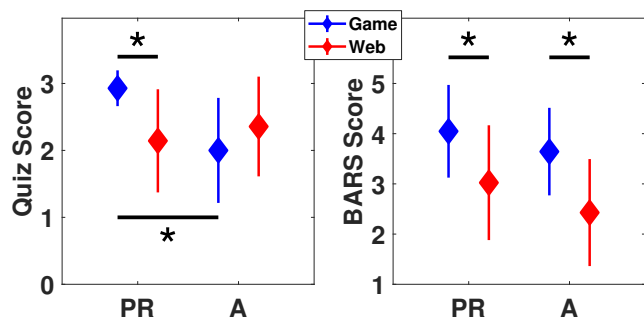
Figure 5: Post-test results for the peer review (PR) and authorship (A) scenarios. Left: The *Academical* group (n = 14, shown in blue) demonstrated significantly better knowledge scores for PR, and no statistical differences on knowledge scores for A. Right: The *Academical* group also demonstrated better moral reasoning skills for both scenarios. Diamonds represent group average scores and error bars indicate SD. Significance was determined by Wilcoxon rank sum tests and Wilcoxon signed rank tests where appropriate and is noted as *p<.05.

first hypothesis, and also falls in line with existing claims [7, 10, 24, 44, 61] and findings [53, 67, 73] that interactive storytelling designs can improve learner engagement and motivation. Additionally, we further extend these findings to illustrate that interactive storytelling games can also increase motivation when learning more ethically complex and ambiguous content—beyond the generally rote material covered in existing STEM [53, 73] and history [7] examples.

6.2 RCR Learning Outcomes

Our study also identified that short term quantitative learning outcomes for knowledge of RCR concepts in *Academical* was neither significantly better or worse for Authorship, and was significantly better for Peer Review. This serves to extend current findings on the learning outcomes of educational interactive storytelling games [17, 53, 67, 70, 73] by providing evidence for the efficacy of such games in teaching knowledge of RCR concepts and ethical decision-making. This confirmed, and even outperformed, our second hypothesis that interactive storytelling games would be as effective as traditional educational materials at developing knowledge of RCR concepts.

Additionally, and perhaps most importantly, we found that the *Academical* group performed significantly better overall on qualitative tests assessing moral reasoning skills. We also found a consistent significant increase for the *Academical* group's performance across all three questions addressing the different aspects of moral reasoning. As a whole, this suggests that a choice-based interactive narrative approach also better prepares students to navigate the key aspects of moral dilemmas and ethical decision-making that are common in research. These medium to large effect sizes (Table 2) indicate that *Academical* may also provide a substantial improvement over existing live action immersive role-play techniques for improving moral reasoning and knowledge of RCR concepts [59].

However, this needs to be further verified through additional studies. Overall, these results are very encouraging considering that various other (non role-playing) educational methods that have been shown to improve knowledge of RCR concepts often report either comparatively weak benefits, no effect, or even harm to moral reasoning skills [1, 11, 46, 58]. Conversely, *Academical* appears to have a significant impact on improving both players' knowledge of RCR concepts and their moral reasoning skills—providing a marked improvement over most existing RCR training tools.

These positive outcomes are also particularly impressive and interesting considering that the traditional web-training course was designed to teach the specific knowledge tested in this study's quantitative quizzes. In comparison, *Academical* immersed players in moral dilemmas that did not explicitly provide instruction about correct moral behavior or RCR concepts, yet they almost always performed better than the web-trained group on the same tests. Future studies are required to determine why *Academical* is seemingly able to provide these strong benefits.

6.3 Relative Difficulty of RCR Topics

Different RCR topics will vary in their perceived complexity, moral ambiguity, and professional relevance. Therefore, applying the same pedagogical methods to widely different subject matter is not guaranteed to be equally effective at teaching those topics [39, 68]. Comparing test results between topics can help educators better understand what information is being taught most effectively. In this study, all participants were trained and tested exclusively on two common yet distinct RCR concepts, peer review and authorship. We showed that the *Academical* group significantly outperformed the web-trained group for both knowledge and moral reasoning skills related to peer review content, demonstrating that the game was the superior tool for teaching that topic. In comparison, while the *Academical* group also did significantly better than the web-trained group on qualitative tests of moral reasoning related to authorship, playing the game did not provide a similar boost to knowledge of the subject. This result suggests that *Academical* participants may have generally performed better on tests about peer review, but struggled as much as the traditional RCR educational approach when learning concepts related to authorship—which could indicate differences in pedagogical efficacy. In order to further explore these results and any potential differences in pedagogical efficacy, we conducted a secondary analysis comparing participants' performance between the two RCR topics.

For the quantitative results assessing RCR knowledge, we used non-parametric signed rank tests to analyze within-subject quiz results for each training group. We found that the *Academical* group had significantly better scores for knowledge of peer review than for authorship (p = .0007, r = .62), while the web-trained group's results were comparable between the two topics (p = 0.47, r = .14). Considering that the *Academical* group's peer review knowledge scores were also significantly better than those of the web-trained group (p = .002, r = .56), this suggests that *Academical* was more effective at teaching peer review material compared to the authorship content (Figure 5, left).

For the qualitative results assessing moral reasoning skills, we first calculated each participant's average BARS score for the peer

review and authorship test scenarios and compared training group performance within each topic. Independent-samples t-tests revealed significantly higher overall moral reasoning scores for the *Academical* group in both scenarios (Figure 5, right; Peer Review: p = .015, r = .44; Authorship: p = .0028, r = .53). However, this result does not indicate if participants performed better or worse on one topic over another within the *Academical* or web groups. Therefore, in order to explore any differences in the pedagogical efficacy of developing moral reasoning skills for different scenarios, we conducted a paired-sample t-test across all participants and compared their scores between the two RCR topics. The aggregated scores showed that everyone had significantly better moral reasoning skills for the peer review scenario than the authorship scenario (All participants: p = .012, r = .21). Performing paired t-tests at the group level also found that each group individually showed a similar but not significant trend towards better overall scores on peer review than authorship (*Academical*: p = 0.098, r = .22; Web: p = 0.068, r = .26). These results suggest that teaching moral reasoning skills with respect to the RCR topic of authorship is harder to teach than peer review overall—regardless of which educational RCR tool was used. However, because everyone read and responded to the two test scenarios in the same order (with authorship last), it is still somewhat unclear whether these overall differences in moral reasoning skill between the two RCR concepts are due to pedagogical efficacy or simply performance fatigue.

Considered together, the results of this secondary analysis support the idea that *Academical* is more effective at teaching the tested peer review material over the authorship content, and that authorship content is substantially more difficult to teach in general regardless of the tool. Future work is required to more explicitly explore this and the relative difficulty of all other RCR topics. Overall, these observations at least indicate that the quantitative and qualitative test measures used for this study are sensitive enough to detect significant differences in performance across topics after training.

6.4 Role-Playing

Given that a choice-based interactive storytelling design approach is both more engaging than traditional RCR materials and equally/more effective for both quantitative and qualitative learning outcomes, *Academical* is ultimately a useful tool to address the engagement and critical thinking needs of current RCR education [19]. Studies have shown that live action interactive role-play can help students practice moral reasoning skills, but when compared to playing a computer game, it is a relatively resource-intensive activity in terms of the time and energy needed to facilitate and evaluate the training process. Furthermore, role-playing with others in the physical world can be an uncomfortable experience for some people, potentially compromising the learning experience [59]. In comparison, *Academical* is an engaging single-player role-playing experience that carries no social pressure, allowing students to explore multiple perspectives at their own pace. Furthermore, its digital nature means that all students can play through the same training scenarios with the same dialogue options, and consequently their learning progress and progression through the stories can be tracked far

more easily. Critically, the improved convenience of using *Academical* for ethical training has the potential to reach a far broader audience than live action role-playing, as well as enable larger and more controlled studies of its effects on RCR learning outcomes.

7 STUDY LIMITATIONS AND FUTURE WORK

Despite *Academical*'s encouraging effect on engagement and RCR learning outcomes, one notable limitation of this study is the small sample size of participants. Additionally, the training and test procedures were not randomized, making it somewhat more difficult to explain differences in performance between topics. Furthermore, this study only measured short-term learning resulting from a single session of training. Overall, these positive results are quite valuable given the relatively poor state of current RCR education [19, 20], but further work is needed to assess long-term skill retention and engagement. Another potential disadvantage is that, since we did not test an untrained group of participants, we could not report how much of an effect RCR training in general had on our learning outcomes. Finally, improvements in the above learning outcomes do not necessarily lead to better attitudes or moral behavior [48], therefore the impact of *Academical* on such factors needs to be explored in future studies as well.

Specifically, future work will include longitudinal studies that measure long-term learning outcomes and improvements to RCR practices over time. We plan to achieve this by embedding the game content into relevant university courses. Future studies will also examine whether improvements in RCR learning outcomes from training with *Academical* can generalize to untrained content. We are also interested in determining which design aspects of *Academical* best contribute to learning, and similarly want to better understand how different player types can affect engagement and learning outcomes. Lastly, future studies will include an additional assessment to determine how different training methods affect a player's attitude about the importance of moral conduct in research.

8 CONCLUSION

In this paper we described the design of *Academical*, a choice-based interactive storytelling game for RCR education that enables players to experience a story from multiple perspectives. We also presented results from an initial study comparing *Academical* with traditional web-based educational materials from an existing university RCR course. The initial study results highlighted that a choice-based interactive story game design is effective for an RCR education tool—with significantly higher engagement, better scores overall for qualitative tests of moral reasoning skills, and significantly better scores for some quantitative tests of RCR knowledge and neither significantly better or worse scores for others.

REFERENCES

[1] Alison L. Antes, Xiaoqian Wang, Michael D. Mumford, Ryan P. Brown, Shane Connelly, and Lynn D. Devenport. 2010. Evaluating the effects that existing instruction on responsible conduct of research has on ethical decision making. *Academic Medicine* 85, 3 (2010), 519–26.

[2] Ruth S Aylett, Sandy Louchart, Joao Dias, Ana Paiva, and Marco Vala. 2005. FearNot!–an experiment in emergent narrative. In *International Workshop on Intelligent Virtual Agents*. Springer, 305–316.

[3] Mathieu Bouville. 2008. On using ethical theories to teach engineering ethics. *Science and Engineering Ethics* 14, 1 (2008), 111–120.

[4] Tharrenos Bratitsis. 2016. A digital storytelling approach for fostering empathy towards autistic children: Lessons learned. In *Proc. International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*. 301–308.

[5] Bradley J Brummel, CK Gunsalus, Kerri L Anderson, and Michael C Loui. 2010. Development of role-play scenarios for teaching responsible conduct of research. *Science and Engineering Ethics* 16, 3 (2010), 573–589.

[6] Janelynn Camingue, Edward F. Melcer, and Elin Carstensdottir. 2020. A (Visual) Novel Route to Learning: A Taxonomy of Educational Visual Novels. In *Proceedings of the 15th International Conference on the Foundations of Digital Games*.

[7] Dimitrios Christopoulos, Pavlos Mavridis, Anthousis Andreadis, and John N Karigiannis. 2011. Using Virtual Environments to Tell the Story:" The Battle of Thermopylae". In *2011 Third International Conference on Games and Virtual Worlds for Serious Applications*. IEEE, 84–91.

[8] Polina Danilicheva, Stanislav Klimenko, Yury Baturin, and Alexander Serebrov. 2009. Education in virtual worlds: Virtual storytelling. In *2009 International Conference on CyberWorlds*. IEEE, 333–338.

[9] Lucas Pfeiffer Salomâo Dias, Jorge Luis Victoria Barbosa, and Henrique Damasceno Vianna. 2018. Gamification and serious games in depression care: A systematic mapping study. *Telematics and Informatics* 35, 1 (2018), 213–224.

[10] Michele D Dickey. 2006. Game design narrative for learning: Appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. *Educational Technology Research and Development* 54, 3 (2006), 245–263.

[11] James M. DuBois, Jeffrey M. Dueker, Emily E. Anderson, and Jean Campbell. 2008. The development and assessment of an NIH-funded research ethics training program. *Academic Medicine* 83, 6 (2008), 596–603.

[12] Jane Friedhoff. 2013. Untangling Twine: A Platform Study. In *Proc. DiGRA*.

[13] Stefan Göbel, André de Carvalho Rodrigues, Florian Mehm, and Ralf Steinmetz. 2009. Narrative game-based learning objects for story-based digital educational games. *narrative* 14 (2009), 16.

[14] Stefan Göbel and Florian Mehm. 2013. Personalized, adaptive digital educational games using narrative game-based learning objects. In *Serious Games and Virtual Worlds in Education, Professional Development, and Healthcare*. IGI Global, 74–84.

[15] Matthew Guzdial, Brent Harrison, Boyang Li, and Mark Riedl. 2015. Crowdsourcing Open Interactive Narrative.. In *Proc. Foundations of Digital Games*.

[16] Rania Hodhod, Paul Cairns, and Daniel Kudenko. 2011. Innovative integrated architecture for educational games: challenges and merits. In *Transactions on edutainment V*. Springer, 1–34.

[17] Rania Hodhod, Daniel Kudenko, and Paul Cairns. 2009. AEINS: Adaptive Educational Interactive Narrative System to Teach Ethics. In *AIED 2009: 14 th International Conference on Artificial Intelligence in Education Workshops Proceedings*. 79.

[18] Michael Kalichman. 2013. A brief history of RCR education. *Accountability in Research* 20, 5-6 (2013), 380–394.

[19] Michael Kalichman. 2014. Rescuing responsible conduct of research (RCR) education. *Accountability in research* 21, 1 (2014), 68–83.

[20] Michael Kalichman. 2016. Responsible Conduct of Research Education (What, Why, and Does It Work?). *Academic medicine: journal of the Association of American Medical Colleges* 91, 12 (2016), e10.

[21] Michael W Kalichman and Dena K Plemmons. 2007. Reported goals for responsible conduct of research courses. *Academic Medicine* 82, 9 (2007), 846–852.

[22] Greg Kearsley and Ben Shneiderman. 1998. Engagement theory: A framework for technology-based teaching and learning. *Educational technology* 38, 5 (1998), 20–23.

[23] Oleksandra Keehl and Edward Melcer. 2019. Radical tunes: exploring the impact of music on memorization of stroke order in logographic writing systems. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*. 1–6.

[24] Michael D Kickmeier-Rust, Stefan Göbel, and Dietrich Albert. 2008. 80Days: Melding adaptive educational technology and interactive storytelling in digital educational games. In *Proceedings of the First International Workshop on Story-Telling and Educational Games (STEG'08)*. 8.

[25] Vykinta Kligyte, Richard T Marcy, Sydney T Sevier, Elaine S Godfrey, and Michael D Mumford. 2008. A qualitative approach to responsible conduct of research (RCR) training development: Identification of metacognitive strategies. *Science and Engineering Ethics* 14, 1 (2008), 3–31.

[26] Hartmut Koenitz, Gabriele Ferri, Mads Haahr, Diğdem Sezen, and Tonguç İbrahim Sezen. 2015. *Interactive digital narrative: history, theory and practice*. Routledge.

[27] Matthew Lombard, Theresa B Ditton, and Lisa Weinstein. 2009. Measuring presence: the temple presence inventory. In *Proceedings of the 12th Annual International Workshop on Presence*. 1–15.

[28] Matthew Lombard, Lisa Weinstein, and Theresa Ditton. 2011. Measuring telepresence: The validity of the Temple Presence Inventory (TPI) in a gaming context. In *ISPR 2011: The International Society for Presence Research Annual Conference*. Edinburgh.

[29] Brian Magerko. 2007. Evaluating preemptive story direction in the interactive drama architecture. *Journal of Game Development* 2, 3 (2007), 25–52.

[30] Rosa Mikeal Martey, Kate Kenski, James Folkestad, Laurie Feldman, Elana Gordis, Adrienne Shaw, Jennifer Stromer-Galley, Ben Clegg, Hui Zhang, Nissim Kaufman, et al. 2014. Measuring game engagement: multiple methods and construct complexity. *Simulation & Gaming* 45, 4-5 (2014), 528–547.

[31] Peter Mawhorter, Michael Mateas, Noah Wardrip-Fruin, and Arnav Jhala. 2014. Towards a theory of choice poetics. In *Proc. Foundations of Digital Games*.

[32] Peter Mawhorter, Carmen Zegura, Alex Gray, Arnav Jhala, Michael Mateas, and Noah Wardrip-Fruin. 2018. Choice poetics by example. In *Arts*, Vol. 7. Multidisciplinary Digital Publishing Institute, 47.

[33] Peter Andrew Mawhorter. 2016. *Artificial intelligence as a tool for understanding narrative choices*. Ph.D. Dissertation. UC Santa Cruz.

[34] Edward F Melcer, Victoria Hollis, and Katherine Isbister. 2017. Tangibles vs. mouse in educational programming games: Influences on enjoyment and self-beliefs. In *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*. 1901–1908.

[35] Edward F Melcer and Katherine Isbister. 2018. Bots & (Main) frames: exploring the impact of tangible blocks and collaborative play in an educational programming game. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.

[36] Edward F. Melcer, Truong-Huy D. Nguyen, Zhengxing Chen, Alessandro Canossa, Magy Seif El-Nasr, and Katherine Isbister. 2015. Games Research Today: Analyzing the Academic Landscape 2000-2014. In *Proceedings of the 10th International Conference on the Foundations of Digital Games*.

[37] Edward F. Melcer, James Ryan, Nick Junius, Max Kreminski, Dietrich Squinkifer, Brent Hill, and Noah Wardrip-Fruin. 2020. Teaching Responsible Conduct of Research Through an Interactive Storytelling Game. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*.

[38] Brian Moriarty. 2015. I Sing the Story Electric. PRACTICE. http://ludix.com/moriarty/electric.html Accessed Jan 2 2020.

[39] Michael D. Mumford, Shane Connelly, Ryan P. Brown, Stephen T. Murphy, Jason H. Hill, Alison L. Antes, Ethan P. Waples, and Lynn D. Devenport. 2008. A sensemaking approach to ethics training for scientists: Preliminary evidence of training effectiveness. *Ethics and Behavior* 18, 4 (2008), 315–339.

[40] John Thomas Murray. 2018. *Telltale Hearts: Encoding Cinematic Choice-based Adventure Games*. Ph.D. Dissertation. UC Santa Cruz.

[41] Truong-Huy D Nguyen, Edward Melcer, Alessandro Canossa, Katherine Isbister, and Magy Seif El-Nasr. 2018. Seagull: A bird's-eye view of the evolution of technical games research. *Entertainment computing* 26 (2018), 88–104.

[42] Heather L O'Brien and Elaine G Toms. 2008. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology* 59, 6 (2008), 938–955.

[43] National Institutes of Health et al. 1992. Reminder and update: requirement for instruction in the responsible conduct of research in national research service award institutional training grants. *NIH Guide for Grants and Contracts* 21, 43 (1992).

[44] Natalia Padilla-Zea, Francisco L Gutiérrez, José Rafael López-Arcos, Ana Abad-Arranz, and Patricia Paderewski. 2014. Modeling storytelling to be used in educational video games. *Computers in Human Behavior* 31 (2014), 461–474.

[45] Nathan Partlan, Elin Carstensdottir, Erica Kleinman, Sam Snodgrass, Casper Harteveld, Gillian Smith, Camillia Matuk, Steven C Sutherland, and Magy Seif El-Nasr. 2019. Evaluation of an automatically-constructed graph-based representation for interactive narrative. In *Proc. Foundations of Digital Games*. ACM.

[46] Dena K. Plemmons, Suzanne A. Brody, and Michael W. Kalichman. 2006. Student perceptions of the effectiveness of education in the responsible conduct of research. *Science and Engineering Ethics* 12, 3 (2006), 571–82.

[47] SH Plimpton. 2009. NSF's Implementation of Section 7009 of the America COMPETES Act. *Fed. Regist* 74, 160 (2009), 42126–42128.

[48] Sean T. Powell, Matthew A. Allison, and Michael W. Kalichman. 2007. Effectiveness of a responsible conduct of research course - a preliminary study. *Science and Engineering Ethics* 13, 2 (2007), 249–64.

[49] Teddy Pozo. 2018. Queer games after empathy: Feminism and haptic game design aesthetics from consent to cuteness to the radically soft. *Game Studies* 18, 3 (2018).

[50] Mark O Riedl, Andrew Stern, Don Dini, and Jason Alderman. 2008. Dynamic experience management in virtual worlds for entertainment, education, and training. *International Transactions on Systems Science and Applications, Special Issue on Agent Based Systems for Human Learning* 4, 2 (2008), 23–42.

[51] Jake Rossen. 2014. A Brief History of "Choose Your Own Adventure". *Mental Floss* (Apr. 10 2014).

[52] Jonathan P Rowe and James C Lester. 2010. Modeling user knowledge with dynamic Bayesian networks in interactive narrative environments. In *Sixth Artificial Intelligence and Interactive Digital Entertainment Conference*.

[53] Jonathan P Rowe, Lucy R Shores, Bradford W Mott, and James C Lester. 2011. Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education* 21, 1-2 (2011), 115–133.

[54] James Ryan. In press. A Garden, A Forking Path: Interactive Branching Narrative in *The Lady of May* (1578). In *Proc. 1st Workshop on the History of Expressive Systems*.

[55] Anastasia Salter. 2014. *What is your quest?: from adventure games to interactive books*. University of Iowa Press.

[56] Anastasia Salter. 2016. Playing at empathy: Representing and experiencing emotional growth through Twine games. In *Proc. International Conference on Serious Games and Applications for Health*. IEEE, 1–8.

[57] Ben Samuel, Jacob Garbe, Adam Summerville, Jill Denner, Sarah Harmon, Gina Lepore, Chris Martens, Noah Wardrip-Fruin, and Michael Mateas. 2017. Leveraging procedural narrative and gameplay to address controversial topics. In *Proc. International Conference on Computational Creativity*.

[58] Karen B Schmaling and Arthur W Blume. 2009. Ethics instruction increases graduate students' responsible conduct of research knowledge but not moral reasoning. *Accountability in Research* 16, 5 (2009), 268–283.

[59] Stephanie N Seiler, Bradley J Brummel, Kerri L Anderson, Kyoung Jin Kim, Serena Wee, CK Gunsalus, and Michael C Loui. 2011. Outcomes assessment of role-play scenarios for teaching responsible conduct of research. *Accountability in Research* 18, 4 (2011), 217–246.

[60] Adil E Shamoo and David B Resnik. 2009. *Responsible conduct of research*. Oxford University Press.

[61] Qiulian Song, Ling He, and Xiaoqiang Hu. 2012. To improve the interactivity of the history educational games with digital interactive storytelling. *Physics Procedia* 33 (2012), 1798–1802.

[62] Ulrike Spierling. 2008. 'Killer Phrases': Design Steps for a Game with Digital Role-Playing Agents. In *Transactions on edutainment I*. Springer, 150–161.

[63] Katryna Starks, Dakoda Barker, and Alayna Cole. 2016. Using Twine as a Therapeutic Writing Tool for Creating Serious Games. In *Proc. Joint International Conference on Serious Games*. Springer, 89–103.

[64] Telltale Games. 2012. *The Walking Dead*. Telltale Games.

[65] Telltale Games. 2013. *The Wolf Among Us*. Telltale Games.

[66] Pengcheng Wang, Jonathan Rowe, Bradford Mott, and James Lester. 2016. Decomposing drama management in educational interactive narrative: A modular reinforcement learning approach. In *International Conference on Interactive Digital Storytelling*. Springer, 270–282.

[67] Scott Watson, Natalie Vannini, Megan Davis, Sarah Woods, Marc Hall, Lynne Hall, and Kerstin Dautenhahn. 2007. FearNot! an anti-bullying intervention: Evaluation of an interactive virtual learning environment. *Artificial Intelligence and Simulation of Behaviour (AISB), April* 24 (2007).

[68] Logan L. Watts, Kelsey E. Medeiros, Tyler J. Mulhearn, Logan M. Steele, Shane Connelly, and Michael D. Mumford. 2017. Are ethics training programs improving? A meta-analytic review of the past and present ethics instruction in the sciences. *Ethics and Behavior* 27, 5 (2017), 351–384.

[69] Sebastian A Weiß and Wolfgang Müller. 2008. The potential of interactive digital storytelling for the creation of educational computer games. In *International Conference on Technologies for E-Learning and Digital Entertainment*. Springer, 475–486.

[70] Jui-Feng Weng, Hsiu-lien Kuo, and Shian-Shyong Tseng. 2011. Interactive storytelling for elementary school nature science education. In *2011 IEEE 11th International Conference on Advanced Learning Technologies*. IEEE, 336–338.

[71] Peter Weyhrauch. 1997. *Guiding Interactive Fiction*. Ph.D. Dissertation. Carnegie Mellon University.

[72] Caroline Whitbeck. 2001. Group mentoring to foster the responsible conduct of research. *Science and Engineering Ethics* 7, 4 (2001), 541–558.

[73] Lei Zhang, Doug A Bowman, and Caroline N Jones. 2019. Exploring Effects of Interactivity on Learning with Interactive Storytelling in Immersive Virtual Reality. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, 1–8.

[74] Alexander Zook, Stephen Lee-Urban, Mark O Riedl, Heather K Holden, Robert A Sottilare, and Keith W Brawner. 2012. Automated scenario generation: toward tailored and optimized military training in virtual environments. In *Proceedings of the international conference on the foundations of digital games*. 164–171.

A STUDY MEASURES

A.1 Peer Review RCR Quiz

The post-test RCR peer review quiz questions. Questions were taken from an existing RCR course at the University of Utah:

(1) "According to the study materials, peer reviewers are asked to make judgements about the quality of a proposed or completed project. This certainly includes all EXCEPT the following:"
(Multiple Choice)

   (a) Making sure the conclusions are supported by the evidence presented.
   (b) Checking calculations and/or confirming the logic of important arguments.
   (c) Assessing whether the research methods are appropriate.
   (d) Confirming that the relevant literature has been consulted and cited.
   (e) Verifying the qualifications of graduate students.

(2) "If you can figure out the authors of a paper you are peer reviewing after conflicts of interest are disclosed, should you still review the paper"
(Yes/No)

(3) "There is no simple solution to the problem of bias in peer review. However, researchers can lessen the impact of bias by writing transparent reviews."
(True/False)

A.2 Authorship RCR Quiz

The post-test RCR authorship quiz questions. Questions were taken from an existing RCR course at the University of Utah:

(1) "When should authorship for a paper be discussed?"
(Multiple Choice)

   (a) Just before submitting the paper.
   (b) Before starting the paper.
   (c) Throughout the process of working on the paper.

(2) "Which of the following is NOT considered a contribution to a paper?"

(Multiple Choice)

(a) Drafting and editing manuscripts.
(b) Developing methodology.
(c) Defining problems.
(d) Presenting or interpreting theories or ideas.
(e) Being the leading researcher in the field.

(3) "There is disagreement over whether authorship should be limited to individuals who contributed to all phases of a publication or whether individuals who made more limited contributions deserve authorship credit."
(True/False)

A.3   Qualitative Moral Reasoning Assessment

The post-test RCR qualitative moral reasoning questions. Questions were phrased to match the content of the three scoring categories from the RCR BARS rubric described in [59]:

(1) "What are the issues in this scenario? Please write a paragraph."
(2) "Describe the various viewpoints in this scenario. Please write a paragraph."
(3) "What would you propose as a solution? What problems might occur in resolving these issues? Please write a paragraph."

B   STORIES FOR QUALITATIVE ASSESSMENT

B.1   Peer Review Scenario

Professor John Slater is supervising a research project conducted by Alice Parker, a graduate student in Slater's lab. Parker is trouble-shooting a protein purification protocol; she wants to use the protocol to purify a recombinant form of a mammalian protein growth factor expressed in bacteria. Parker needs the purified protein to complete the final experiment required to prove her experimental model. Parker and Slater intend to submit a manuscript based on this model to The Journal of Cool Results.

While Parker is trouble-shooting the protocol, The Journal of Cool Results sends Slater a manuscript to review; he is asked to return the manuscript with his comments and recommendation for publication. The manuscript turns out to be from a competitor's lab, and the title indicates that the work closely resembles the work Parker and Slater intend to publish.

Slater considers the situation. He decides that he can be objective in his review, and he proceeds to read and evaluate the manuscript. After his initial review, he asks Parker for her comments on the manuscript, as the work falls within her field of expertise. Slater and Parker agree that the data are not convincing and that the paper should not be accepted for publication. Slater returns the manuscript to the editor of The Journal of Cool Results, with his recommendation that it not be accepted for publication.

After reviewing the manuscript, Slater and Parker note that the authors use a recombinant form of the protein growth factor that they purified from yeast using a novel technique. Slater suggests that Parker apply this technique to her purification protocol. The revised protocol works well, and Parker is able to complete the final experiment.

Source: To Review or Not - Reviewing the Competition. https://www.onlineethics.org/Resources/gradres/gradresv2/review.aspx

B.2   Authorship Scenario

Mike is a bright, young post-doc working in a big research group in the physics department at Bambuka University. His life-long career goal is to conduct research in a leading research university as a professor. During one of his job interviews, he had a discussion about a particular problem in his field of expertise with the interviewer. In the course of the interview, he was not able to satisfactorily prove his point, because his theoretical arguments did not convince the interviewer. Upon returning to his lab, Mike decides to pursue the matter further and conduct an experiment to verify his argument. Mike's experimental background is not sufficient to obtain the desired results.

Lisa, Mike's friend, is a fourth-year graduate student working on her PhD in the same lab. She volunteers to help Mike with the experiment. Lisa is a talented experimentalist, and she successfully completes the experiment. Mike sends the results to the interviewer, thereby proving his point.

While working on this small experiment, Mike gets an idea for an interesting study, which, if done correctly, could yield a good publication in an important journal. But Mike is discouraged, because he knows he cannot handle the complicated experiment alone. Lisa encourages him to proceed with the idea and promises to design and complete the experimental aspect of the project. Mike agrees and while he is working on the theory, Lisa designs and builds the experiment. Mike is very excited about his theoretical results and shares them with his adviser. The professor likes Mike's ideas and tells him that it is time for Mike to get his name noticed in the scientific community. He encourages Mike to publish the results in a famous journal. The adviser also suggests that it would be better for Mike's career if he publishes the work in a single author paper. He says, "You worked on it exclusively, and it would be a wonderful opportunity to write a paper by yourself. It will be a stronger paper if you could validate your theory with experimental data."

Mike likes the suggestion. He doesn't mention that Lisa has already done a significant amount of work on the project. He tells Lisa that his adviser recommended his publishing results in a single author paper, and says, "I really think that this would help my career, plus that's what our adviser wants. How cheated will you feel if I publish this paper alone using all the data that your experiment provides?"

Lisa and Mike are good friends, and she feels obligated to help him. Even though Lisa is disappointed, she tells Mike to do whatever he feels is right. Mike decides to submit the paper as the sole author.

After this conversation, Lisa stops working on the experiment and Mike takes over. He did not design the experiment; therefore he cannot manage to get it to work and does not make any progress. Lisa does not offer any more help, and Mike doesn't ask her for any. Finally, Mike decides to submit the paper without the experimental part. It will be an interesting theoretical investigation, but it will not have the scientific impact that it could have had with the experimental validation.

Source: A Single Author Paper.
https://www.onlineethics.org/Resources/gradres/gradresv4/single.aspx