

Resolving Conflicts in Educational Game Design Through Playtesting

by Brian Winn and Carrie Heeter

The multidisciplinary nature of educational game design requires multidisciplinary teams for product development. Typically such teams consist of game designers, pedagogy experts, and content experts, each of whom must resolve significant and often fascinating ideological disagreements resulting from disparate disciplinary values, vocabulary, and culture. Game designers hope to create a highly interactive, compelling experience that is also fun to play. Pedagogy experts insist that the game must be an effective teacher. Content experts expect the game to include accurate, richly detailed content, ideally inspiring at least some of the passion they feel for their field of study. Game design teams that include experts in each of these areas face the challenge of coordinating their expertise to produce a product that ideally meets their respective priorities and criteria.

Mishra and Koehler (2006) define technological pedagogical content knowledge (TPCK) as an emergent form of knowledge that goes beyond the three individual components of technology, pedagogy, and content to yield a result that is more than the sum of its parts. This concept is particularly relevant for understanding the ultimate goal of educational game design as well as the challenges that collaborative teams face as they try to meet this goal. An award-winning game designer is highly unlikely to have a doctorate in education and be a leading paleobiologist at the same time, and is therefore unlikely to produce an educational game grounded in TPCK. Because each team member prioritizes different goals at the inception of the designing process, tensions between TPCK perspectives are necessary and important.

To be sure, such tensions will vary depending on the nature of the educational game in question. In the case of exogenous educational games, conflicts between pedagogy, content, and gameplay rarely arise since the game mechanics and pedagogy are already defined at the outset. Such games separate learning content and game mechanics (Halverson [2005](#); Malone and Lepper 1987). Designers of exogenous educational games typically reuse successful game mechanics, such as hangman, a Jeopardy-style game show, or a Space Invaders-style shooter, inserting the content to be learned into the preexisting game structure and rules. Content is often the only new input, and the pedagogy tends to be limited to reinforcing knowledge recall.

In contrast, endogenous educational games target more complex learning goals beyond memorization and do so in part by integrating learning content into the structure of the game (Halverson [2005](#)). Like exogenous games, endogenous games frequently adopt familiar game genres such as role play or adventure games, board or card games, or puzzles. The defining characteristic of endogenous games, however, is that the gameplay itself informs the pedagogical method and embodies the learning concept. By requiring players to explore the game space and to use their knowledge to meet game challenges, designers of endogenous games promote active problem-solving and reinforce context-specific learning goals. Consequently, endogenous educational game designers begin with a more or less blank slate. They seek an idealized convergence of content, pedagogy, and game mechanics that achieves the hypothetical potential of games to promote advanced forms of learning. This ill-specified design problem has infinite possible solutions. (For discussion of games and learning, see Prensky 2001; Gee 2003, 2005; Van Eck [2006](#); and Kirriemuir and McFarlane [2004](#).)

This article describes a sequence of conflicts encountered during the game design process and the way in which each was resolved through iterative playtesting to arrive at convergent, amicable game design features that balanced pedagogy, content, and gameplay. In outlining the step-by-step process whereby such conflicts were resolved, we hope to provide a collaborative, test-based model of educational game design that will be useful for other multidisciplinary teams as they face their own distinctive challenges.

Case Study

The Michigan State University ([MSU](#)) Games for Entertainment and Learning ([GEL](#)) Lab created a Web-based endogenous learning game designed to teach a subset of the National Science standards on adaptation and evolution (American Association for the Advancement of Science 1994; Center for Science, Mathematics, and Engineering Education 1996). Funded by the [National Science Foundation](#), the game targeted 7th through 9th grade students and was intended to be playable within a 45-minute class period. Our case study of this game design process illustrates how

playtesting resulted in progressive revisions that brought pedagogy, content, and gameplay goals into ever closer alignment.

At the 2005 Designing User eXperience ([DUX](#)) conference, we presented a detailed discussion of this study's game design and prototype-playtesting progression (Heeter, Winn, and Greene [2005](#)). In this article, we will highlight some of the key moments in resolving pedagogy-content-gameplay conflicts ([Exhibit 1](#)). In doing so, we will delineate a nine-step progression of design modifications (Figures 1-9 below) through three successive prototypes of the learning game (Exhibits 2-4 below). For clarity, we will outline the design process as though a single individual team member represents each individual area of expertise.

Preliminary Development and Design Considerations

To illustrate the dilemma of open-ended learning game design (which can also be enthusiastically described as "the opportunity"), consider an overly simplistic goal of creating a game to teach only one of the 10 national science standards related to evolution that our game targeted: "The basic idea of biological evolution is that present day species developed from earlier distinctly different species" (Center for Science, Mathematics, and Engineering Education 1996). Like the other national science standards, this relatively simple statement encapsulates a highly abstract concept. Identifying even a specific abstract conceptual learning goal does little to narrow the scope of game design, pedagogy, or specific science content possibilities.

The content and pedagogy experts easily converged on the desire to base the game on realistic, detailed exemplars of now-extinct creatures from Earth's past as a mechanism for teaching evolution concepts. The gameplay expert accepted this idea as a central design specification without objection ([Figure 1](#)).

The game designer initially included [levels](#) as an abstract design feature because they provide players with subgoals and a means to gauge their progress through the game. However, levels also create barriers to advancement in the game. If players are unable to beat level 1, they will not be allowed to play level 2. The pedagogy expert deemed this unacceptable. From her point of view, she did not want the game to prevent learners from accessing learning content. Therefore, levels conflicted with the pedagogy perspective, and the feature was eliminated ([Figure 2](#)).

Before the team converged sufficiently on any design concept to be able to create a prototype, one or more of the experts categorically vetoed other team members' game ideas. The game designer immediately dismissed game ideas proposed by the pedagogy expert and scientist because they did not constitute "a game." The scientist immediately rejected many of the game designer's ideas because the science was incorrect.

Most initial conflicts centered on the game's story. Two types of storytelling can occur in games: the designer's story and the player's story (Rouse 2001). The designer's story is the storytelling built into the game. Story can be used to set the stage (often called backstory), provide purpose and engagement, and convey content as well as other things. Endogenous games often use a designer's story that is closely tied to the learning content. The player's story is the story of the interactions and choices the players make as they play the game. Some games have stronger designer's stories, such as adventure and role-playing games, while others have little to no designer's story, such as classic arcade games like *Pacman* and puzzle games like *Tetris*. However, all games have a player's story.

At this early stage of development, the gameplay experts were proposing designers' stories that were rejected by the science experts because they did not reflect accurate science. After several attempts at formulating a designer's story, the consideration of story was temporarily shelved ([Figure 3](#)).

The collaboration between scientists who did not understand games and game designers who did not understand evolution was not working. The entire team delved into the science concepts, gaining a more sophisticated understanding of the content domain. The game designer also taught team members basic concepts about games and game design. Although mutual education helped, it did not always yield agreement. Therefore, the most hotly disputed design features were implemented into a playable prototype, and the first playtest was conducted.

First Prototype: The Critter Card Game

In accordance with the precepts of [iterative game design](#), we conceived of each prototype as a hypothesis to be tested, revised, and tested again. Designers arrived at the first playable prototype—*The Critter Card Game* or *CCG*—through a mixture of compromise and inspiration ([Exhibit 2](#)). The gameplay designer, pedagogy expert, and scientist agreed on a set of design concepts and features. *CCG* used a variation on familiar card game mechanics. Endogenous to the content domain, the game included one virtual deck of critter cards and one virtual deck of "Who am I" adaptation challenges. Players earned points by matching critter cards with adaptation challenges. Each critter card included extensive learning content about an animal from Earth's past. More rare adaptation challenges were worth more points. Players who wanted to maximize their score strategically selected high value matches. In theory, as players match critters and adaptations, successful matches exemplify national science standard concepts about evolution. In a typical early playtest scenario, designers recruited a small sample of relatively convenient strangers. Five male and five female college students playtested the *CCG* prototype, one at a time. Researchers observed players' engagement as well as how they progressed through the game and then interviewed them about what they learned.

Playtesting the first prototype disproved three assumptions. The first was the scientist's expectation that extinct creatures from Earth's past would be so intrinsically interesting that the players would be motivated to read and explore as much as possible. In fact, the playtesters mostly guessed, and only one of the 10 playtesters clicked to reveal any of the detailed critter content. The entire design team concluded that including less detail would actually increase the amount of detail players noticed. Furthermore, detail should be immediately visible without relying on optional player actions to reveal extra content ([Figure 4](#)).

The design team also reconsidered the initially vetoed concept of levels to use as a possible means to increase player motivation. Male playtesters easily but without enthusiasm traversed the first prototype from beginning to end, but all five female playtesters appeared to be lost and unmotivated, making little headway. No female playtester reached the midpoint of the game, and all of them scored far below the male playtesters. Levels were reintroduced, accommodating both the pedagogical and scientific perspectives. Levels allowed designers to organize the science learning logically by breaking the game into three ages: Paleozoic (age of amphibians), Mesozoic (age of dinosaurs), and Cenozoic (age of mammals). Players would hopefully be motivated to advance to the next level. To mitigate pedagogical concerns, any player could eventually succeed in beating a level. Levels were created to entice players, not to lock learners out ([Figure 5](#)).

Replayability is a characteristic of a good game (Prensky 2001, 179-180). This means that a player can play the game many times, ideally without end, and have a different experience each time. Games achieve replayability by providing many ways to accomplish the game's goals and subgoals and/or by introducing a random element to the game. Initially, the pedagogy expert accepted the gameplay expert's advocacy of replayability as a design goal. However, in playtesting the *CCG* prototype, sometimes cards were randomly drawn in a sequence that happened to be better for learning than other random sequences. Given playtesters' failure to learn and the intent to create a game for one-time classroom use, the pedagogy expert changed her mind about replayability, deciding that the best pedagogy could be achieved through purposive ordering of carefully selected challenges. The card metaphor was abandoned, and play was situated inside of the Tree of Life ([Figure 6](#)).

Second Prototype: The Tree of Life Game

The second playtest was conducted on the modified prototype, now called *The Tree of Life Game* or *TOLG* ([Exhibit 3](#)), which embodied the design convergences described above (Heeter, Winn, and Green [2005](#)). Five male and five female college students playtested this prototype.

The *TOLG* prototype hit its mark. Player engagement improved, and players appeared to be learning some of the intended concepts. However, female players, in particular, were uninterested in the amphibians of level 1, and all five of them ran out of time during the playtest without advancing beyond the first round of the age of dinosaurs. Designers decided to eliminate the first level and further reduced the amount of irrelevant (though to some of them, fascinating) content in the game ([Figure 7](#)).

With basic game mechanics in place, designers could now invent a backstory that was consistent with (and could ideally even contributive to) the science and learning goals and that would increase engagement and motivation for both play and learning. Content, pedagogy, and gameplay enthusiastically embraced adding an alien invasion round at the end of level 1 (the age of dinosaurs) and level 2 (the age of mammals). Each alien invasion round challenged the player to identify which creatures from Earth's past would be most affected by each of two imaginary invading alien species ([Figure 8](#)).

In the *TOLG* prototype, a guide provided basic feedback (right-wrong) about the player's actions. For the third prototype, feedback was customized and expanded, offering longer explanations for each right and wrong choice to better scaffold learning. Points and a score replaced the simple count of number right and wrong to reintroduce the feeling of competitive play. Transitional scenes between rounds reinforced learning and contributed to storytelling ([Figure 9](#)).

Third Prototype: Life Preservers

The design team felt that the third prototype, now called *Life Preservers* or *LP*, was good enough to test with target users ([Exhibit 4](#)). Designers recruited local 7th and 8th grade Girl Scouts to playtest this prototype. The *LP* playtest validated design decisions. All six players appeared to be strongly engaged throughout the game. They were disappointed when they made mistakes and spent time carefully considering the game's challenges. After playing the game, when asked what they had learned, playtesters volunteered national science standards, almost verbatim.

LP acted as a good teacher, guiding learners to think about questions of evolution and adaptation in a carefully constructed order. Gameplay took place inside of the Tree of Life. Although gameplay involved answering questions, the pedagogy was more complex than a trivia game. Learners explored a carefully selected content domain (part of the Tree of Life containing only selected critters). Questions in each round guided players to explore real scientific examples from Earth's past that illustrated key concepts in evolution. Inferring from the structure of the Tree of Life that the first bird adapted from theropod dinosaurs around 150 million years ago may be interesting, but the science standard to be learned was that "the basic idea of biological evolution is that present day species developed from earlier distinctly different species" (Center for Science, Mathematics, and Engineering Education 1996). A narrated, animated transitional scene reiterated the key learning concepts and reinforced each round.

The third prototype validated convergence on the basic game mechanics and pedagogy ([Figure 10](#)). After the playtest, two months of subsequent polishing and refinement ensued, improving the game's visuals and sounds, science content, wording of questions and feedback, and so on. Significant ideological conflicts did not arise during these refinements because the team agreed on the overall plan for what the game was and how it would be played. At times, pedagogy and content experts offered opinions about game design refinements to optimize learning. Quantitative surveys and in-game data collection with 150 players helped show where the science content needed finetuning. An external scientist validated the learning content when the game was complete. Nearly 1,000 6th through 9th graders have now played *LP* at school. Teachers have observed high levels of engagement (Heeter and Winn, forthcoming).

Conclusion

Iterative prototypes and playtesting are critical to the design of a great commercial game (Salen and Zimmerman 2006). They are even more necessary for designing great learning games. Playtesting not only helps designers refine the game mechanics, but it can also help resolve conflicts among pedagogy, content, and gameplay by moving disagreements from theoretical stances to demonstrated success or failure of design concepts.

The holy grail of educational game design is finding an optimal convergence of these perspectives for a particular content domain and set of learning goals. As Gee (2006) pointed out at the GameDevelopers conference, game designers working with instructional designers can often be overheard complaining about "those damn academics" while academics complain about "those damn designers." Although uncomfortable, these conflicts draw attention to critical game design decisions. Gee remarks, "If we do not learn that game design and learning design are compatible, serious games will remain a small space" (2006). Great educational game design teams give voice to all sides of TPCK and

make good choices about which perspective prevails for each issue based on playtesting and the overarching goals and constraints of the game.

[This article was modified from a presentation called [Using Player Research to Mediate Battles Between Pedagogy, Learning Content, and Fun](#), given at the Serious Games Summit during the Game Developers Annual Conference in San Jose, CA, March 2006. The conference audience was gameplay designers. The presentation was modified to address pedagogy and content experts who may or may not have personal experience with game design.]

References

- American Association for the Advancement of Science. 1994. *Benchmarks for science literacy*. Cary, NC: Oxford University Press.
- Center for Science, Mathematics, and Engineering Education. 1996. *National science education standards*. Washington, DC: The National Academies Press.
- Gee, J. 2003. *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.
- Gee, J. 2005. *Why video games are good for your soul*. Common Ground Publishing: Victoria, Australia.
- Gee, J. 2006. *What's wrong with serious games*. Panel Session at the Serious Games Summit 2006, Game Developers Annual Conference, San Jose, CA, March.
- Halverson, R. 2005. What can K-12 school leaders learn from video games and gaming? *Innovate* 1 (6). <http://www.innovateonline.info/index.php?view=article&id=81> (accessed November 28, 2006).
- Heeter, C., B. Winn, and D. Greene. 2005. Theories meet realities: Designing a learning game for girls. Paper presented at DUX (Designing the User eXperience) conference, San Francisco, November. <http://commtechlab.msu.edu/publications/files/dux2005.htm> (accessed November 28, 2006).
- Heeter, C., and B. Winn. (Forthcoming). Implications of gender, player type, and learning strategies for the design of games for learning. In *Beyond Barbie and Mortal Kombat: New Perspectives on Gender and Computer Games*, eds. Y. Kafai. C. Heeter, J. Denner, and J. Sun. Cambridge, MA: MIT Press.
- Kierrimuir, J., and A. McFarlane. 2004. *Literature review in games and learning*. NESTA FutureLab Series, Report 8. http://www.futurelab.org.uk/research/reviews/08_01.htm (accessed November 28, 2006).
- Malone, T., and M. Lepper. 1987. Making learning fun: A taxonomic model of intrinsic motivations for learning. In *Aptitude, learning, and instruction*, volume 3: *Cognitive and affective process analysis*, ed. R. E. Snow and M. J. Farr, 223-53. Hillsdale, NJ: Erlbaum.
- Mishra, P., and M. Koehler. 2006. Technological pedagogical content knowledge: A new framework for teacher knowledge. *Teachers College Record*. 8 (6):1017-1054.
- Prensky, M. 2001. *Digital game-based learning*. New York: McGraw-Hill.
- Rouse, R. 2001. *Game design theory and practice*. Plano, TX: Wordware Publishing.
- Salen, K., and E. Zimmerman. 2006. *The game design reader*. Cambridge, MA: MIT Press.
- Van Eck, R. 2006. Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE Review* 41 (March/April): 16-30. <http://www.educause.edu/apps/er/erm06/erm0620.asp> (accessed November 28, 2006).

This article may be reproduced and distributed for educational purposes if the following attribution is included in the document:

Note: This article was originally published in *Innovate* (<http://www.innovateonline.info/>) as: Winn, B., and C. Heeter. 2006. Resolving conflicts in educational game design through playtesting. *Innovate* 3 (2). <http://www.innovateonline.info/index.php?view=article&id=392> (accessed December 2, 2006). The article is reprinted here with permission of the publisher, The Fischler School of Education and Human Services at Nova Southeastern University.

To find related articles, view the webcast, or comment publically on this article in the discussion forums, please go to <http://www.innovateonline.info/index.php?view=article&id=392> and select the appropriate function from the sidebar.