4.

Inciting Out-Of-Game Transfer

Adapting Contrast-Based Instruction For Educational Games

Catherine C. Chase (Teachers College, Columbia University), Erik Harpstead (Carnegie Mellon University), & Vincent Aleven (Carnegie Mellon University)

Abstract

We adapted a successful instructional principle – contrasting cases – to create an educational game to teach young children physical principles of stability. Our goal was to design a game that would promote transfer – extending the reach of the educational game beyond the game itself. In Study 1, we compared a "standard" version of the game to a "contrast" version that contained contrasting case levels designed to help learners notice the principles underlying game content. In Study 2, we augmented the contrast version of the game with induction levels that focused learners on abstracting general principles from sets of contrasting cases. In both studies, we found that contrast versions of the game facilitated transfer, while standard versions did not. Students found contrast versions of the game highly enjoyable, just as enjoyable as the standard game. Findings have implications for the design of educational games that are instructive yet fun.

Introduction

While research findings on the effectiveness of educational games are still somewhat mixed (Honey & Hilton, 2011), there is now mounting evidence suggesting that educational games can produce greater learning than traditional instruction (Wouters et al., 2013). But how exactly does one design an effective educational game? Some have attempted to identify design guides for educational games, but few of these guides describe specific forms of learning activities (Clark et al, 2009; Mayer, 2014). We posit that one successful approach for designing effective educational games is to adapt existing instructional principles as game mechanics (Aleven et al, 2010). Learning science research has identified several instructional principles that enhance learning and transfer in a wide variety of contexts.

We are chiefly interested in producing transfer – the application of content learned in one context to a different context. Since our focus is on how games can support classroom instruction, we explore transfer from game play to academic tasks. Many have explored this same question but with a focus on learning rather than transfer. We have chosen to adopt a transfer perspective because we see the challenge of learning from games as a transfer problem at its core; game contexts differ fundamentally from traditional academic contexts. For example, games provide rich narratives, situated representations of content, and play goals, while school contexts tend to be devoid of narrative, provide abstract representations, and invoke learning goals.

An overlooked principle that promotes deep learning and transfer is the use of “contrasting cases.” Contrasting cases are sets of examples that share many similarities but differ on key features, which are tied to deep domain principles (Bransford et al, 1989; Gibson & Gibson, 1955). There are two key processes by which contrasting cases provoke transfer. First, the contrasts help learners *notice* deep principles in the context of specific examples. Second, learners are pushed to form an *abstract* understanding of the deep principles that run throughout a set of cases, which can be flexibly adapted to novel contexts. Contrast-based instruction has been used effectively to achieve both noticing and abstraction (Gick & Patterson, 1992; Schwartz et al, 2011).

While a great deal is known about instructional principles that help people learn, it is less clear whether these same principles will work well in a game context. Unfortunately, many educational games come across as “chocolate-covered broccoli” where standard learning activities are “gamified” by the addition of simple game features like points and fancy graphics. While this kind of learning environment can enhance learning, it separates the instructional and game play elements, and in the end, may not feel like a real game. One reason we chose contrasting cases as our core pedagogy is because they share many characteristics with common games, such as being highly visual, player-controlled, intuitive tasks. We hoped that by adapting a game-like instructional technique into a game, it would make for a meaningful play experience and a strong integration of instruction with play.

In this paper, we describe our efforts to build and evaluate an educational game that pushes students to notice and abstract underlying principles through the application of contrasting cases, an often-overlooked learning science principle. We integrated contrasting cases into *RumbleBlocks*, a game that was designed to teach elementary-aged children about the physics of stability (Christel et al., 2012). To our knowledge, this is the first educational game that explicitly incorporates the instructional principle of contrasting cases. In Study 1, we compared a standard version of the *RumbleBlocks* game, with no explicit educational scaffolds, to a contrast version of the game, imbued with several contrasting case levels. In Study 2, we compared the standard version to a contrast version that was augmented with new induction levels, which encouraged learners to abstract the deep principles that ran throughout a set of cases. In both studies we investigated the following hypotheses:

1. Students who play contrast versions of the game should demonstrate greater out-of-game transfer than students who play the standard game.
2. Students who play contrast versions of the game should demonstrate similar in-game learning as students who play the standard game. Previous work has found that contrasting case-based instruction specifically aids transfer, but not always immediate learning or performance (Roll, Aleven, & Koedinger, 2011; Schwartz & Martin, 2004).
3. Students who play contrast and standard versions of the game will report similar levels of game enjoyment. We were careful to design contrasting case levels that seamlessly integrated instruction and play.

Study 1 Method

Game Design

The *RumbleBlocks* game is designed to teach three principles of stability. Structures tend to be more stable when they have (1) a wide *base* (2) *symmetrical* sides and (3) a *low center-of-mass*. These concepts fall under PS 2.C "stability and instability in physical systems" in the National Research Council's framework for K-12 science education (National Research Council, 2012).

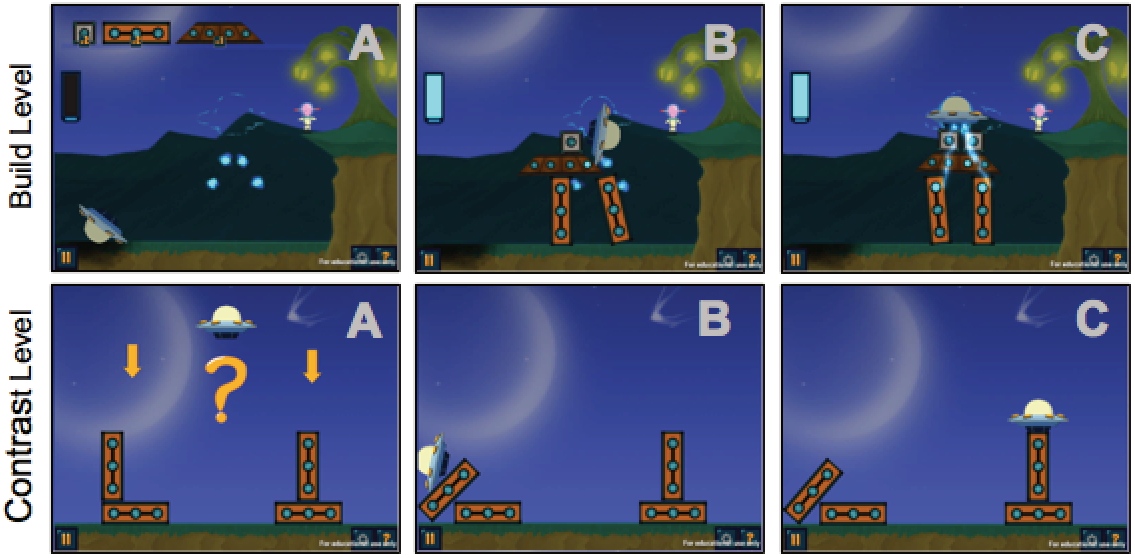


Figure 1. Example "Build" and "Contrast" Levels for Symmetry. A) Start state, B) Failed attempt, C) Successful attempt.

RumbleBlocks is couched in an alien-themed narrative. An alien has been stranded on a foreign planet and separated from its spaceship. In each level, the player's goal is to reunite the alien with its spaceship (Figure 1). The alien is located on the edge of a high cliff, with the spaceship located in a chasm below. The main mechanic of the *RumbleBlocks* game is block building. Each level varies the given set of blocks and the target height of the structure. Through the strategic placement of "energy dots" that towers must cover, levels scaffold students in building structures that have low centers of mass, wide bases, or symmetrical sides. The player's job is to build a structure that will reach the high cliff, cover the energy dots, and place the spaceship at the top of the structure. After that, an earthquake hits, and if the structure stays standing, the alien can now reach its ship and fly home and the level is won. The game was created in the Unity game engine, which has a well-developed physics engine, allowing the game to accurately simulate the laws of two-dimensional rigid body physics. Thus, the blocks and towers behave realistically, responding to gravity, friction, and torque.

The "standard" version of the game contained only the build levels described above (see Figure 1). The "contrast" version of the game contained build levels interspersed with "contrast" levels (see Figure 1), where players predict which of two towers will stay standing in an earthquake (to help the alien spaceship find a stable place to land). The towers differ on a single feature that maps to a specific principle (e.g. towers are identical except one has a wider base or lower center-of-mass). After players predict, the earthquake knocks one of the two towers down, providing feedback (while also adding suspense to the game), and then students move on to the next level. The "contrast" levels were designed to help students attend to features of the towers that relate to principles of stability.

Participants

Study participants were 166 children in grades K-3 at two public elementary schools in Pennsylvania. The schools were high-performing (only 3% and 7% of 3rd graders were not proficient on state tests in math and reading, respectively), moderate SES (27% economically disadvantaged), and predominantly Caucasian (3% minority students).

Procedure

Students in each class were randomly assigned to either "standard" (n = 90) or "contrast" (n = 76) conditions. Conditions were implemented via two different game versions. The standard version was designed to mimic many common games, where players apply the principles of physics to achieve game goals, but those principles are never highlighted in any way. In the "standard" version, students played only "build" levels of the game. The build levels were blocked by principle, such that students played a series of build levels about symmetry, then base width, then center-of-mass. The "contrast" version of the game was identical to the "standard" version, except that it contained additional "contrast" levels, designed to focus students' attention on a specific principle, in isolation.

Transfer beyond the game was assessed by out-of-game paper-and-pencil pre and posttests that asked students to apply knowledge of stability principles to a more traditional schoolish test (scores were converted to percent correct). Learning was measured by in-game pre and post assessment levels, which contained no scaffolding. Each structure that students built in these levels was assessed in two ways: (1) success: whether or not the structure stayed standing in an earthquake, max score=3 and (2) principle application scores: standardized Z scores reflecting the base width, degree of symmetry, and height of center-of-mass of each structure, converted so that positive scores indicate greater application of each principle. Game enjoyment was measured by a 3-point "smiley face" likert scale that asked students how much they liked the game.

The study took place over the course of four consecutive days, with 30-minute sessions each day. On Day 1, students took the out-of-game transfer pretest. On Days 2-3, students played the *RumbleBlocks* game on desktop computers. Total game play time averaged 52 minutes in both conditions. At the beginning and end of game play, students took the in-game learning pre (Day 2) and posttests (Day 3), which were embedded in the game. On Day 4, students completed the out-of-game transfer posttest and the survey of game enjoyment.

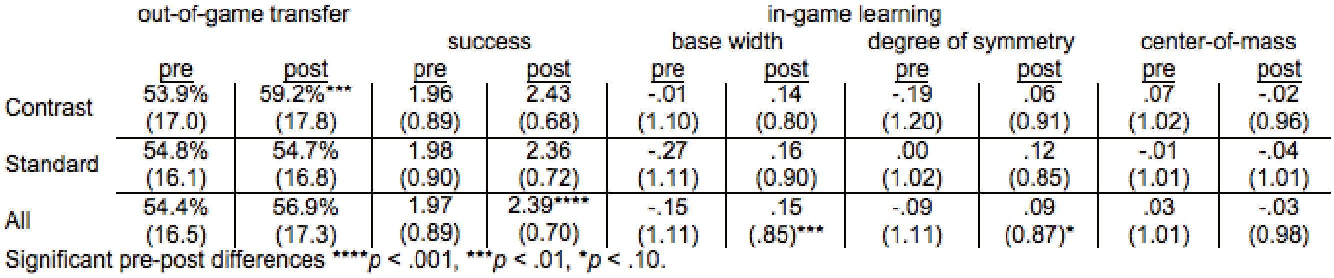
The study took place over the course of four consecutive days, with 30-minute sessions each day. On Day 1, students took the out-of-game transfer pretest. On Days 2-3, students played the *RumbleBlocks* game on desktop computers. Total game play time averaged 52 minutes in both conditions. At the beginning and end of game play, students took the in-game learning pre (Day 2) and posttests (Day 3), which were embedded in the game. On Day 4, students completed the out-of-game transfer posttest and the survey of game enjoyment.

Results

There were no pre-existing differences on out-of-game pretest scores across conditions or schools, *p*'s >

.11. However, there were significant differences by grade, F(3, 150) = 10.13, p < .001. To test for gains on the out-of-game transfer test, we conducted a mixed ANOVA with grade as a covariate, condition as a between-subjects factor, time as a within-subjects factor, and test score as the dependent measure. There was a strong main effect of grade, F(1, 163) = 58.99, p < .001, demonstrating that students in higher grades scored higher on both tests. There was no main effect of time, but there was a significant interaction of time by condition, F(1, 163) = 3.76, p = .05. Pairwise comparisons revealed that the contrast condition increased its score from pretest to posttest, p = .01, ηp2 = .04, whereas no statistically significant difference between pretest and posttest was detected for the standard condition, p = .93. In other words, students who played the contrast version transferred their learnings beyond the game, while those who played the standard version did not (see Table 1).

The in-game learning pre and posttests were scored for the success of students' structures (i.e., whether structures remained standing when the earthquake hit). To test for differential effects of condition on in-game learning, a mixed ANOVA was conducted with time as a within-subjects variable and condition as a between-subjects factor. There was a large main effect of time, F(1, 164) = 25.50, p < .001, ηp2 = .14. This demonstrates that students across both conditions made sizeable gains in their success from pre to post build levels, with an average gain of 14% and a medium effect size. However, the interaction effect of time by condition was not significant, p = .57, indicating that both conditions got better at playing the game over time (see Table 1).



	out-of-game transfer		in-game learning							
			success		base width		degree of symmetry		center-of-mass	
	pre	post	pre	post	pre	post	pre	post	pre	post
Contrast	53.9% (17.0)	59.2%*** (17.8)	1.96 (0.89)	2.43 (0.68)	-.01 (1.10)	.14 (0.80)	-.19 (1.20)	.06 (0.91)	.07 (1.02)	-.02 (0.96)
Standard	54.8% (16.1)	54.7% (16.8)	1.98 (0.90)	2.36 (0.72)	-.27 (1.11)	.16 (0.90)	.00 (1.02)	.12 (0.85)	-.01 (1.01)	-.04 (1.01)
All	54.4% (16.5)	56.9% (17.3)	1.97 (0.89)	2.39**** (0.70)	-.15 (1.11)	.15 (.85)***	-.09 (1.11)	.09 (0.87)*	.03 (1.01)	-.03 (0.98)

Significant pre-post differences ****p < .001, ***p < .01, *p < .10.

Table 1. Adjusted mean scores (and SD) for in-game learning an out-of-game transfer.

We also examined the level of principle application in pre and posttest game levels. To test for differential learning gains across conditions by principle, a mixed ANOVA with condition, time, and principle as independent factors and Z-scores as the dependent variable was conducted. The interaction of time by condition was not significant, and neither was the interaction of time by condition by principle, indicating that performance did not differ by condition. However, there was a marginally significant main effect of time F(1, 164) = 3.04, p = .08, which was largely driven by specific principles, as evidenced by the marginally significant interaction of time by principle, F(2, 328) = 2.85, p = .06. No other main effects or interactions were significant. Pairwise comparisons (with Bonferroni correction) between in-game pre and posttest scores revealed that students built structures with wider bases, p = .01, that were marginally more symmetrical, p = .07, but, the structures' centers-of-mass did not change, p = .66. Students improved greatly at applying the base width principle, somewhat on the symmetry principle, and not at all on the center-of-mass principle (see Table 1). However, there were no significant differences between conditions.

Gain scores on the out-of-game transfer test were significantly correlated with gain scores for success on the in-game learning test (r = .22, p = .005), but not with principle application scores. This suggests that successful game play does not fully predict transfer outcomes, underscoring the need for out-of-game measures to assess whether in-game learnings transfer out of the game.

Children reported equally high levels of enjoyment for both versions of the game, and in all grades. A univariate ANOVA with condition and grade as independent variables and enjoyment ratings as the dependent variable found no significant main effects of condition or grade, nor any significant interaction, *p*'s > .31. Students in both conditions gave high average enjoyment ratings (Mcontrast = 2.90, SDcontrast = .39; Mstandard = 2.88, SDstandard = .38 on a scale of 1-3), and 91% of all students gave the highest rating of 3 ("I liked it!"). Children in our study greatly enjoyed playing the *RumbleBlocks* game, and the addition of contrasting case levels did not reduce that level of enjoyment.

Study 2 Method

Game Design

In Study 2, we augmented the existing "contrast" game with new induction levels that were designed to encourage learners to abstract the principles underlying sets of contrasting cases. Typical induction prompts ask learners to explain the similarities and differences between a set of cases (Alfieri, Nokes-Malach, & Schunn, 2013), or sometimes, learners are asked to generate an equation or graph that demonstrates the pattern across a set of cases (Schwartz & Martin, 2004; Schwartz et al., 2011). However, writing and generating equations are not feasible activities for very young children, and they violated our design objective of avoiding instructional elements that "smell like school."

Our solution was to design levels where players select visual representations or "goggles" that depict the pattern of principles imbued in a set of cases. This mechanic honors the visual nature of game tasks, and created a selection task that was intuitive and developmentally appropriate for young players. In an induction level, players were instructed to pick the goggles that "explain why some towers fell while other stood when the earthquake hit," encouraging learners to think abstractly about what might affect stability across several different types of towers. Players could view the same set of structures through different principle-focused lenses (see Figure 2). Players were given two chances to identify the right "goggles," and this was followed by a short movie that gave a brief explanation of the relevant principle and how it impacts stability. This general sequence of having learners attempt to induce the relevant principle shown in multiple cases followed by expository instruction on the correct principle was modeled after activities that prepare people to learn from future instruction (Schwartz & Martin, 2004; Schwartz et al., 2011).

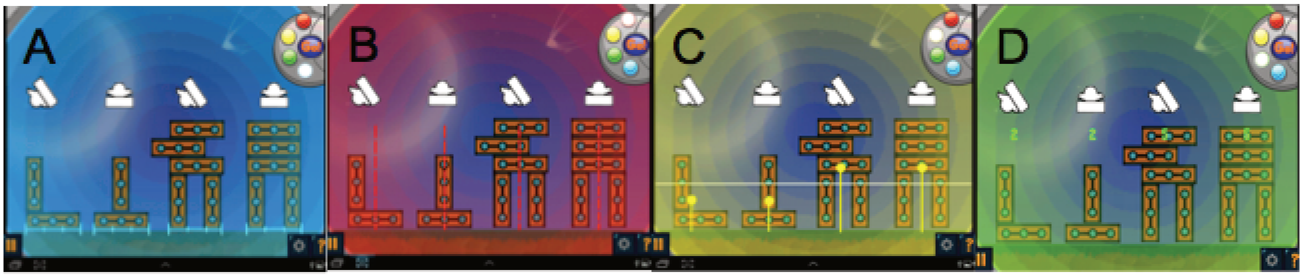


Figure 2. "Goggles" in introduction level. A) Base width, B) Symmetry, C) Center-of-mass, D)Number of blocks goggles (a distractor choice.)

Participants

Study participants were 83 first and second graders at a public elementary school in Pennsylvania. The school was low-performing (38% and 46% of 3rd graders at the school were not proficient in reading and math respectively, on statewide tests), low SES (87% economically disadvantaged students), and predominantly African American (70% African American, 20% Caucasian).

Procedure

Students were randomly assigned within each class to either “standard” (n = 43) or “contrast + induce” (n = 40) conditions. The contrast + induce condition was essentially the same as the contrast game used in Study 1, but each block of contrast levels was followed by an additional induction level. The standard version of the game was very similar to the standard game in Study 1, with the addition of 12 more build levels, to accommodate an additional session of game play.

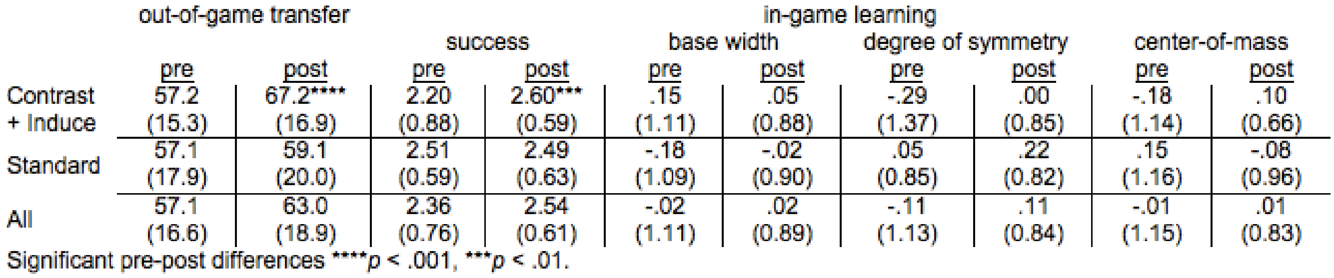
The procedure was roughly the same as in Study 1, with one additional day of game play and slightly longer daily sessions. The study took place over the course of five consecutive days, with 45-minute sessions each day. Measures were highly similar to the ones used in Study 1. The most significant changes were that the out-of-game transfer test was computerized and game enjoyment measures were converted to a more sensitive 5-point likert scale.

Results

There were no pre-existing differences on out-of-game pretest scores across conditions or grades, *p*’s > .53. To test for differential gains on our out-of-game transfer test, we conducted a mixed ANOVA with condition as the between-subjects factor, time as a within-subjects factor, and total test score as the dependent measure. There was a significant effect of time, $F(1, 81) = 10.31$, $p = .002$, which was largely driven by the significant interaction of condition by time, $F(1, 81) = 4.55$, $p = .04$, $\eta p2 = .05$. There were no other significant effects, *p*’s > .23. Pairwise comparisons revealed that the contrast + induce condition increased its score from pretest to posttest, $p < .001$, $\eta p2 = .15$, while the standard condition’s score remained the same over time, $p = .44$. Similar to the findings in Study 1, students who played the contrast + induce version of the game made measurable transfer gains, while those who played the standard version did not (see Table 2).

To test for differential effects of condition on in-game learning (success), a mixed ANOVA was conducted with time as a within-subjects factor and condition as a between-subjects factor. There was a marginal main effect of time, $F(1, 81) = 3.47$, $p = .07$ and a significant interaction of time by condition, $F(1, 81) = 4.39$, $p = .04$, indicating differential growth by condition. Pairwise comparisons revealed that the contrast + induce condition increased their score from pre to post, $p = .007$, $\eta p2 = .09$, while the standard condition’s score did not change over time, $p = .87$ (see Table 2).

We also explored whether students improved in their application of the targeted instructional principles on the in-game learning pre and posttest (see Table 2). To test for differential gains across conditions by principle, a mixed ANOVA with condition, time, and principle as independent factors and Z-scores as the dependent variable was conducted. There was no significant effect of time, and there was no significant interaction of condition by time, nor were there any main effects or interactions with principle, *p's* > .18. Results indicate that neither condition improved in their application of the target principles from pre to posttest, and there were no differences by principle type.



	out-of-game transfer		in-game learning							
			success		base width		degree of symmetry		center-of-mass	
	pre	post	pre	post	pre	post	pre	post	pre	post
Contrast	57.2	67.2****	2.20	2.60***	.15	.05	-.29	.00	-.18	.10
+ Induce	(15.3)	(16.9)	(0.88)	(0.59)	(1.11)	(0.88)	(1.37)	(0.85)	(1.14)	(0.66)
Standard	57.1	59.1	2.51	2.49	-.18	-.02	.05	.22	.15	-.08
	(17.9)	(20.0)	(0.59)	(0.63)	(1.09)	(0.90)	(0.85)	(0.82)	(1.16)	(0.96)
All	57.1	63.0	2.36	2.54	-.02	.02	-.11	.11	-.01	.01
	(16.6)	(18.9)	(0.76)	(0.61)	(1.11)	(0.89)	(1.13)	(0.84)	(1.15)	(0.83)

Significant pre-post differences ****$p < .001$, ***$p < .01$.

Table 2. Adjusted mean scores (and SD) for in-game learning and out-of-game transfer.

As in Study 1, there was a positive relationship between the gain in success on the in-game learning test and the out-of-game transfer gain, $r = .24$, $p = .03$, but not with principle application scores on the in-game test. Again, this underscores the fact that in-game measures of learning do not fully capture the degree of transfer beyond the game.

Children reported equally high levels of enjoyment for both versions of the game, though 2nd graders liked the game better than 1st graders. A univariate ANOVA with condition and grade as independent variables and enjoyment ratings as the dependent variable found a significant effect of grade, $F(1, 71) = 3.85$, $p = .05$, (M2nd = 4.75, SD2nd = 1.02; M1st = 4.28, SD1st = 1.02), though both grades gave very high ratings. However, there was no significant main effect of condition, nor any interaction of condition by grade, *p*'s > .53. Students in both conditions rated the game as highly enjoyable on a scale of 1-5 (Mcontrast+induce = 4.44, SDcontrast+induce = 1.03; Mstandard = 4.59, SDstandard = 1.04), and 75% of all students gave the highest rating of 5 ("I liked it a lot"). Children in our study greatly enjoyed playing the *RumbleBlocks* game, and the addition of contrast and induction levels did not reduce enjoyment.

General Discussion

In this paper, we have demonstrated the value of adapting contrast-based instruction for game contexts. Two studies using very different student populations demonstrated that an educational game can be enhanced by the addition of contrasting cases. The addition of contrasting case levels to a standard game led players to develop flexible understandings of stability principles that transferred beyond the game to an out-of-game test that asked learners to reason about the principles in novel contexts. In contrast, we did not find evidence that children who played a standard game transferred what they learned beyond the game context. Findings as to whether players of the standard game learned to build more stable structures in the game itself are mixed; in Study 1 we found evidence that they did, while in Study 2 they did not. Regardless, playing a standard game did not help learners to apply their knowledge outside the confines of the game. This finding also emphasizes the importance of measuring transfer in studies that assess the effectiveness of educational games. Measures of learning derived from pure gameplay metrics, such as number of levels beaten, may falsely inflate our estimates of the learning students actually take away from the game (cf. Long & Aleven, 2014).

Another important metric when considering a game's effectiveness is how much players enjoyed playing the game. Too often, games are implemented as “chocolate covered broccoli” where the unpleasant experience of learning is masked by the enjoyable experience of play. We explicitly chose to adapt contrasting cases as a game mechanic (out of many other successful instructional principles) because of their game-like affordances, hoping that the experience of the contrasting case activity would make learning and play one and the same. Moreover, to build *RumbleBlocks*, we forged a close collaboration between learning specialists and game designers from the start. Our approach paid off – in both studies, adding carefully designed contrast and induction levels did not change children’s liking of the game and may have even contributed to it.

This set of studies contains several limitations that prompt ideas for future research. First, we cannot make strong claims about the added value of the induction game levels, over and above the contrasting case levels. Future research could isolate the effects of the induction levels on transfer outcomes. Moreover, it is impossible to tell whether the “build” levels were at all instrumental in the transfer that resulted from either of the contrasting case versions of the game. Would students have transferred at the same rate had they only played the contrasting case levels? Or were the build levels necessary for helping learners to apply the principles they were noticing and reasoning about in the contrast and induction levels? Future research could explore how build and contrasting case activities mutually influence the learnings that players build in the game.

The larger message for educational game designers is to incorporate instructional methods that specifically impact transfer. Our work underscores the fact that students playing standard game levels can get better at playing the game and can effectively apply the target content in achieving game goals, without demonstrating transfer from the game to more standard school tasks. A final takeaway is that game mechanics must help players notice and reflect on the content-based principles they are applying to achieve game goals. We have illustrated one way to do this through the use of contrasting cases with prompts to induce the deep principles that underlie a set of cases.

References

Aleven, V., Myers, E., Easterday, M., & Ogan, A. (2010). Toward a framework for the analysis and design of educational games. In G. Biswas, D. Carr, Y. S. Chee, & W. Y. Hwang (Eds.), *2010 third IEEE international conference on digital game and intelligent toy enhanced learning* (pp. 69-76). Los Alamitos, CA: IEEE Computer Society. doi:10.1109/DIGITEL.2010.55.

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: a meta-analytic review. *Educational Psychologist*, *48*(2), 87-113.

Bransford, J.D., Franks, J.J., Vye, N.M., and Sherwood, R.D. (1989). New Approaches to Instruction: Because wisdom can’t be told. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning.* Cambridge University Press: New York.

Christel, M. G., Stevens, S. M., Maher, B. S., Brice, S., Champer, M., Jayapalan, L., Chen, Q., Jin, J., Hausmann, D., Bastida, N., Zhang, X., Aleven, V., Koedinger, K., Chase, C., Harpstead, E., & Lomas D. (2012, July). RumbleBlocks: Teaching science concepts to young children through a Unity game. In *Computer Games (CGAMES), 2012 17th International Conference on* (pp. 162-166). IEEE.

Clark, D., Nelson, B., Sengupta, P., & D’Angelo, C. (2009, October). Rethinking science learning through digital games and simulations: Genres, examples, and evidence. In *Learning science: Computer games, simulations, and education workshop sponsored by the National Academy of Sciences, Washington, DC*.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment. *Psychological Review*, 62, 32–41.

Gick, M. L., & Paterson, K. J. (1992). Do contrasting examples facilitate schema acquisition and analogical transfer? *Canadian Journal of Psychology*, 46, 539–550.

Honey, M. A. & Hilton, M., (Eds.). (2011). *Learning science through computer games and simulations*. Washington, DC: National Academies Press.

Long, Y., & Aleven, V. (2014, January). Gamification of Joint Student/System Control over Problem Selection in a Linear Equation Tutor. In *Intelligent Tutoring Systems* (pp. 378-387). Springer International Publishing.

Mayer, R. E. (2014). *Computer games for learning: An evidence-based approach*. MIT Press.

National Research Council. (2012). *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, D.C.: National Academies Press.

Roll, I., Aleven, V., & Koedinger, K. R. (2011). Outcomes and mechanisms of transfer in invention activities. In *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2824-2829). Cognitive Science Society Austin, TX.

Schwartz, D.L., Chase, C.C., Oppezzo, M.A., and Chin, D.B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*, 4, 759-775.

Schwartz, D.L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*(2), 129-184.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249.