Tuning the Knobs and Dials: Empirically Maximizing Features for Serious Games

Rita Bush, Intelligence Advanced Research Projects Activity
Carl Symborski, Leidos
Rosa Mikeal Martey, Colorado State University
Emilie T. Saulnier, 1st Playable Productions
Elizabeth S. Veinott, Applied Research Associates (ARA)

Games research has often treated the game as a black box; we introduce the game into a situation, observe the effects, and declare success. But we don't know *why* the game worked. What was it about the game that made it a powerful tool for learning? Can games be used to teach not just concepts and knowledge, but to also spark changes in reasoning, judgment, and decision-making? This panel will address these questions.

Five game development and research teams were charged with creating 5 games to teach intelligence analysts to learn to recognize and mitigate 3 cognitive biases: Confirmation Bias (CB), Fundamental Attribution Error (FAE) and Bias Blind Spot (BBS). Cognitive biases occur in situations where evidence is incomplete or uncertain and time pressures rise. Heuristics provide a quick solution to the problem under consideration, but "heuristics gone astray" result in biases may lead to errors in judgment. The judgment and decision-making literature has shown that overcoming biases is difficult (Kahneman, 2011).

Each of the teams went through 3 rounds of prototyping and testing their games over an 18-month period. Each round varied one or more game features that were used in a controlled experiment to determine the effectiveness of the particular feature to teach players to recognize and mitigate their own cognitive biases. For example, one version of a game examined the use of a student model to guide game play, which was compared to a second version of the game without a student model. Participants in the experiments were randomly assigned to play one version of the game or to watch a 30-minute training video, which served as the control condition. Game variables examined by the teams included student models, priming, visual fidelity, and repetition.

Three of those games are discussed here. The 3 games employed a variety of narratives and gameplay styles; we had no preconceived notions about what would make a successful game. Thus the game spectrum ranged from casual, puzzle-style games with minimal narrative structure, to immersive, 3D game environments with strong storylines. We now describe the 3 games in more detail, the game features examined, and the results of the empirical tests of game effectiveness on recognizing and mitigating cognitive bias.

Heuristica (ARA) is a 3D immersive video game developed using Unreal3. Designed with a space station narrative, Heuristica's gameplay is driven by exploration, problem solving, and includes two phases for learning, a training phase and an action phase. The player assumes the role of a human astronaut on a space station where a new starship crewed by a team of androids is about to launch their first expedition. A small team of humans must compete to win command over the androids on-board and the player is competing for one of these spots. There is a core difference between the minds of androids and humans. Humans use heuristics, leaving them vulnerable to biases, while androids do not. To be successful leading the androids, the player must demonstrate his or her ability to recognize and mitigate several cognitive biases.

Over the course of 3 experiment cycles, we examined the effects of 10 different game variables (e.g., real time reward, time pressure, session duration, repeated sessions of game play, 3rd person perspective, student model/intelligent tutoring) on learning and retention with Heuristica (Veinott et al, 2013). We found that certain versions of the video game improved participants' immediate knowledge of the cognitive biases and their ability to mitigate them. The size of the improvement depended on the game variable being examined. Two game conditions that showed the greatest improvement in learning, and significantly more than the training video, were the repeated session game (2 sessions over 3 days) and the single session game using a student model for intelligent tutoring. Furthermore, these learning improvements were retained after an 8-week delay.

The ***CYCLES*** game (Albany Team) is a 2D casual game in which players navigate an avatar through a slightly sinister "Bias Training Center" comprised of a series of rooms with puzzles that teach about the three biases and how to mitigate them. A humorous storyline involving an evil genius provides a basic story and context for the player's predicament. Players receive infographics before each room to define and describe biases and mitigation strategies as well as short quizzes after each room to reinforce each lesson and transfer to real-world scenarios. Our team studied the effects of three game variables (Martey et al., forthcoming). First, we examined avatar customization by varying whether the player could choose their avatar's shape and uniform. We found no significant differences in learning between the game versions. Our next experiment looked at the effects of visual and narrative detail. To do so, we used two art styles: a detailed condition with full-color, rich texture and shading, and realistic detailing; and a minimalistic condition that was largely monochrome with minimal shading and almost no textures. We also used two narratives styles for the game text: a light narrative in which the player was only told they were in a training center; and a rich narrative that explained a plot with backstory and specific character motivations, as well as an opening cut scene. Because narrative is often conveyed via visual details in commercial games and thus closely associated with each other, we combined these variables to create four versions of the game. This allowed us to examine the impact of art and narrative combinations as well as individually using statistical analyses.

Our finding was that the minimalistic art conditions did result in greater bias mitigation than a detailed art game. Somewhat similarly, adding more narrative through text did not improve mitigation. Our experiments also show the final *CYCLES* game trains players well in bias recognition and mitigation, outperforming the training video control condition. In a final round of experiments, we also found that replaying the game 5-7 days later greatly improves learning retention 8 weeks later. We found this repetition was far more effective than a longer version of the game with additional levels.

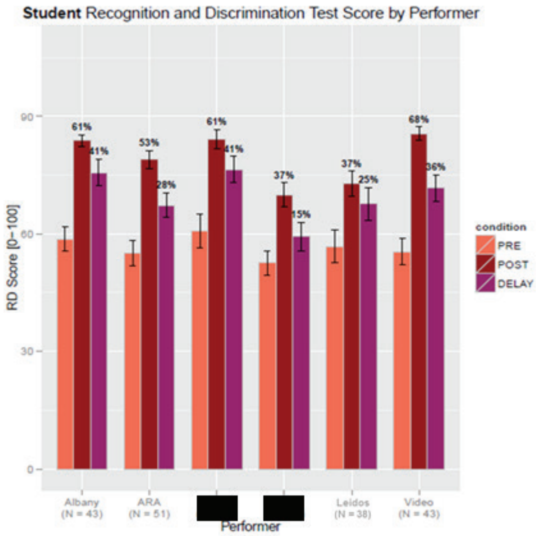
Missing: The Pursuit of Terry Hughes (Leidos Team) is designed in the style of an adventure game that combines the rich immersive qualities of entertainment software with a host of training activities on cognitive bias recognition and mitigation. The story develops over the course of three episodes, during which the player completes a series of tasks and interactions with game characters, all in pursuit of resolving the mystery at the center of the story. The player is exposed to specific bias-invoking situations in the form of "bias vignettes," where cognitive biases exhibited by the player are measured. After each episode there is an After Action Review that teaches about specific biases and offers feedback on player performance. The features examined include session duration, either 30 or 90 minutes; type of narrative, varied by including or excluding cognitive reinforcing stories; point of view, either first or third person; and communication style, varied by providing or not providing hints during the game.

Results obtained by the Leidos team showed that the game *Missing* improved participants' ability to recognize and mitigate the targeted cognitive biases. Studies of the individual game variables showed they had varying effects on the game efficacy to teach bias recognition and mitigation, depending on the bias. The session duration game variable had a significant effect, particularly for the mitigation of Confirmation Bias, when it was set to 90 minutes. The inclusion of cognitive reinforcing stories had a positive effect for the mitigation of both Confirmation Bias and the Bias Blind Spot. First person point of view had a large mitigation effect on Confirmation Bias. Finally, providing hints during the game had a mitigation effect on all three biases. These studies resulted in the final game variable selection of 90 minute duration, providing reinforcing stories, first person point of view, and providing hints.

At the conclusion of the 3 rounds of prototyping and experimentation, the 5 teams sent the "best and final version" of their games to a third party, for independent validation and verification of game effectiveness. Participants were randomly assigned to play one of the 5 games or to watch the educational video. Two studies were conducted; one with college students and one with practicing intelligence analysts. Clear differences in game effectiveness emerged from these studies.

The independent validation showed that 3 of the 5 games resulted in participants with significant improvement in *knowledge* about biases, even after a delay of 8 weeks (Figure 1). We also assessed *behavior change*, that is, did the participant show less biased decision-making when presented with a series of judgment and decision-making tasks? Participants who played one of the 5 of the games, or who watched the video, showed immediate improvement in bias mitigation; for the games (but not the video) this improvement persisted even after a delay of 8 weeks (Figure 2). Four of the 5 games beat the video in the amount of participant bias mitigation.

This research program has shown the value of having game developers and researchers conducting more systematic examination and testing of game features. Some of the game variables that we expected to be effective were not. Some of the game storylines, which intuitively should have worked very well, did not. The heuristic judgments of the game designers were sometimes off-target; empirical testing revealed these flaws and showed us what did work well in accomplishing the goals of the games. Game designers and researchers must beware of their own biases!



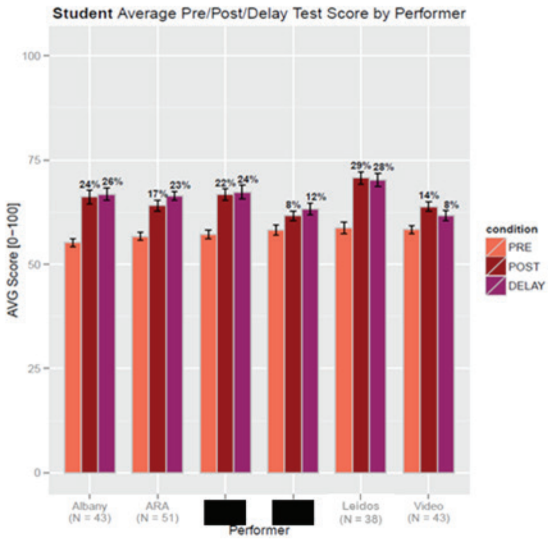


Figure 1: Recognition/ Discrimination (Knowledge) Scores.

Figure 2: Bias Mitigation (Behavior Change) Scores.

References

Kahneman, D. (2011). *Thinking, fast and slow*. London: Macmillan.

Martey, R., et al. (forthcoming). Five approaches to measuring engagement: Comparisons by video game characteristics. *Simulation and Gaming*.

Veinott, E. S., Leonard, J., Lorince, J., Hotaling, J., Ross, T., Mayell, S. ,Todd, P.M., Busemeyer, J., Castronova, E., Hale, C., Catrambone, R., Whitaker, E., Fox, O., Flach, J., & Hoffman, R. R. (2013). The effect of camera perspective and session duration on training decision making in a serious video game. In *Games Innovation Conference (IGIC), 2013 IEEE International* (pp. 256-262). Vancouver, BC: IEEE.

Quinn, M. M., Symborski, C., Barton, M., Korris, J., Falstad, T., & Granato, S. (2013). Methodology for Playtesting Computer Games: A Mixed Method Approach. In *Games Innovation Conference (IGIC), 2013 IEEE International* (pp. 222-227). Vancouver, BC: IEEE.

Acknowledgments

This work was accomplished in support of the Intelligence Advanced Research Projects Activity (IARPA) Sirius Program, BAA number IARPA-BAA-11-03. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA or the U.S. Government.