

Training Anchoring and Representativeness Bias Mitigation through a Digital Game

Citation: Lee, Y.-H., Dunbar, N., Miller, C., Lane, B., Jensen, M., Bessarabova, E., Burgoon, J., Adame, B., Valacich, J., Arterburn, E., Bostwick, E., Piercy, C., King, S., Elizondo, J., & Wilson, S. (2016). Training anchoring and representativeness bias mitigation through a digital game. *Simulation & Gaming*, 1-29. DOI: 10.1177/1046878116662955.

Abstract

Humans systematically make poor decisions because of cognitive biases. Can digital games train people to avoid cognitive biases? The goal of this study is to investigate the affordance of different educational media in training people about cognitive biases and to mitigate cognitive biases within their decision-making processes. A between-subject experiment was conducted to compare a digital game, a traditional slideshow, and a combined condition in mitigating two types of cognitive biases: anchoring bias and representativeness bias. We measured both immediate effects and delayed effects after four weeks. The results suggest that the digital game and slideshow conditions were effective in mitigating cognitive biases immediately after the training, but the effects decayed after four weeks. By providing the basic knowledge through the slideshow, then allowing learners to practice bias-mitigation techniques in the digital game, the combined condition was most effective at mitigating the cognitive biases both immediately and after four weeks.

Keywords: digital game, heuristics, anchoring, representativeness, cognitive bias

Introduction

Can digital games train people to avoid the use of problematic heuristics and cognitive biases in decision-making? There has been increasing interest in many fields to study the effects of digital games on people's cognition and behaviors (e.g., Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Eden, Maloney, & Bowman, 2010; Green & Bavelier, 2003; Peng, 2009; Peng, Lee, & Heeter, 2010; Reinecke, 2009; Ritterfeld, Cody, & Vorderer, 2009). Digital games, as an interactive medium, can afford experiential learning that may be especially effective in changing individual's cognitive processes and decision-making. The first goal of this study is to examine the use of experiential learning through a digital game to train people to avoid cognitive biases. The second goal is to compare a digital game to a more traditional training method, and determine the best-practices for improving people's cognitive processing during decision making.

Traditional decision-making theories often assume people make rational decisions by weighing the costs and benefits of their choices. However, despite people's best efforts to be rational, empirical evidence has shown people routinely make biased decisions, especially under conditions of uncertainty (Tversky & Kahneman, 1974). Decision makers often take into account irrelevant information that appear just prior to their decisions (Wegener, Petty, Blankenship, & Detweiler-Bedell, 2010). They frequently misunderstand the nature of statistical chances, believing their luck should change after a series of losses (Tversky & Kahneman, 1971). Moreover, people tend to heavily rely on stereotypical information over statistical base rates

(Argote, Devadas, & Melone, 1990). These are examples of cognitive biases that hinder rational decision-making.

Cognitive biases can be caused by overreliance on heuristics. Heuristics are generally conceptualized as cognitive shortcuts that help reduce people's cognitive load by decreasing the amount of information to process (Chaiken, 1980; Gigerenzer & Selten, 2002; Kahneman, 2011). According to Kahneman and Tversky (1973): "In making predictions and judgments under uncertainty, people do not appear to follow the calculus of chance or the statistical theory of prediction. Instead, they rely on a limited number of heuristics which sometimes yield reasonable judgments and sometimes lead to severe and systematic errors" (p. 237). In other words, heuristics do not always lead to biased decisions, but they can lead to less desirable outcomes in some situations.

Experts can also be susceptible to cognitive biases, and their biases can sometimes affect important decisions. The naturalistic decision making tradition posits that experts accumulate large amount of expert knowledge that allows them to quickly recognize problems and retrieve matching mental models. Expert chess masters or firefighters are able to make fast decisions intuitively based on their experience. These expert heuristics or intuitions can be highly accurate when the decision-making contexts are generally regular and the experts have sufficient skills to recognize the regular cues (Kahneman & Klein, 2009). However, when the problems are irregular, or when the decision-maker has not acquired sufficient skills, their familiar heuristic can lead to poor decisions with severe consequences. For example, Levinson (2006) found that cognitive bias can mislead judges and jurors into forming biased case-related memories. Cognitive biases can also affect medical doctors' diagnostic decisions concerning what types of tests to order, and what treatments to prescribe (Hicks & Kluemper, 2011).

Cognitive biases are difficult to avoid because people are often not aware of their existence. Even when people are aware of their biases, people rarely make efforts to inhibit them, given the convenience that heuristics offer in reducing cognitive deliberation (De Neys, Vartanian, & Goel, 2008). Although there are many studies identifying the causes and effects of cognitive biases (e.g., Epley & Gilovich, 2001; Strack & Mussweiler, 1997), there are relatively few studies on how to train and mitigate them. Those that do focus primarily on immediate effects rather than long-term mitigation.

To address this problem, this study seeks to examine the effectiveness of a digital game designed to promote experiential learning about cognitive biases, and evaluate the game's ability to mitigate cognitive biases in decision-making—both immediately and after a four-week delay. We argue that because cognitive bias functions as a heuristics, people need to not only learn about the biases, but also repeatedly practice avoiding the biases in order to change their heuristics. A traditional lecture can communicate knowledge but does not provide hands-on practice. A digital game provides both the experiential learning and practice. We conducted an experiment to investigate the effect of a game focusing on two specific cognitive biases: anchoring and representativeness bias. This study also contrasts different bias training methods by comparing the digital game to a traditional lecture, and to a combined condition.

Heuristics and Cognitive Biases

Human decisions are comparative in nature. When making decisions, we evaluate new information in relation to our existing beliefs, past experiences, and/or social norms (Mussweiler, 2003). Anchoring and representativeness bias are two common biases that can predispose human judgments by leading individuals to overemphasize certain seemingly-related information, while overlooking differences and other information that is both important and readily available.

Anchoring bias

Anchoring refers to the human tendency to rely on the first, most accessible piece of information available as a reference point, especially in numerical estimates. When the initial information (i.e., the “anchor”) is irrelevant or misleading, anchoring can lead to over- or under-estimation. For example, Tversky and Kahneman (1974) provided a random number between 0 to 100 by spinning a wheel to participants. After which they were asked to estimate the percentage of African countries within the United Nations. Even though participants were aware that the initial number was randomly generated, those receiving higher numbers gave higher estimates, whereas those receiving lower numbers gave lower estimates.

There are at least two explanations for why people are prone to use an initial value as a reference anchor. First, the anchoring-and-adjustment perspective argues an anchor value acts as a starting point for assessment, people then adjust their estimates towards a more plausible direction, but stop adjusting when the value reaches the edge of a plausible range, often insufficient to reach the actual value (Jacowitz & Kahneman, 1995). The second explanation, selective accessibility (Strack & Mussweiler, 1997), or the confirmatory search mechanism (Chapman & Johnson, 1994), holds that an anchor tends to be considered a plausible value, thus individuals are more likely to retrieve information confirming the anchor, while selectively making the confirmatory information more accessible in their evaluations.

Anchoring has been shown to be a robust effect that influences human judgment both in controlled laboratory settings and instances of real-life decision making such as price negotiation (Chapman & Johnson, 1999). Expertise in the subject matter does not seem to mitigate anchoring to any significant extent. For example, Northcraft and Neale (1987) gave professional real estate agents all the information commonly used to make property estimates, yet the agents' estimates were still influenced by the initial listing price: Those who received higher initial listing prices gave higher estimates relative to those who received lower ones.

Furthermore, anchoring occurs even when the anchors are obviously extreme and implausible. For example, one study asked participants to estimate the year Albert Einstein first visited the United States. After initially exposing them to an extremely low (1215) or high (1992) anchor, and although most participants knew the anchors were ridiculously implausible, their estimates were nevertheless still affected by them (Strack & Mussweiler, 1997). Similar effects were observed even when people were forewarned about anchoring bias and explicitly instructed to avoid using the anchor as a reference (Wilson, Houston, Etling, & Brekke, 1996).

Mitigating anchoring effects

Only a few studies have examined how anchoring might be avoided. Since anchoring occurs when people focus their attention on the initial number and use it as a reference point, then reminding them to consider alternative information while downplaying the importance of the initial anchor could be a useful countermeasure. Lord, Lepper, and Preston (1984) explicitly reminded subjects in their study to reevaluate each step of their judgments “had the results been the opposite” of what was presented. They found this *consider-the-opposite* reminder effective at mitigating biased processing relevant to several different types of decision-making errors.

Mussweiler, Strack, and Pfeiffer (2000) specifically tested the consider-the-opposite strategy on mitigating anchoring bias. In their experiment, they asked subjects whether the probability of Republicans winning the next presidential election is higher or lower than the last two digits of their social security number (an irrelevant anchor). The first group was then asked to list one reason why Republicans will win, the second group was asked to list one reason why Republicans will not win, and the third (control) group was not asked to list any reasons. The results showed subjects who listed reasons why Republicans will not win were least susceptible to anchoring bias because they considered the possibility of an opposite outcome, while the group that listed supporting reasons and the control group were both susceptible to anchoring.

The researchers then tested the consider-the-opposite strategy in a real-world scenario, asking car experts to estimate the value of a used car with information such as mileage, make, and also provided the actual car for the experts to inspect. Half of the experts were asked to list reasons why they think the asking price is too high or too low; the other half did not list the reasons. Again, the experts who listed reasons were less susceptible to the anchor. Mussweiler et al. (2000) argued that considering opposite explanations activated anchor-inconsistent information normally overlooked, which balanced out the anchor-consistent information that was activated by the initial anchor. According to the heuristics and bias explanation, anchoring bias occurs when individuals process information with their fast, intuitive system 1 (Kahneman, 2011). Consider-the-opposite may disrupt the fast heuristic processing of system 1 and activate system 2, which requires more cognitive efforts and information elaboration. These findings suggest making people aware of anchoring, and motivating its avoidance, can be successful at mitigating the anchoring effect.

Previous studies induced their consider-the-opposite manipulation by asking participants to actively generate anchor-inconsistent information. Another option that has thus far not been explored is whether consider-the-opposite may be equally effective if the anchoring-inconsistent information was generated by the researcher. If the researcher-generated inconsistent information is effective in mitigating anchoring bias, it would suggest that merely disrupting heuristics processed through system 1 can prompt individuals to process the information more carefully through system 2 and avoid making biased decision. This has great implications for mitigation trainings as the researcher or other-generated information requires much less motivation than the self-generated information. The present study sought to investigate whether consider-the-opposite strategy could effectively mitigate anchoring bias when the opposite options were generated by the researchers and were presented to the participants within a decision-making game. We also test whether the consider-the-opposite strategy has long-term mitigation effects.

Representativeness bias

Representativeness bias is the tendency to ignore probability while making judgments of people and situations. Tversky and Kahneman (1974) describe representativeness bias as the tendency to judge the probability of a hypothesis by considering how the hypothesis resembles available schema, rather than analyzing useful descriptive information. As Tversky and Kahneman (1974) note, most probabilistic questions tend to be framed in the form of “What is the probability object Y belongs to class Z?” or “What is the probability that event Y originates from process Z?” To answer such questions, people typically assess probabilities by the degree to which Y resembles or is representative of Z, so that to the extent Y is highly representative of Z, the probability that Y originates from Z will be judged to be high. Conversely, if Y does not appear to be similar to Z, the probability that Y originates from Z will be judged to be low—

regardless of a number of statistical indicators that could specify otherwise. Because heuristics are used in simplifying the process of assessing probabilities and predicting values, they tend to replace more reliable methods of calculating potential outcomes, which leads to several key deficits in decision-making, including (1) stereotypes, (2) insensitivity to sample size, (3) disregard for prior probability of outcomes (aka, base-rate fallacy), and (4) misconceptions of chance (aka, gambler's fallacy) (Kahneman, 2011; Kahneman & Tversky, 1972).

First, concerning stereotypes, if people make predictions solely in terms of the applicability of a stereotypic description, their predictions will be insensitive to the reliability of the evidence and to the expected accuracy of the prediction. For example, even if one knows that there is only a 10% chance of an older woman being a grandmother, but then hears her described as "warm and caring with a great love of young children," representativeness bias would encourage the assumption she is indeed a grandmother despite the 90% chance she is not.

Second, insensitivity to sample size misleads people to overlook errors when making decisions. A smaller sample will have more variance and is more prone to errors, whereas a larger sample has less variance and is more reliable. Even experienced researchers are often vulnerable to this bias, whereby small samples are deemed as representative of the populations as a large sample, leading to expectation that a population will be represented by a statistically significant result in even the smallest sample (Kahneman, 2011). For example, if on average, one out of five people (20%) are over six feet tall, is it more likely to obtain an average height of over six feet from a sample of 100 or 1000? Most students responded that the probabilities are the same (20%) when in reality, it is much more likely in the smaller sample (100) since it has more variance. The probability in the larger sample would resemble the actual probability more.

A third type of representativeness bias is base-rate fallacy, which causes people to ignore base rates and use current descriptive information to evaluate probabilities. For example, if there is a 10 percent chance of a student having a cold (the base rate), you select two samples of three students, and learn that one out of three students in the first sample has a cold. When estimating the probability of a student from the second sample having a cold, representativeness bias would influence you to ignore the base rate of 10% and estimate 33% (i.e. one out of the three). When a descriptive information is available, people tend to believe the description over their knowledge of the base rate (Kahneman, 2011).

A fourth type of representativeness bias, often referred to as the gambler's fallacy, involves misconceptions of chance. People tend to expect a sequence of events generated by a random process to represent the essential characteristics of that process. Such a misconception commonly views chance as a self-correcting process (Kahneman, 2011). For example if a fair coin lands heads-up nine times in a row, the next time it is tossed, gambler's fallacy would lead one to expect a higher chance for tails-up—rather than the correct 50-50 chance.

Mitigating representativeness bias

Although representativeness bias can be resistant to mitigation, literature suggests several methods of alleviating each of its aspects. For example, Tremblay (2007) found two techniques, a scenario-based approach (in which decision makers assess “what if?” scenarios) and benchmarking (where a decision-rule is established prior to decision making) to be effective in aiding decision makers by mitigating insensitivity to sample size. Another method involves prompting a search for disconfirming evidence (Einhorn & Hogarth, 1978), and still another, within the context of repeated judgments, entails making the potential for representative errors salient. Similarly, other strategies include highlighting the potential unreliability of a description,

searching for disconfirming evidence, searching for anomalies, considering specific exceptions, and making the bias salient by reminding decision makers about its deleterious effects on judgment (Silverman, 1992).

Digital games for mitigating cognitive bias

Cognitive biases are incredibly difficult to mitigate because most people are unaware of their own biases (Ehrlinger, Gilovich, & Ross, 2005; Kahneman & Klein, 2009). Even if people are aware of their biases, avoiding bias takes effort and practice, which is why simply reminding people about their biases has little effect on mitigating them (Wilson & Brekke, 1994).

We propose that digital games have great potential as a training medium for mitigating cognitive bias in decision making. From a cognitive learning perspective, the interactive nature of digital games require learners to actively process the information which promotes problem-solving transfer to real decision-making scenarios (Wouters, Van Nimwegen, Van Oostendorp, & Van Der Spek, 2013). Digital games can also provide immediate feedback for learners to assess and correct their decisions through multiple representations (Mayer & Johnson, 2010; Moreno & Mayer, 2007)

However, as an interactive medium, digital games can also demand a large amount of cognitive resources to process the interacting controls, narratives, and feedback (Moreno, 2004). Without sufficient knowledge or support, learners can easily be overwhelmed by the amount of information communicated simultaneously through multiple modes. In comparison to digital games, traditional lectures may offer more direct communication of bias-relevant knowledge since students are usually familiar with such methods of learning. However, when dealing with hard-wired heuristics, these traditional modes of training may not provide learners with sufficient opportunities to reflect on personal biases, or practice their bias mitigation techniques. Thus, they

may result in superficial, short-term understanding, rather than deeper internalization. Are digital games more effective than traditional lectures? Meta-analyses which made this type of comparison between interactive games and traditional learning have found that mixed results (Wouters et al., 2013). This reasoning suggests the following research questions:

RQ1. Can a digital game be effective at mitigating anchoring bias?

RQ2. Can a digital game be effective at mitigating four aspects of representativeness bias:

(a) stereotypes, (b) insensitivity to sample size (c) base-rate fallacy, and (d) gambler's fallacy?

While digital games applied alone may or may not be more effective than traditional instructional methods. Meta-analyses results have suggested that digital games supplemented with traditional instructional methods often yield higher learning gains than when the game is the only source of learning (Wouters et al., 2013). The improved effects are especially significant when the game is designed for skill-based training, and when the supplemental instructions help learners identify new relevant information (Wouters & Van Oostendorp, 2013). We believe a combination of passive (i.e., lecture) and interactive (i.e., digital game) training modes should be most effective, by first providing a framework the cognitive biases through the lecture, learners can gain a basic understanding of the biases and learn to identify bias-related information from non-related ones. After the lecture, the training is most effective followed up by a digital game capable of offering multiple scenarios for the learner to observe the consequences of their biased decision-making while practicing mitigation strategies. Such a combined approach should be most effective at activating bias awareness and mitigation, thus we posit the following hypotheses:

- H1.** Relative to slideshow lecture or digital game alone, their combination (a lecture followed by game play) is more effective at mitigating anchoring bias.
- H2.** Relative to slideshow lecture or digital game alone, their combination (a lecture followed by game play) is more effective at mitigating four aspects of representativeness bias: (a) stereotypes, (b) insensitivity to sample size (c) base-rate fallacy, and (d) gambler's fallacy.

Method

Participants and procedure

A three-way between-group (training modes: slideshow vs. game only vs. combined) experimental design was utilized to test the hypotheses. A priori power analysis was conducted with G*power, with $\alpha = .05$ and power $(1-\beta) = .80$, the projected sample size needed with this effect size was approximately 158. A total of 335 university students were offered extra course credit for their participation in the study. Upon arriving in the computer lab, participants were assigned to a computer cubicle. After giving informed consent, they were directed to complete a pretest questionnaire measuring their anchoring judgment skills and proneness to representativeness bias. The survey included questions concerning participants' experience with strategy and simulation games, their computer comfort level, and demographic items such as gender, age, and education. Participants were then randomly assigned to one of the three training modes: Those in the slideshow condition watched a 14-minute presentation explaining the biases along with examples of mitigation strategies. Those in the game-only condition played a digital game designed to train them about the biases and bias mitigation for four gameplay scenarios or until 60 minutes passed. Those within the combined condition (slideshow plus game) watched the slideshow first, then played the game.

Although the total exposure time of 14 minutes in the slideshow condition was much shorter than the 60 minute of maximum possible exposure in the game condition, the game involved interactive elements such as moving the avatar to navigate the game environment which was not directly involved with communicating bias-related information of mitigation strategies. It should be noted the slideshow content was comparable to the educational content in the game, the same description of the biases and similar examples were used in both the slideshow and the digital game. Therefore, the content of the training was matched between the game and the slideshow, as were the number of examples presented, even though the game took longer to deliver the content. As a result, the bias-related content is controlled for, and the comparison is of the different instruction modes.

After the training, all participants filled out an immediate posttest survey with another set of anchoring and representativeness bias measures. Four weeks later, participants received an email containing a link to the delayed posttest survey assessing the same constructs as the immediate posttest survey.

Stimuli

The digital game

The digital game used in this study was designed by the research team to mitigate anchoring bias and the four different forms of representativeness bias (See Figure 1 for a screenshot of the game). In the game (name removed) players take the role of an intelligence analyst who is in charge of directing a field agent. The field agent is tasked with investigating a series of international trafficking cases, and must search for useable intelligence about a potential threat. Players direct the agents to points of interest within the facilities and receive intel from the field agents. Players must then evaluate the intel to determine if it is biased, and identify the

bias involved, whereupon they must instruct the field agent to collect only unbiased intel. Once enough unbiased intel is collected, players are tasked with considering the collected intel and making a determination as to the threat status of the subject being investigated. Feedback about the biases is provided to the players during the game when they make decisions about whether to include or reject a piece of intel. The game provides a short description of the biases at the beginning to provide some basic background on their nature.

Anchoring bias training is presented in the form of a multiple-choice question. Before making an estimate, players receive a question asking them to consider opposite possibilities. One of the options is *yes*; the initial anchor is consistent with the target estimate. The other three options are *no*, with potential reasons why that initial anchor is unrelated to the target. This process is similar to the consider-the-opposite strategy, the only difference being the items of anchor-inconsistent information are not self-generated by the participant, but rather provided to them for their consideration (See Figure 2).

Two strategies were employed to mitigate the four aspects of representativeness bias. The first strategy addressed base rate fallacy by having players read descriptions of suspicious intel in which each description included a base rate number and a description as to why the item was suspicious. After reading these descriptions, players were encouraged to consider the base rate rather than the description of the item. A second strategy addressed both gambler's fallacy and insensitivity to sample size: Players were given an example of two similar situations, then given a conclusion that followed the pattern of the previous two situations. Players were then encouraged to reject that conclusion based on small sample size and to consider the probabilities of various patterns, thereby addressing gambler's fallacy as well.

The slideshow lecture

A 14-minute slideshow lecture was created to simulate a more traditional training method, and to act as a comparison for the interactive learning experience provided by the game. To make the slideshow comparable to the game, it included definitions of the biases and examples matching those used in the game to ensure the message content was comparable. To further ensure comparability of the two conditions, the slideshow also included practice scenarios in which the voiceover explained strategies for mitigating the biases similar to the instructions and feedback delivered in the game.

The combined condition

Many previous studies have found that digital games are most effective when supplemented with a knowledge framework to help learners identify relevant information and reduce cognitive demands (Mayer & Johnson, 2010; Moreno, 2004; Wouters & Van Oostendorp, 2013). Participants in the combined condition were instructed to watch the slideshow lecture first, followed by the game.

Measurements

Anchoring bias measures

Similar to previous studies (Jacowitz & Kahneman, 1995; Kahneman & Tversky, 1979; Mussweiler, Englich, & Strack, 2004), anchoring bias was measured by first asking participants to evaluate an anchor. For example, “is the tallest tree in the world taller or shorter than 3,000 feet?” After exposure to the anchor, participants were asked to make a numerical estimate. In this experiment, the targets consisted of pictures of a jar of jellybeans, a jar of pennies, and a swimming pool, and participants were asked to estimate the number of jellybeans (pretest), the amount of pennies (immediate posttest), and the amount of water in the swimming pool (4-week posttest). The anchors were intentionally selected to be irrelevant to the estimated target. We

chose irrelevant anchors because any reliance on them would clearly demonstrate bias. The target pictures were also selected because they did not have correct answers, thus eliminating the effects of previous knowledge on elicited estimations. The reliability of these three picture tests was assessed via a pilot test, which revealed a Cronbach's α of .78, demonstrating good reliability across the three tests.

Representativeness bias measures

Representativeness bias mitigation was assessed using four instruments measuring gambler's fallacy, insufficient data, base-rate fallacy, and stereotypes. The first measure of representativeness bias assessed gambler's fallacy with four items adapted from Cox and Mouw (1992). Participants were asked to select the most likely option after reading a scenario depicting 50:50 odds. The four items were summed with higher numbers indicating less bias. Item reliability measured with Cronbach's α was .43 in the pretest, .66 in the immediate posttest, and .78 in the 4-week posttest. Note that reliabilities improved after training in the two posttests, which suggests participants were guessing randomly in the pretest, as would be expected.

The second representativeness bias measure assessed base-rate fallacy with four items also adapted from Cox and Mouw (1992). Participants were asked to select the most likely option of three choices after being presented with a scenario. For example:

A group of 100 professionals takes a personality questionnaire. The group has 20 used car salespeople and 80 museum curators. You pick out one person and see that the person is extroverted and aggressive, places a high value on his appearance, and enjoys debates.

Which of the following is more likely?

- He is a used car salesperson. (1)
- He is a museum curator. (2)

- Need more information. (3)

Items were scored such that the biased response, or the stereotypical option, was coded as a 1, the response with the largest base-rate presented in the scenario was coded as 2, and the option to collect more information was coded as 3 (because it indicated systematic decision-making). The mean of all four items was computed. Again, improving reliabilities suggest guessing in the pretest: $\alpha = .42$ (pretest), $\alpha = .71$ (immediate posttest), and $\alpha = .71$ (4-week posttest).

The third measure of representativeness bias tested participants' proneness to settle for insufficient data by presenting them with scenarios of individuals making poor decisions based on insufficient data. For example:

Jim went golfing for the first time and finished playing 18 holes under par. Jim's wife signs him up for a local golf tournament a couple of weeks later, certain that he will win the \$20,000 first prize. How confident are you in her evaluation?

Participants were then asked how confident they were in the individual's decision using a 5-point Likert scale; whereupon the mean score was calculated for inclusion within the analysis. Item reliability measured with Cronbach's α once again improved in the posttests: $\alpha = .58$ (pretest), $\alpha = .72$ (immediate posttest), and $\alpha = .75$ (4-week posttest).

Finally, the fourth measure of representativeness bias assessed stereotypes. Because no reliable, valid measure for assessing the stereotype aspect of representativeness bias exists, a measure was developed specifically for this study. This new scale offered a brief scenario followed by six response options, three of which indicated diagnostic responses (coded +1), and three of which indicated stereotypical responses (coded -1); scores ranged between -3 and +3, with lower scores denoting higher levels of representativeness bias. We will refer to this new representativeness bias scale as NewRep. See Appendix A for the full NewRep scale items.

Results

Thirty four participants who did not complete the required training were removed, only the 301 participants ($M_{\text{age}} = 21.00$, $SD = 3.15$; 54.5% female) who completed all three questionnaires were included in the analyses.

Anchoring Mitigation

To test RQ1 and H1 concerning anchoring bias mitigation, a mixed-model analysis of covariance (ANCOVA) was conducted with training mode entered as a between-subjects factor, and time period (pretest, immediate posttest, and 4-week posttest) entered as a within-subjects factor; standardized anchoring bias scores were used as the dependent variable. The anchoring bias measures were centered to the anchors so that deviation from the anchor indicated bias mitigation since moving away from the anchor is desired. The values were standardized to z values to allow comparison between the three time periods. ANCOVA results showed a significant main effect for time period, $F(2, 208) = 3.39$, $p = .038$, $\eta_p^2 = .07$. There was also a significant main effect for training mode, $F(2, 208) = 2.82$, $p = .031$, $\eta_p^2 = .03$. Pairwise comparisons using Scheffé's correction indicated significant differences between the combined slideshow plus game condition, and the slideshow only condition in the immediate posttest ($p = .039$), and between the combined condition and the other two conditions (slideshow only [$p = .039$] and game only [$p = .041$]) in the 4-weeks posttest. See Table 1 for the means and standard deviations.

As Figure 3 illustrates, mitigation effects were not significant for the slideshow or digital game conditions alone. However, the combined application of the slideshow lecture and the digital game reduced anchoring bias significantly immediately after the stimuli, and the mitigation effects did not diminish over time.

[Insert Figure 3 here]

Representativeness Bias Mitigation

To test for representativeness bias mitigation (RQ2 and H2), four separate mixed-model ANCOVAs were conducted for each dependent variable: the NewRep (measuring stereotypes), insufficient data scores, base-rate fallacy scores, and the gambler's fallacy scores. All analyses included training mode (slideshow, digital game, and slideshow plus game) as the between-subjects factor and test period (pretest, immediate posttest, and 4-week posttest) as the within-subjects factor.

The first analysis tested for mitigation of stereotype. Although all the conditions showed improvement in bias mitigation over the time periods ($F [2, 291] = 3.38, p = .018, \eta_p^2 = .02$). There was no significant difference between the three conditions in the immediate posttest. However, after four weeks, the mitigation effects diminished in both the slideshow only and the digital game only conditions, but not within the combined slideshow plus game condition, $F (2, 292) = 5.18, p = .003, \eta_p^2 = .03$. See Table 1 for the means and standard deviations. The effect size was small but significant. See Figure 4.

The second analysis tested for mitigation of insensitivity to sample size, and again, the effect of training mode was significant with a small effect size, $F (2, 285) = 3.02, p = .025, \eta_p^2 = .02$. Both the digital game only and the combined condition improved mitigation significantly in the immediate posttest, but the effects diminished slightly after four weeks. Surprisingly, participants within the slideshow only condition performed worse in the immediate posttest, although the detrimental effect appeared to have diminished after four weeks. See Figure 5 and Table 1 for mean comparisons.

The third analysis tested the mitigation effect of base rate fallacy, and results showed that although there was a main effect in time periods, $F(2, 285) = 54.57, p < .001, \eta_p^2 = .16$ (medium effect size), there was no significant difference between the conditions, $F(2, 285) = .28, p = .376$. (See Figure 6 and Table 1), indicating the three conditions were not significantly different in mitigating base rate fallacy.

The fourth analysis examined mitigation of gambler's fallacy, and the effect of training mode was again significant but small, $F(2, 285) = 6.20, p = .002, \eta_p^2 = .04$. Pairwise comparisons using Scheffé's corrections indicated there were no significant differences between the three conditions in the immediate posttest. Nevertheless, there was a significant difference between the game-only ($M = 2.97, SD = .10$) and the combined condition ($M = 3.63, SD = .18$) in the 4-week posttest, $t(119.02) = -3.75, p < .001$, indicating the combined condition performed better than the game-only condition in the 4-week posttest. Unfortunately, there were no significant differences between the three time periods, $F(2, 285) = .14, p = .71$. The findings suggest that while the combined condition seemed to have performed better than the game-only condition in mitigating gambler's fallacy in the 4-week posttest, the amount of mitigation did not appear to be significantly different across the three time periods (See Figure 7 and Table 1). In other words, the different trainings were not effective in mitigating gambler's fallacy.

Discussion

The goals of this study were twofold. First, we attempted to investigate the effectiveness of a digital game at mitigating anchoring and representativeness bias. Second, we explored the effectiveness of different modes of training by comparing the digital game to a slideshow lecture and a combined condition in which learners were exposed to the slideshow lecture followed by the digital game. We expected the slideshow lecture to provide learners with a framework for

bias mitigation, but insufficient practice. On the other hand, we predicted that a digital game would provide better context for practicing bias mitigation techniques, but perhaps at the risk of overloading learners with too many bias-unrelated information.

Our research questions asked whether the digital game could be effective at mitigating anchoring bias and four different forms of representativeness bias. Results showed the digital game alone to be effective at mitigating three of the four types of representativeness bias immediately after the training (with the exception of gambler's fallacy), but not anchoring bias. Similarly, the slideshow condition mirroring the traditional training was effective at mitigating the same three types of representativeness bias immediately after training, but this effect was short-lived, decaying markedly after four weeks.

Perhaps one reason for the non-significant effects of the digital game alone can be attributed to its non-linear, multimedia nature. Digital games can support multiple representations of decision-making scenarios for learners to practice bias mitigation techniques, but this process of navigating multiple representations can be burdening to the learners' cognitive ability (author citation). In particular, learners with little prior knowledge of the biases may have difficulty identifying the important bias-related concepts in the game from game-related information unrelated to cognitive bias. Studies have shown that learners with little prior knowledge often focus on surface features and have trouble identifying conceptually relevant items in a multimedia representation (Lowe, 1996). Repeated gameplay sessions should reduce this factor as players become more familiar with the game and can better distinguish bias-related information from gameplay information unrelated to training.

The traditional slideshow lecture was an effective mode of training at mitigating three types of representativeness bias in the immediate posttest, and its effect did not differ from the

game-only condition for mitigating anchoring bias. This may not be surprising for two reasons: First, a lecture supported by slideshows is a familiar mode of learning for students. Unlike digital games, most undergraduate students have taken lecture classes that utilize slideshows to illustrate different concepts. Therefore they know how to search and identify information communicated through this mode, making it more effective immediately. The second reason is that a slideshow lecture provides more explicit explanations about biases without the mechanics, narrative, and compound representations of a digital game. This should reduce the cognitive load placed on learners by only presenting biases-related knowledge and their mitigation techniques. But while the mitigation effect of the slideshow lecture was effective immediately after the training, it also decayed the most after four weeks. Simply learning about the biases was not enough to change people's decision-making process in the long run. People are also less likely to watch a slideshow lecture repeatedly compared to playing a digital game.

We hypothesized that combining the affordances of the slideshow with the digital game would be more effective at mitigating cognitive bias than the slideshow or the game alone. This is because a slideshow lecture can provide learners with basic concepts, whereas a game can further provide opportunities for exploration and practice in engaging the mitigation techniques in simulated decision-making scenarios. The findings from this experiment provide support for our assumption: The combined slideshow plus game condition significantly reduced anchoring bias and three types of representativeness bias (except gambler's fallacy) immediately after training. Moreover, the mitigation effects persisted even after four weeks. Consistent with H1, the slideshow plus game condition was significantly better than the slideshow alone and digital game alone conditions at mitigating anchoring bias both in the immediate posttest and at the 4-week posttest. This most effective method of training combines the best of both, offering the

direct learning style along with reduced cognitive load in playing the game because players are already familiar with the content and can focus on practicing decision-making through the game.

In terms of representativeness bias mitigation, the mitigation effects of the combined condition actually increased after four weeks. We believe this may be because learning through the problem-solving experiences afforded by digital games requires active construction of mental representations, which may take place during gameplay or after the game through reflection (Dunbar et al., 2014). As a result, the mitigation effect immediately after gameplay was not significantly different than from the more static slideshow method. However, the reflection and active construction process allowed learners to gain a deeper understanding of the biases and modify their decision making process, which led to significantly greater mitigation after four weeks. Our overall results were consistent with both H1 and partially consistent with H2.

Limitations

This study has several limitations. First, the comparison between the three modes of training is a practical comparison. We chose to compare the digital game to a slideshow lecture because lectures are a common method used for training. Although we controlled content equivalence by using comparable content in both the digital game and the slideshow, these two communication modes are nevertheless different in many other ways. Therefore, the findings from this study should not be interpreted as resulting from any single construct unique to the two communication modes, but as an overall comparison of the different affordances of the media. Second, we did not test for the condition in which the game is introduced before the slideshow lecture, or a repeated gameplay condition, which would be more amenable to the nature and advantage of the digital game medium. Testing such conditions would provide more insight into the best practice for decision-making training, and also further tease apart the theoretical nuances

involved in starting learners with an informational framework, and then engaging them in more interactive problem solving activities. Thirdly, since the study made multiple comparisons among the variables, this procedure can risk increasing the chances for Type 1 error. In order to address this risk, we used Scheffé's method for comparison to correct the significance level.

Conclusion

The goal of this study was to investigate the effectiveness of a digital game in mitigating anchoring and representativeness biases, as well as recommend best practices for applying games for bias mitigation training. By comparing a digital game to a slideshow, and a combined condition, we were able to demonstrate the benefits of a multimodal approach. Overall, the findings suggest a digital game can effectively mitigate cognitive bias, especially when learners have a sufficient understanding of the cognitive biases involved in the gameplay. Compared to a slideshow lecture, the digital game used in this study was not more effective as a standalone training tool. Without sufficient knowledge about the biases, players are likely to feel overloaded and distracted by the more complex game elements, and find it more difficult to identify bias-relevant information. This is consistent with research on learning in multimedia environments, suggesting learners may not be able to take advantage of the multiple problem representation without prior knowledge to help them navigate the environment and identify relevant information (Moreno, 2004; Rey & Buchwald, 2011). Although the traditional slideshow lecture training seems to be effective in mitigating several representativeness biases immediately after the training, the mitigation effects were short-lived and decayed drastically after four weeks. Thus, future studies testing cognitive bias effects and mitigation would be advised to measure both short and long term effects.

Combining the slideshow lecture with the digital game led to the most effective bias mitigation, which did not decay after four weeks. One possible explanation might be that the slideshow lecture provided learners with a basic understanding of the biases and the techniques for bias mitigation, which reduced strain on cognitive resources when that information is brought into the interactive environment of the digital game. Such an approach allows for more effective identification of biased decision-making within the various scenarios, along with the practice of more effective mitigation techniques. Future studies should incorporate measures of cognitive load to determine more effective ways of communicating the complex bias-related information to learners without overloading their cognitive capacity.

One key question that needs to be addressed is whether the mitigation skills acquired from the game or combined training can be transferred to actual decision-making scenarios. We believe the answer is yes, but to a certain degree. Whether a heuristic intuition leads to fast, accurate judgements or biased decisions depends on the context and available cues (Kahneman & Klein, 2009; Shanteau, 1992). Individuals who routinely face regular decision-making scenarios can gradually build heuristics that allow them to make fast, accurate decisions without much thought. However these heuristics can mislead their judgements if they are faced with an unfamiliar scenario or misleading cues. For example, Ricks, Turley-Ames, and Wiley (2007) found that experts were more likely to provide incorrect answers when the problems seemed familiar, but were in fact different. Non-experts on the other hand made more accurate decisions. This may be because experts quickly recognized the seemingly familiar scenario and made the decisions based on their heuristics, failing to recognize that the problem was actually different. In comparison, since non-experts do not have existing heuristics, they were more likely to process the problems cautiously, consider more information, and answer correctly. The stimulus

materials tested in this study were designed to mimic intelligence analysts' work, which involves processing multiple sources of intel and deciding whether to seek more information. The decision to mirror intelligence analysts' work was a deliberate one. The goal is to develop skills in digital media that can be transferred to improve intelligence analysts' decision-making process. By providing multiple scenarios for players to practice repeatedly, we hope that they will learn the potential biases in their actual work, learn to recognize cues that can trigger cognitive bias, and apply the strategies they learned in the trainings to avoid making biased decisions. However, this does not suggest that the players will be bias-free. They may still fall for cognitive biases in other decision-making contexts that are dissimilar to their work, or when the cues are less available. In other words, we argue that training which combines digital games can mitigate context-specific biases by teaching players how to recognize cues and offer experiential learning opportunities to practice bias mitigation strategies within similar scenarios.

In practical terms, this study demonstrates how cognitive biases can be effectively mitigated through interactive media representations of decision-making scenarios. The findings also provide further support for the effectiveness of the consider-the-opposite technique at mitigating anchoring bias. Prior studies have asked decision-makers to generate opposite possibilities to mitigate anchoring effects. Our study showed that consider-the-opposite was effective even when the anchor-inconsistent information was not generated by the decision-maker but provided to them. Unlike self-generated inconsistent information, which requires higher motivation and capacity. This finding suggests that merely disrupting heuristic processing can mitigate anchoring bias to a certain degree. From a training or operational perspective, the findings imply that designing prompts to remind decision-makers to consider-the-opposite can be an effective way to help decision-makers avoid anchoring bias.

From an educational-game design perspective, these results indicate the digital game was only effective in the long run when paired with a traditional slideshow lecture. This finding underlines the importance for game designers to consider the context in which their games will be applied, along with the players' level of prior knowledge of the subject matter. For training material that is new and complex, game designers are advised to provide some introductory background materials to help fortify learning before gameplay begins. Along similar lines, game designs should include more linear and direct instructional materials built into their games upfront that can provide learners with a basic understanding of the underlying concept. This basic understanding may help prevent player disorientation due to excessive cognitive load in an interactive multimedia environment.

Reference

- Argote, L., Devadas, R., & Melone, N. (1990). The base-rate fallacy: Contrasting processes and outcomes of group and individual judgment. *Organizational behavior and human decision processes*, 46(2), 296-310. doi:10.1016/0749-5978(90)90034-7
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752-766. doi:10.1037/0022-3514.39.5.752
- Chapman, G., & Johnson, E. (1994). The limits of anchoring. *Journal of Behavioral Decision Making*, 7(4), 223-242. doi:10.1002/bdm.3960070402
- Chapman, G., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational behavior and human decision processes*, 79(2), 115-153. doi:10.1006/obhd.1999.2841
- Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2), 661-686. doi:10.1016/j.compedu.2012.03.004

- Cox, C., & Mouw, J. T. (1992). Disruption of the representativeness heuristic: Can we be perturbed into using correct probabilistic reasoning? *Educational Studies in Mathematics*, 23(2), 163-178.
doi:10.1007/BF00588054
- De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter Than We Think When Our Brains Detect That We Are Biased. *Psychological science*, 19(5), 483-489. doi:10.1111/j.1467-9280.2008.02113.x
- Dunbar, N. E., Miller, C. H., Adame, B. J., Elizondo, J., Wilson, S. N., Lane, B. L., ... & Lee, Y. H. (2014). Implicit and explicit training in the mitigation of cognitive bias through the use of a serious game. *Computers in Human Behavior*, 37, 307-318.
- Eden, A., Maloney, E., & Bowman, N. D. (2010). Gender attribution in online video games. *Journal of Media Psychology: Theories, Methods, and Applications*, 22(3), 114. doi:10.1027/1864-1105/a000016
- Ehrlinger, J., Gilovich, T., & Ross, L. (2005). Peering Into the Bias Blind Spot: People's Assessments of Bias in Themselves and Others. *Personality and Social Psychology Bulletin*, 31(5), 680-692.
doi:10.1177/0146167204271570
- Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological review*, 85(5), 395. doi:10.1037/0033-295X.85.5.395
- Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological science*, 12(5), 391-396. doi:10.1111/1467-9280.00372
- Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox*: Mit Press.
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, 423(6939), 534-537. doi:10.1038/nature01647
- Hicks, E. P., & Kluemper, G. T. (2011). Heuristic reasoning and cognitive biases: Are they hindrances to judgments and decision making in orthodontics? *American Journal of Orthodontics and Dentofacial Orthopedics*, 139(3), 297-304. doi:10.1016/j.ajodo.2010.05.018.

- Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, 21, 1161-1166. doi:10.1177/01461672952111004
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.
- Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree. *American Psychologist*, 64(6), 515. doi:[10.1037/a0016755](https://doi.org/10.1037/a0016755)
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3), 430-454. doi:10.1016/0010-0285(72)90016-3
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review*, 80(4), 237. doi:10.1037/h0034747
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.
- Levinson, J. D. (2006). Forgotten racial equality: Implicit bias, decision-making and misremembering. *bePress Legal Series*, 1630.
- Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: a corrective strategy for social judgment. *Journal of Personality and Social Psychology*, 47(6), 1231-1243.
- Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European journal of Psychology of Education*, 11(4), 377-397. doi:10.1007/BF03173279
- Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research*, 42(3), 241-265.
- Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science*, 32(1-2), 99-113.
- Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, 19(3), 309-326. doi:10.1007/s10648-007-9047-2
- Mussweiler, T. (2003). Comparison processes in social judgment: mechanisms and consequences. *Psychological review*, 110(3), 472. doi:10.1037/0033-295X.110.3.472

- Mussweiler, T., Englich, B., & Strack, F. (2004). Anchoring effect. *Cognitive illusions—A handbook on fallacies and biases in thinking, judgment, and memory*, 183-200.
- Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, 26(9), 1142-1150. doi:10.1177/01461672002611010
- Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational behavior and human decision processes*, 39(1), 84-97. doi:10.1016/0749-5978(87)90046-X
- Peng, W. (2009). Design and evaluation of a computer game to promote a healthy diet for young adults. *Health communication*, 24(2), 115-127.
- Peng, W., Lee, M., & Heeter, C. (2010). The Effects of a Serious Game on Role-Taking and Willingness to Help. *Journal of Communication*, 60(4), 723-742. doi:10.1111/j.1460-2466.2010.01511.x
- Reinecke, L. (2009). Games and recovery: The use of video and computer games to recuperate from stress and strain. *Journal of Media Psychology: Theories, Methods, and Applications*, 21(3), 126. doi:[10.1027/1864-1105.21.3.126](https://doi.org/10.1027/1864-1105.21.3.126)
- Rey, G. D., & Buchwald, F. (2011). The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied*, 17(1), 33 - 48. doi:10.1037/a0022243
- Ricks, T. R., Turley-Ames, K. J., & Wiley, J. (2007). Effects of working memory capacity on mental set due to domain knowledge. *Memory & Cognition*, 35(6), 1456-1462. doi:10.3758/BF03193615
- Ritterfeld, U., Cody, M. J., & Vorderer, P. (2009). *Serious games: Mechanisms and effects*: Taylor & Francis.
- Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational behavior and human decision processes*, 53(2), 252-266. doi:10.1016/0749-5978(92)90064-E
- Silverman, B. G. (1992). Modeling and critiquing the confirmation bias in human reasoning. *Systems, Man and Cybernetics, IEEE Transactions on*, 22(5), 972-982.

- Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3), 437.
- Tremblay, M. C. (2007). *Uncertainty In The Information Supply Chain: Integrating Multiple Health Care Data Sources*. (PhD), University of South Florida. Retrieved from <http://en.scientificcommons.org/56477363>
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, 76(2), 105. doi:10.1037/h0031322
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124-1131. doi:10.1126/science.185.4157.1124
- Wegener, D. T., Petty, R. E., Blankenship, K. L., & Detweiler-Bedell, B. (2010). Elaboration and numerical anchoring: Implications of attitude theories for consumer judgment and decision making. *Journal of Consumer Psychology*, 20(1), 5-16. doi:10.1016/j.jcps.2009.12.003
- Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: unwanted influences on judgments and evaluations. *Psychological bulletin*, 116(1), 117. doi:10.1037/0033-2909.116.1.117
- Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 125(4), 387. doi:10.1037/0096-3445.125.4.387
- 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.
- Wouters, P., & Van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, 60(1), 412-425. doi:10.1016/j.compedu.2012.07.018