CHAPTER 15.

SHIFTING IMPPLICIT BIASES WITH GAMES USING PSYCHOLOGY

The Embedded Design Approach
BY MARY FLANAGAN AND GEOFF KAUFMAN

We know ablehe drill: video games don’t do a good job on the diversity front yet. They don’t, for example, excel in representing female characters fairly. Case in point: Grand Theft Auto V simulates sex with prostitutes, introduces players to the denizens of seedy strip clubs, and showcases scantily clad vixens wandering the streets; these, the only female characters, have little agency, and exist merely as potential sexual partners or targets of violence. Exposure to such sexualized characters has been shown to produce a number of disturbing effects, including increasing players’ acceptance of myths and false assumptions about rape (Fox, Ralston, Cooper, & Jones, 2014). Black characters in games don’t fare any better—often they are depicted as gang members and/or criminals. These skewed representations foster negative attitudes and reinforce negative stereotypes toward African American males encountered outside the context of the game (Yang et al., 2014).

Much recent scholarship is uncovering just how potent such oversimplified, homogeneous representations across media forms can be. Stereotypic representations of race, sexuality, and other social identity groups are highly troubling; in games, these limit players’ perceptions of members of those groups and their abilities (Sherman & Zurbrigggen, 2014). The findings that have emerged from these studies are not surprising, given that there is a continued lack of diversity within the game industry itself. While digital game audiences have been diversifying—in 2014, 48% of US players were female (ESA, 2014) and Hispanic players outnumber non-Hispanic players, particularly in the use of console games (and Hispanic players show an interest in a wider range of gaming platforms than do their non-Hispanic counterparts) (Mintel, 2014)—games are still typically created by
white male designers (IGDA, 2014). Homogeneity among game creators, as well as the repetition of existing stereotypical depictions of characters, cultures, and worlds sets in motion a vicious cycle of implicit bias that pervades game design, game play, and the broader game culture.

Game makers are awkwardly located at the crux of the matter. Imagine if looking at images were so socially out of place, so culturally loaded, that only specific groups made or were empowered to look at images. Game making and games journalism can sometimes feel this way. This is unfortunate, for sales figures, societal practices, and public discourse reveals that games are a key 21st century art form—right now, they are media culture. Therefore, not only can designers be mindful of the ways their creations may (often unintentionally and subtly) reinforce implicit bias, but, indeed, designers can (and should) design in ways that can effectively counteract stereotypes and discrimination through play. Can a psychological understanding of bias help designers expand their perspectives to make diverse games for diverse audiences?

There have been relatively few attempts to amass, synthesize, and share what has been learned from psychological studies relating to diversity and apply these to the domain of game design. This chapter begins to fill this gap with an overview of what we currently know—and what we ought to know as participants in, and contributors to, game culture. There are several approaches to promoting ‘diversity’ in games, from increasing the audience for games to attempting to shift biases in games—and in players themselves. What are the ways that we can think scientifically about bias and design to shift biases, both in who designs and plays games and in the content and potential outcomes for that play? Established theories and decades of empirical work in psychology provide an illuminating knowledge base from which to examine issues of diversity in games—as well as an accessible set of techniques and tools for the conscientious design of games to counteract or circumvent social biases.

One key root cause of the imbalanced participation and lack of equity and diversity inherent in game play, game design, and game culture, is the phenomenon of implicit bias and the psychological obstacles that it entails. By definition, implicit bias refers to an unconscious negative evaluation or association that gets incorporated into one’s mental representation (or “schema”) of a particular concept (be it a person, group, place, event, idea, value, etc.). What’s particularly insidious about implicit bias is that it does not depend on (and often occurs in spite of) an explicit endorsement or intention to have a negative attitude toward a particular target (Blair, 2002).
No one is immune to implicit biases. Even those who genuinely and legitimately consider themselves to be non-prejudiced and committed to instilling human values such as egalitarianism at a conscious level could, at the same time, hold a negative association toward individuals of another group at an implicit level (Beattie, et al., 2013; Flanagan & Nissenbaum, 2014). In fact, it is likely that most biases go unnoticed and even “unbelieved” (Pronin et al., 2002).

In order to craft and disseminate those new messages effectively, we—as serious game designers, researchers, players, and consumers—must have a clearer understanding of how implicit biases manifest and operate in game play, game design, and game culture. While implicit biases emerge from the world around us and alter people’s experiences and ideas about themselves and others, the extraordinary thing about them is that they are not fixed. As we shall explain in this chapter, biases can be countered, and even dramatically reduced. We will show how new messages in games can alter or counteract biases.

**Game Play.** As the examples earlier on in this paper illustrate, games can transmit and perpetuate implicit biases through their depiction of (and/or their non-depiction of) social identity categories like gender, race, and sexual orientation. But biases may not be so overt in game play. Implicit bias can also affect the types/genres of games that we prefer (or think we prefer) or believe we’d be adept at. Young girls or players of color may shy away from heavy strategy games or games that involve science or math components due to unconscious societal stereotypes questioning their ability to excel in those domains. Implicit bias can profoundly affect the verbal and nonverbal behaviors of players toward each other, particularly behaviors that often go left unchecked or remain outside of our conscious control, such as nonverbal cues (facial expressions, tone of voice, word choice, etc.) and shape differential patterns of expectation or response toward fellow players.

**Game Design.** Implicit bias can affect the audiences designers envision when conceptualizing and crafting a game. Such bias can affect designers’ playtesting process (e.g., preventing designers from seeking out diverse samples of testers or altering their perceptions of players’ behaviors and responses based on their expectations). These unconsciously flawed practices can trigger confirmation bias and self-fulfilling prophecy, by which designers’ expectations of who would prefer to play a game and/or how different demographic groups would perceive or approach a game influence the way they see or interpret players’ responses—thus unwittingly providing false evidence of their own expectations.
**Game Culture.** The culture around games and gameplay has its own biases. Games are still largely seen as a male domain, even though adult women constitute about half of all digital game players (ESA, 2014). The percentage of women in the US game industry has doubled since 2009 but still only stands at 22%, and people of color continue to be marginalized (IDGA, 2014). Over half of those working in games report that sexism and violence in games, as well as sexism in the game development workforce, leading to negative perception of games and the game industry overall (IDGA, 2014). In the larger culture, recent articles show that even unexpected audiences such as elderly women are entering the game space in record high numbers (Newman, 2014). Yet the environment for women, racial minorities, and LGBT players in online multiplayer games continues to be challenging, and such communities have not yet evolved to invite and embrace the growing diversity of audiences.

**THEORETICAL FRAMEWORK**

Unconscious biases originate from a number of sources (Dasgupta, 2013). For example, repeated exposure to stereotypical depictions or prejudicial attitudes toward individuals of another race in one’s personal life (e.g., growing up constantly hearing racist views or jokes at home or school) or in the broader culture or media (e.g., seeing caricatured representations of a group’s members in films or TV shows).

Implicit bias is a likely culprit that explains imbalanced representations in both game design and game play. It also explains why those who try to improve on gaming diversity often fail. In addition to in-game representation and questions of agency and empowerment for minority and female characters, widespread stereotypes of “gamer” and “game designer” in the broader culture have likely instilled a stronger association between both categories and “heterosexual, white male” than with categories such as “female,” “Latino/a,” “African American,” “gay,” or “transgender.” Implicit associations have been shown to be automatically triggered by exposure to a target and, as a result, often affect individuals’ judgments and behaviors outside of their awareness (e.g., Bargh & Chartrand, 1999). So, for example, implicit bias regarding the proficiency of girls and women in gaming and game design could lead peers, parents, and teachers to form gender-stereotypical expectations, which they likely transmit subtly and unwittingly to girls (e.g., in the level of attention or encouragement they provide to girls versus boys who express an interest in games and game design). This may happen even if parents, teachers, and peers do not consciously endorse these stereotypes (e.g., Jacobs & Eccles, 1992). Even subtle reminders of beliefs about their identity group can trigger *stereotype threat*, by which
the mere activation of a stereotyped identity can elicit performance-debilitating anxiety and preoccupation about confirming a stereotype about their group (e.g., Logel et al., 2009). In other words, implicit bias has a deleterious effect on attitudes held both toward and by members of a group for which a negative stereotype is known or has been propagated.

Implicit biases also manifest in everyday microaggressions: the “subtle, stunning, often automatic, and non-verbal exchanges” (Pierce, Carew, Pierce-Gonzalez, & Willis, 1978, p. 66) that reflect and transmit one’s unconscious associations toward a group. Such responses are often overlooked or glossed over for their unintentional and seemingly innocuous nature, but research has shown that “micro-inequalities” have a profound detrimental effect on the motivation and performance of members of targeted groups (Sue, 2010). For example, a number of learning theorists and researchers have described how microaggressions create and perpetuate the “chilly climate” faced by women and other underrepresented groups in game design and game play contexts, just as they have in STEM learning contexts (Wasburn & Miller, 2005).

The significance of understanding implicit bias—and combating it—is eloquently expressed by social psychologist Laurie Rudman, who stated that “for a deep and lasting equality to evolve, implicit biases must be acknowledged and challenged; to do otherwise is to allow them to haunt our minds, our homes, and our society into the next millennium” (Rudman, 2004, p. 139).

PSYCHOLOGICAL MECHANISMS FOR REDUCING AND REVERSING IMPLICIT BIASES

Fortunately, though implicit biases are well-ingrained by powerful forces of socialization and inequitable media representations, they are by no means impermeable to change. Attacking implicit bias requires reversing or retraining the unconscious negative or stereotypical associations individuals hold—that is, replacing or overriding existing automatic associations with new associations. Research conducted over the past several decades has revealed a number of psychological mechanisms that have proven to be quite effective at reducing implicit biases and reversing implicit associations, at least temporarily. Empirical work in this space has shown that the most successful of these techniques include the following:

1. **Counter-stereotypical Examples.** Exposure to numerous counter-stereotypical role models (e.g., having girls read about successful female scientists and engineers) helps counteract the abundance of stereotypical exemplars who
occupy existing mental representations of a group or category (e.g., Dasgupta & Asgari, 2004).

2. **Counter-Stereotype Training.** Counter-stereotypic training (e.g., having individuals press a button labeled “No” every time a picture of a group member paired with a stereotypical trait—such as an image of a female paired with the word “weak”—appeared on a computer screen) reduced the automatic activation of the stereotype (Stout et al. 2011; Forbes & Schmader, 2010).

3. **Perspective-Taking.** Guided perspective-taking activities and exercises that encourage simulating the experience of “outgroup” members (e.g., imagining a “day in the life” of a member of another race) effectively reduce biases and stereotypes, in part by forging greater psychological overlap between one’s representation of “self” and “other” (Kaufman & Libby, 2012; Todd, Bodenhausen, Richeson & Galinsky, 2011).

4. **Social Norms.** The fostering of egalitarian norms and motivations, when activated, can counteract or prevent the automatic activation of implicit negative stereotypes (Moskowitz, Gollwitzer, Wasel & Schaal, 1999).

The success of these interventions rests on the assumption that they can change existing mental representations and enable individuals to “relearn” a new, counter-stereotypical association. As summarized by social psychologist Patricia Devine: “inhibiting stereotype-congruent or prejudice-like responses and intentionally replacing them with non-prejudiced response can be likened to the breaking of a bad habit” (Devine 1989, p. 15). Further, sufficient “intention, attention, and time” are required to strengthen new, non-stereotypical responses (Devine, 1989, p. 16). Because implicit biases tend to be firmly ingrained and entrenched, “de-biasing” devices must constantly battle the impact of sociocultural factors (such as the transmission of stereotypical messages in popular media) that, in essence, “re-bias” through their perpetuation of negative stereotypes (Kang et al., 2012, p. 1170).

**AN “EMBEDDED DESIGN” APPROACH FOR ACTIVATING PSYCHOLOGICAL MECHANISMS IN GAMES**

The great news is that there are clear mechanisms that designers can glean from psychology to counteract biases through their work. But how can such direct, explicit, largely instructional mechanisms be most effectively integrated within—or activated by—a game? What our game design and research experience has taught us first-hand is that implementing these mechanisms often requires great nuance and finesse. Just as an “aggressive investment strategy” can lead to great wealth or bankruptcy, employing psychological mechanisms to counteract the effects of bias is
not a foolproof venture and must be employed with proven, evidence-based approaches. Designers of “games for change” largely do not consult the psychological literature for these techniques—and those who do often rely on mostly a superficial skimming of the surface or implement techniques without full consideration or empirical verification of their effectiveness and impact. Implicit biases are challenging, because even those designers who wish to alter this status quo by providing new models can overlook how their own work may inadvertently reinforce or perpetuate biases.

Using a psychological approach can benefit designers as key players in creating both games and the culture around them, but they must be used with care. A major caveat to consider when implementing these techniques is that their effectiveness can be limited by executions that are perceived to be too overt or forceful in their intention to shift mindsets, attitudes, or behaviors. The theory of reactance—one of social psychology’s oldest and most well-established formulations—stipulates that any persuasive attempt that is subjectively experienced as coercive triggers an aversive state of arousal to the belief that one’s freedom to think or act freely is being threatened (Brehm & Brehm, 2013). This aversive state motivates individuals to resist a persuasive message as a means of restoring that freedom—even if they would normally be in agreement with that message. Thus, if players are too cognizant of a designer’s intention to promote diversity or to de-bias players’ implicit associations, then they are likely to be less receptive to such messages (and, indeed, less likely to engage with the toy or game in the first place).

So, for example, when learning that exposing individuals to counterstereotypical role models is an effective de-biasing technique, the temptation might be to throw as many role models at your target audience as possible. When deciding on the cast of characters in one’s game, a designer may err on the side of putting white males in the minority to counterbalance the norms in media representations. However, doing so would likely only put a spotlight on the game’s intention to present positive exemplars from underrepresented groups and, at best, reduce players’ engagement with what they perceive to be an agenda-driven game and, at worst, trigger reactance and make players less amenable to accept or to be persuaded by the game’s de-biasing content.

In contrast to such explicit, overt approaches to presenting persuasive content in games, we have proposed an “Embedded Design” approach to game design that relies on a much higher level of subtlety—and, in some cases, subterfuge—to create games that have a greater potential to be impactful without sacrificing player enjoyment or
immersion (Kaufman & Flanagan, 2015; Kaufman, Flanagan, & Seidman, 2015). In this way, potentially sensitive, controversial, or counterattitudinal ideas or themes in games can be crafted in a way that is less overt and less obviously didactic or “message-driven.” In contrast to explicit instructional strategies for promoting prosocial attitude and behavior change, the Embedded Design approach offers strategies for designing persuasive and impactful game experiences that are not overly explicit in their goal to change players’ attitudes or beliefs. The Embedded Design approach rests on the foundational premise that the persuasive potential of games can be significantly augmented by interweaving an intended message or theme within a game’s content, mechanics, or context of play—rather than making that message or theme an overt and explicit focal point. We believe that the techniques can work across various types of media forms, as well. Our design and research team’s efforts to create and study games for social impact have allowed us to propose and test a set of best practices and concrete design techniques (with empirically demonstrated effectiveness) for increasing the persuasive impact of our games.

First, the “best practices” that are derived from the Embedded Design approach include the following basic tenets backed by empirical research:

- **Combine on-topic and off-topic content.** Interweave playful content with serious, educational, or ‘message-related’ content. When making a game promoting diversity and egalitarianism, designers may be (and often are) inclined to “overstuff” their game with the key pro-diversity message. We found that too much of a good thing—in this case, too much “on-message” content that game creators wish to communicate—can be damaging. To illustrate, in a set of games we designed to combat social stereotypes and biases, we have shown that “overstuffed” versions of the game are not only ineffectual for reducing bias, but players enjoy them less than versions that utilize the “intermixing” method (Kaufman & Flanagan, 2015). Moreover, one of the key reasons for this pattern of results is that game versions that “overload” their content with on-topic materials arouse player’ ire and frustration, rendering them less receptive to the games’ pro-diversity aims and themes (Kaufman & Flanagan, 2015; Kaufman & Flanagan, under review B)

- **Avoid imbalanced representations.** Avoiding an imbalance or overrepresentation of counter-stereotypical examples is another technique we have uncovered through our research. Too much of a good thing can seem unattainable. Female players for example, who were presented with a majority-female science team in a strategy game, versus players who were given a balanced gender team, actually were less immersed and invested in the game, reducing the impact of any potentially empowering messages they might discover through play (Kaufman &
Flanagan, under review A). Over-representations of female role models not only made the success of women in science seem more unrealistic or unachievable, they also likely made the intention of the game too obvious or overt (and, by putting gender in more apparent focus, may have triggered stereotype threat or primed stereotypes related to gender in STEM contexts).

- **Design for repetition.** There’s a need for repetition over time to reinforce any potential psychological effects in games (or prevent effects from evaporating), so making sure gameplay is compelling is a key goal. In fact, repetition is a particularly key ingredient to psychological change when it comes to reversing well-ingrained implicit attitudes or associations (Devine, 1989). Most established interventions for reducing implicit bias (e.g., completing a version of the Implicit Association Test that has a respondent repeatedly categorize counterstereotypical stimulus pairs, such as women’s names and STEM fields, using the same key) rely on repetition to build new mental connections and/or weaken existing, stereotypical ones.

- **Be less obvious.** In general, we have found that simply making the purpose or aim of a “device” less explicit or obvious increases player’s pleasure, interest in repeat play, and openness to change (Kaufman & Flanagan, 2015). In general, this is one of the great benefits of well-designed games: players are motivated to play again and again— and one of the pitfalls of well-intentioned but poorly designed games is that they prioritize the message over the experience of the player.

To provide a more concrete depiction of these best practices of the Embedded Design approach as we have enacted and researched them, the following table 15.1 presents a set of easily implementable design techniques (and illustrative examples of their effectiveness) from our own game creations and accompanying empirical studies investigating the games’ efficacy and impact.
### Embedded Design Technique

<table>
<thead>
<tr>
<th>Technique</th>
<th>Description of Technique</th>
<th>Example of Implementation / Empirical Evidence of Techniques Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermixing</td>
<td>Combining “on-topic” and “off-topic” game content in order to make the focal message or theme less obvious and more accessible</td>
<td><em>Awkward Moment</em>, a party card game for pre- and early-teens that has players reacting to social scenarios on Moment Cards by submitting Reaction Cards to a Decider. The game intermixes “on-topic” Moment Cards pertaining to social biases (e.g., “While stopping at the mall, you see a store is selling t-shirts for girls that say Math is hard!”) with a majority of “off-topic” Cards unrelated to bias (e.g., “While bending over to tie your shoes in the school hallway, you feel your pants split wide open.”) Our research has shown that the “intermixed” version of the game increases players’ association between women and science threefold over a “neutral” version of the game (where all Moment cards present “off-topic” content. An intermixed version of the game, but not an “overloaded” version (with a majority of bias-related Moment Cards, increases players’ perspective-taking abilities. (Kaufman &amp; Flanagan, 2015).</td>
</tr>
<tr>
<td>Obfuscating</td>
<td>Using game genres that direct players’ attention or expectations away from the game’s true aims.</td>
<td>Framing <em>Awkward Moment</em> groups as a game that deals with “social stereotypes,” compared to a neutral frame describing the game as pertaining to “social situations,” led youth participants to evaluate the game as less fun and immersive and, furthermore, prevented the game from strengthening players’ association between women and science (Kaufman &amp; Flanagan, 2015). In other words, obfuscating the game’s persuasive aims (by using a description that diverted attention away from the game’s intent) was key to ensuring the game’s impact.</td>
</tr>
<tr>
<td>Distancing</td>
<td>Employing fiction or metaphor to increase the psychological gap between players’ identities and beliefs and the game’s characters and persuasive content.</td>
<td>We have compared two versions of our public health board game <em>POX</em>: <em>Save the People</em>, which is intended to promote positive attitudes towards vaccination and concern about those infected with disease: one version (<em>POX</em>) utilizing a more realistic narrative about disease spread, and another (<em>ZOMBIEPOX</em>) sharing the same game mechanics but presenting them in the context of a “zombie plague” narrative. Results showed that players of the zombie-themed version of the game reported higher levels of empathy towards individuals with infectious diseases, and that this outcome was mediated by players’ level of enjoyment and psychological immersion during play (Kaufman &amp; Flanagan, under review D).</td>
</tr>
<tr>
<td>Delayed Reversal</td>
<td>Withholding potentially threatening or alienating information until players have</td>
<td>In research on <em>Monarch</em>, a strategy board game that positions players as sibling princesses, with male youth participants, we found that slightly delaying the gender of players’ characters significantly increased players’...</td>
</tr>
</tbody>
</table>

*Table 15.1. Embedded Design Approach*

As these examples demonstrate, the Embedded Design approach promises to revolutionize the ways that game developers tackle serious content issues and make...
more effective and more enjoyable games that are remarkably more transporting and impactful compared to games in which the persuasive message or material is presented more overtly or directly.

WHAT CAN DESIGNERS TAKE AWAY TO DIVERSIFY GAMES?

It is finally time for a change in gaming, but the change has to carefully address implicit biases in order to have impact. Designing for diversity means expanding the audience who plays and who is excited about playing—and creating—games. It means improving the social landscape in the development and design of games. Finally, it means making a more equitable and just space for games in popular culture. In this chapter we’ve outlined some concrete techniques to design to address or combat biases. These are now tangible, evidence-based approaches that designers can create games with diversity in mind from a psychological perspective. We wish to emphasize that these can be techniques used by any game designer in any circumstance—they are just as suitable when designing characters for a learning game about math as they are when nudging the content of World of Warcraft to be about science to change conversation to be about women in science). Implicit biases yield stereotypical or non-diverse representations, reinforce categorical dichotomies in game design and marketing, and cater to an unnecessarily limited demographic. The de-biasing devices offered here can – and should – inform both designers and their designs. At the same time, the importance of empirical evidence to test the impact/efficacy of game and the chosen mechanisms and techniques of implementation should not be overlooked. Even the best-intentioned and most well-designed games can unwittingly perpetuate biases—in their representation of characters, their intended audiences, and their positioning in the broader cultural context. In light of this fact, the Embedded Design approach offers key strategies for tackling social issues and including persuasive content in a game in ways that circumvent players’ psychological defenses, trigger a more receptive mindset for internalizing the game’s intended message, and do so without sacrificing players’ enjoyment or the game’s replayability. Indeed, the power of this approach is that it offers design solutions that have the potential to be equally effective for both individuals who may already endorse a particular stance as well as those who may initially be opposed or indifferent to it. These strategies can enable games that address social issues, such as issues related to diversity and bias, to have a much broader and more long-lasting impact.
References


International Game Developer’s Association (IGDA). (2014). Developer Satisfaction


Yang, G. S., Gibson, B., Lueke, A. K., Huesmann, L. R., & Bushman, B. J. (2014). Effects