A psychologically “embedded” approach to designing games for prosocial causes

Geoff Kaufman¹, Mary Flanagan²

¹ Human-Computer Interaction Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, United States
² Tiltfactor Laboratory, Department of Film and Media Studies, Dartmouth College, Hanover, NH, United States

Abstract

Prosocial games often utilize a direct, explicit approach to engage players with serious real-life scenarios and present information about key societal issues. This approach, however, may limit a game’s persuasive impact and ability to produce beneficial outcomes, particularly when the apparent aims of the game trigger players’ psychological defenses or reduce players’ potential engagement with – and enjoyment of – the game experience. In contrast, the “Embedded Design” approach that we introduce here offers effective, evidence-based strategies for more stealthily or covertly delivering persuasive content in a game in a fashion that circumvents players’ psychological defenses and triggers a more receptive mindset. This paper provides an in-depth exploration of two key Embedded Design strategies: (1) intermixing: combining “on-topic” and “off-topic” game content in order to make the focal message or theme less obvious and more accessible and (2) obfuscating: using game genres or framing devices that direct players’ attention or expectations away from the game’s true aims. To illustrate the implementation and effectiveness of these strategies, we detail the design of two games that utilize a number of these techniques to reduce stereotypes and biases and present the methods and results of a set of empirical studies testing the prosocial impact of these games. In addition, we introduce a number of other Embedded Design strategies that have emerged in our work and discuss the most viable contexts for the use of this design approach.

Keywords: Embedded Design, game design, prosocial games, stereotypes, prejudice, perspective-taking

Introduction

For thousands of years games have reflected profound elements of culture and the struggles inherent in the human condition. Indeed, by leveraging the unique benefits and affordances of games and play to facilitate personal growth and to address pressing social concerns, game designers and researchers have the ability to make a tangible lasting impact on society through their work. Rather than exploring and elucidating the transformative positive effects of games, however, media effects research over the past several decades has devoted considerably more attention to uncovering the potential negative impact of game play, particularly violent game play, on aggression, hostility, and antisocial behaviors. This controversial line of inquiry has, on the whole, yielded inconclusive results (e.g., a recent meta-analysis by Ferguson (in press) found minimal evidence of video game effects on children and adolescent players), leading to a recent push toward less sensationalism and greater objectivity and criticality in scrutinizing the methods, results and interpretations of investigations of media effects (see Ferguson, 2007, 2013; Ferguson & Olson, 2013; Granic, Lobel, & Engels, 2014; Markey, Markey, & French, 2014).

As the violent video game debate has evolved, recent years have also seen the emergence of encouraging empirical work exploring the potential for games to promote and instill prosocial attitudes, behaviors, and emotional responses. For instance, cooperative game play, even in games containing violent content, has been shown to reduce intergroup hostility and prejudice (Adachi, Hodson, Willoughby, & Zanette, 2014)
and promote helping behavior (Ewoldsen et al., 2013), and other work has shown that games can help facilitate emotional regulation (Russonielo, O’Brien, & Parks, 2009) and foster self-actualization (Przybylski, Weinstein, Murayama, Lynch, & Ryan, 2011). While this small but growing body of empirical evidence helps substantiate the long-standing view that games can encourage positive shifts in the hearts and minds of players, a key question that must be addressed is: how can designers effectively craft games that both provide an meaningful immersive experience and produce beneficial effects for individuals and society at large? There remains a critical gap in knowledge concerning what strategies for design and what dynamics of game play might make for the most effective persuasive games (Bogost, 2007), set the stage for the most significant attitude and behavior changes to occur through play, and support the intended values of the game initiative (Flanagan & Nissenbaum, 2014).

In surveying the growing list of ‘games for impact’ that have appeared in the last decade, it becomes readily apparent that a majority of them have attempted to tackle a plethora of societal problems head-on, using play to raise consciousness and social awareness. One strategy commonly used in such games is to inform the player about real-life issues, such as facts about energy use (e.g., CEO2 (Allianz Knowledge, 2010) and ElectroCity (Genesis Energy, 2007)). Another strategy is to build empathy to encourage players to understand the circumstances of real-life individuals facing dire circumstances (such as underprivileged women in Half the Sky Movement: The Game (Half the Sky Movement LLC, 2013), or refugees in the case of Darfur Is Dying (mtvU, 2006)). Some games urge players to role play and practice social responses to personal or public health threat scenarios (as in the HIV prevention game, PlayForward: Elm City Stories (Play2Prevent Lab, 2013)) or make decisions in situations involving biases faced by members of stereotype-targeted groups (such as those faced by an African-American graduate student in Fair Play (GLS Studios, 2013)). The common tactic utilized by these games is to rely on an overt, data-driven, fact-specific model, with the underlying assumption that more information (and more explicit information) is better for players – and that being more informed will change player’s beliefs, actions, and emotional responses.

As research in psychology has shown, however, being informed does not necessarily lead to beneficial outcomes. One reason that this direct, informational approach can backfire is the well-established finding that exposure to descriptive social norms regarding the prevalence of undesirable behaviors, such as littering, more salient to individuals sometimes increases their likelihood of committing the same actions (Cialdini, Reno, & Kallgren, 1990). As an illustration of this phenomenon, a field study conducted at Arizona’s Petrified Forest National Park showed that exposure to posted messages highlighting the commonness of theft of petrified wood resulted in higher levels of theft of the wood among visitors (Cialdini et al., 2006). As these and other studies demonstrate, more information about a topic does not necessarily lead to more prosocial behavior. Indeed, it can, under some circumstances, produce perceptions and actions that are detrimental to self and society.

Another potential pitfall of more explicit approaches to shifting prosocial attitudes and behaviors is that their very directness is liable to trigger psychological defenses that can blunt their persuasive potential – or, worse, inspire beliefs and actions counter to the intended outcomes. As a vast body of research stemming from reactance theory has shown, whenever we perceive that our freedom to form our own attitudes or decide for ourselves how to feel about or act on a particular issue is threatened, this triggers an aversive state of arousal that motivates us to defend that freedom – typically by resisting the persuasive attempt (Brehm, 1966; Brehm & Brehm, 2013). Importantly, this defensive motivation persists even in situations in which we agree with the position being advocated (Worchel & Brehm, 1970). In the domain of pro-health interventions, for example, prior work has shown that more obvious, forcefully worded persuasive messages advocating beneficial behaviors such as flossing and moderating alcohol intake were much less effective compared to more subtle, gently worded messages - in large part because they triggered more anger and aversive responses on the part of message recipients (Dillard & Shen, 2005; see also Grandpre, Alvaro, Burgoon, Miller, & Hall, 2003; Quick & Considine, 2008). The mere perception of attitude change pressure from some external source is sufficient to activate reactance (Knowles & Linn, 2004).

In addition to individuals’ likely resistance to overt or obvious persuasive attempts, another obstacle to prosocial interventions is the denial of one’s own need for such interventions in the first place. Research on the bias blind spot, for instance, has revealed that we are prone to believe that others are much more susceptible to a host of cognitive and social biases than we are ourselves (Banaji & Greenwald, 2013; Pronin, Lin, & Ross, 2002). Thus, when devising and implementing interventions that target unconscious processes, such as implicit social or cognitive biases, signaling their intent to intended recipients may, in fact, render them less open to attitude or belief change. Moreover, when the effects of a particular...
prosocial intervention are able to occur at least partially through unconscious mechanisms, making individuals aware of the intervention’s aim can actually attenuate its positive impact. For instance, informing individuals of the intended self-esteem boosting effects of a self-affirmation exercise was shown to diminish its efficacy (Sherman et al., 2009).

In response to these psychological patterns, and aligned with a belief that games can enhance both our individual and collective quality of life, we, in collaboration with our team of designers and psychologists, have crafted and tested the foundations of a theory for effective and playful ways to shift player psychology in a more subtle (and more impactful) fashion. In the present work, we introduce the concept of “Embedded Design,” through which potentially sensitive, controversial, or counterattitudinal ideas or themes in games are 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. Below, we will outline several distinct but overlapping strategies of Embedded Design and illustrate their utility and effectiveness in practice. We detail two focal examples of games (and accompanying research) from our own work, both of which employ several of these strategies to present content intended to reduce social stereotypes and prejudice, and present empirical findings from a set of randomized experimental studies testing the efficacy and impact of these games (and, more broadly, of the Embedded Design approach).

The Embedded Design Approach

The Embedded Design approach to crafting prosocial games is grounded on the premise that the persuasive potential of prosocial 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 (Kaufman, Flanagan, & Seidman, 2015). 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 Embedded Design strategies with empirically demonstrated effectiveness at increasing the impact of our games on players’ prosocial attitudes and perceptions. Here, we focus primarily on two of these strategies: (1) intermixing: combining “on-topic” and “off-topic” game content to make the focal message or theme less obvious and more accessible and (2) obfuscating: using game genres or framing devices that direct players’ attention or expectations away from the game’s true aims.

**Strategy 1: Embedding through intermixing.** One way that potentially sensitive or serious persuasive content can be embedded within a game’s design is a technique we refer to as “intermixing”. The strategy of intermixing involves placing on-message, focal content alongside non-focal content that either distracts attention away from the game’s persuasive goals or renders the persuasive content more approachable or palatable. The balance of on-topic and off-topic elements helps counteract the serious or sensitive nature of the “on-message” elements of the game with elements whose tone is humorous or “off-topic”.

One game that we have designed that exemplifies the intermixing technique is *Awkward Moment*, a party card game that aims to reduce gender stereotypes and biases in science, technology, engineering, and math (STEM) domains. In *Awkward Moment*, players begin by drawing a hand of five “Reaction Cards” (see Figure 1). These cards describe potential responses to the game’s titular “moments,” including actions (e.g., “Scream your head off”,” Write a blog post about it”, “Talk it out”), exclamations (e.g., “Rats!”,” OMG!”,” No way!”), and frames of mind (e.g., “Get serious”,” Relax”,” Channel your inner warrior”). During each round, one player serves as the “Decider” and draws a “Moment Card” that poses a hypothetical situation (e.g., “Somebody hacks your Facebook account and changes your status to ‘Girls are stupid.’”), to which the other players respond by submitting a Reaction Card face-down. The Decider then reads each of the submitted cards aloud and deems the Reaction he/she believes to be the best response to the Moment as the winner.

In line with the “intermixing” approach outlined above, included within the full set of Moment cards, a majority of which describe general stressful or humiliating social occurrences (e.g., “An embarrassing picture of you is being sent around school.”), is a subset of cards that present situations involving bias against girls in STEM (e.g., “While shopping at the mall, you notice a store is selling t-shirts for girls that say, ‘Math is hard.’”) or a lack of gender equity in STEM domains (e.g., “Your school’s math team is 100% boys.”). Moreover, these cards are intermixed with other bias-relevant Moment Cards that address other negative beliefs, behaviors, or evaluations toward a number of social targets, based on dimensions such as ethnicity, body type, and nonconformity to traditional gender norms.
A pair of randomized experimental studies tested the general impact of the game on youth players’ gender stereotyping in STEM and responses to bias (Study 1A) and the effectiveness of the game’s intermixing strategy in increasing players’ levels of perspective-taking (Study 1B).

**Study 1A: Testing the Effectiveness of the Intermixing Strategy in Awkward Moment**

To test the general effects of the game on youth players, we created two versions of the game that participants were randomly assigned to play. One version preserved the game’s intermixing strategy, with five of the twelve Moment cards used in the game depicting gender bias scenarios (the Bias Game condition). For comparison, a second version of the game contained twelve “neutral” Moment cards describing situations unrelated to gender bias (the Neutral Game condition). Following gameplay, participants completed measures assessing their likelihood of associating women and STEM and their level of assertiveness in response to occurrences of bias.

**Method**

**Participants.** A sample of 309 participants (179 females and 130 males between the ages of 12 and 14) were recruited from public middle schools and high schools in New England in the United States. For this study (and all studies to be reported), all materials and procedures were approved by the Committee for the Protection of Human Subjects at Dartmouth College, and the written consent of the participants (and, in the case of youth samples, the written consent of their parents or guardians) was obtained.

**Materials and procedure.** Participants were assigned to either the Bias Game condition, the Neutral Game condition, or to a control condition, in which participants completed our dependent measures (described below) before playing the neutral version of the game. In all conditions, participants played the game in randomly determined mixed-gender groups ranging in size from four to six. To ensure consistency in exposure to the game’s content, the same sequence of Moment Cards was used for all playgroups in each condition. In addition, to maintain a higher level of uniformity between the two game conditions, the seven “neutral” Moment Cards contained in the intermixed deck in the Bias Game condition were also used in the Neutral Game condition deck (in the same position within the sequence of twelve cards in the deck).
Two primary measures of gender bias were included in a paper-and-pencil questionnaire administered to participants immediately following gameplay. Participants were separated from their fellow players to complete the questionnaire. The first measure presented a set of six images of faces (3 female and 3 male) that were identified as characters for a new game. Participants were asked to assign each character a job from a set of six that contained both male-stereotypic options (pilot, soldier, scientist) and female-stereotypic options (nurse, secretary, teacher). In our analyses, the percentage of participants who assigned “scientist” to one of the three female characters, as an indication of a rejection of gender stereotypes in STEM, was of particular interest.

The second measure had participants read a set of five hypothetical situations in which an individual was the target of bias (e.g., witnessing a female classmate’s brother seeing her A+ mark on a math test and saying, “Not bad for a girl!”) and rank, in order of likelihood, a set of five responses (which emerged in prior works by Czopp, Monteith, & Mark, 2006, and Hyers, 2007). The nature of the options ranged from strong and assertive responses that directly negated the act of bias (e.g., explaining that girls are just as good at math) to unassertive responses that passively expressed disapproval (e.g., giving him a “dirty look”) to trivializing responses that implied approval of the biased act (e.g., laughing along). Of particular interest was the percentage of participants in each condition who selected an assertive response to bias as their most likely.

Results and Discussion

The study utilized a nested design, with participants assigned to playgroups with three to five other participants. Thus, for all analyses, the play group (rather than the individual participants) was used as the unit of analysis; that is, the average response to the outcome measures reported by participants in a given playgroup was used for purposes of data analysis. This resulted in a total of 56 distinct groups (20 each in the Bias Game and Neutral Game conditions and 16 in the Control condition).

Results revealed that participants who played the Bias version of Awkward Moment showed a much stronger anti-bias response on both of these measures compared to participants in the other two conditions (see Figures 2 and 3). First, a significantly higher percentage of participants assigned the occupation of “scientist” to a female character in the Bias Game condition (58%) compared to the Neutral Game condition (18%) and the Control condition (25%), χ² (N = 56) = 6.77, p < .04 (see Figure 2). In addition, a significantly higher percentage of participants identified an assertive response as their most likely reaction to occurrences of bias in the Bias Game condition (68%) compared to the Neutral Game and Control conditions (35%), χ² (N = 56) = 5.59, p < .05).\(^1\)

![Figure 2. Percentage of participants in each condition who assigned “scientist” to a female character.](image)
These results confirm that despite the fact that only a minority of the Moment cards in the Bias Game condition directly addressed gender bias in STEM, the game was nonetheless successful at strengthening youth players’ association between women and science and inspiring more assertive responses to multiple forms of social bias. As a more direct test of the intermixing strategy, a follow-up study was conducted to compare the impact of versions of the game that varied their ratio of bias-related to neutral Moment Cards. In line with the intermixing strategy, we predicted that presenting a lower (versus higher) ratio of bias-themed to non-bias-themed Moments would prove much more effective at shifting players’ attitudes and perceptions. Thus, in Study 1B, we directly compared an “Intermixed” version of the game (with a balance of on-topic and off-topic content) with an “Imbalanced” version of the game (with on-topic content outweighing off-topic-content) to determine if the latter would prove to be less impactful. This methodological choice was intended not only to demonstrate the benefits of the Embedded Design approach but also to provide evidence against a plausible alternative explanation for the results of Study 1A: namely, that participants in the Bias Game condition discerned the intention of the game (and of the study) and accordingly adjusted their responses to the outcome measures employed in the study. Showing that an “Imbalanced” version of the game (which would be more explicit in its focus and intention) produced less movement on post-game measures would provide evidence against this possible interpretation.

**Study 1B: Comparing Intermixed and Imbalanced Versions of Awkward Moment**

This study compared the impact of two versions of Awkward Moment – an “intermixed” version (with five of twelve Moment Cards related to bias, as in Study 1A) and an “imbalanced” version (with nine of twelve Moment Cards related to bias) – on players’ post-game levels of perspective-taking (i.e., players’ inclination and ability to adopt another person’s point of view).

**Method**

**Participants.** A sample of 232 youth (124 females and 108 males between the ages of 11 and 14) from public middle schools and high schools in New England participated in this study.

**Materials and procedure.** Participants were randomly assigned to one of three game conditions: (1) the Intermixed Game condition; (2) the Imbalanced Game condition; or (3) a control condition, in which participants played Apples to Apples, a game that is mechanically similar to Awkward Moment, in that it involves the private submission and judgment of cards among a group of players, but contains no content that is overtly bias-related. In all conditions, participants played the game in mixed-gender groups ranging in size from four to six.

Two primary measures of perspective-taking were administered in paper-and-pencil questionnaires immediately after gameplay. The first measure asked participants, who completed the measures separately and out of sight of one another, to complete a number of drawing and writing tasks, with the
focal task of drawing a capital letter “E” on their foreheads. The extent to which participants drew the “E” from an “other-oriented” perspective (i.e., legible to another person viewing the drawn letter) represents an established measure of individuals’ perspective-taking ability (e.g., Galinsky, Magee, Inesi, & Gruenfeld, 2006; Hass, 1984). In the second measure, participants were asked to read a scenario in which a target character’s friend recommends a restaurant at which the character has a terrible dining experience; afterward, the character sends an email to the friend who recommended the restaurant with the text, “About the restaurant, it was awesome. Just awesome.” Participants were then asked to judge how sarcastic the character’s friend (who was unaware of the particulars of the dining experience) would perceive the email to be. Successful perspective-taking is indicated by lower ratings of perceived sarcasm (Galinsky et al., 2006).

Results and Discussion

The average responses on the outcome measures reported by participants in the 48 distinct playgroups for both measures showed that the “intermixed” version of Awkward Moment (in which the ratio of bias-related to neutral Moment Cards reflects the ratio in the full deck) elicited the highest levels of perspective-taking, compared to the levels exhibited by participants who played either the “imbalanced” version of Awkward Moment or Apples to Apples (see Figures 4 and 5). First, the percentage of participants who drew their “E” from an other-oriented perspective was significantly higher in the Intermixed Game condition (89%), compared to the percentages in the Imbalanced Game condition (63%), and the Apples to Apples control condition (29%), χ² (N = 48) = 5.38, p < .05 (see Figure 4). A similar pattern of response emerged for the sincerity measure, with participants in the Intermixed Game condition reporting the highest level of perceived sincerity (M = 5.01, SD = 1.09) compared to participants in the Imbalanced Game condition (M = 3.30, SD = 1.16) and Apples to Apples control condition (M = 2.29, SD = .99), F(2, 46) = 3.15, p < .05.
These findings demonstrate that the Intermixed version of the game, despite featuring far less focal content related to social biases, was clearly and substantially more effective at strengthening players’ perspective-taking than the comparatively more explicit Imbalanced version of the game. In other words, in this case, less is more. This pattern of results affirms the efficacy of the intermixing strategy as utilized in the game and reinforces our general belief that more information in a game is not necessarily better for producing positive or prosocial effects through gameplay. In more recent work, we have further investigated the distinct effects of “intermixed” versus “imbalanced” ratios of bias-related to neutral content in a version of Awkward Moment for adults called Awkward Moment at Work. This research has shown that not only is the “intermixed” version of the game more effective at increasing players’ motivation to combat stereotypes and discrimination but, further, that the “imbalanced” version of the game triggered negative emotional responses in players, indicative of higher levels of reactance to the game’s more apparent persuasive aims (Kaufman & Flanagan, under review).

**Strategy 2: Embedding through obfuscating.** A second Embedded Design strategy that has emerged in our work is “obfuscating”: obscuring or disguising the true intent of a game and using more covert means of presenting or modeling a game’s underlying persuasive themes or content. One primary means of obfuscation that we have employed is to utilize a game genre whose perceived goals or anticipated intentions do not include the aim to change attitudes or beliefs. To this end, the choice to use the party game genre for Awkward Moment was a wholly intentional one. The success of the game was likely in large part due to players’ anticipation (and experience) of a fun, engaging play experience with high levels of levity and humor (achieved in part through the game’s “intermixing” of Moments and Reactions that range from silly to serious in tone). The game’s balance of social and humorous elements, thus, creates a “safe” space for players to perceive and react to the game’s on-message content with greater comfort and candor.

In developing a second game aimed at reducing players’ social stereotypes and biases, **Buffalo: The Name Dropping Game**, our design team took an even more extreme approach to using the party game genre to obfuscate the game’s underlying goals: namely, creating a game whose purpose was even more “hidden” to players. In **Buffalo**, players simultaneously flip cards from two decks. The “orange deck” contains cards that list adjectives, including ones based on age (e.g., young, old), race (e.g., Hispanic, Caucasian, multiracial), physicality (e.g., tall, unattractive, blond), personality (e.g., strong, corrupt, funny), and ideology (e.g., spiritual, eco-friendly). The “blue deck” contains cards listing nouns, including professions (e.g., scientist, supermodel, talk show host), roles/types (e.g., grandparent, superhero, homemaker), and ideological or interest group affiliation (e.g., environmentalist, animal expert, feminist). Using the noun-adjective combination formed by the cards drawn, players race to collect the cards by identifying a real person or fictional character who matches the revealed pair (see Figure 6).

As play progresses, **Buffalo** activates a wide array of combinations of social categories and attributes, including pairings that are often novel, surprising, or unexpected (e.g., tattooed visionary, South American philosopher, mischievous professor) and, in some cases, directly counterstereotypical (e.g., rugged fashion designer, wheelchair-dependent athlete, saintly merchant). In this way, the game aims to expand players’ mental representations of numerous social categories with a host of new exemplars that play groups offer. At the same time, the intended persuasive impact of **Buffalo** was designed to be largely unrecognized by players, at least on a conscious level (indeed, our extensive playtesting of the game revealed that all but a small minority of players failed to connect the game in any way to stereotypes or prejudice).

![Figure 6. Sample adjective cards and noun cards from Buffalo: The name dropping game.](image-url)
A pair of randomized experiments tested the predicted effects of the game among older adolescent and young adult participants. The first study (Study 2A) investigated the impact of *Buffalo* game play on players’ representations of social identity categories and levels of prejudice following play; a second study (Study 2B) utilized a measure of players’ motivation to recognize and control their own biases as the focal outcome variable.

**Study 2A: Testing the Effectiveness of the Obfuscating Strategy in Buffalo**

**Method**

**Participants.** A sample of 193 young adults (102 females and 91 males between the ages of 18 and 26) from public and private colleges and universities in New England participated in this study.

**Materials and procedure.** Participants were randomly assigned to one of two experimental conditions: a Game condition, in which groups played a full session of *Buffalo* (featuring 25 noun-adjective pairings) before completing a questionnaire containing the outcome measures (described below) or a no-game control condition, in which participants completed the measures prior to game play. The card pairings used in the Game condition were randomly drawn from the game’s full deck of 256 noun and adjective cards; to maintain uniformity in exposure to the game’s content, all groups in the Game condition were given the same pre-stacked set of cards.

Two primary outcome measures were administered to participants in a paper-and-pencil questionnaire immediately following gameplay. The first measure assessed participants’ level of *social identity complexity*: the breadth and inclusiveness of their representation of their self-identified social identity groups (Roccas & Brewer, 2002). This measure, which has been shown to be a reliable predictor of intergroup tolerance (Roccas & Brewer, 2002), asks participants to list the four social identity groups (e.g., gender, race, ethnicity, religion, profession or area of study, hobby group, etc.) that are most important to them. Then, participants are asked to rate how similar the average member of each group is to the average member of each of the other three groups, on a scale from 1 (not at all similar) to 9 (extremely similar). Responses to each of the six resulting comparisons are averaged to form an overall similarity rating, with lower levels of perceived similarity corresponding to higher levels of social identity complexity.

Second, participants completed the Universal Orientation Scale (Phillips & Ziller, 1997), an established and validated measure of “general non-prejudice” that assesses the degree to which respondents see their social world as diverse and inclusive. This scale presents respondents with a set of twenty statements (sample items include: “There is potential for good and evil in all of us” and “I can see myself fitting into many groups”), to which participants rate their agreement on a scale from 1 (strongly disagree) to 9 (strongly agree).

**Results and Discussion**

As in Studies 1A and 1B, the responses of participants within each of the 37 distinct play groups in the study (19 in the Game Condition, 18 in the Control Condition) were averaged, so that the play group was the unit of analysis for the study. First, on the social identity complexity measure, participants’ responses were all reverse-coded, so that higher scores would correspond to higher levels of social identity complexity. Results revealed that that participants in the Game condition reported higher social identity complexity scores ($M = 6.5$, $SD = 1.27$) than did participants in the no-game Control condition ($M = 4.8$, $SD = 1.19$), $F (1, 36) = 4.87$, $p < .04$ (see Figure 7).

A parallel pattern of results emerged for participants responses to the Universal Orientation Scale, with participants in the Game condition, on average, reporting higher scores on this measure ($M = 6.92$, $SD = 1.44$), compared to participants in the Control condition ($M = 6.02$, $SD = 1.31$); this difference obtained marginal statistical significance: $F (1, 36) = 4.87$, $p < .06$ (see Figure 8).
These results suggest that just a single play session was sufficient to trigger a higher degree of change in intergroup perceptions and judgments among participants in the Game condition, compared to the baseline levels reported by participants in the Control condition. As a follow-up, a second study utilized the same basic procedure as Study 2A to test the impact of *Buffalo* game play on younger players’ motivation to avoid being prejudiced.

**Study 2B: Additional Evidence for the Effectiveness of the Obfuscating Strategy in *Buffalo***

**Method**

**Participants.** A sample of 201 high school students (90 females and 111 males between the ages of 15 and 18) from public and private schools in New England participated in this study.

**Materials and procedure.** As in Study 2A, participants were randomly assigned to one of two experimental conditions: (1) a Game condition, in which groups played a full session of *Buffalo* (with each group given the same randomly pre-drawn set of 25 noun-adjective pairings) before completing a questionnaire containing the outcome measures (described below) or (2) a no-game control condition, in which participants completed the measures prior to play. In both conditions, participants played the game in groups ranging in size from four to six. Following game play, participants completed the Internal Motivation to Respond without Prejudice Scale (Plant & Devine, 1998), which presents twelve items assessing respondent’s intrinsic desire to recognize and control their bias against social outgroups (e.g., “I attempt to act in non-prejudiced ways toward people from other social groups because it is personally important to me”), to which respondents rate their agreement on a scale from 1 (strongly disagree) to 9 (strongly agree).
Results and Discussion

The average scores for the items on Internal Motivation to Respond without Prejudice Scale reported by participants within each of the 40 distinct play groups in the study were calculated and used for analysis. Results revealed that participants in the Game condition reported a significantly higher score on this measure ($M = 7.36$, $SD = 1.40$) than did participants in the no-game Control condition ($M = 6.35$, $SD = 1.28$), $F (1, 39) = 5.11, p < .03$ (see Figure 9).

![Figure 9. Average score on the internal motivation to respond without prejudice scale by condition.](image)

Taken together with the results of Study 2A, these findings reveal that Buffalo game play effectively promoted broader and more conclusive perceptions of social groups and raised players’ concern about their own potential biases, compared to the baseline levels reported by Control condition participants in both studies. Nonetheless, although the true purpose of the game was never revealed to participants in any fashion, we cannot entirely rule out the possibility that participants discerned the prejudice-reducing aims of the game and that their responses to the utilized measures represent an artifact of a social desirability bias on the part of participants. Our extensive playtesting of the game during its formative stages provide some level of evidence against this alternative explanation: in one-on-one interactions with both youth and adult testers, the vast majority of players did not connect the game directly to biases and stereotypes; when asked what they thought the game’s purpose was, many identified such goals as testing or expanding one’s memory or knowledge of historical or pop culture figures. However, more systematic evidence against a social desirability interpretation of the results from Studies 2A and 2B would be offered by demonstrating that disclosing the game’s purpose lowers the impact of the game – in other words, that not obfuscating the game’s aims actually blunts its persuasive impact.

To this end, we have started to investigate the impact of revealing (versus obfuscating) the bias-related content or aims of both Awkward Moment and Buffalo by varying the description of the game provided to participants before play. Specifically, in a pair of pilot studies, we investigated the impact of framing Awkward Moment and Buffalo in a manner that explicitly identified their relevance to issues of bias (specifically, labeling Awkward Moment as a game pertaining to “awkward social stereotypes” and Buffalo as a game pertaining to “pop culture stereotypes”) or described the game’s content in a neutral fashion (i.e., Awkward Moment as a game pertaining to “awkward social situations” and Buffalo as a game pertaining to “pop culture knowledge”). In a pair of pilot studies, we randomly assigned groups of participants (ages 12-14) to play one of the two games under one of these two framing conditions. Our hypothesis, in line with the basic premise of the Embedded Design approach, was that the “bias frame” would reduce the efficacy of the games, in part by activating the assumption that a game about stereotypes would be less enjoyable or by arousing players’ psychological reactance about the game’s persuasive intent.

The findings from this pilot work supported this hypothesis. In a study involving Awkward Moment (with 20 groups of youth participants), play groups who randomly received the neutral frame for the game prior to play, compared to groups who received the bias frame, rated the game as more fun ($M_{\text{neutral}} = 6.58$, $M_{\text{bias}} = 5.60$, $p < .04$) and more interesting ($M_{\text{neutral}} = 5.57$, $M_{\text{bias}} = 5.10$, $p < .10$), reported having their mind wander less during play ($M_{\text{neutral}} = 3.08$, $M_{\text{bias}} = 4.60$, $p < .06$), and showed a stronger association between women and science on the post-game job assignment measure described earlier for Study 1A (57% of Neutral Frame participants assigned “scientist” to a female character compared to 30% in the
Bias Frame condition, $p < .09$). A parallel study involving *Buffalo* (with a separate sample of 26 groups of youth participants) revealed that groups who randomly received the neutral game frame reported higher scores on the Universal Orientation Scale after playing the game ($M_{\text{neutral}} = 6.05$, $M_{\text{bias}} = 5.11$, $p < .06$). In both studies, these differences emerged despite the fact that the only difference between the game conditions was the one-phrase frame used to describe the game. These results confirm that both games’ impact rests on players not perceiving them to be overtly didactic or “lesson-heavy.”

**Other embedded design strategies.** Along with the *intermixing* and *obfuscating* techniques, a third Embedded Design strategy that we have explored in our work is the use of psychological *distancing* (Trope & Liberman, 2010) to offer a safe space or buffer between players and the serious or uncomfortable themes or topics explored or modeled by a game. For example, many of our games utilize the distancing mechanism of *fictionalizing* real-life issues and themes - that is, embedding those elements within more abstract or fantastical modes of representation. To cite one illustrative example, our team has designed and studied two versions of our public health board game *POX: Save the People* (Tiltfactor Laboratory, 2011), which is intended to promote positive attitudes toward vaccination and concern about those infected with disease: one version (*POX*) utilizing a more realistic narrative about disease spread, and another (*ZOMBIEPOX*) sharing the same game mechanics but presenting them in the context of a “zombie plague” narrative. A randomized experiment comparing the impact of the two games (see Kaufman & Flanagan, 2013) showed that players of the zombie-themed version of the game reported higher levels of empathy toward individuals with infectious diseases, and that this outcome was mediated by players’ level of enjoyment and psychological transportation (i.e., absorption; Green & Brock, 2000).

In addition, we have begun to conceptualize and test a number of new manifestations of Embedded Design. These include the use of: (1) *embodied cognition* (Wilson, 2002) to trigger emotional or cognitive shifts unconsciously or surreptitiously through the semantic or metaphorical link between bodily experience and abstract concepts; (2) *priming* (Doyen, Klein, Simons, & Cleeremans, 2014), through which strategically chosen or positioned game elements (e.g., graphics or text) activate particular mindsets, motivations, or concepts in a surreptitious manner; (3) *strategic inversion*, the deployment of irony or satire to represent desired themes or topics in subversive ways (e.g., using parodies of social issues as a means of persuasion: see Nabi et al., 2007); and (4) *delayed reveal*, the choice to withheld potentially threatening or alienating information (e.g., the outgroup membership of a protagonist) until players have formed a psychological connection with the world of a game and its characters (see Kaufman & Libby, 2012 for an illustration of this approach with characters in fictional narratives).

**General Discussion**

The present work introduced the Embedded Design approach to the creation of effective prosocial games and presented in detail the design and empirical research for two games, *Awkward Moment* and *Buffalo*, that exemplify key Embedded Design strategies, such as *intermixing* on-message and off-message content, *obfuscating* their true persuasive aims, and *distancing* players from real-life situations and their own identities, preconceptions, and beliefs. The results of our empirical studies involving both games attest to the efficacy of the games for decreasing players’ social biases and promoting more egalitarian, diversity-embracing mindsets among players. Moreover, we have shown that versions of the game that utilize Embedded Design techniques to a greater extent are significantly more effective at changing players’ hearts and minds than versions that make the games’ persuasive goals more explicit or transparent.

One unique strength of games (digital or non-digital games alike) as prosocial interventions is that they allow a high degree of flexibility in the manner in which they can present information or model individual-level processes or broader, complex systems. As the results of the present research demonstrate, the embedding of a game’s persuasive content or intentions is a powerful design strategy that effectively circumvents players’ psychological defenses and reduces the likelihood of resistance to the game’s message. Prosocial games that take too explicit or direct an approach in their handling of serious or sensitive topics likely have inherently less capacity to persuade players because they afford too low a level of concealment or distance to explore those topics in a non-threatening fashion. Indeed, these potential pitfalls of traditional approaches have led researchers increasingly to endorse the use of entertainment media as alternative avenues for persuasion (Bandura, 2002; Slater & Rouner, 2002).

**When is embedding a more (or less) viable design strategy?** One of the hallmarks of the Embedded Design approach is its ability to make prosocial games more engaging for a wider range of players,
including both those who are sympathetic or in agreement with a particular aim or goal as well as those who are not. Thus, strategies derived from the Embedded Design approach will be particularly effective for topics that are likely to be controversial, counterattitudinal, or in some way threatening to players. Secondly, as discussed earlier, some psychological mechanisms that promote attitude or behavior change operate unconsciously and can be disrupted or short-circuited by conscious deliberation or reflection (e.g., Sherman et al., 2009; Wilson & Brekke, 1994). Whenever a process is more effective when it remain obscured or concealed, embedding is likely to be particularly effective.

There are ways to use the strategy incorrectly, of course. If creators simply seek to integrate “fun” elements within an already problematic game with this technique, they may not be successful. The same goes for those who have a “fun” game but wish it to be more pro-social. There are no statistics yet for the use of these strategies among other designers of prosocial games, but there are some promising examples emerging. For example, the digital app Dumb Ways to Die (Metro Trains Melbourne, 2013) could be read to use the Embedded Design approach: the game intermixes a small number of “mini-games” involving train safety (e.g., stepping too close to a track of an oncoming train) among many other mini-games that depict perilous, potentially fatal scenarios unrelated to train safety (e.g., swimming with piranhas, inserting a metal fork into a plugged-in toaster, or poking a grizzly bear with a stick).

On the other hand, the Embedded Design approach is less suitable for prosocial games that aim to provide factual information or model specific, readily transferable sequences of action for players to enact outside of the game. For example, if the sole purpose of a game is to provide facts or statistics about a particular real-life issue or event, or to teach a specific behavior, a more direct, informational approach may be more suitable (though, as discussed above, these types of interventions run the risk of inadvertently activating descriptive norms for a prevalent social problem). Embedding may also be less viable (and less necessary) in domains that are not particularly sensitive or with intended audiences who are assumed to be inherently amenable to a game’s persuasive intent.

In the end, Embedded Design approaches and more explicit approaches to crafting games for prosocial impact need not be viewed as opposing alternatives, but rather as potentially complementary and mutually reinforcing. In particular, as our own research has shown, games utilizing Embedded Design strategies may be especially effective at creating more open, expansive, and empathetic mindsets in players, one that makes them more likely to consider or accept the information contained in more direct or more obviously didactic interventions. Games that embed their prosocial aims and intent can thus be potentially powerful precursors to explicit games that share the same persuasive goals. Moreover, the two approaches are by no means mutually exclusive. As Figure 10 illustrates, the strategies suggested by the Embedded Design approach can be implemented within games that utilize a more direct, fact-based design format (e.g., trivia games presenting statistics or objective facts about a social issue or training modules that describe real-life occurrences and model appropriate responses to them) to deliver persuasive content to players in a less overt or obvious fashion (i.e., in a fashion less likely to trigger psychological defenses or mindsets that reduce players’ receptivity of the game’s content or enjoyment of the play experience).

**Conclusion**

Among designers, researchers, and players alike, there exists a growing belief in the capacity of games to address the struggles of the human condition and solve pressing societal issues. For games to be most effective, however, they should take advantage of their unique affordances and flexibilities of representation that give them distinct advantages over traditional, information-based interventions. With repeated evidence demonstrating that more information about a topic does not necessarily lead to more prosocial beliefs and behaviors, here we have presented the foundations of a theory for more subtle means of designing games to shift players’ psychological responses. The techniques that emerge from the Embedded Design approach represent concrete, easily implementable design practices with mounting empirical evidence for their efficacy in stimulating transformative prosocial effects on players.
Figure 10. Graphical depiction of the potential implementations of embedded design strategies within games utilizing explicit, information-based design formats for prosocial content delivery.

Notes

1. There were no significant between-condition differences in the percentage of participants who identified a trivializing response as their most likely ($p > .90$), with only a small percentage of participants selecting those response options in the Bias Game condition (8%), Neutral Game condition (12%), and Control condition (13%).
2. The game also has the potential of revealing noun-adjective pairings that are more consistent with social stereotypes (e.g., socially awkward geek, greedy CEO, Type A lawyer); however, in devising the content of the game, we attempted to include mostly adjectives that would not directly pertain to cultural stereotypes. To estimate the proportion of stereotypical and counterstereotypical card pairings that would typically emerge in game play, three independent coders classified 1000 randomly generated noun-adjective combinations from the full Buffalo deck (the coders exhibited 84% agreement, with all disagreements in classifications resolved by the authors). This analysis revealed that a small proportion of pairings were deemed to be stereotypical (20%) and a larger and nearly equal proportion of pairings were deemed to be counter-stereotypical (39%) or completely unrelated to known stereotypes (41%).

Acknowledgement

This research was supported by grants from the National Science Foundation (#HRD-1137483. #DRL-1420036). We gratefully acknowledge the invaluable contributions of the staff and students at Tiltfactor in all stages of the design and research of the games presented here, as well as the participation of our partnering schools in New Hampshire and Vermont, without whom our research would not be possible.

References


**Correspondence to:**
Geoff Kaufman  
Carnegie Mellon University  
Human-Computer Interaction Institute  
5000 Forbes Ave  
Pittsburgh, PA 15213  
United States  
Email: gfk(at)andrew.cmu.edu

**About authors**

**Geoff Kaufman**, Ph.D., is an assistant professor at the Human-Computer Interaction Institute in the School of Computer Science at Carnegie Mellon University. A social psychologist by training, his cross-disciplinary research investigates the affective, cognitive, and behavioral effects of fictional narratives and games; the design and implementation of games, technologies, and interaction platforms as interventions for promoting greater self-insight and social consciousness; and the unique dynamics and affordances of computer-mediated communication for facilitating interpersonal synchrony and interconnectedness.

**Mary Flanagan**, Ph.D., is the Sherman Fairchild Distinguished Professor of Digital Humanities at Dartmouth College and directs the Tiltfactor game research laboratory there. She studies design for social impact and creates games and interventions backed by empirical evidence using a unique interdisciplinary approach. Her research interests range from the psychological influences behind behavior and belief change to the use of experimental art techniques to invoke new thinking.