Facial features of non-player creatures can influence moral decisions in video games

YLVA FERSTL, ELENA KOKKINARA and RACHEL MCDONNELL, Trinity College Dublin

With the development of increasingly sophisticated computer graphics, there is a continuous growth of the variety and originality of virtual characters used in movies and games. So far however, their design has mostly been led by the artist’s preferences, not by perceptual studies. In this paper, we explored how effective non-player character design can be used to influence gameplay. In particular, we focused on abstract virtual characters with few facial features. In experiment 1, we sought to find rules for how to use a character’s facial features to elicit the perception of certain personality traits, using prior findings for human face perception as a basis. In experiment 2, we then tested how perceived personality traits of a non-player character could influence a player’s moral decisions in a video game. We found that the appearance of the character interacting with the subject modulated aggressive behavior towards a non-present individual. Our results provide us with a better understanding of the perception of abstract virtual characters, their employment in video games, as well as giving us some insights about the factors underlying aggressive behavior in video games.


General Terms: Human Factors

Additional Key Words and Phrases: Moral dilemmas, personality perception, video games, virtual characters

ACM Reference Format:
DOI: http://dx.doi.org/10.1145/0000000.0000000

1. INTRODUCTION

Video games and animated movies often make use of abstract-looking characters that portray complex personalities or, in video games, that interact with the player to influence the gameplay and experience. However, perceptual studies focussing on the visual design of these characters have been limited and there are no perceptually-guided rules as to how an abstract character can be used to believably portray a certain personality. Understanding the impression that a character makes on the audience is crucial for their effective employment in storytelling and for creating engaging content.
Character design can not only influence the quality of storytelling in a video game, but, in the case of self-avatars, can also change the player’s behavior (Proteus effect) [Yee and Bailenson 2007; Fox et al. 2013]. Players assigned an avatar associated with more aggressive behavior have been shown to display more aggression [Peña et al. 2009], as do players with personalized avatars [Hollingdale and Greitemeyer 2013]. Here, we hypothesized that the influence of visual appearance on a player’s behavior may not be limited to self-avatars, but may extend to other characters that the player is interacting with. We expect that an interaction with a more aggressive-looking character could lead to more aggressive behavior from the player.

In humans, facial features have been shown to heavily and reliably impact the perception of aggressiveness [Carré et al. 2009; Willis and Todorov 2006; Bar et al. 2006]. We present an exploratory experiment investigating if and how such features could also influence perception of more abstract, less human-looking characters [Ferstl et al. 2016] (experiment 1). Our results indicate that different rules apply for the perception of personality from facial features for abstract virtual characters, compared to human faces. Our aim is to provide guidelines for artists who wish to create characters that portray certain personality traits.

Using our data from experiment 1, we picked the faces that best elicited the perception of high and low aggressiveness and trustworthiness (four in total) to test the influence on a player’s behavior in a video game. In experiment 2, we created a short video game in which players were presented with moral dilemmas. Based on previous findings that more aggressive-looking self-avatars trigger players to display more aggressive behaviors [Peña et al. 2009], we hypothesized that a player’s perception of the aggressiveness of another character will influence their willingness to make an aggressive choice.

Moral dilemmas are typically presented with two answer options, a deontologically motivated and a utilitarian answer, with the utilitarian choice representing the more aggressive answer option [Choe and Min 2011; Gao and Tang 2013].

We specifically designed four different moral scenarios that fit the setting of our game, each probing the player to make an ethical decision between two alternatives. We found that perceived personality of the character can significantly influence moral choices within a game.

Our characters and the game assets used for the creation of the environment of our moral dilemma game stem from the EU POPULATE project\(^1\) and create a tribal, fantasy forest atmosphere.

2. RELATED WORK

A large number of studies have shown that the perception of certain personality traits is strongly influenced by specific (human) facial features. Amongst these traits are sociosexuality [Boothroyd et al. 2008], trustworthiness [Stirrat and Perrett 2010], aggression [Carré et al. 2009], and personality in general [Kramer and Ward 2010; Little and Perrett 2007]. Our brains seem to have evolved to process these features quickly and to draw immediate conclusions from them, such as whether the observed person is a threat or is a superior that requires subordinate behavior [Sell et al. 2009]. Even more so, we can recognize these traits not only in human counterparts, but also in non-human primates [Kramer et al. 2011; Kramer and Ward 2012], suggesting that this is a mechanism deeply rooted in evolution, happening automatically and on the subconscious level.

One such facial feature associated with trait perception is the facial width-to-height ratio (FWHR), the width of the face divided by the height of the upper face, which has been hypothesized to be a cue of threat and dominance in the human face [Carré et al. 2009]. In a large-scale meta-analysis study, Geniole et al. [2015] concluded that individuals with a larger FWHR were judged by observers as more threatening, more dominant and less attractive, compared to those with smaller FWHRs. Another

\(^1\)http://www.populateproject.eu/
Facial features of non-player creatures can influence moral decisions in video games

A characteristic found to affect personality judgments on human faces is eye size. Eye size has been linked to the perception of dominance [Keating 1985] and honesty [Zebrowitz et al. 1996]. Specifically, people tend to perceive faces with bigger eyes as more honest and less dominant.

A study on the effect of these features on the perception of virtual human faces suggested that these traits might not apply to virtual characters [Wang et al. 2013]. Other studies have shown effects of render style [Zibrek and McDonnell 2014], realism [Hyde et al. 2013], and shape and material [Zell et al. 2015] on judgments of personality of virtual characters. However, previous research has not included extremely abstract-looking characters that have minimal facial features. It is not known whether facial features that have been found to influence perception of human faces could similarly influence the perception of abstract characters. Conceivably, the perception of a virtual character that has only a pair of round eyes on a simple spherical head might rely heavily on the configuration of these limited cues. In experiment 1, we addressed the question of whether findings about personality perception of virtual human faces could be extended to the design of abstract virtual faces [Ferstl et al. 2016]. We found that different rules of perception apply to abstract virtual faces.

Rules for the link between facial features and personality perception can help artists during the creation of abstract characters, especially when aiming for certain personalities, or when trying to affect the player’s choices and game experience. Quality game experience relies on engaging the player, which according to game designer Sid Meier, is providing “a series of interesting choices”, which, in turn should leave players emotionally invested [Sicart 2013]. Following Meier’s guidelines, game designers often use moral dilemmas as scenarios that both trigger emotional involvement and require a difficult choice; the Fallout [Interplay Entertainment 1998; Bethesda 2015] and the Witcher series [Atari 2007; CD Projekt 2015] are just two of the more popular examples of this. Moral dilemmas tend to successfully engage players’ emotions and make for more interesting playing experiences in video games.

In experiment 2, we followed this guideline and utilized moral dilemmas for testing the theory that the visual appearance of an in-game character can influence decisions in a video game. Because moral dilemmas present the subject with two answer alternatives that differ in their associated aggressiveness, they also provide a way to specifically test the aggressiveness of a player’s behavior, as will be discussed later in more detail. Additionally, moral dilemmas are popular in research about external influences on decision processes because they engage competing cognitive and emotional reasoning [Greene et al. 2004; Greene and Haidt 2002; Moll and de Oliveira-Souza 2007].

A game player’s behavior has previously been shown to be influenced by visual perception of a self-avatar [Messinger et al. 2008; Banakou and Chorianopoulos 2010]. In one study, for example, players behaved more confidently when assigned taller avatars, and more verbally intimate when assigned more attractive avatars [Yee and Bailenson 2007]. In another study, female players assigned sexualized avatars reported more body-related thoughts and showed more acceptance of false beliefs about rape that blame the victim [Fox et al. 2013]. Players’ aggressiveness, too, can be influenced by manipulating the appearance of a self-avatar. In-game aggressiveness has been shown to be changeable by assigning players a white- versus a black-cloaked avatar, or an avatar associated with a doctor versus a Ku-Klux-Klan member [Peña et al. 2009], by giving players the chance to customize their avatar [Hollingdale and Greitemeyer 2013], and by manipulating the avatar’s gender [Eastin 2006]. Furthermore, some of these effects may even extend beyond the game environment, inducing changes in self-perception by creating stronger associations between oneself and the played character [Klimmt et al. 2010], and influencing real-life behavior in the same direction as in-game behavior [Yee et al. 2009]. Furthermore, Headleand et al. [2016] have shown that the effects of character appearance are not limited to the design of self-avatars, but that the visual appearance of an icon alone, representing an AI character,
can influence a player’s behavior towards this character; players acted more protectively towards an AI represented by a human head icon versus a robot head icon.

Focusing heavily on visual design of self-avatars, previous research has mostly omitted investigations of effects of non-player character (NPC) appearance on gameplay. We sought to extend the current understanding of character design by testing if a player’s perception of an NPC could influence their behavior in a video game. Specifically, we hypothesized that the perception of aggressiveness and trustworthiness of a character can influence the aggressiveness of the player’s behavior in the game. The aggressiveness of the player’s behavior was measured by the player’s moral choices within the game.

Previous studies have categorized moral dilemmas into one of two main types. The first type is the direct harm dilemmas, also called ‘personal’ dilemmas, where causing harm is used as a means for avoiding greater harm [Greene et al. 2001, 2004; Moore et al. 2011]. In direct harm dilemmas, the subject is presented with a situation that offers the option to harm or kill another person in order to achieve a higher goal; that is, the subject would knowingly generate an action to directly harm another person. An example of a direct harm dilemma is the footbridge dilemma [Thomson and Parent 1986]:

A runaway trolley is headed for five people working on the tracks and oblivious to the imminent danger. If the trolley proceeds on its present course, the five people will be killed. You are standing next to an unusually large stranger on a footbridge above the tracks, between the approaching trolley and the people on the tracks. The only way to save them is to push the stranger off the bridge, onto the tracks below. This means he will die, but his body will stop the trolley and save the other five people. Would you push the stranger onto the tracks in order to save the five others?

The second category is indirect harm dilemmas, also called ‘impersonal’ dilemmas, where harm is an unintended side-effect for avoiding greater harm [Greene et al. 2001, 2004; Moore et al. 2011]. In indirect harm dilemmas, the subject is presented with a situation that offers performing an action that only indirectly harms or kills another person and helps achieve a higher goal, e.g., by redirecting harm from a large number of people towards an individual. In indirect harm dilemmas, the harm caused by the subject is “just” collateral damage, caused by protecting a number of people. An example of an indirect harm dilemma is the trolley dilemma, a variation of the above footbridge dilemma [Thomson and Parent 1986]:

A runaway trolley is headed for five people on the tracks. If the trolley proceeds on its present course, they will be killed. The only way to save them is to hit a switch that will redirect the trolley onto a different course. On this course, one person is working on the tracks and will be killed by the redirected trolley. Would you hit the switch to turn the trolley in order to save the group of five people but kill the single worker?

Choosing to execute the proposed action represents the utilitarian choice, the choice to cause some harm in order to avoid even greater harm. Not performing the proposed action, on the other hand, represents the deontological choice and is led by the evaluation of the ethical correctness of the action itself (“do no harm”). The utilitarian choice is typically viewed as the more aggressive choice, and having an aggressive personality has been linked with more utilitarian choices [Choe and Min 2011; Gao and Tang 2013]. We therefore hypothesized that the aggressiveness of a presented character would influence the subject’s willingness to make aggressive, i.e., utilitarian, choices. We further hypothesized that the perceived trustworthiness of the character presenting the moral choice could influence the subjects decision by providing more or less confidence in the rightfulness of the decision.
Facial features of non-player creatures can influence moral decisions in video games

It has been suggested that the two categories of dilemmas, direct and indirect harm, are processed differently in the brain [Greene et al. 2001] and we therefore chose to look at both categories of moral dilemmas in order to assess any possible effect of dilemma type on the participants playing behavior.

3. SUMMARY OF EXPERIMENT 1: FACIAL FEATURE PERCEPTION

In experiment 1, further detailed in Ferstl et al. [2016], we examined whether facial features of abstract virtual characters could influence people’s perception of the character’s personality traits. We sought to find design rules that would allow us to create abstract characters that convincingly mime a personality type in order to subsequently use them to investigate the influence of an NPC’s perceived personality on player behavior.

We specifically aimed to test the effect of the facial width-to-height ratio (FWHR) and eye size, as previous studies on the perception of human faces have found effects of these facial features on personality perception. Additionally, we manipulated the eye shape, an exploratory choice not linked to previous research due to its restrictions to realistic human features.

We retrieved two abstract virtual characters from the game designed for the EU project POPULATE 2 (Figure 1). Both of the two original characters were manipulated to create three different head shapes (wide, round, narrow), three different eye shapes (almond, round, rectangular), and three different eye sizes (small, medium, large) (for measurements see supplementary material section 2). Combining all eyes and head manipulations, we had a stimulus set of 54 test images (section 1 in supplementary material).

Participants were asked to rate perceived dominance, trustworthiness, aggressiveness, appeal and eeriness of the characters.

We will briefly summarize our results below; for a more detailed analysis and discussion, please refer to Ferstl et al. [2016].

Our results stood in contrast to some of the findings of personality perception for human faces. For human faces, a larger FWHR has been consistently linked with increased ratings of dominance and aggressiveness [Geniole et al. 2015], and may be linked with decreased ratings of trustworthiness [Stirrat and Perrett 2010; Wilson and Rule 2015]. Contrary to this, for our abstract characters, we found narrow faces to be perceived as most aggressive as well as most dominant.

For eye size, human faces with larger eyes have been associated with increased ratings of trustworthiness [Zebrowitz et al. 1996], as well as decreased ratings of dominance [Keating 1985]. However, for our abstract characters, medium-sized eyes, not large eyes, were linked to the highest ratings of trustworthiness, while there was no effect on dominance ratings. Medium eyes were also most appealing and least eerie; large eyes were least trustworthy and appealing, and most eerie. For eye shape, the almond shape, was perceived as most aggressive, most dominant, least trustworthy and most eerie. The classic round eye shape was perceived as the least aggressive and the least dominant.

With our findings, we could validate a set of stimuli to portray the specific personality traits we required for our subsequent study. Our validated set of stimuli was composed of the four characters whose facial configurations were perceived as looking least and most aggressive, and least and most trustworthy (Figure 2). We used this set of stimuli to test effects of personality perception on a player’s behavior in a video game in experiment 2. Previous studies have shown that the appearance of a self-
Avatar influences a player’s behavior [Yee and Bailenson 2007; Fox et al. 2013; Messinger et al. 2008; Banakou and Chorianopoulos 2010], and specifically in-game aggressiveness [Peña et al. 2009; Hollingdale and Greitemeyer 2013]; we designed an experiment to test for similar effects for appearances of non-self characters.

4. EXPERIMENT 2: MORAL DILEMMAS

In experiment 2, we investigated whether perceived personality traits of a virtual character could actually influence players’ choices and actions within a game. We created a short online video game where participants were presented with four different moral dilemmas. Each dilemma was presented by a different character asking the player to solve the situation. We chose the four characters that received the highest and lowest trustworthiness rating, and the the highest and lowest aggressiveness rating in experiment 1 (Figure 2). As in other moral dilemma studies, participants were given a choice to either execute an action that will sacrifice one life in order to save others, or to not execute the proposed action. Executing the action represents a utilitarian choice, following the maxim of saving the most lives. Not executing the action represents a deontological choice, which focuses on the rightness (or wrongness) of the action itself, rather than its consequences.

Because the utilitarian choice represents the more aggressive alternative and is associated with a more aggressive personality [Choe and Min 2011; Gao and Tang 2013], we hypothesized that the presence of a more aggressive-looking character might trigger more aggressive thoughts and behavior, just as more aggressive-looking self-avatars do [Peña et al. 2009], and hence lead to more utilitarian choices. For the dimension of trustworthiness, we were interested to see if a more trustworthy-looking character would lead to feeling more at ease and less aggressive, and hence less utilitarian choices, or alternatively, if a more trustworthy-looking character would make the subject more willing to execute the proposed action because it is suggested by a character that seems honest and helpful. Previous studies have shown the perceived trustworthiness of an opponent’s face can influence decisions in a poker game [Schlicht et al. 2010] and in a trust game [Stirrat and Perrett 2010; Vant Wout and Sanfey 2008]; we were interested in investigating whether a face’s trustworthiness could also influence decisions about a third party.

Moral dilemmas have been assumed to fall into one of two main categories, direct harm dilemmas, where causing harm is used as a means to avoiding greater harm, and indirect harm dilemmas, where harm is an unintended side-effect of avoiding greater harm [Greene et al. 2004; Moore et al. 2011]. In order to account for dilemma categories possibly being influenced differently by character appearance, we therefore chose to phrase the dilemmas in a way that two represented direct, and two indirect harm. We found differing effects for direct and indirect harm dilemmas and discuss possible reasons for our findings.

4.1 Methods

With our data from experiment 1, we determined the four characters that were rated (1) most trustworthy, (2) least trustworthy, (3) most aggressive, and (4) least aggressive (Figure 2). The dilemmas were modelled after the classic trolley and footbridge dilemmas, but adapted in a way that they would fit the fantasy forest atmosphere of the game environment. For each dilemma, the participants had a choice of either executing a proposed action or refraining from doing anything. Executing the proposed action always meant sacrificing a single life to save multiple others (utilitarian choice) which would not be spared when refraining from the action (deontological choice) (for dilemma phrasing see supplementary material section 4).

We collected data from 48 participants (10 females, 1 ‘other’ gender, ages 19-61 years, $M = 26.1$, $SD = 8.87$). The study was run online and participants were recruited via university mailing lists.
Facial features of non-player creatures can influence moral decisions in video games

Each participant completed each of the four moral scenarios once, and each scenario was presented by a different character. The four dilemmas were always presented in the same order, but character to dilemma assignment was manipulated between participants. This character assignment was balanced for dilemma type across participants; overall, each character was paired with each dilemma type 24 times.

The game was created using the Unity game engine. Before the presentation of the dilemmas, the scene was set with a short text introduction describing the participant’s role in the virtual world (supplementary material section 3). Each trial started with one of the four characters running into view, stopping in front of the subject’s viewpoint, and then presenting one of the dilemmas. The dilemmas appeared as text below the character, while the character made unintelligible noises as if the character was talking. This audio was always generated using the same 1-second audio clip and manipulating it by continuously randomizing its pitch during the duration of playing. The audio and talking animation played for the duration of the dilemma text being printed letter-by-letter to the screen (1 letter/.15 seconds). The subject then had to decide on one of the two answer alternatives (executing or not executing the proposed action) by pressing the respective button on-screen. There was no consequence to choosing either answer; in each case, the character then walked off and a time-lapse animation simulated the passing of 3 days before the next character then ran into view. Calm music with some bird chirping played in the background of the whole experiment.

4.2 Results

We created a generalized linear mixed model by maximum likelihood (Laplace Approximation) using glmer from the R lme4 package [Bates et al. 2015]. Our model included the factors character (low/high aggressiveness/trustworthiness), dilemma type (direct/indirect harm), and repetition (repetition 1 and 2 for each dilemma type), and an interaction term for character and dilemma type. Subjects were specified as a random factor. We included the factor repetition in our model to see if subjects answered more aggressively when noticing after the first answered dilemma that there is no visual consequence to their choice. We chose to use a generalized mixed model because of the binary response variable and due to the fact that each participant completed each type of dilemma only with two of the characters. Overall, character to dilemma type assignment was balanced, with each character being paired with each dilemma type 24 times. Response choices were binary (deontological/ utilitarian) and we calculated the proportion of utilitarian choices as the response variable.

ACM Transactions on Applied Perception, Vol. 000, No. 000, Article 000, Publication date: 2017.
Our results showed that the character rated high on aggressiveness was associated with significantly more utilitarian responses compared to the low aggressiveness character \((p < .05)\), and marginally more than the low trustworthiness character \((p = .065)\). There was a significant interaction between the character rated high on aggressiveness and dilemma type \((p < .01)\), with the high aggressiveness character being associated with less utilitarian choices for the direct harm dilemmas, and more utilitarian choices for indirect harm. There was an opposite interaction between the character rated low on aggressiveness and dilemma type \((p < .01)\) (see Figure 3).

Dilemma type had a significant effect \((p < .001)\); subjects made significantly more utilitarian choices in the indirect harm dilemmas. Repetition had a significant effect, with subject giving significantly more utilitarian responses in the second repetition of each dilemma type \((p < .01)\).

A Hosmer Lemeshow goodness of fit test revealed no significant difference between our model and the observed data \((p > .05)\). An ANOVA on our model, showed a significant effect of dilemma type \((\chi^2(1) = 21.61, p < .001)\) and repetition \((\chi^2(1) = 7.00, p < .01)\), as well as a marginally significant interaction between dilemma type and character \((\chi^2(3) = 7.79, p = .05)\). Removing character from the model made it unfit to explain the data, described by a significant difference between the reduced model and the observed data (Hosmer Lemeshow test: \(\chi^2(8) = 17.99, p < .05\)), as did removal of dilemma type \((\chi^2(8) = 21.55, p < .01)\).

To investigate whether the appearance of the first presented character would influence subjects’ decisions in later trials, we modified our model to include a factor identifying the first shown character. We found no significant effect of the first character biasing subsequent moral decisions.

### 4.3 Discussion

Our results show that character appearance can significantly affect moral choices, and that this is modulated by dilemma type. The aggressive-looking character interacted with dilemma type, and the high aggressiveness character being associated with less utilitarian choices for the direct harm dilemmas, and more utilitarian choices for indirect harm. There was an opposite interaction between the character rated low on aggressiveness and dilemma type \((p < .01)\) (see Figure 3).

Dilemma type had a significant effect \((p < .001)\); subjects made significantly more utilitarian choices in the indirect harm dilemmas. Repetition had a significant effect, with subject giving significantly more utilitarian responses in the second repetition of each dilemma type \((p < .01)\).

A Hosmer Lemeshow goodness of fit test revealed no significant difference between our model and the observed data \((p > .05)\). An ANOVA on our model, showed a significant effect of dilemma type \((\chi^2(1) = 21.61, p < .001)\) and repetition \((\chi^2(1) = 7.00, p < .01)\), as well as a marginally significant interaction between dilemma type and character \((\chi^2(3) = 7.79, p = .05)\). Removing character from the model made it unfit to explain the data, described by a significant difference between the reduced model and the observed data (Hosmer Lemeshow test: \(\chi^2(8) = 17.99, p < .05\)), as did removal of dilemma type \((\chi^2(8) = 21.55, p < .01)\).

To investigate whether the appearance of the first presented character would influence subjects’ decisions in later trials, we modified our model to include a factor identifying the first shown character. We found no significant effect of the first character biasing subsequent moral decisions.

**Fig. 3.** Effects of character on utilitarian choices, plotted for each dilemma type separately. Plotted is the overall proportion of utilitarian choices (the likelihood to sacrifice one life to save multiple others).

**Fig. 4.** The indirect harm dilemma type was associated with significantly more utilitarian choices. Utilitarian choices were also more frequent for the second presentation of each dilemma type.
Facial features of non-player creatures can influence moral decisions in video games

modulating response behavior towards more utilitarian choices in indirect harm dilemmas, and towards less utilitarian choices in direct harm dilemmas. In contrast, the non-aggressive-looking character modulated responses towards more utilitarian choices for direct harm dilemmas, and towards less utilitarian choices for indirect harm dilemmas. These effects could possibly be partly due to the aggressive-looking character increasing the perceived aggressiveness of the already very violent proposed action in direct harm dilemmas, while the non-aggressive-looking character had the opposite effect. This explanation, however, is limited to the effects for direct harm dilemmas.

Here, we looked at the influence of the perception of the two dimensions aggressiveness and trustworthiness. Aggressiveness has previously been associated with utilitarian moral choices [Choe and Min 2011; Gao and Tang 2013], and we included trustworthiness as an exploratory variable that we assumed might play a role in a player's in-game moral choices. However, we found no effect of the character’s perceived trustworthiness on moral choices.

We found direct harm type dilemmas to be associated with fewer utilitarian choices, in line with previous findings [Thomson and Parent 1986; Royzman and Baron 2002; Greene et al. 2004; Chan et al. 2016]. Participants were also more likely to choose the utilitarian option the second time they encountered each dilemma type. This might be due to the lack of visual consequences to subjects moral choice, as after the first round it became clear that they would not have to watch the consequences of their choices. It would be interesting to test if visual feedback about the participant’s choice would change the response strategies.

In our experiment, the ethical dilemmas were presented in a text format while the present character emitted unintelligible sounds. If the dilemmas were presented more directly by the visible character, i.e., with an audio stream explaining the dilemma and with lip-synchronisation for the character, the effect of the presenting character might be stronger. The fact that subjects had to read the dilemma description from a box below the character might have additionally caused them to pay less attention to the character, decreasing its effect on the decision process.

We did not run a pre-test on the possible effects of our audio stream, and different audio might affect results. Due to the high randomization of the speech-imitating audio part, variation between trials can be considered indistinguishable, and therefore any audio effects should be limited to across-experiment rather than within-experiment. The constant calm background music could have put subjects in a more relaxed state of mind, which could have reduced the dilemma’s power to induce negative emotions.

Furthermore, we did not subject the dilemma scenarios to any prior testing and we therefore cannot estimate how effective the descriptions were for inducing emotional engagement. While the scenarios were based on the classic trolley dilemma, the original dilemma has been utilized and tested in multiple morality studies, while our dilemmas were reformulated to a degree that they could create a different experience than its template. Our reformulation of the classic trolley scenario gives the advantage, however, of complete novelty for every participant, avoiding the possibility of individual participants being already familiar with the scenario through prior experiment participation.

Our findings could be used to create more provoking moral dilemmas in games, e.g., by using a more aggressive-looking character for scenarios where most people would choose not to execute an utilitarianistic action, in order to create greater conflict.

5. GENERAL DISCUSSION

We ran two experiments, first analyzing which facial features can be used to portray certain personality traits in abstract virtual characters, and subsequently testing whether characters portraying specific personality traits could influence moral decisions in games.
The results from experiment 1 indicate that different rules should be considered when creating abstract virtual characters versus human characters. For example, while wide faces may portray an aggressive and dominant personality for human faces, these traits may be achieved with narrow faces for abstract characters. Our results can be used as guidelines for the creation of abstract characters that are intended to portray certain personality features. With the data from experiment 1, we validated a set of stimuli for our subsequent experiment, composed of the characters rated as least and most aggressive, and least and most trustworthy.

With experiment 2, we sought to examine whether perceived personality traits of an abstract NPC could actually influence gameplay. This aimed to extend previous findings that player behavior can be influenced by the visual appearance of a self-avatar [Yee and Bailenson 2007; Messinger et al. 2008; Peña et al. 2009; Banakou and Chorianopoulos 2010; Fox et al. 2013; Hollingdale and Greitemeyer 2013] and by the appearance of the character that is the beneficiary of a behavior [Headleand et al. 2016]. In our study, we manipulated the appearance of an abstract creature that is both an NPC and not the target of an action. We found that the perceived personality of the character can influence a player's moral choices in a video game. Specifically, the visual appearance of the NPC affected decisions depending on moral dilemma type, with the aggressive-looking character triggering less utilitarian choices in direct harm scenarios, and more deontological choices in indirect harm scenarios. The opposite was true for non-aggressive-looking character. These findings could be used to create more challenging moral dilemma scenarios in video games by selectively employing characters likely to bias decisions towards the less-common response behavior. Whereas we find a specific effect of perceived aggressiveness on players’ moral choices, this shows a general possibility of using impressions of personality of an NPC to selectively influence player experience.

The characters in our studies were limited to the designs of two creatures from the EU POPULATE project3 and future research without such limitations could potentially better control for some design variables. For example, our characters differ in color and in selecting the overall most/least aggressive/trustworthy-looking characters from experiment 1, characters in experiment 2 were not balanced with respect to color.

A future study could explore whether visual consequences will influence subsequent moral decisions. Future research may also investigate the possible influence of other dimensions of personality for other, non-moral gaming scenarios.

REFERENCES

3http://www.populateproject.eu/
Facial features of non-player creatures can influence moral decisions in video games


Janine Willis and Alexander Todorov. 2006. First impressions making up your mind after a 100-ms exposure to a face. *Psychological science* 17, 7 (2006), 592–598.


ACM Transactions on Applied Perception, Vol. 000, No. 000, Article 000, Publication date: 2017.
Facial features of non-player creatures can influence moral decisions in video games


Eduard Zell, Carlos Aliaga, Adrian Jarabo, Katja Zibrek, Diego Gutierrez, Rachel McDonnell, and Mario Botsch. 2015. To stylize or not to stylize?: the effect of shape and material stylization on the perception of computer-generated faces. *ACM Transactions on Graphics (TOG)* 34, 6 (2015), 184.

Supplementary Material

1 Stimuli for experiment 1
Facial features of non-player creatures can influence moral decisions in video games
2 Facial measurements for experiment 1

**FWHR:**
The facial width-to-height ratio was measured by dividing the widest part of the face by the length of the face. Note that this differs from how this measurement is taken on human faces; for human faces, the upper lip and midbrow are used as landmarks for defining the measuring area, but these landmarks are not available on our abstract characters.

<table>
<thead>
<tr>
<th>Character</th>
<th>Narrow face</th>
<th>Round face</th>
<th>Wide face</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (light)</td>
<td>0.6</td>
<td>1.1</td>
<td>1.55</td>
</tr>
<tr>
<td>B (dark)</td>
<td>0.65</td>
<td>0.9</td>
<td>1.3</td>
</tr>
</tbody>
</table>

**Eye size:**
Eye size was measured as percentage of facial area covered by one eye.

<table>
<thead>
<tr>
<th>Character</th>
<th>Eye shape</th>
<th>Small eyes</th>
<th>Medium eyes</th>
<th>Large eyes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (light)</td>
<td>Almond</td>
<td>1.8%</td>
<td>4%</td>
<td>6.6%</td>
</tr>
<tr>
<td></td>
<td>Rectangle</td>
<td>1.3%</td>
<td>3.4%</td>
<td>7.9%</td>
</tr>
<tr>
<td></td>
<td>Round</td>
<td>1.2%</td>
<td>5.2%</td>
<td>9.3%</td>
</tr>
<tr>
<td>B (dark)</td>
<td>Almond</td>
<td>1.0%</td>
<td>2.5%</td>
<td>4.1%</td>
</tr>
<tr>
<td></td>
<td>Rectangle</td>
<td>1.5%</td>
<td>3.1%</td>
<td>6.8%</td>
</tr>
<tr>
<td></td>
<td>Round</td>
<td>0.5%</td>
<td>3.1%</td>
<td>6.2%</td>
</tr>
</tbody>
</table>
Facial features of non-player creatures can influence moral decisions in video games

3 Introduction text for experiment 2

4 Phrasing of dilemma scenarios for experiment 2

**Scenario 1 (indirect harm):**
“The volcano on the neighbouring island is about to break out and the inhabitants of the island won’t have enough time to make it out of harms way quickly enough before the lava reaches them. You are gifted with the ability to cast a spell that will slow the process of the eruption enough for the islands inhabitants to flee in their boats. However, this magic will draw energy from one of the people around you and kill them instantly. Will you do nothing and let the people on the island fall victim to the volcano, or will you cast the spell in order to buy enough time for the islands inhabitants to flee but cause someone in your area to die?”

Option 1 = “Cast the spell.”  
Option 2 = “Don’t do anything.”

**Scenario 2 (direct harm):**
“The altar and its powerful crystals have to be protected at all costs because it keeps the balance of the forces and the nature. A member of one of the tribes has tried to steal crystals from the altar in order to gain more power for their tribe and have an advantage over the other tribal groups. One of
those tribal groups now demand justice by urging you, the protector of the altar, to execute the guilty person. They have already taken 5 hostages of the tribe they suspect as the culprits. It is unknown who really committed the crime. Will you do nothing and be certain that the 5 hostages will be killed as payment, or will you frame a possibly innocent member of the suspected tribe and execute them in order to free the hostages.”

Option 1 = “Frame somebody.”
Option 2 = “Don’t do anything.”

**Scenario 3 (indirect harm):**
“Due to someone trying to tamper with the altar in the holy forest, the balance of the spirits were shortly disturbed and have caused the formation of a poisonous smoke spirit that you can now see drifting fast towards one of the tribal village valleys. You have the power to call upon the wind to deflect the course of its propagation and blow it out towards the sea where it would slowly and safely dissolve. However, that way the smoke would inevitably pass over harbour house where the sea guard is living, and poisoning the person instantly. Will you do nothing and let the smoke continue towards the village valley where it will disperse and poison the valley population, or will you summon the wind and force the smoke out towards the water but across the dock house where it will poison the sea guard?”

Option 1 = “Call the wind.”
Option 2 = “Don’t do anything.”

**Scenario 4 (direct harm):**
“Someone tried to steal from the altar and so the spirit of the altar got angry and is brewing up a small natural disaster, as it has in the past, which we know will lead to the deaths of several people. As the protector of the altar, you can choose to sacrifice someone from the villages to the altar to calm down the spirit, which has worked as a strategy in the past. Will you wait and do nothing and let the spirit take its victims, or will you choose somebody to sacrifice as an offering to the altar in order calm it down?”

Option 1 = “Sacrifice somebody.”
Option 2 = “Don’t do anything.”