

Gender Differences in the Impact of Presentational Factors in Human Character Animation on Decisions in Ethical Dilemmas

Abstract

Simulated humans in computer interfaces are increasingly taking on roles that were once reserved for real humans. The presentation of simulated humans is affected by their appearance, motion quality, and interactivity. These presentational factors can influence the decisions of those who interact with them. This is of concern to interface designers and users alike, because these decisions often have moral and ethical consequences. However, the impact of presentational factors on decisions in ethical dilemmas has not been explored. This study is intended as a first effort toward filling this gap. In a between-groups experiment, a female character presented participants with an ethical dilemma. The character's human photorealism and motion quality were varied to generate four stimulus conditions: real human versus computer-generated character \times fluid versus jerky movement. The results indicate that the stimulus condition had no significant effect on female participants, while male participants were significantly more likely to rule against the character when her visual appearance was computer generated and her movements were jerky.

I Introduction

Modern computer graphics technologies are enabling the creation of animated human characters that can engage people in dialogues that are increasingly natural and complex. Advances in speech synthesis, digital sculpting and rendering, motion capture and synthesis, and the graphics rendering pipeline have significantly enhanced the photorealism, motion quality, and interactivity of these characters, heightening the user's appreciation of their social presence. The benefits of characters that mimic face-to-face communication more closely may include decreased cognitive effort and communication ambiguity (Kock, 2005).

Many studies have investigated how computer-generated (CG) characters in conversational interfaces affect their human interlocutors' attitudes and performance (Dehn & van Mulken, 2000; Yee, Bailenson, & Rickertsen, 2007; Zambaka, Ulinski, Goolkasian, & Hodges, 2007). It has also been established that aspects of a situation that seem to have no principled relevance to a decision

can dramatically affect its outcome (Bateson, Nettle, & Roberts, 2006; Darley & Batson, 1973; Doris & Stich, 2005; Haley & Fessler, 2005; Matthews & Cannon, 1975). However, the influence of a character's presentation on decision making in an ethical dilemma has not been examined before this study.

Two groups that must eventually come to terms with the impact of computer-generated characters on decisions in ethical dilemmas are medical and legal professionals. Doctors and judges are compelled to weigh ethical principles daily. Doctors, specifically, have a *prima facie* duty to treat patients in accordance with the principles of beneficence, nonmaleficence, autonomy, and informed consent (Beauchamp & Childress, 1979). Although we expect doctors and judges to make unbiased decisions in ethical dilemmas, their judgment can be swayed by factors unrelated to the dilemma. Reminders of mortality, for example, can cause medical students to make assessments of cardiac risk biased against religious out-groups, sending Muslim patients home when they would otherwise be kept under observation (Arndt, Vess, Cox, Goldenberg, & Lagle, 2009). Such reminders can also cause judges to set bail nine times higher for the same prostitution case (Rosenblatt, Greenberg, Solomon, Pyszczynski, & Lyon, 1989). As virtual characters begin to appear in sensitive environments, such as hospital interfaces, their effect on people becomes more important to understand and control. The danger of seemingly irrelevant aspects of a character's presentation affecting decisions of moral and ethical consequence has significant implications for designers of avatars, conversational characters, social robots, and other human-looking interfaces. If these effects are to be controlled, it will be necessary to develop an understanding of how presentational factors manipulate decision making on morally significant issues.

The present study examines how presentational factors impact the moral and ethical constructs that drive behavior. These factors are varied in a between-groups experiment on two dimensions: human photorealism and motion quality. In the scenario, a character presents participants with a dilemma in medical ethics to which they respond in the role of consulting physician. The

effects of presentational factors on the participants' decisions are then analyzed.

1.1 How Extraneous Factors Affect Moral Judgment

Virtue ethics, which finds its roots in the philosophy of Aristotle and Socrates, concerns what character traits make for a good person. Throughout the development of Western philosophy until the twentieth century, issues of moral character were deemed integral to the study of what it means to be ethical (Collinson, 1987). Kant, Hume, and Mill were all concerned with the psychological basis of moral character (Beardsley, 1960).

However, the importance of moral character has been undermined in recent years by findings showing that situational factors greatly influence outcomes in moral dilemmas (Doris, 1998, 2002; Doris & Stich, 2006; Harman, 1999; for dissenting views, see Fleming, 2006; Webber, 2006). Recent experiments suggest that a combination of an individual's normative viewpoint and emotion, especially disgust, distinguish moral questions from those that can be answered merely by convention (Nichols, 2002; Nichols & Mallon, 2006). People may rely on visceral reactions to determine right from wrong as much as cognitive reasoning (Schnall, Haidt, Clore, & Jordan, 2008). Areas responsible for emotion in the brain play a key role in morally significant decisions (Bechara, Damasio, & Damasio, 2000; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Luo et al., 2006). Patients with damage to these areas do not hesitate to make decisions most other people would find rational but immoral (Koenigs et al., 2007). Minor differences have also been detected in the moral reasoning of men and women. For example, women may be more likely to make ethical choices (Glover, Bumpus, Sharp, & Munchus, 2002) and reason on the basis of care rather than justice (Jaffee & Hyde, 2000; Wark & Krebs, 1996).

People's emotional state can easily be affected by aspects of a situation that are not directly related to the ethical dilemma; their emotional state, in turn, influences their decision (Isen, 1987; Prinz, 2006). An experiment conducted by Isen and Levin (1972) showed

that 14 of 16 people spontaneously helped a stranger collect the pages of a dropped manuscript, if they had just found a coin in the coin return of a payphone, whereas only one of 24 helped, if they had not. Matthews and Cannon (1975) found that only 15% of participants helped an injured man pick up his books when a lawnmower was running nearby as compared with 80% when the lawnmower was turned off. Gueguen and De Gail (2003) found a passerby would be more likely to help the experimenter pick up some dropped computer disks if a confederate had smiled at the passerby moments earlier. What is fascinating about these studies is that seemingly unconnected events substantially influenced the outcomes of morally significant decisions without the awareness of the decision maker. These studies raise concerns that ethical judgments may depend more on extraneous factors than on moral character or the application of ethical principles.

Researchers have also explored the impact of presentational factors, specifically physical attractiveness, on the ability to perceive, interact with, and judge other people. People are more easily persuaded by an attractive person than an unattractive person (Chaiken, 1979), and they are also more inclined to perform altruistic acts for an attractive person, such as helping to relocate or donating blood (Cunningham, 1986). Attractiveness can influence the decisions of senior nursing students (Peternelj-Taylor, 1989). In addition, medical residents assess the health of less attractive female patients more negatively (Hadjistavropoulos, Ross, & Von Baeyer, 1990). Perceptions of attractiveness can also influence courtroom outcomes (Sigall & Ostrove, 1975). Mock jurors were more likely to recommend psychiatric treatment for unattractive defendants (McKelvie & Coley, 1993). Although 93% of mock jurors claimed physical appearance should not bias their decisions, they were less certain about the guilt of attractive defendants and recommended lighter punishments (Efran, 1974).

The influence of attractiveness on human judgment is not homogeneous with respect to gender. A meta-analysis of previous studies concluded that men are more influenced than women by the physical attractiveness of the opposite sex (Jackson, Hunter, & Hodge,

1995). Other research expands on this to show that, although men are more sensitive to the physical appearance of women than that of other men, women are more consistent in their evaluations (Kniffin & Wilson, 2004). These particular findings are relevant to this study, because a pronounced difference was found in the responsiveness of male and female participants to the stimuli.

1.2 Presentational Factors in Human-Machine Interaction

Studies have found both similarities and differences in the way people interact with computers and other nonhuman media as compared to other people (MacDorman & Ishiguro, 2006; Nass & Moon, 2000; Reeves & Nass, 1996; Shechtman & Horowitz, 2003). As an example that highlights the similarity between human-human and human-agent interaction, participants performed worse on a complex task when they were being observed regardless of whether it was by virtual humans or real humans (Zanbaka et al., 2007).

What complicates the comparison and makes quality interaction with virtual characters difficult to design is the challenge of creating characters that are consistently humanlike. Failure to match the degree of realism in a virtual character's appearance, motion quality, interactivity, and other elements results in a phenomenon known as the uncanny valley (Ho, MacDorman, & Pramono, 2008; MacDorman, Green, Ho, & Koch, 2009). The uncanny valley was first proposed by Mori (1970), who noted that the more human a robot looks, the more likely subtle nonhuman defects are to be perceived as creepy. Numerous studies have examined how features of characters that are nearly human elicit the uncanny valley effect in participants (Green, MacDorman, Ho, & Vasudevan, 2008; Hanson, 2006; Seyama & Nagayama, 2007). Flaws in physical appearance, a lack of contingent interaction, and poor motion quality are major causes of negative evaluations of human-looking robots (MacDorman & Ishiguro, 2006). In particular, excessive jerkiness, rigidity, or wobble can cause a human-looking robot's actions to seem unnatural and even unintentional. Although Mori originally

applied his observations to robotics, the uncanny valley has also been attributed to computer-generated human characters in video games and computer-animated films (Geller, 2008).

Neuroimaging studies have shown that a mechanical-looking form engages brain areas used in reasoning about the intentions of others with less than a human-looking form (Krach et al., 2008; Tai, Scherfler, Brooks, Sawamoto, & Castiello, 2004; however, Gazzola, Rizzolatti, Wicker, & Keysers, 2007, found no such effect). Mechanical-looking movements may have a similar effect, causing human-looking characters that move mechanically to appear disquieting. These neuroimaging studies indicate that people are better able to relate socially to characters that appear to be more human. It follows that people will be more easily persuaded by more humanlike characters in the context of an ethical dilemma.

Despite the problems associated with the uncanny valley, studies have shown that incorporating computer-generated figures into interfaces can make them more engaging (Takeuchi & Naito, 1995). In one study, participants even perceived virtual humans more favorably than real humans (Zanbaka, Goolkasian, & Hodges, 2006). In addition, interfaces that incorporate computer-generated characters with a human-looking appearance, lifelike gestures, or well-synchronized gaze elicit higher task performance than interfaces that incorporate less realistic characters (Dehn & van Mulken, 2000).

In sum, people report more positive interactions with embodied agents that look more realistic (Yee et al., 2007). Guadagno, Blascovich, Bailenson, and McCall (2007) found that the behavioral realism of virtual humans and the belief that they were representations of actual people heightened their influence on people. These results are consistent with the model of social influence in immersive virtual environments proposed by Blascovich (2002; Blascovich et al., 2002).¹ Participants' sense of copresence was highest when the realism

1. MacDorman and colleagues (2005) also found that participants' pattern of breaking eye contact when interacting with a human-looking robot was similar to that of human-human interaction only when they were told the robot was under human control.

of the agent's appearance and behavior were well matched (Bailenson et al., 2005; Nowak & Biocca, 2003).

The hazards of including humanlike characters in ethically charged environments create a pressing need to understand their effects on human judgment. This research seeks to illuminate these effects and to augment recent studies on inter-gender, human-agent interaction. At this time, there is conflicting evidence on whether women are more easily persuaded by female avatars (Guadagno et al., 2007) or avatars of the opposite sex (Zanbaka et al., 2006). However, some studies have found that women are more sensitive to an avatar's nonverbal behavior, copresence, and behavioral realism (Bailenson et al., 2005; Bente, Eschenburg, & Aelker, 2007). This particular finding is not consistent with the results of this study.

1.3 Research Approach

Previous research has shown that experimental manipulations of an embodied agent's movement and contingency during human-agent interaction can affect human reaction to and persuasion by the agent (Yee et al., 2007). What has yet to be determined is whether in artificial environments, aspects of a character's presentation affect decisions of moral consequence.

This study takes as a starting point an ethical dilemma from one of the scenarios of MedEthEx, an online training system for medical ethics (Fleetwood et al., 2000). MedEthEx was developed to train medical students to communicate effectively with patients about ethical issues and to make sound decisions when faced with ethical dilemmas while practicing medicine. Through several iterations of testing, the scenario has been carefully modified for this study based on the results of focus groups to ensure that the participants would be divided in their final decisions, thus increasing the measurement sensitivity of the dependent variables.

This study's scenario requires the participant, acting in the role of consulting physician, to weigh several competing ethical principles. Kelly and Paul Gordon, a married couple, are both patients. The participant interacts with Kelly privately on the day before Paul's sched-

uled physical. Kelly appeals to the principle of confidentiality in trying to dissuade the participant from telling Paul she has contracted genital herpes. The participant may also consider the principle of nonmaleficence, which would justify telling Paul to protect him from harm.

1.4 Hypotheses

Previous work indicates that situational factors affect decisions in ethical dilemmas and that embodied characters can be designed to elicit specific responses. Furthermore, our understanding of the uncanny valley suggests that imperfections in a human-looking character could result in a negative evaluation of the character (MacDorman & Ishiguro, 2006; MacDorman et al., 2009). Hence, we propose two hypotheses to gauge the impact of presentational factors on participants' approach to weighing conflicting ethical principles. The first issue concerns the persuasiveness of a computer-generated humanlike character compared to a real human. It is expected that the filmed human will be better able to persuade the participants by virtue of her greater realism.

H1. Participants are more likely to decide in favor of a female patient when she is represented by a human actress than by a computer-generated human model.

MacDorman and Ishiguro (2006) concluded that poor motion quality was a major cause of negative evaluations in human-looking robots. This was particularly true of jerky movements, which are perceived as unnatural or unintentional. The present study seeks to reproduce jerky movements with a computer-generated and filmed human character, while gauging their effect on the decisions of participants. If a character's motion quality is jerky and unnatural, it will be perceived as unattractive or abnormal, and it should be less persuasive than a character presented with smooth motion quality.

H2. Participants are more likely to decide in favor of a female patient when she is presented with fluid movements than with jerky movements.

Table 1. Number of Participants by Stimulus Condition

Motion quality	Human photorealism	
	Human	Computer generated
Fluid	166	154
Jerky	173	189

2 Methods

2.1 Participants

Participants were recruited from a random selection of undergraduate students and recent graduates of a nine-campus Midwestern university. Of 682 participants, 396 (58.1%) were female, 286 (41.9%) were male, 467 (68.5%) were aged 18 to 25, 637 (93.4%) listed their country of origin as the United States, and 624 (91.5%) were part-time or full-time students.²

Table 1 shows the total number of participants in each stimulus condition. Participants were randomly assigned to the four stimulus conditions: Human Fluid, Human Jerky, CG Fluid, and CG Jerky. The random assignments were made without matching for gender or other demographic variables. The χ^2 test on the gender proportion of participants showed no significant difference among the four conditions, $\chi^2(3, N = 682) = 5.83, p = .120$, indicating that the random assignment technique reduced systematic bias.

2.2 Stimuli

Each participant interacted with a conversational character that, in the role of patient, presented a dilemma in medical ethics. There were four different groups of participants in this between-groups study, and four different types of stimuli were prepared in the video clips with one type corresponding to each group. The

2. The participants reflected the demographics of the university's undergraduate population (80.1% non-Hispanic white, 6.9% African-American, 3.4% Asian, 3.0% Hispanic, and 6.6% foreign or unclassified). With regard to the sample's representativeness of the undergraduate population as a whole, the measurement error range was $\pm 4.94\%$ at a 99% confidence level.



Figure 1. Kelly Gordon is shown explaining why her husband must not be informed of her testing positive for genital herpes. The image on the top is a frame from the Human conditions, and the image on the bottom is a frame from the Computer-Generated conditions.

figure in the video clip varied in its degree of human photorealism (Human versus Computer-Generated) or motion quality (Fluid versus Jerky).

For each of the four stimulus conditions, there were seven video clips, and between each clip the participant selected one of four possible responses as part of the doctor-patient dialogue. Figure 1 shows the Human (top) and Computer-Generated (bottom) stimulus conditions. The same audio track and computer-generated background were used in all four video conditions. (The human was superimposed on the background using green screen technology.) The video clips were designed for a 4:3 aspect ratio at a size of 480 pixels wide by 360 pixels tall. The conversational character, Kelly Gordon, was represented in the following four stimulus conditions.

Human Fluid. A human character lip-synched to a human audio track against a slightly out-of-focus computer-generated background.

Human Jerky. This condition used the video and audio track of Human Fluid, except five of every six video frames were eliminated to simulate jerky movement, presenting each remaining frame for six times its original duration.

CG Fluid. A computer-generated character lip-synched to the same human audio track against the computer-generated background. The animation in CG Fluid did not rely on motion-capture techniques; instead, it is the result of iterative refinements made after each of seven focus group critiques.

CG Jerky. This condition used the video clips from CG Fluid, except five of every six video frames were eliminated to simulate jerky movement, presenting each remaining frame for six times its original duration.

For experimental control, both the human and CG characters were lip-synched to an audio track produced by a human voice actress. People are sensitive to discrepancies in lip-synching, which can result in a negative evaluation (Reeves & Nass, 1996). Because the process of lip-synching the animated character was imperfect, the human character was also lip-synched to avoid giving the CG character a disadvantage. Though noticeable, the lip-synching did not look out-of-synch or disturbing for either the human or CG character.

2.3 Procedures

The recruitment e-mail directed participants to a website that explained the study's purpose and procedures.³ Once participants had consented to participate, the site tested whether their browser could display the video clips and prompted the participants to adjust their volume to a comfortable level. The next screen explained the role of the participant in the interaction, defined some of the medical terms used, and introduced the scenario.

You're a family doctor who has been treating Kelly Gordon, 27, since her marriage to your patient Paul Gordon, 33, five years ago. Last week you visually diagnosed Kelly with genital herpes and provided her

3. <http://research.joccoram.com>

with educational material and counseling. She phoned in for her lab results, which came out positive for genital herpes. You are scheduled to examine Paul tomorrow for his annual physical, and Kelly asked to be squeezed in for an appointment before Paul's visit.

Participants were then presented with the first video clip in the sequence of stimuli. The video clip displayed Kelly Gordon initiating the interaction. After each video clip, participants selected one of four possible responses. The interaction was designed so that the second through seventh clips followed logically from any of the four participant responses preceding it. Thus, all participants in each group were exposed to the same sequence of clips regardless of their specific responses. This approach was taken for experimental control. Table 2 displays the seven steps of the scenario.

After interacting with Kelly, participants answered questions to determine her fate.

Question 1. When you meet Paul Gordon tomorrow, will you inform him of his exposure to genital herpes?

Question 2. If Paul Gordon has genital herpes, will you inform him that Kelly Gordon is the likely source?

Questions 1 and 2 constitute important decisions because of their potential consequences for both Paul and Kelly, and they shall be the focus of this study. Neither a "Yes" nor a "No" response to these questions is to be deemed the more "ethical" decision, because ethical principles, such as nonmaleficence, autonomy, and informed consent, could be used to justify informing Paul or refraining from doing so. For example, on the basis of nonmaleficence, the doctor could justify informing Paul of his exposure to genital herpes to enable Paul to take action to avoid contracting the disease. By contrast, on the basis of informed consent, the doctor could justify *not* informing Paul because, if this breach of doctor-patient confidentiality were made public, other infected individuals might avoid treatment.

Instruments were also administered to ensure that the different stimulus conditions did not elicit mortality salience and that participants in each condition did not differ in their degree of religious fundamentalism. Both mortality salience and religious fundamentalism have a negative ef-

fect on attitudes toward moral transgressors like Kelly Gordon (Altemeyer & Hunsberger, 2004; MacDorman & Ishiguro, 2006; Rosenblatt et al., 1989).⁴ Lastly, demographic data were collected, which included a self-assessment of colorblindness and visual acuity with correction.

2.4 Variables

The independent variables were the stimulus condition. The first independent variable was human photorealism, which was either a real Human character or a Computer-Generated human character. The second independent variable was motion quality, which was either Fluid (i.e., natural) or Jerky. The dependent variables were the participants' decisions concerning what to tell—or not to tell—Paul.

3 Results

Of the 682 participants, 43.3% chose to inform Paul of his exposure to genital herpes (Question 1), and 25.8% would name Kelly as the likely source (Question 2). In addition, 18.2% of participants strongly favored Paul by answering affirmatively to both questions ("Yes-Yes"), while 49.1% strongly favored Kelly by answering negatively to both questions ("No-No").

Table 3 shows that female participants were consistent in strongly favoring Kelly across all stimulus conditions (51.8% selected "No-No" in the CG Jerky condition, 50.0% in CG Fluid, 52.2% in Human Jerky, and 54.3% in Human Fluid); however, far fewer male participants strongly favored Kelly in the CG Jerky stimulus condition (31.2%) as compared to the CG Fluid (46.3%), Human Jerky (53.0%), and Human Fluid (50.0%) stimulus conditions. CG Jerky was, of course, the condition in which both Kelly's appearance and

4. A one-way ANOVA revealed no significant difference in participants' death-related word completions (the indicator of mortality salience) among the four stimulus conditions, $F(3, 678) = 0.658, p = .578$, or by gender, $F(1, 680) = 0.340, p = .560$. A one-way ANOVA also showed no significant difference in the religious fundamentalism of participants among the four stimulus conditions, $F(3, 678) = 1.922, p = .125$.

Table 2. Scenario Presentation

Video	Text of scenario	Possible responses
1	I've been feeling anxious, especially since my test results were reported to the Department of Health. They contacted me about my partners, so I just told them Paul's stationed overseas. And then I remembered Paul is coming in for a physical tomorrow. I'm trying to cope with so many things right now, and I really don't need one more. I just want to make sure that you won't tell Paul about my condition.	Anyone would be upset in your situation. I'm here to help you. I'm sorry about your test results. Let's work together to find what's best for you both.
2	I appreciate your concern, but what I really need right now is for you to promise me not to tell Paul anything.	Do you know how you contracted genital herpes? Do you think you might have contracted genital herpes from your husband? Are you involved in any high-risk activities? Why are you concerned about Paul finding out?
3	I'm sure Paul's not the source. He has very traditional views on marriage. Even when we were just going out, he wasn't the type to date other girls. Okay, I've had a fling or two, but they were just one-nighters, and I wouldn't know how to contact the guys.	Do you and Paul still have sex? Are you sexually active right now? Are you putting Paul at risk? What steps are you taking to protect Paul?
4	Paul and I are still having sex. But I'm not seeing anyone else right now. Paul knows I have an IUD. If I stopped having sex or asked him to wear a condom, he'd know something's up. He'd soon find out what. I'm sure he couldn't handle it. He'd explode! I'd lose my marriage, the house, everything.	Paul could contract herpes. You're placing Paul at risk. Paul has the right to know. You should tell Paul.
5	You're not in a position to judge me. You really don't know how unstable my situation is right now. I realize he may be at risk, but that's a chance I have to take. I'm just asking you to keep my condition confidential.	Let's tell Paul together. You should tell Paul. I'm concerned about your well-being, but I'm also concerned about Paul's. I respect confidentiality, but even confidentiality has limits.
6	Look, I just can't tell Paul right now, and neither can you. If I didn't think I could trust you, I wouldn't have come to you in the first place.	When you think about this carefully, you'll realize you need to tell Paul. Confidentiality should not be used to put others in harm's way. I also have a duty to protect Paul. Paul trusts me too.
7	Really? Then there's nothing left to discuss. Just give me a couple months to get my life sorted out and figure out how to tell him. In the meantime, I'm counting on you to keep quiet about this. <i>The patient walks out of the office.</i>	No response required

Table 3. The Decision Pattern by Gender and Stimulus Condition

		Questions 1 and 2			
		No–No	Yes–Yes	No–Yes	Yes–No
Overall <i>N</i> = 682	Female	206** 52.02%	54 13.64%	31 7.83%	105 26.52%
	Male	129 45.10%	70 24.48%	21 7.34%	66 23.08%
All CG <i>n</i> = 343	Female	108** 50.94%	25 11.79%	16 7.55%	63 29.72%
	Male	49 37.40%	41 31.30%	13 9.92%	28 21.37%
All Human <i>n</i> = 339	Female	98 53.26%	29 15.76%	15 8.15%	42 22.83%
	Male	80 51.61%	29 18.71%	8 5.16%	38 24.52%
All Jerky <i>n</i> = 362	Female	105* 51.98%	26 12.87%	17 8.42%	54 26.73%
	Male	68 42.50%	41 25.63%	8 5.00%	43 26.88%
All Fluid <i>n</i> = 320	Female	101 52.06%	28 14.43%	14 7.22%	51 26.29%
	Male	61 48.41%	29 23.02%	13 10.32%	23 18.25%
CG Jerky <i>n</i> = 189	Female	58** 51.8%	10 8.9%	10 8.9%	34 30.4%
	Male	24 31.2%	29 37.7%	4 5.2%	20 26.0%
CG Fluid <i>n</i> = 154	Female	50* 50.0%	15 15.0%	6 6.0%	29 29.0%
	Male	25 46.3%	12 22.2%	9 16.7%	8 14.8%
Human Jerky <i>n</i> = 173	Female	47 52.2%	16 17.8%	7 7.8%	20 22.2%
	Male	44 53.0%	12 14.5%	4 4.8%	23 27.7%
Human Fluid <i>n</i> = 166	Female	51 54.3%	13 13.8%	8 8.5%	22 23.4%
	Male	36 50.0%	17 23.6%	4 5.6%	15 20.8%

* $p < .05$. ** $p < .01$. *Note.* Nine chi-square tests were performed to compare the decision pattern for male and female participants. In addition to the overall group, the four main conditions (CG Jerky, CG Fluid, Human Jerky, and Human Fluid) were paired by treatment to create four additional named groups. For example, the All CG group contained the CG Jerky and CG Fluid conditions.

movement were the least human. The differences in male and female decision patterns were statistically significant overall, $\chi^2(3, N = 682) = 13.18$, Cramer's $V = .139$, $p = .004$, and in the All CG, $\chi^2(3, N = 343) = 21.92$, Cramer's $V = .253$, $p = .000$, All Jerky, $\chi^2(3, N = 362) = 11.04$, Cramer's $V = .175$, $p = .012$, CG Jerky, $\chi^2(3, N = 189) = 23.89$, Cramer's $V = .356$, $p = .000$, and CG Fluid, $\chi^2(3, N = 154) = 8.18$, Cramer's $V = .230$, $p = .043$ conditions.

3.1 The Effect of Human Photorealism and Motion Quality on Decision Patterns

Based on the above results, which indicate human photorealism and motion quality influence male participants' decision patterns, a one-way ANOVA was conducted to compare by gender the effect of the stimulus condition on favoring Kelly Gordon. In addition, a two-way ANOVA was used to assess the interaction effect between human photorealism and motion quality. Because the original questions were ordinal, the four possible decision patterns were arranged from weakest to strongest to enable a two-way ANOVA analysis of the character. Possible decision patterns from Questions 1 and 2 were coded in the following order: "Yes-Yes" as 1 (favor Paul), "Yes-No" as 2, "No-Yes" as 3, and "No-No" as 4 (favor Kelly).⁵

Figure 2 shows the mean scores for favoring the character by gender and stimulus condition. A one-way ANOVA test indicated a significant effect for male participants, $F(3, 283) = 4.44$, $\eta^2 = .045$, $p = .005$.⁶ Post

5. The reason a "No-Yes" decision pattern shows greater favoritism to Kelly than a "Yes-No" decision pattern is because, given that Paul has been monogamous as Kelly has stated, he may only need to know the answer to Question 1 (i.e., that he has been exposed to genital herpes) to infer that Kelly is the source. A "No-Yes" decision pattern only requires the participant to warn Paul about Kelly if he tests positive for genital herpes at some time in the future, but Paul will already know Kelly is the source from the results.

6. The four decision patterns were treated as four separate dependent variables (DV1-DV4): (DV1) "Yes-Yes" was coded as 1, and the rest of the patterns were coded as 0; (DV2) "No-No" was coded as 1, and the rest as 0; (DV3) "Yes-No" as 1, and the rest as 0; (DV4) "No-Yes" as 1, and the rest as 0. The intention was to determine the extent of variation across each of the stimulus conditions. Four ANOVAs were run for males, and four for females. Therefore, Table 4 combines the p values from eight ANOVAs.

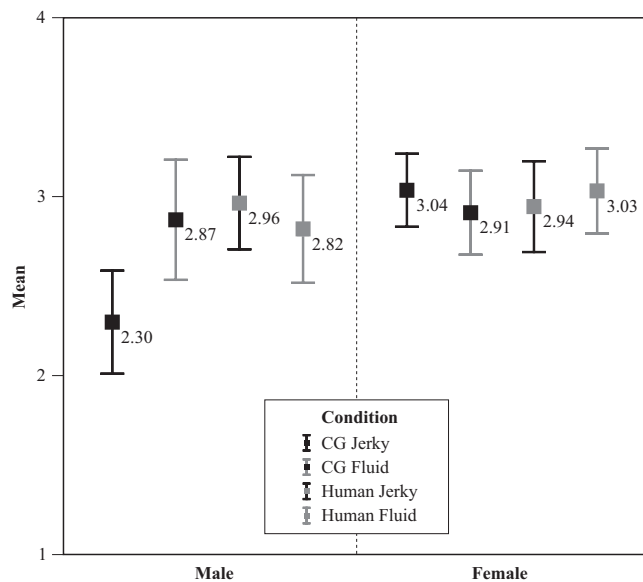


Figure 2. The error bar shows the mean scores for favoring Kelly Gordon by gender and stimulus condition. The scores are derived from the decision patterns for Question 1 and 2. The "Yes-Yes" decision pattern is scored as 1 and signifies strongly favoring Paul Gordon. The "No-No" decision pattern is scored as 4 and signifies strongly favoring Kelly Gordon. ("Yes-No" is scored as 2, and "No-Yes" is scored as 3.) The error bars show 95% confidence intervals.

hoc comparisons using Tukey's test indicated that the mean score for the CG Jerky condition ($M = 2.30$, $SD = 1.27$) differed significantly from the CG Fluid ($M = 2.87$, $SD = 1.23$), Human Jerky ($M = 2.96$, $SD = 1.18$), and Human Fluid ($M = 2.82$, $SD = 1.28$) conditions. However, the CG Fluid, Human Jerky, and Human Fluid conditions did not differ significantly from each other. For female participants there were no significant differences between the CG Jerky, CG Fluid, Human Jerky, and Human Fluid conditions, $F(3, 392) = 0.30$, $\eta^2 = .000$, $p = .826$.

In the male group, the two-way ANOVA results indicate a significant main effect of human photorealism on favoring Kelly, $F(1, 282) = 4.27$, $\eta^2 = .014$, $p = .040$, but motion quality did not reach significance, $F(1, 282) = 2.07$, $\eta^2 = .007$, $p = .15$. In addition, there was a significant interaction effect between human photorealism and motion quality, $F(2, 282) = 5.80$, $\eta^2 =$

.020, $p = .017$. However, in the female group, there were no significant main effects of human photorealism, $F(1, 392) = 0.02$, $\eta^2 = .000$, $p = .896$, or motion quality, $F(1, 392) = 0.03$, $\eta^2 = .000$, $p = .870$, on favoring Kelly, and there was no significant interaction effect either, $F(2, 392) = 0.83$, $\eta^2 = .002$, $p = .362$.

4 Discussion

The results indicated a significant difference in how males and females responded to a female conversational character when presented with differing levels of human photorealism and motion quality. H1 predicted that participants would be more likely to decide in favor of a human female patient than a computer-generated female patient. H2 predicted that participants would be more likely to decide in favor of the patient when presented with fluid movements than with jerky movements. Interestingly, these two hypotheses were supported by the results for males only.

Question 1 asked participants in the role of doctor whether they would inform Paul Gordon of his exposure to genital herpes. A “No” answer is consistent with the wishes of his wife Kelly, who made her case for not telling Paul through the course of the interaction. It could also indicate the participant’s desire to maintain professional objectivity by not getting involved. Question 2 asked participants whether they would inform Paul that Kelly was the likely source of his exposure to genital herpes should he test positive.

The decision pattern of Questions 1 and 2 displayed the greatest difference in responses between male and female participants across the stimuli, with CG Jerky showing the largest difference. In the CG Jerky condition, 37.7% of male participants ruled against the character in both Question 1 and 2 as compared with just 8.9% of females. Males were so affected by the character’s human photorealism that they were more than two-and-a-half times as likely to breach doctor–patient confidentiality to inform Paul about Kelly (“Yes–Yes”) in the CG Jerky (37.7%) condition as in the Human Jerky (14.5%) condition.

Male participants’ sympathy for Kelly appeared to increase significantly with the character’s human likeness. By answering “No” to both Question 1 and 2, participants sympathized more with the character by agreeing not to tell Paul anything about his possible exposure to genital herpes or the disease’s probable source. More than 50% of male participants favored the conversational character by answering “No” to both questions in both the Human Fluid and Human Jerky conditions. However, only 31.2% of male participants favored the character in the CG Jerky condition.

4.1 Relation to Previous Studies

The decision patterns of male and female participants across the stimulus conditions seem contrary to a number of studies on interaction in real and virtual environments. Considerable research on face-to-face interaction has shown that men are less sensitive to nonverbal cues than women (Hall, 1978). The same finding has reappeared in studies on human–machine interaction in recent years. For example, research shows that men are less sensitive to cues like variations in gaze in avatar-mediated conversation (Bente et al., 2007), and men are also less sensitive to the mode of interaction (Guadagno & Cialdini, 2002). Bailenson et al. (2005) found that men respond less strongly than women to the behavioral realism of virtual humans.

Although this study’s results seem contrary to the findings of other studies on gender differences, the study needs to be repeated with additional controls. There was no male counterpart to Kelly Gordon in the study or female counterpart to Paul. If a male character had been used in place of Kelly, the situation may have been reversed. Women might have been more responsive to its degree of realism than men.

A possible explanation for the contrast between male and female participants’ sensitivity to the stimuli is that male participants may have been more sensitive to the photorealism and other aspects of the physical appearance of Kelly Gordon, because she is an attractive member of the opposite sex. As mentioned previously, studies have shown that men are more socially responsive to attractive women (Barocas & Karoly, 1972; Walster,

Aronson, Abrahams, & Rottman, 1966). Furthermore, Feingold (1991) found that men value physical attractiveness more than women.⁷ This particular finding can help to explain the male participants' greater sensitivity to the realism of Kelly Gordon.

The general tendency among female participants to decide in favor of Kelly Gordon agreed with the findings of Guadagno and colleagues (2007), which showed that women were more easily persuaded by avatars of their own gender in immersive virtual environments. This preference, referred to as in-group favoritism, has also been found in male populations, though to a lesser extent (Cameron & Lalonde, 2001; Guadagno et al.). Some studies using the implicit association test have even found that women favor their own gender while men have no gender preference (Nosek & Banaji, 2001; Rudman & Goodwin, 2004). Nevertheless, Zambaka et al. (2006) found that men and women are more easily persuaded by a virtual human of the opposite sex.

We had assumed that male and female participants would be about equally engrossed in the scenario and the conversational character's plea, but male participants may have felt less involved in the scenario than the females. Decisions requiring higher involvement from perceivers can elicit more systematic cognitive processing, which favors the content of an argument over its delivery; decisions requiring lower involvement elicit more heuristic cognitive processing, which can depend more on extraneous factors, such as likeability and perceived physical attractiveness (Chaiken, 1979, 1980). This assessment of attractiveness varies by gender, is often completed in a precognitive split second, and is independent of the subsequent conscious processing of such factors as context (Olson & Marsheutz, 2005;

7. Standards of male and female attractiveness differ in consistent ways across cultures. Female attractiveness is linked to indicators of youth, health, symmetrical features, and features correlated with fertility, such as low waist-to-hip ratio, whereas male attractiveness is also attributed to indicators of ambition, dominance, status, and the ability to provide external resources (Buss, 1989; Cunningham, Roberts, Barbee, Druen, & Wu, 1995; Grammer & Thornhill, 1994; Jones, 1995; Rhodes, Proffitt, Grady, & Sumich, 1998; Scheib, Gangstead, & Thornhill, 1999; Symons, 1995; Trivers, 1972; Townsend & Wasserman, 1998).

Townsend & Wasserman, 1998; Willis & Todorov, 2006). Men and women who are perceived as more attractive are also believed to be more competent (Jackson et al., 1995). If the experimental treatments produced differences in perceived attractiveness, likeability, or competence, members of a group who employed more heuristic-driven cognition would vary more in their decision patterns. Thus, if male participants had a relatively low level of involvement in the interaction, this could have caused them to show greater variability across the different stimuli.

4.2 Limitations and Future Work

The experimental design and crafting of the ethical scenario were some of the strong points of this study's execution. The scenario was a true ethical dilemma that split the participants' decisions overall. However, because the scenario was not gender neutral, the preference that female participants showed for Kelly Gordon as compared to male participants should probably be attributed to the scenario rather than the stimuli. In addition, some aspects of the stimuli limited the possible conclusions of the study. The character's blouse exposed part of her midriff in the CG conditions but not in the Human conditions. The additional display of skin could have been perceived negatively (or positively) by some participants (see Grammer, Renninger, & Fischer, 2004). In particular, participants may have perceived the CG character as being more promiscuous than the human character.

The fact that Kelly Gordon was a member of a racial minority could limit the generalizability of the results to other characters and contexts. Because both the CG character and human actress looked East Asian, the results could reflect stereotypes about East Asian women rather than women in general. Because 96.6% of participants were not Asian, even female participants might tend to perceive Kelly Gordon as an out-group. If Kelly Gordon were played by an actress who was the same race as the majority of the participants, the Human condition would more clearly represent an in-group for female participants. If the Human character were then perceived as an in-group and the CG character were

perceived as an out-group, this could lead participants to favor Kelly Gordon more strongly in the Human conditions as compared to the CG conditions.

The results of the study warrant further exploration, which could focus on potential differences in gender perception. For example, the use of a male character with a female spouse in the same scenario could be used to determine whether presentational factors also had a significant effect on the decisions of women with respect to the opposite sex. Future work could try to isolate other attributes that may contribute to or detract from how the character is perceived, especially along gender lines. These could include varying the character's degree of human likeness to explore the possibility of a threshold for perceiving a human as acceptable. It would also be useful to confirm the reliability of participants' perception of realism across different CG characters. By repeating this study with participants who have already received training in medical ethics (e.g., medical students) and comparing the results, it would be possible to determine whether their training lessened the effect of the stimulus conditions, for example, by encouraging participants to reason from the facts and from ethical principles. Additional information could be gathered from participants to shed light on their exposure to animated human characters in various contexts (e.g., in computer games and films). Likewise, a rating of attractiveness, likeability, and competence of the character used in the scenario would allow us to better understand the factors influencing participants' decisions.

5 Conclusion

This study has shown that aspects of a character's presentation can influence decisions of moral or ethical consequence in gender-specific ways. Although women favored Kelly Gordon more than men overall, only men were significantly affected by presentational factors, reacting negatively to Kelly in the CG Jerky condition.

The fact that males and females react differently to changes in a character's visual presentation has HCI design implications, which could impact future systems

created to facilitate interactions in medical and other scenarios. For example, systems designed to facilitate doctor–patient consultation, or to recreate events at the scene of a crime or automobile collision for courtroom judges, could have substantial effects on human lives because of small presentational factors (Arndt et al., 2009). Medical and legal professionals are tasked with making objective moral and ethical decisions based on their knowledge and interpretation of domain-specific concepts and principles. If extraneous presentational factors prove capable of subconscious influence, careful consideration of character presentation should be a key component of a system's design.

Acknowledgments

Frank Tai at Pixar deserves much appreciation for animating the computer-generated human character and for directing, filming, green screening, and editing the video of the human actress. Thanks go to Colin Allen, Jamie Arndt, Adam J. Burton, Hsin-Liang Chen, Stephen J. Cowley, Robert D. Green, Jacob Faiola, Hadi Kharrazi, Wade Mitchell, Matthias Scheutz, and Jonathan Weinberg for fruitful discussions on machine ethics and critical feedback on an earlier version of this paper. We also greatly appreciate comments from Jeremy Bailenson and two anonymous reviewers. We would like to thank Jesse Lin, the actress who played Kelly Gordon, and Jennifer K. Stewart, the character's voice actress. This research was supported by a Signature Center grant and a Research Investment Fund appropriation from Indiana University–Purdue University Indianapolis.

References

- Altemeyer, B., & Hunsberger, B. (2004). A revised religious fundamentalism scale: The short and sweet of it. *International Journal for the Psychology of Religion, 14*(1), 47–54.
- Arndt, J., Vess, M., Cox, C. R., Goldenberg, J. L., & Lagle, S. (2009). The psychosocial effect of thoughts of personal mortality on cardiac risk assessment. *Medical Decision Making, 29*(2), 175–181.
- Bailenson, J. N., Swinth, K. R., Hoyt, C. L., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and

- interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of co-presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 14(4), 379–393.
- Barocas, R., & Karoly, P. (1972). Effects of physical appearance on social responsiveness. *Psychological Reports*, 31(2), 495–500.
- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2(3), 412–414.
- Beardsley, M. C. (1960). *The European philosophers from Descartes to Nietzsche*. New York: Modern Library.
- Beauchamp, T. L., & Childress, J. F. (1979). *Principles of biomedical ethics*. New York: Oxford University Press.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making, and the orbitofrontal cortex. *Cerebral Cortex*, 10(3), 295–307.
- Bente, G., Eschenburg, F., & Aelker, L. (2007). Effects of simulated gaze on social presence, person perception and personality attribution in avatar-mediated communication. *Proceedings of Presence 2007*, 207–214.
- Blascovich, J. (2002). A theoretical model of social influence for increasing the utility of collaborative virtual environments. *Proceedings of the 4th International Conference on Collaborative Virtual Environments*, 25–30.
- Blascovich, J., Loomis, J., Beall, A., Swinith, K., Hoyt, C., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13, 103–124.
- Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, 12(1), 1–49.
- Cameron, J. E., & Lalonde, R. N. (2001). Social identification and gender related ideology in women and men. *British Journal of Social Psychology*, 40, 59–77.
- Chaiken, S. (1979). Communicator physical attractiveness and persuasion. *Journal of Personality and Social Psychology*, 37(8), 1387–1397.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766.
- Collinson, D. (1987). *Fifty major philosophers: A reference guide*. New York: Croom Helm.
- Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. *Journal of Personality and Social Psychology*, 50(5), 925–935.
- Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C.-H. (1995). Their ideas of beauty are on the whole the same as ours? Consistency and variability in cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology*, 68(2), 261–279.
- Darley, J. M., & Batson, C. D. (1973). From Jerusalem to Jericho: A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27(1), 100–108.
- Dehn, D. M., & Van Mulken, S. (2000). The impact of animated interface agents: A review of empirical research. *International Journal of Human-Computer Studies*, 52(1), 1–22.
- Doris, J. M. (1998). Persons, situations, and virtue ethics. *Nous*, 32(4), 504–530.
- Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. New York: Cambridge University Press.
- Doris, J., & Stich, S. (2005). As a matter of fact: Empirical perspectives on ethics. In F. Jackson & M. Smith (Eds.), *The Oxford handbook of contemporary analytic philosophy* (pp. 114–152). Oxford, UK: Oxford University Press.
- Efran, M. G. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *Journal of Research in Personality*, 8(1), 45–54.
- Feingold, A. (1991). Sex differences in the effects of similarity and physical attractiveness on opposite-sex attraction. *Basic and Applied Social Psychology*, 12(3), 357–367.
- Fleetwood, J., Vaught, W., Feldman, D., Gracely, E., Kasutto, Z., & Novack, D. (2000). MedEthEx Online: A computer-based learning program in medical ethics and communication skills. *Teaching and Learning in Medicine*, 12(2), 96–104.
- Fleming, D. (2006). The character of virtue: Answering the situationist challenge to virtue ethics. *Ratio*, 19(1), 24–42.
- Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *NeuroImage*, 35(4), 1674–1684.
- Geller, T. (2008). Overcoming the uncanny valley. *IEEE Computer Graphics and Applications*, 28(4), 11–17.
- Glover, S. H., Bumpus, M. A., Sharp, G. F., & Munchus, G. A. (2002). Gender differences in ethical decision making. *Women in Management Review*, 17(5), 217–227.

- Grammer, K., & Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology*, *108*(3), 233–243.
- Grammer, K., Renninger, L. A., & Fischer, B. (2004). Disco clothing, female sexual motivation, and relationship status: Is she dressed to impress? *Journal of Sex Research*, *41*(1), 66–74.
- Green, R. D., MacDorman, K. F., Ho, C.-C., & Vasudevan, S. K. (2008). Sensitivity to the proportions of faces that vary in human likeness. *Computers in Human Behavior*, *24*(5), 2456–2474.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.
- Guadagno, R. E., & Cialdini, R. B. (2002). Online persuasion: An examination of gender differences in computer-mediated interpersonal influence. *Group Dynamics: Theory Research and Practice*, *6*, 38–51.
- Guadagno, R. E., Blascovich, J., Bailenson, J. N., & McCall, C. (2007). Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, *10*, 1–22.
- Gueguen, N., & De Gail, M. A. (2003). The effect of smiling on helping behavior: Smiling and good Samaritan behavior. *Communication Reports*, *16*, 133–140.
- Hadjistavropoulos, H. D., Ross, M. A., & Von Baeyer, C. L. (1990). Are physicians' ratings of pain affected by patients' physical attractiveness? *Social Science & Medicine*, *31*(1), 69–72.
- Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, *26*(3), 245–256.
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, *85*, 845–857.
- Hanson, D. (2006). Exploring the aesthetic range of humanoid robots [electronic resource]. In *Toward Social Mechanisms of Android Science, the 28th Annual Conference of the Cognitive Science Society, Long Seminar*, 16–20.
- Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society*, 315–331.
- Ho, C.-C., MacDorman, K., & Pramono, Z. A. D. (2008). Human emotion and the uncanny valley: A GLM, MDS, and ISOMAP analysis of robot video ratings. *Proceedings of the Third ACM/IEEE International Conference on Human-Robot Interaction*, 169–176.
- Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 203–253). San Diego: Academic Press.
- Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology*, *21*(3), 384–388.
- Jackson, L. A., Hunter, J. E., & Hodge, C. N. (1995). Physical attractiveness and intellectual competence: A meta-analytic review. *Social Psychology Quarterly*, *58*(2), 108–122.
- Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin*, *126*(5), 703–726.
- Jones, D. (1995). Sexual selection, physical attractiveness, and facial neoteny: Cross-cultural evidence and implications. *Current Anthropology*, *36*(5), 723–748.
- Kniffin, K. M., & Wilson, D. S. (2004). The effect of non-physical traits on the perception of physical attractiveness: Three naturalistic studies. *Evolution and Human Behavior*, *25*, 88–101.
- Kock, N. (2005). Media richness or media naturalness? The evolution of our biological communication apparatus and its influence on our behavior toward e-communication tools. *IEEE Transactions on Professional Communication*, *48*(2), 117–130.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*, 908–911.
- Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., & Kircher, T. (2008). Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLoS ONE*, *3*(7), e2597.
- Luo, Q., Nakic, M., Wheatley, T., Richell, R., Martin, A., & Blair, R. J. R. (2006). The neural basis of implicit moral attitude: An IAT study using event-related fMRI. *Neuro Image*, *30*, 1449–1457.
- MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in social and cognitive science research. *Interaction Studies*, *7*(3), 297–337.
- MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. T. (2009). Too real for comfort: Uncanny responses to computer generated faces. *Computers in Human Behavior*, *25*(3), 695–710.

- MacDorman, K. F., Minato, T., Shimada, M., Itakura, S., Cowley, S. J., & Ishiguro, H. (2005). Assessing human likeness by eye contact in an android testbed. *Proceedings of the 27th Annual Meeting of the Cognitive Science Society* (41–48). Hillsdale, NY: Lawrence Erlbaum Associates.
- Matthews, K. E., & Cannon, L. K. (1975). Environmental noise level as a determinant of helping behavior. *Journal of Personality and Social Psychology*, *32*(4), 571–577.
- McKelvie, S. J., & Coley, J. (1993). Effects of crime seriousness and offender facial attractiveness on recommended treatment. *Social Behavior and Personality*, *21*(4), 265–277.
- Mori, M. (1970). Bukimi no tani [the uncanny valley]. *Energy*, *7*, 33–35.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Social Issues*, *56*, 81–103.
- Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, *84*(2), 221–236.
- Nichols, S., & Mallon, R. (2006). Moral dilemmas and moral rules. *Cognition*, *100*(3), 530–542.
- Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*(6), 161–176.
- Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism of users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *12*(5), 481–494.
- Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion*, *5*(4), 498–502.
- Peternelj-Taylor, C. A. (1989). The effects of patient weight and sex on nurses' perceptions: A proposed model of nurse withdrawal. *Journal of Advanced Nursing*, *14*(9), 744–754.
- Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, *9*, 29–43.
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press (CSLI).
- Rhodes, G., Proffitt, F., Grady, J. M., & Sumich, A. (1998). Facial symmetry and the perception of beauty. *Psychonomic Bulletin & Review*, *5*(4), 659–669.
- Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T., & Lyon, D. (1989). Evidence for terror management theory: I. The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, *57*(4), 681–690.
- Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic in-group bias: Why do women like women more than men like men? *Journal of Personality and Social Psychology*, *87*, 494–509.
- Scheib, J. E., Gangstead, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London. Biological Sciences*, *266*(1431), 1913–1917.
- Schnall, S., Haidt, J., Clore, G., & Jordan, A. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, *34*, 1096–1109.
- Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: The effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments*, *16*(4), 337–351.
- Shechtman, N., & Horowitz, L. M. (2003). Media inequality in conversation: How people behave differently when interacting with computers and people. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 281–288.
- Sigall, H., & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology*, *31*(3), 410–414.
- Symons, D. (1995). Beauty is in the adaptations of the beholder: The evolutionary psychology of human female sexual attractiveness. In P. R. Abramson & S. D. Pinkerton (Eds.), *Sexual nature, sexual culture* (pp. 80–118). Chicago: University of Chicago Press.
- Tai, Y. F., Scherfler, C., Brooks, D. J., Sawamoto, N., & Castiello, U. (2004). The human premotor cortex is “mirror” only for biological actions. *Current Biology*, *14*(2), 117–120.
- Takeuchi, A., & Naito, T. (1995). Situated facial displays: Towards social interaction. *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, 450–455.
- Townsend, J. M., & Wasserman, T. (1998). Sexual attractiveness: Sex differences in assessment and criteria. *Evolution and Human Behavior*, *19*(3), 171–191.
- Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man* (pp. 136–179). Chicago: Aldine Press.
- Vertegaal, R., & Ding, Y. (2002). Explaining effects of eye gaze on mediated group conversations Amount or synchronization? *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work* (41–48). New York: ACM.

- Walster, E., Aronson, V., Abrahams, D., & Rottman, L. (1966). Importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology*, 4(5), 508–516.
- Wark, G. R., & Krebs, D. L. (1996). Gender and dilemma differences in real-life moral judgment. *Developmental Psychology*, 32(2), 220–230.
- Webber, J. (2006). Virtue, character, and situation. *Journal of Moral Philosophy*, 3(2), 193–213.
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17(7), 592–598.
- Yee, N., Bailenson, J. N., & Rickertsen, K. (2007). A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1–10.
- Zanbaka, C. A., Goolkasian, P., & Hodges, L. F. (2006). Can a virtual cat persuade you? The role of gender and realism in speaker persuasiveness. *SIGCHI Conference on Human Factors in Computing Systems*, 1153–1162.
- Zanbaka, C. A., Ulinski, A. C., Goolkasian, P., & Hodges, L. F. (2007). Social responses to virtual humans: Implications for future interaction design. *SIGCHI Conference on Human Factors in Computing Systems*, 1561–1570.