

Exploring the Effects of Dynamic Avatars on Performance and Engagement in Educational Games

Dominic Kao, Massachusetts Institute of Technology
D. Fox Harrell, Massachusetts Institute of Technology

Abstract: Avatar research has almost exclusively explored avatars that remain the same regardless of context. However, there may be advantages to avatars that change during use. A plethora of work has shown that avatars personalized in one's likeness increases identification, while object-like avatars increase detachment. We posit that in certain situations within a game it may be more advantageous to have increased identification, while in other situations increased detachment. We present a study on *dynamic avatars*, or avatars that change types based on game context. In particular, we investigate what we term the *successful likeness* avatar. The *successful likeness* is an avatar that is only a likeness when the player is in a win state and at all other times an object. Our goal is to determine if this type of avatar can foster an increase in user performance and engagement. Our experiment (N=997) compares four avatars: 1) Shape, 2) Likeness, 3) Likeness to Shape, and 4) Shape to Likeness (*successful likeness*). We found that players using a *successful likeness* avatar had significantly better performance (levels completed) than all other conditions. Players using a *successful likeness* avatar had significantly higher play time (minutes played) than all other conditions. We propose a theoretical model in which identification facilitates vicarious outcomes and in which detachment facilitates outcome dissociation. As performance and engagement are correlated to learning (Harteveld, 2015), *successful likeness* avatars may be crucial in educational games.

Introduction

Over twenty years of research on virtual agents and avatars has revealed that our understanding of how their appearance affects us is still limited. The *persona effect* was one of the earliest studies that revealed that the mere presence of a life-like character in a learning environment increased positive attitudes (Lester et. al, 1997). A wealth of empirical research since then has demonstrated that virtual characters are more influential when they have similar competencies (Kim & Baylor, 2006), a similar gender (Baylor & Kim, 2004; Guadagno et. al, 2007), and a similar ethnicity/race (Pratt et. al, 2007; Rosenberg-Kima et. al, 2010). However, research on the visual form and look of animated agents is sparse; it has been proposed that the following five dimensions are understudied: 1) the degree of "humanoidness," 2) the degree of stableness versus changeability of appearance (morphing), 3) the degree of animation, 4) the degree of 3-dimensionality, and 5) the degree of realism (Gulz & Haake, 2006). The reason for the sparsity of research in this area has been attributed to two possible causes: 1) these questions are difficult to answer using existing methodologies, and 2) people do not readily accept the idea that appearance affects our intellectual processes (e.g., "Don't judge a book by its cover") (Gulz & Haake, 2006). In this work, we propose to explore the *changeability of appearance* aspect of avatars.

In particular, our work is based on increasing evidence that demonstrates that abstract avatars increase player-avatar detachment via low identification (my avatar is *not* me), high sense of control (my avatar is like a tool), low sense of responsibility (my avatar has no needs), etc. (Banks & Bowman, 2013; Bowman, Rogers, & Sherrick, 2013; Kao & Harrell, 2015c). Work in human-computer interaction, psychology, and marketing suggests that within virtual environments, success and failure is attributed to avatars and through them also affects users (Campbell & Sedikides, 1999; Moon, 2000; Wolfendale, 2007; Whang & Chang, 2004; Huff, Johnson, & Miller, 2003). These effects are more powerful with avatars with whom we identify (Vasalou, Joinson, & Pitt, 2007; Duval & Silvia, 2002). Here, we perform the first study to our knowledge on *dynamic avatars*. More specifically, we study what we call the *successful likeness* avatar, an avatar that is normally abstract (e.g., a shape), but that becomes a likeness in win states. Our goal is determining if selectively increasing and decreasing user identification with the avatar during key moments of the game experience can result in increased performance and engagement. We found that participants did not significantly differ in reported engagement between conditions. However, participants using a *successful likeness* avatar both completed significantly more levels, and played the game for a significantly longer period of time,

suggesting greater performance and engagement (see Bauckhage & Kersting (2012) for predicting engagement via play time). Since both performance and engagement have been correlated to better learning outcomes in educational games (Harteveld, 2015; Blumenfeld, Kempler, & Krajcik, 2005), our work has important implications for avatar design in educational environments.

Motivation

The work here is based on the premise that, along with factors such as subject mastery and affect toward the subject, a sense of identity as a STEM learner and doer is necessary for developing literacy and agency in computing (S. Veeragoudar Harrell & D.F. Harrell, 2009). The standard paradigm of computer science education research traditionally focuses almost exclusively on cognitive challenges apparently inherent to particular computational concepts, e.g., (Ben-Ari, 2001). Veeragoudar Harrell states that developments in the learning sciences suggest that computer science curricula should embrace a broader conceptualization of learning: human reasoning, it is proposed, is embodied, distributed, and situated, and learning must be accordingly perceived as inherently collaborative, contextualized, and instrumented (Dourish, 2001; Greeno, Collins, & Resnick, 1996; Hutchins, 2000; Lave, & Wenger 1991). A result of this broader view of human reasoning and learning in the STEM disciplines is the emergence of research on relations between student identity and learning (Gresalfi, Martin, Hand, & Greeno, 2009; Lave, & Wenger, 1991; Nasir, 2002). For example, stereotype threat is the phenomenon in which triggering awareness of a learner's identity results in his or her performance conforming to social stereotypes regarding that identity (Steele, & Aronson, 1995; Shih, Pittinsky, & Ambady, 1999; Gibson, Losee, & Vitiello, 2014; Good, Rattan, & Dweck, 2012). Digital manifestations of such phenomena are important areas for investigation since virtual identities are now frequently used as avatars in videogames, avatars in MOOCs and forums, intelligent tutors, and more.

The Game

The experiment takes place in a STEM learning game called *Mazzy* (Kao & Harrell, 2015e)¹. *Mazzy* is a game in which players solve mazes by creating short computer programs. In total, there are 12 levels in this version of *Mazzy*. Levels 1-5 require only basic commands. Levels 6-9 require using loops. Levels 10-12 require using all preceding commands in addition to conditionals. See Figures 1 and 2. *Mazzy* has been used previously as an experimental testbed for evaluating the impacts of avatar type on performance and engagement in an educational game (Kao & Harrell, 2015a-d; Kao & Harrell, 2016a-c).

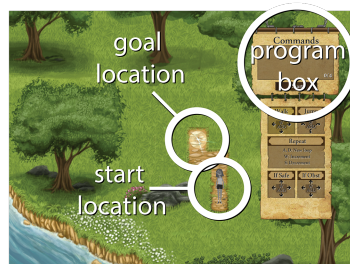


Figure 1. Level 1 in *Mazzy* introduces the basic game mechanics.

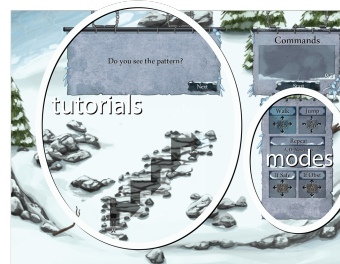


Figure 2. Level 6 introduces looping.

Theoretical Framework

Our work is based on research on avatars, agents, and “blended identities” (Harrell, 2010). Although in this work we are studying avatars, an abundance of work on agents (i.e., virtual pedagogical agents, teaching agents, etc.) helps to guide our study. In particular, a large body of work has shown that avatars and agents that share users’ external characteristics (e.g., age, gender, race, clothing, etc.) are more influential and are linked to better learning outcomes (Kim & Baylor, 2006; Baylor & Kim, 2004; Guadagno et. al, 2007; Pratt et. al, 2007; Rosenberg-Kima et. al, 2010; Johnson, Didonato, & Reisslein, 2013; Arroyo, Woolf, Royer, & Tai, 2009; Bailenson, Blascovich, & Guadagno, 2008). This is posited to be a result of similarity-attraction, the theory that people are attracted to similar others (Byrne, & Nelson, 1965; Isbister, & Nass, 2000). Functional neuroimaging has found that perceived similarity is an important factor in a person’s ability to simulate the internal state of

¹ Gameplay video: <http://tinyurl.com/mazzyquick>

another person (Mitchell, Macrae, & Banaji, 2006). Likewise, Mobbs et. al found that when a participant watched a game show contestant with high perceived similarity, the participant experienced significant increases in both subjective and neural responses to vicarious reward (2009). Other work suggests that what is experienced by an avatar is also experienced by its user (Campbell & Sedikides, 1999; Moon, 2000; Wolfendale, 2007; Whang & Chang, 2004; Huff, Johnson, & Miller, 2003). This effect is more powerful via avatars that we identify with (Vasalou, Joinson, & Pitt, 2007; Duval & Silvia, 2002), identification being positively correlated to such factors as representation of emotions and intent (Hamilton, 2009), physical resemblance (Maccoby, & Wilson, 1957), and avatar customization (Turkay, 2014).

In the past decade, it has become apparent that avatars play an important role in affecting our behaviors. The Proteus effect describes an individual's tendency to conform to behavior typically associated with how an avatar appears (Yee, & Bailenson, 2007). For example, two of the earliest studies found that participants with taller avatars were more aggressive, and that participants with more attractive avatars were more confident. Since avatars affect us in a subtle way, they are a form of "embedded content" (Kaufman, Flanagan, & Seidman, 2015), which studies have shown is more effective than "message-driven" agendas (Brehm, 1966). Avatars, or "blended identities," (Harrell, 2010), can be pivotal in enabling our capacities to put ourselves inside other identities. However, the unfortunate consequence is avatars can also be used to reinforce stereotypes, perpetuate hegemonic views, etc., e.g., women as victims of violence, etc. Fortunately, some representations can begin to combat these stereotypes, e.g., playing a computer science learning game as Marie Curie (Kao & Harrell, 2016a). For instance, research has shown that abstract (or object-like) representations, such as a geometric shape, lead to detachment with the avatar and outcomes associated to the avatar (Banks, & Bowman, 2013; Bowman, Rogers, & Sherrick, 2013; S. Veeragoudar Harrell, & D.F. Harrell 2009; Kao & Harrell, 2015c). Because of the potential usefulness in exploring the dichotomy between identification and detachment, we investigate the *successful likeness*. This avatar is abstract (shape) when the player is not in a win state (to facilitate detachment), and a likeness (Mii) when the player is in a win state (to facilitate identification). Our goal is to test if this type of avatar can enhance player performance and engagement.

Experiment

Our experiment consisted of a between-subjects design. Our goal was to measure performance and engagement across conditions.

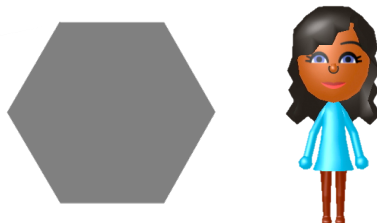


Figure 1: A sample Shape avatar (left); A sample Likeness avatar (right).

Conditions: Our four avatar conditions were:

- 1) Shape
- 2) Likeness
- 3) Likeness to Shape
- 4) Shape to Likeness

Participants were all told that they would be playing a game. No other details were specified. Players were asked to use a publicly available customization system to create a Mii (the *Likeness*). A Mii is a type of avatar developed by Nintendo, chosen since Miis were designed with the intention that most users would create likeness avatars (the word "Mii" is a blend of "Wii" and "me"). Furthermore, players were told to create an avatar that looked like themselves. Players then picked out of eight possible geometric shapes (the *Shape*). Every player created a *Likeness* avatar and selected a *Shape* avatar (see Figure 1). If a participant was assigned to Condition 1, their avatar was always a shape. In Condition 2, their avatar was always a Mii. In Condition 3, their avatar was normally a Mii, but when a level was successfully won, the avatar became a shape. In Condition 4, their avatar was normally a

shape, but when a level was successfully won, the avatar became a Mii (*successful likeness*). The 'winning' avatar (a shape in Conditions 1 & 3, and a Mii in Conditions 2 & 4) was displayed centered in the middle of the screen. All other aspects of the experiment were identical across conditions.

Measures: Our performance measures consist of levels completed and time played, while our engagement measure is the Game Experience Questionnaire (GEQ) (IJsselsteijn, De Kort, Poels, Jurgelionis, & Bellotti, 2007).

Participants: 997 participants were recruited through Mechanical Turk. The data set consisted of 560 male, and 437 female participants. Participants self-identified their races/ethnicities as white (665), Asian Indian (163), black or African American (55), American Indian (14), Chinese (13), Filipino (13), Korean (10), Japanese (6), Vietnamese (4) and other (54). Participants were between the ages of 18-72 ($M = 30.1$, $SD = 8.2$). Participants were reimbursed \$1.50 to participate in this experiment.

Design: Our design was a between-subjects design: avatar condition was the between-subjects factor. Participants were randomly assigned to a condition.

Protocol: Prior to starting the game, players were informed that they could exit the game at any time via a red button in the corner of the screen. When participants were done playing (either by exiting early, or by finishing all 12 levels), participants returned to the experiment instructions, which prompted them with demographics.

Analysis: Data was analyzed in SPSS using analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA). We ran one ANOVA using *levels completed* as the dependent variable, and one ANOVA using *time played* as the dependent variable. Our MANOVA used *GEQ items* as the dependent variable. In all cases, our independent variable is *avatar condition*. To be aligned with our research question, we asked participants after the experiment to rate how similar they felt their Mii was to themselves (1: *Very Dissimilar* to 5: *Very Similar*). We removed participants that reported a similarity less than 3 (189). Additionally, we removed 35 outliers according to the criteria in Hoaglin (1987). These 224 participants were excluded from further analysis. Prior to running our MANOVA model, we checked the assumption of homogeneity of variance by Levene's Test of Equality of Error Variances, and the assumption was met by the data ($p > .05$). All reported p-values are two-tailed.

Results: An ANOVA revealed that *levels completed* was significantly different across avatar conditions, $F(3, 769) = 3.02$, $p < .05$. Post-hoc comparisons (LSD) revealed that the condition **Shape to Likeness** significantly outperformed *Likeness*, $p = .017$. The condition **Shape to Likeness** also significantly outperformed *Likeness to Shape*, $p = .007$ (see Figure 2).

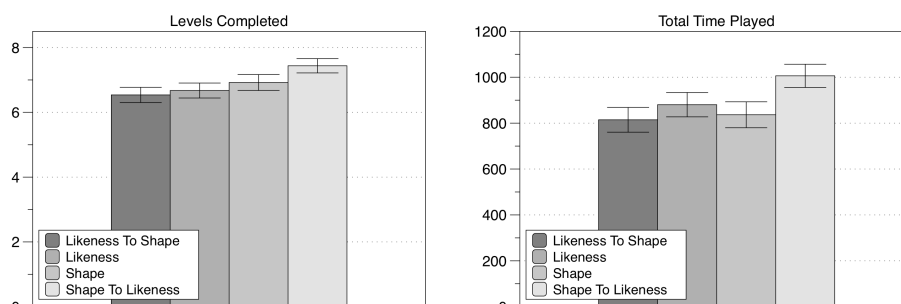


Figure 2: Levels Completed (left); Total Time Played (right).

Similarly, an ANOVA revealed that *time played* (seconds) was significantly different across avatar conditions, $F(3, 769) = 2.69$, $p < .05$. Post-hoc comparisons (LSD) revealed that the condition **Shape to Likeness** had significantly longer play time than *Shape*, $p = .019$. The condition **Shape to Likeness** also had significantly longer play time than *Likeness to Shape*, $p = .010$. The condition **Shape to Likeness** had marginally longer play time than *Likeness*, $p = .072$ (see Figure 2).

A MANOVA revealed that there was no statistically significant difference in GEQ responses across avatar conditions, $F(126, 2190) = 1.02$, $p = .43$; Pillai's Trace = .17, partial $\eta^2 = .055$. See Figure 3.

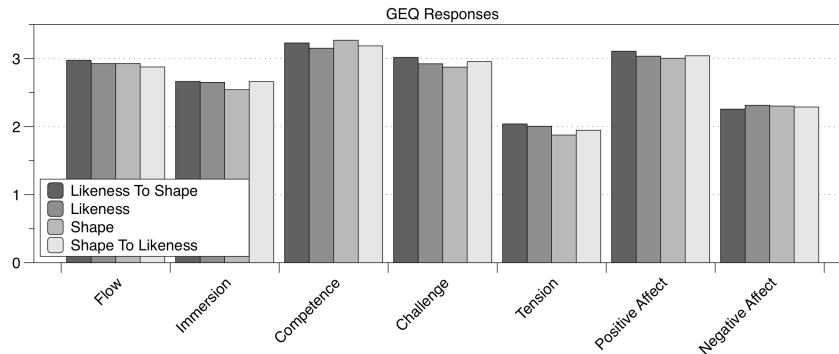


Figure 3: Game Experience Questionnaire (GEQ) responses for all participants.

None of the participants correctly guessed the purpose of the experiment.

Discussion

We first summarize our findings:

- Shape to Likeness participants were **highest performing**.
- Shape to Likeness participants spent the **most time in game**.
- **No significant differences** in GEQ responses.

We found that the *Shape to Likeness (successful likeness)* condition had significantly increased **Levels Completed** and **Time Played**. GEQ responses did not significantly differ. These results support our initial hypothesis that having a shape avatar (greater detachment) when the player is not in a win state, and having a likeness avatar (increased identification) when the player is in a win state, would outperform other avatar types. The worst performing condition was the inverse condition: *Likeness to Shape*.

What do these results mean? For example, the *successful likeness* condition participants on average completed about 1 more level and played for about 3.2 minutes longer than *Likeness to Shape* condition participants. Longer game playtime can be used as a measure of engagement (Bauckhage & Kersting, 2012). Moreover, both increased game performance and engagement have been correlated to better learning outcomes in educational games (Harteveld, 2015; Blumenfeld, Kempfer, & Krajcik, 2005). Therefore, **these results are suggestive that, over longer periods of time, dynamic avatars could have beneficial effects on players.**

Why did this happen? Multiple disciplines have independently demonstrated that avatars with a higher degree of perceived similarity may better facilitate vicarious experiences, positive or negative (Campbell & Sedikides, 1999; Moon, 2000; Wolfendale, 2007; Whang & Chang, 2004; Huff, Johnson, & Miller, 2003; Vasalou, Joinson, & Pitt, 2007; Duval & Silvia, 2002; Kim & Baylor, 2006; Baylor & Kim, 2004; Guadagno et. al, 2007; Pratt et. al, 2007; Rosenberg-Kima et. al, 2010; Johnson, Didonato, & Reisslein, 2013; Arroyo, Woolf, Royer, & Tai, 2009; Bailenson, Blasovich, & Guadagno, 2008; Byrne, & Nelson, 1965; Isbister, & Nass, 2000). Moreover, neural imaging has demonstrated that watching a similar person experience a reward also increases our own vicarious reward (Mobbs et. al, 2009). Effects of an avatar similar to oneself may persist even after the experiment. Fox & Bailenson found that watching one's avatar exercising resulted in significantly more exercise on the part of the participant, 24 hours later, as compared to participants that watched one's avatar loitering or a virtual other exercising (2009). Lastly, abstract (or object-like) avatars can better facilitate detachment and may play a role in helping users dissociate from failure outcomes, such as in cases requiring "debugging" (Banks, & Bowman, 2013; Bowman, Rogers, & Sherrick, 2013; S. Veeragoudar Harrell, & D.F. Harrell 2009; Kao & Harrell, 2015c).

How generalizable are these results? The work here was a single study of how dynamic avatars affected engagement and performance for 997 participants in a coding game. While we feel the results are well supported by the literature, there should be additional investigation of the specific physiological effects of dynamic avatars. While one possible approach is to ask participants questions from, e.g., the Player-Avatar Interaction (PAX) questionnaire (Banks, & Bowman, 2016), we feel that post-game surveys will be a difficult approach given the rather subtle differences in the experience

between conditions. While these subtle differences manifested as tangible differences in performance, they did not manifest in tangible differences in reported engagement. Even if players differed on some self-report (e.g., “This avatar understands me.”), it’s not readily apparent how we can disambiguate, in the dynamic avatar case, between the non-win state avatar, the win state avatar, or some combination. Because these avatars are different than any avatar previously studied, we will need new methods to study them. We find some parallels in work on multiple agents; for instance, it has been found that multiple virtual pedagogical agents with ‘compartmentalized’ roles (e.g., one agent provides confidence-boosting messages, another provides information support, etc.) provide significantly better learning outcomes than a single agent (Baylor, & Ebbers, 2003; Baylor, & Kim, 2005; Odell, Parunak, & Fleischer, 2003). Here, we instead have multiple avatars, and we are facilitating either greater identification or greater detachment depending on the game context. We plan to further investigate this phenomenon in the near future. We are partnered with a non-profit and will be studying Computer Science learning using these avatars in Cambridge schools. We aim to use EEG devices, e.g., the EPOC+, to measure brain activity of participants over the course of game play. Ultimately, such an approach would help us determine the specific physiological influences of these *dynamic avatars*.

Conclusion

Increasingly, there has been research on the different external characteristics of avatars and agents and how they affect users in educational environments (Lester et. al, 1997; Kim & Baylor, 2006; Baylor & Kim, 2004; Guadagno et. al, 2007; Pratt et. al, 2007; Rosenberg-Kima et. al, 2010). However, their visual form and look has been understudied, including avatars that change from one form to another (morphing) (Gulz & Haake, 2006). Here, we provide the first study to our knowledge on *dynamic avatars*, or avatars that are different depending on whether the user is in a win state or not. We found that the dynamic avatar, *successful likeness*, outperformed all other conditions in terms of levels completed. These same participants also played the game significantly longer. We posit that this is a result of shapes (abstract) as avatars leading to more detachment and Mii avatars (likeness) leading to more vicarious experience. Educational systems and games could benefit greatly from such a model of representation, shielding users from internalizing failure, and basking them in self-success-identification.

References

- Arroyo, I., Woolf, B. P., Royer, J. M., & Tai, M. (2009). Affective gendered learning companions.
- Bailenson, J. N., Blascovich, J., & Guadagno, R. E. (2008). Self-rep. in immersive virt. environments.
- Banks, J., & Bowman, N. D. (2013). Close intimate playthings? Understanding player-avatar relationships as a function of attachment, agency, and intimacy.
- Banks, J., & Bowman, N. D. (2016). Emotion, anthropomorphism, realism, control: Validation of PAX.
- Bauchhage, C., & Kersting, K. (2012). How players lose interest in playing a game. *CIG*, 139–146.
- Baylor, A. L., & Ebbers, S. J. (2003). Evidence that Multiple Agents Facilitate Greater Learning.
- Baylor, A., & Kim, Y. (2004). Pedagogical agent design. *Intelligent Tutoring Systems*.
- Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents.
- Ben-Ari, Y. (2001). Developing networks play a similar melody. *Trends in Neurosciences*.
- Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2005). Motivation and Cog. Eng. in Learning Envs.
- Bowman, N. D., Rogers, R., & Sherrick, B. I. (2013). In Control or In Their Shoes?
- Brehm, J. W. (1966). *A theory of psychological reactance*. New York.
- Byrne, D., & Nelson, D. (1965). Attraction as a linear function of proportion of positive reinforcements.
- Campbell, W. K., & Sedikides, C. (1999). Self-threat magnifies the self-serving bias.
- Dourish, P. (2001). Seeking a Foundation for Context-Aware Computing. *HCI*.
- Duval, T. S., & Silvia, P. J. (2002). Self-awareness, prob. of improvement, and the self-serving bias.
- Fox, J., & Bailenson, J. N. (2009). Virtual Self-Modeling. *Media Psychology*, 12, 1–25.
- Gee, J. P. (2007). *What Video Games Have to Teach Us About Learning and Literacy*.
- Gibson, C. E., Losee, J., & Vitiello, C. (2014). A replication attempt of stereotype susceptibility.
- Good, C., Rattan, A., & Dweck, C. S. (2012). Sense of belonging and women’s rep. in mathematics.
- Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning.
- Gresalfi, M., Martin, T., Hand, V., & Greeno, J. (2009). Constructing competence.
- 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), 1–22.
- Gulz, A., & Haake, M. (2006). Design of animated pedagogical agents - A look at their look.
- Hamilton, J. (2009). Identifying with an avatar: a multidisciplinary perspective.
- Harrell, D. F. (2010). Toward a Theory of Critical Computing. *CTheory*.

- Harteveld, C., & Sutherland, S. (2015). The Goal of Scoring: Exploring the Role of Game Performance in Educational Games. *CHI 2015*.
- Hoaglin, D. C., & Iglewicz, B. (1987). Fine-Tuning Some Resistant Rules for Outlier Labeling.
- Huff, C., Johnson, D. G., & Miller, K. (2003). Virtual harms and real responsibility.
- Hutchins, E. (2000). Distributed Cognition. *Representations*.
- IJsselsteijn, W., De Kort, Y., Poels, K., Jurgelionis, A., & Bellotti, F. (2007). Characterising and Measuring User Experiences in Digital Games.
- Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters.
- Johnson, A. M., Didonato, M. D., & Reisslein, M. (2013). Animated agents in K-12 engineering outreach: Preferred agent characteristics across age levels. *Computers in Human Behavior*.
- Kao, D., & Harrell, D. F. (2015a). Exploring construction, play, use of virtual identities in STEM learning. *Jean Piaget Society Annual Conference*.
- Kao, D., & Harrell, D. F. (2015b). Exploring the Impact of Role Model Avatars on Game Exp. in Educational Games. *Computer-Human Interaction in Play (CHI PLAY)*.
- Kao, D., & Harrell, D. F. (2015c). Toward Avatar Models to Enhance Performance and Engagement in Educational Games. In *Computational Intelligence in Games*.
- Kao, D., & Harrell, D. F. (2015d). Toward Evaluating the Impacts of Virtual Identities on STEM Learning. *Foundations of Digital Games*.
- Kao, D., & Harrell, D. F. (2015e). Mazzy: A STEM Learning Game. *Foundations of Digital Games*.
- Kao, D., & Harrell, D. F. (2016a). Toward Understanding the Impacts of Role Model Avatars on Engagement in Computer Science Learning. *The Annual Meeting of the American Educational Research Association (AERA)*.
- Kao, D., & Harrell, D. F. (2016b). Exploring the Effects of Encouragement in Educational Games. *ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI 2016)*.
- Kao, D., & Harrell, D. F. (2016c). Exploring the Impact of Avatar Color on Game Exp. in Educational Games. *ACM Conf. Ext. Abstracts on Human Factors in Computing Systems (CHI 2016)*.
- Kaufman, G., Flanagan, M., & Seidman, M. (2015). Creating Stealth Game Interventions for Attitude and Behavior Change: An "Embedded Design" Model. *DiGRA 2015: Diversity of Play*.
- Kim, Y., & Baylor, A. L. (2006). Pedagogical agents as learning companions.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*.
- Lester, J. C., Converse, S. a., Kahler, S. E., Barlow, S. T., Stone, B. a., & Bhogal, R. S. (1997). The Persona Effect: Affective Impact of Animated Pedagogical Agents. *CHI*.
- Maccoby, E. E., & Wilson, W. C. (1957). Identification and observational learning from films.
- Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable Medial Prefrontal Contributions to Judgments of Similar and Dissimilar Others. *Neuron*, 50(4), 655–663.
- Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., ... Dalgleish, T. (2009). A key role for similarity in vicarious reward. *Science*, 324(5929), 900.
- Moon, Y. (2000). Intimate Exchanges: Using Computers to Elicit Self-Disclosure From Consumers.
- Nasir, N. S. (2002). Identity, Goals, and Learning. *Mathematical Thinking and Learning*.
- Odell, J. J., Parunak, H. V. D., & Fleischer, M. (2003). Roles in Des. Effective Agent Organizations.
- Pratt, J. A., Hauser, K., Ugray, Z., & Patterson, O. (2007). Looking at human-computer interface design: Effects of ethnicity in computer agents. *Interacting with Computers*, 19(4), 512–523.
- Rosenberg-Kima, R. B., Plant, E. A., Doerr, C. E., & Baylor, A. (2010). The influence of computer-based model's race and gender on female students' attitudes and beliefs towards eng.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype Susceptibility. *Psychological Science*.
- Steele, C., & Aronson, J. (1995). Stereotype Threat and the Int. Test Perf. of African Americans.
- Turkay, S. (2014). The Effects of Avatar-based Customization on Player Identification in Extended MMO Gameplay. *Games, Learning and Society Conference*.
- Vasalou, A., Joinson, A. N., & Pitt, J. (2007). Constructing my online self: avatars that increase self-focused attention. *Proceedings of ACM CHI 2007 Conference*.
- Veeragoudar Harrell, S., & Harrell, D. (2009). Exploring the Potential of Computational Self-Representations for Enabling Learning. *Digital Arts and Culture Conference*.
- Whang, L. S., & Chang, G. (2004). Lifestyles of virtual world residents. *Cyberpsychology & Behavior*.
- Wolfendale, J. (2007). My avatar, my self: Virtual harm and attachment.
- Yee, N., & Bailenson, J. (2007). The Proteus Effect: The Effect of Trans. Self-Rep. on Behavior.

Acknowledgements

We thank the anonymous reviewers for their valuable feedback. This work is supported by NSF STEM+C Grant #1542970 and a Natural Sciences and Engineering Research Council of Canada (NSERC) fellowship.