The Effects of Badges and Avatar Identification on Play and Making in Educational Games

Dominic Kao
MIT CSAIL
Cambridge, MA
dkao@csail.mit.edu

D. Fox Harrell
MIT CSAIL
Cambridge, MA
fox@csail.mit.edu

ABSTRACT
In our study (N=2189), we divided participants into 6 badge conditions: 1) Role model badges (e.g., Einstein), 2) Personal interest badges (e.g., Movies), 3) Achievement badges (e.g., "Code King"), 4) Choice, 5) Choice with badges always visible, and 6) No badges. Participants played a CS programming game, then used an editor to create their own level. Badges promoted avatar identification (personal interest, role model), player experience (achievement, role model), intrinsic motivation (achievement, role model), and self-efficacy (role model) during both the game and the editor. Independent of badges, avatar identification promoted player experience, intrinsic motivation, and self-efficacy. Additionally, avatar identification promoted greater overall time spent in both the game and the editor, and led to significantly higher overall quality of the completed game levels (as rated by 3 independent externally trained QA testers). Our study has implications for the design of badge systems and sheds new light on the effects of avatar identification on play and making.

ACM Classification Keywords
K.8.0 **Personal Computing**: General – Games

Author Keywords
Avatar Identification; Badges; Educational Games; Making; Level Creation; Virtual Identity; Avatar

INTRODUCTION
While the motivational potential of badges has been explored extensively in the last few years [164, 144, 166], badge studies almost exclusively focus on badges as an achievement representation (e.g., [114, 7, 140, 147, 83, 176, 160, 111, 6, 68, 144, 69, 66, 125, 39]). Yet a plethora of studies have demonstrated that being personally associated [2, 46, 189] and having role models [132, 25, 127] is crucial, even affecting career choices in mathematics and CS. As such, we believe that badges representing concepts that are associated with the self, or badges representing relevant role models could enhance motivation, self-efficacy, etc.

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ISBN 978-1-4503-5620-6/18/04. . . $15.00

DOI: https://doi.org/10.1145/3173574.3174174

Badges also have the potential to enhance avatar identification, one facet of game experience that has been a topic of increasing interest. Avatar identification has been positively correlated to motivation [16, 178], enjoyment [73, 17, 169], long-term participation [172, 104], learning-related outcomes [105, 14, 64, 10], and other player experience outcomes. While avatar identification has been studied in games [168, 178, 172, 16], little is known about how avatar identification impacts constructionist environments. With the rise of constructionist learning in CS [145, 153, 188, 82], and with virtually all environments having a user representation (user profile, avatar, etc.), this is an increasingly important topic. For example, to what extent are users identified with Logo's turtle, and does it matter?

In this study, we had 4 research questions:

RQ1: Do badges improve identification, need satisfaction, intrinsic motivation, and programming self-efficacy in the CS programming game and the subsequent game-making task?
RQ2: Does avatar identification improve need satisfaction, intrinsic motivation, and programming self-efficacy in the CS programming game and the subsequent game-making task?
RQ3: Does avatar identification translate into higher motivated behavior (time spent, etc.)?
RQ4: Does avatar identification improve created game levels?

We ran a between-subjects study on Amazon Mechanical Turk. Participants played a CS programming game, then used an editor to create their own level. Participants were divided into several badge conditions, including role model badges (e.g., Einstein), personal interest badges (e.g., Movies), and achievement badges (e.g., "Code King"). Independently from badges, we used hierarchical regression with avatar identification as the predictor.

Badges contributed to greater avatar identification (personal interest, role model), player experience (achievement, role model), intrinsic motivation (achievement, role model), and self-efficacy (role model).

Avatar identification consistently improved player experience, intrinsic motivation, and self-efficacy. Avatar identification promoted greater overall time spent in both the game and the editor, and led to significantly higher overall quality of the completed game levels (as rated by 3 independent externally trained QA testers).

To the best of our knowledge, this is the first study to look at badges that represent completely alternative types. To the best of our knowledge, this is also the first study to look at avatar identification in a making context. Our study has implications for the design of badge systems and sheds new light on the effects of avatar identification on play and making.

RELATED WORK

We present related work on badges and avatar identification.

Badges

What Is a Badge?

Badges summarize achievement and signal accomplishment [60]. Badges are used in games (e.g., Xbox 360 [83]), commerce (e.g., eBay, Amazon), education (e.g., Khan Academy), as physical status icons (e.g., ribbons, medals, trophies), and countless digital applications (e.g., foursquare, Nike+). Educationally, badges can motivate [6, 52, 124], scaffold [84, 87, 6], and credential [122, 142, 4]. In this paper, we are primarily interested in the motivational potential of badges.

In defining badges, it is impossible not to also describe the larger discourse surrounding *gamification*. Gamification, or the use of game design elements in non-game contexts [43], has been studied ubiquitously in education [62, 45, 130], online communities [58, 173, 18], health [171, 159, 79], and innumerable other domains [166]. Gamification has been applied to schools (e.g., Quest to Learn [165]), crowdsourced science (e.g., Foldit [103], Galaxy Zoo [117]), and other domains.

Yet gamification has been contentious [21, 20, 152, 41, 40], critics have argued that its approaches involve "taking the thing that is least essential to games and representing it as the core of the experience" [157]. A centerpiece of this discourse is the notion that extrinsic rewards—e.g., external rewards such as points or in-game currency—can undermine intrinsic motivation [36], i.e., engaging in a behavior because it is satisfying in and of itself [161]. However, meta-analytic reviews to date have not supported this argument [166, 44, 51]. Moreover, this debate is generally centered around non-game contexts (rewards such as badges, leaderboards, points, etc. are viewed as essential game components [43]). Our aim here is to study different types of badges in an educational game.

What Are the Effects of Badges?

Researchers have found that badges increase user activity, e.g., posting trade proposals in a trading service [69, 68] and affect behavior [66, 125, 6], e.g., Q/A behavior on Stack Overflow [125, 6]. Badges have been shown to increase student contributions [39], promote higher grades [45], enjoyability [55], and can affect the behavior of students even when badges have no impact on grading [66]. Researchers caution, however, that badges can have a negative effect on written work [45] and foster undesirable behavior [55].

Researchers have also found that badges increase forum engagement in MOOCs [7] and productivity of Wikipedia contributors by 60% [154]. The motivational effects of badges can vary by person, activity, and badge [1, 22, 54, 148]. Typically, badges are awarded for fulfilling criteria such as accomplishment [45], participation [6], carefulness [66], and behavior change [3, 116]. Isolating badge effects is often challenging-ing, as many researchers leverage multiple game elements [111, 70, 166]. However, a systematic review in education of game elements, including badges, have found mostly encouraging results [44].

What Types of Badges Have Been Studied?

Badges are widely considered synonymous to achievements [8, 194, 185, 6] and are depicted as achievements across virtually all research studies [114, 7, 140, 147, 83, 176, 160, 111, 6, 68, 144, 69, 66, 125, 39]. While badges have also been discussed as a mechanism for feedback [156], guidance [183], etc., these have been secondary. Most commonly, badges are awarded for performance or completion [19]—for example, "Faster than Lightning" [1], "3D Expert" [176], "Y U No Make Mistakes?" [66], "Almost finished!" [72], "Curious" [125], "Question answerer" [39], etc.

Researchers have sought to categorize badge types from several mainstream video games including Mass Effect, Grand Theft Auto 4, and World of Warcraft: Tutorial, Completion, Collection, Virtuosity, Hard mode, Special Play Style, Veteran, Loyalty, Curiosity, Luck, Mini-Game, Multi-Player, Paragon, and Fandom [137]. In all cases, however, badges depict a type of achievement. Other researchers note that badges may vary in their signifier (the actual visual badge), completion logic (the conditional logic required), and reward [67]. In this study, we are interested in comparing typical achievement badges with badges awarded for the same reasons, but that represent other things, e.g., a personal interest.

Personal Interests

Researchers have argued that students do not generally value science as personally relevant [5, 110], and that games should utilize the affordances for personal relevance [57]. Researchers argue that by incorporating personally relevant items, games can help students both develop schematic knowledge while building a personal connection [57]—see [2, 46, 189, 5, 57, 35] for similar arguments. Researchers suggest that badges that utilize personal identity would: 1) build on identity that is more firmly established, and 2) strengthen positive associations between learning and the student's identity [2]. One badge type we study is personal interests.

Role Models

Role model—a reference that occupies a desirable standing—was coined by Robert K. Merton [27]. Role models can increase academic performance [132, 25, 127], even affecting career choices in mathematics and CS. Three factors increase role models' effectiveness: 1) Perception of competence [128], 2) Perception of being ingroup, e.g., shared attributes like gender and race [119, 127], 3) Perception of success [25]. In a CS programming game, participants using role model avatars, e.g., Einstein, have significant increases in flow, immersion, etc. [92, 93, 100]. In a study on group brainstorming using a virtual environment, participants that used scientist-like avatars had more original ideas [65]. Towards understanding if benefits from using role model avatars can be applied in other forms, we investigate role model badges.

Avatar Identification

What Is Avatar Identification?

Avatar identification—sometimes called "player-avatar identification" [113], "character identification" [169], or "avatar-self connection" [85]—is a *temporary alteration of media users' self-concept through adoption of perceived characteristics of a media person* [30]. Building upon Cohen's work [32], Klimmt et al. argue that during exposure to a video game, users become one with their character [30]. Extensive work exists on identification with television characters [77, 76, 33, 170]. However, one important difference is that video games emphasize agency [59]. The active participation in video games is argued to override the distance between user and character [73, 108]. Avatar identification is strongly moderated by similarity, such as demographics, experiences, etc. [32] and other variables, e.g., game type [175]. Other work has shown that users realize their "ideal selves" through games [15], both physically and psychologically [47]. Research suggests that we slowly become more congruent with our virtual identities over time [191, 47, 192, 151].

What Are the Effects of Avatar Identification?

Avatar identification can improve game enjoyment [175, 112, 17, 143], health outcomes [106], intrinsic motivation [16, 178], flow [169], exercise motivation [112, 182], and trust in others [104]. Avatar identification can also reduce self-discrepancy (the distance between one's actual and ideal selves [75]) [15, 115], improve self-esteem [184], game loyalty [172], learning interest [9], game appreciation [23], game motivation [180], decrease deceptive behavior [78], increase willingness to purchase game items [146, 193], and has been associated with aggression [109], addiction [168], depression [131, 15], and increased persuasiveness of messages [141].

In education, there is a long history of work in avatars and pedagogical agents (i.e., virtual pedagogical agents, teaching agents, etc.). 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 [105, 14, 64, 10]—i.e., similarities [175, 179]. This is posited to be a result of similarity-attraction, the theory that people are attracted to similar others [26, 81]. Functional neuroimaging has found that perceived similarity is an important factor in a person's ability to simulate the internal state of another person [135]. 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 [136]. Furthermore, work has suggested that what is experienced by an avatar is also experienced by its user [28, 138, 190, 186]. This effect is more powerful via avatars that we identify with [181, 48], identification being positively correlated to such factors as representation of emotions and intent [71], physical resemblance [123], and avatar customization [177].

For instance, Birk, Atkins, Bowey and Mandryk, divided participants into two groups, one that customized their avatar and another that watched a video of their avatar being customized. Those participants that customized their avatar had increased

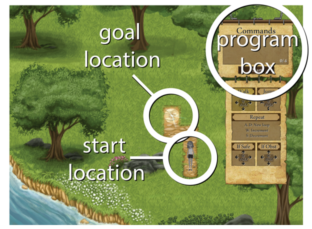

Figure 1. Level 1 in Mazzy introduces the basic game mechanics.

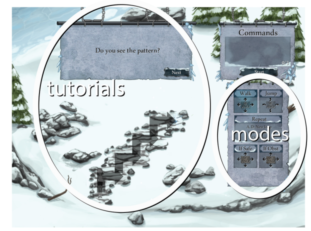

Figure 2. Level 6 introduces looping.

identification. Furthermore, participants' identification with their avatars significantly predicted various measures related to engagement such as affect, immersion, and amount of time playing [16]. While similarity plays a key role in enhancing avatar identification [175, 179], creating ideal versions of ourselves—sometimes referred to as wishful identification [77, 76]—can also be of value [149]. But this same discrepancy between the actual and ideal self is predictive of negative health outcomes, such as depression and anxiety [63, 75].

Through the clear predictive capacity of avatar identification, we seek to study avatar identification in a making context and its relationship to various facets. For instance, with highly active interest in making [88, 38, 24, 90], we have yet to scratch the surface of topics like identification [88]. Does identification with an on-screen object—a profile picture, a Mii, a turtle, an avatar, etc.—enhance users' making experience? Towards investigating this question, we investigate avatar identification in the context of making.

THE GAME

The first phase of our experiment takes place in a STEM learning game called Mazzy [94][1]. Mazzy is a game in which players solve levels 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 [91, 92, 95, 96, 100, 98, 99, 97, 101].

THE EDITOR

The second phase of the experiment takes place in an editor [90]. At a high-level, the editor allows players to create their own Mazzy game levels. Each map consists of a grid of tiles, each of which can be textured separately and modified logically to be a safe or unsafe tile for the player to step on. The maps can be any size (from 1x1 to, e.g., 100x100). See Figure 3. Basic functionalities of the editor include: manipulating the view, creating assets that can be translated, rotated, rescaled, searching for images via a built-in image search that interfaces with Microsoft Bing (see Figure 4), and testing maps by playing them. Although the editor typically provides preloaded images for use, in this study none were provided, i.e., all images were searched for.

[1]Gameplay video: http://youtu.be/n2rR1CtVal8

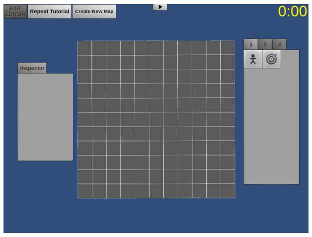

Figure 3. Blank 11x11 map in the editor.

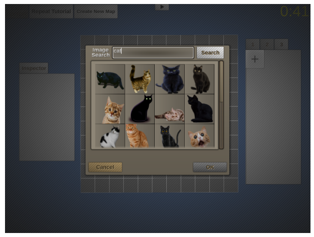

Figure 4. Searching for the image "cat".

METHODS

Creating Badges

Our goal in this step was to create a set of achievement, role model, and personal interest badges. In order to do so, we populated the initial badge list through crowdsourcing. We validated the badges to ensure that they were adequate for use in our experiment also through crowdsourcing. Since our actual study took place on Amazon Mechanical Turk (AMT), these validation studies also used AMT participants.

Initial Population of Interests and Scientists

100 participants were asked to list: 1) 10 of their personal interests, and 2) 10 scientist role models. No restrictions were made as to the types of interests or scientists.

Interest Corrections

The 1000 interests generated from the preceding step then underwent light corrections. During these corrections, similar interests were renamed, e.g., "Video Gaming" and "videogames" were renamed to "Video Games". Typos were also corrected.

Interest Categorization

113 participants categorized the 1000 interests. Users were instructed to write a high-level category for each interest, e.g., for "Playing Drums", they might write "Playing an Instrument" or "Music".

Popularity Ranking of Interests and Scientists

241 unique scientists, 470 unique interests, and 1005 unique categories of interests were ranked by how frequently different users had mentioned them.

Final Scientists

For the final set of 30 scientists, we selected the top 20 most often mentioned scientists, and the remaining 10 were researcher-curated from scientists that were mentioned by at least 2 different users, and were diverse in gender and race. While every scientist mentioned had a record of success (i.e., competence [128, 25]), this curation was done to ensure more coverage and inclusiveness (i.e., in-group potential [119, 127]). See Table 1.

Final Interests

209 participants rated 100 interests (50 most mentioned interests, and categories). Each rating was done with the format "I have an active interest in _________" on a scale of 1: *Strongly Disagree* to 7: *Strongly Agree*. For the final set, we selected the 15 most highly rated interests, and 15 most highly rated categories. We additionally checked that the averages of this final set of 30 did not differ by gender or race to ensure that personal interests were widely represented, p>.05. See Table 1.

Interests: Food Interest, Family Interest, Movies Interest, Technology Interest, Music Interest, Reading Interest, Nature Interest, Games Interest, Television Interest, Comedy Interest, Animals Interest, Traveling Interest, Health Interest, Cooking Interest, Science Interest, Internet Interest, Fun Interest, Life Interest, Knowledge Interest, Money Interest, Intelligence Interest, Self-Improvement Interest, Creativity Interest, Pets Interest, Fiction Interest, Universe Interest, Exploring Interest, Creating Interest, Culture Interest, Society Interest
Scientists: Albert Einstein, Isaac Newton, Stephen Hawking, Nikola Tesla, Marie Curie, Charles Darwin, Galileo Galilei, Thomas Edison, Carl Sagan, Neil deGrasse Tyson, Alexander Graham Bell, Louis Pasteur, Leonardo da Vinci, Niels Bohr, Jane Goodall, Aristotle, Nicolaus Copernicus, Bill Nye, Gregor Mendel, Archimedes, Michio Kaku, George Washington Carver, Rosalind Franklin, Rachel Carson, Mae Jemison, Lise Meitner, Mary Somerville, Ibn al-Haytham, C. V. Raman, Ada Lovelace
Achievements: Baby Steps, Setting Sail, A New World, Spring in Your Step, Scenic Route, Straight Runner, Taking Off, Water Crosser, Zero Gravity, Step Saver, Creative Solution, Fly By, Labyrinth Master, Perilous Pathways, Lucky Leaper, Loophole, Tip of the Iceberg, Round Trip Flight, Fleet Footed, Gaining Traction, Seventh Heaven, Repeat Runner, Sub-Orbital, Making Camp, Mountain Guide, Prodigy Walker, Beep Boop Beep, What a Breeze, Look Out, World Explorer, Conditional Victory, Code Warrior, Extreme Conditions, A New Dawn, Code King, 100% Worth

Table 1. Final set of interests, scientists, and achievements

Finding Badge Images

We used Google to find badge images for each personal interest (searching for "_________ Icon") and for each scientist (searching the name directly). We then took the very first search result and two more results at random from the first 10.

Processing Badge Images

All potential badge images were cropped to be square. All images were converted to 8-bit black and white to normalize color [134, 80, 49, 99].

Rating Badge Images

107 AMT users then ranked the 3 potential badge images for each personal interest and scientist. For personal interests, users were asked to rank the images from best to worst in their representation of the interest. For scientists, users ranked the images from best to worst in their representation of the scientist. If they didn't recognize a scientist, they were asked to find out more about the scientist (a link to the scientist's Wikipedia page was provided). Both question order and image order were randomized. Final rankings did not differ by gender or race, p>.05. Intraclass correlation on the rankings was ICC = 0.94 (two-way random, average measures [167]), indicating high agreement.

Finding Achievement Badge Images

With a graphic design artist, we created or found 5 potential achievement images for each game level. These were created to look like an achievement and match the game level context.

Ranking Achievement Badge Images

122 participants played Mazzy. After each level, users ranked the 5 potential badge images for that level. Users provided

their own captions for the top three badges they ranked. Rankings did not significantly differ by gender or race, p>.05. Intraclass correlation on the rankings was ICC = 0.94 (two-way random, average measures), indicating high agreement. The three highest ranked images for each level were kept.

Achievement Badge Captions
Captions were selected from participants' responses. Captions were selected to match the badge image and the game level context. See Table 1.

Number of Caption Words
We performed an ANOVA on number of caption words by badge type, and found no significant effect, F(2, 93) = 1.42, p = 0.25.

Badge Image Stylization
To ensure that all badges had a uniform design, we first removed transparency and pure white backgrounds replaced by a neutral gray (rgb[212, 212, 212]). We then used a stylization filter called the "Photocopy" filter in Adobe Photoshop CS5. This served both to provide a uniform design and also to normalize average pixel intensity. The average pixel intensity across badge types did not differ, F(2, 93) = 1.39, p = 0.25.

Badge Image Representativeness
100 participants rated the stylized badge images. They were asked "The picture on the right represents the person, icon, or illustration depicted in the picture on the left." on a scale of 1: *Strongly Disagree* to 7: *Strongly Agree*. A one-way ANOVA found no significant effect of badge type on the representativeness of the stylized images, F(2, 93) = 1.39, p = 0.26.

Final Badges
All badge images were then given a circular frame and a ribbon with caption text. This was done by a graphic designer who was instructed to prioritize uniformity, e.g., spacing in the text, etc. All final badges were vetted.

Final Badge-Likeness
103 participants were presented with each final badge and the *original* image—before any changes. They were asked "The image on the right is a good badge representation of the image on the left" on a scale of 1: *Strongly Disagree* to 7: *Strongly Agree*. A one-way ANOVA found no significant effect of badge type on goodness of badge representation, F(2, 93) = 0.56, p = 0.57.

Final Badge Aptness to Game
110 participants watched a gameplay video of Mazzy. For each final badge, they were asked: "This is a suitable badge for Mazzy" on a scale of 1: *Strongly Disagree* to 7: *Strongly Agree*. A one-way ANOVA found no significant effect of badge type on suitability of badge, F(2, 93) = 0.47, p = 0.63.

Conditions
The six badge conditions we tested were:

a. Role model
b. Personal interest
c. Achievement
d. Choice
e. Choice with badges always visible
f. No Badges

Badges are always awarded at the end of each level of the game. See Figure 5. **All badge conditions except *Choice always visible*:** During the game, badges are visible when the player completes a level briefly (1.5 seconds), and briefly again after having chosen a badge (1.5 seconds). See Figure 6. **All badge conditions except *Choice always visible*:** During the editor, badges are always visible in the bottom left. See Figure 6. Badges always appear alongside the avatar. In the editor, the user avatar and badges are positioned such that a distance of 33 pixels exists between the leftmost edge of the window and the nearest badge or avatar pixel. **Choice always visible:** In the *Choice always visible* condition, badges are shown at the bottom of the screen at all times. All other aspects are identical. **No Badges:** In the *No Badges* condition, completing a level shows the Congratulations screen for 3 seconds (with only the player avatar).

Condition specifics: **Role model condition:** Users select from 30 scientists, and can hover over a "Who is this?" button to see a 3-sentence summary taken verbatim from Wikipedia. This is in the form of a semi-transparent black overlay that appears at the bottom. **Personal interest condition:** Users select from 30 interests. **Achievement condition:** Users unlock three achievement badges per level completed. After completing a level, they select from the newly unlocked badges as well as any previously not chosen badges. Newly unlocked achievements appear at the head of the list. There are 36 achievement badges. **Choice condition:** Users choose from all three badge types. Each badge type works the same way as its individual condition. Badge types are presented in randomized ordering, and background images are taken randomly from each individual subset. See Figure 5. **Choice always visible condition:** Identical to *Choice* except for display of badges. **No Badges condition:** No badges. **All badge conditions:** Previously chosen badges are marked with a green checkmark, and cannot be chosen a second time. Badges are randomized in their ordering for each user.

A Note on Choice
We implemented choice in all badge conditions since awarding an appropriate role model and personal interest would require advance knowledge of the user's preferences. The current approach appeared more ecologically valid, e.g., giving your demographic information or your personal interests both could change the outcome of the experiment and is not typical information a game would have. Choice is also generally beneficial, e.g., [162, 126, 86, 107].

Quantitative and Qualitative Measures
Player Experience of Need Satisfaction
We use the 21-item Player Experience of Need Satisfaction (PENS) scale [163] that measures the following dimensions: Competence, Autonomy, Relatedness, Presence/Immersion, and Intuitive Controls. PENS is based on self-determination theory (SDT) [37]. PENS contends that the psychological "pull" of games are largely due to their ability to engender

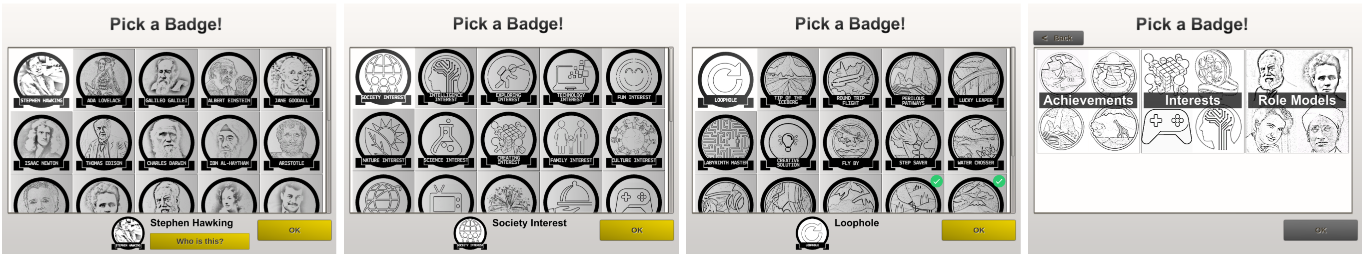

Figure 5. Conditions: a) Role Model, b) Personal Interest, c) Achievement, and d) Choice.

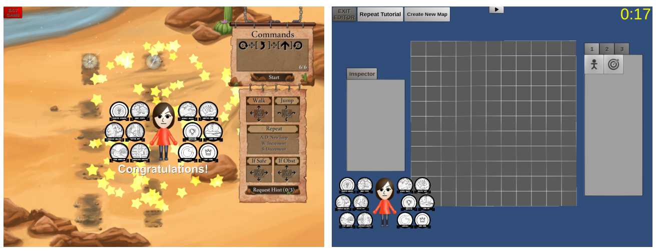

Figure 6. Badges as they appear: a) In-game, and b) In-editor.

three needs—*competence* (seek to control outcomes and develop mastery [187]), *relatedness* (seek connections with others [13]), and *autonomy* (seek to be causal agents [29] while maintaining congruence with the self) [163]. PENS is considered a robust framework for assessing player experience [156, 42].

Computer Programming Self-Efficacy Scale

Self-efficacy represents the belief in one's ability to succeed, either in a particular situation, or at a particular task [12]. The Computer Programming Self-Efficacy Scale (CPSES) is a scale for measuring programming self-efficacy. It consists of a validated 32-item scale that measures the following dimensions: Independence and persistence, Complex programming tasks, Self-regulation, and Simple programming tasks [150].

Intrinsic Motivation Inventory

The Intrinsic Motivation Inventory (IMI) assesses intrinsic motivation using four dimensions: 1) Interest/Enjoyment, e.g., I enjoyed doing this activity very much, 2) Effort/Importance, e.g., I put a lot of effort into this, 3) Pressure/Tension, e.g., I felt very tense while doing this activity, 4) Value/Usefulness, e.g., I believe this activity could be of some value to me [129].

Player Inventory Scale

The Player Inventory Scale (PIS) measures avatar identification [179], which consists of three second-order factors: 1) Similarity identification, e.g., My character is similar to me, 2) Embodied identification, e.g., In the game, it is as if I become one with my character, 3) Wishful identification, e.g., I would like to be more like my character.

Map Quality Ratings

We collected both user and expert quality ratings of the final created game levels. Users were asked to rate their final game level on the dimensions of: "Aesthetic" (Is it visually appealing?), "Originality" (Is it creative?), "Fun" (Is it fun to play?), "Difficulty" (Is it difficult to play?), and "Overall" (Is it excellent overall?) on a scale of 1: *Strongly Disagree* to 7: *Strongly Agree*.

Expert ratings were given by 3 QA testers we hired. All QA testers had extensive games QA experience. The 3 QA testers first underwent supervised training in which they finished the game and created at minimum 3 maps in the editor. QA testers were then given 100 maps at random to establish baseline expectations. Next, QA testers were given another 25 maps at random to rate on the same dimensions as user ratings. Researchers verified the ratings and maps were rescored until there was consensus.

All 3 QA testers were blind to the experiment—the only information they received was a list of maps and links to each game level. They were debriefed on the purpose of their work after they completed all 2189 ratings. The 3 QA testers each spent an average of 109 hours (SD=8.5) over a 1-month period, at $15 USD/hr.

Time Played

We directly measure motivation as operationalized by the amount of time spent playing the game and the editor.

Participants

2189 participants were recruited through Mechanical Turk. The data set consisted of 1001 male, and 1188 female participants. Participants self-identified their races/ethnicities as white (1681), black or African American (177), Chinese (41), Asian Indian (35), American Indian (23), Filipino (22), Korean (17), Vietnamese (13), Japanese (10) and other (170). Participants were between the ages of 18 and 71 (M = 30.1, SD = 9.1), and were all from the United States. Participants were reimbursed $3.00 to participate in this experiment.

Design

A between-subjects design was used: badge condition was the between-subject factor. Participants were randomly assigned to a condition.

Protocol

Players first spent a minimum 4 minutes creating their Mii avatar. See Figure 7. The Mii creator was adapted from a freely available online Mii creator [34]. Options available included avatar/user name, gender, birthday, height, weight, favorite color (6), face shape (8), skin color (6), facial features (12), hair (72), hair color (8), eyebrows (24), eyebrow color (8), eyes

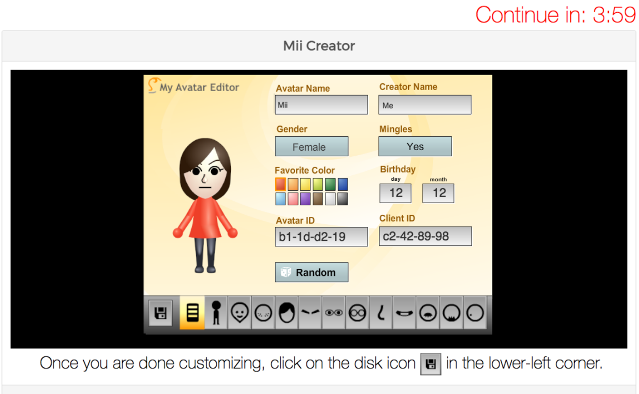

Figure 7. Mii avatar creator.

(48), eye color (6), glasses (9), glasses color (6), nose (12), mouth (24), mouth color (3), mustache (4), beard (4), facial hair color (8), mole (2). There were also miscellaneous other options such as direction of hair part, and positioning, scale, and rotation of various facial elements. Avatar customization has previously been shown to increase, and produce a range of, avatar identification in users [11, 115, 178, 16].

Next, users completed the avatar identification scale. Before proceeding to the game, players were informed that they could exit the game *at any time* via a red button in the corner of the screen. Participants then played Mazzy. When participants were done playing (either by exiting early, or by finishing all 12 levels), participants completed the avatar identification scale, the player experience of need satisfaction scale, the intrinsic motivation inventory, and the programming self-efficacy scale.

Next, users completed a tutorial which introduced them to the editor. The tutorial stepped users through all the interface elements and all editor functionalities. Each of the total 38 steps of the tutorial asked the user to perform an action before they could proceed, e.g., click a highlighted button to test the level. Each tutorial step had an additional help facility that provided additional troubleshooting information. After users completed the tutorial, they were required to spend at least 10 minutes creating a game level. After the 10 minutes passed, they could exit the editor via a red button in the corner of the screen, or continue using the editor until they wanted to quit. Users could repeat the tutorial at any time. After users quit the editor, they took a final screen capture of their level. This was done by positioning the map in the viewport and clicking a “Take Screenshot” button.

Users then completed the avatar identification scale, the player experience of need satisfaction scale, and the intrinsic motivation inventory. Users provided self-ratings on their completed game levels, and filled out demographics.

Analysis

Data was extracted and imported into Statistical Package for Social Science (SPSS) version 22 for data analysis using multivariate analysis of variance (MANOVA). Separate MANOVAs are run for each separate set of items—*PENS, CPSES, IMI, PIS*; with the independent variable—*badge condition*. All the dependent variables are continuous variables. The independent variable badge condition (i.e., 0 = role model, 1 = personal interest, 2 = achievement, 3 = choice, 4 = choice with badges always visible, 5 = no badges) is a sexchotomous variable. To detect the significant differences between badge conditions, we utilized one-way MANOVA. These results are reported as significant when $p<0.05$ (two-tailed). Prior to running our MANOVAs, we checked both assumption of homogeneity of variance and homogeneity of covariance by the test of Levene’s Test of Equality of Error Variances and Box’s Test of Equality of Covariance Matrices; and both assumptions were met by the data ($p>.05$ for Levene’s, and $p>.001$ for Box’s).

To measure the predictive capacity of avatar identification, we used linear hierarchical regression using similarity identification, embodied identification, and wishful identification as individual predictors. Since age and sex have been shown to affect need satisfaction, intrinsic motivation, and other avatar identification-related outcomes during game play [163, 115, 155], we entered age and sex in the first block of the regressions. We use avatar identification to predict game-related PENS, IMI, CPSES scores and editor-related PENS and IMI scores (using the avatar identification recorded pre-gameplay, and pre-editor, respectively). We then use avatar identification to predict play time and other time-related outcomes. Finally, we test if avatar identification can predict both self and expert ratings of final game level quality.

RESULTS

RQ1: Do badges improve identification, need satisfaction, intrinsic motivation, and programming self-efficacy in the CS programming game and the subsequent game-making task?

Avatar Identification

Personal interest and role model badges promoted avatar identification in the game-making task. The MANOVA was not statistically significant across badge conditions in the game, $p>.05$. The MANOVA was statistically significant across badge conditions in the editor, $F(15, 6021) = 4.05$, $p<.0001$; Wilk’s $\lambda = 0.973$, partial $\eta^2 = 0.01$. ANOVAs found the effect to be significant across all three dimensions of identification, $p<.0001$. Posthoc testing was done using Tukey HSD[2]. See Figure 8 and 9.

Similarity Identification:

- Personal Interest > No Badges (*editor*), $p<.05$
- Personal Interest > Achievement (*editor*), $p<.005$
- Personal Interest > Choice Always Visible (*editor*), $p<.001$
- Role Model > No Badges (*editor*), $p<.05$
- Role Model > Achievement (*editor*), $p<.005$
- Role Model > Choice Always Visible (*editor*), $p<.001$

Embodiment Identification:

- No Badge > Choice Always Visible (*editor*), $p<.05$
- Personal Interest > Achievement (*editor*), $p<.005$
- Personal Interest > Choice (*editor*), $p<.05$
- Personal Interest > Choice Always Vis. (*editor*), $p<.0001$

[2]Note that we are calculating the *difference* in identification scores, i.e., post-editor minus pre-editor.

- Role Model > Achievement (*editor*), p <.005
- Role Model > Choice (*editor*), p <.05
- Role Model > Choice Always Visible (*editor*), p <.0001

Wishful Identification:

- Role Model > Achievement (*editor*), p <.05
- Role Model > Choice (*editor*), p <.01
- Role Model > Choice Always Visible (*editor*), p <.01

Player Experience of Need Satisfaction

Achievement badges promoted player experience in the CS programming game. Role model badges promoted player experience in both the CS programming game and the game-making task. The MANOVA was statistically significant across badge conditions in the game and the editor, p <.05. ANOVAs found that the effect was significant across all five dimensions in both the game and the editor, p <.05. Posthoc testing using Tukey HSD found that for **competence:** Role Model > Personal Interest (*game*), p <.05, Role Model > Choice (*editor*), p <.005, Role Model > Choice Always Visible (*editor*), p <.01. For **autonomy:** Achievement > Personal Interest (*game*), p <.05, Role Model > Personal Interest (*game*), p <.05, Role Model > Choice (*editor*), p <.05, Role Model > Choice Always Visible (*editor*), p <.05. For **relatedness:** Achievement > Personal Interest (*game*), p <.05, Role Model > Personal Interest (*game*), p <.05. For **immersion:** Achievement > Personal Interest (*game*), p <.05, Role Model > Personal Interest (*game*), p <.05, Role Model > Choice Always Visible (*editor*), p <.05. For **intuitive control:** Role Model > Choice (*game*), p <.05, Role Model > Choice (*editor*), p <.05. See Figure 8 and 9.

Intrinsic Motivation Inventory

Achievement badges promoted intrinsic motivation in the CS programming game. Role model badges promoted intrinsic motivation in both the CS programming game and the game-making task. The MANOVA was statistically significant across badge conditions in the game and the editor, p <.05. ANOVAs found that the effect was significant for the dimensions of enjoyment (*game*, *editor*), p <.05, usefulness (*editor*), p <.05. Posthoc testing using Tukey HSD found that for **enjoyment:** Achievement > No Badges (*game*), p <.05, Role Model > No Badges (*game*), p <.05, Role Model > Choice Always Visible (*editor*), p <.01. For **usefulness:** Role Model > Choice Always Visible (*editor*), p <.05. See Figure 8 and 9.

Computer Programming Self-Efficacy Scale

Role model badges promoted programming self-efficacy in the CS programming game. The MANOVA was statistically significant across badge conditions, p <.0001. ANOVAs found that the effect was significant for the dimensions of independence (*game*), p <.05, and self-regulation (*game*), p <.05. Posthoc testing using Tukey HSD found that for **independence:** Role Model > Personal Interest (*game*), p <.05. For **self-regulation:** Role Model > Personal Interest (*game*), p <.05. See Figure 8.

Time Played

ANOVAs found no significant effect of badge condition on **game time:** M=1497.06, SD=1952.58, F(5, 2183)=0.64, p=0.67; and **editor time:** M=741.02, SD=723.44, F(5, 2183)=0.88, p=0.49.

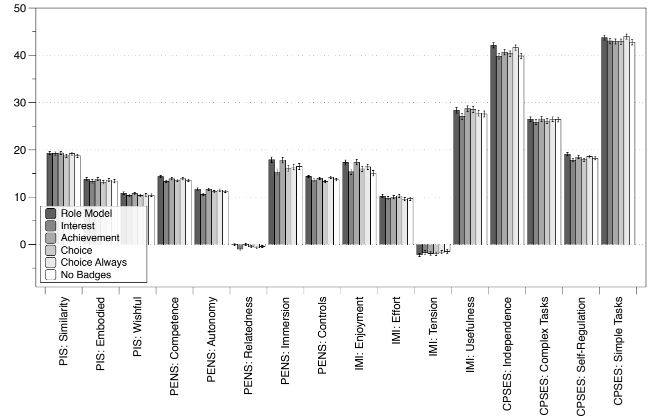

Figure 8. Measures, Post-Game. Error bars are standard error of the mean (SEM).

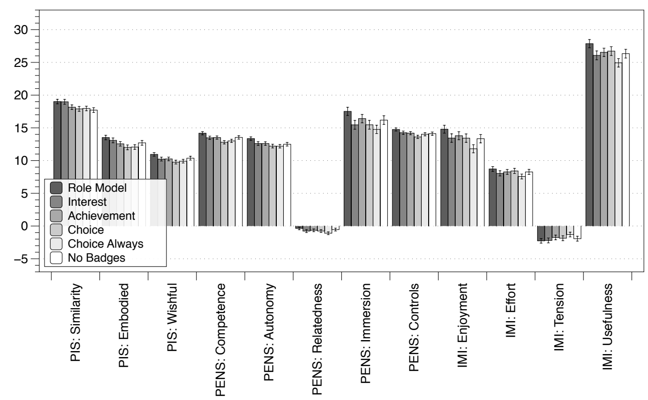

Figure 9. Measures, Post-Editor. Error bars show SEM.

RQ2: Does avatar identification improve need satisfaction, intrinsic motivation, and programming self-efficacy in the CS programming game and the subsequent game-making task?

From the hierarchical regression in Table 2, avatar identification significantly improves need satisfaction, intrinsic motivation, and programming self-efficacy in both the game and the editor. On average, significant R^2 values explain 7.4% of variance.

RQ3: Does avatar identification translate into higher motivated behavior (time spent, etc.)?

From the hierarchical regression in Table 2, embodied identification leads to higher game time and similarity identification leads to higher editor time. All dimensions of avatar identification lead to more time playtesting in editor, and more time spent taking the final screenshot. On average, significant R^2 values explain 1.8% of variance.

RQ4: Does avatar identification improve created game levels?

	Similarity Identification					Embodied Identification					Wishful Identification				
	β	R^2	$R^2(c)$	$F(c)$	$p(c)$	β	R^2	$R^2(c)$	$F(c)$	$p(c)$	β	R^2	$R^2(c)$	$F(c)$	$p(c)$
Player Experience of Need Satisfaction (PENS)															
Competence (*game*)	0.13	0.024	0.016	35.4	**0.00**	0.00	0.009	0.000	0.00	0.96	0.13	0.030	0.022	49.1	**0.00**
Competence (*editor*)	0.25	0.069	0.061	142	**0.00**	0.23	0.060	0.053	123	**0.00**	0.17	0.037	0.028	64.5	**0.00**
Autonomy (*game*)	0.19	0.034	0.034	76.3	**0.00**	-0.01	0.000	0.000	0.05	0.82	0.24	0.059	0.059	136	**0.00**
Autonomy (*editor*)	0.28	0.077	0.075	178	**0.00**	0.32	0.104	0.102	248	**0.00**	0.25	0.065	0.063	148	**0.00**
Relatedness (*game*)	0.18	0.034	0.030	68.2	**0.00**	-0.02	0.004	0.000	0.81	0.37	0.33	0.112	0.109	267	**0.00**
Relatedness (*editor*)	0.21	0.050	0.044	101	**0.00**	0.48	0.234	0.228	650	**0.00**	0.41	0.177	0.171	453	**0.00**
Immersion (*game*)	0.24	0.060	0.059	136	**0.00**	-0.02	0.002	0.000	0.84	0.36	0.40	0.163	0.162	422	**0.00**
Immersion (*editor*)	0.33	0.108	0.108	264	**0.00**	0.63	0.394	0.394	1418	**0.00**	0.51	0.257	0.257	755	**0.00**
Intuitive Control (*game*)	0.10	0.021	0.009	20.9	**0.00**	-0.01	0.012	0.000	0.09	0.76	0.11	0.024	0.012	27.6	**0.00**
Intuitive Control (*editor*)	0.22	0.056	0.048	111	**0.00**	0.20	0.049	0.041	93.7	**0.00**	0.13	0.026	0.018	39.9	**0.00**
Intrinsic Motivation Inventory (IMI)															
Enjoyment (*game*)	0.19	0.041	0.035	79.8	**0.00**	0.00	0.006	0.000	0.01	0.93	0.17	0.035	0.029	65.9	**0.00**
Enjoyment (*editor*)	0.24	0.059	0.058	135	**0.00**	0.29	0.084	0.083	198	**0.00**	0.22	0.051	0.050	115	**0.00**
Effort (*game*)	0.20	0.080	0.038	91.2	**0.00**	0.02	0.042	0.001	1.17	0.28	0.12	0.056	0.015	33.9	**0.00**
Effort (*editor*)	0.23	0.088	0.051	122	**0.00**	0.21	0.080	0.043	103	**0.00**	0.16	0.061	0.026	60.1	**0.00**
Tension (*game*)	0.02	0.004	0.000	1.03	0.31	-0.01	0.003	0.000	0.17	0.68	0.07	0.008	0.005	10.3	**0.00**
Tension (*editor*)	-0.01	0.005	0.000	0.07	0.79	0.12	0.019	0.014	31.4	**0.00**	0.12	0.019	0.014	32.1	**0.00**
Usefulness (*game*)	0.21	0.043	0.042	95.8	**0.00**	0.00	0.001	0.000	0.00	0.99	0.26	0.065	0.0065	151	**0.00**
Usefulness (*editor*)	0.31	0.096	0.094	228	**0.00**	0.38	0.149	0.147	376	**0.00**	0.32	0.106	0.104	254	**0.00**
Computer Programming Self-Efficacy Scale (CPSES)															
Independence (*game*)	0.06	0.026	0.004	8.54	**0.00**	0.01	0.023	0.000	0.31	0.57	0.02	0.023	0.000	0.60	0.44
Complex Tasks (*game*)	0.08	0.035	0.005	12.4	**0.00**	0.01	0.030	0.000	0.07	0.80	0.08	0.035	0.006	12.8	**0.00**
Self-Regulation (*game*)	0.09	0.018	0.009	19.4	**0.00**	-0.01	0.011	0.000	0.43	0.51	0.07	0.016	0.005	11.2	**0.00**
Simple Tasks (*game*)	0.08	0.020	0.006	12.4	**0.00**	0.01	0.014	0.000	0.18	0.67	0.02	0.015	0.000	0.92	0.34
Behavior															
Time Played (*game*)	0.03	0.017	0.001	1.45	0.23	0.04	0.019	0.002	4.20	**0.04**	0.02	0.017	0.000	0.56	0.46
Time Played (*editor*)	0.04	0.002	0.002	4.13	**0.04**	0.02	0.001	0.000	0.71	0.40	-0.02	0.001	0.001	1.10	0.30
Time Testing (*editor*)	0.02	0.026	0.000	0.87	0.35	0.07	0.030	0.005	11.0	**0.00**	0.06	0.029	0.004	8.79	**0.00**
Time Taking Screenshot (*editor*)	0.08	0.017	0.006	12.3	**0.00**	0.06	0.015	0.004	8.24	**0.00**	0.07	0.016	0.005	10.1	**0.00**
Game Map Ratings															
Overall Quality (*self-rated*)	0.20	0.059	0.038	87.2	**0.00**	0.23	0.074	0.052	124	**0.00**	0.20	0.063	0.041	95.2	**0.00**
Aesthetic (*expert-rated*)	0.06	0.021	0.004	7.86	**0.00**	0.07	0.021	0.004	9.45	**0.00**	-0.04	0.019	0.002	3.62	0.06
Originality (*expert-rated*)	0.05	0.021	0.002	5.04	**0.03**	0.05	0.022	0.003	6.44	**0.01**	-0.06	0.023	0.004	9.14	**0.00**
Fun (*expert-rated*)	0.04	0.024	0.002	3.98	**0.05**	0.06	0.025	0.003	6.81	**0.01**	-0.06	0.026	0.004	8.86	**0.00**
Difficulty (*expert-rated*)	0.05	0.037	0.003	6.09	**0.01**	0.03	0.035	0.001	1.98	0.16	-0.06	0.037	0.003	7.41	**0.01**
Overall Quality (*expert-rated*)	0.06	0.027	0.003	7.01	**0.01**	0.06	0.028	0.004	8.30	**0.00**	-0.06	0.027	0.003	7.70	**0.01**

Table 2. Regression properties β, R^2, change in R^2, F, and p from adding identification. Change statistics are marked (c). Significant results are bold.

Intraclass correlation across the three raters on overall quality was ICC=0.83 (two-way random, average measures), indicating high agreement. From the hierarchical regression in Table 2, all three dimensions of avatar identification lead to higher self-perceived quality. Similarity identification and embodied identification lead to increases in actual game level quality. However, wishful identification leads to a decrease in actual game level quality. On average, significant R^2 values explain 3.3% of variance. Average quality as rated by experts was M=3.54, SD=1.15. See Figure 10.

DISCUSSION

We found that each of our research questions could be answered in the affirmatory. Badges promoted avatar identification (interest, role model), player experience (achievement, role model), intrinsic motivation (achievement, role model), and programming self-efficacy (role model) during both the game and the editor. Role model badges were particularly effective during the game making task.

Our results have implications for both play and making. We find that role model badges improve virtually all facets of experience (player experience, intrinsic motivation, self-efficacy) relative to other badge types. This effectiveness was particularly relevant during the game making task. Achievement badges were found to be effective in the game, which would corroborate previous work, e.g., [114, 7, 140, 39]. Personal interest badges were found to only improve avatar identification during the game making task. Both choice conditions did not appear to be effective—possibly indicative of too much choice, or that the choice between badge types was simply not a meaningful one [56, 102, 158, 50].

Badges appeared to differ in their effectiveness based on the game or the editor. Therefore, it's likely that task context is a moderator—e.g., achievement badges earned during the game are less effective in the editor, whereas role model badges may generalize across the two. We also note that for all conditions except the Choice Always Visible condition, badges were only briefly visible after having completed a level in game. It's possible that having badges visible at all times—as during the editor—would further reinforce badge effects.

Additionally, we found that avatar identification positively affects all measures (player experience, intrinsic motivation, programming self-efficacy, overall time) in both play and making. Furthermore, avatar identification leads to higher quality completed maps. Therefore, both badges and avatar identification affect a variety of play and making related outcomes.

One caveat, however, is that wishful identification was actually negatively correlated to map quality. Wishful identification—or wanting to be like a fictional or media character [53, 76, 77, 121]—is correlated with lower psychological well-being [15, 75, 139]. However, wishful identification may be beneficial for self-esteem [47, 15]. This two-sided nature of wishful identification was expressed here as a universally positive effect on outcomes, *except* on actual game level quality. More work on wishful identification is needed to precisely characterize why this was the case.

Our R^2 values range between small (0.01), medium (0.09), and large (0.25) effect sizes. Avatar identification was particularly predictive of player experience, e.g., similarity identification (10.8%), embodied identification (39.4%), and wishful identification (25.7%) were all highly predictive of immersion during game making. Our mean significant R^2 value is 5.9% which we've demonstrated at a scale of N=2189 across many different outcomes, suggesting that avatar identification is an important component to our play and making experience.

Badges Applications

Our results suggest role model badges are effective—similarly to role model avatars [92]—yet badges in contrast may have more general application. We might imagine scientific games that leverage the crowd (e.g., FoldIt [103]), Massive Open Online Courses (MOOCs), digital learning platforms, etc. as being possible beneficiaries of these badges.

Literature suggests that role models are useful outside of academic contexts (e.g., [31])—as long as they are relevant [120]. Therefore, other domains such as business (Steve Jobs, etc.), politics (Barack Obama, etc.), health (Oprah Winfrey, etc.), may also benefit from these badges—a game, an educational platform, a gamified app, etc.—so long as the role models meet the criteria of perceived competence, similarity, and success.

Avatar Applications

Our results suggest that avatar identification can improve time on task, and positively impact player experience, intrinsic motivation, and self-efficacy. Applications range from enhancing motivation in crowdsourced tasks, to serious contexts such as behavior change, education, etc. We extend previous work [16] by showing that avatar identification can impact the quality of created levels. It remains an open and interesting question as to whether the production of other artifacts can also be similarly affected through identification with an onscreen representation—e.g., writing an essay, programming an application, designing a graphic, etc.

With increasing emphasis on making as a pedagogical method, there is ongoing concern about the quality of produced artifacts [89, 38]. Critics of Mario Maker, for instance, have condemned the majority of user created levels as being impossible, gimmicky, and "bafflingly opaque, frenzied contraptions that rarely seem to have a purpose" [74, 174, 38]. Here, we have made a first step towards understanding badges and avatar identification in relation to creation.

LIMITATIONS

Our study consisted of a period of time on the order of hours. However, the interaction between a user and a game often extends into long-term (e.g., in World of Warcraft [118]). Therefore, a longitudinal study could elucidate how our results are moderated by longer term use.

We also took one specific approach to studying badges. For example, we decided that making the badges black and white (to control for color confounds [61, 133, 134]) was a necessary price to pay. However, future studies could introduce color in a controlled way to further understand how color can moderate our findings. Another example was in how we implemented role model badges. We were cautious to ensure that our badge creation process yielded: 1) role models perceived as competent [128], 2) role models perceived as ingroup [119, 127], and 3) role models perceived as successful [25]—the lattermost was additionally reinforced with in-game text. However, how much deviation from these criteria that can still result in effective role model badges remains to be explored.

Figure 10. Example maps rated overall 2 (left), 4 (center), and 6 (right).

CONCLUSION

In this study, we have looked at how badges and avatar identification impact both play and making in an educational game. We found that certain badges could promote avatar identification (personal interest, role model), player experience (achievement, role model), intrinsic motivation (achievement, role model), and programming self-efficacy (role model) during both the game and the editor.

Avatar identification promoted player experience, intrinsic motivation, programming self-efficacy, and the total time spent playing and making. Avatar identification also promoted other meaningful in-editor activity, such as playtesting time, etc. and led to significantly higher overall quality of the completed game levels (as rated by 3 independent externally trained QA testers). Here, we've conducted a first study (N=2189) on alternative badge types, and a first study of badges and avatar identification in a making context. These findings contribute to both the literature on badges and avatars.

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their valuable feedback. Authors were supported by NSERC, the NSF (#1542970), and QCRI-CSAIL.

REFERENCES

1. Samuel Abramovich, Christian Schunn, and Ross Mitsuo Higashi. 2013. Are badges useful in education?: It depends upon the type of badge and expertise of learner. *Educational Technology Research and Development* 61, 2 (2013), 217–232. DOI: http://dx.doi.org/10.1007/s11423-013-9289-2
2. Samuel Abramovich and Peter Wardrip. 2016. *Impact of Badges on Motivation to Learn*. Routledge.
3. Thomas Agoritsas, Emma Iserman, Nicholas Hobson, Natasha Cohen, Adam Cohen, Pavel S Roshanov, Miguel Perez, Chris Cotoi, Rick Parrish, Eleanor Pullenayegum, and Others. 2014. Increasing the quantity and quality of searching for current best evidence to answer clinical questions: protocol and intervention design of the MacPLUS FS Factorial Randomized Controlled Trials. *Implementation Science* 9, 1 (2014), 125.
4. June Ahn, Anthony Pellicone, and Brian S. Butler. 2014. Open badges for education: What are the implications at the intersection of open systems and badging? *Research in Learning Technology* 22, 1063519 (2014), 1–13. DOI: http://dx.doi.org/10.3402/rlt.v22.23563
5. Glen S Aikenhead. 2006. *Science education for everyday life: Evidence-based practice*. Teachers College Press.
6. Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2013. Steering user behavior with badges. *WWW '13 Proceedings of the 22nd international conference on World Wide Web* (2013), 95–106. DOI: http://dx.doi.org/10.1145/2488388.2488398
7. Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2014. Engaging with massive online courses. In *Proceedings of the 23rd international conference on World wide web*. ACM, 687–698.
8. Judd Antin and Elizabeth F Churchill. 2011. Badges in social media: A social psychological perspective. In *CHI 2011 Gamification Workshop Proceedings*. ACM New York, NY, 1–4.
9. Christine M. Bachen, Pedro Hernández-Ramos, Chad Raphael, and Amanda Waldron. 2016. How do presence, flow, and character identification affect players' empathy and interest in learning from a serious computer game? *Computers in Human Behavior* 64 (2016), 77–87. DOI: http://dx.doi.org/10.1016/j.chb.2016.06.043
10. Jeremy N. Bailenson, Jim Blascovich, and Rosanna E. Guadagno. 2008. Self-representations in immersive virtual environments. *Journal of Applied Social Psychology* 38, 11 (2008), 2673–2690.
11. Rachel Bailey, Kevin Wise, and Paul Bolls. 2009. How Avatar Customizability Affects Children's Arousal and Subjective Presence During Junk Food-Sponsored Online Video Games. *CyberPsychology & Behavior* (2009), 1–9. http://online.liebertpub.com/doi/abs/10.1089/cpb.2008.0292
12. Albert Bandura. 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review* 84, 2 (1977), 191.
13. Roy F Baumeister and Mark R Leary. 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin* 117, 3 (1995), 497.
14. Al Baylor and Yanghee Kim. 2004. Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. *Intelligent Tutoring Systems* 1997 (2004), 592–603. DOI: http://dx.doi.org/10.1007/978-3-540-30139-4_56
15. K Bessière, AF Seay, and S Kiesler. 2007. The ideal elf: Identity exploration in World of Warcraft. *CyberPsychology & Behavior* (2007). http://online.liebertpub.com/doi/abs/10.1089/cpb.2007.9994
16. Max V Birk, Cheralyn Atkins, Jason T Bowey, and Regan L Mandryk. 2016a. Fostering Intrinsic Motivation through Avatar Identification in Digital Games. *CHI* (2016). DOI: http://dx.doi.org/10.1145/2858036.2858062
17. Max V Birk, Regan L Mandryk, and Cheralyn Atkins. 2016b. The Motivational Push of Games. *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '16* April 2017 (2016), 291–303. DOI: http://dx.doi.org/10.1145/2967934.2968091
18. Sanat Kumar Bista, Surya Nepal, and Cecile Paris. 2012. Engagement and Cooperation in Social Networks: Do Benefits and Rewards Help?. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*. IEEE, 1405–1410.
19. Lucas Blair. 2016. What Video Games Can Teach Us About Badges and Pathways. *Digital Badges in Education: Trends, Issues, and Cases* (2016), 62.
20. Ian Bogost. 2011. Gamification is bullshit. *The gameful world: Approaches, issues, applications* (2011), 65–80.
21. Ian Bogost. 2013. Exploitationware. In *Rhetoric/composition/play through video games*. Springer, 139–147.
22. Ivica Boticki, Jelena Baksa, Peter Seow, and Chee-Kit Looi. 2015. Usage of a mobile social learning platform with virtual badges in a primary school. *Computers & Education* 86 (2015), 120–136.
23. Nicholas David Bowman, Mary Beth Oliver, Ryan Rogers, Brett Sherrick, Julia Woolley, and Mun-Young Chung. 2016. In control or in their shoes? How character attachment differentially influences video game enjoyment and appreciation. *Journal of Gaming & Virtual Worlds* 8, 1 (2016), 83–99. DOI: http://dx.doi.org/10.1386/jgvw.8.1.83_1

24. Karen Brennan and Mitchel Resnick. 2012. New frameworks for studying and assessing the development of computational thinking. *Annual American Educational Research Association meeting, Vancouver, BC, Canada* (2012), 1–25.

25. Abraham P. Buunk, José Maria Peiró, and Chris Griffioen. 2007. A positive role model may stimulate career-oriented behavior. *Journal of Applied Social Psychology* 37 (2007), 1489–1500. DOI: http://dx.doi.org/10.1111/j.1559-1816.2007.00223.x

26. Donn Byrne and Don Nelson. 1965. Attraction as a linear function of proportion of positive reinforcements. *Journal of Personality and Social Psychology* 1, 6 (1965), 659–663. DOI: http://dx.doi.org/10.1037/h0022073

27. Craig Calhoun. 2010. *Robert K. Merton: sociology of science and sociology as science*. Columbia University Press.

28. Wk Keith Campbell and Constantine Sedikides. 1999. Self-threat magnifies the self-serving bias: A meta-analytic integration. (1999). DOI: http://dx.doi.org/10.1037/1089-2680.3.1.23

29. Valery Chirkov, Richard M Ryan, Youngmee Kim, and Ulas Kaplan. 2003. Differentiating autonomy from individualism and independence: a self-determination theory perspective on internalization of cultural orientations and well-being. *Journal of personality and social psychology* 84, 1 (2003), 97.

30. Klimmt Christoph, Hefner Dorothée, and Vorderer Peter. 2009. The Video Game Experience as "True" Identification: A Theory of Enjoyable Alterations of Players' Self-Perception. *Communication theory* 19, 4 (2009), 351–373.

31. Paul W Clark, Craig A Martin, and Alan J Bush. 2001. The effect of role model influence on adolescents' materialism and marketplace knowledge. *Journal of Marketing Theory and Practice* 9, 4 (2001), 27–36.

32. Jonathan Cohen. 2001. Defining identification: A theoretical look at the identification of audiences with media characters. *Mass communication & society* 4, 3 (2001), 245–264.

33. Jonathan Cohen. 2002. Deconstructing Ally: Explaining viewers' interpretations of popular television. *Media Psychology* 4, 3 (2002), 253–277.

34. Conrad's Home. 2017. Mii Editor. (2017). http://www.conradshome.com/mii/

35. Diana I. Cordova and Mark R. Lepper. 1996. Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology* 88, 4 (1996), 715–730. DOI: http://dx.doi.org/10.1037/0022-0663.88.4.715

36. Edward L Deci, Richard Koestner, and Richard M Ryan. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. (1999).

37. Edward L Deci and Richard M Ryan. 2012. Motivation, personality, and development within embedded social contexts: An overview of self-determination theory. *The Oxford handbook of human motivation* (2012), 85–107.

38. André R. Denham and Kelly W. Guyotte. 2017. Cultivating critical game makers in digital game-based learning: learning from the arts. *Learning, Media and Technology* 9884, July (2017), 1–11. DOI: http://dx.doi.org/10.1080/17439884.2017.1342655

39. Paul Denny. 2013. The effect of virtual achievements on student engagement. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), 763. DOI: http://dx.doi.org/10.1145/2470654.2470763

40. Sebastian Deterding. 2011. Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using game design elements in non-gaming contexts, a workshop at CHI.*

41. Sebastian Deterding. 2012. Coding conduct: Games, play, and human conduct between technical artifacts and social framing. (2012). https://www.slideshare.net/dings/coding-conduct-games-play-and-human-conduct-between-technical-code-and-social-framing

42. Sebastian Deterding. 2015. The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. *Human-Computer Interaction* 30, 3-4 (2015), 294–335. DOI:http://dx.doi.org/10.1080/07370024.2014.993471

43. Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From game design elements to gamefulness. *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11* (2011), 9–11. DOI: http://dx.doi.org/10.1145/2181037.2181040

44. Darina Dicheva, Christo Dichev, Gennady Agre, and Galia Angelova. 2015. Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society* 18, 3 (2015), 75–88. DOI: http://dx.doi.org/10.1109/EDUCON.2014.6826129.

45. Adrián Domìnguez, Joseba Saenz-De-Navarrete, Luis De-Marcos, Luis Fernández-Sanz, Carmen Pagés, and José-Javier Martìnez-HerráIz. 2013. Gamifying learning experiences: Practical implications and outcomes. *Computers & Education* 63 (2013), 380–392.

46. Rosalind Driver, Hilary Asoko, John Leach, Philip Scott, and Eduardo Mortimer. 1994. Constructing scientific knowledge in the classroom. *Educational researcher* 23, 7 (1994), 5–12.

47. Nicolas Ducheneaut, MH Wen, Nicholas Yee, and Greg Wadley. 2009. Body and mind: a study of avatar personalization in three virtual worlds. *CHI 2009* (2009). http://dl.acm.org/citation.cfm?id=1518877

48. Thomas Shelley Duval and Paul J Silvia. 2002. Self-awareness, probability of improvement, and the self-serving bias. *Journal of personality and social psychology* 82, 1 (2002), 49–61. DOI: http://dx.doi.org/10.1037/0022-3514.82.1.49

49. Andrew J Elliot, Markus a Maier, Arlen C Moller, Ron Friedman, and Jörg Meinhardt. 2007. Color and psychological functioning: the effect of red on performance attainment. *Journal of experimental psychology. General* 136, 1 (2007), 154–168. DOI: http://dx.doi.org/10.1037/0096-3445.136.1.154

50. Miriam Evans and Alyssa R. Boucher. 2015. Optimizing the power of choice: Supporting student autonomy to foster motivation and engagement in learning. *Mind, Brain, and Education* 9, 2 (2015), 87–91. DOI:http://dx.doi.org/10.1111/mbe.12073

51. Nickolas J G Falkner and Katrina E Falkner. 2013. "Whither, Badges?" or "Wither, Badges!": A Metastudy of Badges in Computer Science Education to Clarify Effects , Significance and Influence. *Computer and Education* 50, 1 (2013), 127–135. DOI: http://dx.doi.org/10.1002/meet.14505001101

52. Rosta Farzan, Joan M DiMicco, David R Millen, Casey Dugan, Werner Geyer, and Elizabeth A Brownholtz. 2008. Results from deploying a participation incentive mechanism within the enterprise. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 563–572.

53. C. V. Feilitzen and Olga Linné. 1975. Identifying with television characters. *Journal of Communication* 25, 4 (1975), 51–55.

54. Michael Filsecker and Daniel Thomas Hickey. 2014. A multilevel analysis of the effects of external rewards on elementary students' motivation, engagement and learning in an educational game. *Computers and Education* 75, August (2014), 136–148. DOI: http://dx.doi.org/10.1016/j.compedu.2014.02.008

55. Zachary Fitz-Walter, Dian Tjondronegoro, and Peta Wyeth. 2011. Orientation passport: using gamification to engage university students. In *Proceedings of the 23rd Australian computer-human interaction conference*. ACM, 122–125.

56. Terri Flowerday and Gregory Schraw. 2000. Teacher beliefs about instructional choice: A phenomenological study. *Journal of Educational Psychology* 92, 4 (2000), 634–645. DOI: http://dx.doi.org/10.1037/0022-0663.92.4.634

57. Aroutis Foster. 2008. Games and Motivation to Learn Science: Personal Identity, Applicability, Relevance and Meaningfulness. *Journal of Interactive Learning Research* 19, 4 (2008), 597–614. http://ezproxy.deakin.edu.au/login?url=http://search.proquest.com/docview/62000189?accountid=10445

58. Jordan Harris Frith. 2012. *Constructing location, one check-in at a time: examining the practices of Foursquare users*. North Carolina State University.

59. Tracy Fullerton. 2014. *Game design workshop: a playcentric approach to creating innovative games*. CRC press.

60. David Gibson, Nathaniel Ostashewski, Kim Flintoff, Sheryl Grant, and Erin Knight. 2013. Digital badges in education. *Education and Information Technologies* May 2016 (2013), 1–8. DOI: http://dx.doi.org/10.1007/s10639-013-9291-7

61. Timo Gnambs, Markus Appel, and Bernad Batinic. 2010. Color red in web-based knowledge testing. *Computers in Human Behavior* 26, 6 (2010), 1625–1631. DOI:http://dx.doi.org/10.1016/j.chb.2010.06.010

62. Geoff Goehle. 2013. Gamification and web-based homework. *Primus* 23, 3 (2013), 234–246.

63. Melvin E Gonnerman Jr, Christopher P Parker, Howard Lavine, and Joseph Huff. 2000. The relationship between self-discrepancies and affective states: The moderating roles of self-monitoring and standpoints on the self. *Personality and Social Psychology Bulletin* 26, 7 (2000), 810–819.

64. Rosanna E Guadagno, Jim Blascovich, Jeremy N Bailenson, and Cade McCall. 2007. Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology* 10, 1 (2007), 1–22. DOI: http://dx.doi.org/10.108/15213260701300865

65. Jérôme Guegan, Stéphanie Buisine, Fabrice Mantelet, Nicolas Maranzana, and Frédéric Segonds. 2016. Avatar-mediated creativity: When embodying inventors makes engineers more creative. *Computers in Human Behavior* 61 (2016), 165–175. DOI: http://dx.doi.org/10.1016/j.chb.2016.03.024

66. Lasse Hakulinen, Tapio Auvinen, and Ari Korhonen. 2013. Empirical study on the effect of achievement badges in TRAKLA2 online learning environment. *Proceedings - 2013 Learning and Teaching in Computing and Engineering, LaTiCE 2013* (2013), 47–54. DOI:http://dx.doi.org/10.1109/LaTiCE.2013.34

67. Juho Hamari. 2011. Framework for Designing and Evaluating Game Achievements. *Proceedings of DiGRA 2011 Conference: Think Design Play* (2011), 20. DOI: http://dx.doi.org/10.1.1.224.9966

68. Juho Hamari. 2013. Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications* 12, 4 (2013), 236–245. DOI: http://dx.doi.org/10.1016/j.elerap.2013.01.004

69. Juho Hamari. 2017. Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior* 71 (2017), 469–478. DOI:http://dx.doi.org/10.1016/j.chb.2015.03.036

70. Juho Hamari, Jonna Koivisto, and Harri Sarsa. 2014. Does gamification work? - A literature review of empirical studies on gamification. *Proceedings of the Annual Hawaii International Conference on System Sciences* (2014), 3025–3034. DOI: http://dx.doi.org/10.1109/HICSS.2014.377

71. JG Hamilton. 2009. Identifying with an avatar: a multidisciplinary perspective. *Proceedings of the Cumulus Conference: 38º South: Hemispheric Shifts Across Learning, Teaching and Research* November (2009), 1–14. http://eprints.qut.edu.au/29701/

72. Johannes Harms, Dominik Seitz, Christoph Wimmer, Karin Kappel, and Thomas Grechenig. 2015. Low-Cost Gamification of Online Surveys : Improving the User Experience through Achievement Badges. *CHI PLAY'15 Proceedings of the 2nd ACM SIGCHI annual symposium on Computer-human interaction in play.* (2015), 109–113. DOI: http://dx.doi.org/10.1145/2793107.2793146

73. Dorothée Hefner, Christoph Klimmt, and Peter Vorderer. 2007. Identification with the Player Character as Determinant of Video Game Enjoyment. In *Entertainment Computing - ICEC 2007*. 39–48. DOI: http://dx.doi.org/10.1007/978-3-540-74873-1_6

74. Patricia Hernandez. 2015. The Most Popular Levels In Super Mario Maker (So Far). (2015). http://kotaku.com/the-most-popular-levels-in-super-mario-maker-so-far-1730170457

75. E Tory Higgins. 1987. Self-discrepancy: a theory relating self and affect. *Psychological review* 94, 3 (1987), 319.

76. Cynthia Hoffner. 1996. Children's wishful identification and parasocial interaction with favorite television characters. *Journal of Broadcasting & Electronic Media* 40, 3 (1996), 389–402.

77. Cynthia Hoffner and Martha Buchanan. 2005. Young adults' wishful identification with television characters: The role of perceived similarity and character attributes. *Media psychology* 7, 4 (2005), 325–351.

78. Rosalie Hooi and Hichang Cho. 2013. Deception in avatar-mediated virtual environment. *Computers in Human Behavior* 29, 1 (2013), 276–284.

79. Yukari Hori, Yutaka Tokuda, Takahiro Miura, Atsushi Hiyama, and Michitaka Hirose. 2013. Communication pedometer: a discussion of gamified communication focused on frequency of smiles. In *Proceedings of the 4th augmented human international conference*. ACM, 206–212.

80. Bart Hulshof. 2013. The influence of colour and scent on people's mood and cognitive performance in meeting rooms. *Master Thesis* May (2013), 1–97.

81. Katherine Isbister and Clifford Nass. 2000. Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies* 53 (2000), 251–267. DOI:http://dx.doi.org/10.1006/ijhc.2000.0368

82. Mizuko Ito. 2012. *Engineering play: A cultural history of children's software*. MIT Press.

83. Mikael Jakobsson. 2011. The achievement machine: Understanding Xbox 360 achievements in gaming practices. *Game Studies* 11, 1 (2011).

84. Ruth Jarman. 2005. Science learning through scouting: an understudied context for informal science education. *International Journal of Science Education* 27, 4 (2005), 427–450.

85. Seung-A Annie Jin. 2010. "I Feel More Connected to the Physically Ideal Mini Me than the Mirror-Image Mini Me": Theoretical Implications of the "Malleable Self" for Speculations on the Effects of Avatar Creation on Avatar–Self Connection in Wii. *Cyberpsychology, Behavior, and Social Networking* 13, 5 (2010), 567–570.

86. Kristine Jolivette, Joseph H Wehby, Jennifer Canale, and N Gayle Massey. 2001. Effect of choice-making opportunities on the behavior of students with emotional and behavioral disorders. *Behavioral Disorders* 26, 2 (2001), 131–145. http://proxy2.lib.umanitoba.ca/login?url=http://search.proquest.com/docview/620004188?accountid=14569

87. Barry Joseph. 2012. Six ways to look at badging systems designed for learning. *Global Kids: Online Leadership Program,[online] Available at: http://www. olpglobalkids. org/content/six-ways-look-badging-systems-designed-learning* (2012).

88. Yasmin B. Kafai and Quinn Burke. 2015a. Constructionist Gaming: Understanding the Benefits of Making Games for Learning. *Educational Psychologist* 50, 4 (2015), 313–334. DOI: http://dx.doi.org/10.1080/00461520.2015.1124022

89. Yasmin B. Kafai and Quinn Burke. 2015b. Constructionist Gaming: Understanding the Benefits of Making Games for Learning. *Educational Psychologist* (2015).

90. Dominic Kao. 2017. MazeStar: A Platform for Studying Virtual Identity and Computer Science Education. In *Foundations of Digital Games ASETGC Workshop 2017*.

91. Dominic Kao and D. Fox Harrell. 2015a. Exploring construction, play, use of virtual identities in STEM learning. *Jean Piaget Society Annual Conference* (2015).

92. Dominic Kao and D. Fox Harrell. 2015b. Exploring the Impact of Role Model Avatars on Game Experience in Educational Games. *The ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play (CHI PLAY)* (2015).

93. Dominic Kao and D. Fox Harrell. 2015c. Exploring the Use of Role Model Avatars in Educational Games. In *Proceedings of the AIIDE Workshop on Experimental AI in Games, co-located with Artificial Intelligence in Interactive Digital Entertainment 2015*.

94. Dominic Kao and D. Fox Harrell. 2015d. Mazzy: A STEM Learning Game. *Foundations of Digital Games* (2015).
95. Dominic Kao and D. Fox Harrell. 2015e. Toward Avatar Models to Enhance Performance and Engagement in Educational Games. *IEEE Computational Intelligence in Games* (2015).
96. Dominic Kao and D. Fox Harrell. 2015f. Toward Evaluating the Impacts of Virtual Identities on STEM Learning. *Foundations of Digital Games* (2015).
97. Dominic Kao and D Fox Harrell. 2016a. Exploring the Effects of Dynamic Avatars on Performance and Engagement in Educational Games. In *Games+Learning+Society (GLS 2016)*.
98. Dominic Kao and D. Fox Harrell. 2016b. Exploring the Effects of Encouragement in Educational Games. *Proceedings of the 34th Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI 2016)* (2016).
99. Dominic Kao and D. Fox Harrell. 2016c. Exploring the Impact of Avatar Color on Game Experience in Educational Games. *Proceedings of the 34th Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI 2016)* (2016).
100. Dominic Kao and D. Fox Harrell. 2016d. Toward Understanding the Impacts of Role Model Avatars on Engagement in Computer Science Learning. In *The annual meeting of the American Educational Research Association (AERA)*.
101. Dominic Kao and D. Fox Harrell. 2017. Toward Understanding the Impact of Visual Themes and Embellishment on Performance, Engagement, and Self-Efficacy in Educational Games. *The annual meeting of the American Educational Research Association (AERA)* (2017).
102. Idit Katz and Avi Assor. 2007. When choice motivates and when it does not. *Educational Psychology Review* 19, 4 (2007), 429–442. DOI: http://dx.doi.org/10.1007/s10648-006-9027-y
103. Firas Khatib, Frank DiMaio, Seth Cooper, Maciej Kazmierczyk, Miroslaw Gilski, Szymon Krzywda, Helena Zabranska, Iva Pichova, James Thompson, Zoran Popović, and Others. 2011. Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature structural & molecular biology* 18, 10 (2011), 1175–1177.
104. Changsoo Kim, Sang Gun Lee, and Minchoel Kang. 2012. I became an attractive person in the virtual world: Users' identification with virtual communities and avatars. *Computers in Human Behavior* 28, 5 (2012), 1663–1669. DOI: http://dx.doi.org/10.1016/j.chb.2012.04.004
105. Yanghee Kim and Amy L. Baylor. 2006. Pedagogical agents as learning companions: The role of agent competency and type of interaction. *Educational Technology Research and Development* 54, 3 (2006), 223–243.
106. Youjeong Kim and S Shyam Sundar. 2012. Visualizing ideal self vs. actual self through avatars: Impact on preventive health outcomes. *Computers in Human Behavior* 28, 4 (2012), 1356–1364.
107. CK Kinzer, D Hoffman, and S Turkay. 2012. The impact of choice and feedback on learning, motivation, and performance in an educational video game. *GLS* 2 (2012), 175–181. http://gamesresearchlab.com/wp-content/uploads/2015/11/The
108. Christoph Klimmt. 2003. Dimensions and determinants of the enjoyment of playing digital games: A three-level model. In *Level up: Digital games research conference*. 246–257.
109. Elly A Konijn, Marije Nije Bijvank, and Brad J Bushman. 2007. I wish I were a warrior: the role of wishful identification in the effects of violent video games on aggression in adolescent boys. *Developmental psychology* 43, 4 (2007), 1038.
110. Okhee Lee and Aurolyn Luykx. 2006. *Science education and student diversity: Synthesis and research agenda*. Cambridge University Press.
111. Zakkoyya Lewis, Maria C. Swartz, and Elizabeth Lyons. 2016. What's the Point?: A Review of Reward Systems Implemented in Gamification Interventions. *Games for Health Journal* 5, 2 (2016), 93–99. DOI: http://dx.doi.org/10.1089/g4h.2015.0078
112. Benjamin J. Li and May O. Lwin. 2016. Player see, player do: Testing an exergame motivation model based on the influence of the self avatar. *Computers in Human Behavior* 59 (2016), 350–357. DOI: http://dx.doi.org/10.1016/j.chb.2016.02.034
113. Dong Dong Li, Albert Kien Liau, and Angeline Khoo. 2013. Player–Avatar Identification in video gaming: Concept and measurement. *Computers in Human Behavior* 29, 1 (2013), 257–263.
114. Zhuolun Li, Ke-Wei Huang, and Huseyin Cavusoglu. 2012. Quantifying the impact of badges on user engagement in online Q&A communities. (2012).
115. Sohye Lim and Byron Reeves. 2009. Being in the Game: Effects of Avatar Choice and Point of View on Psychophysiological Responses During Play. *Media Psychology* 12, 4 (2009), 348–370. DOI: http://dx.doi.org/10.1080/15213260903287242
116. Raymund J Lin and Xinxin Zhu. 2012. Leveraging social media for preventive care-A gamification system and insights. *Studies in health technology and informatics* 180 (2012), 838–842.

117. Chris J Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, M Jordan Raddick, Robert C Nichol, Alex Szalay, Dan Andreescu, and Others. 2008. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society* 389, 3 (2008), 1179–1189.

118. Ian J. Livingston, Carl Gutwin, Regan L. Mandryk, and Max Birk. 2014. How players value their characters in world of warcraft. *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14* (2014), 1333–1343. DOI: http://dx.doi.org/10.1145/2531602.2531661

119. Penelope Lockwood. 2006. "Someone like me can be successful": Do college students need same-gender role models? *Psychology of Women Quarterly* 30, 1 (mar 2006), 36–46. DOI: http://dx.doi.org/10.1111/j.1471-6402.2006.00260.x

120. Penelope Lockwood and Ziva Kunda. 1997. Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology* 73, 1 (1997), 91–103. DOI: http://dx.doi.org/10.1037//0022-3514.73.1.91

121. Subhash C Lonial and Stuart Van Auken. 1986. Wishful identification with fictional characters: An assessment of the implications of gender in message dissemination to children. *Journal of Advertising* 15, 4 (1986), 4–42.

122. MacArthur Foundation. 2013. Better Futures for 2 Million Americans Through Open Badges. (2013).

123. Eleanor E. Maccoby and William Cody Wilson. 1957. Identification and observational learning from films. *The Journal of abnormal and social psychology* 55, 1 (1957), 76. DOI:http://dx.doi.org/10.1037/h0043015

124. Lena Mamykina, Bella Manoim, Manas Mittal, George Hripcsak, and Björn Hartmann. 2011. Design lessons from the fastest q&a site in the west. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2857–2866.

125. Andrew Marder. 2015. Stack overflow badges and user behavior: An econometric approach. *IEEE International Working Conference on Mining Software Repositories* 2015-Augus (2015), 450–453. DOI: http://dx.doi.org/10.1109/MSR.2015.61

126. James E Martin, Dennis E Mithaug, Phil Cox, Lori Y Peterson, Jamie L Van Dycke, and Mary E Cash. 2003. Increasing self-determination: Teaching students to plan, work, evaluate, and adjust. *Exceptional Children* 69, 4 (2003), 431–447. DOI: http://dx.doi.org/doi:10.1177/001440290306900403

127. D. M. Marx and J. S. Roman. 2002. Female Role Models: Protecting Women's Math Test Performance. *Personality and Social Psychology Bulletin* 28 (2002), 1183–1193. DOI: http://dx.doi.org/10.1177/01461672022812004

128. David M Marx, Diederik a Stapel, and Dominique Muller. 2005. We can do it: the interplay of construal orientation and social comparisons under threat. *Journal of personality and social psychology* 88, 3 (2005), 432–446. DOI: http://dx.doi.org/10.1037/0022-3514.88.3.432

129. Edward McAuley, Terry Duncan, and Vance V Tammen. 1989. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport* 60, 1 (1989), 48–58.

130. Rudy McDaniel, Robb Lindgren, and Jon Friskics. 2012. Using badges for shaping interactions in online learning environments. In *Professional Communication Conference (IPCC), 2012 IEEE International*. IEEE, 1–4.

131. Daniel G McDonald and Hyeok Kim. 2001. When I die, I feel small: Electronic game characters and the social self. *Journal of Broadcasting & Electronic Media* 45, 2 (2001), 241–258.

132. Rusty B Mcintyre, Charles G Lord, Dana M Gresky, Laura L Ten Eyck, and Charles F Bond. 2005. A Social Impact Trend in the Effects of Role Models on Alleviating Women's Mathematics Stereotype Threat. *Current Research in Social Psychology* 10, 9 (2005), 1–26.

133. Ravi Mehta and Rui(Juliet) Zhu. 2008. Blue or Red? Exploring the Effect of Color on Cognitive Task Performances. *Science* 323, February (2008), 1226–1229. DOI: http://dx.doi.org/10.1126/science.1169144

134. M.A. Meier, Russell A. Hill, Andrew J. Elliot, and R.A. Barton. 2015. Color in Achievement Contexts in Humans. *Handbook of Color Psychology* 44, February (2015), 0–103. DOI:http://dx.doi.org/10.1063/1.2756072

135. Jason P. Mitchell, C. Neil Macrae, and Mahzarin R. Banaji. 2006. Dissociable Medial Prefrontal Contributions to Judgments of Similar and Dissimilar Others. *Neuron* 50, 4 (2006), 655–663. DOI: http://dx.doi.org/10.1016/j.neuron.2006.03.040

136. Dean Mobbs, Rongjun Yu, Marcel Meyer, Luca Passamonti, Ben Seymour, Andrew J Calder, Susanne Schweizer, Chris D Frith, and Tim Dalgleish. 2009. A key role for similarity in vicarious reward. *Science* 324, 5929 (2009), 900. DOI: http://dx.doi.org/10.1126/science.1170539

137. Markus Montola, Timo Nummenmaa, Andrés Lucero, Marion Boberg, and Hannu Korhonen. 2009. Applying game achievement systems to enhance user experience in a photo sharing service. *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era* 4, 46 (2009), 94–97. DOI: http://dx.doi.org/10.1145/1621841.1621859

138. Youngme Moon. 2000. Intimate Exchanges: Using Computers to Elicit Self-Disclosure From Consumers. *Journal of Consumer Research* 26, 4 (2000), 323–339. DOI:http://dx.doi.org/10.1086/209566

139. Marlene M Moretti and E Tory Higgins. 1990. Relating self-discrepancy to self-esteem: The contribution of discrepancy beyond actual-self ratings. *Journal of Experimental Social Psychology* 26, 2 (1990), 108–123.

140. Briana B Morrison and Betsy DiSalvo. 2014. Khan academy gamifies computer science. In *Proceedings of the 45th ACM technical symposium on Computer science education*. ACM, 39–44.

141. Emily Moyer-Gusé, Adrienne H Chung, and Parul Jain. 2011. Identification with characters and discussion of taboo topics after exposure to an entertainment narrative about sexual health. *Journal of Communication* 61, 3 (2011), 387–406.

142. Mozilla Foundation, Peer 2 Peer University, and MacArthur Foundation. 2013. Open Badges for Lifelong Learning. (2013).

143. Raymond Ng and Robb Lindgren. 2013. Examining the effects of avatar customization and narrative on engagement and learning in video games. *Proceedings of CGAMES 2013 USA - 18th International Conference on Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational and Serious Games* (2013), 87–90. DOI: http://dx.doi.org/10.1109/CGames.2013.6632611

144. Steffen P. Walz and Sebastian Deterding. 2015. *The Gameful World - Approaches, Issues, Applications*. DOI: http://dx.doi.org/10.1017/CBO9781107415324.004

145. S Papert and Idit Harel. 1991. Situating Constructionism. *Constructionism* (1991).

146. Bong-Won Park and Kun Chang Lee. 2011. Exploring the value of purchasing online game items. *Computers in Human Behavior* 27, 6 (2011), 2178–2185.

147. Marco Piccioni, Christian Estler, and Bertrand Meyer. 2014. SPOC-supported introduction to programming. In *Proceedings of the 2014 conference on Innovation & technology in computer science education*. ACM, 3–8.

148. Jan L Plass, Paul A O'Keefe, Melissa L Biles, Jonathan Frye, and Bruce D Homer. 2014. Motivational and Cognitive Impact of Badges in Games for Learning. October 2015 (2014). DOI: http://dx.doi.org/10.13140/2.1.3209.9842

149. Andrew K Przybylski, Netta Weinstein, Kou Murayama, Martin F Lynch, and Richard M Ryan. 2012. The ideal self at play: The appeal of video games that let you be all you can be. *Psychological science* 23, 1 (2012), 69–76.

150. Vennila Ramalingam and Susan Wiedenbeck. 1998. Development and Validation of Scores on a Computer Programming Self-Efficacy Scale and Group Analyses of Novice Programmer Self-Efficacy. *Journal of Educational Computing Research* 19, 4 (1998), 367–381. DOI:http://dx.doi.org/10.2190/C670-Y3C8-LTJ1-CT3P

151. Rabindra Ratan and Young June Sah. 2015. Leveling up on stereotype threat: The role of avatar customization and avatar embodiment. *Computers in Human Behavior* 50 (2015), 367–374. DOI: http://dx.doi.org/10.1016/j.chb.2015.04.010

152. Michael Resnick. 2012. Still a badge skeptic. *Retrieved June* 2 (2012), 2015.

153. M Resnick and John Maloney. 2009. Scratch: programming for all. *Communications of the ...* (2009). http://dl.acm.org/citation.cfm?id=1592779

154. Michael Restivo and Arnout Van De Rijt. 2012. Experimental study of informal rewards in peer production. *PloS one* 7, 3 (2012), e34358.

155. Mark Rice, Ranieri Koh, Quintessence Lui, Qixiang He, Marcus Wan, Vanessa Yeo, Jamie Ng, and Wah Pheow Tan. 2013. Comparing avatar game representation preferences across three age groups. *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13* (2013), 1161. DOI: http://dx.doi.org/10.1145/2468356.2468564

156. Scott Rigby and Richard M Ryan. 2011. *Glued to games: How video games draw us in and hold us spellbound: How video games draw us in and hold us spellbound*. ABC-CLIO.

157. Margaret Robertson. 2010. Can't play, won't play. *Hide & Seek* 6 (2010), 2010.

158. David H Rose and Anne Meyer. 2002. Teaching Every Student in the Digital Age: Universal Design for Learning. *Ericedgov* (2002), 216. DOI: http://dx.doi.org/10.1007/s11423-007-9056-3

159. K J Rose, M Koenig, and F Wiesbauer. 2013. Evaluating success for behavioral change in diabetes via mHealth and gamification: MySugr's keys to retention and patient engagement. *Diabetes Technology & Therapeutics* 15 (2013), A114.

160. José A. Ruipérez-Valiente, Pedro J. Muñoz-Merino, and Carlos Delgado Kloos. 2016. An analysis of the use of badges in an educational experiment. *Proceedings - Frontiers in Education Conference, FIE* 2016-Novem (2016), 1–8. DOI: http://dx.doi.org/10.1109/FIE.2016.7757424

161. Richard M Ryan and Edward L Deci. 2000a. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25, 1 (2000), 54–67.

162. R M Ryan and E L Deci. 2000b. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American psychologist* 55, 1 (2000), 68–78. DOI: http://dx.doi.org/10.1037/0003-066X.55.1.68

163. Richard M. Ryan, C. Scott Rigby, and Andrew Przybylski. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30, 4 (2006), 344–360. DOI: http://dx.doi.org/10.1007/s11031-006-9051-8

164. Michael Sailer, Jan Ulrich Hense, Sarah Katharina Mayr, and Heinz Mandl. 2017. How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior* 69 (2017), 371–380. DOI:http://dx.doi.org/10.1016/j.chb.2016.12.033

165. Katie Salen. 2011. *Quest to learn: Developing the school for digital kids*. MIT Press.

166. Katie Seaborn and Deborah I. Fels. 2015. Gamification in theory and action: A survey. *International Journal of Human Computer Studies* 74 (2015), 14–31. DOI: http://dx.doi.org/10.1016/j.ijhcs.2014.09.006

167. P E Shrout and J L Fleiss. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin* 86, 2 (1979), 420–428. DOI: http://dx.doi.org/10.1037/0033-2909.86.2.420

168. David Smahel, Lukas Blinka, and Ondrej Ledabyl. 2008. Playing MMORPGs: connections between addiction and identifying with a character. *Cyberpsychology & Behavior* 11, 6 (2008), 715–718. DOI: http://dx.doi.org/10.1089/cpb.2007.0210

169. Alistair Raymond Bryce Soutter and Michael Hitchens. 2016. The relationship between character identification and flow state within video games. *Computers in Human Behavior* 55, December 2015 (2016), 1030–1038. DOI: http://dx.doi.org/10.1016/j.chb.2015.11.012

170. Jocelyn Steinke, Brooks Applegate, Maria Lapinski, Lisa Ryan, and Marilee Long. 2012. Gender Differences in Adolescents' Wishful Identification With Scientist Characters on Television. *Science Communication* 34, 2 (2012), 163–199. DOI: http://dx.doi.org/10.1177/1075547011410250

171. Jennifer N Stinson, Lindsay A Jibb, Cynthia Nguyen, Paul C Nathan, Anne Marie Maloney, L Lee Dupuis, J Ted Gerstle, Benjamin Alman, Sevan Hopyan, Caron Strahlendorf, and Others. 2013. Development and testing of a multidimensional iPhone pain assessment application for adolescents with cancer. *Journal of medical Internet research* 15, 3 (2013).

172. Ching-I Teng. 2017. Impact of avatar identification on online gamer loyalty: Perspectives of social identity and social capital theories. *International Journal of Information Management* 37, 6 (2017), 601–610. DOI: http://dx.doi.org/10.1016/j.ijinfomgt.2017.06.006

173. Jennifer Thom, David Millen, and Joan DiMicco. 2012. Removing gamification from an enterprise SNS. In *Proceedings of the acm 2012 conference on computer supported cooperative work*. ACM, 1067–1070.

174. Michael Thomsen. 2015. "Super Mario Maker" is an engine for circulating horrible new "Mario" levels. (2015). https://www.washingtonpost.com/news/comic-riffs/wp/2015/09/15/super-mario-maker-is-an-engine-for-circulating-horrible-new-mario-levels/?utm

175. Sabine Trepte and Leonard Reinecke. 2010. Avatar creation and video game enjoyment. *Journal of Media Psychology* (2010).

176. Torrey Trust, Robert W. Maloy, and Sharon Edwards. 2017. Learning through Making: Emerging and Expanding Designs for College Classes. *Tech Trends* (2017). DOI: http://dx.doi.org/10.1007/s11528-017-0214-0

177. Selen Turkay. 2014. The Effects of Avatar-based Customization on Player Identification in Extended MMO Gameplay. *Games, Learning and Society Conference* 6, March (2014), 227–234. DOI: http://dx.doi.org/10.4018/ijgcms.2014010101

178. Selen Turkay and Charles K. Kinzer. *The Relationship between Avatar-Based Customization, Player Identification, and Motivation*. 48–79 pages. DOI: http://dx.doi.org/10.4018/978-1-5225-1817-4.ch003

179. Jan Van Looy, Cédric Courtois, Melanie De Vocht, and Lieven De Marez. 2012. Player Identification in Online Games: Validation of a Scale for Measuring Identification in MMOGs. *Media Psychology* 15, 2 (2012), 197–221. DOI: http://dx.doi.org/10.1080/15213269.2012.674917

180. Eva A Van Reijmersdal, Jeroen Jansz, Oscar Peters, and Guda Van Noort. 2013. Why girls go pink: Game character identification and game-players' motivations. *Computers in Human Behavior* 29, 6 (2013), 2640–2649.

181. Asimina Vasalou, Adam N Joinson, and Jeremy Pitt. 2007. Constructing my online self: avatars that increase self-focused attention. *Proceedings of ACM CHI 2007 Conference on Human Factors in Computing Systems* 1 (2007), 445–448. DOI: http://dx.doi.org/10.1145/1240624.1240696

182. T Franklin Waddell, S Shyam Sundar, and Joshua Auriemma. 2015. Can customizing an avatar motivate exercise intentions and health behaviors among those with low health ideals? *Cyberpsychology, Behavior, and Social Networking* 18, 11 (2015), 687–690.

183. Hao Wang and Chuen-Tsai Sun. 2011. Game reward systems: Gaming experiences and social meanings.. In *DiGRA Conference*.

184. Melissa Watts. 2016. *Avatar Self-Identification, Self-Esteem, and Perceived Social Capital in the Real World: A Study of World of Warcraft Players and their Avatars*. University of South Florida.

185. Kevin Werbach and Dan Hunter. 2012. *For the win: How game thinking can revolutionize your business*. Wharton Digital Press.

186. Leo Sang-min Whang and Geunyoung Chang. 2004. Lifestyles of virtual world residents: living in the on-line game "lineage". *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society* 7, 5 (2004), 592–600. DOI: http://dx.doi.org/10.1109/CYBER.2003.1253430

187. Robert W White. 1959. Motivation reconsidered: The concept of competence. *Psychological review* 66, 5 (1959), 297.

188. Uri Wilensky and I Evanston. 1999. NetLogo: Center for connected learning and computer-based modeling. *Northwestern University, Evanston, IL* (1999), 49–52.

189. Keith M. Williamson, Lee Land, Beverly Butler, and Hassan B. Ndahi. 2004. A structured framework for using games to teach mathematics and science in K-12 classrooms. *The Technology Teacher* (2004).

190. Jessica Wolfendale. 2007. My avatar, my self: Virtual harm and attachment. *Ethics and Information Technology* 9, 2 (2007), 111–119. DOI: http://dx.doi.org/10.1007/s10676-006-9125-z

191. Nick Yee and Jeremy Bailenson. 2007. The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Comm. Research* (jul 2007). DOI: http://dx.doi.org/10.1111/j.1468-2958.2007.00299.x

192. Nick Yee, Nicolas Ducheneaut, Mike Yao, and Les Nelson. 2011. Do Men Heal More When in Drag? *CHI 2011* (2011), 1–4.

193. Zhi-Jin Zhong and Mike Zhengyu Yao. 2013. Gaming motivations, avatar-self identification and symptoms of online game addiction. *Asian Journal of Communication* 23, 5 (2013), 555–573.

194. Gabriel Zicherman and Christopher Cunningham. 2011. Gamification by design. *Sebastobol, CA: O'Reilly Media* (2011).