MS_ID: engage_26

engage_26_Martey_five-approaches_to_measuring

Word Count: 6768

Title: Five Approaches to Measuring Engagement: Comparisons by Video Game Characteristics

Contact Details:

Rosa Mikeal Martey	rosa.martey@colostate.edu; (970) 491-6469, Colorado State University, CD 1785, Fort Collins, CO 80523-1785
Kate Kenski	kkenski@email.arizona.edu, (520) 626-3066, Department of Communication 1103 E University Blvd, P.O.Box 210025, Tucson, AZ 85721
James Folkestad	James.Folkestad@colostate.edu, (970) 491-7823, 209 Education Building, Colorado State University Fort Collins, CO 80523-1588
Laurie Feldman	laurie.b.feldman@gmail.com, (518) 442-4842, Department of Psychology, Social Sciences 399, 1400 Washington Avenue, Albany, NY 12222
Elana Gordis	egordis@albany.edu, (518) 442-4820, Department of Psychology, Social Sciences 399, 1400 Washington Avenue, Albany, NY 12222
Adrienne Shaw	adrienneshaw13@gmail.com, 215-204-8422 , School of Media and Communication, Annenberg Hall, 2020 North 13th Street, Philadelphia, PA 19122
Jennifer Stromer-Galley	jstromer@syr.edu, (315) 443-1823 , School of Information Studies, Hinds 220, Syracuse University, Syracuse, NY 13244
Ben Clegg	Benjamin.Clegg@colostate.edu; (970) 491-2211; 1876 Campus Delivery, Colorado State University, Fort Collins, CO 80523-1876
Hui Zhang	huizhang@rams.colostate.edu, Colorado State University, CD 1785, Fort Collins, CO 80523-1785
Nissim Kaufman	nissimkaufmann@yahoo.com, University at Albany, ILS Room 262B, Social Science Building, 1400 Washington Avenue, Albany, NY 12222
Ari N. Rabkin	arabkin@albany.edu, Department of Psychology, Social Sciences 399, 1400 Washington Avenue, Albany, NY 12222
Samira Shaikh	262B, Social Science Building, 1400 Washington Avenue, Albany, NY 12222 @gmail.com
Tomek Strzalkowski	tomek@albany.edu, (518) 442-2608, University at Albany, ILS Room 262B, Social Science Building, 1400 Washington Avenue, Albany, NY 12222

++

Five Approaches to Measuring Engagement: Comparisons by Video Game Characteristics

Rosa Mikeal Martey[1]
Kate Kenski[3]
James Folkestad[1]
Laurie Feldman[2]
Elana Gordis[2]
Adrienne Shaw[4]
Jennifer Stromer-Galley[5]
Ben Clegg[1]
Hui Zhang[1]
Nissim Kaufman[2]
Ari N. Rabkin[2]
Samira Shaikh[2]
Tomek Strzalkowski[2]

1. Colorado State University, CO, USA
2. University at Albany, SUNY, NY, USA
3. University of Arizona, AZ, USA
4. Temple University, PA, USA
5. Syracuse University, NY USA

Abstract:
Engagement has been identified as a crucial component of learning in games research. However, the conceptualization and operationalization of engagement varies widely in the literature. Many valuable approaches illuminate ways in which presence, flow, arousal, participation and other

concepts constitute or contribute to engagement. However, few studies examine multiple conceptualizations of engagement in the same project. This paper discusses the results of two experiments that measure engagement in five different ways: survey self-report, content analyses of player videos, electro-dermal activity, mouse movements, and game click logs. We examine the relationships among these measures and assess how they are affected by the technical characteristics of a 30 minute custom-built educational game: use of a customized character, level of narrative complexity, and level of art complexity. We found that the five measures of engagement correlated in limited ways, and that they revealed substantially different relationships with game characteristics. We conclude that engagement as a construct is more complex than is captured in any of these measures individually and that using multiple methods to assess engagement can illuminate aspects of engagement not detectable by a single method of measurement.

Key words: art; attention; character customization; EDA; educational games; engagement; interactivity; measurement; mouse movements; narrative; presence; transportation;

Research on engagement has taken many forms and been assessed many ways; however, few studies examine multiple measures of engagement simultaneously. There remain questions in the literature about whether different measures capture distinct aspects of engagement. To understand how different measures relate to each other and to determine whether and which elements of the stimuli affect engagement, we conducted two experiments that measure engagement when playing a computer-based video game in five different ways: survey self-report, attention to the stimulus, electro-dermal activity (EDA, previously known as Galvanic Skin Response), mouse movements, and game click logs. We first examine the relationships among these measures, then we analyze which characteristics of an educational game best promote engagement: use of a customized character, level of narrative complexity, and level of art complexity. We found the game manipulations had little to no reliable effect on our various engagement measures. We found evidence, however, to suggest that different measures to assess engagement reveal important player differences that are inaccessible with a single measure approach.

Conceptualizations and Measures of Engagement

Engagement with media and games has been defined many ways, including as involvement (Zaichkowsky, 1985), immersion (Biocca & Levy, 1995), arousal (Ravaja, Saari, Salminen, Laarni & Kallinen, 2006), attention (O'Brien & Toms, 2008), interest (Dow et al., 2007), identification (Brusselle and Bilandizc, 2009), enjoyment (Mayes & Cotton, 2001), involvement (Klimmt & Vorderer, 2003), effort (Dow et al., 2007), and flow (Csikszentmihalyi, 1997). A popular conceptualization of engagement is the notion of presence or a sense of "being there" in a media experience. Although, there are different definitions of social presence, Lombard and Ditton (1997) offer one of the more common approaches, especially in media effects scholarship. They identify six dimensions of presence that are often used separately or together as measures of engagement in video games (see McMahan, 2003; Nowak, Kremar & Farrar, 2008). These dimensions, although linked, identify distinct aspects of presence that have different relationships with media experiences (See, Brusselle & Bilandzic, 2009; Wise, Chang, Duffe, & del Valle, 2004). Most measures that approach engagement as immersion, presence, flow, effort, and enjoyment use self-reports on questionnaires.

In the context of video games research, physiological measures of engagement conceptualize it as increases in arousal and attention. O'Brien and Toms (2008) use physical as well as subjective measures, including attention to the screen and clicking on objects, to capture cycles of engagement and disengagement. They explain that emotional responses, attention, and interactivity increase and decrease throughout engaged play. Ravaja et al. (2006) measured player EDA and facial electromyography (EMG) and found reliable relationships between game events and emotional valence. Lim and Reeves (2010) found increased EDA during game play when players interacted with human-controlled avatars than computer-controlled ones, suggesting EDA is sensitive to measuring arousing events during game play.

Few studies have measured engagement through activity or behavior in video games. One potential way of measuring it is through eye tracking and attention to the screen (see, for example, Ismail, Sidek, Khan, Jalaludin, 2012; Renshaw, Stevens, & Denton, 2009; Teixeira, Wedel, & Pieters, 2012). Another way is to observe behavior of people while they are exposed to a stimulus (Annetta et al., 2009), by examining facial expressions (Karimi & Lim, 2010) or choices people make to engage or not with particular characters (Van Vugt et al., 2006). O'Brien and Toms (2008) advocate that interactivity should be a measure of engagement when the stimulus allows for interaction, such as clicking in a video game. Indeed, mouse movements have been shown to be a valid measure of decision-making (Dale, Kehoe, & Spivey, 2007), and given O'Brien and Toms directive, could be considered a measure of interactivity in games that indicates engagement.

For the purposes of our research, we define engagement as the degree of activity or attention someone gives to a person or object over some period of time. We adopt O'Brien and Toms (2008) operationalization that engagement is a multi-dimensional phenomenon with the following attributes: "challenge, pleasure, endurability, attention, aesthetic and sensory appeal, feedback, variety/novelty, interactivity, and perceived user control" (p. 941). As a multi-dimensional or meta-construct (Appleton, 2008; Fredricks, Blumenfeld, and Paris, 2004), we include self-reports, physiological, and behavioral measures of engagement, and as described next, manipulate game characteristics to determine how measures interrelate.

Engagement and Game Characteristics

Research suggests that technological characteristics of the medium can contribute to – or hinder – engagement (see, for example, O'Brien and Toms 2008). This study examines three game characteristics that may influence player engagement in a video game: player avatar, visual realism, and narrative.

Levels of engagement in games have been linked to identification between player and avatar (Gee, 2003). Character customization may enhance a sense of presence in the game space (Bailey, Wise & Bolls, 2009), enjoyment (Trepte & Reinecke, 2011), motivation to play (Turkay & Adinolf, 2010), and engagement (Shaw, 2010; Trepte & Reinecke, 2011). This may be because avatars are an extension of agency necessary to navigate a game space (Klimmt, Dorothée & Vorderer, 2009). Other research suggests that strong feelings of presence in games can also emerge when characters are a fixed element of the game setting, rather than customized by the player (Shaw, 2010; Gee, 2003; Juul, 2005). Avatar attachment is frequently, although not always, associated with customizing or creating characters as a form of self-representation or to fulfill specific goals (Merola & Peña, 2010).

Some researchers argue that visual realism enhances feelings of co-presence (Bailenson et al., 2005) and avatar credibility (Nowak et al., 2008). Others suggest that less realism leads players to be more engaged in games because they are more open to interpretation, allowing greater identification with the setting (Wolf, 2003) or characters (Shaw, 2010). Research from Jones (2008) demonstrated that the higher complexity and realism of a film actually produced lower levels of presence than a comic book telling the same story. This may be because perceptions of presence are caused by heuristics evoked by specific cues in the digital environment. Sundar, Oeldorf-Hirsch, and Garga (2008) propose that presence in communication technologies is mediated by user interpretations of particular cues, including user agency and interactivity, and navigability. In particular, as a consequence of presence is the sensation that the medium or technology has become 'invisible' to players (Lombard & Ditton, 1997), it is possible that greater visual detail can lead to a decreased awareness of the medium and a consequently increased sense of presence. Yet, not all research supports this. Whitbred and colleagues (2010) found that use of paper versus video communication of a university's mission statement had no influence on involvement with the statement, and that the richer channel resulted in lower feelings of presence. They suggest that increased cues can become a distraction or increase expectations of the audience, leading to a decreased sense of involvement. Other research suggests that in some cases level of visual abstraction has no effect on a sense of presence (Dinh et al., 1999; Zimmons & Panter, 2003). Overall, although most scholars argue visual richness and complexity increase engagement, questions remain about whether this is always the case.

Defined as a story expressed through representations of a series of events (Abbott, 2002), narrative has been identified as when one is absorbed by media content (Green & Brock, 2002), though not all researchers agrees on this. Use of a story can increase learning and engagement, and researchers suggest that rich narratives invite exploratory learning (Mott, Callaway, Zettlemoyer, Lee & Lester, 1999), motivate players (McQuiggan, Rowe, Lee & Lester, 2008), lead to transportation (the sense one has been teleported into the story) and engagement (Gerrig, 1993), and increase arousal and presence (Schneider et al 2004). Narratives in games, however, are not universally viewed as important to engagement (Juul, 2005). This may be because, as Green (2004) argues, transportation into a narrative world may be mediated by perceived realism of the story. Importantly, Green and Brock (2002) argue that transportation emerges most strongly when audiences are fully focused on the events of the narrative. In educational games however, narrative often competes with learning content, perhaps negating effects of transportation as well as distracting players from critically processing game content (Green, Brock & Kaufman, 2004).

The Present Study

The present study examines five different measures of engagement: self-reported presence, the physiological measure of Electro-Dermal Activity (EDA), and behavioral measures of attention to the stimulus, mouse clicks, and mouse rests. These measures capture multiple levels of engagement with varied types of data. We compare these measures and examine how three game characteristics theorized to generate different levels of engagement correspond with these engagement measures. This article discusses only the game engagement portion of that study. We hypothesize the following based on the literature reviewed above:

- H1: Customizing a character in a game increases player engagement more than assigning a character.
- H2: Richer, more complex narrative increases player engagement more than lighter narrative.
- H3: More detailed art style increases player engagement more than simpler art style.

We also aim to answer the following:

- RQ1: What is the relationship between these different measures of engagement?

Methods

Conditions and Stimuli Material

Study 1 used a one factor (2 level) plus control design in which participants could customize or were assigned an avatar. Study 2 used a 2 x 2 plus control full factorial between subjects design with conditions rich versus light narrative and detailed versus simple art.

Both studies used the CYCLES game, a 2D puzzle-based point-and-click game built by a professional game company in Flash and played on a desktop computer, with text, audio, and voice narration. Throughout the game, a trainer speaks to the player to provide instructions and learning content displayed as subtitles with voiceover (VO) audio. In the game, players move from room to room solving puzzles by identifying and mitigating specific cognitive biases, such as confirmation bias. For example, players must develop a hypothesis and then test which robots are fire proof. They achieve success when they test robots that disconfirm their hypothesis. Between puzzle rooms, transition areas present the player with core learning content, reviews, and quizzes.

The Study 1 game varied whether or not participants could customize their avatar by selecting a helmet shape (e.g., round, square,), suit pattern (e.g., striped, plain), gender (male or female), and a name. After the customization screens, game play was identical across the two conditions.

The Study 2 game used an assigned character, where players selected only avatar sex and a name. The game conditions varied by art (detailed vs. simple) and narrative complexity (rich vs. light). Perspective and functionality of the game were the same, but the detailed art condition was full-color, with rich textures and shading. The simple art condition was largely monochrome with minimal shading and almost no textures (see Figure 1). Game play, objects, and text were the same across the two art conditions.

For each art condition, a light and a rich narrative version of the game was built. The rich narrative version positioned the player as the child of a bias reduction expert who once worked with the CYCLES Training Center owner, Dr. Ohm. The Center's top trainer helps the player steal Ohm's techniques to share them with the world. In the process, the player discovers that Dr. Ohm is using human brains to create robots, and the mission is to shut down the evil operation. Most of the story is told in an introduction, in text added to transition rooms, and in the game ending. For example, in the light narrative condition, the player is welcomed with the text, "Welcome to the CYCLES Training and Mitigation Center." In the rich narrative condition, the player first hears, "Psst! Can you hear me? Good. I know you want to get your mother's bias training back from Ohm, but be careful!"

FIGURE 1 ABOUT HERE

Figure 1: Simple and Detailed Art in a Puzzle Room

Measures

Survey self-reported engagement. Four of the scales from Lombard and Ditton's (1997) *Temple Presence Inventory* (TPI) were used in a survey administered immediately after playing the game: *Social Richness*, which identifies how responsive and sociable the experience felt (reliability $\alpha = .88$) ; *Social Realism*, which assess how similar game experiences are to "the real world" (reliability $\alpha = .90$); *Active Interpersonal Presence*, which assesses how often players smiled, spoke, or laughed in response to the game (reliability $\alpha = .73$); and *Mental Immersion*, which identifies how involving and exciting the experience felt (reliability $\alpha = .89$). Except for *Social Richness*, scale answers were on a 7-point Likert scales from *strongly disagree* to *strongly agree*, and some items were reversed. *Social Richness* used semantic differentials on a 7-point scale. We also used the *Narrative Transportation Scale* (Green & Brock, 2000), which assesses the extent to which individuals feel transported “inside” a media experience with questions such as, “I could picture myself in the game” (reliability $\alpha = .81$) Answers were on a 5-point Likert scale from *strongly disagree* to *strongly agree*.

Attention-Based Engagement (ABE). We developed a measure based on engagement measures developed by Read, MacFarlane, and Casey (2002). We video recorded the faces of a 30% sample of participants and content analyzed them to identify time spent looking at the screen. Research assistants were trained for approximately 30 hours using a 10% sample of identical session videos to reach an intercoder reliability of .89 using Krippendorff’s (2004) alpha. Seconds looking away from the screen were subtracted from total game play time and used to calculate the percent of total time participants were looking at the screen, this time on task was ABE. It was positively correlated with how “involving” participants rated the game ($r = .401$, $p < .05$) and negatively correlated with finding the game frustrating ($r = -.376$; $p < .05$). These relationships suggest ABE is capturing valid aspects of engagement, but as a new measure, results are interpreted with caution.

Electro-Dermal activity (EDA). Prior research suggests that EDA, which varies with sympathetic nervous system activation, can serve as a measure of cognitive engagement to external stimuli such as games (Mandryk & Inkpen, 2004). To measure EDA, we connected participants to a Q pod (Affectiva) EDA hand unit with Ag/AgCl dry electrodes. We measured EDA in microsiemens, sampled at a standard frequency of 16 Hz. We use the latter 5 minutes of a 10-minute inactive period before the first survey to establish a baseline EDA level. For our analyses, we log transformed baseline EDA and EDA during game play to attenuate skew. We calculated *EDA Change* as log (EDA during game play) minus log (baseline EDA).

Mouse clicks. The game was developed to include automatic logging capabilities that track player clicks on interactive objects. The measure *Active Mouse Clicks* is the sum of clicks on interactive puzzle objects such as robots and on puzzle room floors that move the avatar. These were selected because they reflect differences in puzzle and movement styles across participants. Clicks on the screen that have no effect in the game (e.g., on non-interactive text or images) were excluded.

Mouse movement. Patterns of mouse movements have been documented to provide measures of cognitive function, including engagement and decision making (Dale, Kehoe, & Spivey, 2007). We recorded mouse location in x,y at 20 millisecond (ms) intervals. The variable in the present analysis is *Mouse Rest Percent*, calculated as the percent of total game play time in which the mouse is not moving. To focus on the parts of the game with the most interaction, we included only movements in puzzle rooms, excluding transition room data. Study 1 included only mouse movements in three rooms, and Study 2 included mouse movement in all puzzle rooms. Measures were normalized for differences between players and differences between rooms and adjusted to normalize the distribution.

Procedures

For both studies, procedures were the same. First, we recruited college students from three universities using psychology research pools and communication classrooms. Participants came into a university computer lab where they used a computer to take a 35-minute pre-session survey that measured game and computer experience, psychological characteristics, cognitive biases, and demographics. They were randomly assigned to a condition. Immediately after completing the game, they answered a 25-minute post-session questionnaire measuring cognitive biases and engagement.

Participants at one university were measured for EDA (roughly 30% of the total N of the sample). This changed their procedures slightly. Before they took the first questionnaire, they were fitted with a Q pod (Affectiva) EDA hand unit with Ag/AgCl dry electrodes, and then sat quietly for 10 minutes to establish a baseline EDA level. After taking the pre-session questionnaire, they sat again for 5 minutes as a rest time, and then played the game and took the post-session survey. They were also video-taped using cameras on the computers. In addition, 35% of a second university's participants were video-taped for analysis.

Debriefing

Because we conducted controlled experiments focused on teaching about cognitive biases, we provided a debriefing, as best practices of teaching in simulations and gaming scholars suggest (Lederman, 1992). Ours consisted of a description of the objectives of the game and provided recommended readings about cognitive biases. The objective was to help participants understand what they had experienced and why.

Participants

Study 1 participants were 280 undergraduate and graduate students from three universities recruited from summer classes. Participants were 41% male and 59% female, and 64% were white. Participants had an average age of 21, and 81% were in their third or last years of school. Nearly all (96%) participants considered themselves to have intermediate or better expertise with computers, 58% agreed or strongly agreed they "like playing video games," 22% considered themselves "gamers," and 60% said they play games 1 day or more per week.

The sample in Study 2 consisted of 480 college students from three universities. They were 35% male, 65% female, and 80% were white. Their average age was 19.5 years, and 67% were in their first or second year of school. Their computer knowledge was high, with 87% reporting intermediate or higher expertise. A majority (59%) agreed or strongly agreed they "like playing video games," 16% considered themselves "gamers," and 61% said they play games 1 day or more per week.

Results

To examine the three hypotheses, we first examined descriptive statistics on each of the measures (see Tables 1 and 2). We then conducted Pearson correlations on the standardized (Z-score) values of the nine engagement measures used here (see Tables 3 and 4).

TABLES 1 AND 2 ABOUT HERE

TABLES 3 AND 4 ABOUT HERE

Analyses of Variance for Study 1

Self-Reported Measures of Engagement. A MANOVA employing the five scale measures as dependent variables, and character customization (customized, assigned) as the independent variable found no effect of customization (*Wilks' Λ*=.03, $F(5,156)=<1$, $p<.95$). Examining the effects of character customization on each measure alone showed no significance for main effects (all p's > .3). H1 was not supported using self-reported measures.

Other Measures of Engagement. A MANOVA employing the four objective measures of engagement as dependent variables, and character customization as the independent variable found no main effect of customization (*Wilks' Λ*=.59, $F(4,7)<1.24$), indicating H1 was not supported using external measures.

Analyses of Variance for Study 2

Self-Reported Measures of Engagement. A MANOVA employing the 5 scale measures as dependent variables, and art (detailed, simple) by narrative (rich, light) as independent variables found a significant main effect of narrative (*Wilks' Λ*=.88, $F(5,375)=10.74$, $p<.001$), but no main effect of art (*Wilks' Λ*=.99, $F(5,375)<1$), and no interaction (*Wilks' Λ*=.99, $F(5,375)<1$).

The main effects on narrative in the multivariate analysis were mirrored on the Social Realism scale alone ($F(1,379)=47.48$, $p<.001$; light narrative mean=4.4; rich narrative mean=3.3), and marginally non-significant on the Transportation scale ($F(1,379)=3.55$, $p=.06$; light narrative mean=2.9; rich narrative mean=2.7). These results indicate that light narrative demonstrated greater engagement than the rich narrative for both measures, which is the opposite direction than predicted. None of the other univariate analyses approached significance for either main effect or for interactions (all p's>.15). H2 and H3 were not supported using self-reported measures.

Other Measures of Engagement. Separate analyses were conducted for each of the three different types of objective measures (EDA, mouse clicks and rests, and ABE).

EDA. Change in EDA relative to the player's own baseline showed a marginally significant interaction of art by time (*Wilks' Λ*=.95, $F(1,79)=3.94$, $p=.51$), a significant effect of art by narrative by time (*Wilks' Λ*=.89, $F(3,79)=3.81$, $p<.05$), and a main effect of time (*Wilks' Λ*=.88, $F(1,79)=10.48$, $p<.01$), that were evidence of interactions, with no main effects or other interactions approaching significance (all p's>.3). Paired sample t-tests showed significant increases from baseline in EDA in the rich narrative, detailed art condition ($t(20)=3.85$,$p<.05$; baseline mean =1.2, game mean=1.6); and in the simple art, light narrative condition ($t(20)=2.82$,$p<.05$; baseline mean =1.3, game mean=1.5), but no significant changes in the detailed art, light narrative condition ($t(20)=1.16$,$p>.05$; baseline mean =1.1, game mean=1.2), or in the simple art, rich narrative condition ($t(20)=-.84$,$p>.05$; baseline mean =1.6, game mean=1.5). Neither H2 nor H3 were supported.

ABE. An ANOVA looking at the amount of time spent attending to the screen found no main effects for art ($F(1,58)<1$) or narrative ($F(1,58)<1$), nor a significant interaction $F(1,58)=1.76$, $p>.1$).

Mouse use. A MANOVA examined mouse performance metrics for mouse rest time and clicks, but showed no main effects for art (*Wilks' Λ*=.99, $F(2,329)=2.11$, $p>.1$) or narrative (*Wilks' Λ*=1.0 $F(2,329)<1$), nor a significant interaction (*Wilks' Λ*=1.0, $F(2,329)<1$).

Engagement Measure Interrelationships

Although the game conditions differed from Study 1 to Study 2, correlations among the self-reported engagement measures were somewhat consistent across the studies. All self-reported measures were positively associated with one another as expected, but at modest levels of correlation, corresponding with research that shows these measures tap into aspects of engagement that are related but not synonymous (Lombard & Ditton, 1997). In Study 1, there was very little evidence that subjective, behavioral, and physiological measures of engagement are related, with no correlations among the measures at $p < .05$. In Study 2, time looking at the screen and mouse rests were related to some subjective measures, and EDA pre-post change was related to Active Interpersonal ($r = -.102$) with marginal significance.

Discussion

Overall, these patterns suggest several things. First, as Study 2 indicates, looking more at the screen (ABE) is related to increases in some types of presence (Transportation, $r = .397$; Social Richness, $r = .344$; and Mental Immersion, $r = .545$), but not all. This reinforces the distinctions among the self-reported measures (Lombard & Ditton, 1997) and suggests that ABE captures a distinct dimension of engagement that may be related to some aspects of presence. Interestingly, ABE was unrelated to all other behavioral and physiological measures (EDA, mouse rests, mouse clicks), suggesting that it is also distinct from actions taken in the game and from arousal. It is possible that looking away from the game is driven by external factors such as distractions in the room rather than game activities. If as suggested by Csikszentmihalyi (1997) participants in a state of flow are less likely to look away at such distractions, these findings fail to document a relation between flow and either actions taken in the game or to arousal.

Second, the lack of relationship between the presence measures and mouse clicks suggest that presence cannot be associated with actions taken during the game. In addition, the fact that EDA is unrelated to all other measures in both studies inidcate that arousal is a substantially different dimension of engagement than attention, presence, and mouse use. We draw this conclusion with caution, however, as the high variability typical of EDA measures (in both studies the SD was higher than the mean) may make such relationships more difficult to detect.

Finally, the negative relationships among mouse rests and presence measures suggest that higher levels of presence correspond with more active mouse use. Players with a greater sense of presence may tend to cover more distance when they move their mouse around the game space. Mouse movements may be indicative of using the mouse cursor as a tool for visual exploration. It may also indicate greater presence and thus greater motivation to look around. Future research should pair mouse movements with eye-tracking data to further explore this pattern.

Engagement and Game Characteristics

Analyses of Study 1 revealed that character customization did not increase engagement on any of our nine engagement measures. H1 was therefore not supported. Given that our nine measures tapped into different dimensions of engagement, we are confident that customization was not a consequential game feature for engagement. This may be because of the short duration of our game (approximately 30 minutes), and it may be the case that avatar customization affects engagement but only for players who have spent substantial time with a game. Alternatively, it may be that customization is less consequential for educational games than for other types.

Study 2 suggested that narrative and art may affect some aspects of engagement. First, two of the presence measures were related to narrative, but in the opposite direction than predicted: richer, more complex narrative corresponded with *lower* levels of Social Realism and (at marginal significance) Transportation. Although narrative has been a well-established technique for generating engagement (Green & Brock, 2002; Green, 2004), our study imposes qualifications on that general finding. In this educational game, the narrative was necessarily combined with learning content, which may have made players feel the narrative was distracting, confusing, or conflicting in relation to the game's goals. It is also possible that the game's puzzle genre was relevant. Research showing the importance of game narratives to engagement generally examine narrative-oriented genres, such as adventure games. The puzzles and learning content may have kept players from fully focusing on the events of the narrative, reducing engagement (Green & Brock, 2002). A sense of realism may have only emerged in the light narrative condition, where the game's real-world examples and emphasis on decision-making were not overshadowed by the story in the rich narrative condition.

Self-reported measures showed no difference by level of art complexity, but EDA measures provide some evidence that art matters to engagement. The significant effects of time on EDA found in only the most complex condition (rich narrative, detailed art) and least complex condition (light narrative, simple art) and the significant interaction among art, narrative and time suggest that combinations of art and narrative influence participants in different ways. Considering also the significant impact of art alone, this suggests that complex art increases engagement, but only when it is combined with similarly complex narrative. This may be because participants expect greater story richness when detailed art is used. A complex narrative with simple art and vice versa might generate conflicting expectations, causing players to disengage, lowering arousal.

Differences in the subjective, behavioral, and physiological measures of engagement documented in this study may be in part due to how participants remember their experience. Subjective measures were assessed by survey after game play, whereas the behavioral and physiological measures were assessed continuously during game play. It is possible that the frequent changes to complex images in the detailed art condition produce arousal, but do not stand out in participants' memories during the post-test. Narrative, in contrast, may be more memorable (Green, Brock and Kaufman, 2004). This corresponds with research from Dinh et al. (1999) that found that level of visual detail had no effect on recall of elements of a virtual environment. To the extent that greater visual and narrative detail led to an increased awareness of the game as a game, these factors may inhibit feelings of presence even if they increase arousal (Sundar et al., 2008; Whitbread, 2010).

Conclusions

We tested a particular genre of game focused on specific learning objectives. We found that manipulated game characteristics of art, narrative, and character customization had little to no effect on engagement. Different measures seemed to tap distinct dimensions of engagement. Some of the dimensions of self-reported presence correlated with ABE, but had no relationship with EDA. Mouse action and presence measures were positively related. Our results support the idea that engagement is a multi-dimensional construct, and suggests the need for more research to systematically examine the relationships between measures of engagement.

Our research has limitations. First, the small sample size of some of our measures and the college student population used in this study necessarily require caution in generalizing our results. Second, ABE is an effort to measure attention without the expense and complexity of eye-tracking devices. Because this is a new measure, results should be understood as preliminary. One current limitation of ABE is that it has a small variance, which we aim to correct as we refine the measure. Third, our experiments were on an educational game that we developed, and it might be an outlier in terms of engagement. In the future, we aim to conduct an experiment using our battery of engagement measures to compare a popular commercial game, a commercial educational game, and our game. Fourth, the lack of differences in our experimental conditions on engagement may be due to not enough distinction between the conditions. Experiments that create conditions that are even further at the extremes of game characteristics may be needed to see meaningful differences in engagement.

Overall, the complex relationships found among game factors like presence and factors widely identified as enhancing engagement (art and narrative) suggest that it is not uniform in all games. Our research suggests that using different methods to assess engagement can illuminate important patterns in player reactions that are inaccessible through a single approach alone.

Acknowledgements

We would like to thank the three anonymous reviewers for their care and attention in their review of the manuscript. We also thank the many lab assistants who helped us to collect the data in this study including Summer Alameel, Daniela Castillo, Bethany Conway, Jennifer Crowley, Christine Filer, Michelle Hook, Scott Hawks Lee, John O'Connell, Xiaoai Ren, Stephanie Train, and Di Wang.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

This research was supported by the Intelligence Advanced Research Projects Activity (IARPA) via the Air Force Research Laboratory. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, AFRL, or the U.S. Government.

References

Abbott, H. P. (2002). *The Cambridge introduction to narrative.* Cambridge University Press.

Annetta, L., Mangrum, J., Holmes, S., Collazo, K., & Cheng, M. (2009). Bridging reality to virtual reality: Investigating gender effect and student engagement on learning through video game play in an elementary school classroom. *International Journal of Science Education, 31,* 1091-1113.

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45*, 369-386. doi: 10.1002/pits

Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 14, 379-393.

Bailey, R., Wise, K., Bolls, P. (2009). How avatar customizability affects children's arousal and subjective presence during junk food-sponsored online video games. *Cyberpsychology, Behavior, and Social Networking* 12, 277-283.

Biocca, F. & Levy, M. R. (1995). *Communication in the age of virtual reality.* Hillsdale, N. J.: Lawrence Erlbaum Associates.

Brusselle, R. & Bilandzic, H. (2009). Measuring narrative engagement, *Media Psychology*, 12, 321-347.

Csikszentmihalyi, M. (1997) *Finding flow: The psychology of engagement with everyday life*. Basic Books, New York.

Dale, R., Kehoe, C., & Spivey, M. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory and Cognition*, 35, 15-28.

Dinh, H. Q., Walker, N., Song, C. Kobayashi, A. & Hodges, L. F. (1999). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. Proceedings of the IEEE Virtual Reality (VR '99). Washington, DC: IEEE Computer Society.

Dow, S., Mehta, M., Harmon, E., Macintyre, B., & Mateas, M. (2007). Presence and engagement in an interactive drama. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59-109.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading.* New Haven, CT: Yale University Press.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*, 701-721.

Green, M. C., & Brock, T. C. (2002). In the mind's eye: Imagery and transportation into narrative worlds. In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 315-341). Mahwah, NJ: Lawrence Erlbaum Associates

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, 14, 311-327.

Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes*, 38, 247-266.

Ismail, E., Sidek, S. N., Khan, M. R., Jalaludin, N. A. (July 2012). Analysis of Engagement Factor in Trajectory Tracking-Based Experiment. 2012 International Conference on Computer and Communication Engineering (ICCCE), Kuala Lumpur, Malaysia.

Jones, M. T. (2008). *Found in translation: Structural and cognitive aspects of the adaptation of comic art to film.* Philadelphia: Temple University.

Juul, J. (2005). *Half-real: Video games between real rules and fictional worlds*. Cambridge, MA: The MIT Press.

Karimi, A., & Lim, Y. P. (December, 2010). Children, engagement, and enjoyment in digital narrative. Proceedings of ASCILITE 2010, Sydney, Australia. Retrieved June 28, 2013 from http://www.ascilite.org.au/conferences/sydney10/procs/Karimi-full.pdf.

Klimmt, C. & Vorderer, P. (2003). Media Psychology ‘is not yet there’: Introducing theories on media entertainment to the Presence debate. *Presence: Teleoperators and Virtual Environments*, 12, 346-359.

Klimmt, C., Dorothée, H., & Vorderer, P. (2009). The video game experience as ‘true’ identification: A theory of enjoyable alterations of players' self-perception. *Communication Theory*, *19*, 351-373.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.,

Lederman, L. C. (1992). Debriefing: Twoard a systematic assessment of theory and practice. *Simulation & Gaming, 23*, 145-160. doi: 10.1177/1046878192232003.

Lim., S., & Reeves, B. (2010). Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player. *International Journal of Human-Computer Studies, 68,* 57-68.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication,* 3(2). Retrieved from http://dx. doi. org/10.1111/j. 1083-6101.1997.

McMahan, A. (2003). Immersion, engagement, and presence: A method for analyzing 3-D Video games. In Wolf, M. J. P., and Perron, B. (Eds.) *The video game theory reader* (pp. 67–86). New York: Routledge.

Mandryk, R. L., & Inkpen, K. M. (2004). Physiological indicators for the evaluation of co-located collaborative play. CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work.

Mayes, D. K., & Cotton, J. E. (2001). Measuring engagement in video games: A questionnaire. Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting.

Merola, N. A., & Peña, J. (2010). The effects of avatar appearance in virtual worlds. *Journal of Virtual Worlds Research,* 2(5). Retrieved from https://journals.tdl.org/jvwr/article/view/843

Mott, B., Callaway, C., Zettlemoyer, L., Lee, S. & Lester, J. (1999). Towards narrative-centered learning environments. Proceedings of the AAAI Fall Symposium on Narrative Intelligence.

Nowak, K. L., Kremar, M., & Farrar, K. M. (2008). The causes and consequences of presence: Considering the influence of violent video games on presence and aggression. *Presence: Teleoperators and Virtual Environments*, 17, 256-268.

O'Brien, H. L. & Toms, E. G. (2008), What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science*, 59, 938–955.

Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic emotional reactions to video game events: A psychophysiological investigation. *Media Psychology*, 8, 343-367.

Read, J., MacFarlane, S. & Casey, C. (2002). Endurability, engagement and expectations: Measuring children's fun. *Proceedings of Conference on Interaction Design and Children* (pp. 189–198). Eindhoven: Shaker Publishing.

McQuiggan, S., Rowe, J., Lee, S., & Lester, J. (2008). Story-based learning: The impact of narrative on learning experiences and outcomes. In B. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.) *Intelligent tutoring systems, vol. 5091 of Lecture notes in computer science* (pp. 530-539). Berlin, Heidelberg: Springer.

Renshaw, T., Stevens, R., Denton, P. D. (2009). Towards understanding engagement in games: An eye-tracking study. *On the Horizon, 17*, 408-420.

Schneider, E. F., Lang, A., Shin, M., & Bradley, S. D. (2004). Death with a story: How story impacts emotional, motivational, and physiological responses to first-person shooter video games. *Human Communication Research, 30*, 361-375.

Shaw, A. (2010). Identity, identification and media representation in video game play: An audience reception study. (Unpublished Dissertation). University of Pennsylvania, Philadelphia.

Sundar, S. S., Oeldorf-Hirsch, A., & Garga, A. (2008). A cognitive-heuristics approach to understanding presence in virtual environments. In A. Spagnolli and L. Gamberini (Eds.), *PRESENCE 2008: Proceedings of the 11th Annual International Workshop on Presence* (pp. 219-228). Padova, Italy: CLEUP Cooperativa Libraria Universitaria Padova.

Teixeira, T., Wedel, M., Pieters, R. (2012). Emotion-induced engagement in Internet video advertisements. *Journal of Marketing Research, 49*, 144-159.

Trepte, S., & Reinecke, L. (2011). The pleasures of success: Game-related efficacy experiences as a mediator between player performance and game enjoyment. *Cyberpsychology, Behavior, and Social Networking,* 14, 555–557.

Turkay, S., Adinolf, S. (2010). Free to be me: a survey study on customization with World of Warcraft and City Of Heroes/Villains players. *Procedia - Social and Behavioral Sciences*, 2, 1840-1845.

Van Vugt, H. C., Hoorn, J. F., Konijn, E. A., & Dimitriadou, A. (2006). Affective affordances: Improving interface character engagement through interaction. *International Journal of Human-Computer Studies, 64*, 874-888. doi: 10.1016/j.ijhcs.2006.04.008.

Whitbred, R., Skalski, P., Bracken, C., & Lieberman, E. (2010). When richer is poorer: Understanding the influence of channel richness and presence on the introduction of a mission statement. *Psychology Journal*, 8, 115 – 139.

Wise, A., Chang, J., Duffy, T., & del Valle, R. (2004). The effects of teacher social presence on student satisfaction, engagement, and learning. Proceedings from the 6th International Conference on Learning Sciences (ICLS '04).

Wolf, M. J. P. (2003). Abstraction in the video game. In M. J. P. Wolf & B. Perron (Eds.), *The video game theory reader* (pp. 47-65). New York: Routledge.

Zaichkowsky, J. L. (1984). Measuring the Involvement construct. *Journal of Consumer Research*, 12, 341 -352.

Zimmons, P., & Panter, A. (2003). The Influence of rendering quality on presence and task performance in a virtual environment (Poster Abstract). Proceedings of IEEE Virtual Reality 2003. Los Angeles, CA: IEEE Computer Society.

Tables and Figures

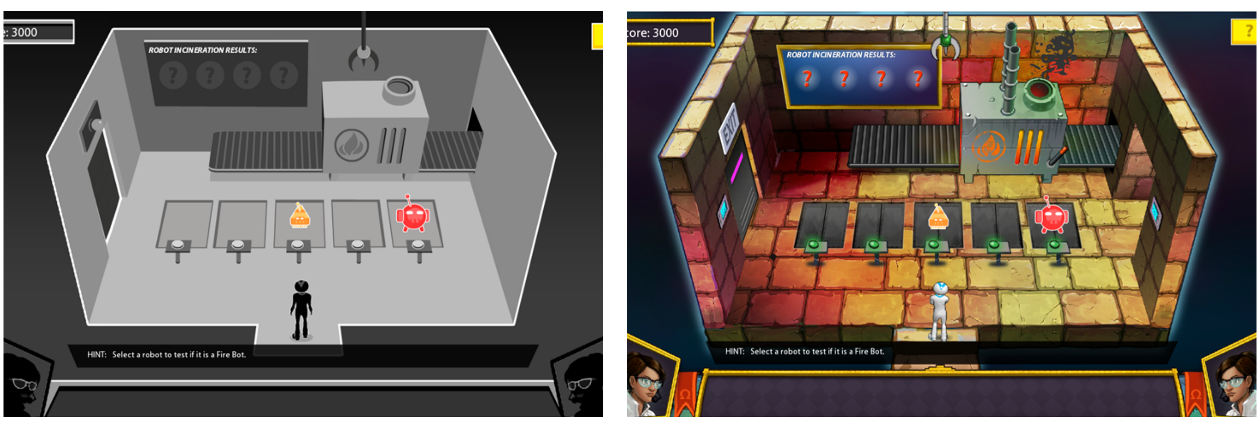

Figure 1: Simple and Detailed Art in a Puzzle Room

Table 1

Engagement Measure Descriptive Statistics Study 1

	N	Minimum	Maximum	Mean	SD
Self-Reported Measures					
Transportation	229	1.27	4.45	3.1	.57
Active Interpersonal	229	1	7	3.0	1.39
Social Richness	229	1	7	4.5	1.16
Social Realism	229	1	7	5.1	1.52
Mental Immersion	229	1	7	4.3	1.27
Objective Measures					
ABE	24	.88	1	0.97	.03
EDA Change	38	-.51	1.96	0.48	.60
Active Mouse Clicks	160	141	1245	218.7	105.08
Mouse Rest Percent	158	39.6	86.1	68.4	9.20

Table 2

Engagement Measure Descriptive Statistics Study 2

	N	Minimum	Maximum	Mean	SD
Self-Reported Measures					
Transportation	474	1.27	4.45	2.9	.60
Active Interpersonal	476	1	6.33	2.3	1.32
Social Richness	472	1	7	4.2	1.17
Social Realism	475	1	7	4.3	1.81
Mental Immersion	473	1	7	4.0	1.22
Objective Measures					
ABE	74	81.61	99.95	96.2	2.22
EDA Change	83	-1.8	1.9	.19	.53
Active Mouse Clicks	373	134	741	194.9	48.04
Mouse Rest Percent	340	33.6	80.7	64.14	8.49

Table 3
Standardized (Z-score) Engagement Measure Correlations, Study 1

	Trans-portation	**Active Inter-personal**	**Social Richness**	**Social Realism**	**Mental Immersion**	**ABE**	**EDA Change**	**Active mouse clicks**
Self-Reported Measures								
Transportation								
Active Interpersonal	.376***							
Social Richness	.589***	.370***						
Social Realism	.352***	.171**	.320***					
Mental Immersion	.733***	.407***	.673***	.349***				
Objective Measures								
ABE				-.423*				
EDA Change								
Active Mouse Clicks								
Mouse Rest Percent			.153*					-.179**

*** p < 0.01; ** p < 0.05; * p < .1

Table 4
Standardized (Z-score) Engagement Measure Correlations, Study 2

	Trans-portation	**Active Inter-personal**	**Social Richness**	**Social Realism**	**Mental Immersion**	**ABE**	**EDA Change**	**Active mouse clicks**
Self-Reported Measures								
Transportation								
Active Interpersonal	.324***							
Social Richness	.535***	.322***						
Social Realism	.435***	.172***	.236***					
Mental Immersion	.655***	.308***	.647***	.368***				
Objective Measures								
ABE	.397***		.344***		.545***			
EDA Change			.205*					
Active Mouse Clicks							-.208*	
Mouse Rest Percent	-.111**	-.102*			-.114**			

*** p < 0.01; ** p < 0.05; * p < .10

Bios:

Rosa Mikeal Martey is an Associate Professor in the Department of Journalism & Communication at Colorado State University. Her research focuses on games, identity, social interaction and digital media, and can be found in *New Media & Society*, *Information, Communication & Society*, and *Games Studies*, among others. Contact: (970) 491-6469, Colorado State University, CD 1785, Fort Collins, CO 80523-1785, USA; rosa.martey@colostate.edu

Kate Kenski (Ph.D., University of Pennsylvania) is an Associate Professor in the Department of Communication and School of Government & Public Policy at the University of Arizona. She has published over thirty articles, books, and book chapters on political communication, public opinion, and research methods. Contact: Department of Communication, University of Arizona, 1103 E. University Blvd., Communication Bldg #25, Tucson, Arizona, 85721, USA; 520-626-3066 (w); kkenski@email.arizona.edu; http://www.dingo.sbs.arizona.edu/~kkenski

James Folkestad is an Associate Professor in the School of Education at Colorado State University. His work focuses on educational technology, activity systems for learning, innovative inference game-based learning, and social network analysis. Contact: (970) 491-7823, 209 Education Building, Colorado State University, Fort Collins, CO 80523-1588, USA; James.Folkestad@colostate.edu

Laurie Feldman is a Senior Scientist at Haskins Laboratories in New Haven, CT and Professor of Psychology at The University at Albany, SUNY. She studies how people use and understand written and spoken languages, both native and non native. Contact: Department of Psychology, SS 399, The University at Albany, SUNY. Washington Ave. Albany, NY 12222, USA. 518 442 4842. lfeldman@albany.edu

Elana B. Gordis studies the effects of family and community violence on later development, with a particular focus on the role of physiological stress response systems including sympathetic, parasympathetic, and hypothalamic-pituitary-adrenal axis activity. She received her PhD in Clinical Psychology from the University of Southern California. Contact: Department of Psychology, University at Albany, SUNY, Social

Sciences 399, 1400 Washington Ave, Albany NY 12222, USA, 518-442-4820 (o); 518-442-2867 (f); egordis@albany.edu, http://www.albany.edu/psychology/20917.php,

Adrienne Shaw has been studying video games since 2005. Her research has been published in *Ada: Journal of Gender, New Media, and Technology, New Media and Society*, *Critical Studies in Media and Communication*, and *Games and Culture*, among others. Contact: Department of Media Studies and Production, 2020 N. 13th Street, Annenberg Hall, Philadelphia, PA, 19122, USA; (215) 204-6201; adrienne.shaw@temple.edu

Jennifer Stromer-Galley (PhD, University of Pennsylvania, 2002) is an Associate Professor in Information Studies. Her expertise is on communication processes and effects through digital communication channels. She has over 30 publications that focus on dimensions of digital media around influence, leadership, political communication, and training. Contact: School of Information Studies, Hinds 220, Syracuse University, Syracuse, NY 13244, USA; jstromer@syr.edu, +1-315-443-1823.

Ben Clegg is a Professor of Cognitive Psychology at Colorado State University. His research is on aspects of human performance, particularly how to apply principles from cognitive psychology to real-world tasks and skills. His work examines issues such as training, automation, workload, and situation awareness, sequence learning, and implicit learning. Contact: (970) 491-2211; 1876 Campus Delivery, Colorado State University, Fort Collins, CO 80523-1876, USA; Benjamin.Clegg@colostate.edu

Hui Zhang is a doctoral student in the Department of Journalismm & Technical Communication at Colorado State University. Her research focuses on health communication and information-seeking in digital media. Contact: Colorado State University, CD 1785, Fort Collins, CO 80523-1785, USA; huizhang@rams.colostate.edu

Nissim Kaufman is a Ph.D. candidate in mathematics at the University at Albany in mathematics and statistics. His work is on parametric models, cluster analysis, and Bayesian inferences, among other topics, and he has contributed to projects funded by DARPA, IARPA and others. Contact: University at Albany, ILS Room 262B, Social Science Building, 1400 Washington Avenue, Albany, NY 12222, USA; nissimkaufmann@yahoo.com

Ari N. Rabkin is currently a doctoral candidate with the Department of Psychology and research project assistant with the the Institute for Informatics, Logics, and Security Studies (ILS) at the University at Albany, SUNY. His research focuses on the intersection of biological/physiological measurement and psychological functioning, including user engagement during video games. Contact: Department of Psychology, Social Science 399, University at Albany, SUNY, 1400 Washington Avenue, Albany, NY, 12222, USA. +1-518-442-4684 (w); +1-518-442-4867 (f); arabkin@albany.edu.

Samira Shaikh has a predilection for puzzle games. A computational socio-linguist by training, she holds degrees in Computer Engineering and Computer Science. Her current research focus is on natural language processing, specifically aimed at understanding social roles and phenomena in social media conversations and artificial chat agents. Contact: SS 262, State University of New York - University at Albany, Albany, NY 12222, USA; +1 (518) 437 3606 (w) & +1 (518) 368 2175 (h); fax +1 (518) 442 2605; samirashaikh@gmail.com; http://www.ils.albany.edu.

Tomek Strzalkowski is Director of the Institute for Informatics, Logics and Security Studies and Professor of Computer Science at SUNY Albany. He has done research in computational linguistics, socio-linguistics, information retrieval, question-answering, human-computer dialogue, and formal semantics of natural language. He has directed research projects sponsored by IARPA, DARPA, EC and NSF, and published over a hundred papers and books. Contact: (518) 442-2608, University at Albany, ILS Room 262B, Social Science Building, 1400 Washington Avenue, Albany, NY 12222, USA; tomek@albany.edu