CHAPTER 10

USING GAME-BASED APPROACHES TO INCREASE LEVEL OF ENGAGEMENT IN RESEARCH AND EDUCATION

Joseph C. Toscano¹ , Andrés Buxó-Lugo¹,², and Duane G. Watson¹,²

"Minecraft is coming into different venues than what we thought was possible"

- Carl Manneh, Mojang, Co-founder and Managing Director

And Now for Something Completely Different

[A note from the editor...]

What you are about to read is a bit of a departure from the rest of the book. So far each chapter has focused on some element of teacher appropriation or evaluation of Minecraft. In the following pages, the authors have generously volunteered to publish their research as as a chapter in this book. Normally this kind of piece would be found in a journal and it will read as such. That said, the authors provide us glimpse of the range of impact Minecraft has in other fields, and I believe opens up an exciting host of possibilities for classroom use of Minecraft as a tool for action research, scientific method, and data collection. In that light, though different, this chapter fits perfectly.

¹ Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign

² Department of Psychology, University of Illinois at Urbana-Champaign

Using game-based approaches to increase level of engagement in research and education

Abstract

Improving student engagement in the classroom has the potential to increase academic success. Similarly, increasing participant engagement in experiments may allow researchers us to better study behavior. As a result, both educators and researchers would like to have tasks and activities that are interesting and engaging, and game-based approaches have been proposed as one way to achieve this goal. We describe a paradigm based on Minecraft that we have been using for basic research in psycholinguistics, the study of the cognitive processes underlying language use. This approach provides us with a task that is interesting for participants, elicits naturalistic language, and allows us to run controlled lab experiments. Moreover, it demonstrates that level of engagement can have large effects on behavior. This illustrates how platforms like Minecraft can provide us with a useful new tool for creating studies that are more engaging, allowing us to examine research questions that we have not been able to address using other techniques.

Context

Many researchers have argued that more engaged students show better learning outcomes (Skinner & Belmont, 1993; Finn & Rock, 1997; Marks, 2000), and as a result, teachers want to ensure that students remain engaged in the classroom. Level of engagement also plays a role in basic research: laboratory experiments are often needed to address certain questions, but these studies typically lack the naturalness of real-world environments and abstract away from many details (Clarke, 1997; Wertsch, 1998). This can make the experiment much less engaging, and it can unintentionally change participants' behavior (Orne, 1962). Thus, a major challenge for both researchers and teachers is to find ways for participants and students to be highly engaged.

Several approaches for increasing student engagement, in particular, the use of technology in classrooms (Kulik & Kulik, 1991) and game-based approaches to teaching and learning (VanSickle, 1986; Moeller, Cootey, & McAllister, 2006), have been proposed. Though debate about the effectiveness of some techniques continues (see Randel, Morris, Wetzel, & Whitehill, 1992, and Vogel, Vogel, Cannon-Bowers, Bowers, Muse & Wright, 2006, for reviews), previous work has shown that game-based approaches can increase engagement and lead to better learning outcomes (Squire, 2003; Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005). Given this, we were interested in whether game-based tasks also provide a way to increase participant engagement in the lab.

To test this, we developed a platform for running experiments using Minecraft (see Chapter 2). Our research focuses on language use, and we used the game to create an experimental task that is more interesting and engaging than those that are typically used in our field. Often, in lab-based studies, participants do not use natural language–for example, their speech lacks the typical prosody (details below) we see in real-world conversations. This creates a challenge for researchers who, on the one hand, are interested in studying natural language, and on the other, need to have the precise experimental control offered by a laboratory setting. Game-based tasks provide a potential solution to this problem, and as we show here, they allow us to create conditions where participants are more likely to use natural language (i.e., they produce speech that is more characteristic of everyday language use).

This study is not about Minecraft, per se, but rather, it provides an illustration of how Minecraft (and games like it) can be used as a tool for research (on language use, or on other topics), allowing us to study cognitive processes that we could not examine using typical laboratory approaches. Although our study did not focus specifically on learning, it may provide insights for educators who want to increase student engagement using similar approaches, as well as ideas for researchers that struggle with engaging study participants.

In the following sections, we discuss 'level of engagement' as it pertains to education and research in cognitive science, with a particular emphasis on our own field of psycholinguistics, the study of how humans produce and understand language. Then, we describe our gamebased task to illustrate how it allows us to run controlled experiments in the lab, while creating an environment that affords natural language use. Our study highlights how games like Minecraft provide a platform for achieving higher degrees of engagement in the laboratory, and more generally, how game-based approaches give us a tool for studying phenomena that we would not have been able to study any other way.

A Discussion on Validity in Psychological Research

One reason we are particularly interested in game-based approaches is that they can provide experimental control, and simultaneously, allow for more natural behavior. This allows us to study a variety of cognitive processes usings tasks that participants find interesting. Similarly, one way that teachers might increase student engagement is to make learning environments richer and more stimulating. Indeed, the environment can affect learning in a number of ways, and many researchers have discussed the importance of context on learning outcomes (e.g., Brown, Collins, & Duguid, 1989). This ties into the idea of studying phenomena in situ, in the way they occur in the real world, something that has been advocated by researchers in a number of fields (Clark, 1997; Hutchins, 1995), including education (Brown et al., 1989; Halverson & Clifford, 2006; Squire & Dikkers, 2012). Although the research questions we focus on here require studies in the lab, game-based approaches, in general, may also give us measures that more closely reflect behavior outside the lab.

The challenge faced by researchers is that it is often difficult to design studies that are constrained enough to give us experimental control, yet flexible enough that they reflect complex, real-world conditions. This is a general issue with trying to maximize the validity of a study–the extent to which we can draw valid conclusions from it–since different aspects of the study allow for different types of validity. For example, internal validity refers to the extent to which the results are not explained by other factors, something that is typically achieved via the precise control offered by lab experiments. Contrast this with external validity: the extent to which results generalize to other situations, something that is typically achieved through fieldwork and studies in real-world environments.

Internal and external validity are opposed to each other such that changes to a study that increase one tend to decrease the other. As a result, there is no "perfect" experiment that can completely maximize both. However, it is possible to find a balance between the two. Game-based approaches, in particular, provide a way to do this, since they allow us to create controlled virtual environments that afford more natural behavior than what we see in typical lab settings. These issues apply to research studies in a number of fields, and they have been particularly relevant in psycholinguistics. Research in this area spans a range of topics, from understanding how listeners recognize speech, to understanding how sentence structure is interpreted during reading, to understanding the representations and processes that underlie discourse and conversation.

A great deal of psycholinguistic research has used laboratory experiments. However, language produced in the lab can differ from real-world language in a number of ways. For example, conversational speech depends on a collaborative effort between interlocutors that take one another's perspective and goals into account (Schober & Clark, 1989; Hanna, Tanenhaus, & Trueswell, 2003; Nadig & Sedivy, 2002; but also see Keysar, Lin, & Barr, 2003; Yoon, Koh, & Brown-Schmidt, 2012), but speakers' goals may be very different in the real world than they are in a constrained lab experiment. In a conversational setting, speakers are generally motivated to be appropriately informative, since their communication has real-world consequences. On the other hand, in a laboratory setting they might be more motivated to simply accomplish the task quickly.

Given this, we need experiments that allow for natural language use in the lab. Virtual environments provide a platform that may allow us to achieve this goal. Part of the appeal of a sandbox-type game (like Minecraft), in particular, is that it allows us to quickly develop complex environments with minimal programming.

A number of previous studies have used various types of virtual environments to study cognitive processes. Virtual reality has been used to investigate phenomena ranging from visual and tactile perception (Atkins, Jacobs, & Knill, 2003) to the ability to detect changes in the environment (Triesch, Ballard, Hayhoe, & Sullivan, 2003). Game-based approaches, in particular, have been used to study processes like selective visual attention (Green & Bavlier,

2003). Other researchers have explored language use in virtual environments: Steinkuehler (2003), for example, describes how multiplayer online games can provide a tool for studying language. Together, these results suggest that virtual environments may offer a viable platform for studying cognition in a way that allows for more participant engagement. As we describe below, we found that they do.

Using a Computer Game for Research

We developed a technique using Minecraft (Persson & Bergensten, 2011)–specifically, a customized version for education called MinecraftEdu (Koivisto, Levin, & Postari, 2012)– to accommodate the need for greater participant engagement in our experiments. The goal was to create a richer environment and more interesting task, while still allowing us to address the research questions we were interested in. The game provides a platform for doing this, since it gives us an immersive environment that can be precisely controlled. Specifically, in our study, Minecraft provides a context in which we can observe particular aspects of language use.

We used this approach to investigate how certain characteristics of *prosody* are used in conversational speech (see Buxó-Lugo, Toscano, & Watson, submitted, for more details). Prosody includes several aspects of language that describe *how* words are produced (as opposed to *what* the words are), including properties like word stress and rhythm. For example, the statement "We're going to the store *now*?" is heard as a question when the talker's fundamental frequency (roughly corresponding to the pitch of her voice) rises for the word *now*, whereas the same sequence of words, "We're going to the store now.", is heard as a statement when it does not.

In Buxó-Lugo et al., we investigate how talkers convey information about *discourse status* in a conversation by making certain words prosodically prominent (i.e., emphasizing them). Discourse status refers to whether particular information is new in a conversation or has been mentioned previously. For example, a talker might say "Give me the *red* one", emphasizing the word *red* to contrast it with some other color or to indicate that they are referring to a new item that has not been mentioned previously in the conversation.

We were interested in which acoustic cues talkers use to convey discourse status in conversations. That is, how does the speech signal change when a talker wants to indicate to their partner that a particular piece of information in the conversation is new? Answering this question is important, not only for understanding the processes underlying language production and comprehension, but also for developing computer systems that take discourse status into account when recognizing speech.

Previous work has suggested that there are a number of acoustic cues that the talker might use and that listeners might take advantage of in interpreting discourse status (Wagner &

Watson, 2010). These include word duration (talkers tend to indicate new information by making words longer), intensity (discourse-new information tends to be perceived as louder), and different aspects of the talker's fundamental frequency (discourse-new information tends to be produced with higher fundamental frequencies). However, previous studies that have used typical laboratory tasks have found inconsistent effects when measuring these cues. Thus, it seems as though talkers do not reliably convey discourse status, and as a consequence, we might conclude that listeners would not be able to infer information about prosody in a conversation.

Our goal was to see whether this was really the case, or whether talkers simply do not produce realistic language in typical lab-based tasks. Specifically, we were interested in whether talkers would convey discourse status more reliably if the task was more engaging, making it more akin to natural conversation. Indeed, studies analyzing spontaneous speech, which lacks the control needed to satisfy internal validity, have found that talkers use these cues to a certain extent, though there has been disagreement about which cues provide the most information (Kochanski, Grabe, Coleman, & Rosner, 2005; Kim, Yoon, Cole, & Hasegawa-Johnson, 2006). Nonetheless, with more natural tasks, talkers may more reliably convey discourse status information to a conversational partner.

Experiment Design and Results

The experiment consisted of two tasks. One was a picture description task, in which participants saw sequences of colored squares on a computer display, and they were asked to read the sequence aloud from left to right (e.g., "red, black, pink"). Then, they saw another set of colors that was either (1) identical to the first, (2) a completely different set of colors, or (3) a set where only the second color changed (e.g., "red, **green**, pink"). This allows us to manipulate the discourse status of specific words (i.e., whether they were repeated or not) and measure the acoustic properties of the words to see which cues (if any) talkers use to convey differences in discourse status. This task is typical of those commonly used in lab experiments studying prosody.

We contrasted this with a second task using Minecraft. We created a Minecraft world in which two participants had to work together to cooperatively solve puzzles. The two "players" were seated in physically different rooms and communicated through headsets. The world was designed so that the players proceeded sequentially through a series of rooms, each of which contained a puzzle that they needed to solve. Each room had a locked door, and after solving the puzzle, the door would open, allowing the participants to proceed to the next room.

Participants were given a brief (5-10 minute) introduction to the game and a tutorial on how to control their character and interact with the environment. Then, they completed a series of 36 puzzles in the game. They had as much time as they needed to complete each puzzle,

and they continued until they reached the end or had been playing for one hour. Half of the puzzles (18 trials; 9 unique puzzles) served as fillers meant to keep the game interesting. To solve them, one participant generally had information that they needed to provide to their partner.

Figure 1 shows an example of one of these puzzles. Here, the two participants have to work together to move a block to a particular location. Pressing the colored buttons on the wall will move the block in different directions. In this case, the participant has to tell their partner to press the *blue*, *pink*, and *yellow* buttons in sequence to move the block to the correct location. After they do this, the door will open and they can proceed to the next puzzle. Other filler trials included games where participants needed to complete a maze, fire an arrow at a target, or guide a minecart to a particular location. Critically, each of these puzzles needed to be solved cooperatively by having the participants exchange information with each other, creating situations that allowed for natural conversation.



Figure 1. Example of a puzzle in the experiment from one of the participant's perspectives. The other participant's character (labeled as Player 1) can be seen in the top center of the figure.

The critical experimental trials consisted of puzzles where we manipulated the information status of specific words in the discourse, just as we did in the picture description task. These took the form of "combination locks" in which one player was given a code (a set of colors) and the other player needed to enter it in the correct sequence to unlock the door and exit the room. This allowed us to create trials in the game that corresponded to those in the more typical (but less interesting) picture description task described above.

Figure 2 provides an example of one of the "combination lock" trials. The two participants' characters are located in different rooms and cannot see each other. The participant on the left (Player 1) sees the sequence of *red*, *black*, and *pink* blocks on the wall. Her partner (Player 2) sees a set of colored buttons on his wall. To solve the puzzle and open the door, Player 1 tells her partner the first sequence ("red, black, pink"). After Player 2 presses those buttons in the correct order, a new set of colors appears on the wall in Player 1's room. She gives this new sequence to her partner ("*red*, *green*, *pink*") who presses the corresponding buttons, solving the puzzle and triggering the doors to open. Both players then move on to the next puzzle.

Importantly, we varied the color sequences in the same way as in the picture description task. The same colors, color sequences, and trial orders were used in both tasks. Thus, the only thing that differed was the task itself, where one task led to low levels of engagement (picture description), and the other was much more interesting (the Minecraft task).

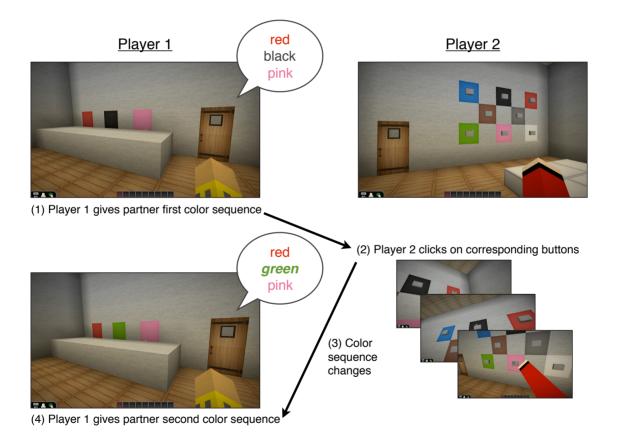


Figure 2. Example of the sequence of events for one of the "combination lock" (discourse status) puzzles in the experiment.

The transcript below illustrates a typical exchange between two participants in the Minecraft task on a trial in which the colors were different for each sequence (making all the colors in the second sequence new to the discourse):

Player 1: Ok, now my order is *black*, *red*, and umm, *blue*.
Player 2: Ok.
[The color sequence changes for Player 1.]
P1: Ok, now, I have *white*, *brown*, and *green*.
P2: Ok.
[The door to the room opens after Player 2 enters the second sequence.]
P1: Ok, off we go!

Here, we are interested in how the participants emphasize the second set of colors to distinguish it acoustically from the first set (e.g., the acoustic differences between the word *brown* in the example above and the word *brown* spoken on a trial in which that color was repeated).

Twenty-four pairs of participants were given the Minecraft task, and another 24 were given the picture description task. We found that in the (more typical) picture description task, participants did not provide very reliable information about discourse status. In contrast, in the Minecraft task, participants provided information to their partner using a number of acoustic cues. For example, differences in duration between discourse-new and discoursegiven words were larger in the Minecraft task (98 ms) than in the picture description task (30 ms; see Buxó-Lugo et al. for details about the results). Thus, talkers provided more information about the discourse status of these words in the game.

These results demonstrate that level of engagement matters. By providing participants with a more interesting task, language use was more natural, and we were able to measure properties of spoken language that we could not measure using typical laboratory approaches.

Discussion and Application

The game-based task provided us with the balance we needed in the experiment: it was constrained enough that we expected participants to say particular words (allowing us to make the necessary acoustic measurements) but unconstrained enough that participants engaged in natural conversations about the game. Although our task isn't completely free of any constraints (since participants are still instructed to solve the puzzles in the game), it offers them much more flexibility. This allows us to have control over the stimuli and conditions, but gives participants a more natural, and importantly, a more engaging task. The study described above represents a middle-ground between more typical laboratory experiments and analyses of real-world language that have no experimental control, satisfying (to some degree) the demands of both internal and external validity.

How do these results inform educational research and the utility of game-based approaches? The current example provides preliminary evidence that more engaging environments can have large effects on behavior, since we found that talkers provide information about prosody more reliably in the game-based task than in the less interesting picture description task. Given this, and the observation that level of engagement is related to student performance (Skinner & Belmont, 1993; Finn & Rock, 1997; Marks, 2000), similar approaches might translate to a variety of settings in education. For instance, our study could be used in a science class to teach students how to design an experiment: it illustrates concepts such as experimental control, independent and dependent variables, and randomization. The game provides a concrete way to teach students about these ideas, and this may make them more memorable.

Many of the other advantages that this approach provides, such as ease of use and flexibility for a variety of tasks, may also be useful for teachers (Table 1). As we have shown here, computer game approaches provide researchers with a tool for increasing participant engagement, which in turn, allows us to study basic cognitive processes. For teachers, increasing the level of student engagement is, in itself, beneficial and may lead to better learning outcomes.

	LABORATORY RESEARCH	CLASSROOM USE
Level of engagement	Higher	Higher
Environment	Richer stimuli	More stimulating learning context
Ease of use	High	High
Effects	More naturalistic behavior	Improved learning
Flexibility/Adaptability	Can be used to study a range of cognitive processes	Can be used for a range of subjects

Table 1. Game-based approaches in research and classroom settings.

The results of the experiment also mirror work in education on game-based approaches (Squire, 2003; Barab et al., 2005). With respect to questions about language and literacy, previous work has suggested that when students are able to choose topics in a game-based task, they perform at a higher reading level (Steinkuehler, Compton-Lilly, & King, 2010). The Minecraft task used in our experiment may have had a similar effect for us: by allowing participants to choose their topic of conversation within the game, they may have been more engaged and produced more natural language (in contrast to the picture description task, where their speech was much more constrained by asking them to simply read the color sequences).

This approach could be further extended to learning contexts. The experiment itself could provide a useful platform for teaching students about language processing. By participating in the experiment, students can gain a better understanding of the cognitive processes that underlie language production, as well as a better sense of how researchers study these processes.

These types of tasks could also be useful in language learning situations. If they produce a higher level of engagement than other approaches, they could lead to more successful outcomes. For researchers, this paradigm could also be adapted for studying other aspects of language processing, and experiments using this approach could be constructed to study many other cognitive processes (e.g., attention, working memory, executive function). Because of the flexibility of computer-game platforms like Minecraft, a wide range of environments can be constructed that are suited to studying specific phenomena. Thus, many of the same benefits that we have outlined here can be gained by researchers in other fields as well.

To conclude, we have demonstrated that a game-based approach can be used to increase participant engagement in laboratory studies, providing a valuable experimental paradigm for researchers. Thus, games like Minecraft are useful not only as teaching tools, but also for scientific research. They offer a flexible, easy to use platform for increasing level of engagement both inside and outside the classroom, and they open up a number of possibilities for studying behavior in a way that few other techniques can.

References

Atkins, J.E., Jacobs, R.A., & Knill, D.C. (2003). Experience-dependent visual cue recalibration based on discrepancies between visual and haptic percepts. *Vision Research, 43*, 2603-2613.

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology, Research, and Development, 53*, 86-107.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*, 32-42.

Buxo-Lugó, A., Toscano, J.C., & Watson D.G. (submitted). Effects of participant engagement on prosodic prominence.

Clark, H.H. (1997). Dogmas of understanding. Discourse Processes, 23, 567-598.

Finn, J.D., & Rock, D.A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology*, 82, 221-234.

Green, C.S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534-537.

Halvorson, R.R., & Clifford, M.A. (2006). Evaluation in the wild: A distributed cognition perspective on teacher assessment. *Educational Administration Quarterly, 42,* 578-619.

Hanna, J.E., Tanenhaus, M.K., & Trueswell J.C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49, 43-61.

Hutchins, E. (1995). Cognition in the Wild. Cambridge, MA: MIT Press.

Keysar, B., Lin, S., & Barr, D.J. (2003). Limits on theory of mind use in adults. *Cognition, 89,* 25-41.

Kim, H., Yoon, T., Cole, J., & Hasegawa-Johnson, M. (2006). Acoustic differentiation of L- and L-L% in switchboard and radio news speech. In R. Hoffman & H. Mixdorff (Eds.) *Proceedings of the Third International Conference on Speech Prosody 2006*, Dresden, Germany.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America, 118*, 1038-1054.

Koivisto, S., Levin, J., & Postari, A. (2012). MinecraftEdu [computer software]. Joensuu, Finland: Teacher Gaming.

Kulik, C.C., & Kulik, J.A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior, 7*, 75-94.

Marks, H.M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal, 37*, 153-184.

Moeller, R.M., Cootey, J.L., & McAllister, K.S. (2006). "The perpatos could not have looked like that," and other educational outcomes from student game design. In B.E. Shelton & D.A. Wiley (Eds.), *Educational Design & Use of Computer Simulation Games* (pp. 129-152). Rotterdam: Sense Publishers.

Nadig, A.S., & Sedivy, J.C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*, 329-336.

Orne, M.T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776-783.

Persson, M., & Bergensten, J. (2011). Minecraft [computer software]. Stockholm: Mojang.

Randel, J.M., Morris, B.A., Wetzel, C.D., & Whitehill, B.V. (1992). The effectiveness of games for educational purposes: A review of recent research. *Simulation Gaming*, *23*, 261-276.

Schober, M.F., & Clark, H.H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*, 211-232.

Skinner, E.A., & Belmont, M.J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85, 571-581.

Squire, K. (2003). Video games in education. International *Journal of Intelligent Simulations and Gaming, 2.*

Squire, K., & Dikkers, S. (2012). Amplifications of learning: Use of mobile media devices among youth. *Convergence, 18,* 445-464.

Steinkuehler, C.A. (2003). Massively multiplayer online video gaming as a participation in a discourse. *Mind, Culture, and Activity, 13,* 38-52.

Steinkuehler, C.A., Compton-Lilly, C., & King, E. (2010). Reading in the context of online games. In S. Goldman, J. Pellegrino, K. Gomez, L. Lyons, & J. Radinsky (Eds.), *Proceedings of the 9th International Conference of the Learning Sciences* (pp. 222-229).

Triesch, J., Ballard, D.H., Hayhoe, M.M., & Sullivan, B.T. (2003). What you see is what you need. *Journal of Vision*, *3*, 86-94.

VanSickle, R.L. (1986). A quantitative review of research on instructional simulation gaming: A twenty-year perspective. *Theory and Research in Social Education, 14,* 245-264.

Vogel, J.J., Vogel, D.S., Cannon-Bowers, J., Bowers, C.A., Muse, K. & Wright, M. (2006). Computer gaming and interactive simulations for learning A meta-analysis. *Journal of Educational Computing Research*, *34*, 229-243.

Wagner, M., & Watson, D.G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes, 25*, 905-945.

Wertsch, J.V. (1998). Mind as Action. New York: Oxford University Press.

Yoon, S.O., Koh, S., & Brown-Schmidt, S. (2012). Influence of perspective and goals on reference production in conversation. *Psychonomic Bulletin and Review*, *19*, 699-707.