<u>What are mixed reality and virtual "embodied" lessons?</u>
<u>And while we're at it...how do I design one?</u>

Dr. Mina C. Johnson-Glenberg and the Embodied Games for Learning Lab
Arizona State University
mina.johnson@asu.edu;
or mina@eglgames.com www.eglgames.com

First published as invited article on ARVEL SIG SUPERNEWS Website in 2012, updated August 2014 for ARIEL Special Interest Group website.

What is a Mixed Reality learning environment?

Embodiment and Mixed Reality
As touchscreen technologies and new immersive learning environments come on the market, the term "embodied" is heard more often, yet its meaning is usually left opaque. We feel it is time for the education community to clarify the meaning of embodied, especially as it relates to emerging learning technologies or 'new media". To this end, this article first defines the term "embodied" and how our lab uses it in relation to Mixed Reality environments and to purely digital, or virtual, environments. The paper then expands on a taxonomy for embodiment, our design principles for embodied lessons, and ends with assessment ideas and some questions to guide future research.

The emergence of new educational technologies and interfaces that accept natural physical movement (i.e., gestures, touch, body positioning) as input into interactive digital environments is an exciting development. One such category of new technologies is referred to as "Mixed Reality" (MR), this involves the "merging of real and virtual worlds" (Milgram & Kishino, 1994). That is, digital components like projected graphics on a floor or wall are synced with real world tangible objects, e.g., trackable hand-held wands (see Figure 1) and students interact with both.

Figure 1 Rigid Body wand moving virtual wavelengths on floor projection of SMALLab.

Figure 1 displays an example of a student using a 3D-printed rigid body tracking wand to control floor projections. In this case, two students work together to control the length of a light wave. Our lab, and others, have cited the strong potential of these technologies to engage learners of all ages in immersive experiences that enhance education (Birchfield & Johnson-Glenberg, 2010; Chang, Lee, Wang, & Chen, 2010; Hughes, Stapleton, Hughes, & Smith, 2005; Johnson-Glenberg, Birchfield, & Uysal , 2009; Johnson-Glenberg, Birchfield, Megowan-Romanowicz, Tolentino, & Martinez, 2009; Johnson-Glenberg, Koziupa, Birchfield, & Li, 2011; Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2014; Lindgren & Moshe1, 2011).

Embodiment and Cognition.
The working hypothesis in our lab is that human cognition is embodied cognition. This means that cognitive processes are *deeply rooted and come from the body's interactions with its*

physical environment (Wilson, 2002). Multiple research areas now support the tenet that embodiment is an underpinning of cognition. The various domains include (but are not limited to): neuroscience (Rizzolatti & Craighero, 2004), cognitive psychology (Barsalou, 2008; Glenberg & Kaschak, 2002; Glenberg, 2010), math (Lakoff & Nunez, 2000), gesture (Hostetter & Alibali, 2008; Goldin-Meadow, Cook, & Mitchell, 2009), expert acting (Noice & Noice, 2006), and dance (Winters, 2008). It follows that all thought - even the most abstract - is built on the foundation of physical embodiment. Pulvermüller and Fadiga's (2010) review of fMRI experiments demonstrate that when reading words related to action, areas in the brain are activated in a *somatotopic* manner. For example, reading "lick" activates motor areas that control the mouth, whereas reading "pick" activates areas that control the hand. This activation is part of a parallel network representing 'meaning' and shows that the mappings do not fade once stable comprehension is attained. Motoric codes are still activated during linguistic comprehension in adulthood. If these codes are still active in the adult brain, then we think a good design principle would be one that included the modality of gesture or kinesthetics to help people remember new content.

What exactly is embodied learning?

The vague and fluctuating terminology associated with embodiment in Mixed Reality and Virtual education does not help the cause. Every tap on a screen is not "embodied". The domain of education could benefit from a taxonomy of embodiment. We created one as a guide for educational new media designers who wish to incorporate components of embodiment into their curricula. It has been designed so that researchers and users of these systems can use a common language to discuss how embodied certain platforms or modules are.

At the simplest level to say that a lesson is highly embodied means the human **gesture or whole body movements are aligned with the mediated content to be learned.** This alignment, or what others call congruency (Segal, 2011), means the movements map to key concepts. While bodily activity has the potential to strengthen memories and seed learning (Cook, Mitchell, & Goldin Meadow, 2008), these activities or movements must be *designed* such that they facilitate the desired instances of understanding. If you are designing a lesson on gears in the mouse-driven world it seems natural to use a "click" to start the gear train system turning. However, when using the Microsoft *Kinect* joint-tracking sensor as input students can now spin their hands in the direction the input gear should go. Students can intuitively control the speed of the gear train. This "spin" movement is congruent with the material to be learned. If we had designed the interaction with a default type of "push" gesture on the *Kinect* that would not be congruent. Figure 2 is a live shot of two players playing the biking game "Tour de Force" created by SMALLab Learning. By spinning the tracked wrist joint around the pivot point of the shoulder students are able to change the diameter of the input gear. The goal is to finish the course first; however, along the route gear sizes must be changed. Imagine when a steep hill is reached. Many of our younger students hold the misconception that "bigger is always better". With practice they realize that the largest gear will not get them up the steep hills. This is the real world (ok...virtual world) introduction into why they should care about the concept of mechanical advantage. In Tour de Force, gears changed diameter but direction was always the same. In the Winching Game (Figure 3) the goal was to winch up boulders from a

Figure 2 Spinning the arm while the Kinect uses the gesture to drive the bike gears.

mine and the diameter and *direction* of the spin was crucial for winning, thus Winching was harder to master.

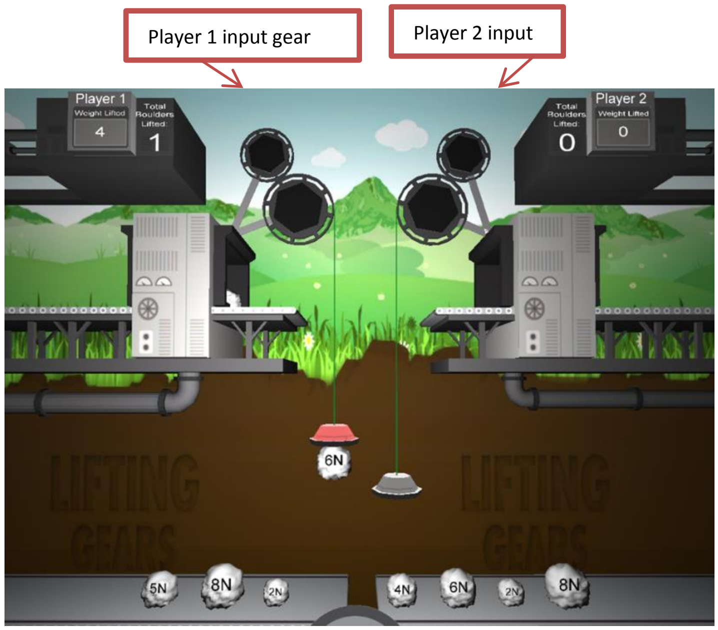

Figure 3. The Winching Game. Top gear is input gear and learners spin their right arms to winch up boulders of varying mass from below with the red magnet.

The Taxonomy.

Designers of new media using new types of input devices must make decisions about how embodied a lesson is. In the gears examples, learners' body movements mapped to the direction and speed of the gears. The gear gesture for diameter – distance between right wrist to right shoulder - is congruent to the educational insight that is crucial to the lesson, i.e., mechanical advantage and the fact that smaller input gears are needed to get up steeper hills and lift heavier rocks. While bodily activity has the potential to seed learning or act as a prime (Hostetter & Alibali, 2008), we should remember embodiment is not binary (either present or not). The degree of embodiment in an entire lesson will lie on a continuum. We propose a tractable taxonomy with four degrees. The 4th degree is the highest. More detail about the taxonomy can be found in the *Journal of Educational Psychology* (Johnson-Glenberg, Birchfield, Koziupa, & Tolentino, 2014). The proposed four degrees of embodiment depend on the magnitude of three key components:

1) Amount of motoric engagement.
2) Gestural congruency– how well-mapped the evoked gesture is to the content to be learned.
3) Perception of immersion – this is often correlated with perceived size of visual display (e.g., screen size).

These three components can range from low to high, and their interactions result in four levels of embodiment. The overlap intersections are fuzzy and are not meant to be overly restrictive—this is more of a common lexicon for researchers and designers. The degrees should not be viewed as four rigid bins. First, we describe the highest.

4th degree= Content at the 4th degree includes locomotion which results in a high degree of motoric engagement. Gestures within an environment that includes locomotion need to be consistently mapped congruently to the content being learned, and the learner **perceives** the environment as very immersive. The learner would report feeling "place illusion" (Slater, 2009). This is the strong illusion of being in a place in spite of the sure knowledge that you are not there. …it is a qualia …" (p. 3550). This sensation could be achieved in *SMALLab* with its large floor projection, or with wrap around head-mounted displays (HMD), or even in a CAVE environment (Cruz-Neira, Sandin, & DeFanti, 1993) or larger portable domes.

3rd degree = Content at the 3rd degree does not include sustained locomotion, however, the whole body could still be engaged while in the same location (as in working with a *Kinect* and a large Interactive Whiteboard- IWB). There must be some degree of gestural congruency in the system. The learner perceives the environment as somewhat immersive, thus the display would need to be larger than a computer monitor.

2nd degree = Content at the 2nd degree is usually designed for the seated learner, there is upper body movement but it is constrained to fingers and arms. The interfaces should be highly interactive (i.e., learner affects actions on screen), however, gestural congruency is not always a given in this level. The smaller monitor-sized display means that the learner does not experience Place Illusion or "presence". We created a centripetal force lesson once where the learner spun the virtual tethered yo-yo on the screen with the mouse as the input. Learners watched on a typical-sized computer screen, with the movement congruency it was at the high end of the 2nd degree but because the participant was seated and not really swinging a tracked yo-yo overhead it would not be a 3rd or 4th degree lesson (see Johnson–Glenberg, Birchfield, et al., 2014 for more).

1st degree = Content at the 1st degree is also designed for the seated learner. There is primarily only finger movement for a mouse or keyboard-driven system. At this level the learner primarily observes videos/simulations. There is no real gestural congruency, e.g., hitting the space bar starts the simulation. With a typical monitor as the visual display there is no sense of Place Illusion. However, a *well*-designed simulation that includes animations from a human POV and/or human-like avatars is embodied in that humans appear to possess a mirror neuron system designed to learn from such visual input and many of the same sensory and emotional channels are activated when watching as when doing an action.

Example of Mixed Reality. We are most familiar with the mixed reality environment of *SMALLab* (Situated Multimedia Arts Learning Lab) first created at the Arts, Media and Engineering department at ASU. This is a large scale 15 x15 x 15 foot platform with 12 OPTITRACK infrared cameras that can track up to four handheld objects. In Figure 4, two students are exploring chemistry titration- click link to see a [video]() of a high school teacher using this scenario. The students use the wands in an embodied up-and-down motion like pipettes to drop molecules into a virtual flask in the center.

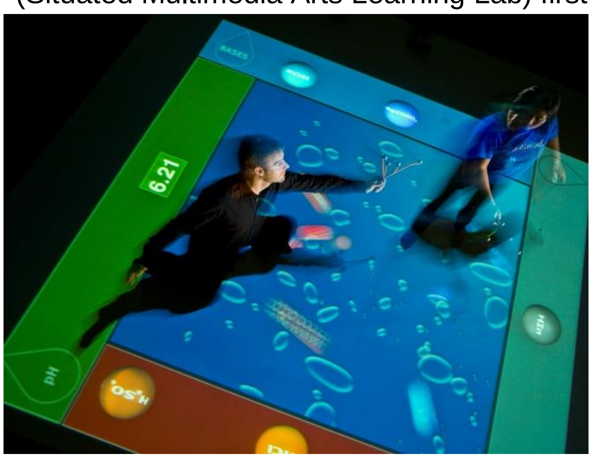

Figure 4. Students in the Titration Chemistry scenario of SMALLab.

What is special about locomotion? SMALLab allows participants to walk or locomote. Campos et al. (2000) contend that locomotion is still important and hugely relevant in the adult world. The broad-based and context-specific psychological reorganizations set in place via toddler locomotion have powerful consequences, and after infancy "can be responsible for an enduring role in development by maintaining and updating existing skills". With parallax and full body physical engagement a powerful signal is attached to the content to be learned, that is going to be part of the new memory trace. Although *SMALLAb* engenders a high degree of embodiment (usually 4th), new affordable motion capture sensors are now making it possible to quickly create applications for embodiment and get those out to schools at scale. Thus, the

Embodied Games for Learning lab has begun to focus on more cost-effective and portable motion capture technologies.

Multimedia design for embodied lessons.

We design for a new generation of motion capture technologies like the *Kinect* or the soon-to-released embedded 3D cameras in laptops. The lab adheres to several design principles honed over the years.

Make it

- **Embodied** – with as much gestural congruency as possible
- **Socio-collaborative -** build for the observing students as well, build in places for reflection
- **Generative** – learner is interactive and constructive
- Wrap it in **narrative** – make them care
- Give **immediate performance feedback**
- Level up in **cycle of expertise**
- Design for **user-created content -** Students should be producers not just consumers. (This is only moderately hard to build into a game – but it is *very* hard to get the teachers to use in-game editors and insert student-created content. This thorny subject is for another manuscript...)

Example of a virtual embodied educational game- Alien Health. We provide a short example of an embodied game created for a whole classroom to instruct in nutrition science. The full article with significant gains seen for the embodied group (on the delayed tests) can be found at Games for Health Journal (Johnson-Glenberg, Savio-Ramos, & Henry, 2014). Goldin-Meadow (Goldin-Meadow, 2011) posits that additional motor traces in memory may be created during gesture-based instruction. Perhaps it is these multi-modal traces that enhance the retention of new information? Our hypothesis is that practicing better food choices while actively choosing via congruent gestures in a fun, game-like environment would lead to increases in learning and retention. We built the first version of Alien Health in *SMALLab* and then transferred the game to a vertical one-wall projection and used the *Kinect* as the motion-capture sensor.

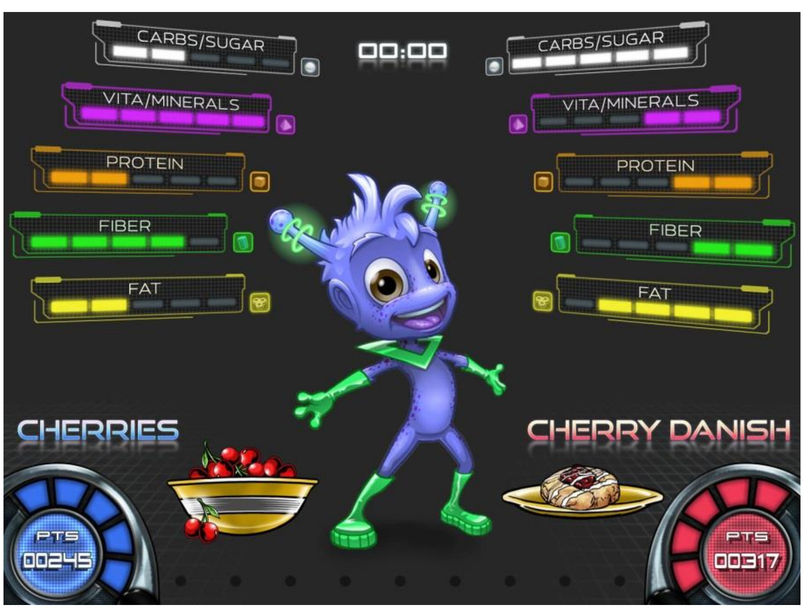

Figure 5. Screenshot from the Alien Health exergame.

Figure 5 shows the interface from the latest Level One of the game. The game was designed to teach about the five nutrients and optimizers in common foods and also to encourage discourse between the players. This level is based on a classic forced choice task. The backstory (narrative) is, “You have awoken to find an alien under your bed. He is hungry but you cannot communicate. You must figure out which foods make him feel better”. In the opening shot the alien is hanging his head and looking tired. The two players must come to the same conclusion regarding which of two foods to place in the alien's mouth. The *Kinect* tracks the placement of their hands. The act of placing the virtual food item in the alien's mouth has gestural congruency. After feeding him, players perform several exercises (e.g., jumping jacks, arm circles, etc.) and the *Kinect* gives them feedback on the quality of the exercise. In Level Two, they pick nonfoods and in Level Three they are introduced to a dynamic USDA *MyPlate* icon. The free beta version of the game and professional development kit can be downloaded from http://eglgames.com/product/alien-health/

How should we assess learning? What about in-game or process assessments?

We have used paper and pencil tests for decades and are very ready to move away from them. Although these tests have helped us to discover that embodiment may aid with **retention** of knowledge, and that learning effects might not be present on immediate posttests. For example, Johnson-Glenberg et al. (submitted) found that when randomly assigned participants showed up one week later for a retention test, those who had been in the **high** embodied condition retained significantly more physics knowledge. The learning platform did not matter. This difference was not evident on immediate posttest. Interestingly, our studies in classrooms have shown immediate effects for learning in *SMALLab* (Johnson-Glenberg, Birchfield, et al., 2014), but content was not always held exactly constant.

With the Alien Health game randomized control lab trial we used a traditional paper and pencil test as well to assess nutrition knowledge. Again, we found no significant difference at immediate posttest between the control versus embodied conditions (see test in the appendix of the article). However, we did see a crossover interaction from posttest to follow-up that approached significance, $F_{(19)=}$ 3.96, $p < .058$. Again, the embodied experimental group outperformed the control group for knowledge retention. We believe this may have to do with memory consolidation and sleep (Walker & Stickgold, 2004), but more research needs to be conducted to confirm how embodiment interacts with retention effects.

Even though paper and pencil tests still have uses, the field should be designing new metrics for new media. Our engaging, immersive simulations and games can remain engaging and immersive and still contain components of assessment (this is sometimes called stealth assessment). As a pilot study, we looked at the results of two small classes (with one teacher) as they played through the series of gears games. These are front-of-the-class *Kinect* games where dyads come up to perform. Day one was spent on a warm up gears game to learn the mechanics of the arm spin. Day two was spent on the Tour de Force biking game (Figure 2). Day three was spent on the Winching game (Figure 3). Students spun their arms around the pivot of their shoulder to drive the games. There were 22 7^{th} graders who completed both pre and posttests (the paper and pencil kind) assessing gears and mechanical advantage knowledge. The in-game process data were how many times each student switched gear diameters during play. We assumed that a more stable profile – that is, a player with fewer gear switches would have better comprehension. If they truly understood the principle of mechanical advantage, there was no need to keep changing gear diameter back and forth (wider, smaller, wider...) until the object moved. On the paper and pencil tests the students displayed significant gains in learning immediately after the intervention (there was no control group). We were primarily interested in how the in-game gear switching correlated with the tests. We predicted negative correlations, that is, the high switchers would do worse on the paper and pencil tests.

The valence and magnitude of the correlations between gear switches and test seemed to depend on the time of test and difficulty of the game. For the easier first game (Biking), switching data significantly predicted pretest score, $r = -.41$, but not posttest score. For the harder game (Winching), number of switches were highly negatively correlated with both pre and posttests showing that in-game data can perhaps provide a window into how students might perform on pre and posttests if the game is sufficiently challenging. (Johnson-Glenberg, Birchfield, & Megowan-Romanowicz, submitted). This is a promising intersection of immersive gesture-based STEM instruction using stealth assessments which may one day replace more traditional paper and pencil tests.

Mixed Reality and Virtual learning have made tremendous strides in the past five years and new technology is driving new designs and assessment tools. It is an extremely exciting time to be in the field. There are several questions we have at this juncture and would enjoy hearing your comments or answers; please feel free to email mina@eglgames.com to continue the discussion.

Future Questions

- How do subject variables interact with the in-game data? Is adding the modality of **gesture infelicitous for some learners**? Perhaps it is more difficult for low prior knowledge students to integrate all the input? Aptitude by treatment interactions have yet to be explored adequately in new media immersive learning spaces.
- Are there **limits to what can be embodied**? M. Eisenberg (personal correspondence, 2012) asks, "How come no one has embodied a Laplace transformation?"
- How can we more seamlessly **integrate assessment into the games**? How might learning gains be influenced by game components and competition? How can we assess and model the efficacy and quality of whole class collaboration (this is similar to Agent Based Modeling). Our field needs new metrics to **understand knowledge change at the multiple and simultaneously varying grain sizes and time scales associated with mediated classroom learning**.
- Can we capture the magic of, and then create more, **"aha" moments**?
- How can new media designers make sure there is **time for reflection built into the content?** Should we try to carve out time for solitary reflection? Will it just feel like dead time for some students?
- How can we better support teachers to take the necessary time to **encourage their students to create more media content** (e.g., draw and scan a culturally relevant food item into the Alien Health game)? If a whole class could then play with student-created content all would feel more agency. It is a powerful incentive for learning because no one wants to "be the slacker" in front of peers. For teachers to master these types of in-game editors it takes a different kind of professional development (PD) than we have been able to provide thus far. What is the best way to scale great new media PD for curious teachers?

References

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59,* 617–645.

Birchfield, D., & Johnson-Glenberg, M. C. (2010). A next gen Interface for embodied learning: *SMALLab* and the geological layer cake. *International Journal of Gaming and Computer-mediated Simulation, 2*(1), 49-58.

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., Witherington, D. (2000). Travel broadens the mind. *Infancy, 1* (2), 149-219.

Chang, C-W., Lee, J-H., Wang, C-Y., & Chen, G-D. (2010). Improving the authentic learning experience by integrating robots into the mixed-reality environment*. Computers & Education, 55*(4), 1572-1578.

Cruz-Neira, C., Sandin, D. J., & DeFanti. T. A. (1993). Surround-Screen Projection-based Virtual Reality: The Design and Implementation of the CAVE, *SIGGRAPH'93: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 135–142, DOI:10.1145/166117.166134

Glenberg, A. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 586–596.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science, 20*(3), 267-272.

Hostetter, A. B., Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin and Review, 15,* 495–514.

Johnson-Glenberg, M., Birchfield, D., Tolentino, L., & Koziupa, T. (2014). Technology-based embodied learning environments: Two STEM studies. Journal of Educational Psychology

Johnson-Glenberg, M. C., Birchfield, D., & Uysal, S. (2009). *SMALLab*: Virtual geology studies using embodied learning with motion, sound, and graphics. *Educational Media International, 46*(4), 267-280*.*

Johnson-Glenberg, M. C., Birchfield, D., Megowan-Romanowicz, C., Tolentino, L., & Martinez, C. (2009). Embodied Games, Next Gen Interfaces, and Assessment of High School Physics. International Journal of Learning and Media,1, 2.

Johnson-Glenberg, M. C., Koziupa, T., Birchfield, D. & Li, K., (2011). Games for Learning in Embodied Mixed-Reality Environments: Principles and Results. Proceedings for Games, Learning, and Society Conference, Madison, WI.

Lakoff, G., & Núñez, R. (2000). *Where mathematics comes from.* New York: Basic Books.

Lindgren, R. , & Moshell, J. M. (2011). Supporting children's learning with body-based metaphors in a mixed reality environment. *Proceedings of the Interaction Design and Children Conference,* 177-180.

Milgram, P., & Kishino, A. F. (1994). Taxonomy of mixed reality visual displays. *IEICE Trans. Information and Systems, E77-D* (12), 1321-1329.

Noice, H., & Noice, T (2006). What studies of actors and acting can tell us about memory and cognitive functioning. *Current Directions in Psychological Science*, *15*, 14–18.

Pan, Z., Cheok, A. D., Yang, H., Zhu, J., & Shi, J. (2006). Virtual reality and mixed reality for virtual learning environments. *Computers and Graphics, 30*(1), 20–28.

Pulvermüller, F., & Fadiga, L. 2010. Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11* (5), 351-360.

Rizzolatti G, & Craighero L. (2004). The mirror-neuron system. Annual Review of Neuroscience, (27), 169–192.

Segal, A. (2011). *Do gestural interfaces promote thinking? Embodied interaction: Congruent gestures and direct touch promote performance in math.* (Unpublished doctoral dissertation). Columbia University, New York.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transcripts of the Royal Society (Lond B Biol Sci)* (364) 1535, 3549-3557.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9,* 625–636.

Winters, A. F. (2008). Emotion, embodiment, and mirror neurons in dance/movement therapy: A connection across disciplines. *American Journal of Dance/Movement Therapy, 30*, 84-105.

Goldin-Meadow, S. (2011). Learning through gesture. . *WIREs Cognitive Science, 2*, 595–607. doi: 10.1002/wcs.132

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin and Review, 15*, 495-514.

Johnson-Glenberg, M. C., Birchfield, D., Koziupa, T., & Tolentino, L. (2014). Collaborative Embodied Learning in Mixed Reality Motion-Capture Environments: Two Science Studies. *Jouranl of Educational Psychology, 106*(1), 86-104. doi: 10.1037/a0034008

Johnson-Glenberg, M. C., Savio-Ramos, C., & Henry, H. (2014). ‘‘Alien Health’’: A Nutrition Instruction Exergame Using the Kinect Sensor. *Games for Health Journal: Research, Development, and Clinical Applications, 3*(4), 241-251. doi: 10.1089/g4h.2013.0094

Walker, M. P., & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron, 44*(1), 121-133.