Shortened version of original:
Johnson-Glenberg, M. C. (2019). The Necessary Nine: Design Principles for Embodied VR and Active STEM Education (pp. 83-112). In P. Diaz, A. Ioannou, K.K. Bhagat, & J.M. Spector (Eds.), *Learning in a Digital World: Perspective on Interactive Technologies for Formal and Informal Education.* Singapore, Springer. https://link.springer.com/chapter/10.1007/978-981-13-8265-9_5

The *Necessary Nine*: Design Principles for Embodied VR and Active STEM Education

Mina C. Johnson-Glenberg

Abstract: This chapter explores the two profound affordances of VR for learning; namely 1) the sense of **presence** attendant with immersive VR, and 2) the active learning associated with **movement/gestures in a three dimensional virtual world**. The chapter highlights several theories supporting embodied education and two examples of mediated STEM lessons which have been designed to maximize active learning. The first example explores the journey of redesign when a 2D tablet game is transformed into a 3D immersive VR lesson. The second example highlights how the new generation of hand controllers in VR can be used with constructivism to scaffold complex topics (chemistry and fireworks). The chapter ends with a set of optimal design principles for immersive VR in STEM education. The most important are called the *Necessary Nine*.

Key words: Virtual reality, VR, embodiment, STEM education, multimedia design principles, XR

1. The Two Profound Affordances

For several decades, the primary input interfaces in educational technology have been the mouse and keyboard; however, those are not considered highly embodied interface tools (Johnson-Glenberg, Birchfield, Koziupa, & Tolentino, 2014). Embodied, for the purposes of education, means that the learner has initiated a physical gesture or movement that is well-mapped to the content to be learned. As an example, imagine a lesson on gears and mechanical advantage. If the student is tapping the *s* on the keyboard to make the gear spin that would be considered less embodied than the student spinning a fingertip on a screen to manipulate a gear with a synchronized velocity. With the advent of more natural user interfaces (NUI), the entire *feel* of digitized educational content is poised to change. Highly immersive virtual environments that can be manipulated with hand controls will affect how content is encoded and retained. Now learners can spin a virtual hand crank with full arm movements (circles) and engage with 3D complex gear trains from any vantage point desired. One of the tenets of the Embodied Games lab is that doing actual physical gestures in a virtual environment will have positive, and lasting, effects on learning in the real world. Tremendous opportunities for learning are associated with this latest generation of virtual reality (VR) (Bailenson, 2018) and one of the most exciting aspects of VR is its ability to leverage interactivity (Bailenson et al., 2008).

Immersive and interactive VR is in its early days of educational adoption. Now that many of VR's affordability and sensorial quality issues are being addressed, it is reasonable to assume that VR experiences will become more ubiquitous in educational settings. When the demand comes, the community should be ready with quality educational content. There are few guidelines now for how to make optimal educational content in VR; this chapter will begin by explicating several relevant pedagogical theories. The chapter includes two case studies of lessons that have been built already, and it ends with tenable design principles.

First, what makes VR special for learning? Two attributes of VR may account for its future contributions to education. These we call the two *profound affordances*. The first profound affordance is the *feeling of presence* which designers must learn to support, while not overwhelming learners. Slater and Wilbur (1997) describe presence as the feeling of being there. It is a visceral transportation that, in many individuals, occurs immediately; when surrounded in 360 degrees by the virtualized unreal environment, players often lose sense of time. The second profound affordance pertains to *embodiment and the subsequent agency associated with manipulating content in three dimensions.* Manipulating objects in three dimensional space gives a learner unprecedented personal control (agency) over the learning environment. Gesture and re-enactments using the hand controls (and tracked fingers) should increase agency and positively impact learning. The basis for this prediction is the research on embodiment and grounded cognition (Barsalou, 2008). Although other methods for activating agency can be designed into VR learning environments (e.g., using eye gaze and/or speech commands), it may be the case that gesture plays a special role. Gesture kinesthetically activates larger portions of the sensori-motor system and motoric pre-planning pathways than the traditional modalities for learning (i.e., the visual and auditory). Gesture may lead to stronger memory traces (Goldin-Meadow, 2011). Another positive attribute of engaging the learner's motoric system via the hand is that the use of hand controls is associated with a reduction in cybersickness (Stanney & Hash, 1998).

VR for education should take full advantage of 3D object manipulation using the latest versions of handheld controllers (as well as, gloves and in-camera sensors to detect joints, etc.). The domain of gesture analytics in 3D is an area in need of more research and evidence-based design guidelines (Laviola, Kruijff, McMahan, Bowman, & Poupyrev, 2017)**.** This shortened chapter focuses on design practices that the author has learned from creating content in mixed and virtual realities over the past 12 years. An early, and evolving, set of design principles for VR in education is provided at the end, with the hope is that the guidelines will assist this nascent field as it matures.

1.1 We All on the Same Vocabulary Page?

Below are different terms, used by different communities. It makes sense to make sure we are all on the same page. This section defines some terms that are in flux in the field: VR, presence, agency, and embodiment.

1.1.1 VR

In this chapter, the term VR refers to an immersive, 360° experience, usually inside a headset, where the real world cannot be seen. In VR, the learners can turn and move as they do in the real world, and the digital setting responds to the learner's movements. *Immersive VR* systematically maintains an illusion of presence, such that learners feel their bodies are inside the virtual environment. Being able to see evidence of the real world, even in the periphery, would mean the platform should be deemed either augmented or mixed reality (AR/MR).[1] A

[1] No space is devoted to CAVES in this chapter (environments with projected wall surfaces, or cubes, where reality is never present) because the cost of a CAVE is still prohibitive for most educational settings.

three dimensional object or avatar displayed on a regular-sized computer monitor is never "VR". It is preferred that PC monitor-supported content be referred to as mediated or digital environment, even if the user can scroll the viewer/screen in 360°; the term VR should be reserved for immersive VR experiences where no real world components are visible.

1.1.2 Presence

The term, presence, as it relates to education is also defined in a glossary by Dede and Richards (Dede & Richards, 2017). Presence is a… "particular form of psychological immersion, the feeling that you are at a location in the virtual world" (p.5). The sensations are reported to be quite visceral. In a full immersion headset experience, the feeling of being in a different location is systematic and usually instantaneous. The presence associated with VR is one of the most immediate and well documented phenomena. Thus, **presence is deemed the first profound affordance of VR**. Several surveys are available for assessing the amount of presence in a mediated experience (Makransky, Lilleholt, & Aaby, 2017; Slater & Wilbur, 1997).

1.1.3 Agency

Immersive VR has the ability to immediately transport the user to a heightened emotional space that can have positive effects on attention and engagement; this is one reason why educators believe that learning will be positively affected. Whenever users feel they have control over the environment, they experience *agency*. Agency underpins the second profound affordance of VR. When learners are able to manipulate more objects in the world, with more than a gaze-based signal, we predict more agentic behaviors will emerge. When learners feel they control multiple parameters in the learning scenario, they own the experience and may also take more responsibility for learning. Learning is defined as the building of new knowledge structures. Many researchers hold that to build better knowledge structures one should be more agentic during the act of learning. The term agentic connotes that the user has volition over the individual objects in the environment; agency is considered a 'self-directed construct' per the Snow, Corno, and Jackson (1996) provisional taxonomy of conative constructs.

The newest generation of VR includes synced hand-held controls, these are a more Natural User Interface (NUI), compared to a keyboard. Using them makes it easier to incorporate gesture and to manipulate objects in VR. The second profound affordance of VR is driven by the **ability to gesturally interact with virtual content in 3D** and receive realtime feedback**.** Because of this affordance, NUIs are likely to have long-lasting effects on the types of content, and the quality of the pedagogy, that can be designed into educational spaces.

Evidence continues to accumulate that it is better for learners to be agentic and to kinesthetically engage with tasks rather than watching others engage. As an example in a real world study, two participants were randomly assigned to one of two roles in a learning dyad, either active or observant (Kontra, Lyons, Fischer, & Beilock, 2015). Participants who were active and physically held the spinning bicycle wheel learned more about angular momentum compared to those who observed the spinning wheel (Kontra et al., 2015). In a virtual setting, Jang, Vitale, Jyung, and Black (2016) studied with yoked pairs. One participant manipulated a virtualized 3D model of the inner ear, while the other participant viewed a recording of the 3D interaction. Results indicate that participants in the manipulation group showed greater posttest knowledge compared to the observation group. Results from the Embodied Games lab's previous mixed reality research (Johnson-Glenberg, Birchfield, et al., 2014; Johnson-Glenberg & Megowan-Romanowicz, 2017; Johnson-Glenberg, Savio-Ramos, & Henry, 2014) support the hypothesis that when learners perform actions with agency and can manipulate content during learning, they are able to learn and retain STEM knowledge better compared to learners exposed to low embodied, less agentic content.

1.1.4 Embodiment

Proponents of embodiment hold that the mind and the body are inextricably linked (Wilson, 2002). Varela et al. (1991) describes cognition as an "interconnected system of multiple levels of sensori-motor subnetworks" (p. 206). This author believes that activating these subnets can facilitate STEM learning. Embodied learning theory has much to offer designers of VR content working with NUIs. The strong stance on embodiment and education holds that the body should be moving, not just reading or imaging, for a high level of embodiment to be in a lesson (Johnson-Glenberg, 2017; Johnson-Glenberg & Megowan-Romanowicz, 2017). When a motoric modality is added to the learning signal, more neural pathways are activated and this may result in a stronger learning signal, or memory trace. Several researchers posit that incorporating gesture into the act of learning should strengthen memory traces (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007; Goldin-Meadow, 2011). It may be the case that adding more modalities to the act of learning (beyond the usual visual and auditory ones) will further increase the strength of the memory trace.

Throughout this chapter, the term gesture is used to mean both the movement as a communicative form, and the action used to manipulate virtual objects in the VR environment. [2] Research on non-mediated forms of gesture in the educational arena has also been fruitful. As an example, when teachers gesture during instruction, students retain and generalize more of what they have been taught (Goldin-Meadow, 2014). Congdon et al. (2017) showed that simultaneous presentation of speech and gesture in math instruction supported generalization and retention. Goldin-Meadow (2011) posits that gesturing may "lighten the burden on the verbal store" in a speaker's mind. Gesturing may serve to offload cognition (Cook & Goldin-Meadow, 2006). Gestures may aid learners because learners use their own bodies to create an enriched representation of a problem, which is then grounded in what have been called 'physical metaphors' (Alibali & Nathan, 2012; Hostetter & Alibali, 2008; Nathan et al., 2014). In addition, using gesture requires motor planning and this activates neural activations, and multiple simulations, even before the action is taken. Hostetter and Alibali (2008) posit that gesture first requires a mental simulation before movement commences, at that time motor and premotor areas of the brain are being activated in action-appropriate ways.

1.1.5 Congruent Gestures

[2] The 'gesture-enhancing-the-memory-trace' argument can also be framed as one of levels of processing, which is a well-studied concept in cognitive psychology (Craik & Lockhart, 1972), as is, learning by doing. Further supported by a large body of research on Self Performed Tasks (Engelkamp & Zimmer, 1994). In those studies, when participants *performed* short tasks, the task-associated words were better remembered compared to conditions where the participants read the words, or saw others perform the tasks.

The gesture should be congruent to the content being learned (Black, Segal, Vitale, & Fadjo, 2012; Segal, Black, & Tversky, 2010). That is, the gesture should map to the instructed concept. For example, if the student is learning about the direction and speed of a spinning gear, then it would be important for the student's spinning hand gesture to go in the same direction, and initiate the approximated speed of the virtual gear on screen (Johnson-Glenberg, Birchfield, Megowan-Romanowicz, & Snow, 2015). Gestures may provide an additional code for memory as well as adding additional retrieval cues. Learners with stronger memory traces should do better on post-intervention tests.

In a digital VR world, gesturing with a human-looking avatar hand may have special affordances that further increase the sense of agency. It is known that using one's hands to control screen action can attenuate simulator sickness (Stanney & Hash, 1998). Research further supports that users quickly begin to treat their avatars as their real bodies (Maister, Slater, Sanchez-Vives, & Tsakiris, 2015). With the advent of VR hand controls, where gestures can be fairly easily mapped, and more embodiment can be designed into lessons, it seems timely to revisit and clarify an earlier taxonomy on embodiment for education.

2. Taxonomy for Embodiment in Education

Appendix A goes into more depth on the three constructs in the Taxonomy for Embodiment in Education. The constructs are a) gesture congruency, 2) immersion/presence, and 3) the magnitude of the gesture (how much sensori-motor activation there is). Figure 1 is an updated graphic for embodiment in education that was originally proposed in Johnson-Glenberg, Birchfield, Megowan-Romanowicz and Savio-Ramos (2016). This new model takes into account the continuous nature of the three constructs. The crosshairs in the middle allow the reader the opportunity to partition the space into more tractable low and high construct areas (as opposed to degrees). Because magnitude of the gesture (i.e., the amount of motoric engagement) may prove to be the least predictive construct for content comprehension, it is relegated to the Z axis. (That depth axis is usually more difficult to conceptualize in a 2D graphic.)

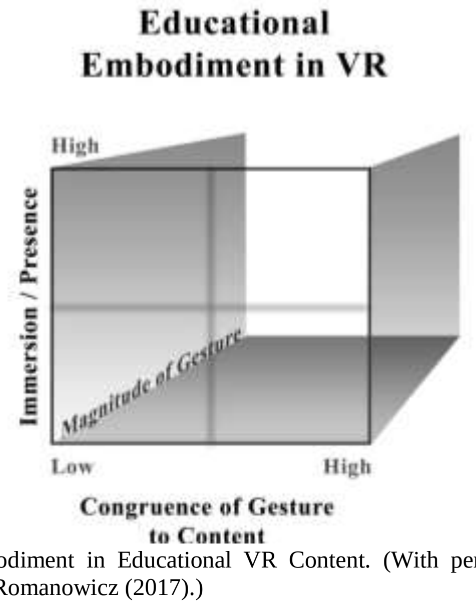

Figure 1. Cube of Embodiment in Educational VR Content. (With permission from Johnson-Glenberg and Megowan-Romanowicz (2017).)

The main take away is that a lesson can be deemed high on the embodiment scale if the gestures are congruent and well-mapped to the lesson's content, and if the lesson induces immersion/presence. In much of the past research on learning in VR, e.g., see Guiterrez et al. (2008), the focus has been on the technology and short shrift has been given to the pedagogies supporting the lessons. Designers and users of VR should be more aware of felicitous learning theories, so short descriptions of three relevant theories (used by this lab) are included in Appendix B. Users and purchasers of VR for education are also in need of rubrics to assess the quality of VR modules on the market and the author is working on one called QUIVRR (Quality of Virtual Reality in Education Rubric; Johnson-Glenberg, in preparation).

3.0 VR Design Guidelines Thus Far

For the most part, immersive VR education studies have occurred primarily in adult populations (Freina & Ott, 2015). Health and medicine appear to be leading the way with VR, from surgical training of craniofacial repairs (Mitchell, Cutting, & Sifakis, 2015) to behavioral change interventions related to PTSD (Rizzo et al., 2010). When the Standalone headsets (e.g., *Oculus GO, Quest*) which do not require phones or separate CPU's, become more affordable, then immersive VR experiences with hand control will hopefully become more popular for classroom use. At that time, educators will ask, "where is the quality content"?

What *will* high quality pedagogy in VR look like? Not everything in 2D needs to be converted to 3D. When designing for VR in education, Dalgarno and Lee presciently published five affordances for three dimensional VR environments (Dalgarno & Lee, 2010). Those five mesh nicely with Bailenson's below (2016). Bailenson posits that VR should be used in situations where it is most advantageous (Bailenson, 2016). Situations that are:

- **Impossible** – For example, you cannot change skin color easily, but in VR you can inhabit avatars with different skin colors with important and intriguing results (Banakou, Hanumanthu, & Slater, 2016; Hasler, Spanlang, & Slater, 2017). You cannot perceive a photon going directly into your eye in the classroom, but the next section describes a VR simulation doing just that.
- **Expensive** – You cannot easily fly your whole school to Machu Picchu.
- **Dangerous** – You would not want to want to train emergency landings by crashing real airplanes.
- **Counterproductive -** You should not cut down an entire forest to instruct on the problems of deforestation.

4. Designing for Embodied VR in Education

Adroitly meshing quality pedagogy with compelling gameplay is a far more arduous and heart-breaking endeavor than one would initially suspect. It is very difficult to create learning games that are both a) educational and b) sustainably entertaining. When these goals collide, the Embodied Games lab has opted to maintain high educational standards, and to let the entertainment aspects wane. This means that the player (student) needs to come to the task with an expectation of perseverance. It also means the starting point of serious game engagement is fundamentally different from the starting point of entertainment gameplay. Educational game designers *do* need to keep the game engaging, but we should rightfully be wary of adopting media design guidelines whole cloth from the entertainment world. Some of the end goals of entertainment are prolonged time and repeated visits. Paradoxically, an effective educational game that instructs well would not necessarily be re-visited multiple times by the same learner (unless the learner needed an occasional refresher). Additionally, learning games should never prompt for in-game purchases.

Porting learning content to the latest XR environment (the accepted term for MR, AR and VR) will add another layer of complexity to all learning games, until the conventions and user experience (UX) components become second nature. This author has created several epically flawed "edu-games" after 20 years of designing and developing. Designers must work against the biases they encountered in the world of 2D, and not bring them into the world of 3D design. One tenet repeated to those of us who have worked on VR entertainment games is that "immersion must never be broken". As a learning scientist, I am not certain that holds true for educational experiences. Inside a mediated learning module, space needs to be created for both 'experiential and reflective' learning (Antle, Corness, & Droumeva, 2009). The research community is still trying to figure out the parameters of this in VR, but it may involve breaking the illusion of immersion (e.g., by conversing with someone outside the headset, by writing content down on paper, etc.).

The original 18 design guidelines in the next section have been pulled together with pedagogy and optimal learning in mind. The final *Necessary Nine* at the end of the chapter have been further culled with financial constraints and development studio realities in mind. And, they can be taped to a wall on one page!

Multiple articles and books addressing principles of multimedia (Mayer, 2009) and how to design games in 2D exist, for examples see Squire (2008), Salen and Zimmerman (2004), and Schell (2014). The set below is one of the first for VR in education, especially with a focus on using hand controls for STEM learning. The focus is on making VR content for STEM that is engaging and embodied. To that end, these design guidelines will continue to be updated and refined as the technology, and its affordances, are updated and refined. An older version of these guidelines appears in Johnson-Glenberg (2018).

5. 1. Education in VR - 18 General Guidelines

➢ **Assume Every Learner is a VR Newbie - *Start slow***

- Not everyone will know the hand controls. Not everyone is a gamer. Not everyone will look around. Users are now in a sphere and sometimes need to be told to turn their heads. *However*, they should not turn too far, nor too quickly. Do not place important interface, HUD components, or actionable items, too far from each other.

- Part of starting slow includes being gentle with the user's proprioceptive system (where the body is in space). For example, if your user captures a butterfly at 10°, then do not force the next capture to be at 190°. Watch out for large body-action disconnects as well, e.g., the learner is standing, but the avatar is running, or lying in a bed. If the content includes varying levels of difficulty, allow the user to choose the level at the start menu. (This also gives a sense of agency.)

➢ **<u>Introduce User Interface (UI) Components Judiciously (fewer is better)</u>**

- Keep the as screen clean as possible. Permanent objects (i.e., a timer that stays center-screen as players turn their heads) will unnecessarily disrupt presence. Be creative about health bars (e.g., when the game *Snow Fortress* ported to VR, the designers got rid of pinned health bars, now the amount of snow accumulating on the users' mittens serves as health state feedback - cool). When users build the first fireworks in the chemistry lesson (next section), they can only make simple one stage rockets. The more complicated multi-stage components are not available in the interface until users show mastery of the simpler content. Designers should add visual complexity to the interface when the user is acclimated and ready (Johnson-Glenberg, Savio-Ramos, Perkins, et al., 2014).

➢ **<u>Scaffold – Introduce Cognitive Steps One at a Time</u>**

- Build up the user interface as you build up in cognitive complexity. This is a form of scaffolding (Pea, 2004). In the electric field series[3] of seven mini-games, users are not immediately exposed to the multi-variable proportionality of Coulomb's Law. Each component, or variable, in the Law is revealed one component at a time and reinforced via gameplay. Users explore, and eventually master each component successively before moving to the final lesson that incorporates all the previously learned content and culminates in the formation of lightning (Johnson-Glenberg & Megowan-Romanowicz, 2017).

➢ **<u>Co–design with Teachers</u>**

- Co-design means early and with on-going consultations. Let the teachers, Subject Matter Experts (SMEs), and/or clients play the game at mid- and end stages as well. Playtesting is a crucial part of the design process. Write down all comments made while in the game. Especially note where users seem perplexed, those are usually the breakpoints.

- Working with teachers will also ensure that your content is properly contextualized (Dalgarno & Lee, 2010), i.e., that it has relevance in, and is generalizable to the real world and to relevant educational content standards.

➢ **<u>Use Guided Exploration</u>**

- Some free exploration can be useful in the first few minutes for accommodation, and to incite curiosity, but once the structured part of the lesson begins, it is your job to *guide* the learner. Guide using constructs like pacing, signposting, blinking objects, constrained choices, etc. To understand why free exploration as an instructional construct has not held up well in STEM education, see Kirschner, Sweller, & Clark (2006).

➢ **<u>Minimize Text Reading</u>**

- Rely on informative graphics or mini-animations whenever possible. Prolonged text decoding in VR headsets causes a special sort of strain on the eyes, perhaps due to lens muscle fatigue or the vergence-accomodation conflict. In the *Catch a Mimic* game (described next section), players do not read lengthy paragraphs on butterfly cocoons and emerging, instead a short 2-second cut-scene animation of butterflies in chrysalis and then emerging is displayed.

➢ **<u>Build for Low Stakes Errors Early On</u>**

- Learning often requires errors to be made. Learning is also facilitated by some amount of cognitive effortfulness. In the *Catch a Mimic* game, the player must deduce which butterflies are poisonous, just like a natural predator must. In the

[3] Also www.embodied-games.com or https://www.youtube.com/watch?v=eap7vQbMbWQ

first level, the initial butterflies that appear on screen are poisonous. Eating them is erroneous and slightly depletes the player's health score, but there is no other way to discern what is toxic, some false alarms must be made. In psychology, this is called 'learning from errors' (Metcalfe, 2017); in the learning sciences, it has been called productive failure (Kapur, 2016).

➢ **Playtest Often with both Novices and End-users**

• It is crucial that designers playtest with multiple waves of age-appropriate learners for feedback. This is different from co-designing with teachers.

• Note, playtesting with developers does *not* count. Human brains learn to reinterpret visual anomalies that previously induced discomfort, and over time users movements become more stable and efficient (Oculus, 2018). Developers spend many hours in VR and they physiologically respond differently than your end-users will.

➢ **Feedback - Unobtrusive, Actionable and Well-timed**

• This does not mean giving constant, on screen, feedback (Shute, 2008). Feedback should be high level, and if text is included, it should be evaluative but short. Some proportion of users will be colorblind, so you cannot rely on only red and green colors for feedback. Feedback should be paced because it takes time for the cognitive adjustments to be integrated into the learner's ongoing mental model. This leads to the next guideline on reflection.

➢ **Design in Opportunities for Reflection** (it should not be all action and twitch! Include metacognition.)

• Education game designers are currently experimenting with how to do this in VR. Reflection allows the learner's mental model to cohere. Some ongoing questions include: Should the user stay in the headset or not? How taboo is it to break immersion? Should short quizzes be embedded to induce a retest effect (Karpicke & Roediger, 2008)? Perhaps screencasting/mirroring with dyads where one partner is outside the headset could be conducive to learning and the interleaving of new knowledge.

➢ **Encourage Collaborative Interactions**

• Synced, multiplayer experiences are still expensive, but their creation is a worthy goal. Until the cost drops, designers should explore workarounds to make the experience more social and collaborative. Some ideas include: using a preprogrammed non-player character (NPC), having a not-in-headset partner interact via a screencast, or building sequential tasks that require back-and-forth asynchronous activities.

5.2 Using Hand Controls/Gestures

The final design guidelines (numbers 12 through 18) focus on using the hand controllers in VR for learning.

➢ **Use the Hand Controls to Encourage the Users to be "Active"**

• Incorporate into lessons opportunities for learners to make physical, kinesthetic actions that manipulate content. Where appropriate, try to include representational gestures and/or re-enactments.

• In this lab's previous research, the group that was instructed in centripetal force and made kinesthetic circles (either with the wrist or arm) retained more physics knowledge, compared to the group that made low embodied, less active motions (Johnson-Glenberg, Birchfield, et al., 2014). Active learning has been shown to increase STEM grades by up to 20% (Waldrop, 2015).

➢ **How Can a Body-based Metaphor be Applied?**

- Be creative about ways to incorporate kinesthetics, or body actions, into the lesson. At first blush, it may not be apparent how to make a traditional bar chart become more embodied. But with a VR hand control, the learner can now use a gesture to fill a bar to the correct height. An upward swipe is also congruent with our cultural concept of more (see Dor Abrahamson's work on embodied and mediated examples of proportional reasoning). In the "Catch a Mimic" game, learners are instructed to make a prediction about species survivability using the hand controls (see Figures 5 and 6, next section). Additionally, prediction is a well-researched metacognitive comprehension strategy (Palincsar & Brown, 1984).

➢ **Congruency**

- The gesture/action should be congruent, i.e., it should be well-mapped, to the content being learned (Black et al., 2012; Johnson-Glenberg & Megowan-Romanowicz, 2017). The action to start a gear train spinning should involve moving the hand or arm in a circle with a certain velocity; the action should not be pushing a virtual button labeled "spin" (Johnson-Glenberg et al., 2015).

➢ **Actions Strengthen Motor Circuits and Memory Traces**

- Performing actions stimulates the motor system and appears to also strengthen memory traces associated with newly learned concepts (Refer to the Appendix B on embodiment, or Johnson-Glenberg & Megowan-Romanowicz, 2017).

➢ **Ownership and Agency**

- Gestural control gives learners more ownership of, and agency over, the lesson. Agency has positive emotional affects associated with learning. With the use of VR hand controls, the ability to manipulate content and interactively navigate appears to also attenuate effects of cybersickness (Stanney & Hash, 1998).

➢ **Gesture as Assessment – Both Formative and Summative**

- Design in gestures that reveal the state of the learner's mental model, both *during learning* (called formative or in-process) and *after the act of learning* (called summative).

- For example, you might prompt the learner to demonstrate negative acceleration with the swipe of a hand controller. Does the hand controller speed up or slow down over time? Can the learner match certain target rates? This is an embodied method to assess comprehension that includes the added benefit of reducing guess rates associated with the traditional text-based multiple choice format. For an example of hand movements showing vector knowledge on a tablet, see the *Ges-Test* in Johnson-Glenberg and Megowan-Romanowicz (2017).

➢ ***Aspirational* - Personalized, more Adaptive Learning**

- Finally, try to include adaptivity. This is acknowledged to cost more, but the learning content level should often reside a fraction beyond the user's comprehension state, also known as the learner's Zone of Proximal Development (ZPD) (Vygotsky, 1978).

• Gesture research on younger children shows they sometimes gesture knowledge before they can verbally state it. Gesture-speech mismatches can reveal a type of readiness to learn (Goldin-Meadow, 1997). Thus, gestures can also be used as inputs in adaptive learning algorithms. Adding adaptivity (dynamic branching) based on performance is difficult and time-consuming to design, but it is considered one of the best practices in educational technology (Kalyuga, 2009); it is something to strive for.

These 18 guidelines are condensed into the most important nine in the final section of this chapter.

6. Case Studies: Two VR Examples

This section describes two relevant case studies. The first example showcases some of the design changes that occurred as 2D content was repurposed to a 3D VR lesson that is now available for free in the *Oculus* store. The second example highlights the design techniques of both constructivism and guided exploration in VR.

6.1 Example 1. The Natural Selection Game: Reconceptualizing 2D content into a 3D VR lesson

This project began as a 2D assessment tool to measure knowledge gained after watching a giant screen movie, *Amazon Adventure*.[4] One of the key science topics in the movie was Batesian mimicry. The tablet-based test was designed *not* to *instruct* in the topic, but rather to assess whether players became more adroit at picking out non-poisonous butterflies over time, as the levels increased in difficulty. Design was constrained because we could not include explicit text that described how mimicry occurred. Now that the assessment tool phase is over, the game has been redesigned to be instructional, and text has been added. "Catch a Mimic" is a standalone lesson for middle and high school students now available in the *Oculus* Store for both *Go* and *Rift*[5] platforms.

6.1.1 Tablet version of the Natural Selection Game

The earlier 2D tablet assessment was given at multiple time points associated with movie viewing, i.e., pre-, post-viewing and gain after a two week delay. The butterflies would spawn from the right side of the screen and the background of a forest would scroll to the left. The instructions read, "You are a bird trying to eat as many non-poisonous butterflies as possible". A finger-tap on a butterfly would make it disappear. Immediate feedback was given (visually, not auditorily as it was a test taken by entire classrooms). Persistent feedback was displayed on top of the screen as to whether the selection was poisonous or non-poisonous. As the levels progressed, the non-poisonous butterflies' wing patterns altered to more closely resemble the poisonous butterflies', that is, mimicry occurred. On the *Kindle FIRE 8* tablet, the actionable play space was only 7.0 inches (diagonal) and no distractors were included, i.e., falling leaves, particles in foreground, moving water, etc. To turn the tablet assessment tool into an engaging, instructional game on the topic of natural selection, it was first ported onto a PC – the most common form factor in schools.

[4] The movie, *Amazon Adventure,* was released in 2017. The funding agency for the assessment tool was the National Science Foundation, grant # 1423655. A WebGL version of the game called "Catch a Mimic" can be played at www.embodied-games.com.
[5] *Rift* version https://www.oculus.com/experiences/rift/2656510471032810

6.1.2 PC Version of the Natural Selection Game

In the current version of the natural selection game (v6.0.1 found at https://www.embodied-games.com/games/natural-selection-catch-a-mimic/), the mouse controls the location of a virtual net. A mouse click captures a butterfly, see Figure 2. The new opening narrative changed and states, "You are a zookeeper capturing butterflies to feed to your birds". The scroll to the left mechanic did not feel appropriate for the larger, computer monitor (average diagonal 16 inches), so now the butterflies spawn and fly out of a central bush. Because the game would eventually move to VR, the team decided that flying and swooping as a bird would make the player nauseous, and so the bird POV was abandoned.

This was no longer merely a test so visual elements to enhance engagement were included, e.g., moving waterfalls, particle effects, and audio. Chirping birdsong was added to increase presence. Feedback was handled differently on the PC. In an effort to declutter the UI, the permanent feedback at the top of the screen was removed. Now on the bottom right of the screen (pinned to the world, BUT not to the HUD) the performance feedback was displayed. In addition, audio feedback, as positive or negative sounds upon collision with a butterfly, was added. A green heart or red skull showed up upon collision on the central screen for 1.5 seconds as a form of feedback as well. On the bottom right, persistent (but unpinned) numerical feedback on type of butterfly captured was also displayed. The timer restarted at 60 seconds for each of the six levels.

Figure 2. PC version. Butterflies spawn from a central bush and fly towards the player. The screen no longer scrolls, but the moving waterfall keeps the background from feeling too static.

6.1.3 VR Version of the Natural Selection Game

To move the 2D PC version to 3D VR all moving assets (butterflies, net, etc.) were rendered into 3D. The VR version has graphics (rainforest background) that span 360°. However, all the action is constrained to occur in the 'central play arc' of approximately 170°. Figure 3 shows a portion of the VR playspace. Feedback is now located closer to the spawning bush and is pinned to the world (not HUD) – this means if you turn around 180°, you will only see the forest and stream. The waterfall now continues as a stream that encircles you as the player. Sound is omni-directional. Although players can turn all the way around and see trees, earth, and sky, no butterflies or clickable action content appear "behind" the players because we do not want them to spin around, get dizzy, or become tangled in wires. At the bottom of Figure 3, note the ghostlike avatar hand that is mapped to the human player's hand and wrist movements. In this *Oculus* version, the hand grips around the net handle. The net is fully articulated in three axes.

Figure 3. The VR version with an articulated avatar hand and wrist.

6.2 Creative and embodied assessments

For both the PC and VR versions an interactive assessment was created. Population dynamics can be a difficult concept to teach; we believe that its instruction need not "necessarily be quantitative" as some do (Schaub & Abadi, 2011). Middle and High School students can make inferences and predictions about joint likelihoods without memorizing statistical formulae. The assessment includes prediction which is one of a set of powerful and well-researched comprehension strategies (Rosenshine & Meister, 1994). It is part of metacognition, encouraging learners to think about their thinking. The goal was to include an embodied prediction in the VR environment. It is straightforward to track the hand controllers in 3D. This opens up multiple opportunities to include a large spectrum of gestures and re-enactments for the purposes of assessment (and instruction). A predictive question was created that would adhere to many of the design principles stated in this chapter, including low stake errors with feedback, being active, and using congruent movements (*swipe upwards to connote an increase*). The question's answer should provide a snapshot of the learner's comprehension state, while also encouraging the learner to think deeply about outcomes (i.e., being metacognitive and reflective).

Figure 4 shows the interactive bar chart prompt, after a learner has submitted an answer. The learner must make a best guess as to the survivability of the next four species, i.e., bar #1 is prefilled – learners must drag the blue oval up to show survivability of the poisonous butterfly (#2) and the three non-poisonous butterflies to the right (#3, #4 and #5). When learners are satisfied with their decisions, they click on the submit button. Learners are allowed three incorrect submissions before an animation shows the correct answer.

Figure 4. Interactive assessment in both 2D and 3D VR versions, with two incorrect choices.

6.3 Example 2. A High Embodied VR Lesson with hand controls. Topic: Chemistry/Physics

The second example comes from a multiplayer experience. This module was designed to highlight constructivism and scaffolding, it was included in a multi-player entertainment game called *Hypatia*, available on the Steam store (although later versions may vary). *Hypatia* is an open world primarily built for social entertainment. For the Alpha version of a high school-level chemistry lesson, the author served as a consultant to ensure best pedagogies were used in the module. The highlighted module is called *Kapow Lake*; it was conceived of as a high school lesson using fireworks to instruct in physics and chemistry. Two learning goals were embedded: 1) understand which metal salts burst into which colors, and 2) understand the preliminary physics behind why the burst is perceived as a particular color. Players start on the beginner side of the lake, they can watch fireworks in the sky and are motivated to build some of their own.

One can scaffold cognitive elements, as well as interface elements. As a form of UI scaffolding, light cues, were used to “signpost” players to a certain building. In a sphere, it can be difficult to know where to travel next. With free exploration, precious classroom time could be wasted with students trying out dead-end options. Via the lit doorway, we encourage players to enter the expert’s shed to learn more. See Figure 5.

Figure 5. The invitingly lit expert's shed from *Hypatia*. Enter and learn enough to get to the next level. *Created by www.timefirevr.com.*

In order to construct their own fireworks, players must first master the names of the salt colors. The salts are grey, and names are not readily deducible from their exteriors. Players would grasp the triggers of the hand controls and when their avatar hands collided with a metal salt, the salt could be picked up. The first series of grey metal salts (see Figure 6) did not have the colors on the labels. Thus, players did not know that the salt called strontium would burn red. Via guided exploration, they would place each salt into the flame of the Bunsen burner and note the color that the salt burned.

Figure 6. The strontium Bohr model. Note the red wave. *Created by www.timefirevr.com*

Figure 6 shows the avatar on the left side of the screenshot. The salt labels are now colored and visible (i.e., if strontium burns red, how will copper burn?). After the player places the grey salt over the flame a Bohr atom model of strontium appears on top of the flame.

Recall that the first profound affordance of VR is the immediate presence. Note that the screenshot is taken from the 3rd person POV for the purposes of edification, but the human player, is seeing the atom floating towards her/him in 1st person or a "head on" POV. This is very engaging, but it could be alarming if the object moved too quickly towards one's eye. Multiple sessions were spent playtesting the optimal velocity for this interaction.

After the learner places the strontium over the heat, the outer electron jumps from the stable outer orbit. The unstable orbit is shown briefly as a dotted ring during play (*not* shown in Figure 6). Quickly, the electron falls back to its more stable orbit. As it does so, it releases a packet of energy called a photon. This photon is perceived in the red spectrum. In Figure 6, the photon has been visualized as *both* a red wave and a particle heading towards the eye[6]. The learner perceives the photon as traveling directly towards the eye. (This is perhaps the only thing humans want heading directly towards our eyes!)

The simulation of the photon as a wave reifies the concepts that energy is released by the heat burst, and that the energy is then perceived by the human eye as a visible wavelength. The five other salts release electrons from different orbits, thus creating different wavelengths. Once players are able to match all six metal salts to their colors, the players are 'guided' to exit through the back door to the multi-staging firework building area.

This is where the social and collaborative aspects comes into play, because other experts can be out by the lake building multistage rockets and can give feedback and clap when the final firework version is correct. Building a multistage rocket is complex and so the building was scaffolded. The player is first asked to build a one color firework. Then players are asked to make their rockets burst in a predetermined sequence of multiple colors. The building of the firework rocket is a sequential production. Using the hand controllers, a player must construct in a certain order: tube first, then fins, salts, fuse, then the cone top. This is a constructive, engaging task, but it also serves as a form of stealth assessment (Shute, 2011). Now a teacher, or spectator, can observe whether the student really understands how strontium and copper need to be sequenced to make a red *then* a blue explosion. When a rocket explodes correctly, there are often group shouts of approval - if others are in the space.

7. The Necessary Nine

As the technology moves forward, designers should keep principles of best practices in mind, and instructors should consult the principles when making purchasing decisions. The term "best" is relative. It depends on several constraints including the affordances of the technology (which are constantly changing). This chapter ends with the current top contenders. All designers strive for engagement, so it is not explicitly mentioned as a guideline. If there are only resources to focus on a subset of the main guidelines, then the author recommends the *Necessary Nine.*

[6] In a small usability study, several players reported this model helped them to understand color perception. Whether the task inadvertently supports an incorrect model of "red waves moving through the air" could be explored with a larger and more formalized study. These sorts of issues are always a tension when visualizing abstract phenomena.

THE NECESSARY NINE

1- Scaffold cognitive effort (and components in interface) - one step at a time
2- Use guided exploration
3- Give immediate, actionable feedback
4- Playtest often, *with correct user group*
5- Build in opportunities for reflection
6- Use the hand controls for active, body-based learning
7- Integrate gestures that map to the content to be learned
8- Gestures are worth the time and extra expense - they promote learning, agency, and attenuate cybersickness (be creative about using motion and gesture for assessment)
9- Embed assessment, both during and after the lesson

8. Conclusion

It is an exciting time for education and VR, filled with opportunity and enlivened by a rapidly changing hardware landscape. Besides issues around *how* to design optimal lessons, there are important questions regarding *when* to insert a VR module. Aukstakalnis (2017) shares an anecdote about a student in a design class who regretted designing his first project in a VR headset during the year-long course because he missed watching classmates work in the real world and being able to learn from "his peers' collective mistakes" (p. 306). Like most academic musings, this current chapter ends with a request for more research on learning in VR.

The design guidelines presented here will be refined as the hardware and its affordances change. This chapter focused on the two profound affordances associated with the latest generation of VR for educational purposes: 1) *presence*, and 2) the *embodied affordances of gesture in a three dimensional learning space*. VR headsets with hand controls allow for creative, kinesthetic manipulation of content, those types of movements and gestures have been shown to have positive effects on learning. Hand controllers can be used for innovative types of assessment. Hopefully, the case studies and design guidelines here will help others to create effective immersive VR lessons.

References

Abrahamson, D. (2009). Embodied design: Constructing means for constructing meaning. *Educational Studies of Mathematics, 70*, 27-47. doi:10.1007/s10649-008-9137-1

Abrahamson, D., & Trninic, D. (2015). Bringing forth mathematical concepts: signifying sensorimotor enactment in fields of promoted action. *ZDM Mathematics Education*, *47*(295). doi:https://doi.org/10.1007/s11858-014-0620-0

Alaraj, A., Lemole, M. G., Finkle, J. H., Yudkowsky, R., Wallace, A., Luciano, C., . . . Charbel, F. T. (2011). Virtual reality training in neurosurgery: Review of current status and future applications. *Surgical Neurology International*, *2*, 52. doi:10.4103/2152-7806.80117

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences*, *21*(2), 247-286.

Antle, A. A., Corness, G., & Droumeva, M. (2009). What the body knows: Exploring the benefits of embodied metaphors in hybrid physical digital environments. *Journal of Interactive Computing*, *21*(1-2), 66-75.

Aukstakalnis, S. (2017). *Practical Augmented Reality: A guide to the Technologies, Applications, and Human Factors for AR and VR*. Boston: Addison-Wesley.

Bailenson, J. N. (Producer). (2016, January 20, 2018). The Trials and Tribulations of Narrative in VR. *MediaX*. Retrieved from https://vhil.stanford.edu/news/2016/the-trials-and-tribulations-of-narrative-in-vr-mediax/

Bailenson, J. N. (2018). *Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do*. New York, New York: W.W. Norton & Co.

Bailenson, J. N., Patel, K., Nielsen, A., Bajscy, R., Jung, S., & Kurillo, G. (2008). The effect of interactivity on learning physical actions in virtual reality. *Media Psychology, 11*(3), 354-376.

Banakou, D., Hanumanthu, P. D., & Slater, M. (2016). Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias. *Frontiers in Human Neuroscience, 10*(601). doi:http://doi.org/10.3389/fnhum.2016.00601

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22*, 577-660.

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

Bjork, R. A. (1994). *Memory and metamemory considerations in the training of human beings* Cambridge, MA: MIT Press.

Bjork, R. A., & Linn, M. C. (2006). The science of learning and the learning of science: Introducing desirable difficulties. *Association of Psychological Science Observer, 19*(29), 39.

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. L. (2012). Embodied cognition and enhancing learning and motivation. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments.* NY: Routledge.

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. . *Journal of Experimental Psychology: General, 136*(4), 539.

Congdon, E. L., Novack, M. A., Brooks, N., Hemani-Lopez, N., O'Keefe, L., & Goldin-Meadow, S. (2017). Better together: Simultaneous presentation of speech and gesture in math instruction supports generalization and retention. *Learning and instruction, 50*, 65-74.

Cook, S., & Goldin-Meadow, S. (2006). The Role of Gesture in Learning: Do Children Use Their Hands to Change Their Minds. *Journal of Cognition and Development*, *7*(2), 211-232.

Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11*, 671-684.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology, 41*(1), 10-32. doi:10.1111/j.1467-8535.2009.01038.x

Dede, C., & Richards, J. (2017). Glossary of realities' terms. In D. Liu, C. Dede, & J. Richards (Eds.), *Virtual, Augmented, and Mixed Realities in Education* (pp. 5-11). Berlin: Springer Verlag.

Dewey, J. (1966). *Democracy and education: An introduction to the philosophy of education*. New York: The Free Press.

Dunleavy, M. (2014). Design principles for augmented reality learning. *TechTrends*, *58*(1), 28-34.

Engelkamp, J., & Zimmer, H. D. (1994). Motor similarity in subject-perfromed tasks. *Psychological Research-Psychologische Forschung, 57*(1), 47-53.

Freina, L., & Ott, M. (2015). *A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives*.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 586-596. doi:10.1002/wcs.55

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the revolution to embodiment: 25 years of cognitive psychology. *Perspectives on Psychological Science*, *8*(5), 573 –585. doi:10.1177/1745691613498098

Goldin-Meadow, S. (1997). When gestures and words speak differently. *Current Directions in Psychological Science, 6*(5), 138-143. doi:10.1111/1467-8721.ep10772905

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

Guiterrez, F., Pierce, J., Vergara, V., Coulter, R., Saland, L., Caudell, T. P., . . . Alverson, D. C. (2008). The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. *Medicine Meets Virtual Reality 15: In Vivo, in Vitro, in Silico: Designing the Next in Medicine., 15*, 13.

Hasler, B. S., Spanlang, B., & Slater, M. (2017). Virtual race transformation reverses racial in-group bias. *PLoS ONE*, *12*(4), 4.

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

Jang, S., Vitale, J., Jyung, R., & Black, J. (2016). Direct manipulation is better than passive viewing for learning anatomy in a three-

dimensional virtual reality environment. *Computers & Education, 106*(March), 150-165. doi:10.1016/j.compedu.2016.12.009
Johnson-Glenberg, M. C. (2017). Embodied education in mixed and mediated realities: Principles for content design. In D. Liu, C. Dede, & J. Richards (Eds.), *Virtual, Augmented, and Mixed Realities in Education* (pp. 193-218). Berlin: Springer Verlag.
Johnson-Glenberg, M. C. (2018). Immersive VR and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI*, *5*(81). doi:10.3389/frobt.2018.00081
Johnson-Glenberg, M. C., Birchfield, D., Koziupa, T., & Tolentino, L. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, *106*(1), 86-104. doi:10.1037/a0034008
Johnson-Glenberg, M. C., Birchfield, D., Megowan-Romanowicz, M. C., & Snow, E. L. (2015). If the gear fits, spin it! Embodied education and in-game assessments. *International Journal of Gaming and Computer-based Simulations, 7*(7), 40-65. doi:DOI: 10.4018/IJGCMS.2015100103
Johnson-Glenberg, M. C., & Megowan-Romanowicz, M. C. (2017). Embodied science and mixed reality: How gesture and motion capture affect physics education. *Cognitive Research: Principles and Implications, 2*(24). doi:10.1186/s41235-017-0060-9
Johnson-Glenberg, M. C., Megowan-Romanowicz, M. C., Birchfield, D., & Savio-Ramos, C. (2016). Effects of embodied learning and digital platform on the retention of physics content: Centripetal force. *Frontiers in Psychology, 7:1819*. doi:doi: 10.3389/fpsyg.2016.01819
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
Johnson-Glenberg, M. C., Savio-Ramos, C., Perkins, K. K., Moore, E. B., Lindgren, R., Clark, D., . . . Squire, K. (2014, June, 2014). *Science Sims and Games: Best Design Practices and Fave Flops.* . Paper presented at the The International Conference of the Learning Sciences (ICLS), Boulder, CO.
Kalyuga, S. (2009). Managing cognitive load in adaptive ICT-based learning *Journal of Systemics, Cybernetics and Informatics, 7*, 16-21.
Kapur, M. (2016). Examining Productive Failure, Productive Success, Unproductive Failure, and Unproductive Success in Learning. *Educational Psychologist, 51*(2), 289-299. doi:10.1080/00461520.2016.1155457
Karpicke, J. D., & Roediger, H. L. (2008). The Critical Importance of Retrieval for Learning. *Science, 319*(5865), 966-968. doi:10.1126/science.1152408
Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist, 41*(2), 75-86. doi:10.1207/s15326985ep4102_1
Kontra, C., Lyons, D., Fischer, S., & Beilock, S. L. (2015). Physical experience enhances science learning. *Psychological Science, 26*, 737-749. doi:10:1177: 0956797615569355.
Laviola, J. J., Kruijff, E., McMahan, R. P., Bowman, D. A., & Poupyrev, I. (2017). *3D User interfaces: Theory and Practice* (2nd ed.). Boston: Addison-Wesley.

Lindgren, R., & Johnson-Glenberg, M. C. (2013). Emboldened by embodiment: Six precepts regarding the future of embodied learning and mixed reality technologies. *Educational Researcher 42*(8), 445-452. doi:10.3102/0013189X13511661

Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2015). Changing bodies changes minds: Owning another body affects social cognition. *Trends in Cognitive Sciences, 19*(1), 6-12. doi:10.1016/j.tics.2014.11.001

Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the multimodal presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior, 72*(C), 276-285. doi:10.1016/j.chb.2017.02.066

Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). New York: Cambridge University Press.

Megowan, M. C. (2007). *Framing Discourse for Optimal Learning in Science and Mathematics.* (PhD), Arizona State University, Tempe, AZ.

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology, 68*(1), 465-489. doi:10.1146/annurev-psych-010416-044022

Mikropoulos, T. A., & Natsis, A. A. (2011). Educational virtual environments: A ten-year review of empirical research (1999-2009). *Computers & Education, 56*(3), 769-780. doi:10.1016/j.compedu.2010.10.020

Mitchell, N., Cutting, C., & Sifakis, E. (2015). GRIDiron: an interactive authoring and cognitive training foundation for reconstructive plastic surgery procedures. *ACM Trans. Graph., 34*(4), 1-12. doi:10.1145/2766918

Nathan, M. J., Walkington, C., Boncoddo, R., Pier, E. L., Williams, C. C., & Alibali, M. W. (2014). Actions speak louder with words: The roles of action and pedagogical language for grounding mathematical reasoning. *Learning and instruction, 33*, 182-193. doi:DOI: 10.1016

Oculus. (2018). *Developers: Best Practices*. Retrieved from https://developer.oculus.com/design/latest/concepts/book-bp/

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117-175. doi:10.1207/s1532690xci0102_1

Pea, R. D. (2004). The Social and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. *Journal of the Learning Sciences, 13*(3), 423-451. doi:10.1207/s15327809jls1303_6

Piaget, J. (1997). *The Language and Thought of the Child* (3rd ed.). London: Routledge.

Rizzo, A., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., & Mclay, R. (2010). Early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. *ANNALS OF THE NEW YORK ACADEMY OF SCIENCES, 1208* (Psychiatric and Neurologic Aspects of War Development), 114-125.

Rosenshine, B., & Meister, C. (1994). Reciprocal Teaching: A review of the research. *Review of Educational Research, 64*(4), 479-530. doi:10.3102/00346543064004479

Salen, K., & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Cambridge, MA: MIT Press.

Schaub, M., & Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology, 152*(1), 227-237. doi:10.1007/s10336-010-0632-7

Schell, J. (2014). *The Art of Game Design* (2nd ed. Vol. New York): Taylor Francis.

Segal, A., Black, J., & Tversky, B. (2010). *Do Gestural Interfaces Promotoe Learning? Congruent Gestures Promote Performance in Math*. Paper presented at the 51st Meeting of the Psychonomic Society Conference, St. Louis, Missouri.

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189. doi:10.3102/0034654307313795

Shute, V. (2011). Stealth Assessment in computer-based games to support leanring. *Computer Games and Instruction, 20*, 503-523.

Skulmowski, A., & Rey, G. D. (2018). Embodied learning: Introducing a taxonomy based on bodily engagement and task integration. *Cognitive Research, 3*(1), 6. doi:10.1186/s41235-018-0092-9

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with Immersive virtual reality. *Frontiers in Robotics and AI, 3*(74). doi:10.3389/frobt.2016.00074

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments *Presence: Teleoperators and virtual environments, 6*, 603–616.

Snow, R. E., Corno, L., & Jackson, D. N. (Eds.). (1996). *Individual differences in affective and conative functions*. New York: Macmillan.

Squire, K. D. (2008). Video Games and Education: Designing Learning Systems for an Interactive Age. *Educational Technology, 48*(2), 17-26.

Stanney, K. M., & Hash, P. (1998). Locus of user-Initiated control in virtual environments: Influences on cybersickness. *Presence: Teleoperators and virtual environments, 7*(5), 447-459. doi:10.1162/105474698565848

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: The Harvard University Press.

Waldrop, M. M. (2015). Why we are teaching science wrong, and how to make it right. *Nature News, 523*(7560), 272-275. doi:doi:10.1038/523272a

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review, 9*(4), 625-636.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments, 7*(3), 225-240.

Woolfolk, A. (2007). *Educational Psychology* (10th ed.). Boston: Pearson.

Appendix A

Taxonomy of Embodiment for Education in VR

As with all theories, there are inclusive (weak) ones that start the spectrum, and exclusive (strong) ones that end it. One inclusive theoretical stance on embodied learning would be that any concept that activates perceptual symbols (Barsalou, 1999) is, by its nature, embodied. Following this stance, all cognition is embodied because our earliest knowledge is gathered via the body and its interactions with the environment, even new concepts that are later imagined. The environment's affordances (Gibson, 1979) shape and constrain how our bodies interact, ergo, cognition continues to be formed and expanded by these interactions. In an inclusive interpretation, according to some researchers, cognition would be broadly defined to include all sensory systems and emotions (Glenberg, 2010; Glenberg, Witt, & Metcalfe, 2013). A more exclusionary stance is one that distinguishes between low and high levels of embodiment. For a lesson to be deemed highly embodied, the learner would need to be *physically active;* the learner would have to kinesthetically activate motor neurons. Some principles for designing embodied education into MR platforms have been suggested (Lindgren & Johnson-Glenberg, 2013), and several AR design principles have been proposed (Dunleavy, 2014); however, there are currently no design guidelines for VR that are based on embodiment. Given the new affordances of VR hand controls, it seems timely to reframe some of this lab's previous embodied principles.

A more exclusionary definition of embodiment for education was proposed by this lab in 2014 (Johnson-Glenberg, Birchfield, et al., 2014) and updated recently (Johnson-Glenberg & Megowan-Romanowicz, 2017). That taxonomy posited four degrees of embodiment based on three constructs: a) amount of sensori-motor engagement, b) how congruent the gestures were to the content to be learned, and c) amount of 'immersion' experienced by the user. Each construct will be expanded upon below. Finally, a new cube of embodiment is proposed (See Figure 1).

2.1 Sensori-motor Engagement

In terms of sensori-motor engagement via gesture (construct a), the first distinction relates to the magnitude of the motor signal. This means that a larger movement, e.g., a gross arm movement would activate more sensori-motor neurons a smaller one like swiping a finger across a small screen. The magnitude of the movement should probably be part of the metric, but it is perhaps less important than whether the gesture is well-matched (congruent) to the content to be learned (construct b). A small, yet highly congruent movement may be just as effective as a large one that is only loosely related to the learning concept. That is an experiment that needs to be conducted.

2.2 Congruency of the Gesture

Construct b refers to the congruency of the gesture, that is, the movement should be mapped to, related to, the concept to be learned. The gesture should support the gist of the content and give meaningful practice to the learning goal; however, the movement need not be a perfect isomorphic match. In the spinning gears example, a mediated lesson was created to instruct in mechanical advantage for gear systems (Johnson-Glenberg et al., 2015). The *Microsoft Kinect* sensor was used to capture the direction and speed of the spin of the learner's arm. The learner extended his/her arm in front of the body and rotated it around the shoulder joint. That movement drove the first gear in a simulated gear train. Using distance from shoulder joint to wrist joint, the average diameter of the driving gear was mapped to the learner's body; when the learner altered the size of the physical spins, that action altered the size of the gear on screen in real time. Using the learner's real time wrist speed, the velocity of the gear spin was also mapped in real time. **Congruency means a large overlap between the action performed and content to be learned.** In the above study, the learners who understood mechanical advantage (on a content knowledge test) also showed greater competency during gameplay. The better testers also consistently chose the correct diameter gear during the virtual bike race during play. This is an example of how gesture can be part of both the learning situation and assessment wrapped in virtual gameplay.

2.3 Immersion/Presence

Construct c has been called *sense of immersion* in previous articles describing the Johnson-Glenberg embodiment taxonomy for education (Johnson-Glenberg, Birchfield, et al., 2014; Johnson-Glenberg & Megowan-Romanowicz, 2017). However, Mel Slater's lab posits that immersion is a non-subjective property of the technological system and should not be considered a *sensation*. Immersion is composed up of various system attributes, e.g., Field of View (FOV), fidelity to environment, etc. Slater and Wilbur (1997) distinguish between presence and immersion, positing that presence is what is subjectively felt by the user. Slater and Sanchez-Vives (2016) concede the two terms are "subjective correlates". This author is guilty of often conflating the two terms. Slater and others (Witmer & Singer, 1998) assert that the two terms should be kept separate because presence is always a subjective experience. But, we agree believe the two terms are inextricably "tangled" (Alaraj et al., 2011), and given the high fidelity and immersive affordances of the current spate of immersive VR technologies, it may be appropriate to assume the majority of users will be in high fidelity and highly immersive VR environments (our lab focuses on high-end, non-mobile phone headsets). As the amount of immersivity in the technology begins to asymptote, perhaps we can conflate the two terms into the one called *presence* when assessing psychological/educational experiences? The levels of quality for optics, lag, and audition are impressive; we believe they are sufficient for the majority of users to suspend disbelief and feel deeply translocated.[7]

Thus, the author proposes using the one term *presence* to also connote a very high degree of immersion as well, because the amount of immersion is universally high in the current generation of immersive 3D VR. For VR, this chapter continues with a fusion term of *immersion/presence* to bridge to the future. Under the construct of immersion/presence, there are subsumed other factors or corollaries that are critical to learning, e.g., motivation and prior knowledge, which are clearly important. Although, many of these factors are not under the control of lesson designers. One might experience low presence in a lesson if prior knowledge were extremely low and inadequate for the task.

Several new taxonomies for embodiment are being proposed that do not include the third dimension of immersion/presence (Skulmowski & Rey, 2018). In many

[7] This is not to say the distinction between immersion and presence should never be used for MR and/or AR systems. Playing games on smartphones, which are bordered, small screen experiences (not 360) do seem to still induce hours of "presence" in many users.

ways, a two axes model makes for a tidier taxonomy. However, we believe that to reframe the embodied taxonomy for education for 3D immersive VR, a construct for immersion/presence is crucial because presence is one of the unique and profound affordances of VR. The original table (a 3 X 8 matrix) that partitioned the three constructs into high and low spaces can be found in Johnson-Glenberg & Megowan-Romanowicz (2017).

Appendix B

VR and Education Theories

Scholars have been asking for educational research on VR for some time (Mikropoulos & Natsis, 2011), but the resources and affordable technologies were not readily available. Up until 2016, most of the literature on VR and education was based on proprietary VR software and hardware. The research labs, the military, or the commercial companies created in-house products that were too expensive, and unwieldy for public consumption. In 2016, two sets of high-end headsets with hand controllers (Oculus *Touch* and HTC *VIVE*) came to the market. Studies on gesture in VR are slowly coming to light.

In these early days, trial and error play an outsized role in design. Education researchers borrow heavily from entertainment designers, who focus on engagement, and not necessarily on retention of content. This begs the question of whether some rules in the entertainment domain, like "never break immersion", should be violated if higher order learning is to occur? The two lessons highlighted in the next sections were designed using components of three education theories that lend themselves to creating gesture-controlled multi-media content. The three theories are constructivism, guided inquiry and embodied cognition.

4.1 Constructivist Learning Theory

Constructivism builds off of Dewey's (1966) concept that education is driven by experience. Piaget (1997) further describes how a child's knowledge structures are built through exploratory interactions with the world. Environments such as VR can provide opportunities for learners to feel present in goal-driven, designed activities. Further definitions are culled from a teacher's textbook (Woolfolk, 2007). Common elements in the constructivist perspective include:

1. Embed learning in complex, realistic, and relevant learning environments.
2. Provide social negotiation and shared responsibility.
3. Support multiple perspectives and multiple representations of content.
4. Knowledge is constructed (built upon) – the teaching approach should nurture the learner's self-awareness and understanding of ongoing construction.
5. Encourage ownership in learning. (p. 348)

Point 2 regarding social negotiation is important in education. It should be noted that it is still expensive to implement multiuser, synchronized learning spaces. Educational instances of real-time, multi-user social negotiations in VR are coming though (for an update on multi-user VR in education, see Slater & Sanchez-Vives, 2016). In scaffolded, virtual STEM environments, the learners start with simple models and interact to create more complex ones over time. Learners receive immediate feedback and know they are the agents manipulating the objects. They know they are in charge of the constructing. When a lesson is appropriately designed, with incrementally increasing difficultly, and includes evaluative, real-time feedback, then learners are encouraged to become more metacognitive. Learners become evaluative about their output. They can re-submit or reconstruct models multiple times. In this way, agency and ownership are encouraged. Active learning is especially important in the STEM domain where the majority of young learners drop out from studying that subject area over time (Waldrop, 2015).

4.2 Guided Inquiry

Guided inquiry emerged in the late 1980's as an effective practice because it had been shown that free, exploratory learning, on its own, could lead to spurious hypotheses. Minimally guided instruction is "less effective and less efficient" (Kirschner, Sweller, & Clark, 2006), at least until a learner has a sufficient amount of prior knowledge. Students benefit from pedagogical supports that help them construct conceptual models, or knowledge structures (Megowan, 2007). VR can be an important supportive tool in the guided learning domain because real world distractions are mitigated. Guiding learners towards accurate deductions does not mean hand-holding. It means giving just enough information so that the final deduction is made by the students, and they take ownership over what they have learned. Clearly some cognitive effort is needed for learning "to stick"; these concepts are in line with the desirable difficulties literature (Bjork, 1994; Bjork & Linn, 2006), and levels of processing research.

4.3 Embodied Learning

Human cognition is deeply rooted in the body's interactions with the world and our systems of perception (Barsalou, 1999; Glenberg et al., 2013; Wilson, 2002). It follows that our processes of learning and understanding are shaped by the actions taken by our bodies, and there is evidence that body movement, such as gesture, can serve as a "cross-modal prime" to facilitate cognitive activity (e.g., lexical retrieval) (Hostetter & Alibali, 2008). Several studies by Goldin-Meadow's group have shown a direct effect of gestures on learning (Goldin-Meadow, Cook, & Mitchell, 2009). Recent research on embodied learning has focused on congruency (Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2014; Segal, 2011), which posits an alignment of movements or body positioning (the body-based metaphor – see Lindgren's work) and within specific learning domains (e.g., learning about centripetal force and circular motion by performing circular movements as opposed to operating a linear slider bar (Johnson-Glenberg, Megowan-Romanowicz, Birchfield, & Savio-Ramos, 2016)). Virtual and mixed reality environments afford the opportunity to present designed opportunities for embodied interactions that elicit congruent actions and allow learners opportunities to reflect on embodied representations of their ideas (Lindgren & Johnson-Glenberg, 2013).

Embodied learning is probably most effective when it is *active*, and the learner is not passively viewing the content, or watching others interact with manipulables (Abrahamson, 2009; Abrahamson & Trninic, 2015; Kontra et al., 2015). If the learner is induced to handle the physical content, or to manipulate the content on screen then they must be physically active and moving the body (which activates more sensori-motor areas). The new VR hand controls will allow for enactive engagement and high levels of embodiment in lessons. Using virtual content, teachers will not be less constrained by having to purchase specific physical manipulables. What is needed now is a set of design guidelines for educational content being created for VR.

Acknowledgments

Many thanks to James Comstock, Tyler Agte, Dennis Bonilla, the crew at TIMEFIREVR, and Diane Carlson. The 2D version of Natural Selection "Catch a Mimic Game" was funded by NSF grant number 1423655. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Author Information

Mina C. Johnson-Glenberg, PhD

Embodied Games, LLC, and Arizona State University, Department of Psychology, Tempe, Arizona, USA

Email Minaj@embodied-games.com

Please visit the website of the spin-out company: *www.embodied-games.com*