Chapter 5
The *Necessary Nine*: Design Principles for Embodied VR and Active Stem Education

Mina C. Johnson-Glenberg

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 (spin in directional 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, 2017) 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. It will not prove to be a panacea for every disengaged student (as is sometimes touted in the popular press), nor do we expect future scholars to spend entire days in virtual classrooms [see fiction by Cline (2011)]. However, 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

M. C. Johnson-Glenberg (✉)
Department of Psychology, Embodied Games, LLC, Arizona State University, Tempe, AZ, USA
e-mail: Minaj@embodied-games.com

P. Díaz et al. (eds.), *Learning in a Digital World*, Smart Computing and Intelligence, https://doi.org/10.1007/978-981-13-8265-9_5

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, so 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° 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. We believe that gesture and re-enactments using the hand controls (and tracked fingers) will 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 other two systems, and 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 simulator sickness (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). Because randomized control trials (RCT) are just starting to be published on immersive VR in education, it is not possible to do a review. Thus, this chapter focuses on design practices that the author has learned from creating content in mixed and virtual realities for the past 10 years. An early, and evolving, set of design principles for VR in education is provided at the end, and 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. We should make sure we are all on the same page. This section defines some terms still in flux in the field: VR, presence, agency, and embodiment.

1.1.1 VR

In this chapter, the term VR refers to an immersive experience, usually inside a headset, where the real world is not seen for 360°. 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 three-dimensional object or avatar displayed on a regular-sized computer monitor is never "VR"; we hope that educators soon stop conflating the terms and phenomena. It is preferred that PC monitor-supported content be referred to as virtual environment, or mediated or digital (some is even 2.5 dimensional), and that terms like IVR and VR be reserved for the immersive VR experience afforded by headsets (and CAVE systems with no real world components visible).

1.1.2 Presence

The term, presence, as it relates to education is also defined in a recent 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 limbically 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. The *Google Expeditions* series relies on presence to immediately engage learners. A recent exploratory study explicitly states that the presence afforded by the 3D technology "opens up" the senses and mind for learning (Minocha, Tudor, & Tilling, 2017). Minocha et al. further hypothesize that because the students are in control of where they look and for how long, they can now follow "… their interest and curiosity,

[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.

hence giving them a sense of control and empowerment over their own exploration”. Whenever users feel they have control over the environment, they experience *agency*.

Agency underpins the second profound affordance of VR. Interestingly, this author (Johnson-Glenberg) considers the type of VR experience that is purely gaze-based to afford a relatively low amount of agency. (Although, it is still superior to a pre-programmed linear story.) When learners are able to manipulate more objects in the world, with more than a gaze-based lag time signal, we predict more precise and 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 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 individual (self-initiated) control and volition over the individual objects in the environment. In education, 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. It is easier than ever to incorporate gesture and to manipulate objects in VR using this more Natural User Interface (NUI). The second profound affordance of VR is driven by the **ability to gesturally interact with virtual content in 3D** and receive realtime feedback. Our prediction is that hand controls will have long-lasting effects on the types of content, and the quality of the pedagogy, that can be designed into educational spaces. Instructors and researchers are no longer being constrained by commercially available tangibles or peripherals; it is now possible to build or print almost any desired tangible or vessel. (Need to pour from a specialized beaker in a chemistry experiment? You can 3D print the vessel, place trackers on it, add some actuators, and seamlessly simulate complex fluid dynamics—see http://meteor.ame.asu.edu/).

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, 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 bicycle wheels spinning on an axle learned more about angular momentum compared to those who observed the spinning wheel (Kontra et al., 2015). The second agentic example comes from a Jang, Vitale, Jyung, and Black (2016) study. In their yoked-pair design, one participant manipulated a virtualized 3D model of the inner ear, while another participant viewed a recording of the interaction. Results indicate that participants in the manipulation group showed greater posttest knowledge compared to the observation group.

1.1.4 Embodiment

Proponents of embodiment hold that the mind and the body are inextricably linked (Wilson, 2002). Varela et al. (1991) describe cognition as an “interconnected system of multiple levels of sensori-motor subnetworks” (p. 206). In this current chapter, the focus is on learning the content of science. 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. The modality of interest in this chapter is gesture.

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. 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. 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.

Perhaps only concrete words and actions are recalled via gesture? What about abstract words? Repetto, Pedroli, and Macedonia (2017) instructed participants with 30 novel abstract words in a foreign language. The encoding conditions were (1) verbal—read word out loud, (2) picture encoding—read word with a concurrent static graphic image, and (3) gesture encoding—read word while observing a 2-s-long video of someone performing metaphoric gestures that connoted meaning. The group exposed to the gesture encoding videos displayed better word recall. Repetto et al. stipulate that training should occur over multiple days (many of their experiments repeat over three to five days) because procedural learning via gesture takes long(er) to cohere, compared to other types of encoding. Thus, it may be worth the extra trials and training time, because information retrieval post-gestural encoding was also superior in speed of recall and resistance to decay.

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 Here Is First Mention of Congruent

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 (again, strengthening the trace) 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 be in control of the action on screen 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 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 et al., 2014a) 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 will be proposed.

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 compared to 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, Koziupa, & Tolentino 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 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 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.[2]

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 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

[2]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.

affordances of VR. The original table (a 3 × 8 matrix) that partitioned the three constructs into high and low spaces can be found in Johnson-Glenberg and Megowan-Romanowicz (2017).

3 3D Figures for 3D Learning

The reconceptualization of the graphic for embodiment in VR takes into account the continuous nature of the three constructs. Figure 1 also maintains the concept of immersion/presence. The crosshairs in the middle allow the reader the opportunity to partition the space into more tractable low and high construct areas. Figure 1 could even be imagined as eight sub-cubes. Because magnitude of the gesture (i.e., the amount of sensori-motor engagement) may prove to be the least predictive construct for content comprehension, it is relegated to the Z axis. The Z axis, or depth, is usually more difficult to conceptualize in a graphic. The goal for graphics like these is to aid researchers and designers in visualizing embodiment in educational content and aid the community in using the same terminology. These graphics should also spur researchers to assess the orthogonality of the constructs and to design randomized controlled trials to explore the variance associated with each axis (as well as higher order interactions).

The main take away is that a lesson can be deemed high on the embodiment scale if the gestures are congruent and meaningful, 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 learning pedagogies supporting the lessons. Designers and users of VR should be

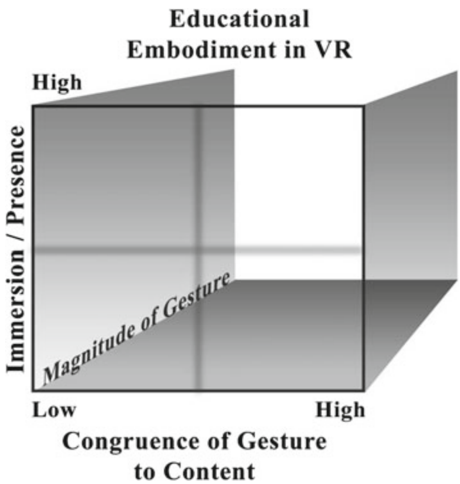

Fig. 1 Cube of Embodiment in Educational VR Content (With permission from Johnson-Glenberg and Romanowicz (2017) from *Frontiers*)

more aware of felicitous learning theories, so a concise summary of three relevant theories, used by this lab, has been included.

4 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 multimedia 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, multiuser social negotiations in VR are coming though (for an update on multiuser 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 1980s 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 now 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.

5 Prudent VR Guidelines Thus Far

For the most part, immersive VR education studies have occurred primarily in adult populations (Freina & Ott, 2015). Health and medicine have been leading the way, with everything from surgical training of craniofacial repairs (Mitchell, Cutting, & Sifakis, 2015) to behavioral change interventions related to PTSD (Rizzo et al., 2010). We are confident that VR will prove to be very useful, but it is currently under-utilized due to price. We predict that when the Standalone headsets (e.g., *Oculus GO*) which do not require phones or separate CPU's, become more affordable, then immersive VR experiences with a hand controller will become 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 for education, Dalgarno and Lee presciently published five affordances for three-dimensional VR environments (Dalgarno & Lee, 2010). We agree with their five and those mesh nicely with Bailenson's below (2016). He 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 profound results (Banakou, Hanumanthu, & Slater, 2016; Hasler, Spanlang, & Slater, 2017). You cannot perceive a photon going directly into your eye in the classroom, but in the next section we describe 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.

Results from the Embodied Games lab's previous mediated research (Johnson-Glenberg 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 content. Thus, our two affordances for education can be meshed with the Bailenson's more general ones. The design principles in the next section highlight how tracked gestures could be incorporated into learning.

6 Design Principles for Embodied VR 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, this 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), and should never prompt for in-game purchases.

Porting learning content to the latest XR environment (the new term for MR, AR and VR) will add another layer of complexity to all learning games, until the conventions and UI 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. Because I am a human, who is shaped as a cognitive psychologist hammer, it is my duty to bang away on the absence of learning scientists on many game design teams. Our absence leads to several easily remediated flaws, such as, pacing issues (typically massed, rarely spaced), ill-conceived reward structures (variable, when interval is more appropriate), and surface-level assessments (often unimaginative, declarative knowledge, dual choice questions at the end of the module). The advice is to procure a strong, multiskilled team and use your funds where they are most needed. In the education domain, probably playtesting is more important than polished graphics.

The original 18 design guidelines in this section have been pulled together with pedagogy and optimal learning in mind. The final *Necessary Nine* at the very end of the chapter have been further culled with financial constraints and development realities in mind. Also, 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 and education, especially with a focus on using hand controls for STEM learning. Our focus is on making VR content 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. Each guideline ends with a sentence of support or a citation should readers want to dive deeper. A version of these guidelines first appeared in Johnson-Glenberg (2018).

6.1 Education in VR—General Guidelines

- **Assume every learner is a VR newbie—*start slow***
  - Not everyone will know the controls. Not everyone knows to 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.
  - The user should not capture butterfly #1 at 10° and then capture butterfly #2 at 190°. Be gentle with users' proprioceptive systems (where the body is in space). Watch out for large body-action disconnects, 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. Our "start slow" advice comes from years of designing educational content.
- **Introduce User Interface (UI) components judiciously, fewer is better**
  - Keep the screen clean. Permanent objects (i.e., a timer that stays center-screen as the player turns head) will break the presence/immersion feel. When users build the first fireworks in our fireworks/chemistry lesson, they can only make one stage rockets. The more complicated multistage components are not available in the interface until users show mastery of the simpler content. Add visual complexity when the user is acclimated and ready (Johnson-Glenberg et al., 2014).
- **Scaffold—introduce cognitive steps one at a time**
  - Build up to cognitive complexity (Pea, 2004) as well. In the electric field series[3] of seven mini-games, users are not immediately exposed to the multivariable 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

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

lesson that incorporates all the previously learned content and culminates in the formation of lightning (Johnson-Glenberg & Megowan-Romanowicz, 2017).

- **Co–design with teachers**
  - 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 important content standards.
- **Use guided exploration**
  - 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, and Clark (2006).
- **Minimize text reading**
  - 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 (next section), players do not read lengthy paragraphs on butterfly cocoons, instead a short cut-scene animation of butterflies in chrysalis and quickly emerging is displayed.
- **Build for low stakes errors early on**
  - Learning often requires errors to be made. Learning is also facilitated by some amount of cognitive effortfulness.
  - In our recent Catch a Mimic game, the player must deduce which butterflies are poisonous, just like a natural predator must. In the 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 toxic from non-toxic without feedback on both types. Some false alarms must be made. Later in the game, errors are weighted heavier. 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.

- **Give players unobtrusive, immediate, and actionable feedback**
  - This does not mean *constant* feedback (Shute, 2008). 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!)**
  - 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 with dyads could work—one partner outside the headset asks questions of the one inside?

- **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 on a handheld, or building sequential tasks that require back-and-forth asynchronous activities. A classroom collaboration and cooperation classic is Johnson and Johnson (1991).

6.2 Using Hand Controls/Gestures

The following design guidelines focus on using the hand controllers in VR for learning.

- **Use the hand controls to encourage the learners 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 or 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 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, it is easy for the learner to use a gesture to fill a bar to the correct height. An upward swipe is also congruent with our cultural concept of higher (see Abrahamson's work for embodied and mediated examples of proportional reasoning https://edrl.berkeley.edu/edrl_publications). In the Munch a Mimic game, students are asked make a prediction about species survivability using the hand controls (see Fig. 5, next section). We also note that prediction is a well-researched metacognitive comprehension strategy (Palinscar & 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, it 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 previous section on embodiment for multiple citations, or read Johnson-Glenberg and Megowan-Romanowicz (2017) for an example of a mixed reality RCT where active and embodied mini-games resulted in significant learning gains and increased engagement scores.
- **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 motion sickness (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 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, this is acknowledged to cost more, but the learning content level should 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 more costly, but it is considered one of the best practices in educational technology (Kalyuga, 2009); it is something to strive for.

7 Case Studies: Two Examples

This section highlights relevant education theories and design principles in two case studies. The first example showcases some of the changes that occurred as 2D content was repurposed to a 3D VR lesson. The second example highlights the design techniques of construction and guided exploration. Both of the VR lessons are based on theories of guided inquiry and embodied learning. Scaffolding, reflection, and creative assessments are strongest in the first lesson on natural selection. Constructivism, agency, and guided exploration will be further discussed in the second lesson on chemistry via fireworks.

7.1 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 to *not 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. This version of the pattern matching game ended with several open-ended and multiple choice questions. Design was constrained because we could not include explicit text that described how mimicry occurred. We have since added text to make this a standalone lesson. This earlier 2D assessment was given at multiple time points around movie viewing: pre-, post-viewing and after a two week delay.

[4] The movie is called *Amazon Adventure*. The funding agency for the assessment tool was National Science Foundation, grant # 1423655. A WebGL version of the game can be played at www.embodied-games.com.

7.1.1 Tablet Version of the Natural Selection Game

The butterflies would spawn from the right side of the screen and the background of a forest would slowly scrolls 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 made it disappear. Immediate feedback was given (visually, not auditorially as this was a test taken by entire classrooms). Persistent feedback was displayed on top of the screen as to whether the choice was poisonous (red outline boxes), or non-poisonous (green outline boxes) showed tapped butterflies. As the levels progressed, the non-poisonous butterflies naturally selected to more closely resemble the poisonous butterflies. On the *Kindle FIRE* 10 tablet, the actionable play space was only 7.0 in. (diagonal) and so no distractors were included, i.e., falling leaves, particles in foreground, moving water, as there was not enough room.

The next step was to turn the assessment tool into an engaging, instructional game on the topic of natural selection. The format should match a ubiquitous digital form factor in schools, but not move to a smaller screen because the game depends on detailed visual pattern matching. Thus, a PC format with mouse input was chosen.

7.1.2 PC Version of the Natural Selection Game

In the current PC version of the Natural Selection Game (release February, 2019), the mouse controls the location of a net. A mouse click captures a butterfly, see Fig. 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 in.), so now the butterflies spawn and fly out of a central bush, and higher resolution is possible. Because the game would eventually move to VR the team decided that flying and swooping as a bird could make the player nauseous, and so the bird POV was abandoned.

The PC version is not constrained to be an assessment tool. The re-design to an instructional lesson includes more embedded text. Appropriate game elements to enhance engagement were included, e.g., moving waterfalls, visual distractors, and audio. Chirping birdsong helps to increase presence. Feedback is handled differently on the PC. We wanted to declutter the screen so we removed the permanent feedback at the top of the screen. Now on the bottom right of the screen (pinned to the world, BUT not to the HUD) is performance feedback that is numerical. In addition, audio feedback, as positive or negative sounds upon collision with a butterfly has been added. Now, gameplay encourages the player to remember the butterfly that was just captured, feedback, as a green heart or red skull, shows up upon collision on the central screen for 1.5 s. On the bottom right is persistent (unpinned) numerical feedback on type of butterfly captured. The ongoing count is displayed next to either

Fig. 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 very static

the green heart icon (non-poisonous) or the red skull (poisonous).[5] The timer restarts at 60 s for each of the six levels.

7.1.3 VR Version of the Natural Selection Game

The background and moving assets (butterflies, etc.) were then rendered into 3D. The VR version has graphics that curve around 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. This maintains presence/immersion. 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 Fig. 3, note the ghostlike avatar hand that is mapped to the human player's hand and wrist movements. In this *Oculus* version *(Rift* and *GO*), the hand grips around the net handle. The net is fully articulated in three axes.

[5]Late stage playtesting with colorblind males, revealed that red and green remained poor choices for feedback. Even though we knew this at the onset and tried to compensate with a second feedback signal of shape, i.e., two different icons: heart versus skull. The images were just too small to be easily distinguished, and this needs to be addressed in the next version.

Fig. 3 The VR version with flowing water and an articulated avatar hand and wrist

7.2 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 must not "necessarily be quantitative" (Schaub & Abadi, 2011). Middle schoolers can make inferences and predictions about joint likelihoods without memorizing statistical formulae. Prediction is one of a set of powerful and well-researched comprehension strategies (Rosenshine & Meister, 1994). Primary school students are capable of making predictions in the middle of STEM texts and then evaluating the outcomes by the end of the text (Palinscar & Brown, 1984). 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 in the 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 predictive).

Figure 4 shows the interactive bar chart prompt, after an active submission. That is, the first green-filled bar is animated to show it filling up and it is then locked in the HIGH position. The learner must make the best guess as to the survivability of the next four species, i.e., the poisonous butterfly (#2) and the three non-poisonous butterflies to the right (#3, #4 and #5). Dragging the blue oval moves the green fill in the bars (which snap to either low, medium, or high positions). When learners are

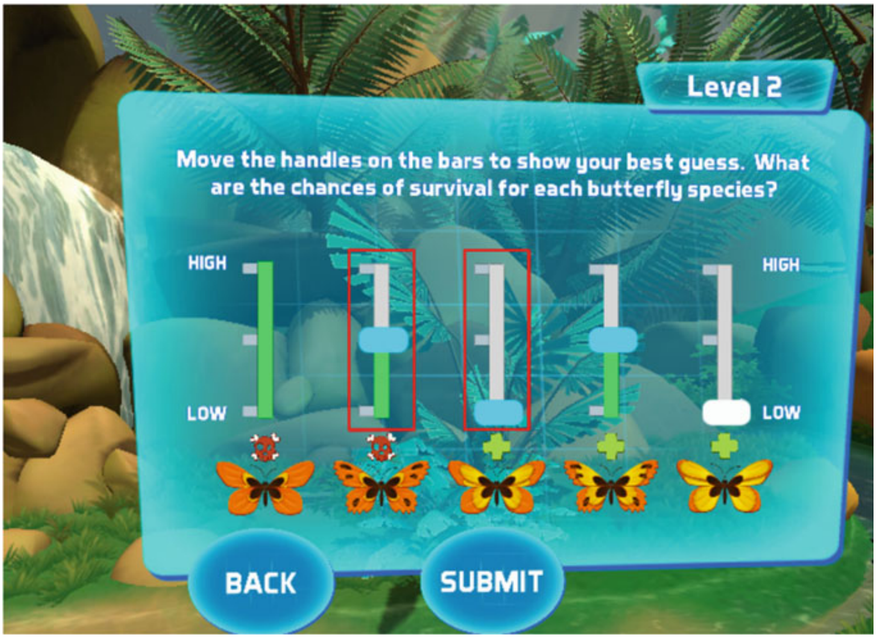

Fig. 4 Interactive assessment in both 2D and 3D VR versions—with corrective feedback. The red boxes around butterflies #2 and #2 connote incorrect

satisfied with their decisions, they click on the submit button. Learners are allowed three incorrect submissions before an animation shows the correct answer.

7.3 *Example 2. A High Embodied VR Lesson with Hand Controls. Topic: Chemistry/Physics*

Example 2. CHEMISTRY—Fireworks
The second example is special because it is multiplayer. Multiplayer mode is still expensive, but it is coming! This module was designed in 2017 to highlight constructivism and scaffolding, it was included in a multiplayer entertainment game that is currently available (although later versions may vary). *Hypatia* is a multiplayer 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 developers at the game company followed the mantra: "never break immersion". But, learning scientists know it is also important to build in time for reflection so that students can create meaning around intense and novel stimuli.

The never-break-immersion guideline from entertainment may not migrate well into the education domain. In a cognitive, goal-driven learning situation, it may be efficacious to request learners remove the headset to make handwritten notes or to engage in face-to-face collaborations/question sessions with a partner. These are empirical questions. We do not yet have the answers.

In the *Hypatia* game world, players first create non-humanesque avatars by choosing from a library of body parts. This module described was 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 one. As a form of UI (user interface) 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.

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 would be picked up. The first series of grey metal salts (see Fig. 5) 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

Fig. 5 The strontium Bohr model. Note the wave heading towards the avatar's eye is a red wave

would place each salt into the flame of the Bunsen burner and note the color that the salt burned.

Figure 5 shows the avatar (Jessica) 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 Jessica 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 Jessica, the human player, is seeing the atom floating towards her in 1st person or a "head on" POV. This is very engaging, indeed it could be alarming if it moved too quickly.

After she 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. Quickly, the electron falls back to its more stable orbit, as it does this a packet of energy called a photon is released. This photon is perceived in the red spectrum. In Fig. 5, the photon has been visualized as *both* a red wave and a particle heading towards the eye.[6] Jessica is watching the dynamic model in 3D and she perceives the photon as traveling directly into her eye. (This is perhaps the only thing humans want heading directly towards our eyes!) The sinusoidal movement was designed to be somewhat slow, so it would not be frightening.

The simulation of the photon as a wave reifies the concepts that energy is released by the heat burst, and that 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 signposted to exit the back door to the multistaging firework building area.

This is where the social and collaborative aspects comes into play, because other experts are often out by the lake building multistage rockets and can give feedback and clap when the final version is correct. We scaffolded the difficulty of building the rocket. The player is first asked to build a one color firework. Then players are requested to make their rockets burst in a predetermined sequence of multiple colors. If a player is having trouble, someone else in the game can come over to help. When a rocket explodes correctly, there are often group shouts of approval. 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. After some minutes of free exploration, they are instructed to build specific multistage, color sequence rockets. This is an 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.

[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.

8 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 rapidly changing. The previous section describes the multiple embodied principles in detail. This chapter ends with the top contenders below. These are easier to recall, and if there are only resources to focus on a subset of the main guidelines, then the author recommends the *Necessary Nine*:

- Scaffold cognitive effort (and the interface)—one step at a time
- Use guided exploration
- Give immediate, actionable feedback
- Playtest often—with correct group
- Build in opportunities for reflection
- Use the hand controls for active, body-based learning
- Integrate gestures that map to the content to be learned
- Gestures are worth the time and extra expense—they promote learning, agency, and attenuate simulator sickness
- Embed gesture as a creative form of assessment, both during and after the lesson.

9 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 over-arching questions regarding *when* to insert a VR lesson. Aukstakalnis (2017) shares an anecdote about a student in a design class who regrets designing his first project in a VR headset during the year-long course because he missed watching his peers work in the real world. The student admitted that he "missed learning from his peers' collective mistakes" (p. 306). This is an instance of the timing being off; perhaps the digitized platform should have been made available after a real world introduction.

Clearly more research is needed on learning in VR, and 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, and the controllers can be used for innovative types of assessment. A new graphic "cube" is introduced to help visualize the amount of embodiment in immersive educational lessons. It is our hope that the case studies and the set of design guidelines will help others to design optimal immersive VR lessons.

Acknowledgements 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.

References

Abrahamson, D. (2009). Embodied design: Constructing means for constructing meaning. *Educational Studies of Mathematics, 70,* 27–47. https://doi.org/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). https://doi.org/10.1007/s11858-014-0620-0.

Alaraj, A., Lemole, M. G., Finkle, J. H., Yudkowsky, R., Wallace, A., Luciano, C., et al. (2011). Virtual reality training in neurosurgery: Review of current status and future applications. *Surgical Neurology International, 2,* 52. https://doi.org/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.

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. (2017). *Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do*. New York, New York: W.W. Norton.

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). http://doi.org/10.3389/fnhum.2016.00601.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and 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.

Cline, E. (2011). *Ready Player One*. New York, New York: Random House.

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 and 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. https://doi.org/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. https://doi.org/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. https://doi.org/10.1177/1745691613498098.

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

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

Goldin-Meadow, S. (2014). Widening the lens: What the manual modality reveals about learning, language and cognition. *Philosophical Transactions of the Royal Society, Biological Sciences, 369*(20130295). https://doi.org/10.1098/rstb.2013.0295.

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 & 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. https://doi.org/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). https://doi.org/10.3389/frobt.2018.00081.

Johnson-Glenberg, M. C., Birchfield, D., Koziupa, T., & Tolentino, L. (2014a). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology, 106*(1), 86–104. https://doi.org/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. https://doi.org/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). https://doi.org/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. https://doi.org/10.3389/fpsyg.2016.01819.

Johnson-Glenberg, M. C., Savio-Ramos, C., & Henry, H. (2014b). Alien Health: A nutrition instruction exergame using the *Kinect* sensor. *Games for Health Journal: Research, Development, and Clinical Applications, 3*(4), 241–251. https://doi.org/10.1089/g4h.2013.0094.

Johnson-Glenberg, M. C., Savio-Ramos, C., Perkins, K. K., Moore, E. B., Lindgren, R., Clark, D., et al., Squire, K. (2014, June). *Science Sims and Games: Best Design Practices and Fave Flops.* Paper presented at the International Conference of the Learning Sciences (ICLS), Boulder, CO.

Johnson, D. W., & Johnson, H. (1991). *Learning together and alone: cooperation, competition, and individualization*. Englewood Cliffs, NJ: Prentice Hall.

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. https://doi.org/10.1080/00461520.2016.1155457.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966–968. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/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. https://doi.org/10.1016/j.compedu.2010.10.020.

Minocha, S., Tudor, A., & Tilling, S. (2017). Affordances of mobile virtual reality and their role in learning and teaching. *Proceedings of British HCI, Digital Make Believe, 1*.

Mitchell, N., Cutting, C., & Sifakis, E. (2015). GRIDiron: an interactive authoring and cognitive training foundation for reconstructive plastic surgery procedures. *ACM Transactions on Graphics, 34*(4), 1–12. https://doi.org/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. 10.1016.

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

Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117–175. https://doi.org/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. https://doi.org/10.1207/s15327809jls1303_6.

Piaget, J. (1997). *The language and thought of the child* (3rd ed.). London: Routledge.

Repetto, C., Pedroli, E., & Macedonia, M. (2017). Enrichment effects of gestures and pictures on abstract words in a second language. *Frontiers in Psychology, 8*(2136), 11. https://doi.org/10.3389/fpsyg.2017.02136.

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. https://doi.org/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. https://doi.org/10.1007/s10336-010-0632-7.

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

Segal, A. (2011). *Do Gestural Interfaces Promote Thinking? Embodied Interaction: Congruent Gestures and Direct Touch Promote Performance in Math*. Dissertation, ERIC. ISBN: ISBN-978-1-1246-2610-9.

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. https://doi.org/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. https://doi.org/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). https://doi.org/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. https://doi.org/10.1162/105474698565848.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

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. https://doi.org/10.1038/523272a.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & 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.