“Games as popular art forms offer to all an immediate means of participation in the full life of a society, such as no single role or job can offer to any man.” Marshall McLuhan¹

With the rapid development of computer and game console hardware, graphics, artificial intelligence and network technologies in recent years, video games are becoming more and more capable of providing a vivid, fictional reality within which players can immerse, and have attracted more diverse demographics. Increasingly, studies have shown that players’ experiences in games can have cognitive an affective impact on their lives outside the game world; skills learned and knowledge gained in games can often be transferred to real life economies. For example, it has been shown that playing games can help individuals practice their eye-hand coordination, and also exercise their social skills and emotional coping strategies. Serious games are designed to systematically leverage this effect and positively influence the players’ lives outside of games.

This paper discusses how serious games can benefit by incorporating frameworks of embodied cognition and how new game peripherals, such as the Microsoft Kinect, can help realize the potential of embodied cognition in games.

1. Serious Games

Serious Games, and its other monikers, Games for Change, Games for Health and Games, Learning and Society, is an exponentially expanding global movement, involving diverse sectors of society from policy makers and educators to funders, game designers and the media. Last summer’s annual Games for Change (G4C) conference alone hosted over 800 participants from around the world and was keynoted by former Vice President Al Gore. Serious game design practitioners and theorists are also being courted at the White House to help mobilize President Obama’s Digital Promise initiative, and invited to other government conferences, such as Tech@State, to shape discourses on the future of education and technology. And this year, the

Games Development Conference (GDC), the world’s largest “professional-only” game industry event, has added a new G4C track, indicating that the gap between serious games and the mainstream industry is closing. Game scholar Stewart Woods has even ventured that serious games are the goal of those within the game industry for the future of games, noting that “many developers wish to create serious content or experiences that are typically represented within traditional narrative forms such as books or film.”

Almost all books on game design cite the Dutch anthropologist, Johan Huizinga, as the first to interrogate, though never quite resolve, the distinction between seriousness and play. But a lesser known artifact is Clark C. Abt’s 1970’s book entitled “Serious Games,” which addresses the use of analog board games in education, science, government and industry. In his first chapter, titled “The Reunion of Action and Thought,” Abt forwards an surprisingly anticipatory definition of serious games: “We are concerned with serious games in the sense that these games have an explicit and carefully thought out educational purpose and are not intended to be played primarily for amusement...this does not mean that serious games are not, or should not be entertaining.”

As early as the 1980s, scholars, too, began to see the potential of video games for education and enhancing learning outcomes. Game assessment research, therefore, primarily targets five main areas: “learning to learn” through novel environmental engagement, skill development, such as inductive discovery and problem-solving, and hand-eye coordination, motivation, memory retention, and utility for special groups such as attention-deficit children, the elderly or differently abled.

For most of these researchers, serious games are designed with the explicit goal of helping students learn about important subject matter in an enjoyable, engaging way, and research measurements narrowly evaluate increases in problem-solving strategies, and cognitive or social skills. Yet at present, there still is not a large body of literature “scientifically” supporting speculative claims that games facilitate learning.

As Arthur Graesser, Patrick Chipman, Frank Leeming, and Suzanne Biedenbach state in their article “Deep Learning and Emotion in Serious Games,” “[u]nfortunately, at this point in the learning sciences very few serious games have been developed that would impress experts in education.” The authors imply that it is not clear if game design is systematically aligned with or incompatible with pedagogy and curricular objectives. To combat this uncertainty, they recommend that:

> [t]he field needs a theoretical framework that maps the game genres and game features onto theoretical components of cognition, emotion, motivation, and social interaction. Deeper levels

---


3 Abt, Serious Games. need full citation.
of learning would involve many different elements, including an analysis of causal mechanisms, logical explanations, creation and defense of arguments, management of limited resources, tradeoffs of processes in a complex system, and a way to resolve conflicts. More shallow level include perceptual learning, motor skills, definition of words, properties of objects, and memorization of facts.⁴

They conclude that aside from the depth of the skills afforded by a serious game, there is the persistent question of their utility and relevance to the real world. As Graesser et al bemoan:

[t]he scientific status of game features proposed by game designers is greatly in need of computational and empirical inquiry. Aside from the computational essence of games, there is the question of what makes them successful psychologically. At this point in the science, there are few firm answers to such questions about the essence of games and their psychological impact.⁵

Even when the empirical data exists, the quality of research design possesses severe shortcomings. In “Evaluating the Potential of Serious Games: What Can We Learn from Previous Research on Media Effects and Educational Intervention,” Marco Ennemoser makes a strong and well-reasoned critique of current “naive” research methodologies employed in studying serious games, and in particular, the over-reliance on “media effects.” Invoking McLuhan, he counsels “[t]he medium is not the message. Don’t study media effects.”⁶ He calls for better assessment of game-playing outcomes, more consideration of the specific mediating processes by which serious games produce effects, and for introduction of variables that moderate the impacts of game playing on individuals into serious games effects theories.

Christoph Klimmt expresses similar disappointment with the limitations of media effects theory when evaluating the accuracy of serious games impact at the level of societal transformation. In his article “Serious Games and Social Change: Why They (Should) Work,” he proffers:

[s]ociety-level and group-level effects of communication campaigns depend on the ability to reach and influence a sufficiently large number of individuals, including opinion leaders, innovators, or influencers. Because social change involves complex processes of knowledge acquisition, attitude change, follow up-interpersonal communication, and collective action, a potentially large number of cognitive/attitudinal, motivational/affective, and behavioral media effects needs to be assumed as part of the communication process of social change, which can complicate serious games research.⁷


⁵ Ibid., 82


Enjoyment, perhaps, is one of the most taken for granted, and often under-investigated assumptions pervading the impact of serious games on learning, development and social change. The player’s reaction to serious games implicitly supports an increase in enjoyment, which translates into a new found interest in a topic, what Csikzentmihaly calls the flow experience. Flow occurs when the learner has such intense concentration that time and fatigue disappear.

Yet in a wide evaluation of “serious games” by Cuihua Shen, Hua Wang, and Ute Ritterfeld reveals that the majority of serious games do not achieve the “threshold of enjoyability.” As Shen et al report:

\[\text{for the serious games examined in this study, we found that most passed the threshold of technological capacity, aesthetic presentation, and game design. A few of the games were less enjoyable due to technological glitches. Most of the other games fell into a second category: they were playable, at rather average fun level, but not highly enjoyable. These games were generally stable technologically, but might have some problems with control or being less sophisticated aesthetically, which inhibited their enjoyability. In order to reach the third stage—to be a highly enjoyable game that is often deliberately selected and played over a longer period—a game must utilize both narrativity and social interaction to promote a player’s emotional engagement and elevate the level of pleasure in game play.}\]

The above researchers call for both an in-depth, scientific analysis of the impact of serious games, and also more effective approaches to designing serious games. This paper, thus, aims to establish a framework for increasing the enjoyability and effectiveness of serious games by integrating theories of embodied cognition with the application of emerging technologies, like the Microsoft Kinect, that afford kinesthetic engagement. In the next sections, we will introduce the current theories of embodied cognition, examine case studies of serious games, which employ embodied technology, and discuss how practitioners and scholars can effectively combine embodied theories and emerging technologies to activate social and behavioral change, thereby increasing social impact.

2. Embodied Cognition
The theories of embodied cognition highlight the close relationship between our body, mind and the environment. Many of our concepts are intimately related to aspects of our body and bodily movement. According to this approach, our ideas and understanding of basic concepts such as “up,” “down,” and “over” are shaped by the nature of our body and its relationship to the environment. In fact, even many of our mathematical concepts, such as the continuity of a function, which are supposed to be highly abstract, can be traced back to very basic spatial, perceptual and motor capacities (Lakoff, 2001). Jean Lave and Etienne Wenger explicitly claim

---

that "there is no [mental] activity that is not situated." Clark, therefore, maps out a more granular depiction of cognition distributed between mind, body and world, which he labels "principles of ecological assembly," wherein he proposes that the embodied agent (what he labels the "canny cognizer") recruits, sometimes, exploits, often on the spot, a mix of problem-solving resources/opportunities provided by dynamic loops of perceptual and motor routines, combined with neural processing and storage, active sensing, and iterated bouts of environmental affordances, which “produce self-stimulating cycles of material scaffolding to yield an acceptable result with a minimum of effort, and no ‘central meaner’” (Clark, 137).

2.1 Decision Making
Embodied cognition theory points out that our decision making process is embodied, which means it does not happen purely in our mind, but leverage resources in the environment and in particular our body-environment relationship. Therefore, physical movements can help us to think, solve problems, and make decisions. This has been evidenced in many scenarios. For example, people often walk around a room while trying to decide where the furniture should go, and we do not rotate Tetris pieces because we know where they can go, but rather as a method to find out where they can go.10

2.2 Emotion
Both our own emotional experience and our ability to understand another’s emotional experience are deeply rooted in our physical body. This can be traced back to the James-Lange theory of emotion, which states that “emotion is the mind's perception of physiological conditions that result from some stimulus.”11 The theory (arrived at by William James and Carl Lange independently in 1884) famously underscores this by way of example, when James remarks: “Common sense says, we meet a bear, are frightened and run...The hypothesis here to be defended says that this order of sequence is incorrect....Without the bodily states following our perception, the latter would be purely cognitive in form. It is rather that we see a bear and run, consequently we fear the bear.”12 More recent studies in neuroscience have identified this phenomenon with mirror neurons, which can be triggered by either imaginary body movements (i.e. simulations in the mind) or the observation of another person’s movements

---


11 James-Lange Theory of Emotion. need full citation.

12 James, William. *What is Emotion?* Mind, ix. pg. 189. need full citation.
(even an avatar on a screen) in the same way as when we move our own body.\textsuperscript{13} This provides a physiological basis for how people can understand each other, have empathy, and treat one's own gesture as a social stimulus.

2.3 Perception
Even our basic perceptual functions are not straightforward and immediate processes. Instead, perception requires interaction with the environment. As Alva Noe, a proponent of the enactive view, observes the "sensorimotor model of perception suggests that an important role is played by embodied action in terms of information pick up and initially tuning circuitry which supports perceptual awareness."\textsuperscript{14} For Noe, sensory perception is not just something that unfolds in the brain, but instead a mode of active and motivated exploration of the environment drawing on an implicit understanding of sensorimotor regularities. To model vision correctly, for example, he believes "we must model it not as something that takes place inside the brain, but as something that directly involves not only the brain, but also the animate body and the world."\textsuperscript{15} When discussing human perception and action, J.J. Gibson similarly establishes that the way our body acts serves two functions simultaneously: one is approaching to accomplish the current task, such as standing still, catching a ball or reaching a destination, and the other is information collection for the next cycle of action planning.\textsuperscript{16}

2.4 Memory
Moreover, our memory is often off loaded to the environment. This can best be illustrated through Gibson’s concept of affordance, which he describes as a person-environment relationship, such as the height ratio of the stairs to a person’s leg length, which can be immediately picked up by the person for deciding how he/she should act. In fact, Gibson’s view of perception and action does not include the construct of memory. Instead, past experience is encoded into various affordances of the environment.

Memory can also be off loaded to body-environment relationships that are created by the individual. Kinesthesia is such an example. In the Goldin-Meadow’s 2003 study, two matched groups of children were asked to memorize a list, and then carry out some mathematical problem solving before trying to recall the list. In the study, one group could freely gesture during an intervening math task, while the other group was asked not to move while conducting the same task. The results showed that the group that was not allowed to gesture


\textsuperscript{14} Noe, Alva. \textit{Action in Perception (Representation and Mind)}. (Cambridge, Ma: MIT Press, 2005), 87.

\textsuperscript{15} ibid, 30.

\textsuperscript{16} Gibson and Gibson, 77. need full citation.
performed significantly lower in the memory recall test than the group that was able to move around during the intervening math task. As Goldin-Meadow concluded, "the physical act of gesturing plays an active (not merely expressive) role in learning, reasoning, and cognitive change by providing an alternative (analog, motoric, visuospatial) representational format."\(^{17}\)

### 2.5 Embodied Cognition Departs from Earlier Theories of Cognition

The theories of embodied cognition are fundamentally different from many other prevailing theories of cognition. For example, René Descartes described a dualistic relationship between our mind and body, which claims that the human mind and consciousness are an abstract entity which is separated from the physical body. Early theories of human cognition, stemming from cybernetics, perceived the mind as a collection of mental processes and functions. Herbert Simon defines cognitive science as "the study of intelligence and intelligent systems, with particular reference to intelligent behavior as computation."\(^{18}\) While this definition does not completely rule out the possibility that our physical body plays a role in consciousness and cognition, it clearly places more emphasis on computation and information processing, and inscribes the computer as a perfect metaphor for how our mind functions. Howard Gardner goes further and argues that it is necessary (and sufficient) to "posit a level of analysis wholly separate from the biological or neurological, on the one hand, and the sociological or cultural, on the other."\(^{19}\) This view suggests that human consciousness and cognition can be simulated or created from non-physiological forms using representations unrelated to our physical body.

Recent studies into cognitive extension, which reify the critical role both the body and the environment play in instigating and sustaining cognitive processes, therefore, offer insight into the potential of using embodied technologies to enhance deeper learning outcomes in serious games. Most existing games rely on generic, mechanical gestures imposed by game controllers, which have little to do with the player’s physical actions in the game world. Embodied technologies, on the other hand, like the Microsoft Kinect, allow for more organic engagement and direct mirroring of in-game movements, thereby increasing cognitive scaffolding, and catalyzing higher levels of learning. As Wayne Gray and V.D. Veksler support, “the degree of embodiment” is directly proportional to our reliance upon external scaffolding; the more immersive the environment (virtual, real or mixed reality), the more we offload. To unpack the interplay between cognitive extension, embodied technology and serious games, we will more closely examine two key features of cognitive scaffolding.

---


\(^{19}\) Gardner, Howard. 1985. pg. 6. need full citation.
3. Cognition Extension Enacted in Serious Games through Embodied Cognition

In the much cited, seminal Tactile-Visual Substitution System study conducted over the course of several years by Bach y Rita, a blind subject is rigged into a head-mounted camera with sensors attached to their thigh. When visual stimulation enters the camera, the images are transduced to trigger an array of vibrations on the subject's thigh, which then activate "quasi seeing" without the parts of the body and brain normally dedicated to seeing.

As Andy Clark acknowledged of the study, such sensory substitution reveals that "even without penetrating the existing surface of the skin and skull, sensory enhancement and bodily extension are pervasive possibilities" for inducing new agent-world circuits. Such cross-modal substitution, however, requires goal-driven sensory-motor engagement for adaptation to be successful. Games, of course, provide this.

Games combined with an awareness of Clark’s extended cognition theory, offer novel opportunities for learning, health-related issues and behavior change. Because games, enhanced by the new affordances of the Microsoft Kinect are grounded through bodily states, such as simulation and situated action, social interaction, affect and emotion, they provide a rich opportunity for social and behavioral change.

Specifically, we suspect serious games can be enhanced by exploiting two concepts derived from cognitive scaffolding onto a simulated environment through kinesthesia: 1) niche construction and 2) the dynamic coupling effect.

3.1 Niche Construction

Defined by Laland et al (2000), niche construction consists of:

> The activities, choices and metabolic processes of organisms, through which they define, choose, modify and partly create their own niches. For instance, to varying degrees, organisms choose their own habitats, mates, and resources and construct important components of their local environments such as nests, holes, burrows, paths, webs, dams, and chemical environments.

By extension, cognitive niche construction is an iterative process whereby we build physical

---

20 Studied conducted in 1972, 83,84, 96.


22 ibid, 61.
structures that transform problem spaces in ways that aid (or sometimes impede) thinking and reasoning about some target domain(s). Such organism-induced modification of virtual spaces in games, which often mimic real spaces, but are far more permeable, offer an alternate mechanism for us to enhance problem-solving and skill attunement through simulation, and alternately try on new forms of thought, reason and even identity construction. As game designer Gonzalo Frasca espouses in his seminal essay, “Video Games of the Oppressed,” game worlds function as a microcosmic sandbox for experimenting with new ways of understanding and re-envisioning complex, macrocosmic systems. Informed by Augusto Baol’s theatre techniques, Frasca writes:

The goal of these [serious] games is not to find appropriate solutions, but rather to trigger discussions...It would not matter if the games could simulate the situation with realistic accuracy. Instead, games would work as metonyms that could guide discussions and serve to explore alternative ways of dealing with real life issues.  

Interestingly, in grounded cognition theories, such as those put forth by Lawrence Barsalou, simulation is a naturally occurring phenomenon, forming a core computation in the brain itself. For Barsalou, simulation is the re-enactment of perceptual, motor and introspective states acquired during our daily experience with the world, body and mind. As an experience occurs, he contends, "the brain captures states across the modalities and integrates them with a multi-modal representation stored in memory. Later when knowledge is needed to represent a category, multi-modal representations captured during experiences with its instances are re-activated to simulate how the brain represented perception, action and introspection associated with it." One might assume, therefore, that our proclivity to rely on simulation as memory reinforcement, might render us susceptible to game simulations as a replacement, or a substitute, for the above described naturally occurring function. We, therefore, hypothesize that if humans are provided a cognitive niche construction in the form of an embodied virtual simulation, where gesture plays an integral role, the tendency would be to somehow shift or reduce aspects of the overall neural cognitive load by displacing processing onto the extended tool, thus freeing up resources for the memory task, and enabling higher assembly processes to take place.

3.2 Dynamic Coupling
In Goldin-Meadow’s study which we mentioned above, it appears that gesture continuously informs and alters verbal thinking, which continuously informs and alters gesture, forming a coupled system, in which the act of gesturing is not simply a motor act expressive of some fully

---


neutrally realized process of thought, but instead a "coupled neural-bodily unfolding that is itself usefully seen as an organismically extended process of thought." Movement, in essence, is a form of thinking.

Interestingly, emotion and affect rely upon a similar dynamic coupling, or feedback loop. Emotions, Damasio claims, are “complex, largely automated programs of ‘actions’ concocted by evolution,” and like other aspects of our brain function and consciousness, emerged, and continue to evolve, as a “life regulator.” The actions, he suggests, are “complemented by a cognition program that includes certain ideas and modes of cognition, but the world of emotions is largely one of actions carried out in our bodies, from facial expressions and postures to changes in viscera and internal milieu.” Thus, affective cues in games, which trigger dopamine, causes an emotion-feeling cycle, which serves to sustain engagement, which in turn deepens learning outcomes.

### 3.3 Realize Embodied Cognition in Serious Games using Kinect

The Microsoft Kinect, which came to market in November 2010, is an add on device for the XBox 360, which consists of software and an on range camera technology, which interprets 3D scene information from a continuously projected infrared structured light. The sensor features a "RGB camera, depth sensor and multi-array microphone running proprietary software, which provides full-body 3D motion capture, advanced gesture and facial recognition and voice recognition capabilities," per the company website. Still fairly primitive, the Kinect can only capture twenty joints per player, and track up two active players, who must activate the sensor through hand gestures for the scene to calibrate. But once activated, the Kinect provides a device-free interface for the player to use their body movements and gestures to naturally interact, thereby embodying communication.

As discerned above, the environment, virtual or physical, in which perceptually guided actions take place offers individuals a material outlet for cognitive scaffolding (defined by Vgotsky as the process by which a learner leverages the environment to aid in learning by offloading low level brain activities). Therefore, if the Kinect's sensor-driven interactive screen, which responds to gesture and voice, serves as both the virtual environment and another layer of feedback on top of the self-stimulation generated by the gesture alone, the screen can potentially function as a quasi, real time cognitive niche construction. In doing so, this malleable and responsive

---


27 ibid, 114.
substitute for a physical environment lends itself to programming--constructing--functionality to enhance learning and modify behavior change.

4. Embodied Learning: A Case Study

While embodied learning practices facilitated by sensor-based technology is still relatively new, a few researchers, like ourselves, have begun to document its promise in educational settings.

A seminal study on "Technology-enabled Embodied Learning Environments" initiated by Mina C. Johnson-Glenberg, David Birchfield, Lisa Tolentino, and Tatyana Koziupa as part of a STEM research grant recently concluded at Arizona State University's Situated Multi-Media Arts Learning Lab (SMALLab). Conducted over two six day sessions, with high school students, the same teacher, and curricular content, the researchers broke the students into two control groups and rotated students between technology-based embodiment learning environment at the SMALLab, and regular instruction in a classroom. In both settings, the study focused on teaching titration (adding unknown solution of acid or base to a known reactant of unknown molarity until the endpoint of reaction occurs), which is part of any traditional chemistry curriculum.

The SMALLab is an educational platform that engages the major sense modalities (i.e. visual, auditory, and kinesthetic) that humans use to learn. It uses infrared motion tracking cameras to send information to computers about where a student is in a floor-projected environment. Students step into the active, defined space and grab a wand (a trackable object) that allows the physical body to now function like a 3D cursor in the interactive space. The space allows for multiple students (up to 4) to be tracked simultaneously. Teachers use wireless peripheral devices (remote controls and game remotes) to control the flow of the dynamic visual, textual, physical and sonic media that students interact with.

Not surprisingly, the study found that placement in the embodied environment consistently led to greater learning gains (effect sizes from .53 to 1.93). The researchers believe the findings are due to three primary features: embodied, kinesthetic learning, the high degree of collaboration designed into the learning platform, and the "novelty effect" of the environmental experience.

The study professes that internal knowledge structures that are gathered in an active multi-modal and kinesthetic manner are learned faster, and that first degree embodied learning is driven by locomotion (increased sensorimotor activation and its varying environmental affordances), perception of immersion, and tangible object manipulation with content-relevant gestures.

Mina Johnson-Glenberg, David Birchfield, Lisa Toledino and Tatyana Koziupa conclude that:
If the cognitive primitives of understanding are embodied, then learning environments that highlight the physical grounded knowledge, and that induce mental simulations of abstract constructs like chemical reactions...may be more powerful learning methods than techniques emphasizing observation and symbol memorization.” 

While such findings are by no means conclusive, and barely scratch the scientific surface, these strategies offer inspiring possibilities for the integration of device-free kinesthetic engagement.

5. Conclusion and Future Work
Multiple research areas now support the tenet that embodiment is a powerful underpinning of cognition. To a lesser extent, some disciplines accept cognitive extension as an active, unhierarchical negotiation between the body, mind and environment that equally optimizes cognition. But when combined with serious game mechanics, which as D.W. Shaffer contends “have tremendous educative power to integrate thinking, social interaction and technology into the learning experience,” opportunities for increasing impact grow exponentially.

Both persuasive, cultural strategies and emerging, embodied technology possess the ability to deeply encode us with unseen moral, social and political structures. Serious game design, which harnesses commercial media mechanics and narrative devices, and applies a deep knowledge of cognitive neuroscience to natural user interfaces can enhance cognitive and affective systems. We argue that such game-based experiences carry the potential to shape and powerfully transform not only individual attitudes and behaviors, but the underlying values that govern our society.

Because of the Kinect’s mainstream popularity, ease of use and affordability, we believe it will not only change who can play games, but will also fundamentally alter the way we engage with games, how we will make games in the future, and the ability of games to effect long term learning, behavior and emotional outcomes.

To further explore how the affordances of kinesthesia, vocal recognition and full body immersion, provided by the Kinect can enhance serious games, we have several projects in production that we are in the process of carrying out over the next two years (of which we may show works in progress), which focus on: 1) increasing conceptual retention when teaching logic, 2) instigating behavioral change in ex-offenders in an attempt to reduce recidivism, and 3) improving social-emotional competence when transitioning 3-7 year old children from the home to the school.
