

Smart 3D collaborative virtual learning environments: A preliminary framework

James Laffey*, Matthew Schmidt, Krista Galyen and Janine Stichter
University of Missouri, 303a Townsend Hall, Columbia, MO 65211, USA

Abstract. With increases in access to powerful computing and high-speed networks, 3D virtual learning environments are being envisioned and developed as places for collaborative learning. These new environments for collaborative learning have promise for great authenticity in experience, great presence with others, and the monitoring/sensing of broad ranges of human cognition and behavior inferred from the actions of avatars. In the process of designing and building such an environment, iSocial, to develop social competency for youth with Autism Spectrum Disorder the authors have explored the potential for developing a smart system. The paper provides a framework for conceptualizing and implementing a smart 3D collaborative virtual learning environments based on 3 key constructs: environmental scaffolds, social affordances and coaching. The framework and constructs are illustrated using experiences and functionality developed for the iSocial system.

Keywords: 3D-VLE, 3D-CVLE, smart-systems, scaffolding, social-affordance, coaching

1. Introduction

Developing interactive learning environments that exhibit intelligent behavior so as to adapt to the specific needs of an individual learner has long been a goal of learning science and computer science. Computer-aided instruction (CAI) became available in the early 1960s; however, a weakness of CAI was its one-size-fits-all approach to presenting material and its inability to adapt to differences between learners [1]. To address these weaknesses, researchers and developers strived to create more intelligent computer aided instruction (ICAI), but early attempts fell short due to the high computing costs involved and the perceived impracticality of incorporating artificial intelligence (AI) into educational software. As computing became more powerful and accessible at the beginning of the 1980s, research again turned to AI as a means to create software models of human tutors [1]. The results of these efforts were intelligent tutoring systems (ITS).

Intelligent tutoring systems have shown success in a variety of learning domains, particularly mathematics, science and technology. Examples of early

successes include the GIL system for learning the LISP programming language [2], the Geometry Tutor [3] and STyLE-OLM, a system for learning scientific and technical terminology [4]. However, the success of these systems can be partially attributed to their limited scope, as they dealt with finite rule sets, relatively narrow problem spaces and well-structured problems. While these systems showed some advantages over traditional teaching modalities, they ultimately fell short of the goal of emulating a human tutor, both because humans and computers operate in different ways and because the problem solving models employed by these systems may not have been sophisticated enough to teach broad, ill-structured domains [5].

Advances in artificial intelligence have paved the way for newer systems to become more sophisticated and to address the weaknesses of prior systems. Examples include the ANDES system for physics [6,7], AutoTutor for networking and computing [8] and Why2-ATLAS [9] for qualitative physics essay writing. While these systems incorporate state-of-the-art AI technologies, natural language processing and statistical techniques to approach the goal of adaptive and responsive attunement to individual learners, they are still beset by a number of challenges. The difficulty of providing explicit representations of

* Corresponding author. E-mail: laffeyj@missouri.edu.

domains increases as the domains and associated contexts become more complex. As problems become more ill-structured, the same is true. In addition, learners are able to take advantage of the brittle nature of these systems to “game” them, that is, to exploit the predictable regularities of the system to complete tasks, as opposed to meaningfully engaging in learning [10,11]. At this point, creating models of known, low-level information and decisions in narrowly scoped domains is relatively easy, whereas modeling the complex interactions between content, instruction, cognitive processing and problem-solving or meaning-making is incredibly difficult [12].

However, as educators start to explore and envision the potential of collaborative 3D virtual environments, now made possible by high-speed network connections and high-end processing capabilities, it is worth re-examining what can be meant by intelligent systems and to re-think the “tutoring” metaphor so pervasive in many of these prior systems. While the tutoring metaphor tends to consider intelligence as something shared between the computer and a single learner, collaborative virtual environments allow for intelligence to be shared not only between learner and system, but also within learner groups and between various objects, actors and actions. Such a conceptualization raises a number of questions. What does it mean for a 3D collaborative virtual learning environment (3D CVLE) to exhibit intelligent behavior? What does it mean for a learning system to be intelligent when learning happens in interactions with varied objects in an environment and through interactions with others, so that the learning is both experiential and collaborative?

The purpose of this paper is to identify and elaborate a framework for intelligence in 3D CVLEs with a vision towards both the environments and the participants exhibiting intelligent behavior within those environments. By intelligent behavior we mean participants adapting themselves to the learning needs and opportunities of the specific context/environment. These adaptations can include collaborative learning with peers, instruction from a teacher, and utilization of objects in the world to invite or constrain behavior. By intelligent behavior by the system we mean the environment adapting and being optimized for the needs of learners and the learning processes. These adaptations can include managing access and interaction with others and having objects transform based on knowledge of behavior patterns, learner models and learning objectives. Thus, we envision environments and associated functions which adapt and can

be adapted through intervention to the needs of learners, and which invite intelligent behavior or constrain unintelligent behavior on the part of learners.

This paper presents a framework for thinking about how 3D CVLE systems can be intelligent across a range of interaction modalities within CVLEs and describes our work to uncover and advance this framework. Our interest and initial steps toward implementing aspects of this framework comes from our research and development of iSocial, a 3D CVLE to help youth with Autism Spectrum Disorders (ASD) develop social competence. Children identified with ASD have deficits in social competence that can lead to problematic social behavior and social isolation [13], which in turn can lead to a lower quality of life, as well as deficits in other developmental areas such as language and cognition. iSocial is a 3D CVLE being built using the Open Wonderland (<http://openwonderland.org>) virtual worlds toolkit, and seeks to use networked-based 3D VLE technology to make a successful face-to-face curriculum [14] typically administered in a clinic-like setting on social competence development available to youth with ASD. Without this transformation to virtual learning most youth who need help with social competence will have no access to the specific lessons provided by instructors who are experts in these approaches. This curriculum includes units on recognizing facial expressions, sharing ideas, turn taking, recognizing emotions and problem solving. The learning objectives within these curriculum areas are both complex and ill-structured. Upon completion, iSocial will include in excess of 30 lessons and approximately 22 hours of instruction. The format of instruction is for groups of four to six youth (aged 11 to 14) to meet online in iSocial with an online guide (OG), who is a specialist in the curriculum and in working with youth with ASD, leading the students through the lesson activities.

To be clear, we do not propose an extension of the rule or neural-net based models that form the basis for tutorial-type learning systems or expert decision-making. Rather than propose a mechanism that makes a system intelligent or in its absence does not represent intelligence, we envision approaches that make a system more intelligent or in their absence less intelligent. In fact we prefer the term “smart system” because while such a system certainly can include the mechanisms others characterize as intelligent systems (*viz.* [6–9]) we mean more of an ecosystem of design and functionality that contributes to

smartness in 3D CVLEs, rather than a single form of intelligent behavior.

How should we think about being smart and where does the smartness reside in 3D CVLEs? Nearly 50 years ago Douglas Engelbart provided a vision for how computing could augment the human intellect. By augmenting he meant “increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular needs, and to derive solutions to the problem” [15]. Engelbart and colleagues provided a pioneering vision for how innovations like the mouse, screen sharing, hypertext and dynamic file linking could support having a computer that was responsive to human needs and worked the way humans worked. Of course, today we need to keep in mind that these innovations were being shown in a context where nearly all human-computer interactions were by professionals and where nearly all computer professionals worked alone and worked by flipping switches, entering commands through a card reader or using a command line on a console. Engelbart’s work preceded and anticipated the ARPA networks (early packet switching networks that stand as fore-runners to the Internet), and he speculated that the networks would allow users to query others on the network for knowledge about and access to services, thus augmenting their own capabilities and intellect. Following a similar interest in having technology support human activity by fitting with the way humans work, Donald Norman’s book, *Things that Make Us Smart*, [16] argues that technology does not simply improve the way we do things, but actually changes what we do. Multiplying with a calculator is a different mental and physical task than is multiplying with paper and pencil. For Norman, computer functionality, as well as the functionality of other devices in our world, needed to be designed to invite behaviors that led to good (smart) outcomes, such as doors that are intended to be pushed not having a handle that invites you to pull them or teapots or pans with handles designed so that the handle does not get as hot as the cooking region of the device. In this way, people will not look or feel dumb by pulling on a door that should be pushed or picking up and dropping a hot pot.

Mark Weiser [17], a former chief scientist at Xerox Parc and largely credited with initiating a vision for ubiquitous computing, had similar beliefs about how computing could augment the human intellect and envisioned ubiquitous computing as enabling a smart environment. A smart environment is a place where people work, learn, play or reside that is augmented by computational resources which provide

information and services when and where desired. Weiser envisioned technologies in the service of natural human and human-to-human action. Abowd and Mynatt [18] further explicate the concept of a smart environment by identifying the need for sensors so the technology can react to humans, for communication from the technology to the human and for mechanisms that integrate between the digital and physical worlds so that the sensing and communicating will be seamless to the human activity. This vision of a smart environment includes functions like sensing human presence, so that the human-computer interaction does not require humans to directly act “on” computers, such as typing on a keyboard, but rather act “with” computers by having the computer monitor natural human activity and act with it. Weiser’s view of ubiquitous computing and a smart environment is human-centric in the sense that it is meant to augment human activity. The smartness resides in the system of sensors dedicated to observing human behavior in natural contexts and communication and services in support of those behaviors. However, the smartness also resides in the design of the human spaces to seamlessly integrate the sensors, feedback and support for human behavior.

This intellectual stream of Engelbart, Norman, Weiser and others has helped frame a view of smartness as distributed between the brain and its external resources. This sense of intelligent behavior as human cognition, moderated and shaped by external resources, makes sense as a way to also think about intelligent behavior of a 3D CVLE. Adding to our perspective on intelligent behavior is the recognition that much of the context into which human cognition is distributed is social. The work of Lave and Wenger [19] provided leadership in showing how professions and disciplines have communities of practice wherein learning takes place through experience and induction takes place through apprenticeship. They argue that learning is not a type of activity, but rather is an aspect of all activity. Wertsch [20] and others taking a socio-cultural approach have shown that the “intelligence” of actions is only meaningfully understood in the context of knowledge about the cultural tools invoked. These works have helped frame a view of cognition as distributed and a stance toward facilitating learning that calls for situated and social practice. For example, Collins, Brown, and Newman [21] showed how novice learners could perform more expertly when scaffolded by having a more experienced performer complete part of a complex task for which a learner is unprepared, thereby allowing the learner to engage in work that would normally be

outside his/her grasp. Socio-behavioral and socio-cultural perspectives show us that what is learned and the quality of performances are influenced by modeling and social interaction.

These seminal and various views of human computer interaction and how computers augment human capabilities provide a basis for us to think about how to develop smart 3D CVLEs. 3D CVLEs potentially afford great authenticity in experience, great presence with others, and the monitoring/sensing of broad ranges of human cognition and behavior inferred from the actions of avatars. How can we best think about unleashing the potential for smart 3D CVLEs? The following section provides a framework for conceptualizing smart 3D CVLEs, accompanied by examples in our work of how we are trying to leverage this potential. We follow up this section with a more detailed look at one aspect of the framework, which is concerned with having the environment adapt to participants. We then conclude with a discussion of key challenges ahead in our efforts to develop a smart 3D CVLE and recommendations for future research.

2. Pedagogy as a context for smart 3D CVLEs

As is described in the introduction, we are interested in a smart system focused on human learning, which, in practice, may have some distinctions from simply smart systems for human behavior. A smart human learning system operates within a pedagogy that implies appropriate or inappropriate choices for action. By appropriate action we mean an action that based on the pedagogy has promise for leading to desired learning outcomes. For example, when a student asks a teacher a question an intelligent system that knows the answer will probably answer the question. However, when a system is driven by a pedagogy the intelligent system might ask the student a question in response to the student's question as a means of drawing out the students available knowledge. Different pedagogies will call for different choices. For example, a cognitive behavioral pedagogy, as is implemented in iSocial, is highly instructive in making connections between cognitions and behaviors, whereas a discovery learning pedagogy may allow for more exploratory behavior and self-assessments on the part of the learner. Smart actions of the learners and the system are in tune with the pedagogical basis for the curriculum. For example, given the highly instructive nature of the cognitive

behavioral approach it is important to complete activities in a timely manner and in the prescriptive ways called for in the curriculum. Thus, smart behavior in our curriculum may be seen as having students ready themselves for instruction whereas in a more discovery-oriented curriculum, having students express individual interests based on curiosities and take personal responsibility for making decisions about what to do next may be more appropriate. In this paper we do not address issues based on differing pedagogies but want to clarify that the definition of "smart" is aligned with behavior that is appropriate in the given curriculum. For example, given the challenges of fitting learning activities into a fixed time period for a lesson, a number of approaches to making the system smart would address having students be ready for instruction and be efficient in various learning activities.

Similarly, having a smart learning system means being adaptive to the personal capabilities and readiness of the student. For example, asking a student who is just beginning to learn addition to solve a polynomial would not seem smart on the part of the system nor allow the student to seem capable in mathematics. The pedagogical aspects of our smart framework include a three-phase model of capability [22] including (a) acquisition, (b) maintenance and (c) fluency. Students in an acquisition phase require more specific, timely and directive shaping of behavior. Students in the maintenance phase have acquired a rudimentary ability so it is appropriate to moderate the specificity, timeliness and directedness of prompts to allow the student more responsibility in their learning. Students who have achieved fluency may need some assistance if they have repeated errors, but the prompts and shaping should be as natural to the performance as possible. The support, prompts and scaffolding fade across the phases: being heavy yet tolerant during the acquisition phase, moderate during the maintenance phase and light during the fluency phase.

In addition, we generally think of intelligent systems as being adaptive to individual learning needs, but this aspect of the framework may need to be expanded to also include being responsive to the needs of the collaborative learning group. In a collaborative learning context the intelligent system may serve to help manage interactions among students and support learning from each other (as well as supporting learning from the instructor). For example, an intelligent system may know that some members are in the fluency phase and others are in maintenance or acquisi-

tion phases and use this knowledge to shape the interaction among the students.

Thus, we view pedagogy to be an essential guide to how a framework for smart 3D CVLEs can be used to make design decisions. The choice of pedagogy such as cognitive-behavioral versus discovery impacts the objectives for what students will do and how they will work in the system. Understanding individual differences among students, such as the phases of acquisition, maintenance or fluency, may be important for customizing the system to each learner. Understanding the makeup and characteristics of the group, such as the mix of phases that the members have attained or the past history of members with other members, may also suggest ways for the system to optimize learning in a group.

3. Framework for smart 3D CVLEs

In order to guide the design and development of a smart CVLE, we developed a framework for supporting not just intelligence, but an ecosystem of “smartness” to facilitate teaching and learning in such an environment. Without a framework, it is easy to lose sight of the big picture of creating the ecosystem for collaborative teaching and learning, and instead focus on mechanistic intelligence for isolated functionality. The objective of this framework is to support and guide creating and understanding 3D CVLEs so as to bring smartness to bear in support of better teaching and learning in 3D CVLEs.

We view our understanding and ability to articulate a framework for smart 3D CVLEs as a work in progress. Fortunately, we are immersed in the development of a 3D CVLE for students who need a virtual and collaborative learning environment and for whom the learning requires careful and sustained instruction, thus providing us with high motivation for continuing to conceptualize and try out ways of making the CVLE smarter. We also see this framework as applicable to not just our 3D CVLE, but any 3D CVLE; as such we believe this framework may be of great use to those who use 3D CVLEs. We base our framework on three forms of managing learning activity: 1) environmental scaffolding, 2) social affordance, and 3) coaching. Each section will first describe the framework element, and then further explicate it by describing how we have implemented that framework element in iSocial.

3.1. Environmental scaffolding

Environmental scaffolding refers to structures in the environment that shape behavior by inviting appropriate behavior or constraining inappropriate behavior. Just as a sidewalk invites people to walk a specified path rather than through the flowerbed, both the design of the world as well as structures that exhibit agent-like functionalities can serve to invite and constrain behaviors. While our following examples are specific to iSocial and are often social in nature, we see environmental scaffolding as applicable to any 3D CVLE design, no matter the user age or learning needs.

Implementation in iSocial

Based on our early field tests of a unit in iSocial, we identified a need to support core aspects of social engagement and interaction. By core aspects we mean fundamentals of social interaction such as, orienting to others when spoken to, transitioning from one task to another, not being easily distracted while in a social interaction, not interrupting others during conversations, etc. Since a key challenge of iSocial is to assist youth with ASD, who have social performance deficiencies, to be social while learning social performance competencies, we identified a special category of scaffolding as *social orthotics* [22] to represent types of structures that might be needed to facilitate social interaction and social learning in iSocial.

While no exact distinction can be drawn between invitations and constraints, these heuristics provide us with two strategies for thinking about shaping behavior. For example, in early versions of our iSocial environment we observed numerous incidences of the students not understanding where to go for a next activity and having problems with navigation such as getting their avatar stuck in walls. A solution we developed using the “invitation” approach was to redesign the learning environment to be more open such as in a garden space where it was easy to see other spaces and pathways for navigation. In this sense, the enhanced design makes the new destination more visible and removes barriers or occasions for error during navigation.

Another example of invitation is the use of personal pods and group spaces as different indicators of what behaviors are appropriate and inappropriate at a given time. For example, during some forms of instruction, students use “personal pods” to assist them

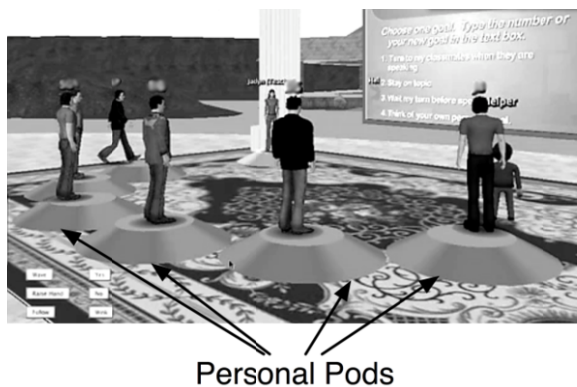


Fig. 1. Pods invite students to park their avatars in a socially appropriate way.

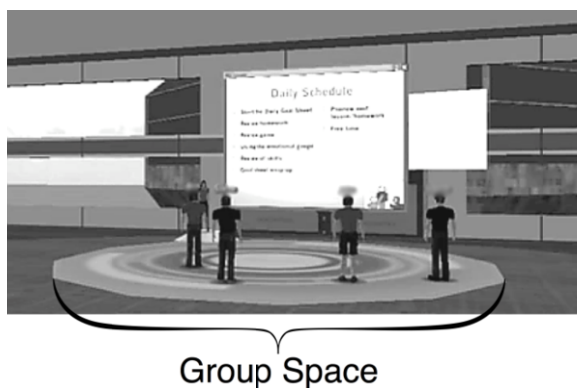


Fig. 2. Group spaces indicate areas in which users may gather for activities.

with orientation to their peers and to instructional materials (see Fig. 1). The placement of the pods in a semicircular fashion invite social behavior related to the curricular content. The pods use color-coding to signify when it is appropriate to enter (green) or when the pods are locked and exiting is not possible (red). The locking of pods or other learning spaces once the youth were in place for an activity restrains the mobility of their avatars and represents a form of the constraint strategy. Figure 2 shows the use of a group space. The group space indicates the area in which it is acceptable to gather, akin to the bounds within a classroom such as a large table or an activity area, and invites users to gather within a certain proximity of each other when there are no pod indicators for where to stand.

While the personal pods and group space indicators were used as invitational forms of inviting grouping behavior, some constraints can be used to constrain users from straying from the group. For instance, a barrier could be used in front of a portal, allowing the online guide or instructor to ensure that

all group members go through the portal at the same time, as having all users present and with each other at all times is an important aspect of the iSocial curriculum.

Another example of a constraint is the instructor's ability to turn off the youth's microphone if other forms of intervention do not work in regulating excessive talk.

Our primary focus in developing social orthotics, based on strategies of invitation or constraint, was to assist the youth in being social and to support the online human guide, whose role it was to manage youth behavior and facilitate learning in the 3D VLE. The social orthotics are designed elements of the environment to invite desired behavior and constrain undesired behavior. To a great extent our application of this process in iSocial addresses social behavior or behavior that orients youth to instruction. Our emphasis on these targets for using scaffolding follows from curriculum objectives and the specific needs of youth with ASD.

We see others using environmental scaffolding, taking their pedagogical context into consideration, as possible important and useful means for supporting users in teaching and learning through inviting and constraining user behavior.

3.2. Social affordance

Social affordance refers to social elements in the environment that shape behavior by inviting appropriate behavior or constraining inappropriate behavior. We have conceptualized two forms of social affordance in the framework. The first form is the social benefits of learning directly with and from others on curriculum activities. In this sense, the students benefit from collaboration such as in needing to explain themselves to others, division of labor in completing tasks and learning from one another as in having a more capable student help a less capable student get started on an activity. The first form also includes having a teacher or guide directly instruct students. The second form of social affordance is the indirect benefit of learning with others, such as comes from learning from modeling or using social navigation to help make decisions about how to proceed. A key aspect of this second form of social affordance is the sense of co-presence with others in the 3D VLE.

Implementation in iSocial

The primary form of direct social learning comes from the instruction provided by the teacher, who we



Fig. 3. Students choose an area to live on a desert island during a “lost at sea” activity.

refer to as the “online guide” (OG). The OG uses audio, gestures and modeling to guide the group through activities where they are learning about a new skill, practicing the skill in structured contexts and then practicing the skill in more naturalistic contexts. The OG serves as a master of social competency and the students have the role of apprentice [21] as they move through the levels of activity. Having a dialog and cooperating with others in the lessons also provides support for acquisition of the new skills. In essence, this direct form of social learning is the implementation of the curriculum and the design challenge is to take the lesson activities that work effectively in face-to-face sessions and translate the cognitive-behavioral requirements into virtual tasks in virtual contexts enacted through avatars.

The design and management of affordances for indirect social learning and co-presence can lead to more appropriate and less inappropriate behavior during and between the lessons. Students learn through modeling by being able to watch the avatars of their peers take on tasks, receive feedback and make corrective adjustments. A key objective of building a smart 3D CVLE is to make the learning “visible” in the sense that behaviors are overt and cognitions are reified through tasks and the use of objects in the environment. For example when students are asked to make a choice, their avatars carry the choice out such as in discussing and selecting items that will be needed for survival on a desert island (see Fig. 3) or when asked to identify forms of interaction the students play a game where labels for forms of interaction are matched to descriptions of

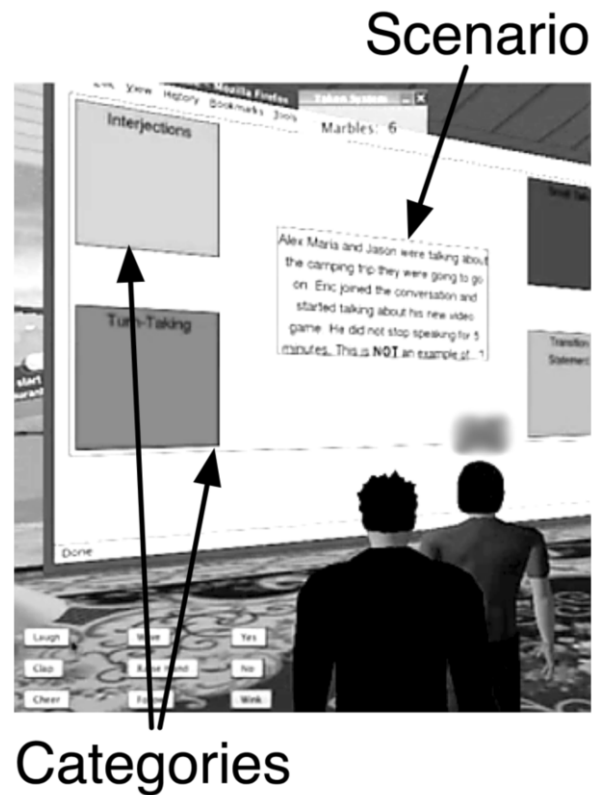


Fig. 4. Students play a game in which they match scenarios with categories in a shared application window.

interactions (see Fig. 4). Thus students are not left wondering if they have understood the concept or skill, but instead they have enacted the knowledge and have a concrete instance of their own behavior and the behavior of others to reference.

There is still much to be learned about being social and social learning in a 3D CVLE. A key direction of our research and development is to innovate in ways that help the OG monitor both the avatar and the physical behavior of the youth. One development that we are testing is to have a snapshot of the youths’ faces every 20 seconds display on the OG’s screen. In this way the OG, if she chooses, can use facial expressions along with vocalizations, voice inflection and avatar movement to be smarter and make better assessments of the involvement and progress on the part of the youth. A second key direction for us is to impact the sense of co-presence that the youth have with the instructor and their peers. For example we notice that gestures are a valuable and valued part of the youth’s repertoire. The youth like to use gestures and movement and have a sense of physicality in the

environment. We hypothesize that giving youth more gestures and movements will serve as a broader vocabulary that resonates with learning in a 3D CVLE and improve their sense of being in the world with others. Theoretically, increasing their vocabulary and sense of co-presence in ways that are in tune with the pedagogy and environment will make them smarter in how they make sense of the learning opportunities, experience the social nature of learning and communicate their achievement.

3.3. Coaching

Coaching refers to dynamic interaction of the 3D CVLE components with individuals and groups in the environment. We envision two forms of coaching in our framework. The first form is the human coaching that an instructor or guide can provide to the students. The second form is agent-like behavior that can be designed into elements of the collaborative 3D VLE.

Implementation in iSocial

The cognitive-behavioral pedagogy implemented in iSocial requires close management and active shaping of the youth activity through human coaching. In a face-to-face session the teacher has many strategies for gaining attention, shaping behavior and intervening when there is inappropriate behavior. In early pilots of iSocial we saw that these strategies fell apart or were hard to implement in a 3D CVLE and that the OG often struggled to fulfill the curriculum in the time allocated. To a great extent the efforts described in the environmental scaffolding and social affordances sections have tried to improve the ability of the OG to guide the youth. Of course, a key part of the OG guidance is through her verbal instructions and remonstrations. In the face-to-face curriculum, the guide uses physical rewards and “strikes” to communicate to individuals and the group that they were on track and doing well or needed to stop an inappropriate behavior. The physical and visible progress cues reminded students and helped them self-regulate as they strove to earn privileges or avoid restrictions.

We sought to mimic this approach in iSocial by providing a score board with a token meter and a behavior meter to provide visible evidence when feedback was given and to provide a summary of the member and group status towards desired outcomes (Fig. 5). The scoreboard eased the work of the OG by providing a visible and convenient way of communicating positive and negative feedback and assisted

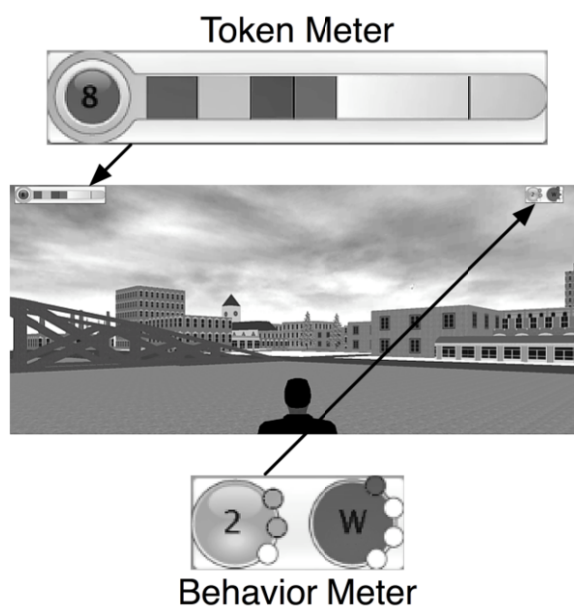


Fig. 5. The scoreboard consists of a token meter (top left) and a behavior meter (top right).

the students in regulating their own behavior. The scoreboard represents one example of how the environment can be designed to facilitate human coaching.

Agent-like behavior in learning environments is often equated with anthropomorphic agents, sometimes referred to as non-player characters (NPCs), cohabitating the learning environment. While we envision a number of possibilities for anthropomorphic agents in iSocial, such as behavioral modeling or having a buddy who keeps track of one’s status and direction, our current research does not focus on NPCs. Rather, we are focused on non-anthropomorphic agent functionality, such as the mechanisms for hints and alerts described in Section 4. Nonetheless, NPCs can play powerful roles in modeling, controlling and communicating pedagogy, and these are characteristics in which we are interested.

Looking to Kim and Baylor’s [23] social-cognitive framework for how pedagogical agents can serve as learning companions (PALS), it is evident that such agents can be used 1) to provoke cognitive conflict, such as bringing up new ideas, 2) to provide scaffolding to extend a learner’s capabilities, such as making a recommendation, and 3) to provide empathic support, such as sharing good emotions or providing comfort after errors. The use of PALS can also be distinguished by who controls the interaction. Does interaction occur when the user requests it (respon-

sive) or when the system identifies a need or opportunity (proactive)? While Kim and Baylor are optimistic about the potential of PALS, they also identify a number of challenges that impede and limit the usefulness of PALS in current learning environments, such as the difficulties in imbuing PALS with intelligence that can be viewed as natural by the students and context specific for the learning objectives. Kim and Baylor's framework informs our own agent design in that we also are interested in the potential pedagogical benefits of agents (although ours are not anthropomorphic) and we are challenged in representing agent feedback as natural and contextually appropriate. The next section provides detail about our design work and exploration of agent-like behavior.

4. Agent-like behavior in smart 3D CVLE

4.1. Agent role

An agent is an entity that acts on behalf of another entity. The term can be ambiguous because it is general and used in everyday language (e.g., real estate agent, travel agent), and even in constrained domains like software engineering, "agent" is an umbrella term [24]. An agent exhibits behavior, and behavior is what differentiates one agent from another.

We anticipate there will be ample opportunity for many forms of agent-like behavior in smart 3D CVLEs: the needs and opportunities are nearly endless. Our focus for a first step in this domain was to integrate pedagogy and concerns for how best to support the human-computer interaction of youth to agent communication. In essence, we were interested in the role of the agent. To this end, we undertook the task of operationally defining the role of our agent by 1) defining the inventory of user actions and behaviors that we wished to affect, 2) identifying the contexts in which those actions and behaviors occur and 3) explicitly defining our target demographic in order to create a preliminary student model. It is the interrelation of these three components that brings the role of the agent into focus, as illustrated in the Venn diagram in Fig. 6.

We started by establishing how the agent might act on behalf of the students or online guide, envisioning the agent's role as one that would fit and reinforce the social competence curriculum the youth were undertaking. For this purpose, it was necessary to

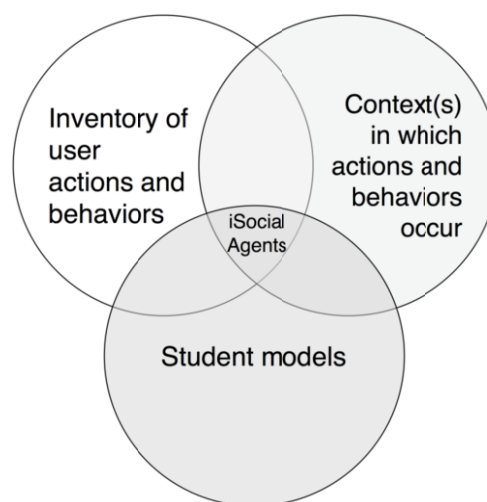


Fig. 6. The role of the agent illustrated as the interrelation between actions and behaviors, context and student models.

identify the actions and behaviors we wished to affect using the agent. We chose the behaviors of orienting one's avatar, distance between avatars and adjacency to others' avatars as a starting point for our exploration. Orientation is the direction an avatar is facing. Distance is how far the avatar is from a speaker or group. And adjacency is how close the avatar is from a speaker or a group.

The context within which the students operate is a key consideration when defining the role of a given agent, as this is the context within which the agent must also necessarily operate. We identified the context as being within lesson activities which are instructor-centered and utilize a question-and-answer format, typically observed during the introduction to a new lesson and while performing wrap-up and review at the end of a lesson. In this context, the agent is triggered by movements of the youth's avatar that are deemed inappropriate for the context, such as not keeping a proper distance between other participants' avatars or exiting a personal pod before the lesson activity is complete.

The final component of defining the agent was to develop a model of our target demographic. This was particularly important since the agent would be required to meet the special needs of an exceptional population of learners. While focusing on the individual student with ASD may have been beneficial due to vast differences in terms of learning styles, individual social deficits and needed supports of the target demographic, the complexities of such an approach made it impractical. Instead, we decided to

initially focus on the broader characteristics of 1) age (between 11–14 years old), 2) having a medical diagnosis of autism determined by the Autism Diagnostic Interview Revised (ADI-R) and/or the Autism Diagnostic Observation Schedule (ADOS), 3) being capable of speech and able to read and comprehend written text and 4) an intelligence quotient within one standard deviation of the mean for the typical range (e.g., a score of 85–115). These learner characteristics form a model of our students, which we used to customize how the agent functions.

While this student model has limitations given that it is so broad, we envision incorporating in future development a student model characteristic for phase of learning (acquisition, maintenance or fluency). So, for example, the agent would perform actions differently for students who have achieved fluency (for example only providing a simple alert), as opposed to a student still in the acquisition phase who might receive more directive hints and instruction.

4.2. Agent functionality

Using the three components described in the previous section to define an agent's role, the agent does not need to constantly monitor all contexts at all times, but instead, only needs to come into play when students enter the context in which the agent operates. Actions and behaviors are constrained to those actions and behaviors that happen within the context, thus allowing the agent to act within a smaller purview. While this constrains the capabilities of the agent, it also allows the agent to be more fine-tuned to the activities for which it is built. The rule sets that it follows to determine a course of action are very specific and purposeful, and are informed by user models. Because the iSocial curriculum follows a pre-defined and specific structure with regimented timings, defining contexts and actions is less difficult than might be the case with less structured pedagogy. Figure 7 provides a general illustration of how the agent functions.

This general description provides a framework for how an agent will function in any specific case, such as the regulation of orientation, adjacency and distance. Consider a student entering a virtual classroom and preparing for introductory activities. That student approaches a personal pod and parks her avatar there. This personal pod is located within the introduction activity context, and the student entrance into the pod triggers the agent to begin monitoring the student's behavior for pre-defined signals, such as moving

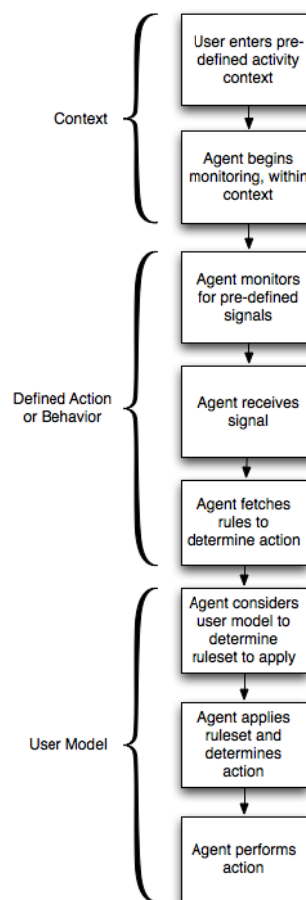



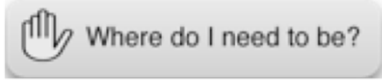
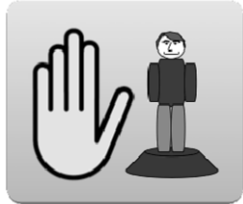
Fig. 7. How an agent functions.

away from the group or not orienting to the speaker or the group. When the student attempts to exit her pod before the introduction activity is completed, the agent fetches a notification of the inappropriate behavior that is worded in a way that takes the student model into account. The agent then acts by sending the student the notification.

4.3. Agent implementation

To investigate our agent design framework, we conducted a three-tier usability study with subject matter experts and youth in which we specifically focused on the type and nature of notifications that the agent would deliver to students. In keeping with the cognitive-behavioral pedagogy and considering our student model, we developed three levels of cognitive hints (notifications) for orientation actions

Table 1
Hints given to a student attempting to leave a personal pod

“Helpful Hint” Level	Design Type 1: Thought-bubble cognitive modeling	Design Type 2: Command-centered notifications	Design Type 3: Visual-only alerts
Thumbnail			
Hint #1 (least assistance)	I am a listener	Where do I need to be?	[Visual of avatar on pod with white indicators]
Hint #2 (moderate assistance)	I need good body control	Have a calm body	[Visual of avatar on pod with orange indicators]
Hint #3 (most assistance)	I need to have a calm body I don't want to bother others	Stay on your pod. Don't bother others	[Visual of avatar on pod with red indicators]

based on the system of least prompts [25,26]. The system of least prompts provides a structured framework to deliver coaching, predominantly seen in the special education domain, wherein the “coach” provides a student prompts or support, commonly over 3 successive hints. These prompts begin by providing the least assistance possible (least intrusive) and progress to providing the most assistance possible (most intrusive). The prompting then ceases when the student demonstrates the desired behavior.

A face-to-face example would be a teacher telling the class to find their seats. One student is still standing so the teacher begins the prompting with “Where should you be right now?” After not responding, she points to the chair and asks the student to sit. After still not complying she may go over and physically assist the child to sit. These commonly go from verbal, then visual and final physical cues. However, given our cognitive behavioral pedagogical framework and the fact that the experience will be in a 3D CVLE (where the “coach” cannot always point, gesture, or model in ways that we can in the real world nor physically make an avatar do something), our conceptualization of the system of least prompts in iSocial were verbally going from general to specific, as can be seen in Table 1.

To improve our understanding of how best to implement this three-hint model of coaching, we designed three possible designs for delivering the coaching feedback for a student trying to behave inappropriately in the environment. Table 2 details the

type of feedback hints for a student attempting to leave a personal pod. The hints follow the system of least prompts model, ranging from a general alert that the student should attempt to recall how to behave appropriately, to a second prompt providing more specific information on what they should be doing, to a final, highly specific notification about what they are doing wrong and exactly how they should correct their behavior. We explored the possibility of three different designs for delivering the hint. Those were: 1) a thought bubble that would appear above the avatar where the wording was based on a cognitive apprentice model, suggesting what the avatar should be thinking at the time, 2) a command-centered notification that would appear next to the avatar and suggest what the avatar should be doing at the time, and 3) a simple symbolic, visual-only cue that represented what the avatar should be doing.

Our usability study [27] included a first round of cognitive walkthroughs in a prototype environment with two experts in special education and two experts in learning technologies, a second round of walkthroughs with two experts in special education and one expert in usability studies and a third round of walkthroughs with four youth in the target age range. The walkthroughs included four design formats (no agent, thought bubble, command-centered notification, and visual-only cue) and four scenarios for each design. Table 2 summarizes the scenarios and illustrates how they are implemented for the thought bubble format. Figure 8 provides an example

Table 2
Scenario descriptions and hints given in the cognitive thought bubble design

Description	Hints Given
<p>Part A: Students are in locked pods. The OG is instructing the students. Tommy tries to run out of his pod three different times.</p>	<p>Hint Level 1 (body movement): "I am a listener." Hint Level 2 (body movement): "I need good body control." Hint Level 3 (body movement): "I need to have a calm body. I might be bothering others."</p>
<p>Part B: Students begin in a large group space where they are supposed to raise their hands before speaking. Tommy misses the mark here and speaks without raising his hand. The students then move to the "small group" spaces where they are to share with each other. Tommy runs around the group space and never shares with his group.</p>	<p>Hint Level 1 (body movement in group space): "I am a listener." Hint Level 1 (not sharing with others during share time): "I am a speaker."</p>
<p>Part C: Students surround a table with food. They are to share their ideas regarding which food they like best. Tommy does not share his ideas with the rest of the group. Tommy then runs around rather than staying with the group. After a while longer, Tommy still has not shared his ideas with the group.</p>	<p>Hint Level 2 (not sharing with others during share time): "I need to share my ideas." Hint Level 2 (body movement in group space): "I need to use good body proximity." Hint Level 3 (not sharing with others during share time): "I need to get people's attention and share my ideas with others."</p>
<p>Part D: Students surround a table and chairs. They are to share what they like about the particular table and chairs. Tommy participates, doesn't move around, and shares his ideas with others.</p>	<p><i>No hints are given as Tommy is demonstrating desirable behaviors.</i></p>



Fig. 8. Students stand on personal pods within a group space, typical of online guide-led discussions in iSocial.

of avatars on pods in a group learning space. This scene illustrates several of the issues in having a smart environment for collaborative learning. One issue is that the instructor needs to gain student attention and have them focus on an object in the world related to their learning. In this case the object is the

whiteboard, which shows slides to facilitate a discussion. Another issue is that while the students need to see both the instructor and the whiteboard for the instruction, they also need to see each other in order to discuss and have a social experience. This area of the iSocial environment is designated as a learning

space so the instructor can keep the students in proximity to herself, the other students and the instructional materials. The area also has learning pods so the students can be assisted or forced into an appropriate orientation and distance to one another while also maintaining an optimal view of the key components of the learning space.

The walkthroughs always started with the “no agent” format to teach the participants about the iSocial experience and provided a baseline for their experience and opinions. Following the first two walkthroughs, improvements were made to the prototype and study process.

4.4. Findings from experts

The experts were concerned that the thought bubbles required too much screen real estate and thus obscured important parts of the regular instruction. Thus, the thought bubbles were made smaller. The experts recognized the value of using color cues with the notices but noted that the color distinctions were not always clear. Thus a color highlight boundary line was provided for all notices. The experts suggested that the wording of short sentences be changed to short phrases and bullet points and be provided in a format consistent with the ways the youth were learning the skills. For example the students were learning skills of being a good listener and being a good speaker so the hints should be given as part of a reminder of how to be a good listener or speaker. Thus hints were started with labels for being a good listener or speaker and were condensed into phrases and bullets. These concerns and others like them suggest that the experts were looking for an optimal state where notices are salient but not too intrusive. In a 3D VLE where the environment is constantly changing and where members’ positions in the world cannot always be anticipated this issue of how to make notices peripheral to the action but still impactful is a key design choice. For example, we had earlier rejected audio cues as those might often overlap with audio instructions from the guide or conversations with peers.

Another issue raised by the experts was a concern that the visual only cues would be too ambiguous and possibly too childish for the youth. The experts concluded that the visual only cues might be a good fit for some circumstances and some populations but were not a good mechanism for our population. Thus we did not include visual only cues as a format in the walkthrough for the youth.

4.5. Findings from youth

Overall the youth preferred the command-centered notifications due to the small footprint of the notice and the ease of translating the notice into action. The youth did have criticisms of the command-centered notifications, such as the colors being too bright, positioning on the screen, and perceptions of the wording being “rude” or “mean.” Even though the thought bubbles had been reduced in size, they still were too large for the youth and were criticized for covering too much of the screen. The youth also found the wording of the thought bubbles confusing, such as one comment, “I don’t really know what it is telling me.” Another comment highlighted the challenge of providing contextually relevant and meaningful hints, “It says I’m a listener, but I wasn’t being a listener. It should say I *should* be a listener.” The wording for the command-centered notices were generally better received as summarized by one youth, “These [command-centered notices] seem more effective than thought bubbles because they tell him [the avatar] what to do and why.”

4.6. Conclusions and next steps

Outcomes from this preliminary work provided guidance for how to represent agent messages and notifications to participants who meet our student model. The next step in our process was to consider the context of activity from a broader perspective and to expand the small inventory of user actions defined for the preliminary usability study into one that more comprehensively encompasses the breadth of activities-in-context.

The usability study looked at the context of a student working within a personal pod. It might seem tempting to think that the context for action or behavior in this case is the act of one’s avatar standing on a personal pod. However, the pod is not a context in itself; rather, the activity that is performed on and around the pod defines the context. Typically, personal pods are used in instructor-centered presentations or question-and-answer sessions, and these activities are usually performed during an introduction to a lesson or a wrap-up at the end of a lesson. In these lesson activities, students may be expected to orient their avatars towards a presenter screen and listen while the OG presents a lesson, to answer questions of the online guide, to orient to other students for discussion and utilize gesturing abilities. These lesson contexts circumscribe a set of behaviors

Table 3
Expanded inventory of user actions and behaviors to be monitored and acted upon by an agent

Desired student action/behavior	Agent trigger	Agent action		
		Social affordances	Coaching	Environmental scaffolding
Students should be orienting to one another and to the guide.	Monitor avatar's field of view; when the field goes outside of a predefined zone, perform action.	Report to OG.	Provide feedback to student.	Rotate students' avatar to orient it to where it should be looking.
Students should stay with the group and not be moving their avatar excessively.	Agent monitors distance moved; once avatar reaches threshold, perform action.	Report to OG.	Provide feedback to student.	Disable movement for a period of time or until the OG re-enabled it, so that student could not move avatar.
Students should not be gesturing excessively.	Monitor total number of gestures performed by a student over time; once threshold reached, perform action.	Report to OG.	Provide feedback to student.	Disable gestures for a period of time or until the OG re-enables. Disable overused gestures for a period of time.
Students should not click on things unless instructed.	Monitor for clicks on nearby objects; once threshold is reached, perform action.	Report to OG.	Provide feedback to student.	Disable mouse clicks for a period of time or until the OG re-enables.

and activities that are germane to the intended outcomes of the curriculum.

The behaviors and activities circumscribed by the lesson context comprise an inventory of desirable user actions. While the complete set of desirable user actions may be large, the actions that are germane for an agent are more constrained, as some user actions are more appropriately supported by the online guide, others are impractical or not possible to be managed by an agent and still others might need to be practiced without intervention. With this in mind, we developed a set of user actions that were appropriate to be monitored and acted upon by an agent. These actions and behaviors are provided in Table 3. This list is not meant to be definitive, but rather to serve as a starting point for implementation of agents. The desired student action/behavior column provides a description of the behavior that the agent is monitoring and can act upon. The agent trigger column describes what it is that the agent is monitoring. The three columns included beneath the agent action heading describe how the agent's action falls into our framework of smart 3D CVLEs. Of note is that actions performed by an agent do not have to be singular, nor do they have to be performed on behalf of a single user. For example, if a student were having problems keeping a calm avatar body, the agent

might do all of the following: report to the OG, provide feedback to the student and/or disable movement of the student's avatar.

5. Discussion

The three-part framework illustrated by Fig. 9 is composed of environmental scaffolding, coaching and social affordances and is preliminary in the sense that while we see these components as valuable for our own design and planning, we also recognize that we and the field are still in the early stages of envisioning the use of 3D virtual environments for collaborative learning. Clearly the examples and description we provide in this paper relate the framework to our curriculum on social competence and the special needs of our learners, but we see the framework as applicable broadly across learning that is considered social and complex, which in our mind is most of the learning that we care about. How the framework is used, however, will need to be adapted to the requirements of the curriculum and the relevant characteristics of the learners.

Further, we see these components of the framework not as individual parts that can be pieced together to create a whole, but rather as an intercon-

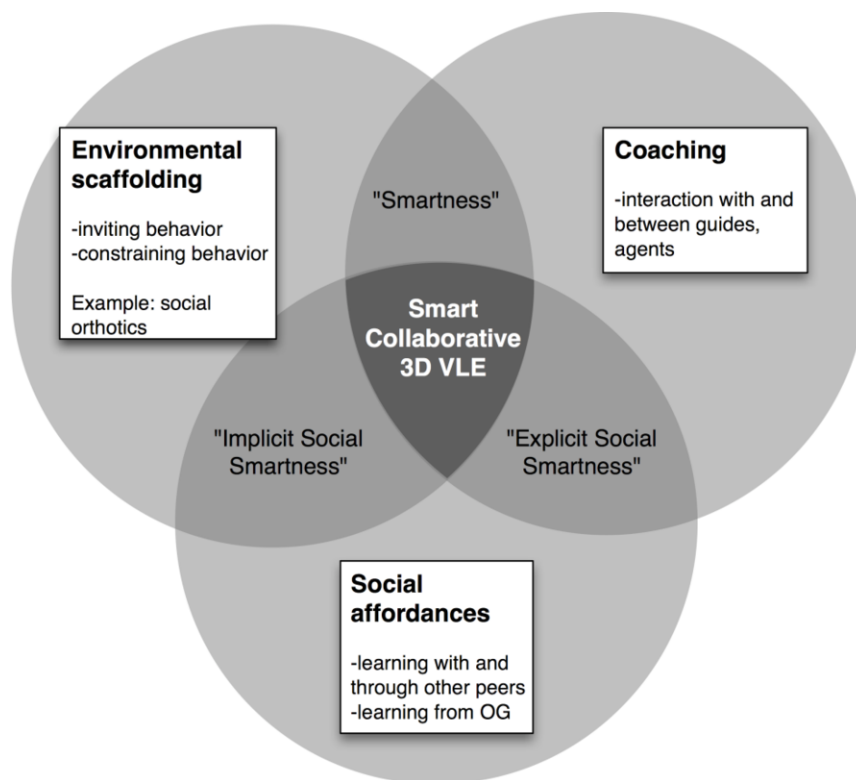


Fig. 9. Framework for smart collaborative 3D VLEs.

nected and symbiotic ecosystem in which all components function in conjunction with one another to exhibit the characteristics of what we envision as a “smart” 3D CVLE. This framework underscores the need for multiple forms of intelligence to realize the potential “smartness” of such systems. Just as Weiser and others have conceptualized “smart environments” as places populated by sensors and intelligence designed around human performance, we also view the 3D CVLE as a place where intelligence can be distributed throughout places, objects and agents to support appropriate cognition and behavior during learning activities.

Our framework envisions environmental scaffolding as a way to both constrain and invite behavior, similar to more traditional ITSs, which aim to constrain undesirable behavior (like gaming the system) and invite desirable behaviors that promote learning. However, our framework differs from traditional systems that conceive of learning as an interchange between a computerized tutor and a single learner, as our vision of learning is social, experiential and distributed. For example, the notion of social orthotics provides social supports in the iSocial system to

promote desirable interaction within groups of peers, between peers and the OG and between peers and the environment. This environmental scaffolding benefits from peers receiving coaching from guides and agents and the interaction between guides and agents working in conjunction, which results in more salient instruction. Learners are able to simultaneously receive explicit guidance from the OG while receiving implicit visual cues from the environment. For example, the design of an environment may invite a behavior (a carpet to designate where to stand or a chair to designate where to sit), while the OG can reinforce learners acting on the environmental prompt. In this way, the OG enhances the smartness of the environment by recognizing and reacting to learners’ behavior, and the environment eases the work of the guide by providing naturalistic prompts for learners to act upon. The current set of environmental scaffolds have limited intelligence, such as to be able to open or lock students and to observe orientation behavior of the youth on a pod, but future versions may also know where the student is on the learning phases and respond according to students prior experience and achievement.

Looking to social affordances, we both recognize and embrace the unpredictable nature of social interaction. Indeed, we see this unpredictability as something that enhances environmental scaffolding. This is because learners in our framework not only interact with a learning environment, but also with and through others who also operate within the environment. The collaborative and interactive nature of social affordances is what enables social learning, making it possible for behavior to be shaped through the social influences of peers in the environment, who are themselves being invited and constrained. This results in a naturalistic experience in which learning is implicit and derived from the synergy between social and environmental factors.

We see the potential of social affordances as an area particularly interesting and important for exploration and innovation for enhancements to 3D CVLE. Using the visibility of others for social modeling or social navigation can provide subtle but easily understood cues for shaping behavior. For example, in sites like Amazon.com we learn that others who bought a certain item also purchased a set of items we have not yet thought about purchasing. This social information predisposes us to consider these new items more so than if there was simply a list of items with no social connection. So, for example, future versions of iSocial or other 3D CVLE we develop may allow students to see traces of behavior from previous sets of students or use activity data to reflect on how their behavior is similar or different from others. Enhancing the use of social information for making smart decisions, whether it be real time modeling or archival processing of trace data from prior users, has substantial potential for supporting collaborative learning.

Coaching, like social affordance, has great potential for 3D CVLE. Currently the OG, a human coach, is a very busy person in iSocial. She implements learning activities, manages a complex technical context, monitors youth behavior, and manages youth behavior. In the Donald Norman sense of “things that make us smart” a substantial amount of our design work goes into trying to help her act smart as she carries out her role. Innovative aspects of the systems, such as scoreboards of youth behavior, personal pods, learning spaces, and video monitoring, are meant to assist the OG in her instructional role. Blending agent-like coaching with human coaching by making learning pods smarter so that they not only lock youth into the space but also help shape their orientation behavior while in the space provide great opportunities for lessening the work of the OG.

Another aspect of helping the OG be smarter is to provide feedback from agent-like coaches to the OG in meaningful ways. For example, currently the scoreboard of behaviors is managed by the OG who adds strikes and merits manually. However we envision agent-like coaches sending reports to the scoreboard about their observations which could help the OG anticipate when youth are starting to misbehave or need some new management approach.

A challenge for coaching and for blending the roles of smart OGs with agent-like coaches is to coach the group, not just the individuals. For example when the OG interacts with one youth to manage behavior it also impacts the others. Understanding how the behavior of some youth impacts the behavior of others can help the OG make good decisions or set up conditions whereby the youth are making their own smart decisions about participating in the learning activities. New types of agents that can analyze social networks and influences among the youth have potential for providing innovative forms of social information back into the learning environment.

As mentioned earlier we see the framework as providing a heuristic for thinking about smart 3D CVLE as an ecosystem rather than a mechanism. The ecosystem has explicit and implicit prompts and invites and constrains. The ecosystem operates through interactions with objects in the environment and through social interactions among peers and with an OG. The environmental scaffolds, social affordances and coaching develop a common vocabulary, skills and ways of working that facilitate communication and enactment in iSocial and the learning of the curriculum objectives.

While the framework we present here has clear potential, there are substantial challenges in implementing environmental scaffolding, coaching and social affordances. One challenge is that while we seek high levels of usability and pedagogical fidelity, there may be conflicts between them. For example, pedagogy may call for messages to be explicit about some aspect of the learning objective, such as our experts guiding us to label hints as being a good listener or speaker, but in practice some students found the labels confusing. Testing is often necessary to develop the best approach because both usability and pedagogical fidelity are required.

Another challenge is how to make the scaffolding, social awareness, and coaching feedback noticed while keeping them peripheral to the learning activities. We do not want the youth self-conscious about standing on a pod, hyper-aware of what others are doing or having hints block their view of instruction-

al materials, yet we need the youth to accept that being on pods helps them succeed in their lesson, to benefit from the presence of others and to get the hint when needed. Learning how to best implement these “smart” affordances is critical to having them benefit the learning process. One area that we envision as having potential for supporting cognition and behavior is by having “smarter” avatars. Currently Open Wonderland provides capabilities for participants to select from a set of possible features of an avatar, such as wearing shorts or long pants or wearing shoes or sandals. However, these options are fairly limited and cannot be dynamically changed during a session. We would like to explore the potential of having dynamic features for an avatar. For example, a youth might start with a green shirt but if they are having behavioral issues and need to self-regulate their actions better, perhaps the shirt changes to red. Alternatively, a youth who is doing very well in a lesson might be rewarded with a cape or star on his shirt.

We take the perspective that the enhancement of collaborative learning through coaching, social affordances and environmental scaffolding most easily focuses on the individual as a part of the group. However, we see a need and opportunity for new conceptualization and research to also focus on how to support a smart group as an entity composed of individuals. In this sense, our framework looks to group behavior as a function of congruent, smart individual behaviors. Congruence and smartness is derived from implicit and explicit coaching and scaffolding and enhanced by the social nature of the collaborative environment. It is the predictable nature of intelligence incorporated into the environment design coupled with the unpredictable nature of social interaction that provides unique opportunities for social and collaborative learning in virtual environments. Additionally, the presence and actions of human and agent-like coaching can promote such learning. As our project moves forward and we continue to develop our understanding of group learning and behavior in 3D CVLEs, we appreciate the need for deeper conceptualizations of what “smartness” might mean for group behavior. Key to developing these concepts will be new knowledge about how to attune scaffolding, social affordances and coaching to group behavior and actions, in contrast to behavior and actions performed by individuals. This remains a direction for further research.

References

- [1] J.R. Anderson, C.F. Boyle and B.J. Reiser, Intelligent tutoring systems, *Science* **228** (1985), 456–462.
- [2] B.J. Reiser, R. Beekelaar, A. Tyle, et al., GIL: Scaffolding learning to program with reasoning-congruent representations, in: *The International Conference of the Learning Sciences: Proc. of the 1991 Conference*, Evanston, IL, 1991, pp. 382–388.
- [3] J.R. Anderson, C.F. Boyle, A.T. Corbett, et al., Cognitive modeling and intelligent tutoring, *Artificial Intelligence Communications* **2**(3) (1990), 277–305.
- [4] V. Dimitrova, J. Self and P. Brna, The interactive maintenance of open learner models, in: *Proc. of the 9th International Conference on Artificial Intelligence in Education*, 1999, pp. 405–412.
- [5] D.C. Merrill, B.J. Reiser, M. Ranney, et al., Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems, *Journal of the Learning Sciences* **2**(3) (1992), 277–305.
- [6] A. Gertner and K. VanLehn, Andes: A coached problem solving environment for physics, in: *Intelligent Tutoring Systems: 5th International Conference*, 2000, pp. 133–142.
- [7] K. VanLehn, C. Lynch, K. Schulze, et al., The Andes physics tutoring system: Lessons learned, *International Journal of Artificial Intelligence in Education* **15**(3) (2005), 147–204.
- [8] A.C. Graesser, K. VanLehn, C.P. Rosé, et al., Intelligent tutoring systems with conversational dialogue, *AI Magazine*, **22**(4) (2001), 39–51.
- [9] K. VanLehn, P.W. Jordan, C.P. Rose, et al., The architecture of Why2-Atlas: A coach for qualitative physics essay writing, in: *Proc. of the 6th International Conference on Intelligent Tutoring Systems*, Berlin, 2002, pp. 158–167.
- [10] R.S. Baker, A.T. Corbett and K.R. Koedinger, Detecting student misuse of intelligent tutoring systems, in: *Proc. of the 7th International Conference on Intelligent Tutoring Systems*, 2004, pp. 54–76.
- [11] R.S. Baker, A.T. Corbett, K.R. Koedinger, et al., Adapting to when students game an intelligent tutoring system, in: *Proc. of the 8th International Conference on Intelligent Tutoring Systems*, Zhongli, Taiwan, 2006, pp. 392–401.
- [12] T. Murray, An overview of intelligent tutoring system authoring tools: Updated analysis of the state of the art, in: *Authoring Tools for Advanced Technology Learning Environments*, T. Murray, S. Blessing and S. Ainsworth, eds, Kluwer Academic Publishers, Dordrecht, 2003, pp. 493–546.
- [13] M.J. Weiss and S.L. Harris, Teaching social skills to people with autism, *Behavior Modification* **25**(5) (2001), 785–802.
- [14] J.P. Stichter, J. Randolph, N. Gage, et al., A review of recommended practices in effective social competency programs for students with ASD, *Exceptionality* **15**(4) (2007), 219–232.
- [15] D. Engelbart, *Augmenting Human Intellect: A Conceptual Framework*, Stanford Research Institute, Menlo Park, 1962.
- [16] D.A. Norman, *Things That Make Us Smart*, Addison-Wesley, New York, 1993.
- [17] M. Weiser, The computer for the twenty-first century, *Scientific American* **265**(3) (1991), 94–104.
- [18] G.D. Abowd and E.D. Mynatt, Designing for the human experience in smart environments, in: *Smart Environments: Technologies, Protocols, and Applications*, D.J. Cook and S.K. Das, eds, John Wiley & Sons, Inc, Hoboken, NJ, 2005, pp. 153–174.

- [19] J. Lave and E. Wenger, *Situated Learning*, Cambridge University Press, New York, 1991.
- [20] J.V. Wertsch, *Mind as Action*, Oxford University Press, 1998.
- [21] A. Collins, J.S. Brown and S.E. Newman, Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics, in: *Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser*, L.B. Resnick, ed., Lawrence Erlbaum, Hillsdale, NJ, 1990, pp. 453–494.
- [22] J. Laffey, J. Stichter and M. Schmidt, Social orthotics for youth with ASD to learn in a collaborative 3D VLE, in: *Handbook of Research on Human Cognition and Assistive Technology: Design, Accessibility and Transdisciplinary Perspectives*, S. Seok, B. Dacosta and E.L. Meyen, eds, Idea Group, New York, 2010.
- [23] Y. Kim and A.L. Baylor, A social-cognitive framework for pedagogical agents as learning companions, *Educational Technology Research and Development* **54**(6) (2006), 569–596.
- [24] H. Nwana, Software agents: An overview, *Knowledge Engineering Review* **11**(3) (1996), 1–40.
- [25] E.A. West and F. Billingsley, Improving the system of least prompts: A comparison of procedural variations, *Education and Training in Developmental Disabilities* **40**(2) (2005), 131–144.
- [26] K. Manley, B.C. Collins, D.M. Stenhoff, et al., Using a system of least prompts procedure to teach telephone skills to elementary students with cognitive disabilities, *Journal of Behavioral Education* **17**(3) (2008), 221–236.
- [27] K.D. Galyen and N. Ding, Usability evaluation study: Agent paper prototypes. Retrieved on Jan. 11, 2011 from <http://social.missouri.edu/iSocial/?q=publications>, 2010.