SUPPORTING COLLABORATIVE LEARNING IN A VIRTUAL REALITY ENVIRONMENT

Shalva S. Landy
Computer Science
Graduate School and University Center, CUNY
New York, NY, USA
slandy@gc.cuny.edu

Lori Scarlatos
Department of Technology and Society
Stony Brook University
Stony Brook, NY
Lori.Scarlatos@stonybrook.edu

ABSTRACT
As computers become more capable of identifying the physical components of human behavior, it becomes less necessary to adjust our behavior to conform to the computer’s requirements. Such capabilities are particularly useful when children are involved, and make using computers in education all the more appealing. Here we look at a system designed to track children as they collaborate on a learning project within a virtual reality environment, giving feedback and limited guidance to encourage progress and problem solving. The learning projects require users to move freely about the environment, thus stimulating the kinesthetic sense. These learning projects encourage participants to access new perspectives, thereby allowing them to perceive the problem in multiple ways. As part of an environment which supports and encourages collaboration, the potential educational benefit is significant.

KEY WORDS
Collaborative Learning, Gesture Tracking, Guide-on-the-Side, Virtual Reality

1 Introduction

In the early 20th century, Maria Montessori introduced the Montessori method, a hands-on, materials-centered approach where children interact with specially designed materials developed to appeal to, and stimulate, the senses[1]. Jean Piaget, considered to be one of the pioneers of child psychology, furthered Montessori’s approach with his advocacy of constructivism[2], which states that people learn through their experiences. Actively doing something provides an experience, whereas passively listening to a lecture does not.

Regardless of whether a materials-centered approach or lecture-based approach is used in the classroom, it is important to reinforce the concepts with plenty of practice. It is widely accepted that having a knowledgeable instructor available on hand during practice is beneficial, especially one that is acting in the role of facilitator, asking relevant and progressively more specific questions to get the students to think about the problem in ways that will lead to understanding and eventual solution. The goal is to provide scaffolding, continually reducing the assistance until the student is capable of solving the exercise unaided.

While human teachers are generally best for conveying new concepts, computers are well equipped to patiently guide and encourage students as they practice applying those concepts. Not only are computers more easily duplicated, updated, and replaced, but without the element of emotion, computers don’t lose patience or grow bored as time passes, and can display a hint or offer encouragement with just as much enthusiasm the tenth time as the first.

By making use of new computer-human interaction techniques that do not rely on good hand-eye coordination or the ability to read, computers can support learners of all ages and abilities. A computer guide-on-the-side provides scaffolding to the user(s), enabling them to discover knowledge not through step by step help, but rather by steering the students with relevant hints at the appropriate times.

Collaborative learning requires those in the group to be responsible for other students’ learning in addition to their own. To be successful, students must join together in analyzing and thinking through a problem, and research shows that children are able to solve more problems when working in groups.[3] Moreover, the discussion that takes place may help promote critical thinking.[4]

The use of computers in supporting team assignments can be successful if “a learning environment that gives all of the children equal access to the data and equal opportunities to manipulate that data” is provided.[5] In other words, true collaboration can only be achieved when all members of the team can explore simultaneously, no turn-taking required. While this is difficult to accomplish with a standard graphical interaction system, other forms of human-computer interaction are perfectly suited to group learning activities; of these, we use virtual reality.

1.1 Virtual Reality

Virtual Reality (VR) is a synthetic environment in which the user is immersed. Inhabitance in virtual environments was originally achieved through the use of a head-mounted display (HMD), a device that displays the virtual environment while simultaneously blocking all external visual stimuli. Because of some of the HMD’s drawbacks (e.g., requiring rendering of such real-world objects as hands and dizziness brought on by the lack of peripheral vi-
sion), other VR systems are on a body scale rather than eye scale, i.e., the user’s body is physically within a room-sized display. One such system is the CAVE™[6], a cube measuring about 10 feet on a side, whose walls and floor are display screens. Participants wear special glasses to view the stereoscopic images. A location sensor is used to enable the viewer’s perspective to be displayed on these screens. Apart from allowing the inclusion of additional users, CAVE™-like systems allow participants to see and use their own hands and other objects in the space.

VR can facilitate constructivist learning activities, which increase motivation and better resist mind-wandering, as children are immersed not only in the environment, but also in the activity. In comparison with a similar paper-based curriculum that included lab time, students in an immersive environment learned as much or more.[7] Experiments in a multi-user virtual environment for teaching science concepts found that typically weak students did as well as their academically stronger peers in the immersive environment; they were able to shed their real-world identities, and step into a successful scientist persona.[7]

Another aspect of VR which makes it so useful as an educational tool is its ability to bypass time and size constraints. As with other computer simulations, we don’t have to wait for things to happen, we can see the results right away, and size can be adjusted to desired dimensions: cells and their parts can easily be viewed or manipulated, and planting a garden results in near instantaneous virtual results.

VR, as well as many non-VR applications, enables the learning experience to be finely tuned to the child’s needs and interests and provides alternative forms of learning which may support different types of learners, such as visually oriented learners.[8] In VR, however, the user is “an integral part of the stimulus flow, [so, if] meaningfulness and active control over a user’s experiences aids learning, then immersive environments likely are better training tools than standard computer-based training environments.”[9]

1.2 Related Work

Early research investigating virtual reality in education typically focused on solitary usage. ScienceSpace[10] is a collection of three worlds created to help students explore difficult physics concepts, such as the relationship among mass, velocity, and energy, while noting the impact of gravity and friction on them.

Cruz-Neira et al.[6] point out that, “One of the most important aspects of visualization is communication. For virtual reality to become an effective and complete visualization tool, it must permit more than one user in the same environment.” Although responses to the ScienceSpace worlds had been primarily enthusiastic, they are inherently individual experiences: the collaborative aspect of learning has been neglected.

The NICE project[11] presents a CAVE™ virtual garden for children within which they can plant vegetables, explore, and meet other children. Available around the environment are “genies” to guide and provide feedback to the children. Although any number of children that will fit inside the CAVE™ may join in, they are collectively represented by a single avatar within the virtual space. This avatar is controlled by a single child in the group who is being tracked. Multiple remote CAVE™s running identical software may network together and appear within the same virtual garden, each represented by its own avatar, thereby providing some level of collaborative ability.

In socio-ec(h)o[12] we see an example of what the authors call “ambient intelligent computing,” which, they state “is the embedding of computer technologies and sensors in architectural environments that combined with artificial intelligence, respond to and reason about human actions and behaviors within the environment.” In socio-ec(h)o, this boils down to using a CAVE™ environment coupled with the Vicon motion capture system to attribute intelligence to the ambient environment. Over the seven game levels that comprise socio-ec(h)o, a group of four players try to uncover the meaning of a word clue that is presented to them. In the authors’ words, “Each level is completed when the players achieve a certain combination of body movements and positions.” As the players move toward completion of the level, the environment changes—the sounds and colors become more intense. Each player is labeled with a different configuration of five reflective markers placed on their back. The Vicon motion capture system tracks the 3D positions of these markers in real-time. The $x$ and $y$ coordinates give the users’ locations on the floor, while the $z$ coordinate indicates whether they are standing or crouching.

In the special education arena, specifically in regard to those with Autism Spectrum Disorders (ASDs), researchers have created “interactive contexts representing a range of social scenarios in which AS1 users can practice social skills.”[13] Virtual environments provide secure, life-like environments in which AS users can learn and practice social skills and rules without the pitfalls of doing it in situ. These environments often have controllable parameters which can adapt to vary and expand the learning experience on a case-by-case basis. Researchers have shown that using VR for social skills training in young adults with high-functioning autism show significant improvements in the clinical measurements of emotion recognition and theory of mind, as well in real life.[14]

As they stand, these social skills training applications are neither collaborative, nor very immersive. Combining the belief that in situ training is best for those with ASDs[15] with their difficulty in generalizing, it may be that using an environment more approximating real-life, such as a CAVE™, may be more relevant and show greater results. Collaboration can be added as more advanced levels for those who have mastered the more basic skills, leading to a more encompassing educational system.

---

1 Asperger’s Syndrome
2 Art Gallery

In our two-stage project, we first detect users’ two-dimensional position on the floor for a simple collaborative application and then combine that data with the three-dimensional location of the users’ body parts (wrist and shoulder) to build a more complex application which relies on users’ moving about and gesturing, working toward a shared goal.

We first implemented an educational game based on the Art Gallery problem, for which we are only interested in the $x, y$ position of each user. This is used as a stepping stone for our second application which involves gestures. The Art Gallery problem is well-known in computer science, and is based on the real world problem of guarding an art gallery. Within a concave polygonal space (i.e., the ‘gallery’), where must guards be stationed so that all parts of the gallery are visible to at least one guard? The goal is to minimize the number of guards needed to protect the gallery, i.e., minimize the number of guards necessary for the entire gallery to be visible by at least one guard.

According to Chvátal’s art gallery theorem, if the museum has $n$ walls, then at most $\lceil n/3 \rceil$ guards are needed to supervise the gallery.[16] Though some concave polygons can be protected with as few as a single guard, we’ve chosen to specialize our implementation for two guards/users. More than two users would require intricate shapes that would not fit comfortably within the allocated area.

2.1 Game Play

For two people, we’ll call them Jamie and Drew, acting as guards for the gallery space shown in Figure 1, we imagine the following interaction may transpire.

1. Jamie and Drew are given instructions and proceed with their mission.
2. They decide to experiment, and Drew goes to the near/bottom left corner while Jamie goes to the diagonally opposite corner.
3. “Hey! We can only see the two end rooms,” they exclaim. “We need someone to protect the center room as well.”
4. They both go into the center room, but now that is all they can see.
5. Drew says, “Let’s see what happens if you go in the corner.”
6. Jamie does so, moving toward the right (near/bottom) corner, and now she can see both the center and right rooms.
7. Jamie says, “Look! We only need to protect that last room.”
8. Drew quickly realizes that he should go (anywhere) in the unprotected room, and does so.
9. They’ve successfully placed themselves to protect all three rooms, and are congratulated.

![Figure 1. Sample art gallery floor, with walls along the edges of the light-colored shape. There are essentially three rooms that need protecting: the left room, the right room, and the connecting center room. The only way to protect this particular layout with two guards is to have one at the juncture between two rooms, and the second anywhere in the third room.](image)

2.2 Technical Details

The specific environment we are working in is a partial CAVE™, with dual 8-foot by 10-foot displays, to give a sensation of partial immersion. The selected polygon representing the shape of the gallery is displayed on the floor of our 12-camera/2-screen Vicon motion capture/CAVE™ system. The wall projection is used to give instructions, as well as hints as needed.

Players wear reflective markers to make them trackable as they move about the gallery trying to place themselves to ensure security of the entire area. For simplification (and as per the theoretical definition), we pretend that our human guards can see 360 degrees about themselves. As they move around, the area that they “protect” gets colored in red on the floor projection, while the rest remains white. When the players (guards) are situated so that the entire polygon is visible, a congratulatory message is displayed on the wall.

Setup involves the users donning a mortarboard-style hat with three reflective markers, the minimum necessary for the Vicon system to recognize as a trackable object. The $x, y$ position of the marker gives the position of the person on the floor.

The application is calibrated by a non-participating party, by selecting each grouping of markers as an object on the tracking computer. This takes about 30 seconds. Having the markers atop a hat helps prevent occlusion.

2.3 Discussion

As discussed in the Introduction, the physical act of moving about within an environment allows a person to perceive things from different perspectives, and that’s what we expected to happen. As players move about, the computer helps them visualize what they would see if it were a real life scenario, thereby enabling them to analyze how the po-
sitions of themselves and others affect what areas are protected. Movement, and the resultant change of perspective, triggers the discussion which ultimately leads to a solution.

It is clear, especially from points (3), (5), (7), and (8) in the sample game play, that the collaboration has fostered the metacognitive behaviors: in (3) and (8) they display a clear sense of understanding of the problem, and in (5) and (7) they plan a course of action which leads to a solution. These metacognitive thought processes help keep the children on track by inducing a higher level of immersion in the activity through the engagement of their intellect.

While we did not bring in groups of children to test this application, three users of varying heights, including a four-year-old child, verified that it will work for users of a wide range of heights. Even with a four-year-old’s attention span, she was able to stand still for the 30 seconds it took to calibrate the system.

Currently, a single hint has been developed, and although applicable at all times, it’s not very specific, merely stating, “Only the red section is visible by the guards. Place yourselves so that the white area is visible, too.” Additional hints would improve the application’s usability, but this was left as a future exercise.

3 Classification Application

Classification is an essential skill that is called into play throughout a child’s education, and is a component of math and science standards. We have therefore designed a system where students sort a group of objects by predefined attributes. Our current implementation has students classify produce as fruit or vegetable, but this can easily be modified to accommodate almost anything—shapes, animals, trees, rocks, artwork, etc.

We now look at a sample round of game play (both generally and in detail), feedback that is given during the game, and the technical details involved in the application. We then discuss the educational benefits of the system.

3.1 Game Play Overview

Let’s look at a scenario where two people, we’ll call them Magenta and Goldenrod, are playing the fruit/vegetable sorting game:

1. At the start of the game, Magenta is standing on a photo of an orange and Goldenrod is standing on the blueberries.
2. Magenta (on the orange) points toward the fruit basket and Goldenrod (on blueberries) points at the vegetable basket.
3. Magenta’s orange is moved from the floor up into the fruit basket. Nothing happens to the blueberries—it remains in its place on the floor—as Goldenrod is not pointing at the correct basket.
4. Magenta suggests to Goldenrod that perhaps blueberries are not a vegetable, but a fruit.
5. Goldenrod realizes Magenta is correct, and points toward the fruit basket . . .
6. . . . while Magenta hops over to the grapes and also points to the fruit basket.
7. Both the blueberries and the grapes move up into the fruit basket.
8. They continue moving about the floor, aiding each other through discussion along the way, and sending the produce to the correct bins, until there are no more left on the floor.
9. At that point, the round of play has been successfully completed.

3.2 Game Play Details

When the application is first started, eight objects to be categorized are displayed on the floor. On the wall are displayed instructions as well as “bins” for the items to be sorted into. In our case, we display a mixture of eight fruits and vegetables on the floor. On the wall are displayed two baskets, one marked “FRUIT” and the other marked “VEGETABLES” along with the instructions, “Stand on a fruit or vegetable, then point to the basket where the item belongs.”

As users move about the floor and point at different locations on the wall, small colored circles are displayed at the position toward which they are pointing.

When the application is started up, placeholders are displayed on the wall above the baskets which make it clear that four items go into each basket. Similar placeholders are displayed on the floor in place of an item once it has been moved into the correct basket for illustrative purposes. These placeholders can easily be removed.
Users continue to play until there are no more items left, at which point they may begin a new round, with eight more randomly selected fruits and vegetables.

### 3.3 Feedback

In any educational activity, on the computer or off, a student needs feedback to learn, and we have incorporated feedback in a few different ways. When you talk to someone, you know they’ve heard you because they respond to you. With a computer, you want to know that they’ve “heard” you and registered your request. For that, we display the pointers mentioned earlier in the game play section. The spots on the wall at which the users are pointing are marked with large dots, using dissimilar colors for each dot.\(^4\)

There may be ambiguity about dot “ownership” when two players are pointing at the same (or nearby) position, but there is only minimal chance that this will create confusion, i.e., if they are near each other, the users are probably pointing at the same thing. Informal testing has not shown differentiation among these dots to be a problem. Another area for possible confusion exists when one or more users is standing too near the screen and can’t easily see their pointing trajectory. To prevent this, we only display floor objects in the far two-thirds of the floor, leaving a large buffer between users and the wall display.

The dots can be a source of other kinds of feedback as well; one variation we tried displayed a check on the dot if pointing at the correct bin, and an X otherwise. This can prevent users from wondering why their item is not being placed into the bin—“Is the system not working or am I wrong?” Ultimately we took this off as we want users to pay attention to the item they are placing and the bins: the primary focus should not be on the dots. The quiver of the dot from a person’s naturally imperfect hand steadiness is usually enough to let the user know that the machine is, in fact, responding.

A fruit-vegetable sorter is trivial as an application: either you know it, or you don’t but guess, and with practice will remember it. With other classifications, such as rocks, neither absolute nor relative size can be ascertained. Texture, too, which is often used as a differentiating factor, is lost. To counteract this problem, actual specimens may be placed near or within the environment, where feasible. Aside from allowing participants to get a better sense of the objects, the objects’ presence is expected to facilitate discussion.

### 3.4 Technical Details

The technical aspects of the system can be broken down into tracking/capture, display, and process, which we discuss here, together with known technical problems, some of which we have addressed, and others which need to be addressed.

#### 3.4.1 Capture

Our system is comprised of a front-projected floor and a single rear-projected wall CAVE\(^\text{TM}\), surrounded by 12 Vicon cameras. These specialized cameras project infrared light which is reflected off the markers worn by the subjects, and filter out all light except infrared. To aid in reflector detection, ambient light must be kept to a minimum. Accordingly, walls are painted black, and dark navy or black clothing should be worn in the environment.

The Vicon motion capture system receives input from the cameras and passes the input along. A server program grabs the data and relays it to another computer where it is picked up and interpreted.

The points we are tracking, each labeled with a set of three reflective markers, include two points per person and two calibration points. The two points we track on each person is their shoulder and wrist of their pointing hand. In our informal tests of three people of heights ranging from under four feet to almost six feet, tracking the shoulder position was more accurate than tracking a point on top of the head. The “feedback dot” appeared to the user to be displayed at the precise location they were pointing toward. Placing the reflective markers on the shoulder and wrist are less likely than the hat to be a distraction for those with ASD-related sensitivities, since the hat is additional attire, while sensors can be placed directly onto the child’s own clothing at the shoulder and wrist positions. Each group of three markers is selected as an object within the Vicon application.

#### 3.4.2 Display and Process

From a group of fruit and vegetables, we randomly select four of each of them, and randomly display them on the

\(^4\)Unfortunately, we did not refer to any guidelines for choosing colors to avoid ambiguity for the color-blind.
floor. Although the proportion does not necessarily have to be 50-50, we chose to display equal numbers of each since they fit nicely on the wall, when categorized: four in the basket on the left, four in the basket on the right. Although we chose to give the learners eight objects to sort for this demonstration, up to fifteen items could be comfortably placed on the floor. These items are spaced far enough apart to decrease the likelihood that the capture system will be confused, and large enough so that even if the user forgets what they are standing on, a quick glance down is all that’s necessary. In most cases, it should not be necessary to move to the side to get a complete visual. As mentioned, items are displayed in the far two-thirds of the screen, leaving a buffer between the players and the wall display.

The wall displays two baskets, but these may be substituted with garages, toy chests, or animal pens, to conform to the objects being classified. Two bins is not an absolute requirement, but we recommend no more than six, in a three across by two down grid, to avoid accuracy issues.

While any number of people that will reasonably fit in the floor area can be set up with markers for game play, two or three seems to be optimal within the proportions of our CAVE™. This allows each person enough space to move around independently without bumping into anyone else, while also giving them the opportunity to sort a moderate number of items. Furthermore, discussions are more focused when they’re one-on-one or in a small group. At the other end of the spectrum, individuals can use the system as well, but working by oneself, one misses the benefits of collaboration.

For every set of user data that arrives, the intersection of the line defined by the two points on the arm and the wall plane is computed. The point of intersection is where that user is pointing, and a large dot is displayed, as discussed. For this application, we ignore any data outside the constraints of the screen. We then determine toward which section of the wall they are pointing, and if it’s correct for the item they’re standing on (as determined by the location of the shoulder), the item is moved into the bin.

We do not require the user to point at a bin for a specified length of time, as we are not waiting for a “click” but rather a “point.” This has its drawbacks, as discussed in the next section.

3.4.3 Known Problems

All vision-based systems are going to be subject to some degree of occlusion. Although increasing the number of markers per object helps insure against this, it can “result in exponentially increasing the ‘confusion factor’, i.e., keeping track of which marker is which.”[17] The Vicon system, suffering from occlusion-related confusion, repeatedly confused objects (a hand for a shoulder, one person’s shoulder for another’s), even when not in close proximity to each other, but especially then. Because this often occurred during calibration, we adjusted our code to account for this. Yet when this confusion occurs during game play, we need to restart the system and re-mark the points to be tracked. This has a tendency to happen more with smaller kids, as their limbs are short, so markers are closer together.

While the wall display was rear-projected, the floor used an overhead projector, and was therefore subject to obscuration, as the users are physically on top of the display surface. For each user, the obscured part would include the section of the floor that they were standing on, plus the shadow of their body and outstretched arm. Narrow arm-shaped shadows were not typically a problem and, as people don’t generally lock their elbows when pointing at something close by, the shadow was shorter than arm’s length. We tried to make the item images large enough to be seen around body shadows, although in general, the centered ceiling projector created shadows outward, which were non-blocking: people standing in positions which otherwise might cast long shadows on those behind—front row center—had significantly shorter shadows because the light source was projecting straight down on them.

Using gestures, while natural to humans, can be physically draining if required over long periods of time[18], especially if they have to be exaggerated for recognition by a computer.

In our informal tests, we have not noticed an incident where the user’s hand crossed over the wrong basket to get to the correct one. This may be because our natural pose is not pointing, and we put our hands down while walking about. When raising the hand to point, it is usually in the predetermined pointing direction. However, since we only recognize a pointing gesture, it is very easy to just move one’s hand around, pointing at different locations on the screen and the item will move into the correct bin at the moment that the pointing is correct. This makes it possible to place all objects in the correct bins without learning, or knowing, anything, if done purposefully.

3.5 Discussion

There are many advantages to this style of learning. Learners do not need to wait for their turn. They move about the workspace, stimulating their kinesthetic sense, thereby engaging more of their body in the process. They work at their own pace, sorting whatever objects they are familiar with, while freely discussing their task with other players or onlookers. Although one player can conceivably classify all of the items while the others watch from the sidelines, we do not include scorekeeping. This is designed to encourage collaboration among the players to work toward a common goal (“clearing” the floor), and not be concerned with how many points they accumulate.

An enforced collaboration option may be useful in some populations, such as among those with ASDs, to require participants to work together. This may work in a couple of different ways: splitting the floor into two identical workspaces of half the number of items to sort, and requiring both players to stand on the same object and classify it together (requiring them to work in concert with one
another), or not allowing the game to proceed until both players have classified an object successfully, thereby encouraging players to assist one another. It may be advantageous to include both of these options.

Whereas collaboration among children sorting fruit and vegetables might seem unnecessary, especially with feedback indicating whether the placement is right or wrong, advanced topics, like sorting rocks for an introductory geology class, are more likely to trigger discussions on what to look for. A rock classifying game may have students sorting rocks into igneous, metamorphic, and sedimentary bins. A sample game with Violet and Hunter may go like this:

1. Violet stands on a photo of sandstone, and Hunter on gneiss.
2. Violet points to place the sandstone into the sedimentary bin, while Hunter points at the igneous bin.
3. The sandstone goes into the sedimentary bin, while Hunter points at the igneous bin.
4. Violet says, “Notice the layers in the rock. I think that’s called foliation, which is an indication that the rock has metamorphosed.”
5. Hunter responds, “I see! It must be metamorphic,” and points to place the gneiss into the correct bin.

In this sample interaction, we see how the two players are collaborating, helping to point out the various factors that go into making a determination on an item’s classification. Metacognitive events show up as analyzing in point (4) and understanding in point (5). Increased immersion in the activity brought about by engaging the child’s intellect—by stimulating metacognitive thought processes through collaboration, combined with arousing the child’s kinesthetic sense, creates what we believe is a synergistic effect on learning.

Since participants in our system are co-located, they can talk naturally without being mindful of the requirements imposed by communicating in real-time across a network, such as facing a camera or taking extra care to speak loudly and clearly for the microphone. Overall, we have tried to keep the focus of the users off the technology, and on task, by avoiding distracting stimuli or requirements.

Although we have not included a scorekeeping feature, this may be an interesting game option to add. Keeping score can encourage collaboration if done across multiple rounds of game play (first one team and then another use the system), particularly if the awarding of points is tied to collaboration-related metacognitive events.

Another improvement would remove the restriction on the quantity of objects that may be classified. Rather than limiting quantity to what can be comfortably displayed on the floor, new objects appear on the floor once a user has correctly placed an object and moved away. One possible application using this feature may be in Geography: classifying countries by continent. Displaying a world map on the wall, with the continent’s location acting as its bin, enables users to sort all the countries of the world (or some subset thereof), as desired by the teacher. To begin, the original eight or so countries are displayed on the floor. Then, as each correct classification is made, the floor is updated from a pool of countries until the pool is depleted, at which point the game is over. One consideration would be how to display all classified items “in” the bins on the wall.

Competitive scorekeeping within a round makes more sense in an unlimited-quantity classification game than it does in a game with all objects pre-displayed. For those on the ASD spectrum, competitive scorekeeping may be a predecessor to working in a more cooperative fashion. Scorekeeping may provide the motivation necessary to maintain interest while the ASD user becomes familiar with the system, an essential step if they are to successfully collaborate in future levels.

Although some researchers are using Microsoft’s Kinect™ sensor for the Xbox® gaming console, to track pairs of participants in learning activities, our system has some significant benefits that give it more capabilities thereby making possible more applications. Although both motion capture systems use infrared, they take different approaches. Kinect’s infrared makes it best for users standing in a gaming configuration about a single display, while Vicon allows players to stand anywhere within the environment (even behind another person or object) and still be tracked.

4 Conclusion

Supporting collaborative learning within the virtual reality environment contributes not only to the child’s academic development, but their social development as well.

Problem solving, in part, requires that people learn that (in all aspects of life) the first approach isn’t necessarily the correct approach. The Art Gallery application provides students with the opportunity to explore how a change in perspective may contribute toward finding a solution to a problem: moving about the gallery, analyzing each perspective, and comparing it with other perspectives. The rudimentary guide-on-the-side supports learning by giving participants food for thought, encouraging metacognitive discussions. Presenting the task as a collaborative one that requires physically active involvement compels all participants to do their share. The resulting full-body engagement in the learning process, coupled with the support of both collaborator and guide-on-the-side, will be the motivation and means to learn.

Although the CAVED/®/Vicon classification system has not been tested on students, we have shown that it is capable of tracking and responding to the gestures of multiple children in a virtual reality learning environment. This enables the creation of learning activities that are truly active and collaborative.

We have put together a number of separate systems (immersive environment, gesture recognition, guide-on-the-side) to form a cohesive system ideal for use in collabo-
rative educational settings. We have also discussed the various considerations that must be taken into account when developing such systems for general, as well as special education, and presented solutions to some of them.

Limitations include the omission of audio (a good way to incorporate it must be found), minimal guide-on-the-side (like audio, various issues need to be taken into account), and environment size and accuracy is constrained by the system.

Possible applications for the classification system specifically range from shape sorters for toddlers (although collaboration is not typical among toddlers) to advanced science concepts for college students. More generally, Gilbert[19] discusses many ways in which movement can be used to teach numerous concepts in subjects like language arts, math, science, social studies, and art. Although fun and educational in a group setting led by the teacher, such learning can be supported by our system, both for single learners as well as collaborators. Additionally, the system can be used for social skills training for those with autism spectrum disorders, as the environment can more closely replicate real-world situations, and can be customized for each child or pair of children, as needed.

Acknowledgements

Thank you to Jim Cox of Brooklyn College, CUNY and Zhigang Zhu of City College, CUNY for their help and support in this research. This work has been supported by the National Science Foundation through grant number CNS-0420996.

References


