Understanding Visual Characteristics in Virtual Digital Assistants that Support Positive Affect in Learning Elementary Science

Eric N. Wiebe
Jennifer London
John Bedward
Marc Russo
Bradford Mott
Robert Taylor
James Lester

North Carolina State University

Presented at the AERA Annual Meeting, Vancouver, BC

April, 2012
1. Objectives or purposes

Computer-based learning environments have become an increasing part of teaching and learning in K-12 classrooms as technology becomes more affordable and more technically suited for educational use. There are several benefits provided by computer-based learning environments, including the potential scalability of tools and resources that are not limited by the purchase and availability of physical textbooks, lab and classroom supplies. In addition, these computer-based tools often seamlessly bridge the time and space gap between learning in and out of classrooms and the traditional school day. A particularly promising emerging capability of computer-based learning environments is the use of virtual digital assistants to provide both affective and cognitive support for students (Boyer, Phillips, Wallis, Vouk, & Lester, 2008; Porayska-Pomsta & Pain, 2004; Wang & Chignell, 2005).

The goal of the Leonardo project is to create a tablet-based elementary science learning environment that includes a virtual digital assistant. Given the importance of the affective role in learning, it was considered central that the embodiment of the virtual digital assistant have physical characteristics that students find to be positive. Of interest in this study were the visual characteristics of the virtual assistant that engender positive affect in students. Also of interest are those visual characteristics that engender negative affect, because not attending to this issue could hamper the overall effectiveness of the tablet-based system. Effectively, even if the cognitive scaffolding designed into the other components of the system and in the dialogue of the digital assistant were well-designed, if the assistant creates negative affect, the cognitive scaffolding could well be muted. This presentation will report on a set of student focus groups designed with the goal of helping unravel what visual characteristics of a virtual digital assistant engender positive affect towards learning science in 4th and 5th grade students.

2. Perspective(s) or theoretical framework

For the purposes of this study, a virtual digital assistant is part of computer software geared for supporting one or more topics or subject areas in the school curriculum. In order to be a cohesive component of the interactive graphical user interface, the virtual digital assistant is represented as an avatar. This avatar is capable of written and/or spoken speech, gesturing, and emotive expression in ways that would be readily comprehensible by the student. Complimenting the indispensable role of the teacher, a digital assistant may be able to monitor student’s work, determine where he or she is having trouble, and provide a targeted response to help the student past the problem. Second, a digital assistant can provide affective support such as encouragement that helps bolster student motivation.

Learners interact with these digital assistants with many of the same kinds of social interaction and expectations that they bring to human-human interactions (Veletsianos, Miller & Doering, 2009). Because of this, well-designed avatars can provoke and support the same kinds of positive affect that humans are capable of. Affective support to increase student motivation and
engagement has been shown to be a key role for virtual assistants in these environments. For example, Craig, Graesser, Sullins, and Gholson (2004) examined several aspects of affect with learning performance during a task using AutoTutor, an intelligent tutor program, and boredom was shown to correlate negatively and flow state positively. In this study, flow was conceptualized as an aspect of affect, but it is also commonly viewed as an aspect of engagement (Csikszentmihalyi, 1991). Furthermore, there has been additional support that adding affective support to the virtual experience results in higher learning performance (Boyer et al., 2008; Porayska-Pomsta & Pain, 2004).

3. Methods, techniques, or modes of inquiry

A total of 24 students from grades 4 and 5 participated in the study. Participant ages ranged from 9 to 12 years and the sample was 54% female and 46% male. The majority of the students were Caucasian (70%), with the others indicating Asian (13%), African American (13%) and Latino (4%).

Procedure

The research team met with two classes of twelve students. The twelve students were seated in groups of six. Each class was shown an iPad® tablet computer and asked to think about the possibility of completing their current paper science notebook tasks on the iPad. They were also told that a digital assistant would be part of this learning environment and designed to help them with their science activities. We asked the students to help think about how the digital assistant should look and act.

Each group of six students had any questions answered related to the opening presentation and then was shown illustrations of 24 potential avatars for the digital assistants and asked to provide impressions of each illustration. Next students were individually asked to sort these illustrations in the order of their preference (i.e., a card sort task). After more discussion of the role of the digital assistant and how it might function, each group was asked to pick their top three to share with the class, after sharing at the class level, each student filled out a demographics survey. The digital assistants varied widely in terms of species, race, and gender. The digital assistants did not vary in terms of artistic style; the same artist created each illustration (see Figures 1 through 3).

4. Data sources, evidence, objects, or materials

Each session of whole class and group activity was video-taped. A two-pronged approach was used for the analysis of this data. First, a research member not affiliated with the data collection team reviewed the video footage with the goal of getting a broad understanding of the students’ verbalized thoughts regarding the virtual assistants. Another independent research team member coded and analyzed the card sort data quantitatively. Finally, all members of the team came together to synthesize the quantitative and qualitative findings.
5. Results and/or substantiated conclusions or warrants for arguments/point of view

The qualitative approach revealed several important factors. First, students verbally indicated that they did not care for avatars who did not have a positive affect. The HiTech Boy (Fig 2), for example, was identified as seeming “blank.” Students also identified the Alien as seeming especially fun, playful, and cool. Another characteristic that students desired is that the assistant be removed from the school environment; they preferred an outsider to a peer. Along those lines, students mentioned that the human avatars seemed boring and average. Finally, students wanted the avatar to have a “history.”

The first aspect we examined was the type of characteristics associated with the most popular avatars. In order to evaluate popularity, the number of each type of ranking for each avatar was tabulated. The most popular avatar, for example, had four students rank it as number one and no students rank it last. The distributions were then divided into quartiles, with six rankings per quartile. Finally, a weighting scheme was applied such that the first and last rankings carried more weight than the intermediate rankings.

According to the rankings, the three most strongly preferred avatars were the Alien, the Fox, and the Lizard (Figure 1). On the other end of the spectrum, the three most strongly rejected avatars were the Caucasian HiTech Boy, the Strange Bug, and the African American Boy (Figure 2). Another set of interesting avatars are those that had little variance; that is, those that a roughly equal number of students ranked high, intermediate and low. Those with the least variance were the Older Girl, the Cat, and Bug #2 (Figure 3).

*Figure 1.* Top three ranked avatars: Alien, Fox, and Lizard.
Figure 2. Bottom three ranked avatars: Caucasian HiTech Boy, Strange Bug, and African American Boy.

Figure 3. Lowest variance avatar group: Older Girl, Cat, and Bug #2.

A series of ANOVAs revealed that certain characteristics of the avatars differed based on the ranking dependent variable, including the species of the avatar ($F(2, 27) = 5.77, p = .008$). A Tukey’s post hoc test revealed that sort priority was not significantly different between humans and bugs. However, significant differences existed between aliens and bugs, aliens and humans, animals and bugs, animals and humans, and robots and humans ($p < .05$). The pattern of mean differences indicated that aliens, animals, and robots tended to score higher than the significantly lower humans and bugs. These differences are illustrated in Figure 4, where a score above the mean indicates a predominance of rankings in the top 12 and below the mean indicates a predominance of rankings in the bottom 12.
Despite many variables being coded, there were few significant correlations with ranking priority. An exception was that dull coloration in the avatar was moderately significant ($r = .44$, $p = .03$). There were no mean differences between high, medium, and low height or weight groups, nor were there differences in color brightness groups.

Although further ANOVAs found no evidence for mean differences between avatar races (Caucasian, African American, and Indeterminate), descriptive nested analyses revealed some interesting relationships when student demographics were taken into account. Namely, students tended to select either their own self-reported race for their top rank choice or an avatar of indeterminate race (see Table 1).

Table 1. Student Ethnicity and VA Ethnicity

<table>
<thead>
<tr>
<th>Ethnicity of Top Ranked VA</th>
<th>Caucasian</th>
<th>African American</th>
<th>Indeterminate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>9 (53%)</td>
<td>0 (0%)</td>
<td>8 (47%)</td>
<td>17</td>
</tr>
<tr>
<td>Other</td>
<td>1 (14%)</td>
<td>1 (14%)</td>
<td>5 (71%)</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>1</td>
<td>13</td>
<td>24</td>
</tr>
</tbody>
</table>
A similar pattern was revealed for gender, wherein males tended to choose males for the top rankings and females for the bottom rankings, and vice versa (see Table 2).

<table>
<thead>
<tr>
<th>Student Gender</th>
<th>Gender of Top Ranked VA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
<td>6 (46%)</td>
</tr>
<tr>
<td>Male</td>
<td>8 (73%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>14</strong></td>
</tr>
</tbody>
</table>

Avatars of indeterminate gender, however, were selected with somewhat equal frequency regardless of student gender.

6. Scientific or scholarly significance of the study or work

In summary, the most important finding was that the top three are all non-human. They seem to have a greater degree of positive affect than the bottom three, and they are more dynamic in their postures. The Alien, which was verbally identified by students in the classroom, was the single most highly ranked character. Meanwhile, the HiTech Boy fell in the bottom three.

The high and low ranked avatars provided important information regarding the likes and dislikes of students, but the low variance group is also of interest. The characters in Figure 3 all embody characteristics identified by students as positive as well as negative. The ranking data suggests that some characteristics are more important to some students than others, and that while these avatars are highly positive for some students, they are negative for others.

The Alien species, and especially the avatar labeled Alien, was by far the most popular type. This may be due to the students’ stated preference for an outsider, which an alien is by definition. Aliens are often portrayed as intelligent in the media, and almost certainly have an interesting history. The aliens depicted in this study tend to be positive and non-threatening. These characteristics help delineate it from the Robot category. The Robots can be perceived as outsiders and perhaps as a pet, but they are also somewhat cold in aspect. Animals were the opposite: they fulfill the desire for being friendly and non-threatening, but not for being an outsider or intelligent. Humans and bugs may not have been perceived as having any of the qualities desired by students other than positivity.

The non-human species may have benefitted from the differential selection of race and gender based on the students’ own traits. For example, while many Caucasian students ranked Caucasian avatars highly, they were somewhat canceled out by the same phenomenon happening with the minority students. The same phenomenon occurred with gender. Because there was no differential selection for the non-human characters, there were no canceling effects. This
evidence points to the unique suitability of non-human avatars due to the fact that they are equally identifiable to a broad demographic.

The analysis was somewhat hampered by a low sample size, but qualitative analysis from the video supported the quantitative analysis. These finding will be used to guide future data collection and analysis as the project moves forward. It will be helpful to have a more balanced sample in terms of facial expressions and posture, as well as student race and ethnicity. Specifically, the study design would benefit from having an overtly happy and smiling group that is roughly equal in size to the neutral group. The number of smiling characters was not sufficient to make any statements regarding the importance of that particular characteristic.

References


