Virtual Human versus Human Administration of Photographic Lineups

Brent Daugherty and Lori Van Wallendael • University of North Carolina, Charlotte

Sabarish Babu • University of Iowa

Brian Cutler • University of Ontario Institute of Technology

Larry F. Hodges • Clemson University

Investigations by the Innocence Project (www.innocenceproject.org) using DNA evidence have led to the exoneration of more than 200 convicted felons, many of whom spent more than 10 years in prison, and some of whom were sentenced to death. The actual number of erroneous convictions is likely orders of magnitude higher than 200 for two reasons. First, almost all the exonerated cases in the Innocence Project involved sexual assault because definitive DNA evidence (contained in semen) is available for these cases. However, according to one recent study conducted in and around Chicago, sexual assault was involved in only 5 percent of all lineups. Furthermore, in most cases involving DNA evidence, this evidence has since deteriorated, was lost, or was destroyed.

The Innocence Project’s research and prior archival studies have established that mistaken identification is the leading precursor of erroneous conviction. For the foreseeable future, eyewitness testimony, although imperfect, will often be the most important evidence in many criminal cases. Recognizing mistaken identification’s role in erroneous identification, some states and many police departments have revised their identification procedures, taking into account scientific research on factors affecting identification accuracy’s suggestiveness. Among the procedural reforms is the use of “blind” administrative procedures (we discuss these procedures in more detail in the next section).

A main concern with blind procedures is their practicality. Some police departments have resisted this reform because they don’t have the personnel to implement it. To address the influence of officers and the limited availability of qualified personnel, research and commercial organizations have introduced software to conduct the lineup autonomously using a standard window-and-mouse interface.

Using a virtual human interface has some advantages in a social collaborative task such as a photographic lineup. Such interfaces can employ multiple modalities (such as gestures, facial expressions, and natural language) to communicate information to the user. (For more on virtual human interfaces, see the related sidebar, next page.) In this article, we describe our virtual officer framework (VOF), which combines cutting-edge speech and

One solution to mistaken identification by a crime’s victims and eyewitnesses is to use a virtual officer to conduct identification procedures. Results from a study comparing a virtual officer with a live human investigator indicate that the virtual officer performs comparably to the human in terms of identification accuracy, emotional affect, and ease of use.
Virtual Human Interfaces

Virtual humans are 3D representations of an underlying, complex system that combines cutting-edge research with paradigms from social psychology that transfer to interactions between humans and virtual humans. Disciplines involved in this research include speech recognition, animation and rendering, planning and discourse modeling, computer vision, and real-time speech synthesis.

Anthropomorphic and Socioemotional Factors

Several studies have shown that people respond socially to computers much like they do to humans and that, when virtual characters are involved, appearance, verbal behaviors, and nonverbal behaviors can affect the interaction.

Andrew Raij and his colleagues examined the similarities and differences in an interpersonal scenario with a real patient and a virtual patient. They observed the following:

- Assessments of virtual interpersonal scenarios’ authenticity should use indirect measures because direct measures invariably lead to a preference for the real condition.
- A virtual human can meet an interpersonal scenario’s goals even if the interaction isn’t equivalent to the real interaction on other measures.
- Virtual characters in interpersonal scenarios must be expressive because expressiveness improves conversation flow and rapport, and elicits more natural behavior from the interlocutor.

Catherine Zanbaka found that a virtual character’s visual appearance isn’t a significant factor in persuading participants. The problem with increasing anthropomorphism is that the increase in the interface’s fidelity corresponds to higher user expectations, which leads to disillusionment when the interface doesn’t meet those expectations. Although realism might not be an added benefit in human-to-virtual-character interaction, evidence suggests that the degree of similarity between the virtual human and the user plays an important role in how the user perceives the virtual human.

Clifford Nass and Kwan Min Lee discovered that participants could identify personality cues through text-to-speech (TTS) and were more attracted to personalities that were closer to their own. Participants were also more likely to buy a product from a computer interface that matched their personality. Interestingly, virtual characters of the opposite gender from the user are more influential than those of the same gender. These results suggest focusing on the individual user’s physical and personality traits to get the most out of the interaction.

Given that the interface agent is comparable to humans, it’s important to avoid side effects that carry over from human-to-human interaction. For instance, social facilitation/inhibition theory states that in the presence of others, people perform simple tasks better and complex tasks worse. Zanbaka recently confirmed this result with virtual humans, with the caveat that simple tasks weren’t reportedly better—although Zanbaka suggests a possible confound in that the task might have been too easy to observe a noticeable effect. This result implies that virtual human applications should be relatively simple. This research’s general implication is the importance of considering human-social-interaction theories when developing interface agents.

Virtual Human Applications

The University of Florida’s Virtual Experiences Research Group has been exploring the use of virtual humans as patients to give medical students the opportunity to practice patient interviewing skills. Aaron Kotranza and his colleagues recently used a mixed-reality virtual patient that provides haptic feedback while students interact with it. When afforded haptic interaction with a mixed-reality virtual patient, users demonstrated interpersonal touch and social engagement similarly to that demonstrated when interacting with a human patient.

A virtual receptionist is one relatively simple application that has matured over time. Marve (Messaging and Recognition Virtual Entity) is the most recent and advanced form of virtual receptionist. Marve uses natural face-to-face communication to take and deliver messages, tell knock-knock jokes, discuss movies, and report the weather. Marve’s developers found that users employed social conversational conventions, perceived and described Marve as a social entity, and—through voluntary repeated interactions—demonstrated a genuine interest in talking with him.

References

Lineup Basics

Law enforcement personnel use photographic lineups (also called photoarrays) to test whether a person suspected of committing a crime is, in fact, the perpetrator. Generally, a police officer shows a set of photographs consisting of the suspect’s photo and photos of fillers or nonsuspects to a victim or eyewitness and asks whether he or she recognizes one of the persons in the photos as the perpetrator. A positive identification of a suspect can lead to an arrest, and the act of identification can later serve as evidence in the defendant’s prosecution.

In a perpetrator-present lineup, the person who committed the crime is in the lineup. In this case, there are three possibilities:

- correct identifications—that is, the eyewitness positively identifies the perpetrator;
- filler identifications—that is, the eyewitness identifies a photo other than the suspect’s; and
- misses—that is, the eyewitness incorrectly claims that the criminal wasn’t in the lineup.

In a perpetrator-absent lineup, the person who committed the crime isn’t in the lineup. In this case, the possibilities differ slightly:

- false identifications—that is, the eyewitness positively identifies an innocent suspect in the lineup;
- filler identifications; or
- correct rejections—that is, the eyewitness correctly claims that the criminal isn’t in the lineup.

In suggestive procedures, the investigator conducting the identification takes actions that increase the likelihood of a positive identification in general or a positive identification of a specific suspect. For instance, a police officer who spends more time on one photo in a lineup than the other lineup photos might suggest to the eyewitness that that photo is the perpetrator. (For more on suggestive procedures, see the “Eyewitness Identification” sidebar.)

In blind procedures, the investigator conducting the identification is unaware of the suspect’s identity. By being blind to the suspect’s identity, the investigator can’t inadvertently or inadvertently influence the eyewitness’s selection of the suspect.

The VOF

We built the VOF on top of a virtual human interface framework. Our virtual Officer Garcia interacts using a combination of spoken natural language and nonverbal cues that include facial expressions and gestures, while maintaining human-like social-communication protocols such as turn-taking, feedback, and repair mechanisms.

Users can change the virtual officer’s gender, race, and language through the VOF. The importance of officer manipulation is apparent in, for instance, cases involving rape (use a same-gender officer) or nonnative speakers (speak in the eyewitness’s native language). However, for this preliminary evaluation, we didn’t change the virtual officer’s appearance or language.

References

We prescribed the nonverbal cues using key-frame techniques afforded by the Haptek motion-generation engine (www.haptek.com). We also used the motion-generation engine to evoke the appropriate nonverbal behavior on the basis of the scripted content. We used the Haptek lip-synch engine for lip synchronization, Dragon NaturallySpeaking 7.3 for speech recognition, and AT&T text-to-speech for speech utterance. A detailed description of the implementation is available elsewhere.\(^5\)

A Pentium IV 2.4-GHz Dell PC with an Nvidia GeForce4 Ti 4200 graphics card served as the graphics generator for the virtual human. The graphics were rendered with OpenGL and displayed on a 19-inch Dell Ultrasharp monitor.

We constructed the VOF iteratively on the basis of feedback from domain experts in computer science, eyewitness memory, and law enforcement.

**The Initial Version**

Initially, we outfitted Officer Garcia in a navy blue police uniform, including a badge, nameplate, two shoulder patches, and a tie. As Figure 1 shows, the background resembled a police office with a projection screen used for the lineup.

The guided conversation closely follows the procedures used in a large metropolitan police department. Specifically, Officer Garcia’s instructions conform to the North Carolina Actual Innocence Commission Recommendations for Eyewitness Identification (see the “Eyewitness Identification” sidebar) with dichotomous questioning (yes or no) to minimize false identifications.

The VOF also lets a real officer upload photos, which Officer Garcia displays sequentially rather than simultaneously—a recommended practice for reducing false identifications. Each time the victim recognizes a photo, Officer Garcia records an audio file of why the victim recognized that photo and the victim’s confidence in the selection. The virtual officer also logs the complete transcript and audio files.

**Iteration One**

The first iteration was a result of feedback from two students and four professors from the Department of Computer Science at the University of North Carolina, Charlotte, and one professor and expert in eyewitness memory from the university’s Department of Psychology. Given that this iteration involved domain experts, we opted for feedback based on first impressions instead of directed questioning. The participants raised the following questions and concerns:

- How will an eyewitness interpret gestures? Could they misconstrue these gestures as biased?
- What if the eyewitness wants to go back in the interaction or the virtual officer incorrectly recognizes a speech command?
- Will the virtual human be effective if the eyewitness is hysterical?
- Is a windows-and-mouse interface a fair comparison to a virtual human interface? Would a

![Figure 1. A sample interaction of the initial version of the virtual officer framework (VOF): Officer Garcia (a) says “hello” and waits for the eyewitness’s response, (b) conducts identification procedures, (c) asks simple yes-or-no questions throughout, (d) and starts the lineup. The eyewitness (e) identifies the photo and (f) records his or her identification confidence level. Garcia (g) continues through all six photos in the lineup and (h) concludes the identification procedure.](image-url)
comparison using touch screens, such as those used for testing at the North Carolina Department of Motor Vehicles, be a more interesting research question than a comparison to a windows-and-mouse interface? The participants also identified words or phrases that were nearly inaudible and could be misinterpreted. (Single-syllable words caused the most problems.)

The claim that a virtual officer could have a bias was surprisingly the most common comment on the system—the very complaint this system was designed to address. However, we had programmed many random gestures into the virtual officer to make him appear more lifelike. Consequently, because a quick blink of the eye, for example, could appear in conjunction with one suspect’s photo but not the others, the officer might appear to project a bias.

As a result of this potential bias and to make the interaction for each suspect consistent, we removed any movement not associated with breathing or talking. During the lineup, the virtual officer completely leaves the viewing area and the screen focuses on the photos (see as Figure 2). The “help” command remains on screen throughout the procedure.

Participants also raised the issue of option selection and error correction. Restricting the eyewitness to two voice commands couldn’t compensate for microphone noise and varying accents. We modified the system so that the eyewitness must confirm the selection before the VOF will continue. This modification also requires the eyewitness to think a moment longer about his or her decision.

A virtual officer’s ability to deal with a hysterical eyewitness is a realistic concern, given that eyewitness lineups are typically reserved for serious crimes. VOF alleviates this anxiety by

- requiring selection confirmation,
- easing into the virtual human interface through small talk and by explaining and demonstrating the process, and
- incorporating the voice command (“help”) to signal the need for a real officer.

Finally, participants suggested that a mouse, keyboard, and window interface isn’t a fair comparison to a multimodal virtual human interface. A touch screen is only one option; many other input modes are possible. The best comparison would be of the virtual officer with a real officer who’s blind to the suspect’s identity.

**Iteration Two**

We based the last iteration on the recommendations of a local police sergeant with little computer experience. He completed an evaluation and discussed with us his experience and concerns with eyewitness interviews. He made these observations:

- The virtual officer’s speech synthesis is choppier than human speech.
- Eyewitnesses won’t take a virtual human seriously. Instead, we should record a TV reporter-like

![Figure 2. A sample interaction of the first iteration of the VOF: Officer Garcia (a) says “hello,” (b, c) introduces the interface via an example “yes/no” question, and (d) informs the eyewitness that a real officer will return if the eyewitness says “help.” He then (e) continues through the procedures, (f) indicates that lineup photos will appear on the virtual projection screen, and (g) leaves the viewing area. After he leaves, (h) the lineup continues as in the initial version. We removed the gestures from this version.](image)
investigator and show the prerecorded video to the eyewitness.

- It was unclear how many photos would be shown and whether the process would continue until the eyewitness identified a photo.
- The instructions could be more concise, but the length was fine (approximately 8 minutes).
- The introduction was essential.
- The commands should stand out, and the background is distracting.
- Uniformed police officers don’t necessarily conduct the investigation, so a police uniform is unnecessary.

We can solve the virtual officer’s irregular-speech problem using human-recorded speech instead of speech synthesis. The framework’s extensible nature lets the virtual officer lip-sync to audio files, making this modification possible. Similarly, using video of real humans would certainly add the missing sense of realism and seriousness.

The procedure’s verboseness is unavoidable because the virtual human leads the conversation and the script must follow a police department’s policy or protocol. The sergeant confirmed that the virtual officer’s introduction to the interface was necessary to complete the procedure. So, although the police sergeant thought the script could be more concise, the opening dialog is important for creating rapport and understanding how to interact.

Figure 3 shows how we implemented his remaining suggestions, which were aesthetic. In this
iteration, we emphasized the commands, removed distractions (the office scene), and changed the virtual officer’s outfit. The identification procedures inform the witness that they’ll view six photos, and the photo number is indicated on each photo. This prevents the eyewitness from thinking that the photos will continue until they make an identification.

Experimental Study
Building on the lessons learned from our preliminary evaluations, we designed a user study to compare the virtual officer with a human investigator who’s blind to the suspect’s identity. Figure 4 shows the two types of investigators.

The study’s participants were 259 students (180 female and 79 male) from the psychology department subject pool at the University of North Carolina, Charlotte. Volunteers received course credit for their participation. Because of a computer malfunction, we lost 20 participants’ postexperiment questionnaires; however, we recorded their lineup identifications. The University of North Carolina, Charlotte’s Institutional Review Board approved the study.

Using crime-simulation methodology, we showed the students videotaped enactments of crimes. We then randomly assigned them to a virtual officer or a live human investigator to attempt identification from a lineup. We used two simulations showing different perpetrators performing the same crime. We also showed two types of lineups: half of the students viewed lineups containing the perpetrator; the other half viewed lineups in which we replaced the perpetrator’s photo with one of an innocent suspect. Thus, the lineups consisted of four conditions:

- virtual officer, perpetrator present;
- virtual officer, perpetrator absent;
- live human investigator, perpetrator present; and
- live human investigator, perpetrator absent.

We hypothesized that there would be no difference between having a virtual officer present the lineup and having a human officer present it. We used three measures to test the hypothesis: identification accuracy, ease-of-use of the identification procedures, and the emotional affect of the participant.

Procedure
Before beginning the experiment, we asked students to complete informed-consent forms and instructed them on the experimental procedures.

Phase 1. Each participant viewed one of two randomly assigned videotaped enactments. In both videotaped crimes, a male enters the scene (with his back to the camera); rummages through a purse, placing items in his pockets; turns toward the camera; and walks away.

Phase 2. Directly after the crime enactment, we randomly assigned each participant to one of the four conditions. In all cases, we showed the lineups sequentially, using a protocol established by a local police department. We warned each participant that the perpetrator might not be in the lineup. The perpetrator of one crime served as an innocent suspect for the other crime, and both lineups used the same fillers.

In the virtual officer condition, the virtual officer appeared on a 19-inch computer monitor, with speakers providing the audio. The virtual officer guided the participant through the identification procedures by asking him or her questions and responding appropriately. The system logged the transcript and audio responses to photos. In the live-investigator condition, the experimenter, sitting on the other side of the computer monitor, guided the participant through the identification procedure using the same script. The lineup photos were at the same resolution, and they faced the eyewitness, as Figure 4 shows.
The perpetrator-present lineup included the perpetrator and five fillers. In the perpetrator-absent lineup, a photo of an individual closely resembling the perpetrator replaced the perpetrator’s photo. We chose the photo order randomly each time, with the caveat that the perpetrator never appeared first, because research has shown that this leads to more accurate identifications.

**Postexperiment.** After a participant made an identification, we asked him or her to complete a questionnaire. This questionnaire contained rating scales and open-ended questions assessing

- basic demographic characteristics (age, gender, ethnicity, and occupation),
- experience with crime and computers,
- positive and negative affects related to the identification procedure,
- comfort with the identification procedure and its ease of use, and
- ratings and suggested improvements.

After the participant finished the questionnaire, we gave him or her a debriefing sheet describing the study. The entire experiment took approximately 30 minutes per participant.

**Data Analysis and Results**

We collapsed data from both crimes for this analysis. We performed chi-square tests to compare the identification accuracy between investigators (virtual versus live) and between lineups (perpetrator present versus perpetrator absent), with an estimated power to detect a medium-size effect \( w = 0.30 \) of 0.99. We also performed a chi-square test between the two mock crime videos to determine whether the perpetrator in one video was easier to identify than the other, with an estimated power to detect a medium-size effect \( w = 0.30 \) of 0.99.

We performed t-tests between investigators (live versus virtual) for questions regarding the identification procedures and the eyewitnesses’ emotional affect score, with an estimated power to detect a medium-size effect \( d = -0.50 \) of 0.98. We used a significance level of \( \alpha = 0.05 \).

**Identification accuracy.** Participants marked their identification for both the live and virtual conditions on a lineup form. They could choose photos 1 through 6 or none of the photos. We compared participants’ identification rates across conditions, as Table 1 shows.

The following results reflect eyewitnesses’ accuracy for perpetrator-present and perpetrator-absent lineups:

- For the perpetrator-present lineups, the percent of correct identifications obtained by the virtual officer \( (M = 31.1 \text{ percent}) \) and the live investigator \( (M = 19.2 \text{ percent}) \) didn’t differ significantly: \( \chi^2 (1, N = 139) = 1.339, p = 0.165, \phi^2 = 0.01 \). (The term \( \phi^2 \) is a measure of correlation between the two categorical variables, where 0 = no correlation and 1 = complete correlation.)
- For the perpetrator-absent lineups, the percent of correct rejections obtained by the virtual officer \( (M = 49.2 \text{ percent}) \) and the live investigator \( (M = 57.9 \text{ percent}) \) didn’t differ significantly: \( \chi^2 (1, N = 120) = 2.290, p = 0.092, \phi^2 = 0.02 \).
- For the perpetrator-absent lineups, the percent of false identifications of the innocent suspect obtained by the virtual officer \( (M = 7.9 \text{ percent}) \) and the live investigator \( (M = 7.0 \text{ percent}) \) didn’t differ significantly: \( \chi^2 (1, N = 120) = 0.04, p = 0.849 \).

Overall, the results indicate that the identification accuracy didn’t vary significantly on the basis of the type of investigator conducting the lineup.

**Mock-crime videos.** To ascertain whether the two mock-crime videos affected the identification results, we performed a chi-square test between the two. We found no statistical difference between the overall correct identifications as a result of the mock crime with perpetrator A \( (M = 48.1 \text{ percent}) \) and the mock crime with perpetrator B \( (M = 51.9 \text{ percent}) \).

<table>
<thead>
<tr>
<th>Eyewitness response</th>
<th>Perpetrator present</th>
<th>Perpetrator absent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Virtual officer (%)</td>
<td>Live investigator (%)</td>
</tr>
<tr>
<td>Correct identifications</td>
<td>31.1</td>
<td>19.2</td>
</tr>
<tr>
<td>Filler identifications</td>
<td>31.1</td>
<td>33.3</td>
</tr>
<tr>
<td>Misses</td>
<td>37.7</td>
<td>47.4</td>
</tr>
<tr>
<td>False identifications</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Correct rejections</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 1. Mean identification results in perpetrator-present and perpetrator-absent lineups. None of the tests were significant at the \( \alpha = 0.05 \) level.
Table 2. Results from the postexperiment questionnaire using a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Virtual officer</th>
<th>Live investigator</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>I found the identification procedure easy to complete.</td>
<td>4.70 1.72</td>
<td>4.36 1.93</td>
<td>1.42 0.158</td>
</tr>
<tr>
<td>I found the identification procedure to be awkward.</td>
<td>3.81 1.78</td>
<td>3.99 1.69</td>
<td>0.78 0.434</td>
</tr>
<tr>
<td>The identification procedure made me uncomfortable.</td>
<td>2.62 1.66</td>
<td>2.72 1.72</td>
<td>0.45 0.261</td>
</tr>
<tr>
<td>I took the identification procedure seriously.</td>
<td>5.35 1.69</td>
<td>5.58 1.78</td>
<td>1.13 0.434</td>
</tr>
<tr>
<td>The identification procedure seemed realistic, like a procedure that an eyewitness would follow.</td>
<td>4.56 1.64</td>
<td>4.59 1.73</td>
<td>1.42 0.158</td>
</tr>
<tr>
<td>For the most part, I forgot that the investigator was present and paid close attention to the lineup.</td>
<td>4.56 1.65</td>
<td>4.54 1.68</td>
<td>0.08 0.935</td>
</tr>
<tr>
<td>I found myself focusing more on the investigator than the people in the lineup.</td>
<td>3.01 1.60</td>
<td>2.51 1.47</td>
<td>2.60 0.016</td>
</tr>
<tr>
<td>I found the identification procedure confusing.</td>
<td>2.52 1.71</td>
<td>3.47 1.88</td>
<td>2.65 0.009</td>
</tr>
</tbody>
</table>

percent): $\chi^2 (1, N = 259) = 0.956, p = 0.328$. This is expected given the perpetrators’ similar appearance and the video’s low resolution.

**Responses to investigator procedures.** We gave participants eight statements about the identification procedures and asked them to respond using a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree). As Table 2 shows, six questions resulted in nonsignificant differences between investigators.

For the two last questions, however, a statistically significant difference exists between the ratings for the virtual and live investigators:

- For the statement, “I found myself focusing more on the investigator than the people in the lineup,” participant responses showed that they focused more on the virtual officer than the live investigator, with $\eta^2 = 0.02$. (The term $\eta^2$ means that the independent variable can explain or account for a certain percentage [2 percent in this case] of the variability in the dependent variable.)
- For the statement, “I found the identification procedure confusing,” responses showed that the procedures for the live investigator were more confusing than those for the virtual officer, with $\eta^2 = 0.02$.

That participants focused more on the virtual officer is surprising, given that Officer Garcia leaves the screen during the lineup. The voice synthesis might have been enough to create a distraction that wasn’t present in the live condition. However, it apparently didn’t impede the procedures.

Although we used the same scripted identification procedures for both investigator types, a human can stumble over words, get distracted, suffer from fatigue, and react to the eyewitness in a way that the virtual officer doesn’t. So, the virtual officer’s consistent presentation likely resulted in the procedures being less confusing.

**Emotional affect.** Positive affect corresponds to positive emotions such as excitement and interest; negative affect corresponds to negative emotions such as stress and anxiety. We asked participants in both the live and virtual conditions to rate their feelings and emotions on 10 positive affects (interested, excited, and so on) and 10 negative affects (distressed, upset, and so on) on a scale from 1 = very slightly or not at all to 5 = extremely. The combination of the 20 affects (positive affects minus negative affects) results in a positive and negative affect scale (Panas score that measures the participant’s overall mood. The Panas score for the virtual investigators ($M = 11.16, SD = 7.59$) and live investigators ($M = 10.91, SD = 7.81$) indicated no statistical difference: $t(231) = 0.249, p = 0.986, \eta^2 < 0.001$.

This result shows that the virtual officer’s emotional influence is akin to a live investigator’s. This is an important result because eyewitnesses might be anxious after witnessing a crime; however, it isn’t clear whether this result would transfer to eyewitnesses to a real crime.

**Responses to Officer Garcia.** Participants ($N = 115$) also rated the virtual officer on seven statements using a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree). Table 3 shows (see next page) the results.

The ratings tend to be greater than the neutral value (4), indicating an overall positive experience as opposed to indifference.

The values were lowest, although still above neutral, when we asked participants to compare a virtual officer with a real officer ($M = 4.03, SD = 1.50$) and with a computer lineup ($M = 4.28, SD = 1.70$). These ratings might suggest that a virtual officer...
Table 3. Participants’ responses to the virtual officer using a 7-point Likert scale (1 = strongly disagree and 7 = strongly agree).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Mean rating (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The virtual officer’s speech was clear.</td>
<td>5.70</td>
</tr>
<tr>
<td>The virtual officer’s response was appropriate.</td>
<td>5.51</td>
</tr>
<tr>
<td>The virtual officer recognized what I told him.</td>
<td>4.88</td>
</tr>
<tr>
<td>The virtual officer’s appearance was realistic.</td>
<td>5.24</td>
</tr>
<tr>
<td>The virtual officer’s behaviors were appropriate.</td>
<td>5.83</td>
</tr>
<tr>
<td>Results of lineups conducted with virtual officers are as valid as lineups conducted with real officers.</td>
<td>4.03</td>
</tr>
<tr>
<td>I would feel more comfortable conducting lineups with real officers than with virtual officers.</td>
<td>5.14</td>
</tr>
<tr>
<td>A lineup with a virtual officer would be more useful than a lineup that requires the eyewitness to use a computer keyboard.</td>
<td>4.28</td>
</tr>
</tbody>
</table>

Participants interacting with the virtual human did no worse than with a blind human experimenter in terms of identification accuracy, ease of completing the identification procedures, and emotional influence. This suggests that if an eyewitness can complete the recommended identification procedures with a human officer, they’ll be able to complete the same procedures with a virtual officer. To take this one step further, if we generalize identification procedures to a series of simple procedures, the results suggest that a virtual human will perform the procedures without impeding memory recall, decreasing usability, or having a negative affect relative to a human performing the same role.

Although we can debate the absolute percentages of correct and incorrect identifications from videotaped simulations versus real crimes, the point of the study isn’t to determine the absolute percentages of correct and incorrect identifications but to compare the human-administered and computer-administered procedures.2

This study’s results are limited to the virtual human, Officer Garcia; they might not generalize to other types of virtual characters. Additional research will need to examine a larger variety of virtual characters. Furthermore, we’ll need to replicate the study with a sample that includes a diverse group of people (beyond college students) to account for the wide range of eyewitnesses expected at a police station. We’re working with local law enforcement to deploy the virtual officer at a police station to perform a pilot field study.

Acknowledgments

We thank the police sergeant, students, and faculty who evaluated the virtual officer and provided the preliminary feedback. Several psychology undergraduates—Britt Nicole Sims, Brandi Jones, Emily Accrocco, Kallyn Falvo, Amber Debruhl, and Samantha Lear—participated in the formal user study. We also thank all the other students who volunteered for the study. We acknowledge Evan Suma and the reviewers for their feedback on this article. A US National Science Foundation award (SES 0552160) partially funded this work. This article is an extension of a preliminary evaluation of the virtual officer framework.7

References


Brent Daugherty is starting a business based on the Virtual Officer. As the cofounder, he’s working with local police departments to realize the integration of the Virtual Officer. His primary interest is transitioning breakthrough interface technologies to industry. Daugherty received his MS in computer science from the University of North Carolina, Charlotte, where his thesis was the development and evaluation of the Virtual Officer. Contact him at brent49@gmail.com.

Sabarish Babu is an assistant research scientist at the University of Iowa’s Hank Virtual Environments Laboratory. His research interests include virtual humans and embodied agents, simulator systems, 3D human-computer interaction, and virtual environments. Babu received his PhD in information technology from the University of North Carolina, Charlotte. Contact him at sabarish-babu@uiowa.edu.

Lori Van Wallenda is an associate professor at the University of North Carolina, Charlotte. Her research interests include cognition, human decision making, hypothesis-testing strategies, juror decisions, eyewitness and earwitness memory, and false memories. Van Wallenda received her PhD in cognitive psychology from Northwestern University. Contact her at lrvanwal@uncc.edu.

Brian Cutler is a professor in the Faculty of Criminology, Justice, and Policy Studies at the University of Ontario Institute of Technology. His research interests include the psychology of eyewitness memory. Cutler received his PhD in social psychology from the University of Wisconsin-Madison. He’s the editor in chief of Law and Human Behavior. Contact him at brian.cutler@uoit.ca.

Larry F. Hodges is the director of Clemson University’s School of Computing. His research interests include virtual reality and 3D user interface design. Hodges received his PhD in computer engineering from North Carolina State University. He’s on the editorial board of IEEE Transactions in Computer Graphics. Contact him at lfh@clemson.edu.