Introducing Animatronics to HCI: Extending Reality-Based Interaction

G. Michael Poor[1] and Robert J.K. Jacob[2]

[1] Computer Science, Bowling Green State University
Bowling Green, OH, 43403
[2] Computer Science, Tufts University
Medford, MA 02155
gmp@bgsu.edu, jacob@cs.tufts.edu

Abstract. As both software and hardware technologies have been improved during the past two decades, a number of interfaces have been developed by HCI researchers. As these researchers began to explore the next generation of interaction styles, it was inevitable that they use a lifelike robot (or animatronic) as the basis for interaction. However, the main use up to this point for animatronic technology had been "edutainment." Only recently was animatronic technology even considered for use as an interaction style. In this research, various interaction styles (conventional GUI, AR, 3D graphics, and introducing an animatronic user interface) were used to instruct users on a 3D construction task which was constant across the various styles. From this experiment the placement, if any, of animatronic technology in the reality-based interaction framework will become more apparent.

Keywords: Usability, Animatronics, Lifelike Robotics, Reality-Based Interaction, Interaction Styles.

1 Introduction

Since the early 1980s, the graphical user interface (GUI) has been the defacto user interface (UI) associated with computing. No matter which platform – Windows, Mac OS, or Linux – when people use a modern computer, as well as most cell phone and handheld devices, they are interacting with a GUI. However, because of the advances in computing technology in both hardware and software during the past few years, there has been increased interest in designing the next "standard" UI. From this interest, researchers have begun developing a broad range of new interaction techniques that split from the "window, icon, menu, pointing device" (WIMP) interaction style that is prevalent today [1]. These post-WIMP interaction styles have been defined by van Dam as "containing at least one interaction technique not dependent on classical 2D widgets such as menus and icons" [2]. Some examples of these post-WIMP interaction styles are tangible user interfaces, as defined by Ishii and Ullmer [3], context-aware interfaces, as defined by Schilit et al. [4]; and a number of other unique systems [5, 6, 7].

J.A. Jacko (Ed.): Human-Computer Interaction, Part II, HCII 2011, LNCS 6762, pp. 593–602, 2011.

Nevertheless, no one next-generation UI has been labeled the next default UI. Researchers at Tufts University have observed that these next-generation interaction styles include many similarities between them [1]; they all appear to use interaction techniques that incorporate reality-based movements, actions, and concepts. User's knowledge gained from experiences in the real world and people's interactions with objects can be incorporated in such a way that people will interact with the next generation of computers in a more natural and less mentally taxing context. This idea led to the framework called reality-based interaction (RBI). According to RBI, when an interaction style incorporates actions based on preexisting knowledge of the nondigital, everyday world, it is easier for users to learn to interact with a new system. The central claim of RBI is that the more people have used or encountered real-world phenomena, the easier it is for the users to call on that knowledge.

RBI theorists have attempted to take the next step in including and characterizing additional kinds of input as well as identifying the trend in the realm of human-computer interaction (HCI) as heading toward more reality-based interaction styles. This study examines whether manipulating the RBI characteristics of an interface has an effect on a user's performance. More specifically, is there a difference in the overall performance of a user's ability to complete a construction task delivered using four interaction styles with varying levels of reality-based interaction techniques?

By reviewing the four themes of reality-based interaction as identified in the RBI framework and identifying potential future interaction styles that would incorporate these four themes, the topic of animatronics and the potential use as an interaction style has become a major subject of interest. Animatronics exhibited many qualities that are reality-based in nature and a number of design process goals were identified:

1. Determine whether an animatronic interface could be used as a reality-based interaction style.
2. Determine how well the animatronic interface compares to other established interaction styles.

In this paper previous work related to the RBI framework and its history are discussed as well as an investigation of the current state of animatronics and the issues pertaining to the technology. This is followed by the description of the experiment; including condition design/creation and addition experimental details. Finally a review of the results is presented in addition to a discussion.

2 Related Works

In this experiment, there are three different types of interaction that the subjects could potentially use; each of which has their own history and bodies of research to explore. Highlights about each type follow.

2.1 Reality-Based Interaction

With the improvements in both software and hardware technology over the past few decades, there has been an increase in the number of new interfaces that have been developed by human-computer interaction researchers. At a cursory glance, these new interaction styles might not appear to have much in common. It has been proposed by

Jacob et al. [1] that "[these new interaction styles] share salient and important commonalities, which can help us understand, connect, and analyze them" (p. 1). One of the biggest commonalities is the increased ability to draw upon the users' preexisting knowledge of the real world, which is a trait that is present throughout all of these new interaction styles. By investigating these commonalities and the various changes that these interaction styles have gone through during their development, RBI would provide researchers with a "lens" through which they could gain insights for design and discover new avenues for research.

When a user attempts to learn a new system, a number of problems must be confronted. According to Norman et al. [8], two important problems will be faced: the "gulf of execution" and the "gulf of evaluation." The gulf of execution is the mental gulf that users must cross to turn their intentions into commands so that the interface can perform a task. The gulf of evaluation is the mental gulf that users must cross to understand the state of the system that results from the interface's feedback (after a task has been executed). According to Jacob et al., when four themes of reality are incorporated, emerging interaction styles attempt to bridge the gap between these two problems. The four themes are:

- *Naive Physics:* the concept that people have commonsense knowledge about the physical world.
- *Body Awareness:* people's awareness of their own physical bodies and their skills for controlling and coordinating their bodies.
- *Environment Awareness:* people's sense of their surroundings and their skills for negotiating, manipulating, and navigating within their environment.
- *Social Awareness:* people generally are aware of others in their environment and have skills for interacting with them.

The bridge that is created by these themes is the result of the users' familiarity with these themes and the ways in which they perform interactions that incorporate them without additional effort. The lack of required interpretation of intentions into the interface's language frees the user to perform the interactions automatically [10]. This type of automation also can be applied to the translation of the system's feedback, which allows the user's preexisting knowledge of the real world to be the predominant evaluation tool [9].

Incorporating interfaces with reality-based interaction characteristics so they completely mimic reality is not an optimal solution. According to this framework, RBI principals are important to consider during development, but there are times when this is not practical. Jacob et al. proposed that if a researcher gives up reality-based interaction, then it should be given up "explicitly and only in return for other desired qualities" (p. 5). Jacob et al. defined these desired qualities as:

- *Expressive Power:* users can perform a variety of tasks within the application domain.
- *Efficiency:* users can perform a task rapidly.
- *Versatility:* users can perform many tasks from different application domains.
- *Ergonomics:* users can perform a task without physical injury or fatigue.
- *Accessibility:* users with a variety of abilities can perform a task.
- *Practicality:* users find the system is practical to develop and produce.

It is through the “lens” of RBI that researchers have been able to analyze and compare various designs, bridge an assortment of problem gaps between seemingly unrelated research areas, and apply the lessons learned in the development of one interaction style to another style.

2.2 Research in Animatronics

Currently, common themes in animatronics research have robots learn from, work collaboratively with, or assist human users [11]. In the area of human-robot partnership, Cynthia Breazeal is attempting to escape from the traditional artificial intelligence goal of creating human-equivalent intelligence in technological systems. Instead she is attempting to create robots that bring value to a human-robot relationship in such a way that humans can appreciate robots for the ways that they enhance our lives and complement people's strengths and abilities. The focus of this avenue of research has been on either the human subject teaching the robot [12] or the robot and human working as a team [13]. The two-way dynamic has not been the animatronic being used as an interactive system to instruct the human subjects.

3 The Study

In order to systematically investigate whether the amount of reality-based interaction has an effect on a subject's ability to interact with a system, four interaction styles were developed, with varying levels of reality-based characteristics. Each of these conditions concerned the same task and the interactions were consistent. Additionally, the number of reality-based attributes used for each condition was augmented for each reality-based interaction style.

Four reality-based attributes were identified as being augmentable:

- *Motion:* utilizing or not utilizing a full range of motion, from static pictures to dynamic movement of the avatar, to the instruction delivery method.
- *Resolution:* the clarity with which the user can see the details of the instruction delivery method ranging from reality (no computer assistance), to 3D graphical rendering, to digital photos viewed on a computer monitor.
- *2D vs. 3D:* the presentation of the information can be varied depending on the instruction delivery method.
- *User-Controlled Depth of Field:* giving the user the ability to change their perspective of the instruction delivery method.

By selecting varying levels of these attributes, we were able to identify four varying levels of reality-based interaction that are used in this experiment. The GUI condition was chosen as the default interaction style because it is currently the most widely used interaction style and hence provides a good baseline. The decision to use the 3D condition was based on the trend in research to use virtual representations of avatars to interact with subjects. The AR condition was then included as an interaction style that had similar properties to the 3D condition, however it included additional reality-based characteristics that were not present in the 3D condition.

3.1 Condition Design

After a series of pilot tests were run using human actors as the instruction delivery system it was concluded that the animatronic condition could incorporate similar interactions. The pneumatically controlled animatronic character (Figure 3) was designed to be in a sitting position with moves that are considered the typical moves for a sitting museum-quality animatronic figure, providing enough body and facial movement to be considered realistic. In order to give the illusion of interactivity for the subject, the "Wizard of Oz" (WoZ) technique was implemented.

In the creation of the GUI condition (Figure 1), it was concluded that using the ZOOB pieces was not an equivalent task to the other three conditions. There was far more information about the specific step being supplied with their inclusion. By taking still images of the animatronic character throughout various points in its movements the interaction that was consistent with the other three conditions. For each step, at least a beginning position and an end position were shown in a comic-book-style format. If the move was so complex that it required additional information, a third picture was included that displayed the position of the character mid-move.

Finally, a 3D representation of the animatronic character was created that when viewed through a head-mounted display would give the appearance of a 3D computer-generated character in the real world for the AR condition. That same character would also be used in a desktop setting for the 3D condition (Figure 2).

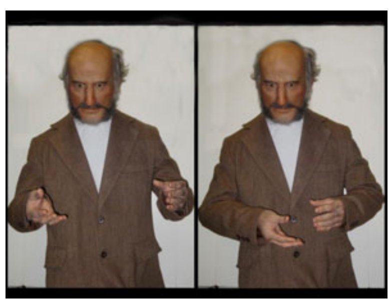

Fig. 1. GUI Condition

Fig. 2. 3D and AR Conditions

Fig. 3. Animatronic Condition

4 Experiment

The study was a between subjects design conducted with 80 undergraduate students randomly assigned, between the ages of 18 and 31, at Tufts University.

Each step was delivered to the subjects via their specific interaction style. To start, the subjects were given all of the individual ZOOB pieces that they would need. There were no distracter pieces included, so the subject would include all supplied pieces in the construction task when it was completed. The instructions were delivered to the subjects one step at a time, giving them as much or as little time as they required to complete the actions.

During step delivery, the subjects were allowed to interact with the pieces however they saw fit. Each step was either a legal connection between pieces that was inherent in the ZOOB, a rotation of a ZOOB piece, a tilt or angling of a ZOOB piece, or some combination of those three. The subsequent step was not provided until the subjects indicated that they had completed the step by placing the semi-finished construction on the yellow "X" and that step had been deemed correct by the researcher. Through the illusion created by the WoZ effect, it would appear to the subject that the system was reacting to their choices.

If a step was completed incorrectly, the interface would supply the subject with the corresponding error instruction, and then it would repeat the incorrect step so that the subject would attempt to perform the step correctly. The subject could attempt to correct the error at any time. An incorrect connection, piece, tilt, or rotation could hinder subsequent steps, so the error messages were continued until either the current step was performed correctly or the researcher deemed the error was "close enough." This means that the step that was performed was correct in regards to the information that was supplied, but might not be correct in terms of the precise step. The alternative option was that the subject could require manual assistance from the researcher.

4.1 Experimental Design

This study employed one independent variable, a covariate (spatial perception), and seven dependent variables. The independent variable was the varying levels of reality-based interfaces. The four different levels of reality-based interfaces included (1) the slideshow on a 2D monitor, (2) the 3D instructions delivered on a 2D monitor, (3) the 3D instructions delivered in an AR environment by an HMD, and (4) the 3D instructions delivered by an animatronic character. A behind-the-scenes person was in control of all the conditions, employing the WoZ technique. The WoZ technique directs the conditions to deliver each of the steps to the subjects, giving the illusion that the computer is adjusting its reactions to the subject's actions. As with all of the conditions, the subjects were not given any way of controlling the interface other than completing each step correctly.

For this experiment, performance was measured in terms of seven dependent variables. These dependent variables were divided into two groups for no other purpose but for ease of presentation. The first group included the number of times a step was repeated, the number of times human assistance was requested, and the total number of errors. The second group included the four types of errors that a subject could perform: incorrect connection, incorrect rotation, incorrect piece selection, and incorrect tilt or angle. A higher occurrence of these dependent variables means a lower performance by the subjects.

5 Results

Once all of the data was collected, it was found that there was a significant difference between interaction styles in all three of the first group's dependent variables. There were also significant differences between interaction styles in two of the dependent variables from the second group, incorrect piece and incorrect angle errors. Thus we are able to say that there is a significant difference in performance of a user's ability to complete a construction task delivered by four interaction styles with varying levels of reality-based interaction techniques. The specifics of the results follow.

The first test performed was a multiple analysis of covariance (MANCOVA) to investigate the effects of the independent variable and the covariate, condition and spatial ability, on a subject's performance as measured by the seven dependent variables. The results of the MANCOVA indicate that the covariate, spatial ability, relates significantly with the dependent variables (Wilks' Lambda, *F* = 4.388$; *p* < 0.001$). Controlling for spatial ability, the effect of the independent variable, interaction style, was still significant for the dependent variables (Wilks' Lambda, *F* = 2.173$; *df* = 3$; *p* < 0.005$). Once the MANCOVA was completed, an analysis of covariance (ANCOVA) was conducted on all seven of the dependent variables, taking into consideration both independent variables.

According to the results of the ANCOVA, a significant difference was found between all of the dependent variables with regards to spatial ability except incorrect piece errors and incorrect angle errors. When controlling for spatial ability, the results of the ANCOVA indicate that the effects of interaction style cause significant differences in the first three dependent variables: repeats (*sig* = 0.012$), outside help (*sig* = 0.020$), and errors (*sig* = 0.042$). Of the remaining four dependent variables, only the incorrect rotation error (*sig* = 0.026$) showed significance when we controlled for spatial ability and took interaction style into consideration.

Given the significant difference between conditions for number of dependent variables when controlling for spatial perception, a multivariate analysis of variance

(MANOVA) was performed to investigate the effect of condition on the seven dependent variables. The results of the MANOVA were significant (Wilks' Lambda, $F(3, 75) = 2.457\$$; $p = 0.001\$$), signifying that across the seven dependent measures, the independent variable of varying levels of reality had a significant effect. This means a between-subjects analysis of variance (ANOVA) can be done.

The ANOVA was performed on all seven of the dependent variables. This test determined whether the independent variable, varying levels of reality-based interaction styles, had any impact on each specific dependent variable. According to the results of the ANOVA, the independent variable had a significant impact on the number of repeat requests ($F(3, 75) = 5.064\$$; $p = 0.003\$$), the number of outside help requests ($F(3, 75) = 4.237\$$; $p = 0.008\$$), the number of errors committed ($F(3, 75) = 4.640\$$; $p = 0.005\$$), the number of incorrect piece errors ($F(3, 75) = 3.909\$$; $p = 0.012\$$), and the number of incorrect angle errors ($F(3, 75) = 4.577\$$; $p = 0.005\$$). Thus there is a significant difference between the four varying levels of reality-based interaction styles in terms of the aforementioned dependent variables. It must be noted, however, that the results of the ANOVA and the ANCOVA are not completely consistent in terms of the types of errors that were performed. It must be noted that the effect that was observed in that portion of the ANOVA could be due to the effect of the subject's spatial ability and the results have to be viewed while taking that into consideration.

6 Discussion

As was shown in the statistical results, there was a significant difference between interaction styles in all three of the first group's dependent variables: steps repeated, assistance requests, errors. There were also significant differences between interaction styles in two of the dependent variables from the second group: incorrect piece and incorrect angle errors. Therefore, given these results, we are able say that there is a significant difference in performance of a user's ability to complete the same construction task delivered by four interaction styles with varying levels of reality-based interaction techniques.

When delving further into the analysis of the overall experiment, there is evidence that users in the AR condition consistently performed much worse than those in other conditions. In both number of repeats and number of errors, the AR condition was significantly different than all of the other conditions. The outside help dependent variable also had a significant difference between the AR condition and the GUI condition. However, the other two conditions, although not significant, were approaching significance when compared to the AR condition.

When this experiment commenced, a formal ranking was not applied to the conditions; however, an informal ranking was conceived. It was concluded that the results would inform the researchers of the ranking after the experiment was completed. The resulting rankings did not coincide with the initial intuition of the researchers. It was believed that the GUI condition would perform the worst, followed by the 3D condition, then the AR condition, with the animatronic condition performing the best. These initial rankings stemmed from extensive dialogue with consultants and experts in HCI at Tufts, as well as experts in HCI and statistics at

Bowling Green State. The logic behind the original intuition was due to the apparent reality-based characteristics that are present in the technologies and varying amounts of the four themes identified in the RBI framework.

However, when realistic budgets and constraints were placed on the technologies, the quality of these technologies and their overall ability was hindered. The rankings that were conceived from the RBI framework at the outset of this experiment did not foresee the technological inhibitors that would be present due to the state of the art in animatronics, and the budget and time constraints. Even with budget, time, and technological inhibitors, the animatronic condition was able to perform at the same level as the top-performing technologies and statistically better than the AR condition.

7 Conclusion

We have presented a study of four interaction styles with varying levels of reality-based interaction techniques each being used to convey instruction for a 3D construction task. It was found that three of the conditions (GUI, 3D, and Animatonic) all performed in a roughly equivalent manner. However, the AR condition was found to perform statistically worse in regards to the dependent variables. In terms of the RBI framework, this was not the expected results. The higher levels of “reality-based” characteristics should have caused a different ranking in regards to performance.

Now why did this happen? One thought is that perhaps the interactions required with AR conditions were too overwhelming for the subjects. In other words, the subjects were perhaps not able to conceive the scope of interaction that was possible in the condition and became intimidated, consequently hindering their performance. Another thought was that even though the AR condition, by its nature, attempted to draw from real-world interactions, the uniqueness of the technology might have made that type of experience almost impossible, given the time that the subjects had to interact with it and become acclimated to it.

The other important point to note is that the animatronic condition performed as well as two interaction styles (3D and GUI) that are widely accepted and in everyday use. It was concluded that given advancement in technology and an increase in budget, the animatronic condition would be able to achieve the higher performance levels suggested by the RBI framework, and has the potential to perform better than these technologies in this type of task. The needed advancements in the technology are still a few years away, but animatronics will become an integral part of the growing robotic world.

References

1. Jacob, R.J., Girouard, A., Hirsheld, L.M., Horn, M.S., Shaer, O., Solovey, E.T., Zigelbaum, J.: Reality-based interaction: A framework for post-WIMP interfaces. In: CHI 2008: Proceeding of the Twenty-sixth Annual SIGCHI Conference on Human Factors in Computing Systems, pp. 201–210. ACM, New York (2008)
2. van Dam, A.: Post-WIMP user interfaces. Communications of the ACM 40, 63–67 (1997)

3. Ishii, H., Ullmer, B.: Tangible bits: towards seamless interfaces between people, bits and atoms. In: CHI 1997: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 234–241. ACM, New York (1997)
4. Schilit, B., Adams, N., Want, R.: Context-aware computing applications. In: Proceedings of Workshop on Mobile Computing Systems and Applications, pp. 85–90 (1994)
5. Boud, A.C., Baber, C., Steiner, S.J.: Virtual reality: A tool for assembly? Presence: Teleoperators & Virtual Environments 9, 486–496 (2000)
6. Iqbal, R., Sturm, J., Kulyk, O., Wang, J., Terken, J.: User-centred design and evaluation of ubiquitous services. In: SIGDOC 2005: Proceedings of the 23rd Annual International Conference on Design of Communication, pp. 138–145. ACM, New York (2005)
7. Tang, A., Owen, C., Biocca, F., Mou, W.: Comparative effectiveness of augmented reality in object assembly. In: CHI 2003: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 73–80. ACM, New York (2003)
8. Norman, D.A., Draper, S.W.: User Centered System Design: New Perspectives on Human-Computer Interaction. Lawrence Erlbaum Associates, Inc., Mahwah (1986)
9. Christou, G.: A Knowledge-Based Framework for the Description and Evaluation of Reality-Based Interaction, Ph.D. thesis, Tufts University (2007)
10. Logan, G.: Toward an instance theory of automatization. Psychological Review 95, 492–527 (1988)
11. Coradeschi, S., Ishiguro, H., Asada, M., Shapiro, S., Thielscher, M., Breazeal, C., Mataric, M., Ishida, H.: Human-inspired robots. IEEE Intelligent Systems 21, 74–85 (2006)
12. Lockerd, A., Breazeal, C.: Tutelage and socially guided robot learning. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2004, vol. 4 (2004)
13. Hoffman, G., Breazeal, C.: Effects of anticipatory action on human-robot teamwork effciency, fluency, and perception of team. In: Proceedings of 24 the ACM/IEEE International Conference on Human-robot Interaction, pp. 1–8. ACM, New York