How to improve the quality of human performance with natural user interfaces as a case study for augmented reality

Citation for published version (APA):
Rauterberg, G. W. M., Stebler, R., & Mauch, T. (1996). How to improve the quality of human performance with natural user interfaces as a case study for augmented reality. In A. Mital, H. Krueger, S. Kumar, M. Menozzi, & J. E. Fernandez (Eds.), *Advances in occupational ergonomics and safety I* (pp. 150-153). International Society for Occupational Ergonomics and Safety.

Document status and date:
Published: 01/01/1996

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 08. Feb. 2025

in: Advances in Occupational Ergonomics and Safety I (vol. 1; A. Mital, H. Krueger, S. Kumar, M. Menozzi & J.E. Fernandez, eds.), Cincinnati (Ohio USA): International Society for Occupational Ergonomics and Safety, 1996, pp. 150-153.

How to Improve the Quality of Human Performance with Natural User Interfaces as a Case Study for Augmented Reality

Matthias RAUTERBERG, Reto STEBLER & Thomas MAUCH

Work and Organization Psychology Unit, Swiss Federal Institute of Technology (ETH)
Nelkenstrasse 11, CH-8092 Zurich, Switzerland, Tel: +41-1-6327082

Abstract. The embodiment of computers in the work place has had a tremendous impact on the field of human-computer interaction. Graphic displays are everywhere, the desktop workstations define the frontier between the computer world and the real world. Users spend a lot of time and energy to transfer information between those two worlds. This could be reduced by better integrating the virtual world of the computer with the real world of the user. The most promising approach to this integration is *Augmented Reality* (AR). The expected success of this approach lies in its ability to build on fundamental human skills: namely, to interact with real world objects! To compare the advantages and disadvantages of a "Graspable User Interface" a field study was carried out. We investigated the effects of four different interaction techniques (command language interface, mouse interface, touch screen interface, and a DigitalDesk) with two different methods: (a) an inquiry with a questionnaire, and (b) a field test for public and anonymous use. (a) *Results of the inquiry:* 304 visitors rated the usability of all four stations on a bipolar scale. The touch screen station was rated as the easiest to use interaction technique, followed by the mouse and DigitalDesk interface; the "tail-light" was the command language interface. (b) *Results of the field test:* The analysis of 9'006 automatically recorded contacts to one of the four stations shows that the highest chance to win the game could be observed for the DigitalDesk interface technique. We conclude that the DigitalDesk as a representative example for AR is a promising candidate for the next user interface generation.

1. Introduction

Today several dialogue techniques are developed and in usage. The following dialogue techniques and dialogue objects can be distinguished with regard to traditional user interfaces: command language, function key, menu selection, icon, and window. These techniques can be summarised into three different interaction styles: *command language, menu selection,* and *direct manipulation.* In all these traditional interaction styles the user can not mix real world objects with virtual objects in the *same* interface space. They also do not take into consideration the enormous potential of human hands and fingers to interact with real and virtual world objects. This aspect was one of the basic ideas to develop data gloves and data suits for interactions in a virtual reality system (VR). The other basic idea, to realise VR systems, was the 3D output capabilities in the usage of head mounted displays. We restrict the notion of "virtual reality system" in the context of this paper to systems with head mounted display and data gloves or suits. In these VR applications the user has to leave his natural physical and social environment and to immerse in the simulated world. The following two unsolved problems are important: (1) how to simulate tactile and haptic feedback, and (2) how to overcome the social isolation for collaborative tasks. The effect, that the social nearness between real persons is of tremendous importance for collaboration, was investigated and shown in [3]. Enthusiasts of the VR approach believe "that all constrains of the real world can be overcome in VR, and physical tools can be made obsolete by more flexible, virtual alternatives" ([5] p. 87-88).

There is another important direction for new interface technology: augmented reality (AR) or ubiquitous computing, resp. We are looking for a realization of a user interface where the user can control the human-computer interaction by his hands dealing with real and virtual objects in the same interface space. One of the first systems with such a user interface was the DigitalDesk of Wellner [5]. Inspired by the ideas of Newman and Wellner [2], we were interested in a way to test empirically the advantages or disadvantages of the DigitalDesk in comparison with established interaction techniques. The DigitalDesk has the following three important features: (1) it projects electronic images (virtual objects) down onto the desk and onto real objects, (2) it responds to interaction with real objects (e.g., pens or bare fingers: hence *Digital*Desk), and (3) it can interpret the scene on an appropriate semantic level (e.g., read paper documents placed on the desk; cf. [5]).

If we generalize the DigitalDesk approach to augment real-world objects, then it leads us directly to "Graspable User Interfaces" (cf. [1]). We follow the argumentation of Fitzmaurice, Ishii and Buxton [1] that a Graspable User Interface has the following advantages: "(1) It encourages two handed interactions; … (3) allows for more parallel input specification by the user, … (4) leverages off of our well developed skills … for physical object manipulations; … (6) facilitates interactions by making interface elements more 'direct' and more 'manipulable' by using physical artifacts; … (9) affords multi-person, collaborative use" ([1] p.443). These arguments sound convincing. But, all the published approaches--caused by the complexity of the implemented tasks--are research prototypes, so that only the experiences of preliminary user testings are reported (cf. [1], [2]).

2. System description

To run a laboratory investigation or a field study we need a fast, reliable and robust implementation of the whole system. First, we decided to minimize the task complexity and to restrict the user's action space. For public use a simple computer game seems to be best. We implemented the computer game "Go-bang" (five chips in a row win). The user has to play the game by moving a *real* chip on the *virtual* playing field. To compare this dialog technique with the most established techniques we implemented the same game algorithm on three other stations with (1) a command language, (2) a mouse, and (3) a touch screen interface.

Command interface (CI): This station run on a 386er PC with a color screen (17 inch) in an upright position. The user has to enter the co-ordinates of the desired place of a playing field with 12 by 12 positions (e.g. A1, L12). To start a new or to cancel a game she or he has to enter the command NEW. The internal state of the algorithm is presented as text in a special output field (e.g. "Make the next move").

Mouse interface (MI): This interface run on a 386er PC with a color screen (17 inch) in an upright position. To move the user has to click with the mouse on the desired place. To start a new or to cancel a running game she or he has to click on the button NEW. The internal state of the algorithm is presented as text in a special popup window.

Touch screen interface (TI): This station run on 386er PC with a color touch screen (21 inch) in an inclined position of 30 degrees. To make a move the user has to touch with a finger the desired place. To start a new or to cancel a running game she or he has to touch the button NEW. The virtual player was shown on a second colored screen (17 inch) served by a second 386er PC in a client-server architecture.

DigitalDesk (DD): This station was completely realized in C^{++} on standard hardware components: (1) a Pentium PC, (2) an overhead projector of high luminous intensity and the projection panel GehaVision™, (3) a high resolution video camera, and (4) the video board MovieMachinePro™. For the virtual player a second 386er PC was connected in a client-server architecture. A user has to make a move by putting a *real* chip on the desired place of the *virtual* playing field. The computer's output is the projection of a *virtual* chip on the desk. If a user wanted to cancel or restart a game, then he or she had to press the real NEW button in the front of the station.

The Virtual Player: The CI and MI have their output screens in an upright position. This upright position makes it impossible to give an opponent--in the metaphor of a game--an individual representation. The classical solution is, to give feedback about the machine's internal states as text or graphics in defined areas on the screen. This solution leads always to a partition of the screen into a working area (e.g., the playing field) and the feedback area. This superimposing of qualitatively different feedback's in the same output space can be overcome, if we use an additional output device. This can be done, if we separate the working area (the playing field) from the feedback area and if we change the upright position of the working area to a horizontal position. Therefore we composed the system of a flat table to project the playing field onto and of a second screen to present the virtual player. All--from the user's point of view--important internal states of the game algorithm were presented by the virtual player with six different facial expressions and a corresponding sound (partially animated comic strip pictures: "Reasoning" = a face with a balloon of animated turning wheels and a machine-like sound, "Waiting for the next move" = a yawning face with a corresponding sound, "Initial state" = a face with blinking eyes, "Incorrect move" = an angry facial expression with an indignant cry, "Be the winner" = a happy facial expression with an arrogant laughter, "Be the loser" = a shrinking face with a disappointed cry). We call these six comic strip pictures "emoticons" (acronym for "emotional icons"). All emoticons of the virtual player are shown on a second color screen (17 inch).

3. Empirical Validation of the Augmented Reality Approach

To present our four dialog techniques to a broad population of heterogeneous users, four special stations were constructed. All stations were presented in a central exhibition area during five days at the largest computer fair of Switzerland in September 1995 in Basel (CH). The official number of visitors was approximately 70'000. Most of these people passed the exhibition area and many of them came into close contact with one of the stations (e.g., playing at one of the stations or observing other people playing).

3.1. Inquiry with a questionnaire

With an additional questionnaire we got some personal information's of several users of at least one of our four stations. The stand personal was instructed to request users to answer the questionnaire. It was especially necessary to ask women, because they behaved very reserved.

Subjects: The questionnaire was answered by 304 visitors (61 females, 243 males, 5 without personal data). The average age of the women was 31 ± 13 years, and of the men 30 ± 14 years (T-Test: $p \le .724$). *There* was a significant gender difference in computer experience; the computer experience of men was higher than the experience of women (men: 58 ± 22 mm; women: 48 ± 23 mm; T-Test: $p \leq .002$).

Dependent Measures: The questionnaire consisted of two parts: (a) personal data (age in years, gender, computer experience in form of a bipolar rating scale ["no experience" … "expert"], and (b) a bipolar rating scale to estimate the "usability" ["very easy to use" … "very difficulty to use"] for each of the four interfaces. For each of the four interfaces the following five aspects were asked with a multiple choice question: "did you play", "did you loose", "did you win", "did you play draw", "did you cancel". The four *dependent measures* are the number of millimetres of the user's marking on the bipolar rating scale ["very easy to use": 0 … 70 mm: "very difficult to use"]. As a *control variable* the "computer experience" was measured in millimetre on the corresponding bipolar rating scale ["no experience": 0 … 90 mm: "expert"]. We also differentiated among "real station contact" and "only observer status".

Results: All persons, which answered that they had no real contact to one of the stations, had a significant higher amount of computer experience ("no contact": 70 ± 17 mm, N = 10; "with contact": 55 ± 23 mm, N = 276; T-Test: $p \leq .04$). We could not find a significant influence among "no contact" and "contact" people to the estimation of the usability. The usability rating was best for the touch screen interface (median = 3; mean rank = 1.9), followed by the mouse interface (median = 6; mean rank = 2.4) and the DigitalDesk (median = 9; mean rank = 2.5). The "tail-light" was the command language interface (median = 17; mean rank = 3.3). These differences are significant (Friedman Test, df = 3, Chi^2 corrected for ties = 89.1, $p \leq .0001$). One very important result was a significant correlation between "age" and "DigitalDesk's usability". This correlation means that older people prefer significantly more a graspable user interface in form of the DigitalDesk than younger people. ($R = -0.202$, $N = 179$, $p \leq .006$). All other correlation's are not significant.

Discussion of the inquiry: To carry out an inquiry with a questionnaire only for scientific purposes in the context of a commercial fair was more difficulty than we expected. The only motivation for the user to fill out such a questionnaire--in contrast to all the lotteries of the commercial issuers around us with their very attractive prizes--was to bring in his or her personal opinion into an scientific research process. This argumentation was the most convincing reason to participate. Overall we got more filled out questionnaires than we expected, but much less than observed visitors at one of the four stations. To increase the number of answered questionnaires the stand personal has to be active: to go to the visitors and to ask them for participation. We could find two main results: (1) the touch screen interface was estimated as the easiest to use, and (2) the significant correlation between age and the usability ratings for the DigitalDesk. If we assume, that the age structure of the populations in all high industrialised countries will change in the next two or three decades, then this result will be of tremendous importance for the development of modern computer technology! To compare the results of the user's subjective ratings with their concrete performances--as an objective measure--we analysed the outcome of all played games. To do this we carried out a field study.

3.2. Field study

All four stations were ready for unconstrained public and anonymous use during five days. The stand personal was instructed to hold back with help. To test the effect of non-verbal feedback we realized a virtual player on a second screen for the touch screen and the DigitalDesk station.

Subjects: All passing visitors had unconstrained access to all stations. It is difficult to get a valid picture of all users in a situation of public and anonymous use. We can assume that the results of the inquiry give us representative information about the gender and age structure of our visitor population.

Dependent Measures: The *dependent measure* is the "frequency of winnings" for each station separately. This frequency can be interpreted as a measure of the "quality" of the task result. The "frequency of winnings" includes all games where the computer would win in the next move, but in this state the game was canceled by the user to avoid losing. We call these games "quasi winnings". The frequency--of real winnings and of quasi winnings--is a reliable and valid measure to compare precisely the different interaction techniques on the performance level. As a *control measure,* the average computing time per move was recorded and analysed.

Results: We analysed our data with the statistic tool StatView (vers. 4.02). We recorded automatically 9'006 contacts with in all 96'739 moves: 1'128 contacts with CI, 3'645 with MI, 2'881 with TI, and 1'352 with DD. In all we observed 3'801 completed games. Each visitor had three different classes of behaviour to react to each station: (1) to pass; (2) to stand still and to observe what happens; and (3) to interact with one of the stations. For the last class of behaviour we could differentiate five separate actions: (A1) to start a new game even if the system is in the initial state; (A2) to cancel a game that was started --either by himself or by another user before; (A3) to finish a game with a draw; (A4) to finish a game with a winning result, either by the user or by the computer; and (A5) to finish a game with a "quasi winning" situation: a user canceled before the computer could win in the next move. These five actions are not independent of the interaction technique ($Chi^2 = 420$, $p \leq .0001$). In the TI and DD condition the two actions A1 and A2 could be significantly observed more often than expected (observed frequencies: A1 x TI: 60%; A1 x DD: 54%; A2 x TI: 4%; A2 x DD: 13%). On the other side, in the CI and MI condition the two actions A4 and A5 were more often observed than expected (observed frequencies: A4 x CI: 38%; A4 x MI: 41%; A5 x CI: 10%; A5 x MI: 7%). This result does not mean that the user had a better chance to win in the CI or MI condition; it means only that for both conditions a game was more often completed, or nearly completed (the "quasi winnings"). This result can be interpreted in the way that a significant number of users had serious difficulties to detect that the system was in its initial state: the corresponding facial expression of the virtual player was not clear enough.

To answer the question in which condition the user had a better chance to win, we have to take a closer look to all winning games. The user's chance to win a game was 11% in all (428 user's winnings divided by 3'801 winning games overall). *First,* we compared CI and MI, both interfaces without emoticons. If in one of the contingency table's cell the observed frequency is larger than the expected frequency, and the p-value of the corresponding Chi^2 test is less than .05, then the user or the computer has a higher chance to win. The main result of this analysis is, that users with CI (13% winning chance) have a significant higher chance to win than with MI (8% winning chance; $\text{Chi}^2 = 9.2$, $p \leq .002$). *Second*, we compared both interfaces with emoticons: TI and DD. The main result is, that users with DD (20% winning chance) have a significant higher chance to win than with TI (11% winning chance; $\text{Chi}^2 = 20.3$, $p \leq .0001$). This result is *not* caused by different computing times: $\text{TI}_{\text{time}} = 55 \pm 32$ s, $\text{DD}_{\text{time}} = 56 \pm 31$ s, (ANOVA, N = 1488, df = 1, F = 1.68, $p \leq .195$). Therefore we can not assume that the DD interaction technique performs best because an extended computing time gives the user more time to think about the next move than in the TI condition. *Third*, we compared CI and DD to find out which one is the best. The result is, that users with DD have still a significant higher chance to win ($\text{Chi}^2 = 9.2$, $p \leq .002$). Fourth, to estimate the effect of emoticons we compared the two interfaces *without* emoticons (CI plus MI: 9% winning chance) with the two interfaces *with* emoticons (TI plus DD: 14% winning chance). The result is, that users with emoticons have a significant higher chance to win ($\text{Chi}^2 = 19.9$, $p \leq .0001$).

Discussion of the field study: The comparison among the four interface techniques is not quite fair, because the result of this comparison is superimposed by the effect of the virtual player's emoticons of the TI and DD stations. Both effects superimpose each other: (1) the effect of the four different interaction techniques, and (2) the effect of non-verbal feedback (the virtual player's emoticons). So, we have to take this effect into account. To differentiate between the influences of both effects on the game outcome, we calculate the measure "user's chance to win". The global average is 11%, so we have to divide the 9% advantage of the DD technique (= 20% with DD minus 11% global average) into two parts: (a) the emoticon effect of 3% (14% with MI plus DD minus 11% global average), and (b) the interaction technique effect of 6% (= 20% with DD minus 11% global average minus 3% emoticon effect). Overall, the DD interaction technique--adjusted by the emoticon effect--guarantees a significant increase of the user's performance.

4. General Discussion and Conclusion

Informal observations during the fair show that several users of the DD have difficulties to use this dialog technique without any kind of "starting" help: (1) to make the first move and to wait a moment, (2) not to leave one of the hands in the video controlled area; but, if a user overcomes these obstacles she or he was able to play without any further help. Now we can discuss the introduced advantages of graspable interfaces (see [1]). In the DigitalDesk interface we could observe two handed interactions; there were no serious problems to move and remove chips with both hands. These two handed interactions enable more parallel input actions than with traditional interaction techniques. The youngest DigitalDesk player was a four-year old girl carried by her father. This girl--well advised by her father--could interact without any serious problems. This easy to use effect is caused by the transfer of the very early and well-developed skills to manipulate physical artifacts. We could also observe that users came together in larger groups--up to three and four persons--to collaborate and to discuss how to play and to move next. In one situation a circa ten years old boy was very proud to present his family that he was able to win. Emoticons seem to be attractive especially for women. We could observe that several women tried to persuade their male partner to come to one of the emoticon stations, and to start playing. Most of male partners could be persuaded to start the game and to demonstrate their girl friends that it is totally harmless. At this moment the woman took over the leading role and played the game. This effect is in congruence with the criteria of "enjoyable interaction" [4]. This field study gives us a strong empirical evidence that a Graspable User Interface with non-verbal feedback in form of emoticons has serious advantages over the traditional dialog techniques.

References

[1] Fitzmaurice, G., Ishii, H. & Buxton W. (1995) Bricks: Laying the foundations for graspable user interfaces. In *Proc. of CHI'95*, ACM, pp. 442-449.

[2] Newman, W. & Wellner, P. (1992) A desk supporting computer-based interaction with paper documents. In *Proc. of CHI'92*, ACM, pp. 587-592.

[3] Rauterberg, M., Dätwyler, M. & Sperisen, M. (1995) The shared social space as a basic factor for the design of group-ware. In: K. Brunnstein & P. Sint (Eds.) Intellectual Property Rights and New Technologies. (Schriftenreihe der Österreichischen Computer Gesellschaft Vol. 82, pp. 176-181), Oldenbourg.

[4] Tamura, H. (1995) Human information technology for living oriented innovation. In: Y. Anzai, K. Ogawa & H. Mori (Eds.) Symbiosis of Human and Artifact: Human and Social Aspects of Human-Computer Interaction. (Advances in Human Factors/Ergonomics, Vol. 20A, pp. 121-124), Elsevier.

[5] Wellner, P. (1993) Interacting with paper on the DigitalDesk. In *Com. of the ACM.* 36(7), pp. 86-96.