

Results of the VISA-comp evaluation

Citation for published version (APA):

Waterham, R. P., Beumer, J. J., Klimbie, P. B., & Poll, L. H. D. (1993). *Results of the VISA-comp evaluation*. (IPO-Rapport; Vol. 958). Instituut voor Perceptie Onderzoek (IPO).

Document status and date:

Published: 17/12/1993

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.

Institute for Perception Research
P.O. Box 513 - 5600 MB Eindhoven

RPW/rpw 93/16
17.12.1993

Rapport no. 958

Results of the VISA-comp
evaluation

R.P. Waterham, J.J. Beumer,
P.B. Klimbie, L.H.D. Poll

1. Introduction1
2. Time measurements1
3. Different kinds of errors2
4. Questionnaire4
4.1 Multiple choice questions4
4.2 Evaluation of the multiple choice questions.5
4.3 Open questions.5
5. Suggestions for improvements due to the experiment.6
5.1 General improvements:.6
5.2 Improvements to the exploration mode and the user mode..7
6. Suggestions for improvements beyond the experiment.7
7 Conclusions7
Literature8

1. Introduction

In report 912 (Guidelines for evaluating the VISA-comp system) we discussed how the VISA-comp system should be evaluated. In this report we will discuss the results of this evaluation.

This report is divided in several sections. The section 'Time measurements' will give some results of time measurements we made of the experiment. The section 'Different kinds of errors' will describe the errors made by the subjects during the experiment due to aspects of the VISA-comp system. In the section 'Questionnaire' the answers of the subjects on questions with and without scale are discussed. In 'Suggestions for improvements' we will discuss some improvements that the subjects mentioned, and in 'Other suggestions for improvements' we will present some improvements which can be made according to our impressions and literature. In the last section 'Conclusion' we will summarize the results of the evaluation. Note that this evaluation is done by using 10 subjects in the Netherlands, (two pilot experiments and eight real experiments). The evaluation performed by Sensory Visionaid is not included because we did not receive the results of them in time.

2. Time measurements

We made some time measurements to see how much time the subjects needed to perform the tasks listed in report 912.

During the experiments the order in which subjects received their tasks were mixed incorrectly, resulting in the fact that the tasks "find and execute Word Perfect" and "make a new file" were only performed in the User Mode and "execute Solitaire" and "find the option save settings on exit" were only performed in the Exploration Mode.

Table 1 shows the time the subjects needed to perform the following tasks:

Task 1: "Go to the program 'Wordperfect' and start the program. When you are sure that you reached your goal, return to the 'Program Manager'."

Task 2: "Go to the game 'Solitaire'. When you are sure that you have reached your goal, return to the 'Program Manager'."

Task 3: "Look for the possibility to create a new file. When you are sure that you have reached your goal, return to the 'Program Manager'."

Task 4: "Look for the option 'Save settings on exit'. When you are sure that you have reached your goal, return to the 'Program Manager'."

In table 1, column 1 the results are presented for the first part of the task (indicated by 'search') being the the part where a program or option has to be found. In the second column the results are presented for the second part of the task (indicated by 'return') being the the part where the subject has to return to the Program Manager. In the third column the total scores are presented. Each column is separated in the column 'mean' being the time needed to perform the task averaged over the subjects and in the column 'std' being the standard deviation of the mean time. In task 4 the system returns to the program manager automatically so only a total time is presented.

Table 1: time needed to perform a task.

	search		return		total	
task	mean	std	mean	std	mean	std
1	3m00	1m01	2m24	1m13	5m24	1m49
2	6m06	3m19	2m32	1m08	8m38	3m20
3	5m25	1m14	1m55	1m39	7m20	2m02
4					4m07	1m23
					-----	-----
total					25m29	5m04

Roughly it can be said that both versions (user-mode and exploration-mode) do not differ in time to complete the tasks. The results from the experiments in England will have to confirm this, however, before explicit statements can be made. Furthermore this experiment was not perfect with respect to randomisation, so these results can only give an indication.

3. Different kinds of errors

This section describes the errors made by the subjects during the experiments. We have classified the causes presumably responsible for the errors into 3 categories. With respect to these 3 categories the errors and their appearance per task are given per subject.

The presumed causes for errors are classified into 3 categories being:

1. Bugs of the program. As it appeared the software was not yet bug-free. These causes are sub-divided into 4 categories.

The mousepointer was not always operated correctly by the VISA-comp system (B1). Although the presentation of an object seemed in order, the VISA-comp system activated (double-clicked the mouse) at the wrong position.

The OCR program failed to function (B2). In some situations the OCR analysis did not finish, resulting in a necessary re-boot of the VISA-comp system.

The program crashes drastically when the left TASO-slider is used (B3). After activating several program groups and or programs, the left slider can not be used without severely crashing the system.

The VISA-comp system gives the wrong contents of a titlebar (B4). When returning to the Program Manager after executing an application program, the system incorrectly gives the Program Manager as the title of a (still open) program group window.

2. Ergonomic imperfections. These are aspects of the system that are not yet or not yet good enough implemented in the system. We came across two of these imperfections.

The generally considered bad quality of the synthetic speech (E1). Many subjects mentioned the low intelligibility of the synthetic speech. These problems were generally solved by the experimenter repeating the spoken message. In some occasions, however, time was lost (an error occurred) because the subject interpreted the message incorrectly and acted before the experimenter could correct this.

The lack of a 'help' function for the use of the VISA-comp system (E2). No 'help' function is available, which caused errors in the experiments because subjects were did not know how to operate the system. These problems also were solved by the experimenter giving the subjects advice about the operation of the system.

3. Unacquaintedness with Windows (O). The errors are due to the fact that our subjects are not familiar with Windows and are not due any aspect of the VISA-comp system. In general the errors were corrected by the experimenter, but sometimes they resulted in time-loss because it was unclear which aspect of Windows was not clear to the subject.

In table 2 the different kind of errors are presented (vertical) and the tasks in which they are made (horizontal). The table presents the number of appearances of these errors as well as a total over the experiment ('total') and the errors averaged over the subjects ('mean').

Table 2: Errors made during the experiment.

	tasks				total	mean
	1	2	3	4		
B1	2	0	2	1	5	0.6
B2	0	2	0	2	4	0.5
B3	1	1	0	0	2	0.3
B4	0	1	0	1	2	0.3
E1	13	11	22	9	55	7
E2	14	7	5	6	32	4
O	16	10	12	3	41	5

The errors B1, B2, B3 and B4 are bugs (see previous classification) and will not be discussed here, because these bugs do not need to be discussed, but should be fixed. However, it should be noticed that these bugs are very confusing for the subjects and may have influenced the results of their performance.

Also the unacquaintedness with Windows could have influenced the results of the performance of the subjects, but will not be discussed here as we consider this to be an integral part of the system. Furthermore, once you are familiar with the use of the VISA-comp system you can use the windows 'help' facilities to learn about the windows philosophy and operating aspects.

The important errors made by the subjects are those that are caused by the ergonomic imperfections of the VISA-comp system. In the experiments we have seen that the low

intelligibility of the speech synthesizer resulted by the subjects in a need for extra information. With an average of 7 times per experiment, this is a serious imperfection. Also the lack of a 'help' function in and about the VISA-comp system, causing errors with an average of 4 times per experiment, is a serious imperfection.

4. Questionnaire

This section describes the results of the questionnaire that was filled in after the experiment. These questionnaire was divided into 2 categories of questions, one containing all questions that were multiple choice and one containing all open questions.

4.1 Multiple choice questions

In total 22 questions were asked, 11 about the user mode and the same 11 about the exploration mode. The questions are listed in report 912 and the numbers in the columns are the percentages that the answer is given (being (number of times the answer is given)/(number of subjects)*100%). No answer was registered for question 5 for subject 3 in the exploration mode so here the scores are no multiple of 12.5.

Table 3: Scores of the multiple choice questions

A= never, B= some of the time, C= most of the time and D= always.

User mode:

	A	B	C	D
1	0	0	62.5	37.5
2	0	12.5	75	12.5
3	0	12.5	50	37.5
4	0	12.5	12.5	75
5	0	12.5	25	62.5
6	37.5	62.5	0	0
7	37.5	62.5	0	0
8	0	12.5	25	62.5
9	25	37.5	25	12.5
10	0	0	50	50
11	25	12.5	12.5	50

Exploration mode:

	A	B	C	D
1	0	12.5	50	37.5
2	0	37.5	50	12.5
3	0	12.5	50	37.5
4	25	37.5	25	12.5
5	0	28.6	42.8	28.6
6	25	75	0	0
7	25	75	0	0
8	0	0	75	25
9	37.5	25	37.5	0
10	0	12.5	50	37.5
11	37.5	0	12.5	50

4.2 Evaluation of the multiple choice questions

Before we start discussing the results from the multiple choice questions it is vital to remark that these results have to be combined with the results from SEV in order to get statistically relevant results. However we can draw some conclusions given that the results from England point in the same direction.

The subjects were moderately positive about how they could access information on the screen (i.e. they could find the information they wanted). During the experiments the subjects had some problems. A cause for problems was the low intelligibility of the synthetic speech. A problem also was the lack of information about screen changes. After selecting or activating an object it was unclear what changes would occur, so in fact the whole screen had to be examined again.

A preliminary conclusion can also be that our subjects understood the structure of the information better in the user mode.

4.3 Open questions

In this section the open questions, that were asked after the experiment, are discussed. The questions can be found in report 912.

The subjects mentioned some good aspects of the system. Most subjects told that they liked the fact that the window environment is accessible to them now. Furthermore one subject told he liked the sliders: "In the explorative mode I can read all information presented on a line with the vertical slider, while I can read spatial information with the horizontal slider". And one subject (partially sighted) told that he liked the system because he did not need to use his eyes, (he got very tired using his eyes). Also some remarks were made with respect to the Windows environment: "A clear structure of the information presented on screen", "everytime the same position is used for e.g. the closebutton" and "you can enter the position you want step by step".

Bad aspects of the system, the subjects mentioned most of the time, were that they could not understand the speech module very well. One subject mentioned he would like to search with one slider (see 'suggestions for improvements due to the experiment'). It was also

mentioned that the use of the left slider is very tiring for the left arm. One subject wanted to interrupt the speech module, which is not possible.

Confusing and irritating aspects of the system were caused by unfamiliar terminology (e.g. 'listbox' was not understood) and the sound of the speech module. Furthermore in the exploration mode the objects were difficult to distinguish from each other (again: see 'suggestions for improvements due to the experiment').

The changes that should be made to improve the VISA-comp system, suggested by the subjects, were most of the time a better speech module. Furthermore a key that can be pressed when the following (or last) item has to be searched was mentioned. And the possibility to use one slider (instead of two) is wanted. In the exploration mode the use of more different beeps (as in the user mode) is asked for. Last but not least it was mentioned that a help function for the VISA-comp system is not available.

5. Suggestions for improvements due to the experiment.

In this section we will sum up changes that can be made to improve the VISA-comp system. All changes suggested here are related to the experiment (the subjects told us to make this changes to improve the system, or we concluded them from their remarks). We will sum up some general improvements to improve the whole interface of the VISA comp system and we will give some improvements related to the exploration mode and the user mode.

5.1 General improvements:

1. Most important improvement which should be made (according to the subjects) is the speech module. It irritated them and they could not always understand the speech module. It should be remarked however, that all subjects use braille and are not used to synthetic speech. Literature mentions that synthetic speech needs a time to get used to, but after that time speech quality is of lesser importance than for instance adjustable speed.
2. The subjects want the possibility to interrupt the speech module. When you know what the speech module is going to say, you want to interrupt the speech module, because it speeds up the performance.
3. No help is available on how to use the VISA-comp system. If you do not know how to find some kind of information (e.g. the closebutton) or you do not know how to select a menu item, you can not find help about it. No information is available about screen changes when activating or selecting an item. For instance, when an application is activated, it would be nice if the system would give a special sound, so the subject knows when 'a new window appears full screen'. Another example is special sound, when selecting a menu item, that indicates 'a pull down menu appeared'.

5.2 Improvements to the exploration mode and the user mode.

The beeps of the system are helpful to the subjects to find the window objects. However, a lot of comment is given on the beeps. In the exploration mode subjects told that too much beeps were presented and the objects were placed too close to each other, so subjects hear separate beeps of the same frequency, consisting of different objects ('I do not know where the second object begins.')

In the user mode some problems occur with the names of the objects. For instance, the closebutton is a 'general button' and 'menu items horizontal' is the menu line. The technical names ('general button', 'menu items horizontal') are good for a technician, but not for a user. Better names, which the user can recognise, should be chosen. One subject mentioned, that he prefers to work with one slider. This was a subject who has participated in the IPO research how a screen design can be made, using the TASO, that is best suited for VIPS (IPO report 951). The guidelines presented in that report are not yet implemented in the VISA-comp system. We think that these implementation of guidelines will increase the ease of use.

6. Suggestions for improvements beyond the experiment.

In this section some improvements are suggested, that are not given by the subjects. These improvements are suggested after reading literature, observing the subjects and looking critically at how the system should be prepared before it can be used.

1. The irritation of the subjects caused by the speech module could also be influenced by the order in which the information is presented to the user. For instance 'listbox icon word perfect window program manager' could better be replaced by 'word perfect icon window program manager listbox', because subjects are searching for Word Perfect, therefore the latter information is of less importance. The order in which the objects are presented should be tested.

2. Before the exploration mode can be used a monstrous sequence of keys (*3x for every x between 0 and 9) should be typed on the keyboard of the TASO-screen. This was performed by the experimenter during the tests.

3. Selections and activations take a lot of time, because you have to move your hand to the keyboard to type +5. Then you move the hand back to the sliders and read the screen over there. To make a selection/activation button near a slider would speed up performance.

7 Conclusions

First of all, we can say that subjects can work with the VISA-comp system. They can find the information they needed on the screen. Furthermore, the subjects appreciated the accessibility to the Windows environment. However, a lot of improvements to the interface have to be made. We suggested some improvements of the VISA-comp system in this report. Finally we want to state that the report on the simulation of the VISA-comp system already presented some aspects of the system that improve the ease of use.

Literature

1. Card, S. K., Moran, T. P., & Newell, A. (1980). Computer text-editing: An information-processing analysis of a routine cognitive skill. *Cognitive Psychology*, 12, pp. 32--74.
2. Hammond, N. (1987). Principles from the psychology of skill acquisition. In M. M. Gardiner & B. Christie (Eds.), *Applying cognitive psychology to user-interface design*. Chichester: Wiley.
3. Marshall, C., Nelson, C., & Gardiner, M. M. (1987). Design guidelines. In M. M. Gardiner & B. Christie (Eds.), *Applying cognitive psychology to user-interface design*. Chichester: Wiley.
4. Paap, K. R., & Roske-Hofstrand, R. J. (1988). Design of menus. In M. Helander (Ed.), *Handbook of human-computer interaction*. Amsterdam: North-Holland.