Computerized Self-Administered Questionnaires on Touchscreen Kiosks: Do They Tell the Truth?

P.J. Blignaut
Department of Computer Science and Informatics, University of the Free State,
P.O. Box 339, Bloemfontein, 9300, South Africa
pieterb.sci@mail.uovs.ac.za

ABSTRACT
A computerized self-administered questionnaire (CSAQ) was implemented on a touchscreen-based information kiosk. Because of the voluntary nature thereof and uncontrolled circumstances in which respondents could complete the survey, it was essential to determine whether the feedback could be regarded as representing the true feelings of kiosk users. Respondents were categorized according to the number of items completed and the internal consistency of responses within each category was examined. Results from the CSAQ were compared with those from a paper-based survey. It was found that results of a CSAQ can be trusted if they are analyzed correctly.

Categories & Subject Descriptors
H5.2 [Information Interfaces and Presentation]: User Interfaces - Evaluation/methodology

General Terms
Measurement; Reliability

Keywords
Usability analysis; Touchscreen; Information kiosk; Self-administered questionnaires

INTRODUCTION
Information kiosk systems are computer-based information systems in a publicly accessible place, offering access to information or transactions for an anonymous, constantly varying group of users with typically short dialogue times and a simple user interface [3]. The success of such systems depends largely on the attractiveness of their user interfaces, how easily they allow access to information, and how clearly the information is presented [2].

Touchscreen-based information kiosks can bring technology to the people but it is important to determine if the system succeeds in bringing the information to the people. One possible technique to do so is to do a usability study by means of a survey form [5].

Survey forms can either be administered by a surveyor or by the respondent himself. Self-administered questionnaires can, in turn, either be paper-based or computerized. Computerized self-administered questionnaires (CSAQs) [4] are presented on-screen as a specific menu item on an information kiosk and users are invited to complete the questionnaire voluntarily.

Because of the absence of a surveyor and also because no incentives are offered, the validity of such a computerized self-administered questionnaire may be questioned:

- Respondents may play around and press buttons at random just to get a feel of the interface and technology.
- Respondents need not be honest in their feedback, nor do they have to complete all items.
- It is possible that low levels of general and computer literacy can influence the results. Respondents may not understand the questions well or the use of scroll bars and other GUI widgets may not always be obvious.

The goal of this study was, therefore, to determine whether the results from such a CSAQ can be regarded as representing the true feelings of kiosk users.

METHODOLOGY
The study was done for a large public transporting company that transports thousands of commuters daily between their homes and workplaces. Depending on the exact area where a person is living and working, he or she might have to take more than one bus for a one-way trip. Commuters need to be informed about routes, schedules and ticket prices. To date, this information was communicated by means of pamphlets and posters at the terminal building. A touchscreen information kiosk, placed in the foyer of the ticket box complex, was identified as a way to supplement this medium of communication while also serving a purpose to improve the company's corporate image amongst commuters.

A single-screen survey form with multiple choice items was made available on the kiosk where users could give their feedback on the system (Figure 1, next page). The survey form consisted of 15 main items of which two (numbers 5 and 15) had sub-sections. The first 5 items captured user profile details and the rest focused on usability aspects. The maximum number of items that could be completed (including sub-sections) was 18.

Virtually no computer literacy skills were expected from respondents as the survey was presented on a touch screen and respondents did not have to scroll or press buttons to go to another screen. Respondents could determine the scope of the survey with a single glance at the screen and

Copyright is held by the author/owner(s).
CHI 2004, April 24–29, 2004, Vienna, Austria.
ACM 1-58113-703-6/04/0004.

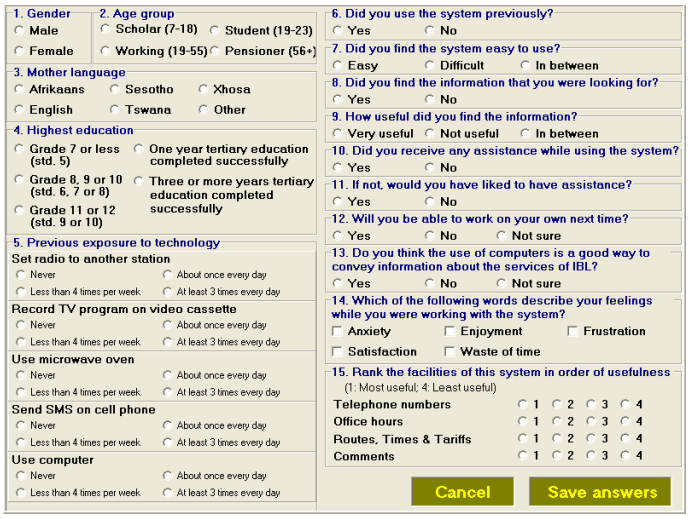

Figure 1. Computerized self-administered survey form

could hardly get lost or confused – one of the explicit guidelines of Norman and colleagues [4]. They also had the freedom to answer the questions in any order or even leave out some questions. This CSAQ was available on the information kiosk for 24 hours per day, seven days per week for a period of 2 months.

As a control experiment, other respondents were approached at random by a human interviewer with a corresponding paper-based survey (PBS) after they used the information kiosk. No respondent did both surveys.

Respondents were divided into categories based on the type of survey (PBS or CSAQ) and the number of items completed in the case of the CSAQ The internal consistency of the surveys was examined for each respondent category and the feedback from the respondent categories was compared with each other.

Items completed	Number of respondents	Time taken per item (s)		
		Avg	SD	SD/Avg (%)
1	163	18.2	15.2	83.5
2	60	12.7	8.4	66.1
3	32	9.7	7.0	72.2
4	30	10.1	7.6	75.2
5	9	10.9	7.8	71.6
6	11	9.4	7.3	77.7
7	10	7.5	4.4	58.7
8	5	15.2	11.3	74.3
9	9	6.7	4.7	70.1
10	3	12.6	5.8	46.0
11	3	6.9	1.8	26.1
12	4	9.4	4.9	52.1
13	7	9.1	7.9	86.8
14	8	8.3	3.3	39.8
15	5	13.9	4.4	31.7
16	17	11.4	3.0	26.3
17	54	10.4	4.5	43.3
18	104	10.1	3.1	30.7
	534	**83.9**	**84.3**	**100.5**

Table 1: Number of respondents per number of items completed on the CSAQ

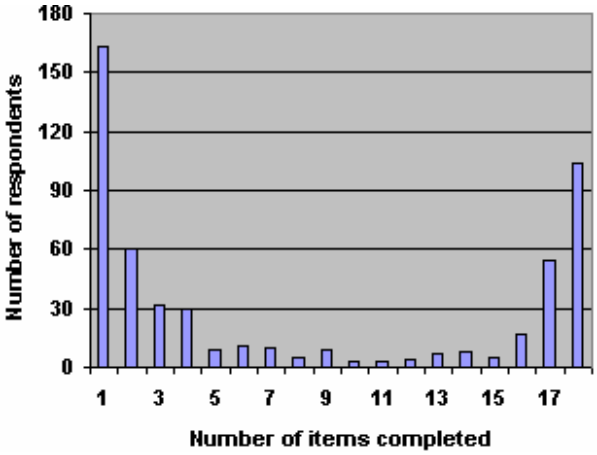

Figure 2: Number of respondents per number of items completed

RESULTS

Respondent categories

Table 1 shows the number of respondents who completed a specific number of items of the CSAQ with details on the time taken to do so. A graphical representation of these results is shown in Figure 2. It is clear that respondents can be divided into four categories:

- A: Respondents who completed the paper-based survey under supervision of a surveyor.
- B: A large number of CSAQ respondents opened the survey screen and completed 4 items or less at random. The large SD/Average ratios suggest a wide spread in the time taken to respond, possibly because respondents either closed the screen after they saw what it entails or left the system without closing the survey screen. (The system was programmed to return to the main screen automatically after 30 seconds after no user activity.)
- C: A small number of respondents completed 5 to 16 items. This is probably because respondents who were serious in their attempt at feedback took their time to complete the entire survey (category D).
- D: Serious respondents filled out the survey form completely (18 items) or missed out on one item only. The average time taken per item of about 10 seconds and the smaller SD/Average ratios indicate that the respondents took more or less the same amount of time to ponder over their responses, thus confirming their honest and serious approach towards the survey.

Internal consistency

In order to test the internal consistency of the surveys for each respondent category, the responses were randomly divided into two test groups. A series of chi-square tests was done to examine the following null-hypothesis:

- There is no difference in outcome for the respective usability items between the two test groups for each respondent category. In other words, the internal

Resp. category	Test group	Ease of use			Results		
		Easy	Difficult	In between	df	c^2	p
A	A	10	2	15	2	1.207	0.547
	B	13	3	11			
B	A	2	2	1	Test inaccurate (low frequencies)		
	B	3	3	1			
C	A	13	5	7	2	1.027	0.598
	B	18	7	5			
D	A	49	13	17	2	1.455	0.483
	B	45	10	23			

Table 2: Contingency tables for users' subjective indication of ease of use

consistency of the surveys for each respondent category is acceptable.

Table 2 shows four contingency tables with their respective chi-square results for the survey item of ease of use for the different respondent categories. Although quite a large number of respondents fall into category B, the number of respondents who completed any specific usability item is hardly ever high enough to justify the inclusion of this category in the analysis. The items that they did complete were mostly the first few on demographical data.

This process was repeated for each one of the usability items. A summary of the results is shown in Table 3. Significant p-values (α=0.1) is underlined. Not enough evidence exists to draw any conclusions for category B. Some of the items in category C proved to have been answered inconsistently. There is no evidence to justify

Item no.	Item	p-values per category			
		A	B	C	D
7	System is easy to use (overall)	0.547	-	0.598	0.483
6 + 7	System is easy to use (first-time users)	0.473	-	-	0.245
6 + 7	System is easy to use (follow-up users)	-	-	-	0.862
8	User found information that he/she was looking for (overall)	0.303	-	<u>0.058</u>	0.183
6 + 8	Found information (first-time users)	0.062	-	0.737	0.118
6 + 8	Found information (follow-up users)	0.393	-	<u>0.019</u>	0.354
9	Information found is useful	0.996	-	<u>0.079</u>	0.307
10 + 11	Would have liked assistance (for those who did not receive assistance)	0.571	-	0.124	0.880
10 + 12	Will be able to use the system next time without assistance (for those who had assistance)	-	-	-	0.619
13	Computers is a good way to disseminate info	-	0.263	0.707	0.107
14	Feelings	0.249	0.767	-	0.503

Table 3: Summary of chi-square results based on a split-half design for internal consistency per survey item

Respondent category	Usefulness of information		
	Very useful	Undecided	Not useful
A	42 (52.5%)	33 (41.3%)	5 (6.3%)
D	82 (52.6%)	43 (27.6%)	31 (19.9%)

Table 4: Contingency table for users' subjective indication of ease of use

rejection of the null-hypothesis with regard to categories A and D and although this does not prove internal consistency for these categories, there is no significant evidence to disprove it.

Comparison of paper-based survey and CSAQ

A series of chi-square tests was done to compare the feedback from CSAQ respondents with those who did the paper-based survey. The following null-hypothesis was examined:

- There is no difference between the outcome for paper-based surveys under supervision of a surveyor and a computerized self-administered questionnaire.

It is a well-known statistical principle that each expected frequency in a 2 x 2 contingency table should be at least 5 to give accurate results [1]. Because of this fact and also because of the dubious nature of their responses, responses in categories B and C were ignored for this analysis.

One of the items in the surveys asked users to give a subjective indication with regard to the usefulness of the information that the system provided (Figure 2, Item 9). The number of users who marked a specific response was entered into the contingency table for each of the categories A and D (Table 4). The overall χ^2 is 9.508 (df=2, p<0.01). This indicates that the null-hypothesis above can be rejected with respect to this item on the level

Item no.	Item	df	c^2	p
7	System is easy to use (overall)	2	9.645	0.008
6 + 7	System is easy to use (first-time users)	2	3.301	0.192
6 + 7	System is easy to use (follow-up users)	2	9.208	0.010
8	User found information that he/she was looking for (overall)	1	6.280	0.012
6 + 8	Found information (first-time users)	1	3.250	0.071
6 + 8	Found information (follow-up users)	1	1.594	0.207
9	Information found is useful	2	9.508	0.009
10 + 11	Would have liked assistance (for those who did not get assistance)	1	5.644	0.018
10 + 12	Will be able to use the system next time without assistance (for those who received assistance)	2	0.593	0.744
13	Computers is a good way to disseminate information	2	8.732	0.013
14	Feelings	4	16.249	0.003

Table 5: Summary of chi-square results for usability items

of significance of α=0.01. The percentages of the total number of respondents are indicated in brackets in Table 4. From these it is clear that the reason for the significant difference was a shift from "Undecided" with the PBS to "Not useful" with the CSAQ.

This process was repeated for each one of the usability questions in the survey forms. The results are summarized in Table 5. For the most of the items, the null-hypothesis that there is no difference between the outcome of a paper-based survey and a CSAQ can be rejected at least at the 5% level of significance. The direction of the shift in Table 4 from "Undecided" with the paper-based survey towards an explicit negative response was repeated for all other items where a significant difference occurred.

DISCUSSION

A computerized self-administered questionnaire can be implemented on a touchscreen-based information kiosk to obtain user feedback on general system usability. Because of the voluntary nature of the survey and the uncontrolled circumstances in which it is mostly completed, the results should be regarded with some circumspection.

It was found that many respondents on a computerized self-administered questionnaire just open the survey screen to see what it entails and then either turn away from the system, leaving it in the middle of nowhere, or close it after some random key presses. Respondents, who are serious and honest in their attempt to provide useful feedback, mostly try to complete the survey to the full. Therefore, only records for respondents who filled out the survey form completely or missed out on one item only should be considered when analyzing the feedback. The fact that no evidence could be provided to reject the hypothesized internal consistency of these respondents' feedback, further supports the view that their feedback may be regarded as valid and reliable.

The feedback of respondents who completed the paper-based survey differed significantly from those who did the CSAQ. The observed direction of the shift towards a more negative response with the CSAQ suggests that this difference can probably be ascribed to the fact that respondents were less honest in their responses in the presence of a surveyor and answered what he/she thought that the surveyor wanted to hear.

To summarize: A computerized self-administered questionnaire can give valuable feedback with regard to the usability of a touchscreen-based information kiosk, provided that only those surveys that are filled out completely are taken into account. One is even tempted to assert that, if analyzed this way; a CSAQ can be even more representative of users' true feelings regarding the system than a paper-based survey under supervision of a surveyor.

ACKNOWLEDGEMENTS

This work would not have been possible without the commitment of the staff and management of Interstate Bus Lines. Their cooperation is highly appreciated.

REFERENCES

1. Berenson, M.L., Levene, D.M. *Basic Business Statistics: Concepts and Applications*. Prentice Hall (1996), sixth edition.
2. Borchers, J., Deussen, O. & Knörzer, C. Getting it across: Layout issues for Kiosk Systems. *SIGCHI Bulletin*, October 1995. 27(4), 68-74.
3. Holfelder, W. & Hehmann, D. A Networked Multimedia Retrieval Management System for Distributed Kiosk Applications. In *Proceedings of the 1994 IEEE International Conference on Multimedia Computing and Systems*.
4. Norman, K.L., Friedman, Z., Norman, K., Stevenson, R. Navigational Issues in the Design of Online Self-Administered Questionnaires. *Behaviour & Information Technology*, 20, 1 (2001), 37-45.
5. Preece, J., Rogers, Y., Sharp, H. *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons (2002).