

Minimizing the errors in the questionnaire and monitoring the survey process

Kimin Kim, Korea Labor Institute

1. Introduction

The Korea Labor Institute (KLI) first started CAPI (Computer Assisted Personal Interviewing) that is the face-to-face interview method by using computer in 2006 on the Korean Longitudinal Study of Ageing (KLoSA) and Workplace Panel Survey (WPS) for the purpose of improving data quality. It was also used for the Korean Labor & Income Panel Survey (KLIPS) in 2007, and today all three panel surveys use the method. Initial efforts were concentrated on ensuring smooth operation of the CAPI system, and it stabilized over the years through several surveys. KLI then started to study ways to improve data quality even further, and focused on the development of programs that could improve the accuracy of the survey and monitor the interview situation.

This paper looks into the two programs thus developed, in terms of how they are being used and the outcome. Chapter 2 reviews the program that allows the survey to be reviewed by displaying the entire content of the blaise datamodel file onto a paper format. Chapter 3 is about the program that analyzes the audit file that stores the entire computer-conducted survey process.

2. Questionnaire Review Program

2.1 Overview

CAPI is a face-to-face interviewing by using computer that replaces the paper questionnaire in a survey. The paper questionnaire must be completely and faithfully implemented onto the blaise datamodel program. The quality of the blaise datamodel program is essential because any errors will severely affect the data. But for such surveys of such complex structure and large size like the KLoSA, KLIPS and WPS, there is a high risk of errors when the questionnaire is implemented onto the blaise datamodel program. To minimize such errors, the questionnaires must be reviewed to the fullest extent possible by allocating much time and human resources.

KLI developed a questionnaire review program to enhance the quality of the blaise datamodel program in a quicker timeframe. How this program works is that it allows everything from the computer-implemented survey to be displayed onto a paper format, including the questionnaire logic, error checking and responses. It can even be downloaded as a .doc file to be printed out. This not only allows better scrutiny of the survey, but also a view of the overall survey flow by observing the responses. For example, the user can check whether the survey was conducted properly by reviewing the responses on some of the questions. This questionnaire review program later added the code book feature, which shows the frequency of a certain response for each question.

2.2 Questionnaire Review

Figure 1 is a screen shot of the questionnaire review program. The program can be reviewed by entering the BDB (Blaise DataBase) file to be searched in the "Database name" field. On right screen all variables can be viewed or different variables can be selected. In particular, certain questions can be categorized as one group, to always query the responses of those questions whenever viewing the same survey. The user can also display the survey according to specific query conditions (see <Table 1>). The response values can be viewed in all or selectively by the survey ID.

Table 1. Query conditions for the questionnaire review program

Query	Description
Questionnaire	To see the questions of the survey only, without the responses
Hard/soft	The Blaise program has "hard check" and "soft check" features. Hard check is for errors that should never happen, and soft check is for potential errors that need to be reviewed by the user. These two will be marked on the survey when this query condition is used.
Block name	When the survey is implemented on the Blaise program, certain chapters of the survey can be separated into blocks. This is the name for each block.
Box by group	To display each block in a separate box.
Responses Question	To see the responses.
Logic	To see the logic of the survey.

Figure 1. Survey Review Program

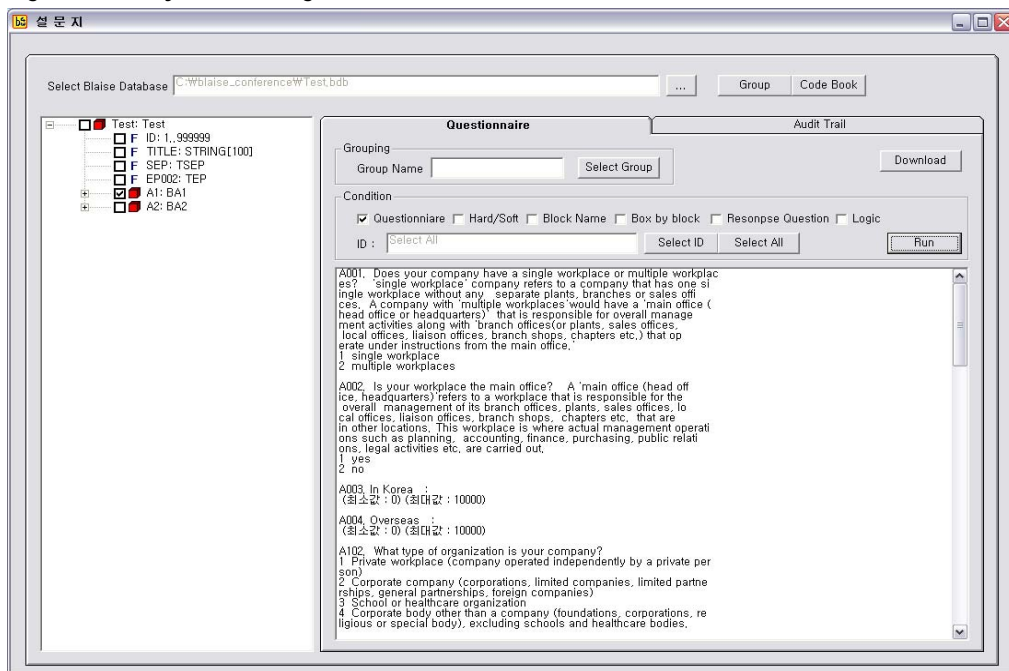
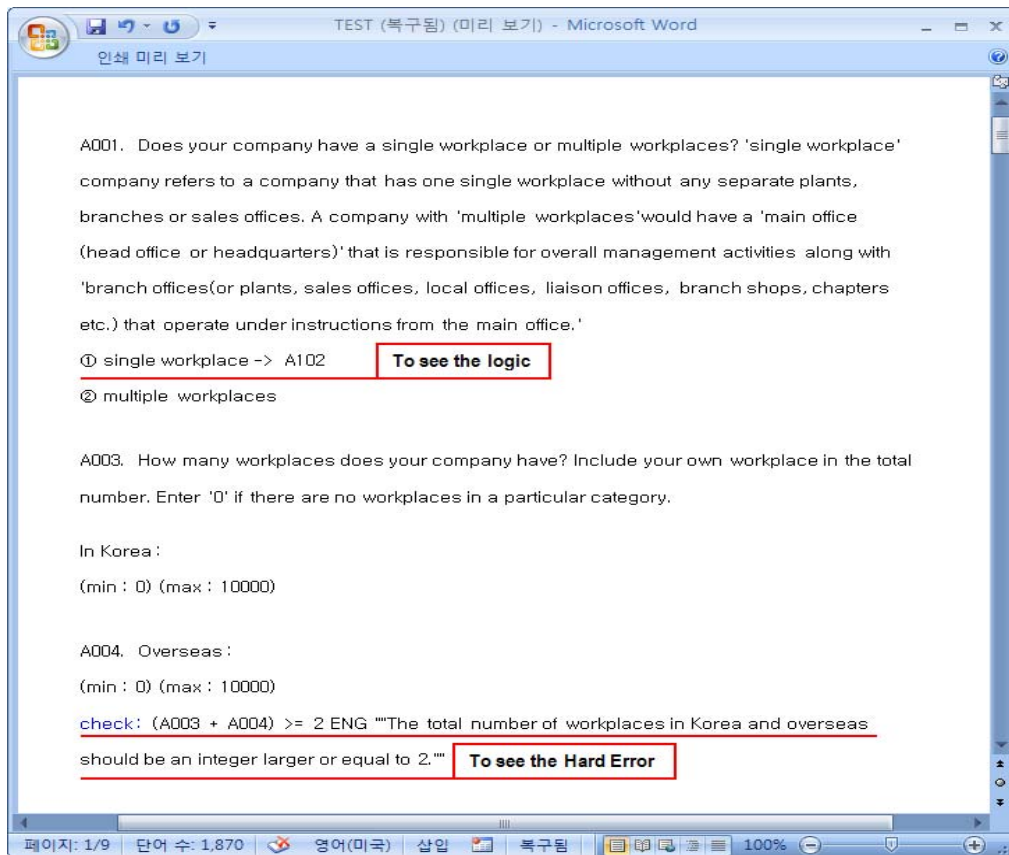


Figure 2 is the screen shot of the questionnaire review program downloaded as a .doc file. This is the result of a query using different variables, survey ID and other query conditions. The .doc file is edited for easier viewing than the computer screen.

Figure 2. Questionnaire in the MS Word Format



2.3 Code Book

The code book feature was added later on because the responses can be queried using the survey files. As can be seen in Figure 3, the code book is designed to show the survey questions and the frequency of the responses in an Excel format. We can know the distribution of answers by question because it is possible to show it

Figure 3. Code book

Fields	Question Text	Category	freq
A001	Does your company have a single workplace or multiple workplaces?	1 single workplace	1,060
		2 multiple workplaces	845
A002	Is your workplace the main office?	1 Yes	514
		2 No	331
A003	In Korea:		845
A004	Overseas:		845
A102	What type of organization is your company?	1 Private workplace (company operated independently by a private	71
		2 Corporate company (corporations, limited companies, limited	1,411
		3 School or healthcare organization	167
		4 Corporate body other than a company (foundations, corporations,	256
A103	Which of the following is most similar to the management system of your company?	1 An ownership management system where the owner has the authority	759
		2 An owner-centric management system where a professional manager	135
		3 A system where much management authority is transferred to a	433
		4 A professional managementsystem that is completely independent	316
		5 None of the above.	262

Sheet1 / Sheet2 / Sheet3 100%

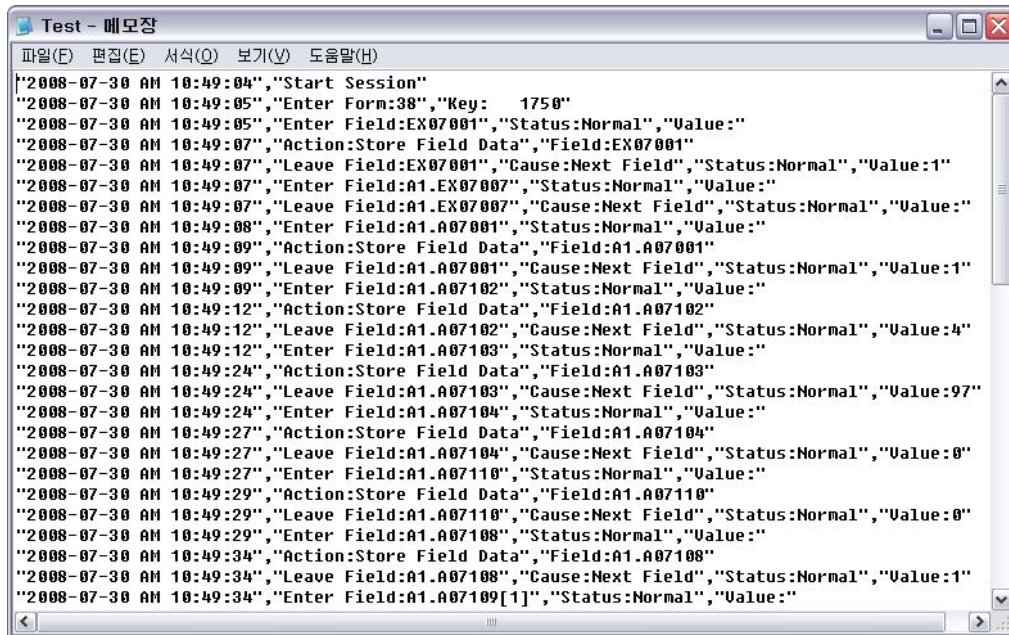
3. Audit File Analysis Program

Although data from a CAPI survey can be sent online immediately after the survey, it is still not possible to monitor all aspects of the survey in real time like CATI (Computer Assisted Telephone Interviewing),

particularly because wireless Internet is still not available anytime anywhere. Blaise makes up for it by allowing indirect monitoring of the survey using the Audit Trail feature. This feature records all movements of the keyboard or the mouse throughout CAPI onto a file named Audit. Figure 4 shows the Audit file format.

The purpose of this program is to indirectly assess the interview situation by analyzing the Audit file. Namely, it shows by survey ID, whether there were any out-of-the-ordinary trends, which questions took unusually long to respond, which questions required repeated corrections, among others. Identifying the difficulties in the survey and their causes will help improve the overall quality of the survey.

Figure 4. Audit file



```
2008-07-30 AM 10:49:04,"Start Session"
2008-07-30 AM 10:49:05,"Enter Form:38","Key: 1750"
2008-07-30 AM 10:49:05,"Enter Field:EX07001","Status:Normal","Value:"
2008-07-30 AM 10:49:07,"Action:Store Field Data","Field:EX07001"
2008-07-30 AM 10:49:07,"Leave Field:EX07001","Cause:Next Field","Status:Normal","Value:1"
2008-07-30 AM 10:49:07,"Enter Field:A1.EX07007","Status:Normal","Value:"
2008-07-30 AM 10:49:07,"Leave Field:A1.EX07007","Cause:Next Field","Status:Normal","Value:"
2008-07-30 AM 10:49:08,"Enter Field:A1.A07001","Status:Normal","Value:"
2008-07-30 AM 10:49:09,"Action:Store Field Data","Field:A1.A07001"
2008-07-30 AM 10:49:09,"Leave Field:A1.A07001","Cause:Next Field","Status:Normal","Value:1"
2008-07-30 AM 10:49:09,"Enter Field:A1.A07102","Status:Normal","Value:"
2008-07-30 AM 10:49:12,"Action:Store Field Data","Field:A1.A07102"
2008-07-30 AM 10:49:12,"Leave Field:A1.A07102","Cause:Next Field","Status:Normal","Value:4"
2008-07-30 AM 10:49:12,"Enter Field:A1.A07103","Status:Normal","Value:"
2008-07-30 AM 10:49:24,"Action:Store Field Data","Field:A1.A07103"
2008-07-30 AM 10:49:24,"Leave Field:A1.A07103","Cause:Next Field","Status:Normal","Value:97"
2008-07-30 AM 10:49:24,"Enter Field:A1.A07104","Status:Normal","Value:"
2008-07-30 AM 10:49:27,"Action:Store Field Data","Field:A1.A07104"
2008-07-30 AM 10:49:27,"Leave Field:A1.A07104","Cause:Next Field","Status:Normal","Value:0"
2008-07-30 AM 10:49:27,"Enter Field:A1.A07110","Status:Normal","Value:"
2008-07-30 AM 10:49:29,"Action:Store Field Data","Field:A1.A07110"
2008-07-30 AM 10:49:29,"Leave Field:A1.A07110","Cause:Next Field","Status:Normal","Value:0"
2008-07-30 AM 10:49:29,"Enter Field:A1.A07108","Status:Normal","Value:"
2008-07-30 AM 10:49:34,"Action:Store Field Data","Field:A1.A07108"
2008-07-30 AM 10:49:34,"Leave Field:A1.A07108","Cause:Next Field","Status:Normal","Value:1"
2008-07-30 AM 10:49:34,"Enter Field:A1.A07109[1]","Status:Normal","Value:"
```

Figure 5 is a screen shot of the Audit file analysis program. For this program as well, the user enters the BDB (Blaise DataBase) to be viewed in the "Database name" field. Specific options can be entered on the screen on the right-hand side to view different time data. The required time data can be viewed in all or in part depending on the query condition (see Table 2). Also, This can be downloaded onto an Excel sheet to allow easy calculation of the time.

Figure 5. Log file analysis program

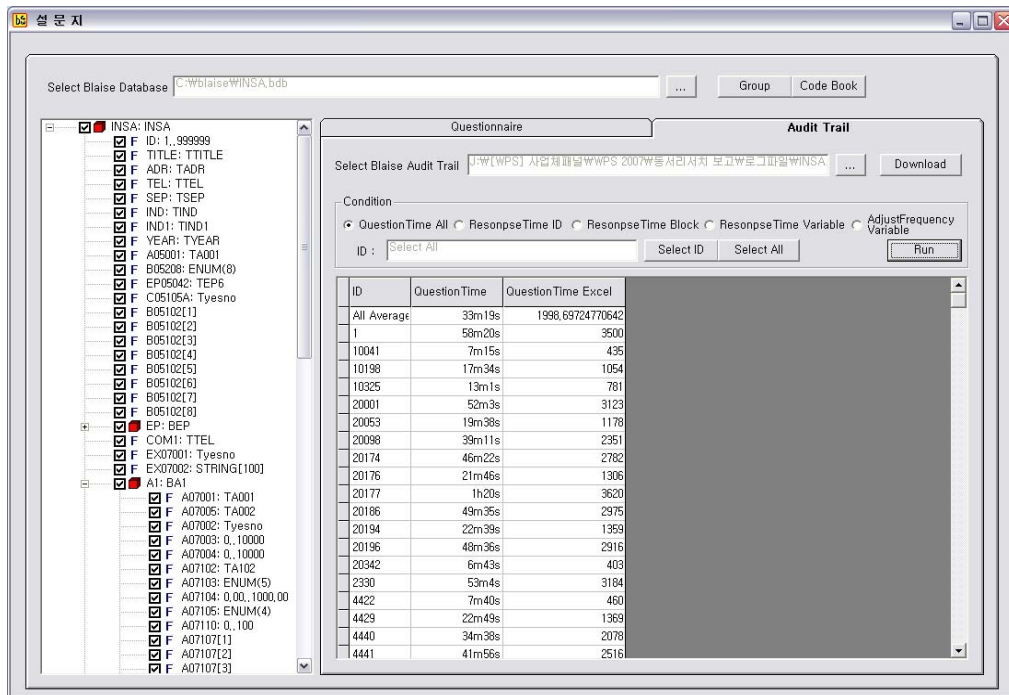


Table 2. Query conditions for the audit file analysis program

Query	Description
Question time all	To calculate between the start time and end time of the survey
Response time ID	To add all the response time for each variable
Response time block	The time required for each block (when the "block" feature is used for the survey program)
Response time variable	Response time by variable
Adjust Frequency Variable	Number of corrections by variable

4. Conclusion

KLI has been producing better quality data by adopting CAPI to conduct its surveys since 2006. As CAPI was smoothly implemented and stabilized, two new programs were developed in 2008 to review the survey tools and interview situation, as a way to enhance data quality even further.

One is a program that allows review of the survey by transferring the blaise datamodel program onto a paper form, to see the overall questionnaire structure and correctness of the response input. Initially, the purpose of this program was to enhance the accuracy of the survey. But not only that, but it also allows the user to see whether the survey was conducted appropriately, by showing the overall status of responses. In addition, the code book feature enables easy viewing of the data.

The other program is designed to enable indirect monitoring of the interview situation by analyzing the files that record the computer movements during the interview. The user can see how long the overall interview took, on which questions the respondents spent more time, how the responses for certain questions were corrected. This is likely to be useful for designing the surveys and training the interviewers. For the questions that took longer to respond, it can be discussed whether they need to be improved, and for those with many corrections, it can be determined whether they are so unclear as to confuse the respondents. Results of these discussions can be reflected to build better questionnaires and train the interviewers, to improve the overall quality of the survey.

For its CAPI, KLI adopted the globally-used Blaise program for the first time in Korea and successfully built a stable CAPI survey system. It continues to work toward improving its data quality, and to develop programs that will achieve this goal.