

주시동양

패널조사에서 PDA 활용 사례연구*

A Case Study on the Practical Use of PDA for Panel Survey

이상준**

Sangjun Lee

1968년 국내 조사기관이 최초로 설립된 이후 우리나라 조사방식은 급속한 발전을 거듭해 왔다. 최근에는 모바일과 인터넷 등 통신기술의 발달로 CATI, 노트북, 인터넷, PDA조사 방식이 진행되고 있다. 이 연구는 한국직업능력개발원이 자체 개발한 KISS 시스템을 이용하여 한국교육고용패널조사에 PDA를 응용한 사례를 중심으로 서술하고 있다. 이 연구에서는 설문조사시 PDA의 장단점 및 활용방안에 대해 살펴보는 것을 목적으로 하고 있다. 이 연구의 주요결과를 살펴보면 PDA조사는 1) 조사에 있어서 필드의 컨트롤이 수월하며, 2)자료검증기간의 단축 및 정확한 자료의 생성, 3) 조사에서 자료완성까지 기간 단축, 4) 소요인원의 최소화 및 장기적으로 비용절감 등의 장점을 가지고 있다. 반면에 단점으로는 1)설문표현의 한계, 2) 조사시간의 증가, 3) 일부지역의 통신장애 등을 들 수 있다.

주제어: 모바일 조사방법, PDA조사, KISS시스템, 조사방법

A survey method in Korea has being developed rapidly since a survey company was established in 1968. In latest, the CATI, internet survey, Notebook, PDA survey methods are being utilized in accordance with communication development. This article focuses on the practical use and the strength and weakness of the PDA survey solution which employs 'KISS' system designed by KRIVET itself. The survey using PDA has, on the one hand, such strengths as follows : 1) efficient field control, 2) reduction of time for verifying data and generation of proper data 3) shortening survey duration

* 패널조사를 위한 PDA 시스템 개발에 많은 분들의 도움이 있었다. 특히 오유성, 최기산 연구원에게 감사할 따르다. 이들이 아니었다면 KISS 시스템은 만들어지기 어려웠을 것이다. 또한 2004년도 11월 학술대회시 논평을 해 주신 통계청의 변재호 선생님과 유익한 논평을 해 주신 익명의 심사위원들에 고마움을 전하고자 한다.

** 교신저자(corresponding author): 한국직업능력개발원 전문연구원 이상준.

E-mail : sjlee@krivet.re.kr

from initial survey to data completion, 4) cutting down survey personnel and cost. On the other hand, survey using PDA has the following weakness: 1) limitation of questionnaire expression, 2) increasing time for survey, 3) communication barrier in some regions.

key words : mobil survey, PDA survey, KISS system, survey method, keep panel

I. 서론

패널조사처럼 대표본과 많은 양의 조사내용을 종이와 연필을 이용하여 조사를 할 경우 조사 의뢰인은 면접원과의 의사소통을 위한 인력과 시간이 필요하다. 면접원 관리를 포함한 필드 컨트롤이 조사의 성패에 영향을 미치기 때문이다. 조사 자체를 전적으로 조사전문기관에 맡기고 최종 결과 자료만을 받으면 이러한 문제는 존재하지 않을 것이다. 그러나 조사 자체의 과정과 오류, 성패여부를 늘 판단하여야 하는 조사의 경우라면 필드 컨트롤은 절대적으로 필요한 사안이라 하겠다. 보통의 경우 몇 명의 인원을 주요 지역에 파견하여 면접원과 슈퍼바이저들과 함께 전반적인 조사에 대해 논의를 하며, 이때 비용과 시간이 많이 들어간다. 조사가 완료된 후에도 자료입력, 코딩 과정 등 상당한 시간과 비용을 요구하는 것이 '종이와 연필조사'이다. 그러나 PDA나 노트북 같은 모바일 조사를 실시하면 인력과 시간, 비용 등에서 많은 효율성 제고 및 절감을 가져올 수 있다.

우리나라는 1968년 국내 조사기관이 최초로 설립된 이후로 지난 35년간 조사방법과 조사내용, 조사에 대한 사회적 인식이 발전과 변화를 거듭해 왔다. 조사내용은 초창기 시장조사를 중심으로 발전해 오던 것이 80년대 초에 들어서면서 정부정책이나 사회적 여론에 대한 조사가 시작되었다. 조사방법은 80년대까지 종이와 연필을 이용한 개별면접조사와 우편조사가 주로 이용되어 왔으나 90년대 초에 들어와 전화조사가 보편화되었으며, 통계청의 국가통계에서 조사자료 수집에 처음으로 컴퓨터를 이

용하기 시작하였다(전준우 2000). 1995년에는 CATI(Computer Assisted Telephone Interview)방식이 처음으로 도입되었다(이홍철 2002). 최근 몇 년 전부터는 인터넷의 확산으로 인터넷조사가 시행되고 있으며 모바일 기술의 발전으로 인해 노트북 및 PDA조사도 진행되고 있다.

노트북 조사의 경우 최근 노트북의 무게가 많이 가벼워졌다 할지라도 조사과정에 많은 어려움이 따른다. 노트북과 선물 등의 무게로 인해 면접원이 응답자를 만나러 가는 이동과정에 장애가 되고 있으며 일부 응답자의 경우에는 취조를 당한다는 느낌과 자신의 정보가 지울 수 없는 곳에 입력된다는 불안감이 존재할 수 있기 때문이다. 그러나 PDA의 경우에는 노트북과는 비교할 수 없는 무게로 인해 휴대하기 간편하며 조사시 응답자에게 접근이 용이하다는 장점이 있다.

이러한 조사방법의 국내외 사례를 보자. 먼저 PDA 조사의 경우는 선거 및 여론조사나, 마케팅조사, 만족도조사에 활용되고 있으며 또한 가정방문 교육관련 산업이나, 가스조사, 배달서비스 등에 현재 활발히 응용되어 사용되고 있다. 외국의 경우를 보면 미국에서는 집세조사를 '98년도부터 조사담당자가 "PEN-PAD" 컴퓨터를 사용하여 조사를 하고 있다. 일본도 도쿄지역의 소매가격조사에 모바일 장비를 이용하는 시스템을 2001년부터 도입하고 있으며 호주는 PDA를 이용하여 소비자 물가지수를 조사하고 있다(전준우 2002).

우리나라의 국가 통계에서 노트북을 이용한 조사는 1998년도 통계청의 경제활동인구조사가 처음이며, PDA 조사는 통계청의 2000년 물가조사에서 사용되었다. 그러나 이 조사는 대량의 조사이기보다는 조사표 형태의 조사에서 이루어져 온 것이다. 외국의 경우도 물가조사같은 단순한 조사표의 경우 PDA를 활용하고 있을 뿐이다. 그러나 연구기관의 패널조사에서 PDA를 활용한 조사는 한국직업능력개발원의 한국교육고용패널조사(KEEP: Korean Education and Employment Panel)가 처음이라 할 수 있을 것이다.

본 연구는 패널조사나 대형 횡단면 조사에서 종이와 연필을 이용한 조사방법을 탈피하여 PDA를 이용할 경우 나타나는 조사의 장단점과 시스템 설계 그리고 기술적 문제 등을 종합적으로 살펴보는 것을 목적으로 하고 있다. 따라서 PDA 조사를 위한 시스템 구축과정에 필요한 세밀한 정보통신공학적 이론과 논의는 이 글의 범위 밖의 문제이다.

이하에서는 「한국교육고용패널」 조사에서 활용한 PDA 조사방법을 중심으로 패널조사에서 PDA의 활용방안에 대해 살펴보고자 한다. II절에서는 PDA 조사를 위한 한국직업능력개발원의 KISS 시스템을 개괄적으로 살펴보고, III절에서는 PDA 활용의 장점 및 단점을 알아본다. 마지막 절에서는 향후의 개선방안을 제시하고자 한다.

II. PDA 조사를 위한 시스템 개괄

PDA를 통한 패널조사는 일반적인 횡단면 조사와 방법상에서 커다란 차이가 있는 것은 아니다. 일반적인 조사와의 차이점은 한번 조사한 사람을 지속적으로 추적하여 조사한다는 것과 첫 번째 조사의 성공이 패널조사의 성공여부를 판가름하기 때문에 기준 년도(base year)의 조사는 상당히 정확하게 이루어져야 한다는 것이다. 따라서 시스템을 설계할 때 특히 주의해야 할 것은 설문내용의 정확성 및 다양성, 응답자 관리, 항목간의 제약조건 부여, 조사의 신속성들이다.

과거 일부 설문조사에서 상용프로그램을 이용하여 PDA 조사를 실시하였을지 모르나 자체 솔루션을 설계하여 PDA 조사를 처음 실시한 것은 통계청의 물가조사가 할 수 있다. 통계청의 PDA 설문구성은 클라이언트 서버(Client Server) 환경을 기반으로 구성하고 있으며 전송에서는 TCP/IP(Transmission Control Protocol / Internet Protocol) 방식을 채택하고 있다. 반면에 한국직업능력개발원의 KISS 시스템의 설문

〈표 1〉 KISS와 통계청 PDA 비교

	통계청 PDA	KISS
설문구성 환경	CS(Client Server) 환경	웹 환경
전송방식	TCP/IP	다중코드분할접속방식(CDMA) 방식
프로그램 언어	Embedded Visual C++	Embedded Visual C++
백업장치	SD 메모리 이용	SD 메모리 이용
설문입력 장치	별도의 설문프로그램을 이용하여 설문을 작성한 후 PDA에서 다운로드	웹 환경에서 작성하여 PDA에서 다운로드
저장장치	Oracle 9i	Oracle 8i
PDA 개수	600여 개	80여 개
특징	<ul style="list-style-type: none"> · 문항과 응답란이 나누어져 있음 · 최종문항 응답 후 조사의 일관성 로직 체크 · 사용자 중심 설계 · 설문항목이 작음 	<ul style="list-style-type: none"> · 문항번호에 응답을 직접체크 · 문항별로 로직 체크 · 기능 중심으로 설계 · 설문항목이 많음

구성은 웹 환경으로 이루어지고 있으며 데이터의 전송방식은 클라이언트 서버 환경하에서 다중코드분할접속방식(code division multiple access) 을 채택하고 있다¹⁾.

통계청과 한국직업능력개발원의 시스템 간의 가장 큰 차이점은 전송 방식에 있다. 전자는 저장된 데이터를 컴퓨터에서 다운로드한 후 서버로 전송시키는 방식이며, 후자는 즉석에서 온라인으로 중앙 서버에 전송하고 있다. 또 하나의 차이점은 전자는 적은 설문량을 주로 사용함에

1) 클라이언트 서버 방식과 웹 방식의 가장 큰 차이점은 별도 프로그램의 설치를 필요로 하는가이다. 즉 클라이언트 서버 방식의 경우는 다운로드와 업로드를 컴퓨터를 통해서 해야 하기 때문에 반드시 별도의 서버접속 프로그램을 필요로 하며 이에 따라 장소의 제약과 사용자의 제약이 존재하는 단점이 있다. 그러나 웹 기반의 경우는 인터넷만 되는 곳이라면 별도의 프로그램을 필요로 하지 않고도 데이터의 전송, 수정 등의 작업이 가능하다. 그러나 단점으로는 웹이라는 제약사항 때문에 클라이언트 서버에 비해 상대적으로 많은 기능을 사용자에게 제공하지는 못한다.

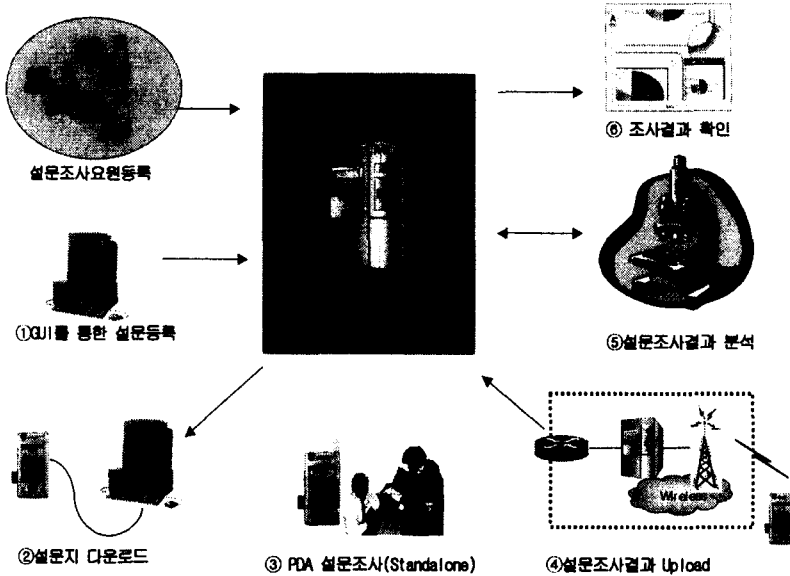
따라 사용자 중심으로 PDA 환경을 구성하고 있으며 후자는 다량의 설문문항 때문에 기능 중심으로 짜여 있다. 후자의 시스템이 사용자 편의보다 기능 중심으로 간 이유는 면접원의 컨트롤이라는 문제에 기인한 면이 크다.

통계청의 경우 응답자와 친숙하고 조사에 숙달된 정규직의 조사원이 책임을 지고 조사를 진행하고 있다. 그러나 일반 리서치 회사 면접원의 경우에는 근로형태가 비정형근로자이기 때문에 통계청 조사원만큼의 책임성과 충실함을 담보할 수 없다. 따라서 가급적 조사 진행시 면접원의 주관적 판단을 배제시키기 위하여 기능 중심의 설계가 필요하다. 이러한 관점 속에서 볼 때 KISS의 전송방식을 어디서든 쉽게 전송할 수 있는 온라인 무선방식으로 채택한 이유도 설명할 수 있다. 통계청의 경우 경제활동인구조사 또는 가계조사는 매월 동일한 조사를 하기 때문에 신속하게 자료를 전송하지 않고 전송 전 별도의 프로그램을 통해 조사된 내용의 논리적 체크를 한 후 성공적인 것만을 자료로 전송하고 있다. 또한 조사된 내용 중 오류는 면접원이 다시 확인 또는 일부 자료의 클리닝 과정을 거친 후 전송한다.

그러나 리서치 회사의 경우 조사원들이 이만큼 설문문의 내용에 대해 자세히 알기 어렵다. 설사 확인과정을 허용하더라도 면접원의 주관적 판단이 들어가 자료의 객관적 사실을 해칠 수 있기 때문에 조사 후 바로바로 전송하여 연구진들의 빠른 검토와 피드백이 중요해질 수밖에 없다.

이하에서는 한국직업능력개발원에서는 PDA를 이용한 「한국교육고용패널」 조사를 위해서 개발한 KISS(KRIVET Information Survey System) 시스템²⁾을 중심으로 PDA 조사에 필요한 솔루션을 서술하고자 한다.

2) KISS 시스템은 비단 패널조사에서의 PDA 활용만을 고려하여 개발된 시스템은 아니다. 일반적인 여론조사나 횡단면 조사에서도 활용이 가능하도록 설계되었으며, CATI 조사나, 인터넷조사까지 할 수 있도록 하고 있다. 또한 조사완료 후 평균, 빈도 정도의 기본적인 분석까지 이루어질 수 있도록 만들어졌다. 완성된 자료 중 사용자가 원하는 변수만 선택하여 추출한 후 다운로드 받을 수 있으며 여



〈그림 1〉 KISS 시스템에서 PDA 조사의 기본 흐름도

1. 설문지 등록 및 PDA 구현

데스크 탑 컴퓨터, 또는 노트북을 이용하는 CAPI(Computer Assisted Personal Interview)나, CATI, PDA를 이용하는 조사(PDA Assisted Personal Interview)는 기본적으로 '종이가 필요없는' 조사이다. PDA상에서는 직접적인 설문등록이 불가능하다. 따라서 다른 방식의 형태로 먼저 설문지를 등록한 후 PDA상에서 구현하는 방안을 고려해야 할 것이다. 먼저 설문지를 웹에서 입력하여 HTML 문서나 기타 PDA와 호환이 가능한 데이터베이스로(DB) 변환시켜 해당 서버에 저장한 후 PDA에서 다운로드 받아 사용하는 방안을 생각할 수 있

기에는 변수별 설명과 빈도, 평균, 무응답 이유 등이 담겨 있는 설명서와 변수의 길이(column), 넓이(width) 등 변수명의 정보를 함께 제공하고 있다.

다. 응답자의 유형이 하나가 아니라 다양한 유형을 대상으로 조사하는 방식이라면 유형별로 다운로드가 가능해야 한다. 아니면 설문문항이 응답자 유형별에 따라 분기를 타는 형태로 구성할 수도 있는데 이 경우 설문량이 많아지며 조사완료 후 전송시간이 길어지는 단점이 나타날 수 있다. 온라인상에서 전송시간의 증대는 PDA 관리비용의 증대를 가져올 수 있다.

PDA로 패널조사를 실시하는 데 있어 설문형태에 제약은 존재하지 않는다. 객관식, 주관식, 순위별 조사, 복수응답 조사 등 모두 PDA 조사가 가능하나 종이에 표현하는 방식과는 상당한 차이가 일어난다. 예를 들면 학생들의 학교환경에 대한 만족도 조사를 하는 경우 종이로 조사를 하면 하나의 표 형태로 물어볼 수 있으나 PDA의 경우에는 학교 환경 문항 각각에 대해 한 개의 화면으로 제시해야 한다. 노트북, PDA, 웹상에서 실시하는 설문조사의 장점 중 하나는 설문항목간의 분기 또는 설문항목 이동에 대해 면접원이나 응답자가 신경을 쓰지 않아도 된다는 점이다.

KISS 시스템에서 원활하고 효율적인 설문조사가 가능하도록 설계된 대표적인 기능으로는 1) XYZ 유형 문항 처리 2) 설문항목간의 GOTO 기능 구현, 3) 항목간의 논리오류 체크 기능, 4) 문항간 이동시간 측정(응답소요시간 측정) 기능이 있다. 1)의 경우는 소득이나 지출 같은 물음에 응답자가 대답을 꺼려할 경우 X: 잘 모름(I don't know), Y: 응답거절(refusal) Z: 해당사항 없음(nonresponse)의 유형을 확인하고 다음 항목으로 넘어가도록 설계하고 있다. 2)의 경우는 설문조사가 어느 정도 진행된 후 앞 문항의 오류가 있음을 인지하였을 때 그 항목으로 이동(goto)하여 다시 수정한 후 제자리로 돌아와 설문조사를 진행하는 기능이다. 이 경우 앞의 문항이 분기를 타게 되는 경우 기존에 조사된 문항이 분기의 논리구조에 모순되는 경우에는 전부 지워지게 된다. 예를 들면 1번 문항의 ①을 답한 사람은 3번으로 가고 ②를 답한 사람은

1. 학교생활 만족도 1. 귀하는 재학하고 있는 학교 생활 전반에 대해 만족하고 있습니까?

... 소문함 등록 ...

문항	소문함	문항	소문함
보기 1: 전혀 그렇지 않다	0	0	적관석
보기 2: 그렇지 않다	0	0	적관석
보기 3: 보통이다	0	0	적관석
보기 4: 그렇다	0	0	적관석
보기 5: 매우 그렇다	0	0	적관석

번호	변수	값	타입	설명	우선순위
번호	BY902001	2	NUMBER	학교생활 전반 만족도	1

〈그림 2〉 설문지 입력 도구

2번으로 가는 경우에 응답자가 ②로 답한 것을 다시 ①로 수정하였다면 2번 문항은 자동적으로 지워지게 된다.

3)의 항목간의 논리오투 체크의 경우를 예로써 설명하면 다음과 같다. 소득은 100만원인데 총지출이 300백만원으로 나타난다면 그 즉시 경고 창을 띄워 왜 그런지 정확히 확인하고 다음 항목으로 넘어갈 수 있는 장치이다. 마지막으로 4)의 기능은 전체 설문조사가 어느 정도 시간이 걸렸는지, 어떠한 항목이 평균적으로 가장 긴 시간을 소요했는지를 파악하고자 하는 기능이다. 또한 현 응답자에서 다음 응답자를 인터뷰하기까지의 이동시간을 체크할 수 있는 화면이다.

〈그림 2〉는 위에서 말한 내용을 웹으로 구현한 것이다. 중요한 부분만을 설명하자. ①은 문항 보기개수로 임의의 보기개수를 지정한다. ②는 응답이 복수문항일 경우 두 개의 답이 나올 수 있으므로 응답 가능한 답의 수를 정하는 기능이다. ③은 무응답 사유를 알고 싶은 항목, 예를

The image shows a screenshot of a survey system interface. It contains two question screens. The top screen is for question 7, '학교 외 일상생활' (Daily life outside school), with options (1) 어머니, (2) 아버지, (3) 형제자매, (4) 친인척, (5) 아무도 없음, (6) 기타. The bottom screen is for question 24, '귀하의 보호자는 누구입니까?' (Who is your guardian?), with options (1) 아버지, 어머니(계부, 계모 포함), (2) 아버지(계부 포함), (3) 어머니(계모 포함), (4) 할아버지(할대니), (5) 할(오빠), 누나(언니), (6) 삼촌(외삼촌), 고모(이모) 등 친인척.

〈그림 3〉 설문지 제약조건 구성화면

들면 소득같은 항목에 지정할 수 있는 기능이다. ④는 제약조건으로 문항의 보기에 따른 제약조건을 관리한다. ⑤는 문항복사 기능으로 공통적인 질문을 다른 설문지에 복사하거나 타 설문지에서 문항을 복사할 수 있는 기능이다. 끝으로 문항분기 항이 있는데 이는 보기1에 응답한 사람은 몇 번으로, 보기2는 몇 번으로 지정만 하면 조사시 PDA에서 자동적으로 항목이동이 이루어지게 한다. 그림 맨 아래의 '변수'란에는 연구자가 변수명과 길이, 숫자인지, 문자인지를 지정해 주고 있으며 변수명에 대한 짤막한 설명을 할 수 있게 하고 있다.

〈그림 3〉은 KISS 시스템에서 채용하고 있는 설문항목간의 제약조건을 부여하는 화면이다. '귀하의 보호자가 누구냐'는 질문에 부모가 없어 형, 오빠, 누나랑 같이 산다고 대답하였다면 '집에 가면 누가 있냐'는 대답에 아버지라고 대답할 수 없도록 제약을 두는 화면이다. 이러한 경우 위배되는 이유를 PDA 화면상에서 설명하고 둘 중 어느 것이 맞는 것인지 물어본 후 둘 중의 하나를 고칠 수 있게 유도하고 있다. 또한 소득은

100만원인데 지출이 200백만원으로 응답한다면 제약조건을 위배하나 응답자가 둘 다 맞다고 하면 다음으로 넘어갈 수 있도록 허용하고 있다.

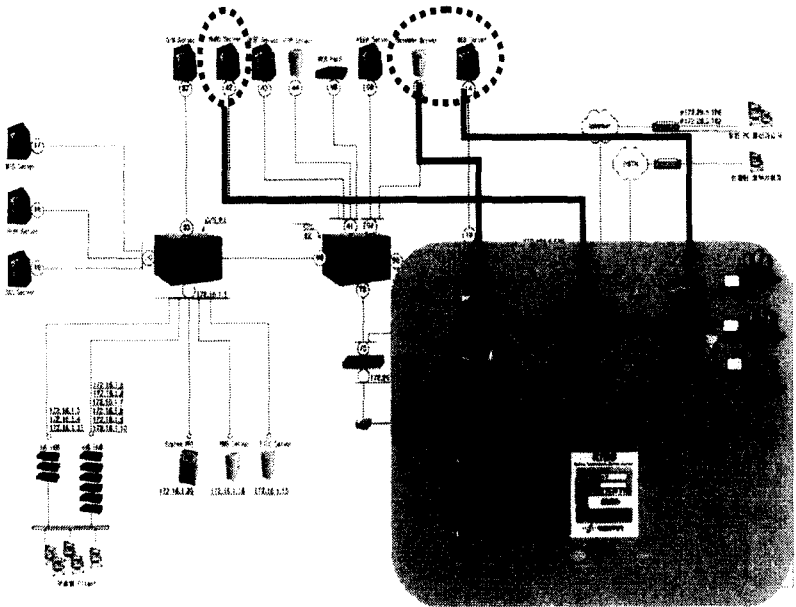
2. PDA 데이터 저장 및 전송을 위한 서버 구축

PDA로 조사한 설문결과는 두 가지 방식으로 저장될 수 있다. 하나는 현장에서 조사완료시 무선통신을 이용하여 중앙서버로 직접 전송하는 방식과 PDA 외부 메모리에 저장한 후 제3의 장소 컴퓨터에서 PDA 크래들(cradle)의 TCP/IP 통신을 이용하여 중앙서버에 저장하는 방식이 있을 수 있다³⁾. 전자의 경우에는 신속하게 자료가 저장됨에 따라 자료 확인이 가능하고 시정조치가 빨리 이루어질 수 있다. 그러나 통신장애가 있는 지역의 경우에는 전송이 어렵거나 불가능해지며 PDA 통신 비용이 많이 든다는 단점이 있다. 후자의 경우에는 통신비용은 적게 드나 전자보다 신속하지 못하다는 점이 있다. 「한국교육고용패널」 조사에서는 이 두 가지 방식 모두를 채택하고 있다.

〈그림 4〉는 「한국교육고용패널」 조사를 위해 구축한 서버의 흐름도이다. 여기서 DBMS(Data Base Management System)는 응답자 정보와 설문결과를 저장하는 서버로 사용하고 있으며 게이트웨이(gateway)는 PDA와 DBMS를 연결하여 PDA에 내장되어 있는 자료를 서버에 올리거나 DBMS의 자료를 PDA에 보내는 역할을 하고 있다. 웹서버는 일반 컴퓨터에서 사이트에 접속하여 설문조사를 실시하는 기능을 담당

3) 이 부분에서 통계청의 방식과 KISS 방식에는 다소 차이가 있다. 통계청의 경우에는 컴퓨터에 별도의 전송 프로그램을 설치한 후 PDA에 입력된 자료를 파일형태로 다운받고 이를 앞의 프로그램상에서 인터넷망을 이용하여 중앙서버에 올리는 방식이다. 반면에 한국직업능력개발원의 KISS에서는 인터넷의 망만을 빌리는 개념이기 때문에 파일형태로 컴퓨터로 다운로드받는 것이 아니라 PDA의 크래들(cradle)을 컴퓨터의 USB 포트와 연결한 후 PDA 화면상에서 업로드를 클릭만 하면 된다. 물론 여기에도 PDA와 컴퓨터 간의 모바일 장치를 인식하기 위해서는 'active sync' 라는 프로그램이 필요하다.

하고 있다. 이는 일반적인 인터넷조사와 함께 CATI 방식의 설문조사가 이루어지도록 설계되었다. 또한 유닉스 SAS 통계프로그램과 연결하여 자료의 가공 및 코드변환을 SAS의 SQL을 통해 편리하게 사용할 수 있도록 하고 있다. 한편 조사결과의 전송상에서 발생할 수 있는 자료의 손실을 막기 위한 백업장치로서 조사한 자료는 PDA의 램메모리에 (RAM)에 저장됨과 동시에 외부메모리(SD memory)에 자동적으로 저장되도록 하고 있다. KISS 시스템에서는 만일 데이터를 업로드한 후 PDA를 분실하거나 파손되었을 시에도 응답자를 다시 조사하지 않아도 된다. 조사한 로그(Log) 파일이 서버에 저장되도록 설계하였기 때문에 분실 및 파손상황에서도 데이터를 복구할 수 있다. 한편 저장된 파일은 윈도우 데이터베이스(WINDOW CE DB) 파일 형식으로 저장이 되며 이 때 서버 데이터베이스는 Oracle 8i를 이용하고 있다.



〈그림 4〉 서버 시스템 흐름도

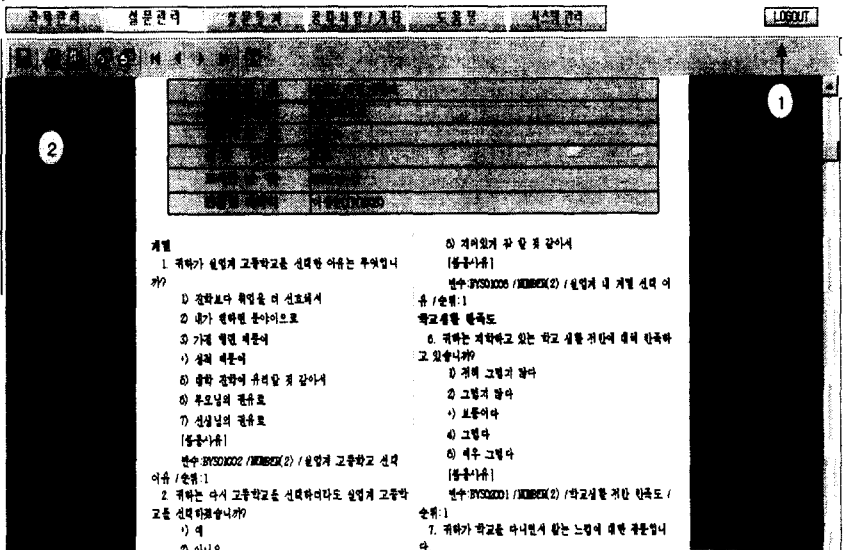
3. 데이터 검증 및 분석

설문조사를 위해 필드에 나갈 때 응답자의 식별과 해당 설문지의 저장이 필요하다. 무작위로 아무나 조사하는 경우라면 조사순서 또는 저장순서로 응답자의 아이디를 부여하면 되겠으나 패널조사처럼 특정인을 조사할 경우라면 사전에 응답자에 대한 아이디를 부여해야 한다. 이때 응답자의 식별은 사전적으로 면접원에게 알려주거나 PDA에 내장되어야만 결과 값의 혼선을 방지할 수 있다. 어떤 지역에 A라는 사람을 조사하는데 이 사람의 식별번호와 설문유형-취업자인지 실업자인지-에 따라 설문을 실시하여야 할 것이다. 이를 위해 현장에서 사전에 부여된 식별번호, 즉 응답자 할당 리스트와 설문지를 다운로드 받거나 아니면 이 둘을 외부메모리(SD memory)에 저장하여 설문조사를 실시한다.

조사가 완료되면 자료의 검증과 분석이 필요하다. 종이로 하지 않기 때문에 현장에서 자료검증 작업을 할 수 없으나 자료전송이 서버로 이루어졌다면 KISS 시스템에서는 접속이 가능한 컴퓨터에서 출력을 하여 확인할 수 있도록 설계되었다. 설문지를 출력하면 PDA에 응답한 결과가 번호대신 *로 마킹되어 나타난다. 이를 통해 면접원이나 검증원, 연구진이 기본적인 자료검증을 실시할 수 있다. PDA 조사는 자료입력 과정에서 예시번호를 벗어나 입력할 가능성이 없으며 자동적으로 응답자에게 필요한 설문문항으로 넘어가기 때문에 종이조사와는 달리 검증 작업은 논리적으로 설문응답이 맞는지를 확인하기만 하면 된다.

패널조사를 장기간 하게 되면 동일질문에 대해 과거 응답과의 일관성을 유지해야 하는 문제가 있다. 이에 대해 KISS 시스템에서는 패널 자료 클리닝 과정에서의 고질적인 문제인 동일질문의 시점 간 차이를 확인할 수 있는 장치를 마련하고 있다. 2~3년 간의 자료의 경우 이러한 문제를 직접 눈으로 확인하는 것이 그다지 어려운 문제는 아니나 10

년 이상 될 경우 자료 클리닝의 기간은 크게 늘어날 수밖에 없다. 이러한 문제를 본 시스템에서는 조사 현시점에서 원천적으로 해결하고 넘어갈 수 있도록 해주고 있다. 예를 들면 전년도 조사에서 학력을 전문대 졸업이라 응답하였는데 올해 조사에서 대학졸업으로 응답을 하게 되면 시스템에서 전년도 응답과 비교를 통하여 경고 창에서 확인하도록 하고 있다. 횡단면 검증 조건으로는 객관식 대 객관식 문항 간, 그리고 주관식 일지라도 소득 문항과 같은 주관식 문항과 객관식 문항 간의 비교를 통한 제약을 두어 원천적인 자료의 클리닝을 할 수 있게 하고 있다. 한편 PDA에서 조사된 자료는 실시간으로 DBMS 서버에 저장된다. 따라서 모든 조사의 검증작업이 끝나지 않아도 일정 정도의 자료만 축적되면 기본적인 분석이 가능하다.



〈그림 5〉 설문결과와 출력화면

III. PDA 활용의 장점 및 단점

1. PDA 조사의 장점

PDA 조사의 장점은 1) 조사에 있어서 필드 컨트롤(field control), 2) 자료검증 기간의 획기적 단축 및 정확한 자료의 생성, 3) 조사에서 자료완성까지의 기간 단축, 4) 소요인원의 절감, 5) 비용절감을 들 수 있다. 1)의 경우에는 조사에 임하는 면접원에게 PDA를 통하여 필요한 공지사항을 실시간으로 연락함으로써 조사의 민첩성을 강화할 수 있다. KISS에서는 면접원에 대한 일일 조사 모니터링이 가능함에 따라 조사의 주의점을 전체 공지나 개별 공지를 통하여 시행하고 있다. 면접원의 이동시간 체크와 항목별 평균시간을 산정하여 평균치에 크게 벗어나는 경우 그 이유를 파악하여 면접원의 필드 컨트롤(field control)를 시행하고 있다. 2)는 앞에서 말한 것처럼 응답자에게 필요치 않은 항목은 자동적으로 건너(skip) 뛰게 하고 항목간의 제약조건을 설정하여 자료의 검증 및 데이터 클리닝 기간을 단축시킬 수 있다. 또한 자료의 신뢰도를 높일 수 있게 하고 있다. 3)은 전체 조사기간에서 데이터 클리닝과 데이터를 서버에서 바로바로 추출하여 분석할 수 있기 때문에 조사에서 최종수요자에게까지 가는 기간을 단축할 수 있다. 패널조사는 면접원을 제외하고 자료의 검증확인과 질 관리에 인원이 많이 소요되는 조사이다. 종이로 할 경우에는 각 지역별로 그 날 조사된 설문지를 슈퍼바이저와 연구진이 최종적으로 검증하여 조사의 성패를 결정하는 것이 일반적이다. 그러나 PDA 조사의 경우에는 조사기간 동안 지역별로 인원을 파견할 필요가 없이 중앙에서 연구진이 확인하면 된다. 끝으로 일반적으로 조사비용에서 10~15% 정도 소요되는 자료입력비와 편차비 등을 절약할 수 있다. 또한 인원절감으로 인건비를 절약할 수 있다.

2. PDA 조사의 단점

일반적인 설문지를 PDA에 맞게 설계하는 데 있어서 가장 큰 문제는 설문지 보기문항 수의 제약과 개방형 질문이다. 글자 포인트 크기에 따라서 PDA에 구현되는 보기문항 수는 다를 수 있으나 일반적으로 한글 8포인트 크기를 기준으로 하면 KISS 시스템에서는 50개 이상 입력이 불가능하며 7개 이상 되면 스크롤 바가 생성된다. 이때 면접원이 이 부분에 대하여 주의를 기울이지 않으면 맨 아래에 있는 보기문항을 간과할 수 있기 때문에 가급적 8개 이상의 보기문항을 두지 않는 것이 좋다. 개방형 질문의 경우에는 원칙적으로는 조사가 가능하나, 키보드 아이콘이 작아 노트북만큼 글자 입력이 수월하지 않으며 오자가 나타날 확률이 상대적으로 크다. 또한 화면에 자필로 문자를 직접 쓸 경우 PDA는 자신이 인식할 수 있는 글자 유형만 읽기 때문에 이 방식의 사용은 더 더욱 어렵다. 한편 자료의 클리닝을 위한 설문지 항목 간, 시점 간의 제약조건은 설문지 입력단계부터 세심한 주의가 요구된다.

조사 현장에서 온라인으로 PDA 조사를 실시할 때 나타날 수 있는 유일한 문제점은 통신환경에 제약을 받는 지역에서의 데이터 업로드이다. 데이터를 업로드하려 할 때 그 위치가 도심에서 벗어난 외지의 경우 일부 통신접속 에러가 발생할 수 있다⁴⁾. 그러나 즉시 자료를 전송해야 할 필요가 없다면 설문결과를 외부메모리(SD memory)에 저장시켜 조사지역의 통신장애나 PDA 방전등으로부터 데이터를 보호한 후 안전한 곳에서 자료를 업로드 시킬 수 있다.

또 하나의 문제점을 꼽으려면 1:1 대인면접을 기준으로 할 때 한 화면에 설문 하나가 보임에 따라 응답자에게 전체적인 설문시간이 상대적

4) PDA는 기본적으로 데이터를 받거나 전송할 때만 통신망을 이용하기 때문에 기본적으로 온라인 상태가 아닌 오프라인 상태에서의 설문진행이 가능하다.

으로 길게 느껴지게 한다는 점이다. 실제적으로도 PDA 조사가 시간이 더 걸린다. KEEP 패널조사용 학생 설문지를 대상으로 PDA와 종이 조사를 비교해 본 바에 따르면 전자의 경우는 25~30분 걸리는 반면에 종이의 경우에는 20~25분 정도 걸리는 것으로 나타났다. 그러나 PDA의 통신장애같은 오류만 없다면 면접원이나 응답자 모두 PDA를 선호하는 것으로 보고되었다.

IV. 개선방안 및 결론

PDA 조사의 성공의 핵심은 같은 질문이라도 PDA에 적합한 설문 문항을 만드는 일과 면접원과 응답자 모두가 모바일에 친숙해져야 한다는 점이다. 「한국교육고용패널」 조사를 PDA로 이용하였을 때 패널 연구진은 전국각지의 면접원을 대상으로 설문지 설명회와 함께 총 3회 가량의 PDA교육을 실시하였다. 나이가 많은 면접원들의 경우에는 컴퓨터나 PDA에 대한 두려움이 높았던 것은 사실이나 이보다 더 큰 문제는 PDA의 글씨가 잘 안 보이는 것이었다. 이에 조사 전에 가급적 젊은 면접원으로 교체를 실시하였다. PDA 입력시에는 정확성이 요구되기 때문이다. 그렇지 않으면 일종의 편칭에러로 인해 자료의 정확성을 해칠 수 있다. 소득변수처럼 연속변수인 경우에는 최소치와 최대치를 부등호 형태의 제약조건으로 걸어두기 때문에 이 부분에 대해서는 정확성을 담보하고 있으나 만족도 조사문항의 경우에는 입력에러를 응답자에게 다시 물어보지 않는 한 오류를 확인할 방법은 없다.

PDA 조사를 실시할 경우 종이로 조사를 실시하는 조사내용과는 다른 형태의 조사내용 양식이 필요하다. 예를 들면 종이로 조사를 할 경우 개방형 질문이 자유롭게 이루어질 수 있으나 PDA의 경우에는 문자 입력시간이 상대적으로 오래 걸리고 오타가 나기 쉽기 때문에 가급적

지양해야 할 방식이다. 또한 지나친 개방형 질문은 조사완료 후 데이터의 정확성을 평가하는 논리적 내부 검사를 방해할 수 있다. 일부 설문에서는 특정 문항에 대해 응답자가 가지고 있는 정보를 다 적게 하는 경우가 있는데 이러한 설문내용은 PDA에서는 가급적 사용을 피하는 것이 좋다. 또한 앞에서도 지적한 것처럼 선택문항이 많아지면 스크롤바가 생겨남에 따라 맨 아래 선택항목의 경우는 못 보고 지나칠 수 있다.

PDA의 자료가 전송될 때 해킹의 위험이 발생할 수 있다. 정말로 중요한 고급정보를 조사한 후에 해킹의 위험으로부터 안전해지고 싶을 경우에는 별도의 VPN(Voucher Private Network) 클라이언트를 두어 해킹에 대한 안전장치를 마련할 수 있다. 모든 설문조사에서도 마찬가지지만 특히 PDA 설문에서 중요히 여겨야 할 것은 조사항목 간의 논리적 구조를 사전에 철저히 파악하여야 한다는 점이다. 조사항목에서 논리적 분기를 타는 경우 PDA 화면상에서 보여주지 않고 있어 면접원들이 이 부분을 체크할 수 없기 때문이다. 끝으로 PDA 이용에서 장애가 되는 문제는 PDA 비용이다. 노트북의 경우는 조사기간 동안 대여를 할 수 있으나 PDA는 현재 대여제도가 없어진 관계로 직접 구입을 하여야 한다. 통계청의 경우도 약 600여개의 PDA를 이용하고 있으며 「한국교육고용패널」 조사에서는 80여대를 별도로 구입하여 조사에 이용하고 있다.

결과적으로 필자가 보기에 소득, 시간, 비용같은 연속변수가 많은 조사나 개방형 질문이 많은 조사에는 PDA 사용이 적합해 보이지 않는다. 물론 조사량이 절대적으로 적은 경우는 PDA의 사용에 있어서 이러한 조사항목이 큰 문제가 되지는 않으나 조사량이 많은 경우는 조사시간의 증가와 이로 인한 집중력 저하, 입력오류 등의 문제가 발생하여 자료의 질을 저하시킬 확률이 높다. 위에서 지적한 바와 같은 문제만 해결된다면 노트북을 이용한 조사보다 PDA를 이용한 조사가 더 효율적일 수 있다. PDA 조사의 장점부분에서 언급한 것처럼 장기적으로

보았을 때 비용측면에서도 PDA 조사가 다른 방식의 조사보다 더 유리할 수 있다. 그러나 무엇보다도 중요한 것은 조사내용과 목적에 맞는 방식을 채택하는 것이라 할 수 있다.

끝으로 PDA는 핸드폰으로 사용하기에는 너무 크고 컴퓨터로 사용하기에는 너무 작아 기기 활용도에 상당한 문제점을 안고 있었다. 그러나 최근 들어 고객관리, 재고관리, 상품의 입출 확인, 선거, 여론조사 등에 활용되면서 상당한 사업적 가치를 증폭시키고 있다. 따라서 PDA 활용도 차원에서 위의 활용사례에 대한 통신공학적인 연구가 필요하다고 하겠다.

참고문헌

- 이계오 외. 2001. <<인터넷조사>>, 나남출판사.
- 이흥철. 2002. “한국 조사업계의 발달사.” <<조사연구>> 3(2): 123-153.
- 전준우. 2000. “자료수집 방법의 변화와 구현.” <<통계분석연구>> 제5권 제1호, 통계청.
- _____. 2002. “현장조사방법 개선사례.” <<통계분석연구>> 제7권 제2호, 통계청.
- 통계청. 2003. “통계청 PDA 사용자 매뉴얼.” 내부자료.
- 한국직업능력개발원. 2004. “KISS 시스템 매뉴얼.” 내부자료.
- 한국직업능력개발원. 2004. “KISS 시스템 설계서.” 내부자료.
- 김해동·이상준 외. 2001. <<청년층 교육·고용 패널자료 구축사업(II)>>, 한국직업능력개발원.
- 채창균·이상준 외. 2003. <<한국교육고용패널조사>>, 한국직업능력개발원.