

Internet Survey Methodology

Hae Yong Lee¹⁾, Kee Whan Kim²⁾

Abstract

Since early 1960s, when the telephone survey was used in the research area for the first time, there has been existed various methods to gather the information by survey. The existing survey methodology called PAPI(Paper-And-Pen Interview), due to the appearance of Personal Computer, might well be developed progressively. Mid-1980s, Internet was advanced remarkably in terms of technology. From early 1990s, in addition it served as a stepping-stone for progressive collecting method. Internet Survey is now called WWW Survey and expected that it will substitute for most surveys from now on. We explain the role and the characteristics for Internet Survey as one of the various data collecting methods. Furthermore, we draw the futures about questionnaires, data collecting and statistical analysis with it.

Keyword : Internet Survey; CADAC; OLAP

1. 들어가는 말

인터넷 서버이란 인터넷 사용자들을 대상으로 웹 또는 전자메일을 이용하여 설문을 진행하고 응답하는 일련의 행위를 말한다. 이는 인터넷이라는 커뮤니케이션 수단을 이용하여 설문지를 작성·제시하고 응답자를 선정하며, 응답자의 응답내용을 자료로 만들고, 그 자료를 이용하여 원하는 통계분석 결과까지 해결하는 원스톱(one-stop) 개념의 통합조사 형태로 확장하여 생각할 수 있다. 또한 인터넷 서버는 현실세계가 아닌 인터넷이라는 사이버 공간에서 이루어지는 설문조사로써 설문지를 하이퍼텍스트(hypertext)라는 웹 페이지의 형태로 사이버 공간에 위치시킴으로써, 전세계의 인터넷을 이용하는 모든 사람들을 대상으로 공개되는 설문조사 방법이다. 이러한 인터넷 서버는 인터넷이라는 새로운 커뮤니케이션 방법의 장점에 의존하고 있다. 인터넷의 보급과 이용이 일반화된 시기가 1990년대 초부터임을 감안할 때 인터넷 서버의 시작은 그보다 뒤인 1990년대 중반에야 본격적으로 활용되었음을 알 수 있다. 우리 나라의 경우 언론사나 기업체에서 포털 사이트(portal site)를 이용하는 방법으로 1997년 3월 KISSNET(<http://www.kissnet.co.kr>)을 필두로 시작하여 최근에는 각 언론사는 물론 방송사 및 여러 여론조사기관에서 활용하고 있거나 본격적인 준비를 서두르고 있다. 비록 시험적인 조사이기는 하였지만 2000년 4.13 총선 여론조사는 우리 나라에서 인터넷 서버를 선거여론조사에 활용한 원년으로 기록되게 되었다.

1) Professor, Department of Statistics, Sungshin Women's University, Seoul. 136-742
E-mail: hylee@cc.sungshin.ac.kr

2) Department of Statistics, Korea University, Seoul. 136-701 E-mail: korpen@hanmail.net

인터넷 서베이는 인터넷의 장점을 고려해 보더라도 기존의 조사방법들에 비하여 많은 장점들을 갖게 된다는 것을 쉽게 예상할 수 있으며 또한 단점에 대해서도 생각할 수 있다. 그러나 인터넷 서베이의 장단점을 논하기 이전에 기존의 조사방법들이 갖는 자료수집방법(Data Collection Method)의 관점에서 인터넷 서베이를 조망한다면 인터넷 서베이의 장단점 뿐 아니라 향후 인터넷 서베이의 역할에 대해서도 구체적인 이해가 가능하리라고 생각된다.

2. 자료수집과정의 구분 및 발전과정

일대일 면접으로 대표되는 전통적인 자료수집방법의 변화는 자료수집에 전화를 이용하면서부터 시작되었다고 할 수 있다. 전화를 이용하여 자료조사를 실시하는 전화조사법의 활용은 미국에서 시작되었으며 전화조사법으로 인해 시작된 변화는 자료수집에 컴퓨터를 이용하면서 속도를 더 하게 되었다. 이를 CADAC(Computer Assisted Data Collection), CASIC(Computer Assisted Survey Information Collection), 또는 CAI(Computer Assisted Interviewing)라고 부르기도 한다(Weeks, 1992). CADAC은 크게 세 가지로 구분할 수 있는데, 첫 번째가 CATI(Computer Assisted Telephone Interviewing)이다. 이는 컴퓨터를 이용한 전화조사법의 전형으로 조사자는 모니터에 나타난 설문을 전화로 질문하게 되고 조사대상자의 응답은 조사자에 의해 컴퓨터로 입력되게 된다(Nicholls & Groves, 1986). 두 번째 방법은 CAPI(Computer Assisted Personal Interview)로 조사자는 노트북 컴퓨터를 갖고 조사대상자의 집을 방문하게 되고 일대일 면접을 컴퓨터를 이용하여 진행하게 된다. 수집된 결과는 모뎀이나, 전자메일, 디스켓을 이용하여 취합되게 된다(Baker, 1992 ; Martin & Manners, 1995). 세 번째 방법은 CASI(Computer Assisted Self Interviewing)이다. 앞의 두 방법과의 차이는 조사원이 없다는 것이다. 즉 조사대상자가 스스로 질문을 읽고 응답을 직접 입력하게 되는 것이다. 이 방법은 조사자가 함께 있지 않고 응답자에게 모든 것을 맡겨야 한다는 것이 문제점이 될 수 있다. 이에 대한 대안으로 개발된 것이 CASI에서 조사자가 도움을 주는 CASIIP(Computer Assisted Self Interviewing with Interviewer Present, Scherpenzeel, 1995)나 디지털 음성이나 테이프를 읽어 주는 Audio-CASI 등이 있다. 컴퓨터를 이용한 자료수집방법은 비용이나 시간을 줄일 수 있다는 장점을 갖고 있다. 그러나 '수집된 자료의 질적인 측면은 어떠한 것인가?'에 의문을 갖지 않을 수 없다. PAPI(Paper-and-Pen Interviewing)에 의한 일대일 면접에서 얻은 자료의 질보다 CADAC에 의해 수집된 자료의 질이 떨어진다면, 우리는 이런 자료수집방법을 사용하는데 심각한 우려를 표명하지 않을 수 없다. CADAC에 의해 수집된 자료의 질에 대한 연구는 Leeuw and Nicholls II(1996)에 의해 수행되었으며, 그의 연구에서 CADAC에 의한 자료수집은 자료의 질을 높이는데 긍정적인 역할을 하고 있는 것으로 밝히고 있으며, 연구결과를 5가지로 항목으로 분류하여 자세히 설명하고 있다.

컴퓨터를 이용한 자료수집의 또 하나의 영향은 우편조사에서도 나타났다. 이는 두 가지로 구분할 수 있는데 DBM(Disk By Mail)과 EMS(Electronic Mail Survey)이다. DBM 방법은 조사대상자에게 설문작성 프로그램이 들어있는 디스켓을 보내게 되고 조사대상자는 이 프로그램을 실행시켜 응답을 한 후 다시 우편으로 되돌려보내는 방법이다(Higgins et al., 1987). EMS는 말 그대로 전자메일에 의한 조사방법으로 조사참여의사를 전자메일로 확인한 후 참여의사가 있는 경우 설문작성 프로그램을 받아 응답을 한 후 다시 전자메일로 보내는 방법이다(Fisher et al, 1995). 위의 두 방법은 조사 대상자들이 컴퓨터를 보유하거나, 네트워크(network)을 사용하는 사람들로 제한된다는 단점을 갖고 있다.

자료수집과정의 구분이라는 관점에 의하면 본 고에서 다루려 하는 인터넷 서베이는 CASI와 EMS가 인터넷에 의해 적절히 결합된 형태임을 알 수 있다. 사실 앞의 설명에서 알 수 있듯이 인터넷 서베이는 자료수집과정 발전단계의 하나라는 것을 주지할 필요가 있다. 전화, 컴퓨터의 도입으로 인한 자료수집방법의 발전과정에 인터넷이 결합됨으로써 자연스럽게 나타난 조사 방법이라고 할 수 있다. 이런 관점이라면 차후 인터넷 서베이가 서베이 방법의 주류를 이루게 될 것은 자명한 일이다. 그러나 현재 인터넷 서베이는 표본조사론(Sampling theory)관점에서 문제를 갖고 있다. 인터넷 서베이는 인터넷을 사용하는 사람들에게만 조사를 진행할 수 있기 때문에 이들이 아직 가능한 모든 모집단을 대표할 수 없다는 것이다. 전화조사법 역시 초기에 이런 문제점이 제기되었으나 전화의 보급과 아울러 해소되었다. 현재 우리 나라의 경우 아직 인터넷 이용자에 대한 통일된 정의는 없지만 한국전산원 기준에 의하면 '월 1회 이상 정기적으로 인터넷을 이용하는 사람'이라고 되어 있으며, 국내 인터넷 사용자 수를 2000년 6월 기준 약 1600만 명으로 추산하고 있다. 여러 기관들의 자료에 의하면 국내 인터넷 사용자는 급격한 증가추세를 보이고 있어 곧 2000만 명을 넘어설 것으로 예상되고 있다. 또한 향후 생활가전을 통한 인터넷 사용이 가능해질 것이므로 우리 나라의 인터넷 이용자는 계속적으로 확대될 것이며 표본의 대표성 문제 역시 머지않아 해결될 것으로 예상된다.

3. 인터넷 서베이의 장점과 단점

인터넷 서베이의 장점은 크게 다섯 가지로 볼 수 있다. 첫째는 조사기간의 단축을 들 수 있다. 단기간에 많은 표본의 확보가 가능하며 많은 응답자를 실시간으로 동시에 조사할 수 있기 때문이다. 둘째로는 조사비용의 절감이다. 별도의 조사원 확보나 교육이 필요하지 않으며 온라인(on-line)의 특성상 별도의 부가 비용이 필요하지 않기 때문이다. 셋째로 다양한 형태의 조사가 가능하다. 웹의 멀티미디어(multimedia) 특성을 활용한 다양한 형태의 조사를 대규모로 실시할 수 있기 때문이다. 넷째로 시간이나 지역의 제약을 극복하고 조사를 실시할 수 있다. 다섯 번째로 구조화된 설문을 통하여 조사의 편리성을 도모할 수 있다. 인터넷 서베이에서 사용되는 설문지는 기본적으로 프로그래밍에 의하여 구현되므로 응답자 특성에 맞는 설문을 선별적으로 웹 페이지를 통하여 보여줄 수 있어 응답자들이 설문을 빠르고 편리하게 따라올 수 있다. 특히 다섯 번째 장점으로 인하여 설문지 작성에서 연유될 수 있는 비 표본오차를 상당부분 줄일 수 있게 된다.

인터넷 서베이가 이상과 같은 장점을 지니고 있는 반면에 단점으로는 앞 절에서도 언급했듯이 인터넷 사용자들의 가능한 모집단에 대한 대표성 부족과, 인터넷 사용자의 모집단 정보를 아직 알 수 없다는 것이다. 먼저 첫 번째 문제의 경우는 인터넷 사용자가 계속 늘어난다면 해결될 수 있는 문제이다. 그러나 현시점에서는 인터넷 사용자들이 10, 20대에 집중되어 있으며, 성비도 7:3 또는 6:4 정도로 남성에 치우쳐 있으므로 당분간은 인터넷 서베이를 일부 특별한 영역에만 활용할 수 있는 실정이다. 두 번째 문제는 확률추출법을 사용할 수 없다는 것을 의미한다. 모집단에 대한 구체적인 정보가 없이는 표본추출틀(sampling frame)을 작성할 수가 없기 때문이다. 사실 표본추출틀 작성의 어려움에 대한 근본적인 이유는 인터넷 사용자가 빠르게 증가하고 있어 모집단 정보를 수집하기 위하여 조사를 하더라도 조사기간 중 모집단의 특성이 변화하기 때문이다. 이 문제 역시 시간이 지나야 해결될 수 있을 것으로 보인다. 인터넷 서베이의 또 다른 문제점들은 전통적인 서베이 방법과 비교하여 역사가 짧기 때문에 응답자의 태도에 관한 연구가 부족한 실정이다. 자기선택(self-selection)으로 인해 발생하는 응답의 문제점에 대한 연구와 인터넷 환경 하에서 응

답자에 대한 보상제공의 어려움을 극복하기 위한 연구 등이 진행되어야 할 것이다. 이 외에도 응답자의 프라이버시 보호나 통신예절 등의 문제도 해결해야 할 점으로 지적되고 있다.

4. 기존 조사법과 인터넷 서베이의 차이

CADAC을 이용하는 조사방법까지를 기존 조사방법이라고 하였을 때 인터넷 서베이와의 차이점에 대해서는 앞 절의 내용에서 개괄적으로 설명하였지만, 불충분한 것이 사실이다. 두 방법에 대한 구체적인 차이를 조사진행에 관련된 세부적인 항목에 따라 비교하는 것은 두 조사방법이 조사 형식이나 조사 진행환경 등이 다르기 때문에 매우 힘든 일이다. 정확한 비교를 할 수는 없지만 그 동안 인터넷 서베이의 실시를 통해 얻어진 경험을 바탕으로 각 세분 항목별 차이를 <표 1>에 제시하였다. 이 결과는 경험에 의한 것이며, 웹을 통하여 자기기입식으로 진행되는 인터넷 서베이에만 해당됨을 밝힌다.

< 표 1> 기존 조사방법과 인터넷 서베이의 차이

	기존 서베이	인터넷 서베이(자기기입식)
비 용	표본 증가에 비례	상대적으로 적은 비용
조사 기간	2주 ~5주	1일 ~ 10일
구조화 설문	어려움	가능
설문 형식	문자 기반	멀티미디어 기반
실시간 분석	불가능	가능
조사의 연속성	어려움	가능
응답자 접근성	제한적	시간·공간적 제한이 없음

기존 조사법과 인터넷 서베이에서의 차이는 자료의 분석에서도 찾아볼 수 있다. 제한이 있기는 하지만, CADAC을 효과적으로 사용하면 자료의 수집은 거의 실시간으로 이루어 질 수 있다. 그러나 인터넷 서베이와 구분되는 점은 수집된 자료의 분석부분이다. CADAC의 경우는 수집된 자료가 모두 취합된 상태에서 자료의 분석을 실시한다. 물론 부분적인 취합상태에서 자료분석이 가능할 수 있지만, 매번 자료의 취합과정을 거쳐야만 한다. 반면에 인터넷 서베이의 경우는 자료의 취합과정이 자동으로 처리되며, 자료의 취합과정 중에도 자료에 대한 실시간 분석이 가능하다. 즉 조사 진행자가 조사 진행과정 중 수집된 자료에 대하여 항상 실시간으로 자료의 검색과 분석을 할 수 있다는 것이다. 이러한 실시간 분석기능은 단순히 인터넷을 이용한 서베이기 때문만은 아니며, 취합과정에 있는 데이터 베이스와의 원활한 연결기능 및 웹 OLAP(On-line Analytical Process)이 구현되기 때문에 가능한 것이다. 이러한 OLAP 기능은 ASP, PHP, JAVA Servlet 등과 같은 다수의 언어로 구현될 수 있다. 이렇듯 인터넷 서베이는 단순히 웹을 이용하여 응답을 받고 자료를 저장하는 것 이상을 의미하며, 데이터베이스 연동기술, 실시간 자료처리 기술 등이 융합된 토탈 솔루션이라고 할 수 있다.

5. 인터넷 서베이의 설문작성 유형

인터넷 서베이에 사용되는 설문지는 웹을 통하여 볼 수 있으며, 응답자가 자기기입식에 의하여 응답을 선택하도록 하고 있다. 조사원의 도움이 없이 응답자 스스로 질문에 대한 응답을 하기 때문에 인터넷 서베이에 있어서의 설문지 작성은 자기기입에 의해 발생할 수 있는 오류를 최소화 할 수 있도록 작성되어야 하며, 이를 위해서는 설문의 종류, 화면의 구성 및 디자인 등과 같은 점들을 응답자 입장에서 세세히 배려하여야 한다. 아래의 <표 2>에 인터넷 서베이에 사용되고 있는 설문지의 유형과 이에 대한 평가를 제시하였다(<http://www.researchinfo.com>).

<표 2> 설문유형에 따른 평가

설문 유형	예제	평가	평가 이유
단일응답, 이분적 또는 다분적인 질문	다음의 보기 중 하나를 선택하십시오.	A	웹 설문지는 응답자에게 보기로부터 오직 하나의 응답만을 요구할 수 있다.
척도형 질문(명목형, 순서형 자료)	다음의 보기(척도)에 따라 평가하십시오.	A	응답자는 척도에서 하나의 응답만을 선택할 수 있다.
쌍 비교	2개중에 당신은 어느 것을 좋아하십니까?	A	위와 마찬가지로 웹 설문지는 응답자에게 보기로부터 오직 하나의 응답만을 요구할 수 있다.
다중 비교	이들 중 당신이 가장 좋아하시는 것은 무엇입니까?	B	복잡함은 응답오류를 증가시킨다.
다중 응답	다음 목록으로부터 하나 또는 둘 이상을 선택하십시오.	B-	응답자들이 전체 보기를 읽는데 부담을 느낀다.
순위형 응답	다음 목록으로부터 첫 번째 선택, 두 번째 선택, ..., 을 하십시오.	B-	응답자들이 부담을 느낄 수 있다. 이 설문유형을 효과적으로 응답받기 위해서는 색깔이 있는 그래프나 그림을 첨가시킨다.
주관식 응답(단문형식)	그렇게 생각하는 이유를 쓰십시오. (한줄 정도)	C	무응답 오류를 범하기 쉽다.
주관식 응답(장문형식)	그렇게 생각하는 이유를 자세히 서술하십시오.	D	매우 높은 무응답 오류를 범하기 쉽다.

앞서 전술한 내용에 의하면 인터넷 서베이에 사용되는 설문지 작성은 프로그래밍에 의하여 만들어진다고 하였다. 그러나 설문지를 작성하기 위하여 매번 프로그래밍을 하여야 한다면, 이에 대한 전문지식이 없는 사람들에게는 매우 어려운 일이 아닐 수 없다. 이런 문제를 해결하는 대안으로 많은 경우에 사용되고 있는 것은 조사자가 메뉴를 선택하는 식으로 설문지를 작성하는 방법이 있으며, 드문 경우이지만 인터넷 서베이 전용 설문 워드프로세서를 제공하기도 한다, 그러나 이 방법은 사용자가 새로이 워드프로세서 사용법을 익혀야 한다는 것이 단점이 될 수 있다. 마지막으로 다양한 경우의 설문지를 준비하고 이를 수정하여 사용하도록 하는 방법이 있다.

6. 인터넷 서베이에서의 통계분석

기존의 서베이는 설문 계획, 작성, 배포, 수거, 자료코딩, 분석과정을 거친다. 조사원이 설문 배포 및 수거를 통해 자료를 모으고 이를 통계전문가가 SAS, SPSS와 같은 통계패키지를 이용하여 분석하는 과정을 따르게 되는 자료수집과 통계분석이 분리된 조사방법이다. 반면 인터넷 서베이는 조사자가 직접 인터넷을 이용하여 웹 설문지를 만들고 앞서 언급했던 DBM, EMS와 같은 방법으로 설문을 배포하거나 웹을 이용하여 직접 응답을 하게 하여 수집된 자료를 웹을 이용한 실시간 분석(Web-Based OLAP)시스템을 이용하여 분석할 수 있다. 이 과정에서 자료분석을 위한 방법론 세 가지 정도로 정리할 수 있다. 첫 번째는 실시간(real time) 분석기능을 사용하지 않고 기존의 통계패키지(SPSS, SAS 등)를 이용하는 것이다. 기존의 통계패키지들은 이미 다양한 분석방법론을 제공하고 있어 깊이 있게 자료를 분석할 수 있으나, 실시간 분석이라는 인터넷 서베이의 장점을 포기하게 되는 단점이 있다. 두 번째는 실시간 분석기능을 제공해주는 통계패키지를 사용하는 것이다. 이를 위해서는 웹 기반에서 자료를 누적하고 분석할 수 있는 환경을 제공하는 통계패키지를 선정해야 한다. 기존의 다양한 분석방법을 사용할 수 있고, 인터넷 서베이의 장점을 모두 활용할 수 있어 가장 좋은 선택이라고 볼 수 있으나, 실시간 분석기능을 실제로 구현하기 위해서는 추가적인 과정이 필요하므로 개발기간이나 개발비용이 문제가 될 수 있다. 세 번째는 인터넷 서베이를 위해 개발된 전용 소프트웨어를 사용하는 것이다. 외국의 경우 인터넷 서베이 전용 소프트웨어가 이미 개발되어 판매되고 있으며, 질문지 작성, 자료 누적, 실시간 분석 및 보고서 작성의 기능등을 갖고 있다. Survey Solution, WebSurveyor, Survey select등 몇 개의 제품들이 있다. 국내의 경우는 아직 인터넷 서베이 전용 소프트웨어가 개발·판매되고 있지는 않다. 이러한 전용 소프트웨어는 저렴한 비용에 인터넷 서베이의 장점을 사용할 수 있지만, 자료분석 방법이 제한적이라는 단점이 있다. 외국에서 판매되는 제품들을 보더라도 주로 기술통계치, 반도분석, 교차분석, 분산분석, 회귀분석등을 제공하고 있으며, 그 이상의 고급통계분석방법은 제공하고 있지 않다. 이는 고급통계분석기법을 프로그램화하는 어려움과 아울러 실시간 분석의 제한점이라고 생각된다. 고급통계분석을 하기 위해서는 사용자의 주관적인 판단이 요구되므로 모든 것을 자동화하는 것은 어려움이 따르게 된다. 향후 인터넷 서베이와 관련된 기술이 더욱 발전하더라도 실시간 분석은 주로 자료의 집계와 탐색적 분석 위주로 이루어져야 할 것으로 생각되며, 고급통계분석은 자료수집이 완료된 후 체계적인 분석을 하는 데 사용해야 할 것으로 생각된다. 그러므로 인터넷 서베이에서의 자료 분석은 온라인(on-line)분석과 오프라인(off-line)분석을 병행하는 것이 바람직하다고 하겠다.

현재 국내에서는 인터넷 서베이 전용 소프트웨어가 개발되어 판매되고 있지는 않지만 인터넷 서베이를 대행해주는 회사들이 활발히 활동하고 있다. 약 30개사 이상이 있는 것으로 파악되고 있으며, 대표적으로 아이클릭(<http://eyeclick.co.kr>), 오타스커뮤니케이션(<http://www.survey.co.kr>), 인터넷메트릭스(<http://www.internetmetrix.co.kr>), 한국인터넷서베이&리서치연구소(<http://www.kisri.re.kr>), 한국패널리서치KPR(<http://www.panel.co.kr>), 이손C&CI(<http://www.best.co.kr>), 아하넷(<http://www.ahanet.co.kr>)등이 있다. 이들이 서비스하고 있는 형식은 인터넷을 이용한다는 것은 같지만, 서베이 내용이나 형식에서는 서로 차이를 보이고 있다. 그러나 통계분석의 관점에서 보면 대부분 인터넷 상에서 웹설문지를 이용하여 조사하고 오프라인에서 통계처리 한 후 그 결과를 다시 웹상에 올려놓는 방법을 택하고 있다. 그러나 최근 들어서는 실시간 분석기능을 제공하는 사이트들이 늘어나고 있는 추세이다. 실시간 분석 기능을 제공하는 회사들 중 다양하고 안정적인 기능과 함께

국내 최초로 질문지 작성, 자료누적 및 실시간 분석 시스템으로 특허를 획득한 곳은 아하넷이다. 이 사이트에서는 실시간 자동분석 시스템으로 조사 진행 중이나 종료 후 수집된 데이터에 대하여 분석하고자 하는 변수와 분석방법을 선택하면 분석결과를 웹 상에서 실시간으로 생성된다. 가능한 통계분석기법으로는 빈도분석, 교차분석, 빈도교차분석과 같은 기초통계분석에서 회귀분석, 다차원 척도법과 같은 분석방법이 제공되며 자료의 일차적인 형태를 파악할 수 있는 그래프를 제공받을 수 있다.

7. 결론 및 인터넷 서베이에서의 주의점

인터넷 서베이는 앞서 설명에서 알 수 있듯이 매우 편리한 조사방법으로 인터넷 뿐만 아니라 여러 전산관련 기술들이 합쳐져 다양한 가능성을 제시하고 있다. 그러나 인터넷 서베이를 갑작스럽게 출연한 혁신적인 방법론으로만 보아서는 안될 것이다. 즉 인터넷 서베이는 인터넷과 전산관련기술로만 이루어진 조사방법은 아니라는 것이다. 인터넷 서베이는 조사방법의 발전 단계의 관점에서 보면, 기술의 발전과 아울러 자연스럽게 나타난 조사방법이며 기존의 조사방법과 많은 부분을 공유하고 있다. 특히 자료수집 과정 및 자료분석 과정에서 많은 시간과 비용을 줄일 수 있도록 해주고 있지만 이러한 시간 및 비용의 절감은 자료수집과정을 발전시켜 나가는 기존 노력의 결과라고 보아야 할 것이다. 이와 더불어 세심한 주의가 요구되는 분야는 바로 자료의 분석분야이다. 자료의 분석은 자료의 집계와 통계분석으로 나누어 볼 수 있다. 자료의 집계는 인터넷 서베이에서 매우 효과적으로 이루어 질 수 있는 부분이며, 별다른 문제가 없다고 하겠다. 그러나 자료의 통계분석은 세심한 주의가 필요하다. 특히 인터넷 서베이에서의 분석기능은 수집된 자료를 정형화되어 있는 메뉴에 넣어서 결과를 얻게되는 기계적인 과정을 거치게 되므로 이 경우 잘못된 통계치의 산출이나, 잘못된 주장으로 이어질 수 있기 때문이다. 특히 이 부분은 현재 모집단의 대표성 문제로 인하여 제한된 영역에만 사용 가능한 인터넷 서베이에서는 심각한 문제가 될 수 있으며, 분석의 깊이가 깊어질수록 이런 문제점들 역시 비례적으로 증가할 것이다. 비록 완벽하지는 않겠지만 기술적으로는 얼마든지 고급통계분석기법을 실시간으로 구현하여 제공할 수 있다. 그러나 이에 앞서 고급분석기능들이 인터넷 서베이 시스템에 제공되었을 때 발생할 수 있는 역기능에 관한 문제 인식이 선행되어야 할 것이다.

앞으로 더 빠른 인터넷 관련 기술의 개발로 인하여 더욱 다양한 조사와 새로운 조사영역이 만들어 질 것이다. 우리는 이러한 인터넷 서베이의 가능성을 인정하고 인터넷 서베이가 향후 서베이의 전형으로 자리잡을 수 있도록 기존의 조사방법과 관련된 연구결과 등을 바탕으로 지속적인 연구를 해야 할 것이다.

마지막으로 인터넷 서베이를 진행하면서 주의해야 할 점을 일곱 가지로 정리하면서 마무리를 하고자 한다.

(1) 응답자들의 시간과 시각을 존중하라.

웹 설문지의 작성은 서베이에서 알고자하는 사항에 초점을 맞추어야 하며 예비조사를 실시하여 그로부터 시간과 구조에 대한 평가를 얻어야 한다. 화면의 구성에서 밝은 색이나 차가운 색을 배경으로 하는 것을 피하고, 진행방식에서는 응답자들이 마우스를 끌어 내리면서 설문에 응답하도록 웹 설문지를 구성하는 것은 피해야 한다. 설문이 길다면 큰 웹페이지를 여러 개 사용하는 것 보다

작은 페이지들로 구성된 화면을 제시하는 것이 효율적이다.

(2) 인터넷 서버에 전화조사 방식을 그대로 이용하려고 하지 마라.

대부분의 연구자들은 전화조사 방식으로 실시되었던 설문내용을 HTML형식의 웹 설문지로 대체하여 사용하는 경우가 많다. 전화조사는 조사원에 의해 조절이 가능한 반면 인터넷 서버는 자기입식에 의하여 진행되므로 조절이 불가능하다. 전화조사를 인터넷 서버로 대체하기 전에 우선 사전조사를 실시해 보는 것이 바람직하다. 인터넷 서버의 설문지는 무엇보다도 간결한 느낌이어서 하고 간단하게 보여야 응답자의 오류 방지 및 도중 이탈을 방지할 수 있다.

(3) 인터넷 서버와 일반 서버 방법을 혼합하여 사용하라.

인터넷을 기반으로 한 서버는 표본추출에 있어 편향적인 성질을 가지고 있으며 이는 데이터의 질을 떨어뜨리는 변수로 작용한다. 또한 인터넷 서버의 커다란 단점이기도 하다. 가능하다면 전화나 일대일 면접과 같은 방법을 이용하여 응답을 얻는 것을 병행하여야 한다. 그리고 응답결과를 반드시 확인해야 한다.

(4) 브라우저와 통신속도를 고려하라.

조사자가 자바나 자바 스크립트 또는 VB 스크립트를 이용한 그래프 또는 동영상을 사용한다면 응답자들이 가지고 있는 브라우저는 이를 구현시킬 수 없을 수도 있다. 느린 모뎀을 보유하고 있는 응답자는 설문조사 시 많은 그림들이 뜨면 거부감을 느낄 수 있다. 거부감은 편향된 응답을 유도할 수 있으며 응답을 그만두게 되는 빈도를 높게 된다. 만약 서버에 그래픽이나 동영상을 삽입해야 한다면 응답자의 브라우저 타입과 연결속도에 따른 응답의 차이를 감안해야 한다.

(5) 정형화되어있는 설문을 사용하는 것을 피하라.

대부분의 인터넷 서버에서 설문지를 구성하고 있는 설문들은 전문가에 의해 계획된 것들이다. 그러나 성별, 나이, 직업과 같은 인구통계학적 요인을 제외하고는 이런 설문을 믿고 그대로 사용하는 것은 삼가 해야 한다. 조사 목적에 가장 중요한 정보를 담고 있는 설문을 신중하게 작성하지 못하고 정형화된 설문을 그대로 사용한다면 데이터의 질을 떨어뜨릴 수 있기 때문이다.

(6) 설문의 동기유발은 많은 응답자를 모은다.

인터넷을 사용하고 있는 사람들에게 설문 참여하도록 요청하는 것은 흥미로운 TV 프로그램을 시청하고 있는 시청자들에게 설문참여를 요구하는 것과 다를 바 없다. 조사자가 설문 응답을 요구하려면 설문하고 있는 주체가 누구인지 무엇을 하려는가를 확실히 밝혀 인터넷 사용자들의 동기유발을 유도하여야 한다.

(7) 응답자들의 사적정보를 존중하라.

무작위 추출로 응답자들에게 전자메일을 보내서 설문조사를 실시한다면 응답자들의 이름과 전자메일 주소 등을 다른 용도로 사용하거나 공개하지 않아야 한다. 만약 사람들이 설문에 응답해준 대가로 스팸메일(spam mail)을 받는다면 조사자 자신을 포함해서 인터넷을 이용하여 조사영역 자체에 큰 타격을 가져올 수 있다(<http://www.researchinfo.com>).

참고문헌

- [1] 한국 인터넷 백서 (2000). 한국전산원.
- [2] Baker, R.P. (1992). New technology in survey research : Computer assisted personal interviewing (CAPI), *Social Science Computer Review*, vol. 10, pp. 145-157.
- [3] Baker, R.P. and Bradburn, N.M. (1992). CAPI : Impacts on Data Quality and Survey Costs, *Information Technology in Survey Research Discussion Paper 10*.
- [4] Edith, D. L. and Nichols, W.L. II. (1996). Technological Innovations in Data Collection : Acceptance, Data Quality and Costs *Sociological Research Online*, Vol. 1, no 4.
- [5] Fisher, B., Resnick, D., Margolis, M., and Bishop, G. (1995). Survey Research in Cyberspace : Breaking Group on the Virtual Frontier, in *Proceedings of the International Conference on Survey Measurement and Process Quality ; contributed papers*. Alexandria: American Staistical Association.
- [6] Higgins, C. A., Dimink, T. P., and Greenwood, H. P. (1987). The DISKQ Survey Method, *Journal of the Marketing Research Society*, vol. 29, pp. 437-445.
- [7] Leeue, E. D. and Nicholls, W. L. II (1996). Technological Innovations in Data Collection : Acceptance, Data Quality and Costs, *Sociological Research Online*, vol. 1, no. 4.
- [8] Martin, J. and Manner, S, T. (1995). Computer Assisted Personal Interviewing in Survey *Information Technology for the Social Scientist*. London : UCL Press.
- [9] Nicholls, W. L. II and Groves, R.M. (1986). The Status of Computer Assisted Telephone Interviewing : Part 1 - Introduction and Impact on Cost and Timeless of Survey Data, *Journal of Official Statistics*, no. 2, pp. 93-115.
- [10] Scherpenzeel, A.C. (1995). A Question of Quality : Evaluating Survey Questions in Multitrait-Multimethod Studies. Leidschendam : Royal PTT, Netherlands.
- [11] Weeks, M. F. (1992). Computer-Assisted Survey Information Collection : A Review of CASIC Methods and their Impication for Survey Operations, *Journal of Official Statistics*, vol. 4, pp. 445-466.
- [12] <http://eyeclick.co.kr>
- [13] <http://www.ahanet.co.kr>
- [14] <http://www.best.co.kr>
- [15] <http://www.internetmetrix.co.kr>
- [16] <http://www.kissnet.co.kr>
- [17] <http://www.kisri.re.kr>
- [18] <http://www.panel.co.kr>
- [19] <http://www.researchinfo.com>
- [20] <http://www.survey.co.kr>