

패널을 이용한 온라인 서베이와 오프라인 서베이 데이터 사이의 신뢰도에 대한 탐색적 연구*

권익현
동국대학교 경영학부 교수
(kwonih@dongguk.edu)
이상원
시너지21(주)
(swlee@synergy21.net)

본 연구에서는 마케팅조사 및 여론조사의 조사방법으로서 그 비중이 점점 증대되는 온라인 서베이에 의한 데이터의 신뢰도에 대하여 살펴보았다. 본 연구에서는 온라인 서베이와 오프라인 서베이를 패널 구성원을 대상으로 동일설문을 이용하여 2주일의 간격을 두고 실행해서 두 데이터 간의 신뢰도를 재시험신뢰도(test-retest reliability)를 이용하여 검증하였다. 분석결과 온라인 서베이 데이터는 오프라인 서베이 데이터와 매우 높은 상관관계를 보여서 그 신뢰도가 수용 가능한 수준임이 밝혀졌다. 다만 여러 세부항목으로 구성된 리커트척도에 의한 문항은 그 구성형식이나 설문지에서의 위치에 따라 신뢰도가 달라질 수 있는 가능성이 있는 것으로 나타났다. 따라서 온라인 서베이에서는 설문지의 디자인이나 구성형식 등에 대해 오프라인 서베이에 비해서 상대적으로 더 많은 주의가 기울여져야 하는 것으로 나타났다.

1. 서론

우리나라 인터넷 이용인구가 처음으로 70%(6세 이상 인구대비 기준)를 돌파했다고 한다. 정보통신부는 2004년 12월 기준으로 전 국민의 70.2%인 3158명이 인터넷을 사용하는 것으로 집계됐다고 2004년 12월 31일 밝혔다. 이는 이용률로는 한해 전보다 4.7% 증가한 것이며 우리나라의 이용률은 아이슬란드에 이어서 세계 2위를 달리고 있다고 한다(조선일보, 2005.2.1). 이러한 인터넷의 급속한 보급과 이용률의 증가는 많은 산업에서 인터넷을 매개로 하는 크고 작은 변화를 가져왔는데, 시장조사와 사회조사 등 조사 산업도 예외는 아니었다. 전통적으로 설문조사는 대인 면접이나 우편조사의

형태에 의해서 진행되다가 기본적 형식의 전화조사가 활용되기 시작하였으며, 그 후에 컴퓨터를 이용한 전화면접(CATI: Computer Assisted Telephone Interview) 및 컴퓨터이용 대인면접(CAPI: Computer Assisted Personal Interview) 등이 새로운 조사유형으로 주목받았고 근래에 와서는 인터넷 상에서 웹 또는 전자우편을 이용한 온라인 서베이가 크게 주목받고 있다.

온라인 서베이는 기존의 설문지를 하이퍼텍스트(hypertext)형태로 사이버 공간에 위치시켜서 인터넷을 이용하는 사람들을 대상으로 하는 서베이를 말한다(김광용, 김기수, 1999). 온라인 서베이의 유형은 크게 이메일 서베이, html형식을 기반으로 하는 웹서베이, 온라인FGI(Focus Group Interview)의 세 가지로 분류될 수 있는데(Batagelj et al

1998: Comley 1998), 이 중에서 이메일 서베이가 가장 신속하고 간편하며, 여러 계층의 사람을 대상으로 실시할 수 있다는 장점을 가지고 있다. 그러나 이메일 서베이는 복잡한 설문형식을 사용하기 어렵다는 단점 때문에 보다 복잡한 설문형식에도 이용 가능한 html 형식의 설문지를 이메일로 발송하거나 완전하게 html 형식을 기반으로 하는 웹서베이 유형이 많이 활용되고 있다(김중훈, 류진화 2002).

이러한 온라인 서베이는 기존의 오프라인 서베이와 비교할 때, 그 기본적인 방법론은 유사하고 인터넷이라는 환경적 특성으로 인해서 보다 즉각적이고 쌍방향적인 상호작용이 가능하고, 멀티미디어적이고 다양한 디자인 요소가 활용될 수 있다는 측면에서 차이가 있다(Solomon 1995). 온라인 서베이는 모집단에 대한 표본의 대표성에 한계가 있다는 문제점과(Comley 1998; Schuldt & Totten 1994; Tse 1998) 응답자의 인터넷에 대한 친숙도에 따른 표본의 대표성 손상문제(Opperman 1995; Schuldt & Totten 1994) 등 작지 않은 문제점을 가지고 있기 때문에 온라인 서베이는 이러한 문제점이 해결될 때까지는 민간 영역의 특수한 목적을 위한 단순한 조사에만 그 적용이 제한되어야 한다는 주장도 있다(명승환과 최영훈 2002). 반면에, 온라인 서베이는 서베이 방법론에 대한 중요한 평가기준인 신속성과 비용의 저렴성 면에서는 장점과(Comley 1998; Schaefer & Dillman 1998; Sheehan & Hoy 1999; Tse 1998) 그 서베이 결과가 오프라인 서베이 방법인 전화조사나 우편조사에 의한 서베이 결과와 비교할 때 그 신뢰도가 만족할 만한 수준을 나타내서(Bachmann et al 1996; Kiesler & Sproull 1986; Taylor 2000; Tse et al 1995) 향후 온라인 서베이의

활용도가 더욱 증대될 것이라는 견해도 있다(박무익, 2000).

본 연구에서는 온라인 서베이에 의한 데이터의 신뢰도를 오프라인에 의한 데이터와의 비교를 통하여 살펴봄으로써 온라인 서베이가 오프라인 서베이를 대체할 수 있는 가능성을 고찰하였고, 나아가서는 온라인 데이터의 신뢰도를 제고할 수 있는 방안을 모색하였다. 구체적으로 본 연구에서는 패널을 구성하여 패널 구성원을 대상으로 온라인 서베이와 오프라인 서베이(대인 면접법)를 실시하여 두 가지 방법에 의한 데이터 간의 신뢰도를 살펴 보았으며 설문지의 구성 형식을 다섯 가지로 나누어 설문지의 구성형식에 따라서 온라인 서베이 데이터의 신뢰도에 차이가 발생하는지를 살펴봄으로써 온라인 서베이 데이터의 신뢰도를 제고하는 하나의 방안으로서 설문지의 구성형식의 영향을 살펴보았다.

II. 관련 문헌 연구

온라인 서베이와 관련하여 진행된 연구들은 크게 두 가지의 영역으로 나누어서 살펴볼 수 있다. 먼저 온라인 서베이에 대한 과학적이고 객관적인 조사방법으로서의 평가를 위하여 기존의 전통적인 오프라인에서의 조사방법과 온라인 서베이와의 비교를 통해 그 신뢰성과 타당성을 고찰하는 연구들이 다(Bachmann et al 1996; Kiesler & Sproull 1986; Stanton 1998; Taylor 2000; Tse et al 1995). 예를 들어 Taylor(2000)는 비록 온라인 서베이가 확률적 표본추출을 하지 못한다는 면에서 한계가 있음을 지적 하였으나 같은 설문지를

이용하여 온라인 서베이와 전화조사를 병행 실시한 결과 두 조사로 부터의 조사 결과가 매우 유사함을 보임으로써 온라인 서베이의 타당성을 보였으며, 이외에도 여러 연구들 (Bachmann et al 1996; Kiesler & Sproull 1986; Taylor 2000; Tse et al 1995)에 의해서 온라인 서베이 데이터의 신뢰도는 믿을만한 것으로 검증되고 있다.

또 다른 연구영역은 온라인 서베이의 데이터 품질에 영향을 미칠 수 있는 요인들에 대한 규명에 초점을 맞춘 연구들이다. 류진화와 김종훈(2003)의 연구가 가장 대표적인데 그들의 연구에서는 관련 과거 문헌들에 근거해서 전자우편을 이용한 설문 조사시에 응답률, 응답속도 및 자료의 질에 영향을 미치는 요소들로서 '설문지의 개인화(personalization)', '조사기관의 성격', '응답자에 대한 보상', '설문형식', '사전 알림' 및 '후속 편지'의 역할을 살펴보았다. 분석결과 종속변수로서 응답속도에 대해서는 영향을 미치는 요인이 없는 것으로 나타났고 다른 두 종속변수인 응답률과 자료의 질에는 영향을 미치는 독립변수들이 있는 것으로 나타났다. 그러한 연구결과를 요약해 보면 다음과 같다.

먼저 Dillman과 Frey(1974)의 연구결과에 근거한 '개인화'에 대한 가설은 응답률과 응답품질에서는 유의한 결과가 나타나서 개인화된 설문지가 응답률과 응답품질에는 긍정적인 영향을 보이는 것으로 나타났다. 둘째, Dillman(1978)의 연구에 근거한 '조사기관의 성격'에 따른 차이에 대한 가설도 응답률과 응답품질에서 유의한 차이를 보였는데 일반기업이 발송한 설문보다 대학이 발송한 설문이 보다 호의적인 결과를 나타냈다. 셋째, Singer et al(1998)의 연구에 근거한 '응답자에 대한 보상'의 영향은 응답품질에서만 부분적으로 지지되었고 넷째, 웹 서베이의 형식을 왼쪽 정렬과 오른쪽 정렬

로 구분하여 실험한 결과 오른쪽 정렬이 더 선호된다는 Bowker 와 Dillman(2000)의 연구와 스크롤바(scroll-bar)로 이동하는 설문 형식이 선호된다는 김광용과 김기수(1999)의 연구결과에 근거한 '설문형식'에 대한 가설도 응답률과 응답품질에서 지지된 것으로 나타났다. 즉 응답자들은 설문지를 단락을 구분하지 않고 하나의 기다란 html 페이지로 구성하여 스크롤바(scroll-bar)로 이동할 수 있는 형식일 때 hypertext 기능의 메뉴를 통해서 여러 페이지의 질문 섹션사이를 이동하게 하는 스크린(screen)형식에서 보다 더 높은 응답률을 나타내었다.

류진화와 김종훈(2003) 연구의 다섯 번째 독립변수인 '사전알림'에 대한 가설은 지지되지 않는 것으로 나타났는데, '사전알림'에 대한 가설은 우편조사에서 사전알림이 응답률을 증가시키는 것을 보인 연구들(Ford 1967; Kanuk & Berenson 1975; Scott 1961)과 사전알림을 실시하면 응답자들이 스팸메일로 간주하여 삭제할 가능성이 줄어들 것이라는 Metha와 Sivadas(1995)의 연구에 근거한 것이었다. 끝으로 후속편지는 응답률에만 긍정적인 영향을 미치는 것으로 나타났는데 이 가설 또한 우편조사에서 후속편지 발송이 응답률을 증가시킨다는 연구들 (Etzel & Walker 1974; Kanuk & Berenson 1975; Peterson 1975; Wiseman 1973)과 이메일 서베이에서도 후속편지 발송이 응답률을 증가시킨다는 연구들(Smith 1997; Schaeffer & Dillman 1998)에 근거한 것이었다.

이상에서 살펴본 바와 같이 많은 연구들이 온라인 서베이에 있어서의 응답률, 응답속도, 응답완성도, 응답의 질 및 설문에 대한 만족도 등에 영향을 미치는 영향요인들을 규명하고자 하였고, 또 다른 연구들은 오프라인에서의 다른 서베이 방법론

(우편조사, 전화조사 등)에 의한 데이터와의 비교·분석을 통해서 온라인 서베이 데이터에 대한 신뢰도를 살펴보았다. 그러나 연구자들이 아는 한 하나의 패널 구성원을 대상으로 동일한 설문지를 이용하여 실시된 온라인과 오프라인 서베이 데이터 사이의 신뢰도를 살펴본 연구는 없었다. 이러한 맥락에서 본 연구에서는 온라인 데이터의 신뢰도를 같은 응답자와 같은 설문지를 가지고 실행된 오프라인 서베이 데이터와의 비교를 통해서 살펴 보았다.

III. 연구방법 및 연구결과

본 연구에서는 국내의 대기업이 운영하는 인터넷 사이트의 패널 구성원을 대상으로 2차에 걸쳐서 연구조사가 진행되었다. 본 연구에서 연구조사가 두 번 실시된 것은 온라인 데이터의 신뢰도를 단일 패널을 대상으로 동일한 설문지를 사용한 오프라인 데이터와의 재시험 신뢰도(test-retest reliability)를 활용하여 검증하려는 목적으로 1차 연구조사를 실행·분석한 결과, 추가적인 규명이 필요하다고 판단되는 연구조사 결과가 도출되었기 때문이었다.

이는 온라인 서베이와 관련해서 진행되어온 두가지의 연구흐름과 맥을 같이 하는 것으로서 결과적으로 1차 연구조사는 온라인 서베이의 조사방법론으로서의 타당성을 오프라인에서의 조사방법론과의 비교분석을 통해 살펴보는 것이고, 2차 연구조사는 온라인 데이터 품질에 영향을 미칠 수 있는 요인들에 대한 규명을 목적으로 실행하였다.

3.1 1차 연구조사

연구문제 1: 온라인 서베이는 오프라인 서베이를 보완 및 대체할 수 있는 조사방법론인가?

앞에서도 살펴보았듯이 온라인 서베이는 표본추출에 있어서의 한계(명수환과 최영훈 2002; Comley 1998; Schuldt & Totten 1994; Tse 1998)와 인터넷에 대한 친숙도와 같은 외생변수의 통제문제(Opperman 1995; Schuldt & Totten 1994) 등 실질적으로 활용 비중이 증가되고 있는 추세와 서베이 조사방법론으로서 잠재적인 문제점을 가지고 있으나 서베이 조사방법론에 대한 중요한 평가기준인 신속성과 저비용성을 감안할 때 온라인 서베이가 오프라인 서베이를 보완하거나 대체할 수 있는 정도의 타당성을 확보하고 있는지를 규명하는 것은 실무적으로는 물론 학문적으로 중요한 출발점이 될 수 있다. 이러한 타당성을 검증하기 위해서 동일한 조사내용을 온라인과 오프라인의 두가지 조사방법론을 통해서 얻어진 두가지 데이터간의 신뢰도를 바탕으로 살펴본 연구들(Bachmann et al 1996; Kiesler & Sproull 1986; Taylor 2000; Tse et al 1995)에 의하면 온라인 서베이 데이터의 신뢰도는 비교적 높은 것으로 나타났다. 본 연구에서는 기존의 연구들과 다른 방법론으로서 동일한 조사내용을 하나의 패널 구성원을 대상으로 온라인과 오프라인에서 시차를 두고 실행하여 얻어진 대상으로 온라인과 오프라인에서 시차를 두고 실행하여 얻어진 데이터간의 재시험 신뢰도(test-retest reliability)에 의해서 살펴보았다.

3.1.1 조사절차

1차 조사연구는 하나의 패널을 대상으로 실행되었는데 온라인 서베이는 2004년 1월 2일부터 2004년 1월 14일까지 13일 간에 걸쳐서 실시하였고, 온라인 서베이가 끝나고 2주 후에 10일간에 걸쳐서 같은 패널 구성원들 즉 동일인들을 대상으로 오프라인 서베이를 대인 면접법에 의해서 실행하였다. 여기서 서베이 사이의 시간 간격을 2주로 한 것은 재시험 신뢰도 (test-retest reliability)를 측정하는데 적절한 시간 간격이라는 판단에 따른 것이었다(채서일, 2004).

3.1.2 조사대상의 특성

1차 조사에 참여한 응답자는 총 412명 이었다. 이 중에서 온/오프라인 설문 결과 불성실한 답변을 한 12명을 제외한 400명의 자료가 분석에 이용되었는데 응답자의 특성은 <표 1>과 같았다.

3.1.3 조사내용

1차 조사에 사용된 설문지는 총 14문항으로 구성되었는데, 세부적으로 살펴보면 명목척도 문항이 7문항 이었고 등급법에 의한 등간척도 문항이 5문항 이었으며 각 6문항과 5문항으로 구성된 리커트

<표 1> 응답자 환경 특성

구 분	분 류	표본수	%
	전 체	412	100.00%
성 별	남 자	223	54.12%
	여 자	189	45.87%
연령별	20세 미만	100	24.27%
	20 - 30세 미만	131	31.79%
	31 - 40세 미만	135	32.76%
	41 - 50세 미만	46	11.16%
직업별	중학생	42	10.19%
	고등학생	58	14.07%
	대학생(대학원생)	74	17.96%
	사무직	95	23.05%
	판매/서비스	40	9.70%
	전문직	18	4.28%
	주 부	85	20.63%
인터넷 사용경력	1년이상 - 3년미만	35	8.49%
	3년이상 - 6년미만	142	34.46%
	6년이상 - 9년미만	149	36.16%
	9년이상 - 12년미만	84	20.38%
	12년이상	2	0.48%

〈표 2〉 척도별 설문문항

척도	설문문항
명목척도 (7문항)	<ul style="list-style-type: none"> · 귀하께서는 XX광고를 보신 적이 있습니까? (예/아니오/모름) · 어디서 보셨습니까? (TV, 잡지, 웹사이트, 기타) · 귀하께서 다음의 광고 중 보신적이 있는 광고를 선택하여 주십시오 (A, B, C, D, E) · 그 광고주에서 마음에 드는 광고는? (A, B, C, D, E) · 마음에 드는 이유 (1, 2, 3, 4, 5) · 귀하께서는 (경쟁사) OO광고를 보신적이 있습니까? (예/아니오/모름) · 어디서 보셨습니까? (TV, 잡지, 웹사이트, 기타)
등급법 (5문항)	<ul style="list-style-type: none"> · 광고물에 대한 전반적 평가(5점척도) · 광고가 제품선택에 미치는 영향정도(5점척도) · 광고 카피가 마음에 드는 정도(5점척도) · 경쟁사 광고물에 대한 전반적 평가(5점척도) · 응답자의 TV시청 정도(5점척도)
리커트척도 (6문항, 5문항)	<ul style="list-style-type: none"> · 응답자에게 노출시킨 두가지 광고에 대한 광고태도 (자사광고: 9번문항 / 경쟁사광고: 12번문항) - 광고카피가 마음에 든다 - 광고의 배경이 아름답다 - 광고가 품격이 느껴진다 - 모델이 마음에 든다 - 배경음악이 좋다 - 표현이 참신하다*

*: 조사를 의뢰한 회사의 광고물은 새로 제작된 것이었기 때문에 조사회사의 요청에 따라 '참신성' 항목을 추가하였음.

척도를 이용한 문항이 2문항으로서 총 질문의 수는 23개였다.

조사내용은 패널을 관리하는 대기업의 광고활동에 대한 효과조사를 위한 설문문항이었다. 각 척도별로 구체적인 문항을 살펴보면 〈표 2〉와 같다.

3.1.4 조사결과

먼저 7개의 명목척도 문항에 대한 온라인 서베이 데이터의 신뢰도를 살펴보기 위해서 온/오프라인 데이터 사이의 재시험 신뢰도(test-retest reliability)

를 명목척도 자료사이의 상관계수(Reynolds 1977)에 의해서 분석한 결과 모두 0.90이상(0.942, 0.914, 0.936, 0.961, 0.908, 0.991, 0.909)으로 나타났다. 다음으로 등급법에 의한 5문항에 대한 신뢰도도 피어슨의 상관계수를 이용하여 분석한 결과 이들 또한 0.90을 훨씬 상회하는(0.967, 0.986, 0.988, 0.952, 0.990) 것으로 나타나서 이들 항목들에 대한 온라인 서베이 자료는 오프라인 서베이 자료에 대비하여 매우 유사한 조사 결과를 보이고 있는 것으로 나타났다. 이는 온라인 서베이가 조사대상이나 조사내용에 따라서는

〈표 3〉 리커트척도 항목의 재시험 신뢰도(test-retest reliability)

9-1	9-2	9-3	9-4	9-5	9-6
0.999	0.992	0.976	0.762	0.786	0.897
12-1	12-2	12-3	12-4	12-5	
0.985	0.968	0.866	0.717	0.721	

오프라인 서베이를 보완하거나 대체할 만한 조사 방법임을 보여주는 또 하나의 연구결과라고 할 수 있다.

그러나 각 6개와 5개의 세부문항으로 구성된 2개의 리커트척도를 이용한 항목에서는 주목할 만한 결과가 나타났는데, 〈표 3〉에 제시되어 있듯이 리커트척도의 후반부를 형성하는 세부 문항들은 온라인 데이터와 오프라인 데이터 사이의 재시험 신뢰도(test-retest reliability)가 통계적으로는 여전히 유의하지만 상대적으로 다른 설문 항목들이나 해당 리커트척도의 전반부 항목들에 비해서 낮은 것으로 나타났다.

물론 9-6항목에 대한 상관계수는 9-4나 9-5 보다 높게 나타나서 일관성 있는 패턴을 보이고 있는 않지만 추가적인 연구조사의 필요성을 제시하는 데는 충분한 근거가 된다고 볼 수 있다. 따라서 이러한 1차 조사의 연구결과에 근거해서 2차 조사에서는 리커트척도의 구성형식과 리커트척도에 의한 항목의 설문지 상에서의 위치에 따라 설문지 유형을 5가지로 구분하여 온라인 서베이에서 리커트척도를 활용할 때 서베이 자료의 신뢰도를 높이기 위한 방안을 추가적으로 살펴보았다.

3.2 2차 연구조사

연구문제 2: 온라인 서베이에서 리커트척도에 의한 데이터는 설문지에서의 응답순서

와 척도점에 대한 설명 방식에 따라 데이터 신뢰도가 달라지는가?

1차 연구조사 결과 리커트척도를 이용한 온라인 서베이 데이터는 신뢰도에서 잠재적으로 문제점이 있을 수 있음이 나타났다. 리커트척도는 Likert(1932)에 대해서 처음 소개된 이후에 대도 측정의 도구로서 사회과학분야에서 매우 활발하게 활용되어 왔다(Likert, Roslow & Murphy 1993; Krosnick 1999). 또한 측정도구로서 리커트척도의 타당도와 신뢰도에 영향을 미칠 수 있는 영향요인으로써 척도점의 수(number of scale point)에 대한 연구(Aiken 1983; Johnson et al 1982; Ko 1994; Oaster 1989)와 척도점에 대한 설명 방식(formats of anchor labels)에 대한 연구(Dickinson & Zellinger 1980; Dixon et al 1984; French-Lazovik & Gibson 1984)등이 진행되어 왔다.

본 연구에서는 온라인 서베이 데이터의 응답품질에 설문의 제시형태가 영향을 미친다는 연구들(김광용과 김기수 1999; Bowker & Dillman 2000)에 근거해서 리커트척도의 구성형식을 네가지로 분류하였고, 리커트척도의 설문지에서의 응답순서가 데이터 신뢰도에 영향을 미칠 수 있다는 Weng과 Cheng(2000)의 연구에 근거해서 설문지에서의 리커트척도의 위치(응답순서)를 실험처리 하였다. 이에 대한 구체적인 설명을 '조사설계'에 제시하였다.

3.2.1 조사 절차

2차 조사는 앞에서 설명되었듯이 1차 조사에서 잠재적인 문제점을 드러낸 온라인 서베이에서의 리커트척도 문항에 대하여 보다 체계적으로 살펴보기 위해서 실시되었다. 다음의 '조사설계'에서 제시된 바와 같이 리커트척도의 구성 형식과 설문지에서의 위치에 따라 5개의 집단으로 나누어서 실시되었는데 1차 조사 때와 같은 인터넷 사이트의 패널 구성원을 대상으로 5개 집단 별로 100명씩 무작위 추출하여 진행하였다. 또한 2차 조사는 조사 인력이 한계 등으로 인하여 5개 집단에 대한 서베이를 동시에 실행하지 못하고 2004년 2월 1일부터 2004년 4월 30일 까지 두 달에 걸쳐서 집단별로 실시하였다. 온라인 서베이와 오프라인 서베이 사이의 시차는 1차 조사 때와 같이 2주일로 하였다. 이러한 조사절차는 실험집단간의 실험처리에 시차가 존재하여 여러 외생변수에 노출될 수 있는 가능성이 있다는 면에서 잠재적 문제점을 가지고 있다고 할 수 있다.¹⁾

3.2.2 조사설계

앞에서도 설명 되었듯이 2차 조사에서는 6개의 세부문항으로 구성된 리커트척도 문항의 구성 형식

및 설문지에서의 위치에 따라서 5가지의 설문 형식을 도출하여 집단별로 나누어서 실행하였다. 리커트척도를 제시하는 형식으로는 첫째 응답설명(전혀 아니다/매우그렇다: 5점척도)을 상단부분에 한번만 제시하고 6개의 세부문항을 하나의 리커트척도로 묶는 가장 일반적인 형식과, 둘째 6개의 세부문항을 세문항씩으로 나누어서 응답설명을 각각으로 나누어서 두 번 제시하는 형식, 셋째 리커트척도를 6개의 단일 문항으로 나누어서 응답설명을 각각의 세부문항 마다 각각 제시하는 등급법의 형태, 넷째는 온라인 서베이의 설문지 디자인 측면의 장점을 활용한 형식으로써 플래쉬를 이용하여 응답설명이 배너(banner)의 형태로 스크롤(scroll)할 때마다 각각의 세부항목을 따라가도록 하는 애니메이션 효과를 사용하였다. 또한 리커트척도의 설문지에서의 위치는 총 총 12문항으로 구성된 설문지에서 앞부분(3번문항), 중간(7번문항), 뒷부분(12번문항)으로 구분하여 설계하였는데,²⁾ 첫 번째 구성형식인 6개의 세부문항을 하나로 묶는 형식만 설문지의 앞쪽과 뒤쪽에 위치시키는 두가지 형식으로 다시 구분하였고 나머지 3가지 구성형식(두 부분으로 구분한 형식, 단문항 형식, 배너 형식)에 의한 리커트척도 항목은 설문지의 중간 부분에 위치시켜서 <표 4>와 같이 총 5가지 형식이 사용되었다.³⁾

오프라인 서베이에서의 설문형식은 온라인과 동

1) 이와 같은 한계는 예상된 것이었으나 조사 내용 및 인력의 한계로 인하여 집단 별로 나누어서 실행하는 것은 불가피한 선택이었다. 이와 관련해서 연구자들은 2개월의 조사기간 중에 해당 인터넷 사이트나 설문 내용과 관련해서 발생할 가능성이 있는 각종 외생변수들(인터넷 사이트 기업에 대한 이미지 등)에 대해서 면밀하게 모니터링 하였는데 연구자들의 인지범위 내에서는 별다른 문제점은 발생하지 않았다. 특히 본 연구는 5개 집단별 차이를 직접 비교하는 것이 아니라, 각각 2주일의 간격을 두고 실시된 온/오프라인 서베이 데이터 사이의 재시험 신뢰도를 통해서 보고 있기 때문에 이러한 한계는 본 논문에 있어서 치명적인 문제점은 아니라고 할 수 있다.

2) 본 연구에서의 설문지는 3page로 구성되었는데 앞부분에 위치시킨 경우는 첫 번째 page에, 중간부분은 두 번째 page에 뒷부분은 마지막 page에 리커트 척도 항목을 위치시킴으로써 조작하였다.

3) 구성형식 4가지와 설문지에서의 위치 3가지를 모두 고려한 4 x 3의 팩토리얼 디자인(factorial design)은 현실적으로 불가능 하였다. 따라서 본 연구에서는 리커트척도의 항목과 기타 항목들 간의 상호작용이라는 순서효과(order effect)가 치명적인 외생변수가 될 수 있는데 본 연구에서 활용된 설문항목들은 이러한 순서효과(설문항목 간의 상호작용 등)에 대한 통제를 고려하여 설정되었다.

〈표 4〉 5가지 설문형식

	설문 형식
A	6문항 묶어서 설문지 앞쪽에 위치
B	6문항 묶어서 설문지 뒤쪽에 위치
C	3문항씩으로 나누어서 설문지 중간에 위치
D	단문항으로 나누어서 설문지 중간에 위치
E	배너형식을 이용하여 설문지 중간에 위치 (오프라인: 6문항 묶어서 설문지 중간에 위치)

〈표 5〉 리커트척도

다음은 프로그램 만족에 관한 내용입니다. 해당되는 것을 선택해 주십시오.

	전혀 아니다	대체로 아니다	보통이다	대체로 그렇다	매우 그렇다
프로그램 수준이 적합하다					
프로그램의 질이 특별히 좋다					
참여로 적절한 시간활용이 되었다					
다른 사람에게 추천할 만한 프로그램이다					
프로그램 참여가 개인적인 도움이 되었다					
전반적으로 프로그램 이용에 만족한다					

일하게 적용하였다. 다만 배너형식의 애니메이션 방식은 오프라인에서는 적용이 불가능하기 때문에 오프라인에서는 6문항을 하나로 묶어서 설문지의 중간에 위치시킨 유형을 이용하였다. 한편, 2차 조사에서 사용된 리커트척도는 '모 프로그램에 대한 만족도'에 대한 것으로서 그 구체적 항목은 〈표 5〉와 같았다.

특히 이러한 4가지 표본 특성 변수와 5개 실험집단 사이의 독립성을 chi-square 검증한 결과 모든 변수에 대해서 통계적으로 유의한 차이를 보이지 않았다. 따라서 본 연구에서의 조사 결과에 있어서 집단별 표본의 특성 차이가 혼재변수(confounding variable)로서 작용했을 가능성은 통제되어 있다고 하겠다.

3.2.3 집단별 조사 대상의 특성

2차 조사에 참여한 조사대상의 특성을 집단별로 구분하여 〈표 6〉에 제시하였다. 〈표〉에서 볼 수 있듯이 본 조사에 참여한 표본은 성별, 연령별, 직업별, 인터넷 사용 경력 별로 고르게 분포되어 있다.

3.2.4 조사결과

2차 조사의 주목적은 1차 조사에서 잠재적인 문제를 보인 리커트척도에 대한 신뢰도를 보다 체계적으로 검증하는 데에 있었다. 따라서 2차 조사에서의 자료에 대한 분석은 5가지유형의 리커트척도 항목에

〈표 6〉 집단별 조사대상의 특성

		조사 대상자의 환경 특성						χ^2	df	P-value
구분	분류	(빈도)%	A그룹	B그룹	C그룹	D그룹	E그룹			
전 체		(500) 100.0%	100 (100.0%)	99 (100.0%)	100 (100.0%)	101 (100.0%)	100 (100.0%)			
성별	남 자	(261)52.2%	(53)53.0%	(51)51.5%	(52)52.0%	(53)52.4%	(51)51.0%	0.099	4	0.999
	여 자	(239)47.8%	(47)47.0%	(48)48.4%	(48)48.0%	(48)48.4%	(49)49.0%			
연령 구분	20세미만	(125)25.0%	(25)25.0%	(25)25.2%	(25)25.0%	(24)23.7%	(24)24.0%	0.576	12	1.000
	20- 30세미만	(154)30.8%	(30)30.0%	(31)31.3%	(31)31.0%	(31)30.6%	(31)31.0%			
	30- 40세미만	(161)32.2%	(32)32.0%	(32)32.3%	(32)32.0%	(32)31.6%	(33)33.0%			
	41- 50세미만	(60)12.0%	(11)12.0%	(11)11.1%	(12)12.0%	(14)13.8%	(12)12.0%			
직업 구분	중학생	(53)10.6%	(10)10.0%	(11)11.1%	(10)10.0%	(11)10.8%	(11)11.0%	1.579	24	1.000
	고등학생	(71)14.2%	(15)15.0%	(13)13.1%	(14)14.0%	(14)13.8%	(15)15.0%			
	대학/원생	(88)17.6%	(17)17.0%	(18)18.1%	(18)18.0%	(19)18.0%	(16)16.0%			
	사무직	(115)23.0%	(23)23.0%	(22)22.2%	(24)24.0%	(23)22.7%	(22)22.0%			
	판매/서비스	(53)10.6%	(10)10.0%	(11)11.1%	(11)11.0%	(10)9.9%	(12)12.0%			
	전문직	(24) 4.8%	(4)4.0%	(5) 5.0%	(4)4.0%	(6)5.9%	(5)5.0%			
	주 부	(96)19.2%	(19)19.0%	(19)19.1%	(19)19.0%	(18)17.8%	(19)19.0%			
인터넷 사용 경력	1년-3년미만	(46)9.2%	(10)10.0%	(10)10.1%	(9)9.0%	(10)9.9%	(8)8.0%	2.048	16	1.000
	3년-6년미만	(165)33.0%	(32)32.0%	(30)30.3%	(34)34.0%	(33)32.6%	(36)36.0%			
	6년-9년미만	(171)34.2%	(34)34.0%	(35)35.3%	(32)32.0%	(35)34.6%	(34)34.0%			
	9년-12년미만	(113)22.6%	(22)22.0%	(22)22.2%	(23)23.0%	(22)21.7%	(21)21.0%			
	12년이상	(5)1.0%	(2)2.0%	(2)2.0%	(1)1.0%	(1)0.9%	(1)1.0%			

대한 전반적인 신뢰도를 살펴보고, 나아가서는 5가지 유형 사이에 신뢰도에서 차이가 존재하는지를 알아보는 데 초점을 맞추어서 진행되었다. 분석결과, 리커트척도 이외의 자료는 1차 조사에서와 마찬가지로 명목 척도 및 등급법에 의한 데이터 모두 온라인과 오프라인 간에 통계적으로 유의한 재시험 신뢰도(test-retest reliability)를 보여서 해당 항목들에 대한 온라인 서베이 데이터의 신뢰도는 검증되었다.

다음으로, 리커트척도 항목의 신뢰도를 내적 일관성 즉 크론바하 알파(Cronbach's α)계수에 의하여 살펴보았다. 〈표 7〉에 제시된 바와 같이 오프라

인 서베이에서는 리커트척도의 구성형식에 관계없이 모두 0.90 이상의 높은 내적 일관도를 보인 반면에, 온라인 서베이에서는 6문항을 모두 묶어서 하나의 응답설명(전혀 아니다 / 매우 그렇다) 아래 제시하는 형식은 0.90에 못 미치는 것으로 나타나서 상대적으로 낮은 신뢰도를 보였다. 특히 6문항을 묶어서 설문지 뒤쪽에 위치시킨 경우는 0.7685로써 절대적인 관점에서도 낮은 내적일관도(채서일, 2004)를 보였다.

이렇게 온라인 서베이에서는 오프라인 서베이에서와는 다르게 설문형식에 따라서 리커트척도의 내적

〈표 7〉 리커트 척도의 내적 일관성(Cronbach's α)

오프라인 서베이(Alpha if item Deleted)					
구분	A그룹	B그룹	C그룹	D그룹	E그룹
	6문항 묶어서 앞쪽에 위치	6문항 묶어서 뒤쪽에 위치	세문항씩 나누어서 중간에 위치	단문항으로 나누어서 중간에 위치	6문항 묶어서 중간에 위치
1	.9637	.9163	.9652	.9699	.9296
2	.9638	.9022	.9597	.9648	.9310
3	.9695	.9093	.9665	.9687	.9381
4	.9694	.9134	.9659	.9710	.9392
5	.9691	.9422	.9663	.9713	.9324
6	.9712	.9426	.9675	.9708	.9490
Alpha	.9738	.9343	.9709	.9744	.9385

온라인 서베이(Alpha if item Deleted)					
구분	A그룹	B그룹	C그룹	D그룹	E그룹
	6문항 묶어서 앞쪽에 위치	6문항 묶어서 뒤쪽에 위치	세문항씩 나누어서 중간에 위치	단문항으로 나누어서 중간에 위치	배너형식 사용하여 중간에 위치
1	.8430	.6533	.9236	.9308	.9201
2	.8304	.6968	.9197	.9272	.9150
3	.8505	.6946	.9254	.9344	.9212
4	.8430	.6901	.9300	.9345	.9224
5	.8816	.7894	.9386	.9353	.9488
6	.9141	.8754	.9466	.9389	.9364
Alpha	.8834	.7685	.9415	.9477	.9376

일관도가 다르게 나타나서 그 원인을 보다 세부적으로 살펴보기 위해서 6개의 세부분항을 문항별로 온라인과 오프라인 서베이 간의 재시험 신뢰도(test-retest reliability)를 분석하였다. 〈표 8〉에 제시된 바와 같이 D그룹(단문항으로 풀어서 제시한 경우)과 E그룹(배너형식을 이용한 경우)을 제외하고는 리커트척도의 뒤쪽에 위치한 세부분항들은 온

라인 서베이와 오프라인 서베이 문항 사이의 재시험 신뢰도가 낮게 나타났다. 특히, B그룹(6문항을 묶어서 설문지 뒤쪽에 위치)의 경우에는 그 문제가 더욱 심각하게 나타났는데 이것이 온라인 서베이에 있어서 6문항을 묶어서 뒤쪽에 제시된 리커트척도의 내적일관도가 낮게 나타난 세부적인 원인으로 분석될 수 있을 것이다. 이는 6문항으로 하나로 묶은 A

〈표 8〉 리커트 척도 세부 문항별 재시험 신뢰도

구분	A그룹	B그룹	C그룹	D그룹	E그룹
	6문항 묶어서 앞쪽에 위치	6문항 묶어서 뒤쪽에 위치	세문항씩 나누어서 중간에 위치	단문항으로 나누어서 중간에 위치	배너형식 사용하여 중간에 위치
1	0.994	0.956	0.970	0.988	0.991
2	0.977	0.917	0.927	0.950	0.959
3	0.933	0.866	0.934	0.906	0.972
4	0.922	0.547	0.872	0.946	0.973
5	0.619	0.468	0.722	0.940	0.968
6	0.717	0.615	0.779	0.956	0.938

그룹과 B그룹에서의 내적일관도 분석에서도 나타나고 있다. 〈표 7〉에서 볼 수 있듯이 오프라인 서베이에서와는 상이하게 온라인 서베이에서는 A그룹과 B그룹의 경우 뒤쪽 문항들인 5문항 6문항을 삭제하면 내적일관도가 오히려 증가하는 것으로 나타나고 있다. 이는 이들 문항들이 리커트척도 항목의 신뢰도에 부정적인 영향을 미치고 있음을 보이는 것이다.

3.3 토의

1차조사와 2차조사를 통하여 살펴본 결과 온라인 서베이 데이터는 같은 패널(동일인)을 대상으로 같은 설문을 이용하여 실시된 오프라인 서베이 데이터와 매우 높은 재시험 신뢰도(test-retest reliability)를 보였다. 이는 온라인 서베이가 마케팅조사 및 여론조사의 자료수집 방법으로서의 오프라인을 통한 서베이 데이터와 유사한 결과를 가져다주는 조사방법으로서 가능성을 보여주는 결과라 할 수 있다.

그러나 본 연구 결과는 온라인 서베이를 통한 데이터의 신뢰도를 높이기 위해서는 오프라인 서베이에

서 보다 상대적으로 높은 주의가 필요하다는 시사점도 아울러 제시하고 있다. 리커트척도를 이용한 항목에서 특히 그러한 것으로 나타났는데, 리커트척도에 의한 항목은 온라인 서베이에서는 단일 항목으로 풀어서 제시하거나 혹은 배너 형식을 이용하여 응답설명이 각각의 세부항목을 따라 내려오게 하는 애니메이션 기법을 활용했을 때 오프라인 서베이에서와 같은 수준의 내적일관도를 보였다. 반면에 온라인 서베이에서는 리커트척도를 형성하는 모든 세부항목을 하나의 응답설명 아래 한꺼번에 제시하는 형태의 설문방식은 상대적으로 낮은 내적일관도를 보였는데 특히 이렇게 묶은 리커트척도를 설문지의 뒤쪽에 위치시켰을 때 이러한 문제는 더욱 심각하게 나타났다. 이러한 연구결과는 설문지의 구성형식이나 디자인에 따라 온라인 서베이의 응답품질 내지는 설문 만족도 등이 달라진다는 기존의 연구들(류진화, 김종훈, 2003; 김광용, 김기수, 1999; Dillmann et al 1998 등)의 연구결과와 리커트척도의 설문지에서의 응답순서가 데이터 신뢰도에 영향을 미칠 수 있다는 Weng과 Cheng(2000)의 연구논리와 일치하는 것인데 그 이유는 온라인 서베이에서는 응답자

들이 컴퓨터 모니터 상에서 설문을 읽게 되기 때문에 설문 디자인이나 설문 제시 방법에 따라서 응답자들이 피로해 하거나 집중력을 잃는 정도가 다르기 때문일 것이다(김광용, 김기수 1999).

IV. 결론

4.1 결론

본 연구에서는 마케팅조사 및 여론조사에서 그 비중이 점점 증대되고 있는 온라인 서베이에 의한 데이터의 신뢰도를 오프라인 서베이를 통해서 얻은 데이터와의 비교를 통해서 검증하였다. 조사는 두 차례에 걸쳐서 진행되었는데 1,2차 조사 모두 동일인을 대상으로 온라인과 오프라인 서베이를 2주일의 간격을 두고 실행하여 재시험 신뢰도(test-retest reliability)를 분석함으로써 살펴보았다.

1차조사와 2차조사 모두 온라인 서베이 데이터와 오프라인 서베이 데이터 사이에 높은 재시험 신뢰도를 보였다. 그러나 1차조사에서는 리커트척도에 의한 문항이 온라인 서베이에서 잠재적인 문제점을 가질 수 있음이 발견되었다. 따라서 2차조사에서는 리커트척도의 구성 형식이나 설문지에서의 위치에 따라서 5가지 유형의 리커트척도 문항을 사용하여 분석하였다. 분석결과 온라인 서베이에서는 리커트척도 문항을 사용할 때 세부분항들을 단문항으로 각각 나누어서 제시하거나 혹은 응답설명(전혀아니다 / 매우그렇다)이 배너형식으로 처리되어서 세부분항을 계속 따라 내려오도록(scroll) 애니메이션 기법을 활용할 때 더 높은 내적일관도(Cronbach's α)와 오프라인 서베이 자료와의 재시험 신뢰도를 보였다.

결과적으로 본 연구에 의하면 온라인 서베이는 설문 디자인과 설문항목의 제시 방법만 체계적으로 관리한다면 오프라인 서베이를 대체할 만한 신뢰도 있는 방법론이라고 할 수 있다. 따라서 온라인 서베이는 표본의 대표성 문제라는 본 연구에서 검증하지 못한 한계에 대한 보완만 이루어진다면 그 신속성과 저렴성이라는 장점을 고려할 때 서베이 방법으로서 활용도가 매우 높은 방법론이라고 할 수 있다.

4.2 연구의 한계 및 향후 연구 방안

본 연구는 탐색적 연구의 수준에서 이루어졌다는 것이 가장 크고 치명적인 한계이다. 즉 온라인 서베이 데이터의 신뢰도에 영향을 줄 수 있는 구성개념(construct)들을 이론적 배경에 근거해서 도출하고 도출된 변수들에 의해서 설정된 연구 모형과 연구가설에 의해서 진행되지 못하였다. 따라서 본 연구에서의 결론 또한 구체적인 연구 변수들에 대한 인과관계(causal relationship)를 밝히지 못하고 단순한 현상만을 제시하고 있으며 그러한 현상에 대한 이유 또한 유추·해석의 수준에 머물고 있다.

그러나 본 연구는 연구자들이 아는 한 동일인들을 대상으로 같은 설문내용을 이용하여 온라인 서베이와 오프라인 서베이를 실행하여 이들 간의 신뢰도를 검증한 최초의 연구라는 면에서 그 의의를 찾을 수 있을 것이다. 따라서 향후 연구에서는 본 연구에서의 연구결과에 근거해서 구체적인 구성 개념들이 제시되고 그들 간의 인과관계를 과학적으로 밝히는 인과관계조사(causal research)가 진행되어야 할 것이다.

참고문헌

- 김광용, 김기수 (1999b), "다양한 인터넷 설문방식 비교와 설문만족에 관한 실증적 연구." 1999년 춘계 경영정보학회 학술대회, 134-141.
- 김중훈, 류진화, (2002), "전자우편 설문조사 반응에 관한 문헌적 고찰." *조사연구*, 제3권 제2호: 91-122.
- 류진화, 김중훈, (2003), "전자우편을 이용한 설문조사시 응답률, 응답속도, 자료의 질에 영향을 미치는 요인들에 대한 연구." *경영학연구*, 제32권 제2호, 475-498.
- 명승화, 최영훈 (2002), "인터넷조사방법론의 가능성과 쟁점: 행정학회 회원들에 대한 설문조사결과를 토대로." *한국행정학보*, 제36권 제2호: 333-351.
- 박무익 (2000), "우리나라 조사산업의 현황과 전망." *한국조사연구학회, 조사연구*, 1권 2호, 161-177.
- 채서일 (2004), *사회과학 조사방법론* 3판, 학현사.
- Aiken, L. R.(1983), Number of Response Categories on a Teacher Rating Scale. *Educational and Psychological Measurement*, 43, 397-401.
- Bachmann, Duane, John Elfrink, and Gary Vazzana (1996), "Tracking the Progress of E-mail versus Snail-mail." *Marketing Research*, 8, 31-35.
- Batagelj, Zenel, Katja Lozar, and Vasja Vehovar. (1998). "Who are Nonrespondents in Web Surveys?" *9th International Workshop on Household Survey Nonresponse Bled*.
- Bowker, Dennis K., and Don A. Dillman. (2000). "An Experimental Evaluation of Left and Right Oriented Screens for Web Questionnaires." <http://survey.sesrc.wsu.edu/dillman/papers.html/>.
- Comley, P., (1998), "On-line Research, Some Options, Some Problem, Some Case Studies." *ASC Conference*.
- Dickinson, T. L., & Zellinger, P. M. (1980), A Comparison of the Behaviorally Anchored Rating and Mixed Standard Scale Formats. *Journal of Applied Psychology*, 65, 147-154.
- Dillman, Don A. (1978), *Mail and Telephone Surveys: The Total Design Method*, New York: John Wiley and Sons.
- Dillman, Don A., and James H. Frey (1974), "Contribution of Personalization to Mail Questionnaire Response as an Element of a Previously Tested Method.", *Journal of Applied Psychology*, 59 (3): 297-301.
- Dillman, Don A., Robert D. Tortora, and Dennis K. Bowker. (1998). "Influence of Plain vs. Fancy Design on Response Rates for Web Surveys." [Http://survey.sesrc.wsu.edu/dillman/papers.html/](http://survey.sesrc.wsu.edu/dillman/papers.html/).
- Dixon, P. N., Bobo, M., & Stevick, R. A.(1984). Response Differences and Preferences for all Category-Defined and End-Defined Likert formats. *Educational and Psychological Measurement*, 44, 61-66.
- Etzel, Michael J., and Bruce J. Walker. (1974). "Effects of Alternative Follow-up Procedures on Mail Survey Response Rates." *Journal of Applied Psychology* 59 (2): 219-221.
- Ford, Neil M. (1967). "The Advance Letter in Mail Surveys." *Journal of Marketing Research* 4(May): 202-204.
- French-Lazovik, G., & Gibson, C. L.(1984), Effects of Verbally Labeled Anchor Points on the Distributional Parameters of Rating Measures. *Applied Psychological Measurement*, 8, 49-57.
- Johnson, S. M., Smith, P. C., & Tucker, S. M. (1982), Response Format of the Job Descriptive Index: Assessment of Reliability and Validity by the Multitrait-Multimethod

- matrix. *Journal of Applied Psychology*, 67, 500-505.
- Kanuk, Leslie, and Conrad Berenson. (1975). "Mail Surveys and Response Rates: A Literature Review." *Journal of Marketing Research* 7(November): 440-453.
- Kiesler, S., and L. Sproull. (1986). "Response Effects in the Electronic Survey." *Public Opinion Quarterly* 50, 402-413.
- Ko, Y-H.(1994), A Search for a Better Likert Point-scale for Mental Health Questionnaires. *Psychological Testing*, 41, 55-72.
- Krosnick, J. A. (1999), Survey Research. *Annual Review of Psychology*, 50, 537-567.
- Likert, R. (1932), A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22, 5-55.
- Likert, R., Roslow, S., & Murphy, G. (1934), A Simplified and Reliable Method of Scoring the Thurstone Attitude Scales. *Journal of Social Psychology*, 5, 228-238.
- Metha, Raj, and Eugene Sivadas. (1995). "Comparing Response Rates and Response Content in Mail versus Electronic Mail Surveys." *Journal of the Market Research Society* 37: 429-39.
- Oaster, T. R. F. (1989), Number of Alternatives per Choice Point and Stability of Likert-type Scales. *Perceptual and Motor Skills*, 68, 549-550
- Opperman, Martin. (1995). "E-mail Survey: Potential and Pitfalls." *Marketing Research* 7, 29-33.
- Peterson, Robert A. (1975). "An Experimental Investigation of Mail-Survey Responses." *Journal of Business Research* 3(3): 199-208.
- Reynolds H. T. (1977). "Analysis of Nominal Data, 2nd Edition Reynolds" Sage Publication.
- Scott, Christopher. (1961). "Research on Mail Survey." *Journal of the Royal Statistical Society* 124: 143-191.
- Schaefer, David R., and Don A. Dillman. (1998). "Development of A Standard E-mail Methodology." *Public Opinion Quarterly* 62 (Fall): 378-97.
- Schuldt, Barbara A., and Jeffrey W. Totten. (1994). "Electronic Mail versus Mail Survey Response Rates." *Marketing Research* 6: 36-39.
- Sheehan, Kim Bartel, and Mariea Grunns Hoy, (1999), "Flaming, Complaining, Abstaining: How Online Users Respond to Privacy Concerns." *Journal of Advertising* 28: 37-52.
- Singer, Eleanor, John van Hoewyk, and Mary P. Maher (1998), "Does the Payment of Incentives Create Expectation Effects?" *Public Opinion Quarterly*, 62, 152-64.
- Smith, Christine B. (1997) "Casting the Net: Surveying an Internet Population." *Journal of Communication Mediated by Computers*.
- Solomon, M. B., (1995), "Marketing Research in Cyberspace.". (<http://www.cyberdialogue.com/press/articles/marketingcyberspace.html>)
- Stanton, J. M. (1998). "An Empirical Assessment of Data Collection Using the Internet." *Personnel Psychology*, Vol, 51(3).
- Taylor, H. (2000) "Does Internet Research Work? Comparing Online Survey Results with Telephone Survey." *International Journal of Marketing Research*, Vol.42 Issue 1.
- Tse, A. C. B. et al., (1995), "Comparing Two Method of Sending out Questionnaires: Email versus Mail.," *Journal of the Market Research Society*, Vol. 37, 429-439.
- Tse, Alan C. B. (1998). "Comparing the Response Rate, Response Speed and Response Quality

- of Two Methods of Sending Questionnaires: E-mail vs. Mail." *Journal of the Marketing Research Society* 40: 353-361.
- Weng, Li-jen & Chung-ping Cheng. (2000). Effects of Response Order on Likert-type Scales. *Educational and Psychological Measurement*, Vol. 60 No. 6(December).
- Wildt, A.R., & Mazis, M. B. (1978), Determinants of Scale Response: Label versus Position. *Journal of Marketing Research*, 15, 261-267.
- Wiseman, Frederick. (1973). "Factor Interaction Effects in Mail Survey Response Rates." *Journal of Marketing Research* 10 (August): 330-333.

An Exploratory Study on Reliability between On-line and Off-line Survey Data Using a Single Panel

Ick-Hyun Kwon* · Sang-Won Lee**

Abstract

This study examined reliability of on-line survey data by checking test-retest reliability between on-line survey data and off-line survey data using a single panel. The empirical results showed that test-retest reliabilities were significantly high enough to conclude that on-line survey provided reliable data in relation to off-line survey.

The use of the Internet as research methodology grew phenomenally since 1990's. With the explosive growth of e-mail and web sites, the Internet could serve as the ideal medium for sending and receiving surveys, potentially replacing off-line media such as mail and telephone. Even with tremendous potential, on-line survey also had been criticized for several sources of survey error. Among quite a few error sources, sampling error including coverage error and nonresponse error and questionnaire design error had been investigated.

The main focus of this study was on the latter one, questionnaire design as the on-line survey error source. In this regards, previous research had examined e-mail only vs. e-mail with Web link, screen vs. scroll layout, plain design vs. fancy design, etc. Another contribution of this study was methodology examining on-line survey data reliability. For the best knowledge of researches, this study was the first study examining on-line survey data reliability by checking test-retest reliabilities between on-line survey and off-line survey data using a single panel with the same questionnaire.

Two separate empirical researches were conducted in this study. The first study exploratory examined test-retest reliabilities of nominally scaled questions and ordinally scaled questions as well as Likert's scale questions. The interval between the on-line survey and off-line survey

* Professor, School of Management, Dongguk University.

** Synergy 21 Co.,Ltd.

had been two weeks considering extraneous variables. The general findings told us that on-line survey data were reliable in relation to off-line survey data. In the meantime, there also were noticeable findings concerning Likert's scale data such that test-retest reliabilities were decreasing from the first question to the last questions.

To investigate this finding, the second research was implemented. In addition to replications of the first research, in the second research, five different forms of questionnaire were manipulated with respect to the way of scale point explanation and the location of Likert's scale in the questionnaire.

The empirical results showed that test-retest reliabilities of all the questions were consistently high when scale point explanation was scrolled along questions as a banner and when each question was separated individually with scale point explanation. When all the Likert's scale questions were aggregated as in a traditional way, test-retest reliabilities between on-line and off-line data were relatively low and showed decreasing pattern.

In conclusion, on-line survey data were reliable in relation to off-line survey data using the same questionnaire with a single panel. Meanwhile the empirical results also showed that researchers should be cautions in designing questionnaire especially for Likert's scale questions in the on-line survey.

Key words: on-line survey, off-line survey, panel, test-retest reliability, likert scale.