



제1장

인터넷조사와 가구방문조사간 데이터 품질 비교

- 사회조사를 대상으로 -

임 경 은

제1절 서론

조사방법은 시대의 흐름에 따라 빠르고 다양하게 변화하고 있다. 대인면접법과 더불어 전화조사나 우편조사 등으로 대표되던 조사방법이 PDA나 컴퓨터를 활용하는 방법으로 발전하고 있으며, 인터넷 이용 인구의 폭발적인 증가에 따라 인터넷조사에 대한 관심도 증폭되고 있다. 인터넷조사는 다른 조사방법에 비하여 조사에 걸리는 시간이 짧아 신속하게 대량의 자료수집이 가능하며, 편한 시간을 이용한 유동적 조사 수행이 가능하고, 조사 비용도 크게 절약할 수 있어 경제적이다. 또한 응답자가 본인이 원하는 시간에 자유롭게 응답할 수 있으므로 다른 조사에 비해 조사환경의 제약을 덜 받게 되며, 개인 정보 노출을 꺼리는 응답자의 참여를 자연스럽게 유도할 수 있어 민감한 문제 등에 대한 조사에 유용하게 이용될 수 있다. 뿐만 아니라 면접조사에서 발생할 수 있는 면접원에 의한 측정오차(measurement error)가 발생하지 않으며, 시청각자료의 활용이 용이함으로 응답에 필요한 정확한 정보 전달이 가능하다. 더불어 복잡하게 설계된 질문지를 이용하여 조사를 진행하는 경우 문항 간 건너뛰기 조절 등 기술적인 지원이 가능하므로 다른 조사방법에 비해 착오에 의한 오류는 적어지고, 자료 입력 단계에서의 입력오류(처리오차; process error)는 발생하지 않는다. 이와 같이 인터넷조사는 다양한 형태의 장점을 가지고 있으며, 이러한 장점들은 기존의 조사방법이 가지고 있는 여러 가지 문제점들에 대한 해결 방향으로 제시되고 있다(Kieslet 외, 1984; Schillewart 외, 1998; Sheehan 2002; 문시연 2005).

그러나 인터넷조사는 표본이 모집단에 대한 대표성을 가지지 못한다는 치명적인 단점도 가지고 있다(Tse, 1998; Hoppe 외, 2001). 즉, 표본조사에서의 목표모집단(target population)과 인터넷조사에서 실제 이용되는 모집단(frame population)이 불일치한다는 문제가 있다. 이는

인터넷조사에 응답하는 응답자들이 자발표본이거나 편의 표집에 의한 경우가 많기 때문에 발생하는 현상으로, 인터넷 이용자들을 대상으로 하는 표본추출틀(sampling frame)이 목표 모집단을 대표하지 못하는 데서 발생하는 문제이다. Couper(2000)는 전통적인 표본조사에서의 표본 대표성 문제를 해결하기 위하여 여러 확률표본추출방법이 개발되었듯이 인터넷조사에서의 표본 대표성 문제를 해결하기 위한 표본추출방법이나 가중치 조정에 관련된 연구가 필요하다고 주장하였다. 그러나 인구사회학적 특성 가중치만을 부여한 인터넷조사 결과를 일반적인 인구의 특성으로 일반화시킬 수 없다는 문제가 많은 연구자들에 의해 지속적으로 제기되었다. Rosenbaum 외(1983)는 이와 같은 문제를 해결하기 위한 방안으로 성향가중모형을 이용한 보정방법을 제안하였으며, 이에 대한 연구가 지금까지 여러 형태로 지속되고 있다. Taylor(2000, 2001)는 Harris 여론조사회사에서 수행한 전화조사와 인터넷조사의 결과를 비교함으로써, 두 조사 사이에 공통의 특성을 이용한 성향가중조정방법으로 목표모집단과의 차이가 거의 사라진다는 것을 확인하였으며, Rubin(1997)과 D'Agostino 외(2000)는 성향가중조정변수로 선정되기 위해서는 예측변수와외의 공변량이 중요하다는 것을 보였다. Tagott(2001)는 다양한 기준에 의한 정확도 평가에서 거의 모든 분야에 걸쳐 인터넷조사가 좋다는 결론을 얻었으며, 이지영 외(2006)는 온라인 패널조사와 오프라인 패널조사의 데이터 품질을 비교하여 온라인조사가 오프라인조사에 대한 대등한 품질을 가지며, 특정 형태의 조사에서는 더 우수한 품질을 가진다는 것을 보였다. 이외 인터넷조사와 관련된 연구로는 인터넷조사의 정확도에 미치는 영향(조성겸 외, 2005), 응답오차에 대한 연구(이윤석 외, 2008) 등이 있다. 인터넷조사는 패널에 대한 기본 정보를 이용하여 매 조사 시 적합한 조사 대상자를 찾아낼 수 있다는 장점이 있으므로, 마케팅 관련 조사나 특정 층을 목표로 하는 조사에서 기존의 조사 방식을 대체할 수 있는 좋은 대안으로 활용될 수 있다(Bandilla 외, 2003; Braunsberger 외, 2007).

2009년 7월 시행된 사회조사에서는 기존의 면접조사방법으로 조사를 진행함과 동시에 인터넷 자원자를 모집하여 인터넷조사를 수행하였다. 본 연구에서는 두 가지 조사방법을 통해 수집된 사회조사 결과를 이용하여 인터넷조사를 통한 자발적 표본의 대표성 문제를 짚어보고, 인터넷조사를 통해 수집한 자료가 기존의 면접조사방법에 의해 수집된 자료에 비하여 통계적 품질 면에서 크게 뒤지지 않는다는 것을 보이하고자 한다. 또한 설문응답시간과 패턴이 응답의 성실성에 얼마나 영향을 미치는지 등을 살펴보아 궁극적으로 국가통계에 인터넷조사방법을 적용하는 문제에 대하여 고민해보고자 한다.

제2절에서는 인터넷조사를 이용한 사회조사에 대하여 알아보고, 제3절에서는 인터넷조사와 기존 면접조사에서의 표본 분포 형태를 비교해본다. 제4절에서는 인터넷조사를 통해 수집된 데이터의 품질 평가를 타당도와 신뢰도 측면에서 점검하게 된다. 이어 제5절에서는 인터넷조사에서의 정확도 관련 요인에 대한 특성과 해결 방안에 대하여 논의



하며, 제6절에서는 설문응답시간과 응답성실성의 관계를 탐색한다. 마지막으로 제7절에서는 분석 결과를 요약하고 연구의 한계점 및 향후 과제를 제시한다.

제2절 인터넷조사를 이용한 사회조사

사회조사는 국민의 삶의 질과 관련된 사회적 관심사와 주관적 의식에 관한 사항을 조사하여 삶의 수준과 사회적 변동을 파악하고 이를 사회 개발 정책 수립의 기초 자료로 제공하기 위한 목적으로 매년 조사되고 있다. 1977년 3월 조사가 시작된 이래 「한국의 사회지표」 체계를 구축하는데 중요한 역할을 하고 있으며, 가족, 소득과 소비, 노동, 교육, 보건, 주거와 교통, 정보와 통신, 환경, 복지, 문화와 여가, 안전, 사회참여 등의 부문으로 나누어 조사하고 있다. 1998년 이후 매년 3~4부문이 3년 주기로 조사되던 사회조사는 2008년부터 부문별 2년 주기로 전환하여 실시되고 있으며, 조사결과는 각 부문별 질문 항목에 대한 응답 비율 등으로 집계하여 인구사회학적 특성별로 공표하고 있다.

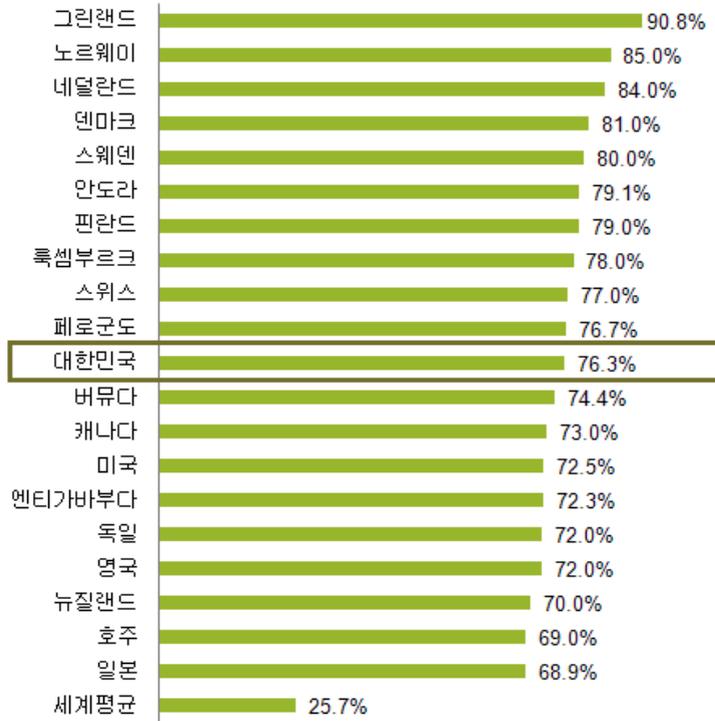
사회조사에서는 확률표본추출방법을 이용하여 약 20,000개의 표본가구를 추출하고, 이들 표본가구를 대상으로 가구방문조사를 수행하고 있다. 그러나 가구방문조사는 비용과 시간이 다른 조사방법에 비해 많이 든다는 단점과 함께 숙달된 면접원의 부족, 1인 가구 및 맞벌이 부부의 증가로 인한 조사 대상자에 대한 접근의 어려움, 조사 거절 등의 문제점도 가지고 있다(강남준 외, 2004, 조성겸 외, 2003, 이홍철, 2002, 홍내리 외, 2001). 이러한 조사환경의 변화로 인한 어려움을 근본적으로 해결할 수 있는 방법은 없는 상황이며, 여러 보정 방법이나 조사 수행에서의 보조 장치 이용 등으로 문제를 최소화하고 있는 실정이다. 따라서 기존의 조사방법을 대신할 수 있는 대안적 조사방법에 대한 관심이 커지고 있으며, 이 중 인터넷조사에 대한 관심은 정보통신분야의 비약적 발전에 힘입어 빠르게 증가하고 있다.

1. 인터넷조사환경

2000년대에 들어서면서 우리나라뿐 아니라 전 세계의 정보통신분야는 눈부신 발전을 이루고 있다. 혁신적인 정보통신의 발전은 사회·문화·경제에 걸쳐 복합적이면서도 빠르게 진행되고 있다. 이와 같은 상황에서 기존의 면접조사나 전화조사 그리고 우편조사 등을 통한 조사는 점진적으로 설 자리를 잃어가고 있으며, 이들에 대한 대안적 조사방법으로 인터넷조사에 관심이 모아지고 있다.

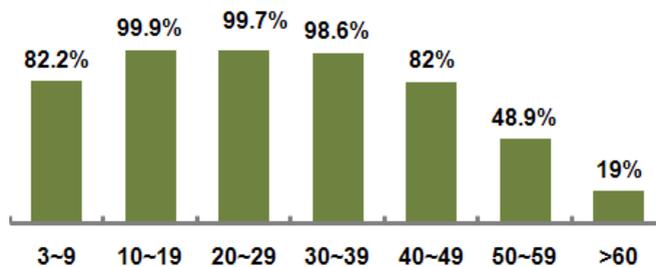
우리나라의 인터넷보급률은 2007년 현재 76.3%에 달하고 있으며, 인터넷 이용자수는

2008년 6월 현재 3,619만 명에 육박하고 있다(ISIS 인터넷통계정보검색시스템).



출처: ISIS 인터넷통계정보검색시스템(2007)

[그림 1-1] 세계의 인터넷 보급률



출처: ISIS 인터넷통계정보검색시스템(2008.6)

[그림 1-2] 연령대에 따른 인터넷 이용률



정보통신기술의 비약적 발전과 더불어 진행되는 조사환경의 변화는 인터넷조사에 대한 관심을 크게 증가시키는 계기가 되고 있으며, 이미 인터넷조사를 활용한 표본조사는 우리 주변에서 쉽게 접할 수 있는 조사방법 중 하나로 자리 잡고 있다. 특히 마케팅 및 사회여론조사, 또는 특정 연령대(특히 30대 이하)를 목표모집단으로 하는 조사의 경우에는 인터넷조사가 적극 활용될 수 있으며, 그 결과의 정합성이나 조사의 효율성이 매우 뛰어난 것으로 보고 있다(김영원 외, 2000). 그러나 표본의 대표성이나 자료의 품질 문제 등에 있어서는 낙관적인 시각과 비관적인 시각이 동시에 공존하고 있는 상황이다. 따라서 본 연구에서는 이번 인터넷조사를 이용한 사회조사 결과를 바탕으로 인터넷조사의 품질에 대한 평가를 수행하고 이어 인터넷조사를 활용한 국가통계작성의 가능성을 확인해보고자 한다.

2. 인터넷조사를 활용한 사회조사

통계청 사회복지통계과에서는 기존의 면접조사와 함께 인터넷 자원자를 대상으로 인터넷을 통한 사회조사(이하 인터넷조사)를 수행하였다. 인터넷조사를 통하여 수집된 자료는 자원자의 성향 분석 및 가중치 산정 방안을 연구하고 면접조사와 인터넷조사의 품질을 비교하는데 이용되며, 해당 연구 결과를 바탕으로 향후 사회조사에의 인터넷조사 적용 가능성을 검토하고자 한다.

본 절에서는 사회복지통계과의 2009 인터넷조사방법 효과분석 보고서의 내용을 토대로 인터넷조사에 관한 전반적인 내용을 소개하도록 한다.

1. 기본계획 수립	<ul style="list-style-type: none"> - 인터넷조사 방안 검토 · 기본방향 확정 · 지방청 및 청내 관계자 의견수렴 	-1~2월
2. 세부실시계획 수립	<ul style="list-style-type: none"> - 세부 시행계획 수립 · 실시과정, 자원자모집방안 - 인터넷조사시스템 개발 	-3~5월
3. 홍보	<ul style="list-style-type: none"> - 보도자료 배포(5.12) - 통계청 관련 사이트 등재·홍보 - 경품 사이트 등록·홍보 	-5~6월
4. 인터넷자원자 모집	<ul style="list-style-type: none"> - 실명인증 - 자원자 가구 및 가구원 정보 등록 - 주소 확보 	-5.12~ 6.10

5. 인터넷표본가구 추출	- 인터넷자원자 표본명부 확정 · 중복응모자 등 제외 - 인터넷표본가구 및 대체 명부 확보	-6.12~ 6.22
6. 인터넷표본가구 선정통보 및 확정	- 인터넷표본가구 선정 알림(이메일, 문자) · 참가여부 회신 시 표본으로 확정 - 참가여부 회신 없는 가구는 동일 특성 가구로 대체 (4차)	-6.23~ 7.3
7. 조사 실시	- 인터넷조사 시작 안내(이메일, 문자) - 입력현황 관리 · 3~4회에 걸쳐 미참여자에게 조사 참여 독려(문자)	-7.6~ 7.20
8. 조사답례품 발송	- 입력가구원수 확인 및 답례품 발송	-7월 말
9. 자료 검토	- 입력내용 검토 - 산업 및 직업 코드 점검 - 결과 집계 및 분석	-8~9월
10. 효과 분석	- 문제점 발굴 및 시사점 제시 - 향후 개선방안 마련	-9~10월

[그림 1-3] 인터넷조사의 추진절차 및 일정

2009년 7월 진행된 인터넷조사는 응답자원자 7,619명 중 중복응모, 무응답자를 제외한 6,856명을 모집단으로 하였으며, 지역, 성, 연령대를 기준¹⁾으로 하는 균형표본추출방법으로 약 1,500가구를 추출하였다. 1차 표본으로 선정된 1,500가구에는 이메일과 문자 메시지를 발송하여 선정 여부를 통보하였으며, 참여 회신을 받으면 인터넷 표본가구로 확정, 그렇지 않은 경우에는 2, 3, 4순위 예비표본가구로 대체 통보하였다. 이러한 과정을 통하여 1,547가구가 최종 인터넷 표본가구로 확정되었으며, 인터넷조사 시스템을 통하여 가구원 정보 및 조사표를 직접 입력하는 자기기입식으로 조사가 진행되었다. 조사 항목은 기본사항 7문항, 문화와 여가 13문항, 소득과 소비 4문항, 사회참여 11문항, 삶의 질 만족도 16문항, 관리사항 6문항 등 총 57개 문항으로 면접조사의 86개 항목에 비해 적은 수로 구성되었으며, 15일 간에 걸쳐 진행되었다. 면접조사와 인터넷조사의 조사방법에 대한 비교표는 <부록 1>과 같으며, 인터넷조사 자원자 모집 결과는 <부록 2>와 같다 (2009 인터넷조사방법 효과분석, 사회복지통계과).

1) 균형기준: ① 광역시·도(16개), ② 동·읍(2개), ③ 성별(2개), ④ 연령대(5개)



제3절 표본 분포

1. 표본의 대표성

인터넷조사의 경우 응답자들의 자발적 조사 참여와 편의표본에 의한 표본의 대표성 문제가 가장 큰 한계점으로 제기되고 있다. 그러나 인터넷조사와 전통적인 표본조사에서의 결과 차이가 모두 인터넷조사의 표본 대표성 문제에서 발생한다고 보기는 어렵다. 왜냐하면 표본 대표성 문제는 인터넷조사뿐 아니라 기존의 전통적인 표본조사에도 발생되기 때문이다. 먼저 각 조사방법에 따른 표본추출방법에 대하여 자세히 알아보고, 이어 면접조사와 인터넷조사에서의 인구사회학적 분포를 인구추계에서의 분포와 비교하여 표본 대표성에 대하여 비교해보도록 한다.

면접조사 방식의 기존 사회조사 표본은 2005년 인구주택총조사 10% 표본조사구 중 아파트와 보통조사구를 대상으로, 전국을 25개 지역²⁾으로 층화한 뒤 분류지표³⁾에 따라 정렬하여 확률비례계통추출방법(PPS_sys: systematic selection with probability proportional to size)에 의해 표본조사구를 추출하게 된다. 추출된 표본조사구 내 가구에는 일련번호가 부여되며, 무작위로 하나의 가구를 추출한 후 해당 가구를 포함한 15가구를 연속적으로 조사하는 방식으로 표본가구를 확정하게 된다. 이 때 조사 대상 가구 내 만 15세 이상 가구원이 조사 단위가 되며, 일부 항목에 대해서는 만 15세 미만 가구원도 포함하게 된다(2008년 사회조사 메타자료).

인터넷조사 방식의 사회조사 표본은 약 한 달(2008.5.12 - 6.10) 동안 인터넷 자원자 표본을 모집하는 형식으로 수집되었다. 모집 결과, 6,856건의 자원자가 응모하였으며, 이 중 1,500가구를 Rim 가중법에 의한 균형계통추출방법(balanced systematic sampling)으로 추출하였다. 이렇게 추출된 1,500개 가구에는 사회조사 참여에 대한 이메일과 문자메세지가 발송되었으며, 이 중 62.27%인 934개 가구가 응답하였다. 일주일 뒤 다시 사회조사 참여에 대한 독려 이메일과 문자메세지가 발송되었으며, 이때도 응답하지 않은 경우에 한하여 하루 간격으로 두 차례의 이메일과 문자메세지가 추가적으로 발송되었다. 그 결과 조사 참여 대상자로 추출된 전체 1,500가구 중 98.47%인 1,477가구가 조사에 참여하였으며, 이 후 추가된 70개 가구를 포함하여 전체 1,547개의 가구를 대상으로 조사가 진행되었다.

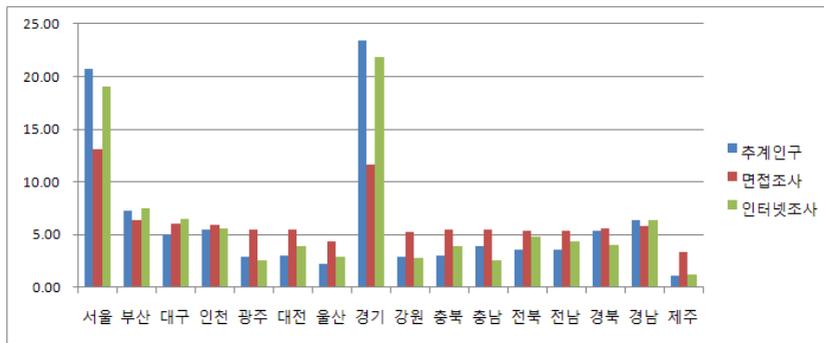
다음은 지역과 조사방법에 따른 인구분포를 나타낸 표와 그래프이다.

2) 16개 시도와 동부 및 읍면부

3) 1차: 주택유형, 2차: 농가 구성비, 3차: 교육정도, 4차: 60세 이상 노인 구성비, 5차: 행정구역 및 조사구번호

<표 1-1> 지역과 조사방법에 따른 인구분포 현황

지역	추계인구*		면접조사**		인터넷조사**	
	추계인구*	추계인구*	면접조사**	면접조사**	인터넷조사**	인터넷조사**
서울	8,328	20.76	4,758	13.06	554	19.08
부산	2,899	7.22	2,330	6.39	217	7.48
대구	2,036	5.07	2,190	6.01	190	6.54
인천	2,196	5.47	2,153	5.91	161	5.55
광주	1,148	2.86	1,987	5.45	74	2.55
대전	1,219	3.04	2,001	5.49	112	3.86
울산	901	2.25	1,588	4.36	85	2.93
경기	9,414	23.46	4,242	11.64	634	21.84
강원	1,178	2.94	1,904	5.23	81	2.79
충북	1,220	3.04	1,994	5.47	114	3.93
충남	1,590	3.96	1,992	5.47	74	2.55
전북	1,418	3.53	1,947	5.34	141	4.86
전남	1,424	3.55	1,965	5.39	127	4.37
경북	2,168	5.40	2,048	5.62	118	4.06
경남	2,553	6.36	2,128	5.84	186	6.41
제주	433	1.08	1,208	3.32	35	1.21
계	40,125	100.00	36,435	100.00	2,903	100.00



[그림 1-4] 지역과 조사방법에 따른 인구분포 현황

<표 1-1>과 [그림 1-4]를 통하여 면접조사와 인터넷조사를 이용한 사회조사의 지역별

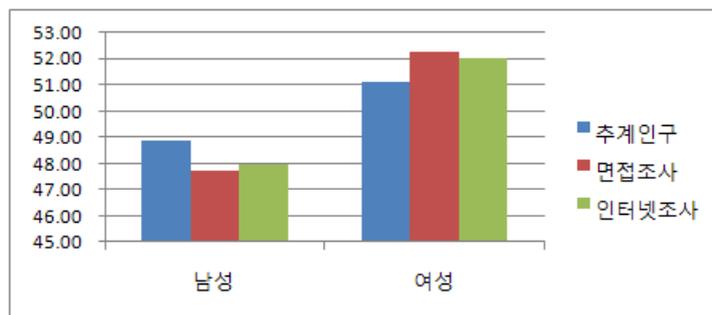


분포를 알 수 있다. 두 조사방법에서의 지역별 분포를 추계인구와 비교하면, 인터넷조사의 지역별 분포가 면접조사에 비하여 추계인구와 더 유사하다는 것을 알 수 있다. 구체적으로 살펴보면, 면접조사의 경우 서울 및 경기지역의 비율이 추계인구나 인터넷조사에 비하여 매우 낮게 나타나며, 광주, 대전, 강원, 충청남·북도의 경우에는 비교적 높게 나타남을 알 수 있다. 이는 면접조사 방식의 표본 추출 시 지역별 층을 이용한 확률비례계통추출방법과 더불어 각 지방청(사무소)의 인력과 예산이 고려되기 때문인 것으로 보인다. 이에 비해 인터넷조사의 경우 균형계통추출방법을 이용하여 표본을 추출하였으며, 자발적으로 참여 의사를 밝힌 조사 희망자를 대상으로 확률표본추출방법을 적용한 것이외의 고려사항은 없었다. 따라서 면접조사에 비하여 인터넷조사의 지역별 표본 분포가 인구추계의 지역별 분포와 더욱 유사하게 나타난 것으로 보인다.

다음은 성별과 조사방법에 따른 인구분포 현황이다.

〈표 1-2〉 성별과 조사방법에 따른 인구분포 현황

성별	추계인구*		면접조사**		인터넷조사**	
	천 명	%	천 명	%	천 명	%
남성	19,614	48.88	17,392	47.73	1,393	47.98
여성	20,511	51.12	19,043	52.27	1,510	52.02
계	40,125	100.00	36,435	100.00	2,903	100.00



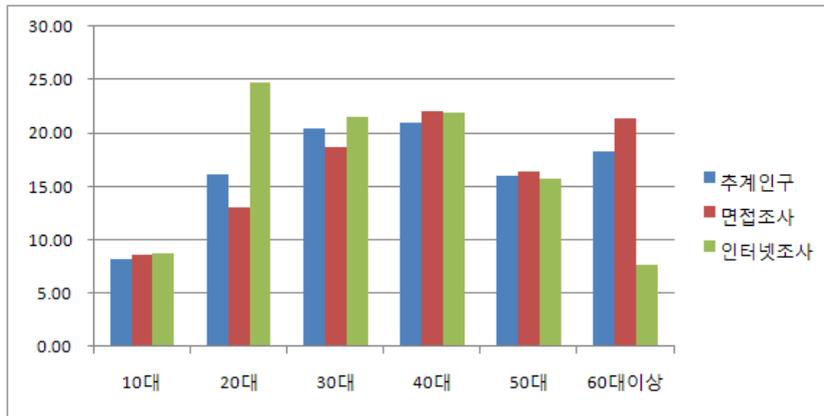
[그림 1-5] 성별과 조사방법에 따른 인구분포 현황

전반적으로 인터넷조사와 면접조사의 성별 분포가 추계인구와는 별다른 차이가 없는 것으로 보이나, 지역별 분포에서와 마찬가지로 면접조사에 비하여 인터넷조사의 성별 분포가 인구추계의 성별 분포와 조금 더 유사하다는 것을 알 수 있다.

다음은 연령대와 조사방법에 따른 인구분포 현황이다.

<표 1-3> 연령대와 조사방법에 따른 인구분포 현황

연령대	추계인구*		면접조사**		인터넷조사**	
	추계인구*	비율(%)	면접조사**	비율(%)	인터넷조사**	비율(%)
10대	3,288	8.19	3,107	8.53	252	8.68
20대	6,490	16.17	4,735	13.00	716	24.66
30대	8,184	20.40	6,818	18.71	623	21.46
40대	8,405	20.95	8,038	22.06	634	21.84
50대	6,414	15.99	5,949	16.33	455	15.67
≥60대	7,344	18.30	7,788	21.38	223	7.68
계	40,125	100.00	36,435	100.00	2,903	100.00



[그림 1-6] 연령대와 조사방법에 따른 인구분포 현황

인터넷조사의 경우 20대의 비율이 다른 연령대에 비하여 높게 나타났으며, 60대 이상의 경우에는 낮게 나타남을 알 수 있다. 이를 통하여 인터넷을 이용한 조사에 연령 편중 현상이 발생하고 있음을 알 수 있으며, 특히 60대 이상의 표본이 작아 해당 연령대에 대한 대표성 문제가 발생할 가능성이 있음을 알 수 있다. 그러나 20대와 60대 이상을 제외한 나머지 연령대는 인구추계나 면접조사와 유사한 비율로 나타났다. 이는 장기적으로 60대 이상의 표본이 작아서 발생하는 대표성의 문제가 해결될 가능성이 있음을 보여주는 것이라고 할 수 있다.

2. 표본 분포 형태

이제 응답자의 특성에 따른 인터넷조사와 면접조사의 분포를 비교해 보기로 하자. <표 1-4>는 조사방법에 따른 성별과 연령대의 관계를 알아보기 위한 분포 현황이다.

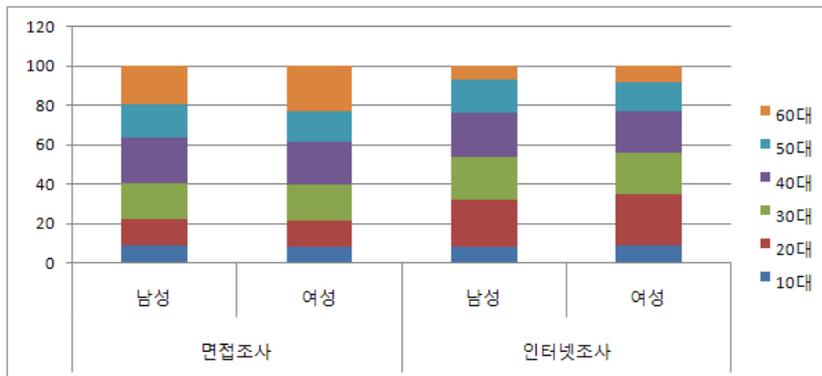


〈표 1-4〉 조사방법에 따른 성별 연령대의 분포 현황

조사 방법	성별	연령대						총합
		10대	20대	30대	40대	50대	60대	
면접 조사	남성	1,582 (9.1)	2,211 (12.71)	3,284 (18.88)	3,968 (22.82)	2,961 (17.03)	3,386 (19.47)	17,392 (100.0)
	여성	1,525 (8.01)	2,524 (13.25)	3,534 (18.56)	4,070 (21.37)	2,988 (15.69)	4,402 (23.12)	19,043 (100.0)
	총합	3,107	4,735	6,818	8,038	5,949	7,788	36,435
	γ	.037**						
인터넷 조사	남성	116 (8.33)	327 (23.47)	304 (21.82)	315 (22.61)	233 (16.73)	98 (7.04)	1,393 (100.0)
	여성	136 (9.01)	392 (25.96)	318 (21.06)	320 (21.19)	219 (14.5)	125 (8.28)	1,510 (100.0)
	총합	252	719	622	635	452	223	2,903
	γ	-.035						

**p<.001

면접조사의 경우 60대 이상의 여성 비율이 23.12%로 같은 연령대의 남성 19.47%에 비해 높게 나타났으며, 성별에 관계없이 40대 비율이 가장 높게 나타났다. 인터넷조사의 경우에는 20-40대의 응답자 비율이 비슷한 수준으로 나타났으며, 60대 이상의 응답자 비율은 남녀 각각 7.04%와 8.28%로 낮게 나타났다. 전반적으로 성별에 따른 차이는 나타나지 않았으며, 면접조사에 비해 인터넷조사에서 20대의 응답 비율이 높고, 60대 이상의 응답 비율이 낮게 나타남을 알 수 있다.



[그림 1-7] 조사방법에 따른 성별 연령대의 분포 현황

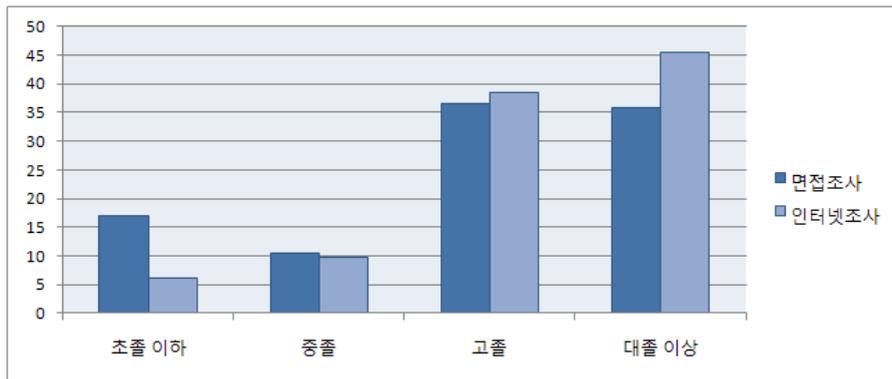
다음은 조사방법에 따른 학력분포 현황이다.

<표 1-5> 조사방법에 따른 학력 분포

학력	면접조사		인터넷조사	
	명	%	명	%
≤초졸	6,219	17.08	180	6.2
중졸	3,851	10.58	284	9.78
고졸	13,273	36.46	1,117	38.48
≥대졸	13,061	35.88	1,322	45.54
χ^2	.229**			

**p<.001

최종 학력이 초졸 이하라고 응답한 면접조사의 응답자는 17.08%로 인터넷조사의 6.2%에 비해 많았으며, 인터넷조사의 응답자 중 45.54%가 대졸 이상이라고 답해 면접조사 35.88%에 비해 높은 비율을 차지하고 있음을 알 수 있다. 즉, 인터넷조사에 참여한 응답자들의 학력이 면접조사에 참여한 응답자들에 비해 전반적으로 높게 나타났다.



[그림 1-8] 조사방법에 따른 학력 분포

조사방법에 따른 가구 소득의 차이를 비교해보면 다음과 같다.

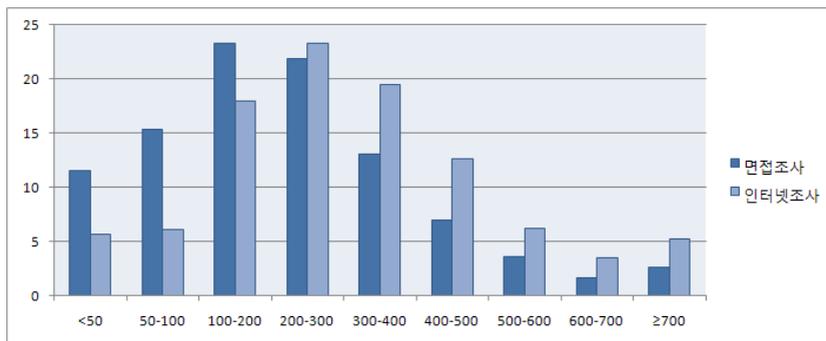


〈표 1-6〉 조사방법에 따른 가구 소득 분포

소득	면접조사		인터넷조사	
	명	%	명	%
<50	1,950	11.59	165	5.68
50-100	2,577	15.31	177	6.1
100-200	3,915	23.27	521	17.95
200-300	3,677	21.85	675	23.25
300-400	2,194	13.04	565	19.46
400-500	1,180	7.01	367	12.64
500-600	611	3.63	179	6.17
600-700	272	1.62	102	3.51
≥700	451	2.68	152	5.24
χ^2	.315**			

**p<.001

조사방법에 따라 가구 소득을 비교한 결과, 가구의 월평균 소득이 200만원 이하라는 응답이 면접조사의 경우 전체의 50.17%로 과반수를 넘었으나, 인터넷조사의 경우에는 29.73%에 그쳤다. 이에 반해 월평균 소득이 400만원 이상이라는 응답은 인터넷조사의 경우 전체의 27.56%, 면접조사의 경우 14.94%로 큰 차이를 보임을 알 수 있다. 즉, 인터넷조사에 참여한 응답자들의 월평균 가구 소득이 면접조사에 참여한 응답자들에 비해 유의미하게 높다는 것을 알 수 있다.



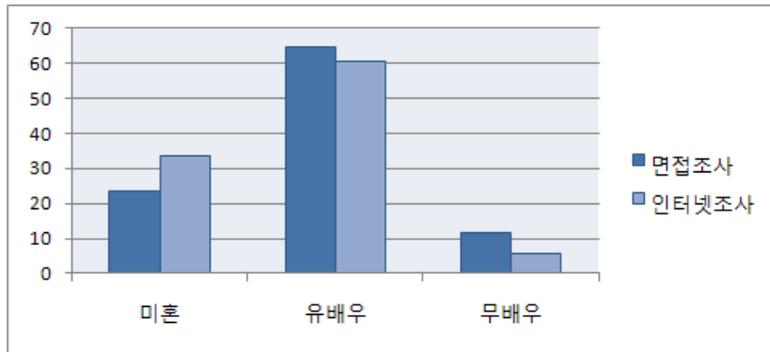
[그림 1-9] 조사방법에 따른 가구 소득 분포

혼인 상태의 경우, 인터넷조사에 참여한 응답자들의 미혼 비율이 33.79%로 면접조사에서의 미혼 비율 23.72%에 비해 높게 나타났으며, 사별이나 이혼 등의 무배우인 경우는 면접조사의 비율이 11.82%로 인터넷조사의 5.51%에 비해 높게 나타났다.

<표 1-7> 조사방법에 따른 혼인상태 분포

혼인상태	면접조사		인터넷조사	
	명	%	명	%
미혼	8,641	23.72	981	33.79
유배우	23,487	64.46	1,762	60.7
무배우	4,307	11.82	160	5.51
γ	-.262**			

**p<.001



[그림 1-10] 조사방법에 따른 혼인상태 분포

조사방법에 따른 직업의 차이를 알아보기 위하여 직업코드의 차이를 살펴보았다.

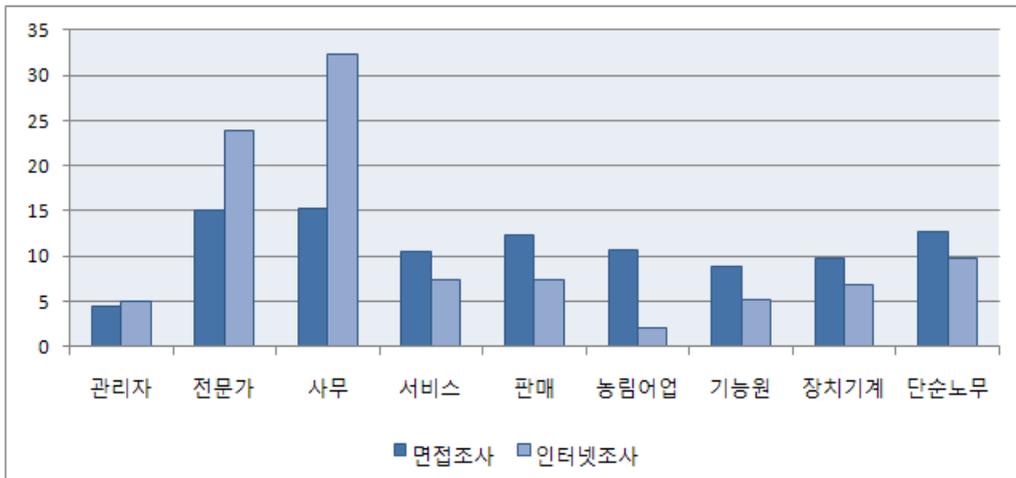
<표 1-8> 조사방법에 따른 직업 분포

직업코드	면접조사		인터넷조사	
	명	%	명	%
관리자	903	4.48	76	5.03
전문가	3,046	15.13	362	23.96
사무	3,085	15.32	489	32.36
서비스	2,132	10.59	112	7.41
판매	2,500	12.42	112	7.41
농림어업	2,166	10.76	30	1.98
기능원	1,779	8.84	79	5.23
장치기계	1,964	9.75	104	6.88
단순노무	2,559	12.71	147	9.73
γ	-.265**			

**p<.001
무직인 경우 제외



조사방법에 따른 직업분포를 살펴본 결과, 인터넷조사에서의 전문가와 사무직의 비율은 각각 23.96%와 32.36%로 인터넷조사 응답자 전체의 과반수를 넘고 있음을 알 수 있다. 이에 비해 면접조사의 경우는 서비스, 판매, 농림어업, 기능원, 장치기계, 단순노무 등의 비율이 인터넷조사에 비해 전반적으로 높게 나타남을 알 수 있다. 즉, 인터넷조사와 면접조사에 참여한 응답자들의 직업에 차이가 있음을 알 수 있으며, 인터넷조사의 경우 전문·사무직의 참여가 두드러지게 많았음을 알 수 있다. 이러한 결과는 두 조사방법에 따른 학력 차이와 무관하지 않으며, 면접조사와 인터넷조사를 구분하지 않은 전체 응답자들을 대상으로 하는 학력에 따른 직업 분포가 <부록 3>에 포함되어 있다.



[그림 1-11] 조사방법에 따른 직업 분포

이상의 비교 결과를 전반적으로 살펴보면, 인터넷조사에 참여한 응답자들의 경우 20대, 여성, 대졸 이상, 400만원 이상의 월평균 가구 소득, 미혼, 전문·사무직인 경우가 많아 면접조사의 응답자 특성과는 차이가 있음을 알 수 있었다. 이와 같은 조사방법별 응답자의 특성 차이는 조사 결과에 대한 편향으로 이어질 수 있으므로, 인터넷조사를 면접조사에 대한 대안조사로 이용하기에는 상당한 어려움이 있을 것으로 예상된다.

제4절 인터넷조사의 타당도와 신뢰도

1. 인터넷조사의 타당도

타당도(validity)란 측정도구가 측정하고자 하는 개념을 정확하게 측정하고 있는지 여부를 판단하기 위한 것으로, 대표적인 타당도로는 개념타당도(content validity), 기준타당도(criterion validity), 구성타당도(construct validity), 안면타당도(face validity) 등이 있다. 개념타당도는 측정도구에 포함된 지표가 측정하려는 속성을 얼마나 잘 반영하고 있는가를 전문가들의 주관적 판단으로 측정하는 방법이며, 기준타당도는 측정결과를 이미 타당성이 경험적으로 입증된 독립된 기준을 적용한 측정결과와 비교하여 나타난 관련성의 정도를 의미하는 것으로 동시타당도와 예측타당도로 나누어 볼 수 있다. 구성타당도는 측정되는 개념을 이론적인 가정들의 틀 속에서 경험적으로 검증하는 방법으로 수렴타당도, 변별타당도 등을 바탕으로 하는 다중속성기법과 요인분석 등을 이용하여 측정할 수 있다. 마지막으로 안면타당도는 검사항목을 일반인들이 보고 얼마나 타당해 보이는가를 평가하는 것으로 개념타당도와 같이 주관적인 판단 기준을 이용하지만 평가자들이 전문가가 아니라는 차이가 있다.

본 연구에서 다루고 있는 사회조사는 국민의 삶의 질과 관련된 사회적 관심사와 주관적 의식에 관한 사항을 조사하여 국민의 삶의 수준과 사회적 변동을 파악하고 이를 사회개발정책수립의 기초자료로 제공하기 위한 목적으로 수행된다. 즉, 연구나 마케팅 등을 위한 조사와는 달리 공익을 위한 목적으로 수행되므로 동시에 두 개의 도구를 이용하여 타당도를 측정하거나 다른 개념의 척도를 동일한 응답자에게 측정하는 등의 설계가 불가능하다. 따라서 사회조사에서의 조사항목들 중 소득 및 소비부문과 삶의 질에 관련된 항목을 중심으로 요인분석을 통한 구성타당도를 측정해보고자 한다.

구성타당도 측정을 위한 요인분석은 여러 문항과 잠재적인 요인 간의 상관관계를 계산하여 상관관계가 높으면 같은 요인으로 묶고, 낮으면 상호 독립이 되는 다른 요인에 묶이도록 하는 분석 방법이다. 요인분석은 개념들 간의 관계를 탐색적 방법으로 파악하는 탐색적 요인분석(exploratory factor analysis)과 이미 선행연구를 통하여 해당 개념을 구성하는 요인들에 대한 합의가 된 후에 이들 개념이 정말 타당한가를 확인하는 확증적 요인분석(confirmatory factor analysis)으로 나누어 볼 수 있다. 삶의 질에 관한 측정 문항의 경우 각 나라마다 분류 기준이 다르며, UN에서 제공하고 있는 기준 또한 서로 상이한 상황이다. 또한 이들 기준은 삶의 질에 대한 각국 전문가들의 판단에 의한 것으로 개념타당도를 확보한 상태이다. 우리나라의 경우 아직 삶의 질에 대한 전문가 기준이 제시되어 있지 않고, 선행 연구를 통한 어느 정도의 합의도 이끌어져 있지 않은 상황이므로,



탐색적 요인분석을 통하여 자료를 기반으로 하는 개념들 간의 관계를 확인하도록 한다. 소득 및 소비 부문의 경우 또한 관련 선행 기준이 마련되어 있지 않으므로, 각 조사방법에 따라 탐색적 요인분석을 개별 수행하고 그 결과를 비교하도록 한다.

조사방법별 탐색적 요인분석 결과를 살펴보면 <표 1-9>와 같다. “스트레스인식”과 “생활시간에 대한 태도”는 스트레스와 관련된 항목으로 응답자의 내적 상태를 나타내고 있다. 이는 국제적인 기준에도 합당한 것으로 면접조사와 인터넷조사에서의 각 요인 내 요인적재량(factor loading)의 순위에만 차이가 있을 뿐 두 항목이 공통의 요인을 설명하는 것은 합당한 것으로 나타났다. “사회 안전에 대한 인식”과 “범죄피해 두려움”, 그리고 “환경오염에 대한 체감”은 응답자가 외부로부터 수동적으로 적용받는 항목들로 요인 내 순위에만 차이가 있을 뿐 두 조사방식에 따른 차이 없이 공통의 요인으로 묶이는데 문제가 없는 것으로 판단된다.

<표 1-9> 조사방법에 따른 삶의 질 만족도의 탐색적 요인분석

조사방법	면접조사				인터넷조사		
	요인1	요인2	요인3	요인4	요인1	요인2	요인3
항목							
거주지역	0.824	0.031	-0.067	0.145	0.716	0.175	0.020
주택	0.815	0.129	-0.114	0.065	0.762	0.075	-0.051
주관적 건강	0.017	0.825	-0.060	0.097	0.547	0.200	-0.007
가족관계	0.398	0.606	-0.071	-0.038	0.694	-0.004	-0.095
스트레스인식	-0.191	-0.157	0.758	0.036	-0.344	-0.116	0.697
생활시간	-0.141	0.553	0.607	0.017	0.142	0.001	0.836
사회안전	0.063	0.129	-0.142	0.783	0.274	0.761	0.020
환경오염	0.291	-0.072	0.263	0.603	0.267	0.419	0.047
범죄피해	-0.127	0.007	0.558	0.523	-0.132	0.786	-0.199
고유값	2.135	1.475	1.102	1.023	2.520	1.275	1.083
누적분산	0.237	0.401	0.524	0.637	0.280	0.422	0.542

“거주 지역 만족도”, “주택 만족도”, “주관적 건강평가” 그리고 “가족관계 만족도”의 경우에는 면접조사의 경우 각각 2개 항목씩 2개의 요인으로 묶였으며, 인터넷조사의 경우에는 4개 항목이 하나의 요인에 묶인 것을 확인할 수 있다. 앞서 밝힌 바와 같이 삶의 질과 관련하여 정확한 기준이 마련되어 있지 않은 상황이므로 둘 중 어느 것이 더 합당한 결과라고 결론내리기는 어려우나, 두 조사방법에 따른 차이가 있음은 알 수 있다.

다음은 소득 및 소비 부문에 대한 탐색적 요인분석 결과이다.

〈표 1-10〉 조사방법에 따른 소득 및 소비 부문의 탐색적 요인분석

조사방법	면접조사	인터넷조사
항목	요인1	요인1
소비생활	0.825	0.819
소득	0.792	0.794
소득분배	0.627	0.688
장래소득	0.488	0.259
고유값	1.938	1.841
누적분산	0.484	0.460

〈표 1-10〉의 결과를 살펴보면, 인터넷조사에서의 장래 소득에 대한 기대의 요인적재량이 .259로 매우 낮게 나타남을 알 수 있다. 보통 요인적재량이 .3 이하이면 공통 속성을 나타낸다고 보기 어려움(Gorsuch, 1983)으로, 인터넷조사에서의 장래 소득에 대한 기대는 소득 및 소비 부문의 만족도를 나타내는 공통적인 특성이라고 보기 어렵다. 이에 반해 면접조사의 경우에는 해당 항목의 요인적재량이 .488로 나타나 해당 요인에 대한 중요한 항목이라고 할 수 있다. 즉, 조사방법에 따른 소득 및 소비 부문의 만족도의 구성타당도에 차이가 있음을 알 수 있다.

삶의 질과 소득 및 소비 부문의 만족도를 대상으로 탐색적 요인분석을 이용한 구성타당도를 비교한 결과, 두 조사방법에 따른 요인 내 요인적재량의 크기에 다소 차이가 있었으며, 요인 구성 항목에서의 차이도 발생함을 확인할 수 있었다. 따라서 조사방법에 따른 구성타당도의 차이를 조정하기 위한 연구가 해당 분야의 전문가와 함께 심층적으로 이루어져야 할 것이다. 그렇지 못한 상황에서 인터넷조사를 면접조사의 대체 조사방법으로 이용하기는 어려움이 있을 것으로 보인다.

2. 인터넷조사의 신뢰도

신뢰도(reliability)란 측정도구가 일관성 있게 특정 개념을 측정하는지 여부를 측정하기 위한 것으로 도구의 안정성을 측정하는 재검사법이나 도구 간의 동등성을 보는 동형검사법, 그리고 도구의 내적 일관성을 측정하는 반분법이나 크론바알파값(Cronbach's α) 등이 이용되고 있다. 타당도에서와 마찬가지로 신뢰도 연구를 위한 별도의 조사표 설계가 어려우며, 이미 조사된 사회조사 결과를 이용하여 신뢰도를 측정해야 하는 한계가 있다. 따라서 본 연구에서는 두 조사방법에서의 소득 및 소비 부문의 만족도와 삶의 질 만족도에 대한 크론바알파값을 비교하도록 한다.



삶의 질 만족도 중 타당도 비교를 통하여 인터넷조사와 면접조사에 차이가 발견된 거주지역, 주택, 주관적 건강, 가족관계 만족도 항목과 소득 및 소비 부문의 경제적 만족도에서 요인적재량이 낮았던 장래소득 항목은 신뢰도 비교 분석에서 제외한다.

다음은 조사방법에 따른 크론바알파값의 비교 결과이다. 인터넷조사와 면접조사의 표본으로부터 얻은 크론바알파값을 비교한 결과, 소득 및 소비 부문의 경제적 만족도와 삶의 질 내 스트레스와 관련된 만족도에는 조사방법 사이에 차이가 없는 것으로 나타났으며, 삶의 질 내 외부 환경 만족도에서는 인터넷조사의 신뢰도가 면접조사에 비해 비교적 높게 나타남을 알 수 있다.

〈표 1-11〉 인터넷조사와 면접조사의 크론바알파값 비교

차원	만족도 항목	인터넷 조사	면접 조사	z
경제적 만족	현재의 전반적인 소비생활에 어느 정도 만족하고 있는가?	0.669	0.658	0.442
	본인의 현재 소득에 만족하고 있는가?			
	현재 우리 사회의 소득분배가 공평하다고 생각하는가?			
스트레스	지난 2주 동안 일상생활에서 스트레스를 어느 정도 느꼈나?	0.378	0.434	-1.510
	평소 바쁘거나 시간이 부족하다고 느끼는 경우가 있는가?			
외부환경 만족	우리 사회가 전반적으로 얼마나 안전하다고 생각하는가?	0.456	0.389	1.837*
	거주지역의 환경오염 정도는 1년 전과 비교하여 어떻게 변했다고 생각하는가?			
	평소 범죄피해에 대한 어느 정도의 두려움을 느끼는가?			

*p<0.05

즉, 인터넷조사를 이용하여 소득 및 소비 부문의 경제적 만족도를 측정하거나 삶의 질과 관련된 만족도들을 측정하는 경우, 인터넷조사를 통해 얻은 자료가 면접조사를 통해 얻은 자료와 동일한 수준의 신뢰도를 제공하거나 더 우월한 것으로 나타났다. Nunnally(1978)는 크론바알파값이 .7 이상이어야 내적 일관성이 있다고 하였으며, 신뢰도 분석을 통하여 얻은 모든 크론바알파값이 .7 이하로 나타나고 있음을 확인할 수 있다. 따라서 크론바알파값을 이용한 신뢰도 비교를 통하여 인터넷조사를 면접조사의 대안 조사로 이용할 수 있는지 여부를 판단하기에는 어려움이 있다.

제5절 인터넷조사의 정확도 관련 요인

1. 인터넷조사의 정확성

인터넷조사와 관련하여 지적되는 여러 문제점들과 적용에의 한계는 많은 연구자들에 의해 다양한 방식의 해결 방안으로 모색되어 왔다. 그러나 인터넷을 이용하여 조사를 진행하는 경우, 인터넷을 이용하지 않는 계층을 포함하지 못함으로써 발생하는 포함오차(coverage error)나 인터넷조사에 자발적으로 참여하는 응답자들로부터 발생하는 자발참여오차(self-selection error) 등의 문제는 여전히 해결하기 어려운 과제로 남아있다. 이러한 문제를 해결하고 인터넷조사의 정확도를 높이기 위한 추정 방법 관련 연구가 크게 두 방향으로 이루어지고 있다. 그 중 하나는 결합추정방법으로 2개 이상의 조사방법을 이용하여 조사를 진행한 후 이들의 결과를 서로 결합하여 이용하는 방법이며, 다른 하나는 성향가중치를 이용한 추정방법이다. 성향가중치를 이용한 추정 방법은 기존의 인구사회학적 특성 변수를 이용한 가중치와 더불어 응답자의 특성을 설명할 수 있는 성향 변수들을 이용한 가중치를 추가적으로 이용하는 것이다. 기존의 인구사회학적 특성 변수만을 이용한 가중치로는 인터넷조사 결과의 정확성을 확보하기 어렵기 때문이다. 성향가중은 인터넷조사에 참여하는 응답자들의 특성을 이미 알려진 모집단의 분포 또는 준거 표본의 분포에 일치시키는 추정방법으로, Rosenbaum 외(1983)가 성향가중모형을 제시한 이래 다양한 분야에 걸쳐 연구가 이루어지고 있다. 그러나 성향가중방법의 실효성과 관련하여 연구자에 따른 낙관적인 시각과 비관적인 시각이 공존하고 있는 상황이다. 2009년 사회조사에서의 성향가중치에 대한 연구는 조성겸 외(2009)에 의해 통계청 표본관리과 용역과제로 수행되었으므로 이 부분에 대해서는 따로 다루지 않기로 한다.

본 장에서는 기존에 알려진 인터넷조사의 정확도 관련 요인의 패턴을 살펴보고, 관련 요인에 따른 소득 및 소비 부문의 만족도와 삶의 질 만족도에 차이가 있는지 알아보도록 한다.

2. 정확도 관련 요인

인터넷조사에서의 정확도 관련 요인에 대해서는 정답이 없으며, 조사 목적, 시기, 대상 등에 따라 달라질 수 있다. 따라서 사회조사를 통하여 조사되는 여러 특성 변수들을 대상으로 정확도 관련 요인 여부를 살펴보아야 한다. 우리는 이미 사회조사에 포함된 인구사회학적 변수와 응답자의 여러 특성 변수들을 대상으로 조사방법에 따른 차이를 비교한 바 있으므로(3절 참고), 본 장에서는 조사 참여의 자발성과 계층의식 그리고 문화와 여가 부문에 대한 내용을 추가적으로 비교하여 정확도 관련 요인을 검토해 보도록 한다.

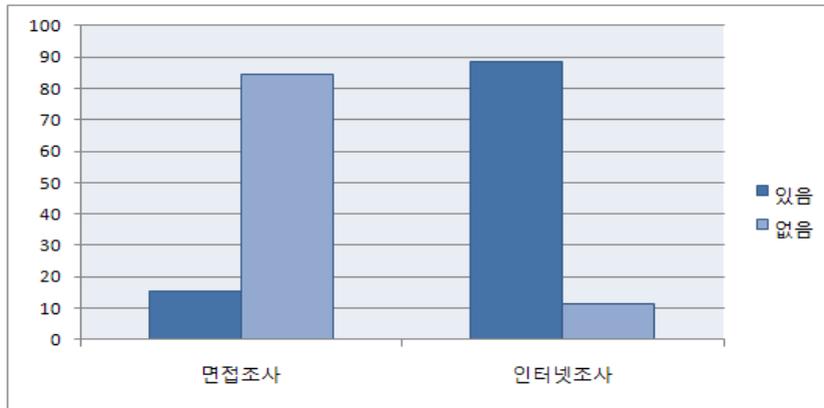


조사방법에 따른 조사 참여의 자발성은 인터넷조사 참여 의향으로 조사하였으며, 그 결과가 <표 1-12>, [그림 1-12]와 같다. 내용을 살펴보면 면접조사의 응답자들 중 84.4%가 인터넷조사에 참여할 의향이 없다고 응답하였으며, 인터넷조사 참여자의 88.67%는 참여할 의향이 있다고 응답하였다. 이는 인터넷조사에 참여한 응답자들의 인터넷조사 참여 의향이 면접조사에 참여한 응답자들에 비해 월등히 높다는 것을 의미한다.

<표 1-12> 조사방법에 따른 인터넷조사 참여 의향의 분포

인터넷조사 참여 의향	명, %			
	면접조사		인터넷조사	
있음	5,685	15.6	2,574	88.67
없음	30,750	84.4	329	11.33
γ	-.954**			

**p<.001



[그림 1-12] 조사방법에 따른 인터넷조사 참여 의향의 분포

인터넷조사에 참여하고자 하는 의향이 인구사회학적 특성 변수에 따라서 차이가 있는지를 확인해 보기 위하여, 면접조사 응답자들을 대상으로 분석을 수행하였다. 면접조사 응답자들을 대상으로 하는 분석에서 인터넷조사 참여 의향에 따른 인구사회학적 특성에 차이가 발견된다면, 이는 조사의 정확도를 떨어뜨리는 요인으로 작용할 수 있을 것이다.

면접조사 응답자들의 인구사회학적 특성에 따른 인터넷조사 참여 의향을 비교한 결과, 여성보다 남성의 참여 의향이 약 4%p 더 높은 것으로 나타났으며, 연령별로는 10~30대의 참여 의향이 각각 24.36%, 26.5%, 24.17%로 40대 이상의 참여 의향에 비해 의미 있게 크게 나타남을 확인할 수 있다. 대졸 이상인 응답자의 참여 의향은 25.71%로 다른

학력에 비해 높게 나타났으며, 미혼인 경우 또한 24.3%로 유배우나 무배우에 비해 높게 나타남을 알 수 있다. 즉, 면접조사 응답자를 대상으로 하는 인터넷조사에의 참여 의향은 10~30대, 남성, 대졸 이상, 미혼인 경우가 높게 나타남을 알 수 있었다.

〈표 1-13〉 면접조사의 인구사회학적 특성별 인터넷조사 참여 의향

					명(%)
특성	항목	참여함	참여하지 않음	전체	γ
성별	남성	3,074 (17.67)	14,318 (82.33)	17,392 (100.0)	.149**
	여성	2,611 (13.71)	16,432 (86.29)	19,043 (100.0)	
연령대	10대	757 (24.36)	2,350 (75.64)	3,107 (100.0)	.437**
	20대	1,255 (26.50)	3,480 (73.50)	4,735 (100.0)	
	30대	1,648 (24.17)	5,170 (75.83)	6,818 (100.0)	
	40대	1,240 (15.43)	6,798 (84.57)	8,038 (100.0)	
	50대	555 (9.33)	5,394 (90.67)	5,949 (100.0)	
	≥60대	230 (2.95)	7,558 (97.05)	7,788 (100.0)	
교육정도	≤초졸	99 (1.59)	6,120 (98.41)	6,219 (100.0)	-.547**
	중졸	255 (6.62)	3,596 (93.38)	3,851 (100.0)	
	고졸	1,946 (14.66)	11,327 (85.34)	13,273 (100.0)	
	≥대졸	3,358 (25.71)	9,703 (74.29)	13,061 (100.0)	
혼인상태	미혼	2,100 (24.30)	6,541 (75.70)	8,641 (100.0)	.392**
	유배우	3,350 (14.26)	20,137 (85.74)	23,487 (100.0)	
	무배우	235 (5.50)	4,072 (94.50)	4,307 (100.0)	

**p<.001

위와 같은 결과로 미루어볼 때, 인터넷조사에 참여할 의향이 있는 사람들을 중심으로 조사할 경우 응답자들의 인구사회학적 편향이 발생하게 되고, 이는 다시 조사 결과의 편



향으로 이어져 통계의 정확성을 떨어뜨리는 요인으로 작용하게 될 것이다. 즉, 자원자를 대상으로 하는 인터넷조사에는 인터넷조사 참여 의향에 따른 편이가 포함되어 있으며, 해당 조사 결과는 자원자들의 특성에 따라 편향되어 있을 가능성이 크다고 할 수 있다.

다음은 인터넷조사 응답자들의 인구사회학적 특성에 따른 인터넷조사 참여 의향을 비교한 결과이다.

〈표 1-14〉 인터넷조사의 인구사회학적 특성별 인터넷 참여 의향

특성	항목	명(%)			γ
		참여함	참여하지 않음	전체	
성별	남성	1,226 (88.01)	167 (11.99)	17,392 (100.0)	-0.063
	여성	1,348 (89.27)	162 (10.73)	19,043 (100.0)	
연령대	10대	214 (84.92)	38 (15.08)	3,107 (100.0)	.228**
	20대	666 (92.63)	53 (7.37)	4,735 (100.0)	
	30대	584 (93.89)	38 (6.11)	6,818 (100.0)	
	40대	570 (89.76)	65 (10.24)	8,038 (100.0)	
	50대	366 (80.97)	86 (19.03)	5,949 (100.0)	
	≥60대	174 (78.03)	49 (21.97)	7,788 (100.0)	
교육정도	≤초졸	139 (77.22)	41 (22.78)	6,219 (100.0)	-.378**
	중졸	227 (79.93)	57 (20.07)	3,851 (100.0)	
	고졸	976 (87.38)	141 (12.62)	13,273 (100.0)	
	≥대졸	1,232 (93.19)	90 (6.81)	13,061 (100.0)	
혼인상태	미혼	881 (89.81)	100 (10.19)	8,641 (100.0)	0.108
	유배우	1,560 (88.54)	202 (11.46)	23,487 (100.0)	
	무배우	133 (83.10)	27 (16.90)	4,307 (100.0)	

**p<.001

분석 결과, 인터넷조사 응답자들의 성별과 혼인상태에 따른 인터넷조사 참여 의향에는 차이가 없었으며, 연령대에서는 2·30대 응답자들의 참여 의향이 각각 92.63%와 93.89%로 다른 연령대에 비해 높게 나타남을 알 수 있다. 특이할만한 점은 면접조사에서의 10대의 인터넷조사 참여 의향이 30대의 참여 의향과 유사하게 나타난 것과는 달리 인터넷조사의 경우에는 84.92%로 2·30대에 비해 비교적 낮게 나타남을 알 수 있다. 또한 40대의 경우에는 인터넷조사에서 참여 의향이 있다고 응답한 경우가 89.76%로 2·30대와 유사한 수준으로 나타났으나, 면접조사의 경우에는 15.43%로 10~30대와는 달리 낮은 비율을 나타냄을 알 수 있다. 이와 같은 현상은 50대와 60대 이상에서도 유사하게 나타나고 있다. 즉, 인터넷조사에 참여했는지 여부에 따라 40대 이상의 인터넷조사 참여 의향에는 큰 차이가 있는 것으로 보인다. 따라서 인터넷조사에 참여한 40대 이상 응답자들의 응답 행태는 참여하지 않은 동일 연령 층과 큰 편의를 가지고 있을 가능성이 크다. 인터넷조사에 참여한 응답자들 중 대졸 이상의 고학력자들의 인터넷조사 참여 의향은 93.19%로 높게 나타났다. 즉, 인터넷조사 응답자를 대상으로 하는 인터넷조사 참여 의향은 2·30대의 대졸 이상인 경우에 높게 나타남을 알 수 있으며, 전반적으로는 면접조사에 비해 큰 차이가 없는 것으로 나타났다.

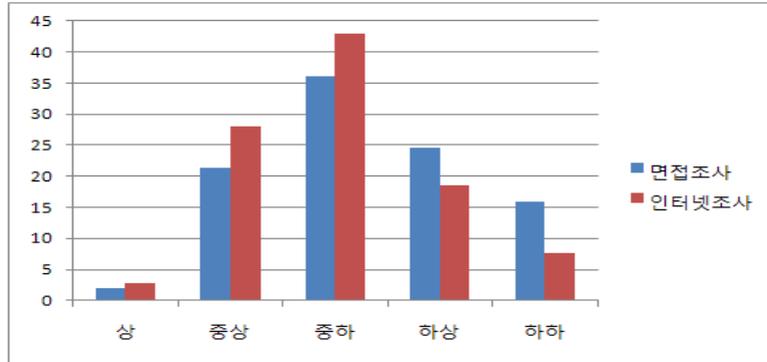
전체적으로 살펴보면, 면접조사에 참여한 응답자들 내에는 인터넷조사 참여 의향에 따른 인구사회학적 편향이 존재하며, 인터넷조사에 참여한 응답자들 내에서는 큰 차이가 발생하지 않는다는 것을 알 수 있다. 이는 인터넷조사에 참여한 응답자들이 그들 고유의 인구사회학적 특성을 가지고 있으며, 해당 특성은 면접조사에 참여한 응답자들과는 유의미한 차이를 가지고 있음을 의미한다.

다음은 조사방법에 따른 계층의식의 분포형태이다.

〈표 1-15〉 조사방법에 따른 계층의식

계층의식	명, %			
	면접조사		인터넷조사	
상	744	2.05	82	2.82
중상	7777	21.34	816	28.11
중하	13146	36.08	1244	42.85
하상	8962	24.6	539	18.57
하하	5806	15.94	222	7.65
γ	-.228**			

**p<.001



[그림 1-13] 조사방법에 따른 계층의식의 분포

조사방법에 따른 계층의식을 살펴본 결과, 사회경제적 지위가 상하 이상(이하 상)이라고 답한 경우는 각각 2.05%와 2.8%로 유사하게 나타났으나, 중상 및 중하라고 답한 경우는 인터넷조사가 면접조사에 비해 각 6.77%p 높게 나타났다. 이에 반해 하상 및 하하라고 답한 경우는 인터넷조사가 면접조사에 비해 각각 6.03%p와 8.29%p 낮게 나타났다. 이와 같은 결과를 통하여 인터넷조사에 응답한 응답자들의 계층의식이 면접조사에 응답한 경우에 비해 전반적으로 더 높다는 것을 알 수 있다. 이러한 차이는 본인의 계층 이동 및 자식세대의 계층이동에 대한 의식에도 차이를 가져올 것이라고 예상할 수 있다.

다음은 조사방법에 따른 계층이동에 대한 의식의 차이를 알아본 표이다.

〈표 1-16〉 조사방법에 따른 계층이동 관련 의식

계층이동	조사방법	명(%)					전체
		매우 높다	비교적 높다	비교적 낮다	매우 낮다	모르겠다	
본인	면접조사	1,057 (2.90)	12,505 (34.32)	12,398 (34.03)	4,065 (11.16)	6,410 (17.59)	36,435 (100.0)
	인터넷조사	134 (4.62)	908 (31.28)	1,262 (43.47)	405 (13.95)	194 (6.68)	2,903 (100.0)
	γ	-.091**					
자식세대	면접조사	1,777 (4.88)	15,656 (42.97)	8,677 (23.82)	2,139 (5.87)	8,186 (22.47)	36,435 (100.0)
	인터넷조사	170 (5.86)	1,196 (41.20)	979 (33.72)	256 (8.82)	302 (10.40)	2,903 (100.0)
	γ	-.085**					

**p<.001

조사방법에 따른 계층이동 의식을 확인한 결과, 본인 및 자식세대의 사회경제적 지

위가 높아질 가능성에 대해 묻는 질문에 인터넷조사를 통한 응답자들의 응답이 좀 더 비관적임을 알 수 있다. 즉, 면접조사의 경우 본인의 사회경제적 지위가 높아질 가능성이 비교적 낮다는 응답이 전체의 34.03%인데 반해 인터넷조사의 경우에는 43.47%로 9.44%p 더 높게 나타났다. 본인의 계층이동에 비해 자식세대의 계층이동에는 두 조사 모두 좀 더 낙관적인 응답을 하였으나, 여전히 비교적 낮다는 비관적인 응답은 인터넷조사가 33.72%로 면접조사 23.82%에 비해 9.9%p 더 높게 나타났다. 또한 모르겠다는 회피성 응답은 두 항목 모두에서 면접조사가 인터넷조사에 비해 10%p 이상 높게 나타남을 알 수 있다. 이와 같은 현상은 조사 방식에 따른 차이에 기인한 것으로 McKane 외(2000)는 Bell Atlantic Inc.의 패 널을 대상으로 하는 전화조사와 온라인조사의 비교를 통하여 인터넷조사 응답자들의 응답의 양이 더 풍부하고, 솔직하며, 문제에 대해 더 깊이 생각한 후 답변하는 경향이 있음을 밝힌 바 있다. 이러한 결과에 비추어볼 때, 응답자가 대답하기 곤란한 회피성 질문이나 민감한 사항에 대한 질문의 경우에 조사원이 있는 면접조사 대신 자기기입식 방식으로 진행되는 인터넷조사를 이용한다면, 항목무응답을 크게 줄이고 질문에 대한 보다 솔직한 답변을 얻을 수 있을 것이다.

다음은 조사방법에 따른 계층이동에 대한 차이 검정 결과이다.

<표 1-17> 조사방법에 따른 계층이동 의식 차이

계층이동*	M±SD		
	면접조사	인터넷조사	t
본인	2.649±.754	2.715±.776	-4.41**
자식세대	2.396±.719	2.508±.760	-7.24**

*모르겠다고 응답한 경우를 제외한 분석한 결과임

**p<.001

본인 및 자식 세대의 계층이동에 대한 생각에 조사방법 간의 차이가 나타났으며, <표 1-16>에서의 결과와 마찬가지로 인터넷조사에서의 응답이 면접조사에 비해 유의미하게 비관적임을 알 수 있다.

이제 조사방법에 따른 문화와 여가 부문에 대한 차이를 알아보자. <표 1-18>은 조사방법에 따른 신문을 보는 횟수의 분포를 나타낸 표이며, <표 1-19>는 TV 시청, 독서 및 레저시설 이용횟수에 대한 분포이다.

<표 1-18>의 결과를 통해 알 수 있듯이 일반신문과 인터넷신문에 관계없이 인터넷조사에 참여한 응답자들의 신문 보는 횟수가 면접조사의 응답자들에 비해 높다는 것을 알 수 있다. 일반신문의 경우 두 조사방법 간 매일 신문을 본다고 답변한 비율의 차이는 14.13%p였으며, 인터넷신문의 경우에는 23.45%p로 나타나, 인터넷신문에서의 조사방



법에 따른 차이가 9.32%p 더 크다는 것을 알 수 있다. 같은 맥락에서 신문을 전혀 읽지 않는 경우는 면접조사가 인터넷조사에 비해 높게 나타남을 알 수 있다. 이 경우 또한 일반신문에서의 두 조사방법 간 차이는 25.88%p, 인터넷신문에서의 차이는 34.58%p로 나타나, 인터넷신문에서의 차이가 8.71%p 더 크다는 것을 알 수 있다. 즉, 인터넷조사 응답자와 면접조사 응답자 간에 신문을 보는 횟수에 의미 있는 차이가 존재한다는 것을 알 수 있으며, 특히 인터넷신문을 보는 횟수에서의 차이가 더 크게 나타남을 알 수 있다.

〈표 1-18〉 조사방법에 따른 신문을 보는 횟수

신문	조사방법	명(%)				
		거의 매일	3-4회/ 1주	1-2회/ 1주	1회/2주	보지 않음
일반신문	면접조사	8,199 (22.52)	3,638 (9.99)	4,407 (12.10)	1,990 (5.47)	18,173 (49.92)
	인터넷 조사	927 (36.65)	352 (13.92)	409 (16.17)	233 (9.21)	608 (24.04)
	γ	-.353**				
인터넷 신문	면접조사	7,913 (21.75)	4,238 (11.65)	3,796 (10.43)	1,612 (4.43)	18,821 (51.73)
	인터넷 조사	1,215 (45.20)	576 (21.43)	311 (11.57)	125 (4.65)	461 (17.15)
	γ	-.513**				

**p<.001

〈표 1-19〉 조사방법에 따른 TV시청, 독서 및 레저시설 이용 경험

항목*	경험여부	명, %				
		면접조사		인터넷조사		γ
TV시청	있다	35,217	96.66	2,731	94.08	.291**
	없다	1,218	3.34	172	5.92	
독서	있다	21,560	59.17	2,301	79.26	-.450**
	없다	14,875	40.83	602	20.74	
레저시설 이용	있다	24,211	66.45	2,193	75.54	-.219**
	없다	12,222	33.55	710	24.46	

*지난 1개월 동안의 TV 시청 여부

지난 1년 동안의 책(교과서 및 참고서 제외) 읽은 경험

지난 1년 동안의 레저시설 이용 경험

**p<.001

조사방법에 따른 TV시청, 독서 및 레저시설 이용 경험에 차이가 있는지 여부를 살펴본 결과, 면접조사에 참여한 응답자들의 독서 및 레저시설 이용 경험이 인터넷조사에 참여한 응답자들에 비해 각각 20.09%p와 9.09%p 더 낮은 것으로 나타났다. 이에 비해 조사방법에 따른 TV시청 경험의 차이는 2.58%p로 비교적 작게 나타남을 확인할 수 있다. 즉, 인터넷조사 응답자들의 독서 경험과 레저시설 이용 경험이 면접조사 응답자들에 비해 높게 나타남을 알 수 있다.

이상의 결과를 종합하여 인터넷조사의 정확도와 관련된 변수를 나타내면 다음과 같다.

- ① 지역, 성, 연령, 학력, 혼인상태 등의 인구사회학적 특성 변수
- ② 인터넷조사 참여 의향(조사 참여의 자발성)
- ③ 계층의식
- ④ 신문(일반 및 인터넷신문) 보는 횟수
- ⑤ 독서 및 레저시설 이용 횟수

이상의 정확도 관련 요인들 중 인구사회학적 특성 변수들을 제외한 인터넷조사 참여 의향, 계층의식, 신문 보는 횟수, 독서 및 레저시설 이용 횟수가 소득 및 소비 부문의 만족도와 삶의 질 만족도 조사 결과에 어떤 영향을 미치게 되는지 살펴보도록 한다. 이와 같은 분석을 통하여 정확도 관련 변수들 중 조사의 정확성을 저해하는 요인이 될 가능성이 있는 변수를 발견할 수 있으며, 분석을 통하여 발견된 정확도 관련 변수는 인터넷조사 결과의 정확도를 높이는 성향가중추정방법 등에 활용될 수 있을 것이다.

3. 정확도 관련 요인과 정확도와의 관계

1) 인터넷조사 참여 의향(조사 참여의 자발성)

조성검 외(2005)는 응답자의 조사 참여에 대한 자발성이 조사의 정확도와 큰 관계가 있음을 보였다. 인터넷조사 표본에 대한 비판적인 시각 중 하나는 인터넷조사에 참여하는 응답자들이 자발적으로 조사에 참여한 자원자 표본이라는 데에 기인한다. 앞서도 언급한 바 있으나, 자발적 참여에 의한 인터넷조사의 응답자들은 보통의 경우에 비해 조사 참여 의향이 높으며 특정 성향을 가진 집단일 가능성이 높다. 따라서 자발성 정도의 차이가 조사의 정확도를 저해할 수 있다는 연구가 이미 많은 연구자들에 의해 이루어졌다.

통계청의 사회조사에서는 인터넷 자원자 표본을 모집하여 인터넷조사를 수행하였다. 따라서 인터넷을 이용한 사회조사에서의 자발성이 조사의 정확도에 어떤 영향을 미치는지 여부를 확인하고, 자원자 표본으로 구성된 해당 자료의 활용 가능성을 확인하는 일은



인터넷조사를 활용한 사회조사의 효용성을 측정하는데 매우 중요한 일이 될 것이다. 앞서 <표 1-12>를 통해 인터넷조사와 면접조사에 참여한 응답자들의 인터넷조사 참여 의향에 큰 차이(73.07%p)가 있다는 것을 발견하였다. 즉, 기본적으로 인터넷조사 응답자들의 자발성이 의미 있게 높다고 할 수 있다.

다음은 면접조사 응답자를 대상으로 자발성이 높은 사람과 낮은 사람 간의 소득 및 소비 부문 만족도와 삶의 질 만족도 차이가 얼마나 발생하는지 알아보기 위한 검정 결과이다.

<표 1-20> 면접조사에서의 인터넷조사 참여 의향에 따른 만족도 차이

항목	인터넷조사 참여 의향		t
	있음	없음	
M±SD			
소득과 소비			
소득만족도	3.28±0.96	3.49±0.92	-12.67**
장래소득기대	2.53±0.96	3.02±0.98	-29.26**
소득분배견해	4.00±0.88	3.91±0.84	6.96**
소비생활만족도	3.21±0.90	3.38±0.86	-12.72**
삶의 질 만족도			
주관적건강평가	2.53±0.83	2.89±0.93	-26.67**
스트레스인식정도	2.42±0.81	2.56±0.82	-11.98**
거주지역만족도	2.71±0.96	2.73±0.92	-1.51
주택만족도	2.81±1.00	2.85±0.98	-1.9
사회안전인식	3.28±0.89	3.36±0.84	-6.91**
범죄피해두려움	2.63±0.96	2.75±0.98	-8.38**
가족관계만족도	2.20±0.94	2.40±0.94	-14.48**
환경오염체감	3.09±0.77	3.13±0.72	-2.95*
생활시간태도	2.33±0.96	2.72±1.10	-27.42**

*p<0.01, **p<.001

면접조사 응답자들을 대상으로 인터넷조사 참여 의향에 따른 소득과 소비 부문 만족도의 차이를 조사한 결과, 참여 의향이 있는 사람들의 소득에 대한 만족, 장래 소득에 대한 기대, 소비 생활 만족도가 더 높게 나타났으며, 소득분배에 대한 견해에 있어서는 좀 더 불공평하다고 느끼는 것으로 나타났다. 삶의 질 만족도에 있어서는 참여 의향이 있는 경우의 주관적 건강평가와 가족관계 만족도가 좀 더 높은 것으로 나타났다. 이에 반해 지난

2주 동안 일상생활에서 느낀 스트레스와 평소 범죄피해에 대한 두려움은 더 큰 것으로 나타났다으며, 평소 바쁘거나 시간이 부족하다고 느끼는 경우도 더 많은 것으로 나타났다.

이와 같은 결과를 통하여 인터넷조사 참여 의향이 있는 사람들을 대상으로 조사를 진행할 경우 조사 결과의 정확도가 저해될 우려가 있음을 확인할 수 있다. 통계청의 인터넷을 이용한 사회조사의 경우 자원자 표본을 모집하여 조사를 진행하였으므로, 해당 조사에서의 자발성 편향은 조사 결과에도 영향을 미쳤을 것으로 판단된다.

2) 계층의식

<표 1-15>를 통하여 조사방법에 따른 계층의식에 차이가 있음을 발견할 수 있었다. 즉, 인터넷조사에 참여한 응답자들의 계층의식이 면접조사에 참여한 응답자들에 비해 의미 있게 높다는 것을 확인하였다. 이와 같은 결과가 자발성에 따른 조사방법 간 차이와 같이 유사하게 나타난다면, 이 또한 조사 결과의 정확도에 악영향을 미치는 요인으로 작용하게 될 것이다. 이에 면접조사에서의 소득 및 소비 부문 만족도와 삶의 질 만족도에 계층의식에 따른 차이를 있는지 여부를 검증하였다.

<표 1-21> 면접조사에서의 계층의식에 따른 만족도 차이

항목	계층의식					M±SD
	≥상하	중상	중하	하상	하하	F
소득과 소비						
소득만족도	2.39±0.95	2.89±0.84	3.36±0.81	3.77±0.83	4.10±0.82	1624.51**
장래소득기대	2.54±0.99	2.61±0.91	2.82±0.95	3.14±0.98	3.39±0.98	484.69**
소득분배견해	3.47±0.98	3.70±0.87	3.90±0.82	4.04±0.80	4.17±0.80	372.24**
소비생활만족도	2.41±0.92	2.85±0.79	3.27±0.75	3.61±0.78	3.93±0.83	2144.51**
삶의 질 만족도						
주관적건강평가	2.32±0.89	2.49±0.84	2.71±0.84	2.99±0.89	3.38±0.99	1066.46**
스트레스인식정도	2.67±0.93	2.64±0.81	2.58±0.78	2.50±0.83	2.38±0.89	98.60**
거주지역만족도	2.22±0.98	2.48±0.91	2.70±0.89	2.85±0.91	2.99±0.95	368.29**
주택만족도	2.09±0.97	2.46±0.90	2.76±0.90	3.02±0.95	3.25±1.02	401.41**
사회안전인식	2.95±1.01	3.19±0.86	3.33±0.82	3.44±0.81	3.52±0.84	207.56**
범죄피해두려움	2.83±1.00	2.74±0.96	2.70±0.94	2.72±0.99	2.81±1.06	13.55**
가족관계만족도	1.79±0.86	2.01±0.85	2.29±0.88	2.53±0.92	2.85±0.99	906.49**
환경오염체감	2.87±0.82	3.05±0.75	3.12±0.71	3.16±0.72	3.20±0.72	65.31**
생활시간태도	2.41±1.11	2.48±1.02	2.56±1.01	2.69±1.09	3.07±1.22	311.35**

**p<.001



면접조사 응답자들을 대상으로 계층의식에 따른 소득과 소비 부문 만족도의 차이를 조사한 결과, 계층의식이 높은 사람들일수록 소득 및 소비 부문에 대한 만족도가 높게 나타났으며, 삶의 질 만족도에 있어서는 주관적 건강평가가 더 좋았으며, 거주 지역 및 주택에 대한 만족도와 가족관계 만족도가 높은 것으로 나타났다. 또한 스트레스 인식 정도가 낮고 평소 바쁘거나 시간이 부족하다고 느끼는 경우가 많이 나타났다.

앞서 살펴본 것과 같이 인터넷조사의 경우 면접조사에 비해 응답자의 계층의식이 높은 경우가 많았으며, 해당 조사에서의 계층 편향은 조사 결과에도 영향을 미쳤을 것으로 판단된다.

3) 신문 보는 횟수

<표 1-18>를 통하여 인터넷조사에 참여한 응답자들의 신문 보는 횟수가 면접조사에 참여한 응답자들에 비해 의미 있게 많다는 것을 확인하였으며, 이와 같은 차이는 인터넷 신문인 경우에 더 뚜렷하게 나타났다. 신문을 보는 횟수는 단순한 생활 습관이 아니라 정치, 사회, 경제에 걸친 관심의 정도를 표현하는 척도가 될 수 있으므로 응답자의 성향을 나타내는데 중요한 특성 변수라고 할 수 있다. 만약 우리 사회가 돌아가는 현상에 관심이 많은 즉, 신문을 자주 보는 사람들의 특성이 특별한 성향을 가지고 있다면 이 또한 조사 결과의 정확도에 악영향을 미치는 요인이 될 가능성이 높다.

<표 1-22>는 면접조사에서의 소득 및 소비 부문 만족도와 삶의 질 만족도의 일반신문을 보는 횟수에 따른 차이를 분석한 결과이다. 면접조사 응답자들을 대상으로 일반신문을 보는 횟수에 따른 소득 및 소비 부문 만족도의 차이를 조사한 결과, 일반신문을 자주 보는 사람일수록 소득 및 소비 생활에 대한 만족도가 전반적으로 높게 나타났으며, 삶의 질 만족도에서는 주관적 건강평가가 높고 더 바쁘다고 느끼는 경우가 많았다.

<표 1-22> 면접조사에서의 신문(일반)을 보는 횟수에 따른 만족도 차이

항목	일반신문 보는 횟수					F
	거의 매일	3~4회/ 1주	1~2회/ 1주	1회/2주	보지 않음	
소득과 소비						
소득만족도	3.23±0.94	3.38±0.91	3.45±0.9	3.50±0.93	3.60±0.91	173.20**
장래소득기대	2.83±0.99	2.79±0.99	2.82±1.00	2.78±1.02	3.08±0.97	102.97**
소득분배견해	3.86±0.88	3.91±0.86	3.95±0.84	4.00±0.84	3.94±0.82	20.79**
소비생활만족도	3.17±0.88	3.27±0.86	3.33±0.85	3.34±0.88	3.45±0.86	158.75**

항목	일반신문 보는 횟수					F
	거의 매일	3~4회/ 1주	1~2회/ 1주	1회/2주	보지 않음	
삶의 질 만족도						
주관적건강평가	2.65±0.83	2.67±0.81	2.69±0.84	2.72±0.88	2.99±0.99	275.46**
스트레스인식정도	2.54±0.81	2.51±0.77	2.49±0.78	2.44±0.81	2.57±0.84	17.48**
거주지역만족도	2.62±0.93	2.74±0.89	2.77±0.89	2.79±0.93	2.76±0.93	35.25**
주택만족도	2.71±0.97	2.86±0.94	2.93±0.95	2.98±0.98	2.90±1.00	36.62**
사회안전인식	3.24±0.86	3.30±0.84	3.33±0.83	3.37±0.86	3.41±0.84	67.43**
범죄피해두려움	2.75±0.94	2.67±0.91	2.67±0.93	2.60±0.95	2.77±1.02	24.23**
가족관계만족도	2.21±0.91	2.30±0.91	2.34±0.93	2.38±0.94	2.45±0.96	98.30**
환경오염체감	3.07±0.75	3.07±0.75	3.11±0.73	3.13±0.72	3.16±0.71	28.06**
생활시간태도	2.49±1.02	2.45±0.95	2.43±0.96	2.43±0.96	2.85±1.15	294.67**

**p<.001

그렇다면 일반신문에 비해 더 큰 차이를 보였던 인터넷신문을 보는 횟수는 조사 결과에 어떤 영향을 미치는지 살펴보도록 하자.

〈표 1-23〉 면접조사에서의 신문(인터넷)을 보는 횟수에 따른 만족도 차이

항목	인터넷신문 보는 횟수					F
	거의 매일	3~4회/ 1주	1~2회/ 1주	1회/2주	보지 않음	
M±SD						
소득과 소비						
소득만족도	3.27±0.95	3.31±0.91	3.37±0.92	3.40±0.93	3.58±0.91	130.95**
장래소득기대	2.46±0.91	2.58±0.94	2.70±0.95	2.72±0.93	3.27±0.92	919.03**
소득분배견해	3.99±0.88	3.95±0.86	3.93±0.85	3.92±0.84	3.89±0.82	21.09**
소비생활만족도	3.20±0.90	3.27±0.87	3.27±0.86	3.32±0.86	3.45±0.85	144.52**
삶의 질 만족도						
주관적건강평가	2.49±0.81	2.54±0.81	2.57±0.81	2.56±0.84	3.12±0.94	1036.49**
스트레스인식정도	2.41±0.82	2.42±0.79	2.45±0.78	2.44±0.81	2.65±0.82	176.83**
거주지역만족도	2.75±0.96	2.79±0.93	2.76±0.91	2.76±0.94	2.69±0.91	14.12**
주택만족도	2.82±0.98	2.89±0.94	2.86±0.98	2.88±0.98	2.84±1.00	1.83



사회안전인식	3.29±0.90	3.35±0.85	3.37±0.83	3.35±0.84	3.37±0.82	12.84**
범죄피해두려움	2.69±0.95	2.61±0.93	2.61±0.92	2.55±0.95	2.82±1.01	87.02**
가족관계만족도	2.21±0.94	2.31±0.92	2.33±0.92	2.34±0.94	2.46±0.94	105.76**
환경오염체감	3.13±0.76	3.15±0.73	3.13±0.73	3.08±0.74	3.11±0.72	2.96
생활시간태도	2.34±0.95	2.40±0.94	2.42±0.93	2.44±0.94	2.91±1.15	591.77**

**p<.001

면접조사 응답자들을 대상으로 인터넷신문을 보는 횟수에 따른 소득 및 소비 부문 만족도의 차이를 조사한 결과, 인터넷신문을 자주 보는 사람일수록 소득 만족도, 소비 생활 만족도 및 장래 소득에 대한 기대가 큰 것으로 나타났으며, 삶의 질 만족도에서는 인터넷신문을 자주 볼수록 주관적 건강평가와 가족 관계 만족도가 의미 있게 높은 것으로 나타났다. 또한 인터넷신문을 많이 보는 사람일수록 더 바쁘다고 느꼈으며, 범죄피해에 대한 두려움을 더 많이 느끼는 것으로 나타났다. 이에 반해 주택 만족도나 환경오염에 대한 체감 정도에는 별다른 차이가 나타나지 않았다.

앞서 살펴본 것과 같이 인터넷조사의 경우 면접조사에 비해 응답자들의 신문 보는 횟수가 더 많았으며, 해당 조사에서의 신문 보는 횟수에 대한 편향은 조사 결과에도 영향을 미쳤을 것으로 판단된다.

4) 독서 및 레저시설 이용 횟수

<표 1-19>를 통하여 조사방법에 따른 여가 시간 활용(TV시청, 독서, 레저시설 이용)에 차이가 있음을 확인하였으며, 이 중 독서 및 레저시설 이용에 따른 소득 및 소비 부문 만족도와 삶의 질 만족도에 차이가 있는지 여부를 면접조사 자료를 이용하여 확인하도록 한다.

<표 1-24> 면접조사에서의 독서 여부에 따른 만족도 차이

항목	지난 1년 동안의 독서 여부		M±SD
	있음	없음	t
소득과 소비			
소득만족도	3.33±0.94	3.63±0.90	-25.03**
장래소득기대	2.69±0.98	3.28±0.91	-47.72**
소득분배견해	3.95±0.86	3.90±0.81	5.58**
소비생활만족도	3.25±0.89	3.49±0.83	-25.95**

항목	지난 1년 동안의 독서 여부		t
	있음	없음	
삶의 질 만족도			
주관적건강평가	2.61±0.84	3.15±0.96	-56.25**
스트레스인식정도	2.48±0.81	2.63±0.83	-16.77**
거주지역만족도	2.73±0.94	2.73±0.90	-0.30
주택만족도	2.82±0.98	2.88±0.99	-3.99**
사회안전인식	3.32±0.87	3.39±0.81	-7.46**
범죄피해두려움	2.63±0.95	2.88±1.01	-24.49**
가족관계만족도	2.28±0.93	2.50±0.94	-21.95**
환경오염체감	3.11±0.75	3.14±0.70	-2.98*
생활시간태도	2.43±0.98	2.98±1.16	-47.25**

*p<.01, **p<.001

면접조사 응답자들을 대상으로 지난 1년 동안의 독서 여부에 따른 소득 및 소비 부문 만족도의 조사한 결과, 독서 경험이 있는 응답자들의 만족도가 전반적으로 더 높은 것으로 나타났으며, 삶의 질 만족도에 대해서는 거주 지역 만족도를 제외한 나머지에서 유의미한 차이가 나타났다. 특히 독서 경험이 있는 경우에 주관적 만족도와 가족관계 만족도가 더 높고 스트레스를 좀 더 많이 느끼며, 보다 바쁘다고 느끼는 것으로 나타났다.

다음은 레저시설 이용에 따른 만족도 차이 검정 결과이다.

<표 1-25> 면접조사에서의 레저시설 이용 경험에 따른 만족도 차이

항목	지난 1년 동안 레저시설 이용 여부		t
	있음	없음	
M±SD			
소득과 소비			
소득만족도	3.34±0.93	3.68±0.90	-26.68**
장래소득기대	2.80±0.98	3.22±0.95	-31.60**
소득분배견해	3.92±0.85	3.94±0.83	-1.92
소비생활만족도	3.24±0.86	3.56±0.85	-33.84**
삶의 질 만족도			
주관적건강평가	2.68±0.86	3.127±0.99	-42.29**
스트레스인식정도	2.55±0.81	2.529±0.85	2.11
거주지역만족도	2.69±0.93	2.793±0.91	-9.89**



주택만족도	2.79±0.97	2.949±1.00	-10.16**
사회안전인식	3.33±0.86	3.395±0.82	-7.43**
범죄피해두려움	2.69±0.96	2.828±1.01	-13.00**
가족관계만족도	2.28±0.92	2.548±0.96	-26.28**
환경오염체감	3.11±0.74	3.143±0.72	-4.05**
생활시간태도	2.53±1.02	2.899±1.18	-29.28**

*p<.01, **p<.001

면접조사 응답자들을 대상으로 지난 1년 동안의 레저시설 이용 여부에 따른 소득 및 소비 부문 만족도의 조사한 결과, 소득분배에 대한 견해를 제외한 나머지 소득 및 소비 관련 요인에서 레저시설 이용 경험이 있는 응답자들의 만족도가 전반적으로 더 높은 것으로 나타났다. 삶의 질 만족도에 대해서는 스트레스 인식 정도를 제외한 나머지에서 유의미한 차이가 나타났다. 특히 레저시설 이용 경험이 있는 경우에 주관적 만족도와 가족관계 만족도가 더 높고 스트레스를 좀 더 많이 느끼며, 보다 바쁘다고 느끼는 것으로 나타났다.

이와 같은 결과를 통해 지난 1년 동안의 독서 경험이 있는 사람들이나 레저시설 이용 경험이 있는 사람들을 대상으로 조사를 진행할 경우 조사 결과의 정확도가 저해될 우려가 있으나, 앞서 살펴본 계층의식이나 신문(일반 및 인터넷) 보는 횟수에 비해 그 차이는 작은 편이었다. 따라서 인터넷조사의 정확도 관련 요인들 중 독서 여부나 레저시설 이용 여부는 조사 결과의 정확도에 미치는 영향이 비교적 작을 것으로 예상된다.

5) 정확도 향상을 위한 노력

앞서 살펴본 것과 같이 조사의 정확도에 영향을 미치는 여러 특성 변수를 살펴본 결과, 이들 특성 변수가 실제 조사에서의 편향을 일으킬 가능성이 있음을 확인할 수 있었다. 이와 같이 응답자의 특별한 성향이 조사의 결과에 영향을 미칠 경우에는 이를 보정하기 위한 가중치 산정 방안에 대하여 고려해보아야 한다.

성향가중치란 인구사회학적 특성 변수를 고려한 전통적인 가중치 외에 관심변수와 관련이 있는 성향 추정 변수를 사용하여 보정하는 방법이다. 그러나 일반적으로 인구사회학적 특성 변수(사회조사에서는 지역, 성, 연령대 고려)는 인구 추계 등 모집단의 분포로 이용 가능한 자료가 있으나, 성향가중치에 이용되는 특성 변수들에는 이와 같은 모집단 분포가 따로 마련되어 있지 않다. 따라서 모집단 분포를 대신할 준거 표본을 이용하는 등의 대체 기법을 이용하여야 한다. 이 때 준거 표본의 정도가 낮거나 모집단과의 편의가 발생한다면 오히려 추정의 정도를 악화시킬 가능성이 있다. 즉, 모집단의 분포를 모르는 상태에서 모집단을 대신할 준거 표본의 선정에 보다 신중할 필요가 있으며, 선택된 준거 표본에 따라 성향보정의 결과도 크게 달라질 수 있음을 감안하여야 한다.

제6절 설문응답시간과 응답성실성의 관계

응답자에 의해 답변 내용이 왜곡되거나 기억의 오류 등 정확하지 않은 정보를 바탕으로 응답이 이루어지는 경우에 발생하는 오차를 응답 오차(response error)라고 한다. 이와 같은 응답 오차가 많이 발생하게 되면 확률에 기반한 표본추출기법으로 표본을 추출하고, 훌륭한 조사원을 이용하여 완벽한 조건 하에서 조사가 진행된다고 하더라도 좋은 결과를 얻을 수 없다. 이와 같은 응답 오차의 중요성 때문에 표본 조사와 관련된 여러 연구 분야들 중 응답 오차에 관한 연구는 여러 연구자들을 통해 지속적으로 이루어지고 있다. 그 중 Krosnick(1991)는 일부 응답자들이 답변을 위해 상당한 수준의 인지적 노력이 필요한 질문에 대략만족적(satisficing) 응답을 하는 경향이 있다고 주장하였으며, 응답과 관련된 정보를 불완전하게 이용하거나 왜곡된 정보를 이용하여 응답하는 것, 또는 응답과 관련된 정보를 전혀 이용하지 않고 임의로 응답하는 행위 등으로 대략만족적 응답을 정의하였다. Krosnick가 제시한 대략만족적 응답 패턴은 보기 선택형 문제에서 첫 번째 보기를 답으로 선택하는 행위, 질문에 대해 긍정적으로 응답하는 경향, 일련의 척도형 질문에서 비차별적으로 똑같이 응답하는 행위, ‘잘 모름’ 응답 행위, 객관식 보기가 주어진 경우 무작위로 답을 선택하는 행위 등이 있다.

인터넷조사의 경우 조사원이 없는 상태에서 스스로 조사표를 작성하는 자기기입식(self-administration) 조사 방식을 따르게 된다. 자기기입식 조사 방식을 이용하는 경우 진실된 응답을 얻을 가능성이 높다는 긍정적인 면과 함께 응답자가 성실하게 조사표를 작성하도록 하기 위한 관리가 어렵다는 부정적인 면도 있다. 특히 조사원이 없는 가운데 진행되는 조사의 경우 응답자의 심신 상태나 조사가 진행되고 있는 시간 등을 통제하거나 필터링 할 수 있는 장치가 없어 불성실한 응답이 발생할 가능성이 높다. 이와 같은 불성실응답은 조사 결과에 심각한 편향을 일으킬 수 있으며, 결국 조사 결과에 대한 신뢰도를 떨어뜨리고 조사의 품질을 크게 손상시키게 된다. 따라서 인터넷을 이용한 조사를 진행할 경우 불성실응답의 발생을 최소화하고, 발생된 불성실응답을 구분해낼 수 있는 장치가 반드시 필요하다.

불성실응답의 발생을 최소화하기 위한 기술적인 장치들은 다양한 형태로 제공될 수 있다. 진행률 표기(progress bar)를 통해 응답자들의 응답을 독려하거나, 응답 간 지나치게 오랜 시간이 경과하는 경우 경고 문구를 팝업창 등을 활용하여 제공하는 방법 등을 적용할 수 있다. 또한 서로 연관된 응답에서 상반된 응답을 하는 경우 불성실응답으로 간주하거나, 가구 단위 조사의 경우 가구주가 한 번 선택되면 다시 선택할 수 없게 하는 등의 방안들도 고려해 볼 수 있을 것이다. 이와 같이 불성실응답을 찾아내거나 발생을 최소화하기 위한 노력들은 소프트웨어의 발달과 함께 빠르게 발전하고 있다.



본 연구에서는 불성실응답의 측정 변수로 설문응답시간을 이용하는 방안에 대하여 알아보고자 한다. 이윤석 외(2008)는 설문응답시간과 응답성실성 간의 관계를 이용하여 온라인 조사에서의 응답 오차에 대한 연구를 진행하였으며, 응답시간에 따른 응답 결과에 차이가 있음을 보였다.

1. 응답시간

응답시간과 응답성실성과의 관계를 밝히기 위하여 응답시간을 기준으로 집단을 구분한 뒤, 구분된 집단 간 응답 결과의 차이를 검정하는 방법으로 분석을 진행하였다. 인터넷을 이용한 설문의 경우 조사원이 없는 상황에서 자기기입식으로 조사가 진행되므로 정해진 응답시간이 있을 수 없다. 따라서 전체 응답자들을 대상으로 어느 정도 수준의 응답 시간을 불성실응답의 기준으로 이용할 것인지를 판단하기 위해서는 다양한 응답 시간을 기준으로 집단을 구분하고, 해당 기준에 의해 분류된 집단 간 조사 결과의 차이를 비교하는 과정을 반복 적용하는 작업이 필요하다.

응답시간을 기준으로 불성실응답 집단을 분류하기에 앞서 인터넷조사에 소요된 응답 시간을 살펴보면 다음과 같다.

<표 1-26> 인터넷조사의 응답시간

(일:시:분:초)				
	평균	표준편차	최소값	최대값
시간	10:31:54	37:23:35	02:07	14:07:06:14

인터넷조사의 시작 시간과 종료 시간을 기준으로 산출한 응답 시간을 살펴본 결과, <표 1-26>과 같은 이해하기 어려운 결과가 도출되었다. 인터넷조사의 응답시간의 최소값은 2분 7초이며, 최대값은 14일 7시간 6분 14초이고, 2900여 명의 인터넷조사 응답자들의 응답시간에 대한 평균과 표준편차는 약 10시간 31분 54초와 37시간 23분 35초로 나타났다. 확인 결과, 사회조사가 시작된 초기에 시스템상의 문제가 발생하여 시스템 접속 후 응답을 완성하지 못하고 있다가 시스템이 안정화된 이후에 설문을 완성한 경우가 있음이 확인되었다. 따라서 전체 인터넷조사 응답자들의 응답시간을 이용하여 불성실응답 여부를 측정하는 것은 의미가 없다. 이에 전체 응답자들 중 응답 시간이 하루(24시간)를 넘긴 경우에는 실제 응답자의 불성실한 응답 때문이 아니라 시스템상의 문제로 인한 불가피한 상황임을 인정하여 응답성실성에 관한 분석에서는 제외하도록 한다.

전체 인터넷조사 응답자 2903명 중 응답시간이 24시간을 초과하는 273명의 응답자(전체 인터넷조사 응답자의 9.4%)와 응답시간이 기록되지 않은 3명의 응답자를 제거한

후 인터넷조사의 응답시간을 살펴본 결과가 <표 1-27>과 같으며, 평균 응답 시간은 약 1시간 5분 35초였으며, 최대값은 23시간 57분 3초로 나타났다.

<표 1-27> 인터넷조사의 응답시간

(시:분:초)

	평균	표준편차	최소값	최대값
시간	01:05:35	03:22:12	02:07	23:57:03

물론 시스템상의 문제로 응답시간이 불가피하게 길어진 경우를 하루 이상인 경우로 제한하는 것은 매우 주관적인 기준이다. 그러나 자기기입식 조사를 수행하게 되는 인터넷조사에서 응답자의 응답 시간에 조사원이 직접적으로 개입하거나 응답하고 있는 당시의 상황을 관찰할 수 없는 상황이므로, 응답 시간에 대한 정확한 기준을 제시하는 것은 한계가 있다. 따라서 응답 시도 후 시스템의 결함으로 응답을 할 수 없는 상황에 처하게 된 응답자는 다음날 다시 시도한다는 가정과 함께 24시간 이내에 응답한 응답자들의 응답시간을 10분위로 나누어 분포를 비교한 후 응답시간의 형태를 파악하여 보다 정확한 기준을 마련하도록 한다.

다음은 24시간 이내에 응답한 응답자들의 응답시간을 10분위로 나누어 각 분위별 응답시간을 살펴본 결과이다.

<표 1-28> 인터넷조사의 10분위별 응답시간

(시:분:초)

	평균	표준편차	최소값	최대값
1분위	5:33	59	2:07	6:50
2분위	7:46	30	6:52	8:38
3분위	9:27	28	8:39	10:12
4분위	11:10	32	10:13	12:04
5분위	13:11	40	12:05	14:14
6분위	15:28	46	14:15	16:48
7분위	18:19	52	16:49	19:50
8분위	22:27	1:36	19:51	25:39
9분위	28:57	10:09	25:41	1:36:24
10분위	8:13:28	30:01	1:36:26	23:57:03

결과를 통하여 알 수 있듯이 9분위까지의 응답시간과 10분위의 응답시간에는 큰 차



이가 있음을 알 수 있다. 이에 10분위 내 응답시간을 대상으로 다시 10분위수를 구하여 비교한 결과, 10분위 내 포함된 응답시간의 5분위수는 3시간 8분 29초이며, 6분위수는 9시간 2분 20초로 각각 나타났다. 즉, 24시간 내 응답한 응답자들을 대상으로 응답시간의 10분위를 구하고, 그 중 10분위 내의 응답자들을 대상으로 다시 응답시간의 10분위를 구한 결과, 10분위 내 5분위수를 기준으로 응답시간이 급격하게 커졌음을 확인할 수 있었다. 이와 같은 결과를 바탕으로 응답성실성을 확인하기 위한 인터넷조사 응답자는 3시간 8분 29초 이내에 응답한 응답자들만을 대상으로 한다.

응답시간과 불성실응답과의 관계를 살펴보기 위하여 응답시간에 따른 집단을 구분한 뒤 각 집단과 해당 집단을 제외한 나머지 집단을 비교하기로 한다. 이와 같은 방법을 이용하면 인터넷조사를 이용한 사회조사 참여자들 가운데 특정 응답시간을 갖는 응답자들의 응답 형태가 나머지 응답자들과 다르게 나타나는지 여부를 탐색할 수 있으며, 이를 통하여 불성실응답 시간의 기준을 마련할 수 있을 것이다.

분석 대상이 되는 인터넷조사 응답자는 2,496명(전체 인터넷응답자의 85.98%)이며, 각 집단 구분은 응답시간을 10등분한 경우와 8등분한 경우, 그리고 5등분한 경우로 나누어 비교하도록 한다.

<표 1-29> 분위별 기준 응답시간

	(시:분:초)		
	10분위	8분위	5분위
기준1	06:44	07:14	08:28
기준2	08:28	09:13	11:42
기준3	10:00	11:16	15:58
기준4	11:42	13:43	22:56
기준5	13:43	16:40	X
기준6	15:58	20:31	
기준7	18:45	28:40	
기준8	22:56	X	
기준9	1:11:17		

분위별 기준 응답시간을 확인해보면 <표 1-29>와 같다. <표 1-29>를 통해 분석 대상인 인터넷조사 응답자들의 80%(10분위수 기준8)가 22분 56초 이내에 응답한 것을 알 수 있으며, 87.5%(8분위수 기준7)가 28분 40초 이내에 응답했음을 알 수 있다. 각 분위별 기준 응답시간을 면밀히 살펴보면, <표 1-30>~<표 1-31>과 같다.

<표 1-30>을 통해 알 수 있듯이 10분위별 응답시간에서 9분위의 평균응답시간은 28분

57초, 마지막 10분위의 평균응답시간은 1시간 44분 23초라는 것을 알 수 있다. 그러나 9분위의 최대값이 1시간 10분 50초로 나타나 9분위 내에도 응답시간이 긴 응답자가 포함되어 있음을 알 수 있다. <표 1-31>을 살펴보면, 8분위별 응답시간의 7분위 평균이 23분 55초였으며, 최대값은 28분 31초로 나타났다. 이에 반해 8분위 평균은 1시간 31분 15초로 나타났다. 두 결과를 종합적으로 살펴보면, 응답시간을 10분위별로 나눌 경우, 9분위에 응답시간이 비교적 긴 응답자가 포함되게 되어 9분위와 10분위의 특성이 모호해질 가능성이 있는 반면, 8분위별 기준을 사용할 경우 응답시간이 비교적 긴 응답자 모두가 8분위에 포함되게 됨으로, 7분위와 8분위에 응답시간에 따른 뚜렷한 차이가 있을 것으로 예상된다.

<표 1-30> 10(5)분위별 응답시간

(시:분:초)

집단	응답자수	평균	표준편차	최소값	최대값
1분위	252	05:30	58	02:07	06:44
2(1)분위	249(501)	07:39	28	06:45	08:28
3분위	248	09:15	26	08:29	10:00
4(2)분위	250(498)	10:51	31	10:01	11:42
5분위	250	12:41	37	11:43	13:43
6(3)분위	251(501)	14:48	38	13:45	15:58
7분위	248	17:22	48	15:59	18:45
8(4)분위	250(498)	20:35	01:15	18:46	22:53
9분위	249	28:57	10:09	22:57	1:10:50
10(5)분위	249(498)	1:44:23	21:37	1:12:19	3:08:29

<표 1-31> 8분위별 응답시간

(시:분:초)

집단	응답자수	평균	표준편차	최소값	최대값
1분위	313	05:47	01:03	02:07	07:14
2분위	311	08:15	35	07:15	09:13
3분위	314	10:14	36	09:14	11:16
4분위	311	12:27	43	11:17	13:43
5분위	313	15:07	51	13:45	16:40
6분위	310	18:26	01:03	17:41	20:29
7분위	312	23:55	02:18	20:31	28:31
8분위	312	1:31:15	33:17	28:41	3:08:29



이제 응답시간을 기준으로 집단을 분류하고, 집단 간 비교를 통하여 유의미한 차이가 있는지 확인하도록 한다. 세 가지 기준을 이용하여 응답시간에 따른 집단을 구성한 후, 각 기준 하에서 응답집단 간의 항목별 차이를 검정하였다. 이를 통하여 특정 응답집단이 다른 집단들과 이질적인 답변을 했는지 여부를 파악할 수 있다. 그러나 이와 같은 세 가지 기준은 임의로 선정한 기준이라는 한계를 가지고 있다. 따라서 보다 객관적인 기준을 마련하기 위한 후속 연구가 필요하다.

2. 집단별 결과 비교

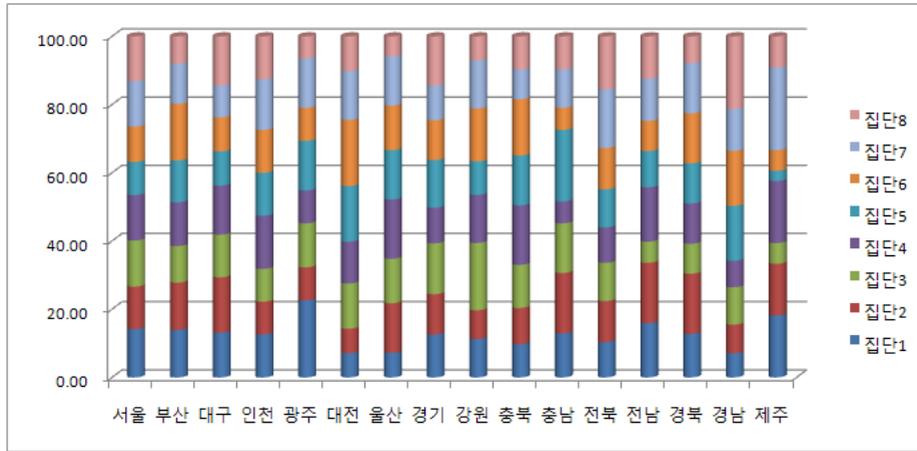
앞서 살펴본 것과 같이 8분위별 응답시간을 기준으로 하는 응답집단이 다른 기준을 이용하는 경우에 비하여 비교적 뚜렷한 차이를 보일 것으로 예상되므로, 본 연구에서는 앞서 소개한 세 가지 응답집단 분류 방법 중 8분위를 이용한 경우의 응답집단 간 차이 검정 결과만을 소개한다.

먼저 인구사회학적 특성 변수에 따른 8분위 응답집단의 분포를 비교해보자.

〈표 1-32〉 지역에 따른 응답집단의 분포

지역	응답집단							
	집단1	집단2	집단3	집단4	집단5	집단6	집단7	집단8
서울	14.16	12.47	13.53	13.32	9.73	10.36	13.32	13.11
부산	13.90	13.90	10.70	12.83	12.30	16.58	11.76	8.02
대구	13.13	16.25	12.50	14.38	10.00	10.00	9.38	14.38
인천	12.59	9.63	9.63	15.56	12.59	12.59	14.81	12.59
광주	22.58	9.68	12.90	9.68	14.52	9.68	14.52	6.45
대전	7.14	7.14	13.27	12.24	16.33	19.39	14.29	10.20
울산	7.25	14.49	13.04	17.39	14.49	13.04	14.49	5.80
경기	12.57	11.85	14.90	10.41	14.00	11.67	10.23	14.36
강원	11.27	8.45	19.72	14.08	9.86	15.49	14.08	7.04
충북	9.71	10.68	12.62	17.48	14.56	16.50	8.74	9.71
충남	12.90	17.74	14.52	6.45	20.97	6.45	11.29	9.68
전북	10.34	12.07	11.21	10.34	11.21	12.07	17.24	15.52
전남	15.93	17.70	6.19	15.93	10.62	8.85	12.39	12.39
경북	12.75	17.65	8.82	11.76	11.76	14.71	14.71	7.84
경남	7.10	8.39	10.97	7.74	16.13	16.13	12.26	21.29
제주	18.18	15.15	6.06	18.18	3.03	6.06	24.24	9.09
γ	0.033							

%



[그림 1-14] 지역에 따른 응답집단의 분포

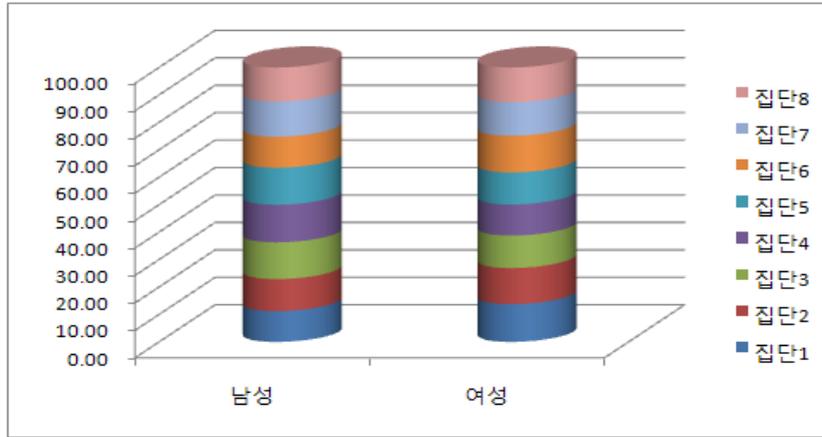
응답집단의 비율이 지역에 따라 차이가 있는지 여부를 확인한 결과, 대전·울산·경남에서 집단1의 비율이 다소 낮게 나타났으며, 광주·제주에서의 비율이 다른 지역에 비해 높게 나타났다. 또한 집단8의 경우는 경남에서의 비율이 다소 높게 나타났다. 그러나 전반적으로 지역과 응답집단 사이에는 큰 연관성이 있는 것으로 보이지는 않는다.

다음은 성별에 따른 응답집단의 분포이다.

<표 1-33> 성별에 따른 응답집단의 분포

성별	응답집단							
	집단1	집단2	집단3	집단4	집단5	집단6	집단7	집단8
남성	11.20	11.71	13.38	13.71	13.46	11.37	12.79	12.37
여성	13.77	13.15	11.85	11.31	11.69	13.38	12.23	12.62
γ	-0.022							

성별에 따른 응답집단에는 통계적으로 의미 있는 차이가 없는 것으로 나타났으며, 이를 그래프로 그려보면 [그림 1-15]와 같다.



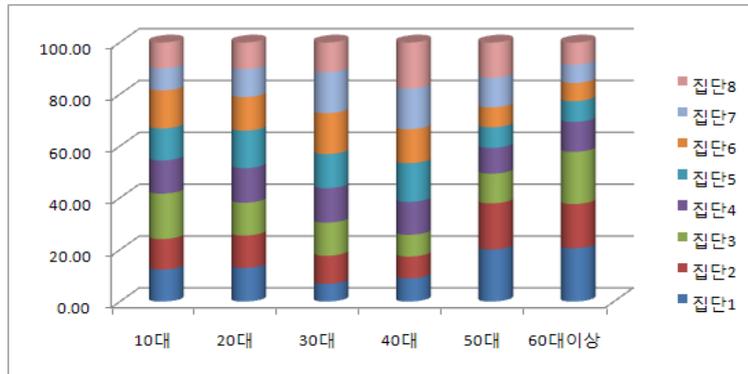
[그림 1-15] 성별에 따른 응답집단의 분포

다음은 연령대에 따른 응답집단의 분포를 확인하기 위한 표이다.

〈표 1-34〉 연령대에 따른 응답집단의 분포

연령대	응답집단							
	집단1	집단2	집단3	집단4	집단5	집단6	집단7	집단8
10대	12.32	11.85	17.54	12.80	12.32	14.69	8.53	9.95
20대	13.00	12.52	12.68	13.32	14.45	13.00	10.75	10.27
30대	6.86	10.86	12.76	13.33	13.14	15.81	15.81	11.43
40대	8.97	8.41	8.41	12.71	14.95	13.08	15.70	17.76
50대	20.10	17.87	11.41	9.93	7.94	7.69	11.41	13.65
≥60대	20.60	17.09	20.10	11.56	8.04	7.04	7.04	8.54
γ	-0.032							

30대의 경우 집단1과 집단8의 비율이 대체로 낮고, 50대와 60대 이상의 경우에 집단1의 비율이 다소 높게 나타남을 알 수 있다. 이는 30대의 성실응답 비율이 다른 연령대에 비해 다소 높고, 50대 이상의 경우 불성실응답 비율이 성실응답 비율에 비해 다소 높게 나타남을 의미한다. 그러나 그 차이가 통계적으로 유의한 수준은 아닌 것으로 보인다.



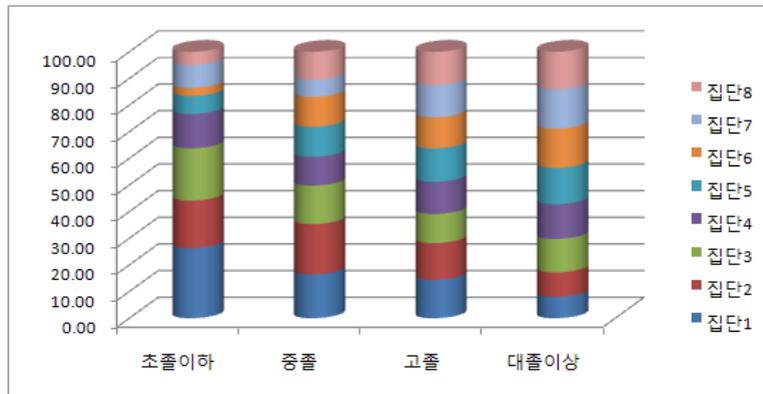
[그림 1-16] 연령대에 따른 응답집단의 분포

다음은 학력에 따른 응답집단의 차이가 있는지 여부를 확인하기 위한 표이다.

<표 1-35> 학력에 따른 응답집단의 분포

학력	응답집단							
	집단1	집단2	집단3	집단4	집단5	집단6	집단7	집단8
≤초졸	26.38	17.79	19.63	12.88	6.75	3.07	8.59	4.91
중졸	16.47	18.88	14.46	10.84	11.24	11.24	6.43	10.44
고졸	14.46	13.74	10.93	12.07	12.59	11.65	12.07	12.49
≥대졸	8.01	9.17	12.56	13.09	13.62	14.69	14.78	14.07
γ	0.203**							

*p<0.05, **p<0.01



[그림 1-17] 학력에 따른 응답집단의 분포



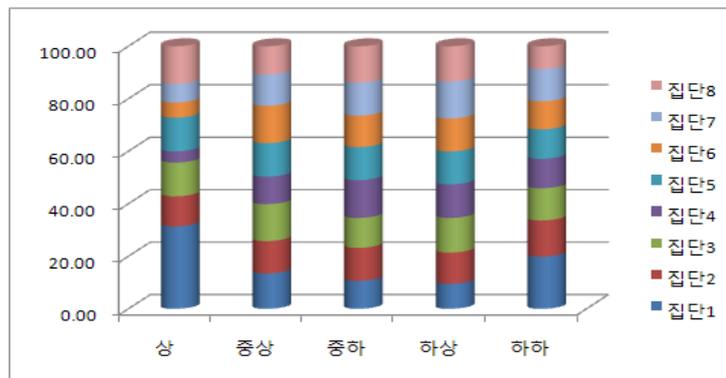
초졸 이하인 경우에 집단1의 비율이 비교적 높게 나타났으며, 대졸 이상의 경우는 가장 낮은 비율로 나타났다. 이와 같은 차이는 통계적으로 유의한 차이이며, 학력이 낮을수록 집단1에 속하는 비율이 높게 나타나는 경향이 있음을 알 수 있다.

즉, 인구사회학적 특성 변수에 따른 응답집단의 형태를 종합적으로 살펴보면, 광주·제주 지역에 거주하는 초졸 이하의 50대 이상 응답자들에게서 집단1(평균응답시간=5분 47초)에 속하는 비율이 비교적 높게 나타났다.

다음은 계층의식과 응답집단과의 관계를 나타낸 표이다.

〈표 1-36〉 계층의식과 응답집단의 분포

계층의식	응답집단								%
	집단1	집단2	집단3	집단4	집단5	집단6	집단7	집단8	
상	31.43	11.43	12.86	4.29	12.86	5.71	7.14	14.29	
중상	13.49	12.34	14.06	10.47	12.77	14.20	11.91	10.76	
중하	10.64	12.61	11.39	14.38	12.61	11.95	12.61	13.82	
하상	9.52	11.90	13.20	12.77	12.55	12.55	14.07	13.42	
하하	19.90	13.78	12.24	11.22	11.22	10.71	12.24	8.67	
γ	0.024								



[그림 1-18] 계층의식과 응답집단의 분포

계층의식이 상인 경우와 하하인 경우에 집단1의 비율이 높게 나타나는 것을 알 수 있으며, 특히 계층의식이 상인 경우에는 집단1에 속하는 비율이 31.43%로 매우 높게 나

타남을 알 수 있다. 그러나 전반적으로는 계층의식에 따른 응답집단 분포에 통계적으로 유의미한 차이는 나타나지 않았다.

이제 응답시간에 따른 계층이동의식에 차이가 있는지 여부를 확인해 보도록 하자. <표 1-37>은 응답집단에 따른 계층이동의식의 차이를 검정한 결과이다.

<표 1-37> 응답집단에 따른 계층이동 의식 차이(t-값)

	계층이동	
	본인	자식세대
집단1 vs. 나머지	0.179	1.152
집단2 vs. 나머지	2.201*	1.832*
집단3 vs. 나머지	-1.013	0.485
집단4 vs. 나머지	0.282	-0.685
집단5 vs. 나머지	-0.39	-0.178
집단6 vs. 나머지	-1.174	-0.269
집단7 vs. 나머지	-0.217	-0.852
집단8 vs. 나머지	0.186	-1.601

*p<0.05, **p<0.01

응답집단에 따른 본인 및 자식세대에 사회경제적 지위가 높아질 가능성의 차이를 알아보기 위하여, 8등분된 응답집단을 집단1과 나머지, 집단2와 나머지 등과 같은 형태로 분류하였다. 이와 같은 과정을 통하여 해당 집단의 응답형태가 다른 집단과 차이를 보이는지 확인할 수 있다(이윤석 등, 2008). 분석 결과, 집단2와 나머지를 비교했을 때 본인 및 자식세대 계층이동에 대한 의식 차이가 발생하는 것으로 나타났다. 반면 집단2를 제외한 나머지 집단들은 해당 집단과 나머지 집단 간의 계층이동에 대한 의식 차이가 나타나지 않았다.

각 집단과 나머지의 응답 평균 차이가 다른 속성에서도 나타나는지 알아보기로 하자. 다음은 전통문화가 본인의 생활에 얼마나 많은 영향을 미치고 있는지를 묻는 항목에서의 응답집단의 차이 검정 결과이다. 전통문화와 관련된 항목은 총 8개의 문항이 리커트 척도 형태로 구성되어 있어, 다른 항목에 비하여 응답자들의 응답 부담이 비교적 크다고 할 수 있다. 이 때 우리는 응답 부담이 큰 경우에 응답에 대한 성실성이 응답 결과에 보다 큰 영향을 미칠 것이라고 예상할 수 있다.



〈표 1-38〉 응답집단에 따른 전통문화와 생활 의식 차이(t-값)

	전통문화와 생활의식							
	생활 의례	인간 관계/예절	민속 놀이	전통 예술	전통 음식	전통 복장	전통 건축	전통 무예
집단1 vs. 나머지	3.810**	6.504**	-1.403	-0.149	4.949**	-1.605	0.453	-2.645**
집단2 vs. 나머지	2.657**	3.990**	1.411	0.275	2.092*	-0.673	0.425	-0.567
집단3 vs. 나머지	1.532	2.299*	0.675	0.210	2.757**	0.808	1.269	-0.612
집단4 vs. 나머지	-0.662	-2.835**	1.231	1.956	-1.922	-1.236	-1.095	1.739
집단5 vs. 나머지	-0.799	-1.934	2.234*	1.978*	0.702	2.106	1.241	2.003*
집단6 vs. 나머지	-3.063**	-5.951**	-0.408	-0.791	-2.895**	-0.530	-1.725	-0.425
집단7 vs. 나머지	-1.637	-2.947**	-1.636	-0.586	-4.606**	0.197	-1.064	0.329
집단8 vs. 나머지	-1.844	-1.998*	-1.756	-2.725**	-1.596	1.209	0.454	0.738

*p<0.05, **p<0.01

전통문화가 생활에 미치는 영향을 묻는 질문에 대하여 집단1과 나머지의 응답을 비교한 결과, 집단1과 나머지의 경우 총 8개 문항 중 4개 문항에서 유의미한 차이가 나타났다. 집단2, 집단5, 그리고 집단6과 각각의 나머지 응답에서는 각각 3개 문항에서 유의미한 차이가 있음을 확인하였다. 이 때 주의하여 보아야 할 것은 “인간관계 및 예절”과 “전통음식”의 경우, 거의 모든 집단에서 차이가 발생하고 있다는 것이다. 즉, 해당 문항에 대한 유의적인 차이는 문항의 특성상 오는 차이로 볼 수 있을 것이다. 그러나 검정 결과를 자세히 살펴보면 집단1~집단3의 경우에는 나머지 응답에 비하여 영향력이 크다고 답한 반면, 집단4~집단8은 나머지 응답에 비하여 영향력이 미치지 않는다고 답한 것이다. 이와 같은 경향은 유의성 여부를 떠나 다른 문항에 대한 검정 결과에도 나타나고 있다. 즉, 집단1과 나머지 응답을 비교한 검정 결과와 집단5와 나머지 응답을 비교한 검정 결과가 서로 상반되게 나타나고 있는 것이다. 두 경우 모두에서 차이를 보인 “전통무예”의 경우 집단1과 나머지 응답 비교에서의 t-값은 -2.645인데 반해 집단5와 나머지 응답 비교에서의 t-값은 2.003으로 나타났다. 이는 집단1의 평균이 나머지에 비해 유의미하게 작고, 집단5의 평균은 나머지에 비해 유의미하게 크다는 것을 의미한다. 여기서 우리는

응답 시간에 따른 응답 태도에 차이가 있음을 확인할 수 있다. 이 때 집단1의 평균 응답 시간은 5분 47초이며, 집단5의 평균 응답 시간은 15분 7초이다.

다음은 사회참여 및 여가생활 만족도, 소득 및 소비생활 만족도, 그리고 삶의 질 만족도에 대한 응답집단의 차이 검정 결과이다.

〈표 1-39〉 응답집단에 따른 만족도 차이(t-값)

	집단1 vs. 나머지	집단2 vs. 나머지	집단3 vs. 나머지	집단4 vs. 나머지	집단5 vs. 나머지	집단6 vs. 나머지	집단7 vs. 나머지	집단8 vs. 나머지
사회참여 및 여가생활 만족도								
주관적만족감	-0.962	0.918	-0.724	0.422	-0.343	-0.870	1.649	-0.085
민원서비스만족도	-0.704	-0.103	0.089	1.374	0.905	-1.709	1.671	-1.570
여가활동만족도	-0.106	-0.478	0.631	1.554	0.692	-2.270*	0.692	-0.722
소득 및 소비생활 만족도								
소득만족도	-1.928	-1.248	-0.118	1.665	-0.715	0.561	0.538	0.978
장래소득기대	0.200	1.590	0.480	1.059	-2.205	-1.439	1.685	-1.140
소득분배견해	-3.006**	-0.780	-2.464*	1.421	0.246	2.046*	2.449*	0.524
소비생활만족도	-1.033	0.010	-1.112	0.262	0.038	0.090	1.001	0.749
삶의 질 만족도								
주관적건강평가	-0.530	2.465*	-0.022	-0.238	-0.561	-0.536	0.429	-0.952
스트레스인식정도	1.463	0.764	-0.728	-0.736	-1.379	-1.389	0.365	1.638
거주지역만족도	-0.223	3.316**	-0.026	0.149	-1.402	-1.528	0.701	-0.774
주택만족도	-1.530	1.853	0.111	-0.294	-0.659	-0.446	1.677	-0.613
사회안전인식	-1.988*	-0.400	0.640	-0.085	-0.606	0.492	1.792	0.157
범죄피해두려움	0.675	2.194*	2.383*	-0.569	0.247	-2.332*	-2.794*	0.213
가족관계만족도	0.185	0.299	0.221	0.345	-0.314	-0.801	0.023	0.023
환경오염체감	-0.202	0.582	0.335	1.342	0.073	0.804	0.017	-2.951**
생활시간태도	2.906**	1.570	-0.895	-0.320	0.045	-1.388	-1.263	-0.654

*p<.05, **p<.01



<표 1-39>는 사회조사에서 조사되고 있는 여러 만족도에 관한 응답에 응답집단별 차이가 있는지 여부를 확인하기 위한 차이 검정 결과이다. 집단1, 집단2, 집단6은 각각 나머지 응답과의 차이 검정에서 3개씩의 문항에 대한 차이를 보였으며, 집단3과 집단7의 경우에는 2개 문항에 대한 유의한 차이를 나타냈다. 각 집단별 차이를 살펴보면, 집단1~집단3이 나머지 응답에 비해 소득분배에 대한 견해에서 더 공평하다고 느끼는 반면, 집단6~집단8은 더 불공평하다고 느끼는 것으로 나타났다. 범죄피해 두려움에 있어서는 집단2와 집단3이 나머지 응답에 비해 두려움을 덜 느끼는 것으로 나타났으며, 집단6과 집단7은 두려움을 더 많이 느끼는 것으로 나타났다. 즉, 응답 시간에 따라 만족도에 유의적인 차이가 나타나고 있음을 확인할 수 있다. 또한 집단4(평균응답시간=12분 27초)와 집단5(평균응답시간=15분7초)에서는 만족도 관련 문항 중 유의미한 차이를 보이는 문항이 발견되지 않았다.

마지막으로 삶의 질 구성영역별 중요도 평가 문항에 대한 응답집단의 차이를 검정해보자. 응답집단에 따른 삶의 질 구성영역별 중요도 평가에 집단에 따른 차이가 있는지 검정한 결과가 <표 1-40>과 같다. 집단1과 나머지 응답과의 차이를 검정한 결과, 소득/소비 등 총 6개의 삶의 질 구성 영역에서 의미 있는 차이가 나타났으며, 집단2와 나머지 응답과의 차이 검정 결과에서도 건강 등 총 6개 영역에서 의미 있는 차이가 발견되었다. 또한 집단7과 나머지 응답과의 차이 검정에서도 노동 등 총 6개 영역에서의 유의미한 차이가 있음을 알 수 있다. 전반적으로 살펴보면 집단1과 집단2와 같은 응답시간이 짧은 경우에 중요도 평가 점수의 평균이 나머지 응답에 비해 낮게 나타났으며, 집단7과 같이 응답시간이 긴 경우에는 중요도 평가 점수의 평균이 비교적 높게 나타나는 것을 알 수 있다. 응답시간이 10분~23분 사이에 들어가는 집단3~집단6까지에서는 별다른 차이가 발견되지 않았다. 집단5의 경우 소득/소비, 건강, 가족에 대한 중요도에서 나머지 응답에 비해 중요도가 높게 나타난 것을 확인할 수 있으며, 이 수치 또한 응답시간이 짧은 경우와는 상반된 결과이다. 즉, <표 1-40>의 결과를 바탕으로 응답시간에 따른 응답 패턴에 차이가 있음을 재확인할 수 있었다.

<표 1-40> 응답집단에 따른 삶의 질 구성영역별 중요도 평가 차이(t-값)

	삶의 질 구성 영역									
	소득/소비	건강	노동	교육	주거/교통	안전	가족	환경	사회통합	여가
집단1 vs. 나머지	-3.705**	-3.913**	-1.823*	-2.472*	-0.119	-2.711**	-3.863**	-1.538	-1.781	-0.541
집단2 vs. 나머지	-0.597	-1.918*	-0.451	-2.238*	-1.917*	-1.838	-2.915**	-3.041**	-2.828**	-1.697

	삶의 질 구성 영역									
	소득/ 소비	건강	노동	교육	주거/ 교통	안전	가족	환경	사회 통합	여가
집단3 vs. 나머지	0.999	0.593	0.208	0.239	0.672	-0.874	0.677	-1.399	-0.990	1.425
집단4 vs. 나머지	1.375	1.957	1.911	1.309	1.689	0.731	1.308	0.076	1.949	1.590
집단5 vs. 나머지	2.959**	2.314*	-0.565	0.418	-0.394	0.891	2.039*	-0.210	-0.451	0.107
집단6 vs. 나머지	-0.129	-0.115	-1.184	0.721	-0.661	0.377	0.942	1.211	0.507	0.853
집단7 vs. 나머지	0.968	1.906	2.819**	2.074*	1.190	2.969**	2.490*	4.918**	2.324*	0.317
집단8 vs. 나머지	-1.191	0.536	-0.781	-0.049	-0.265	0.660	0.539	0.776	1.435	-1.826

*p<.05, **p<.01

이상의 결과를 통하여 응답시간을 기준으로 하는 응답 성실성이 조사 결과에 영향을 줄 수 있음을 확인할 수 있다. 따라서 자기기입방식을 통해 조사에 참여하는 응답자가 성실한 응답을 했는지 여부를 가려내는 것은 인터넷조사에 있어 중요한 과제들 중 하나가 될 것이다. 특히 자기기입식으로 진행되는 인터넷조사에서 테이블형(메트릭스형) 문항 제시 방법을 이용하는 경우(eg. <표 1-38>, <표 1-40>)에는 응답자들의 응답부담이 커지게 되어 불성실 응답이 증가할 가능성이 크다. 이 경우 불성실 응답자들을 구분하지 않고 조사 결과를 그대로 이용하게 되면, 해당 조사 결과는 응답자들의 성실성 여부에 기인한 응답오차로 인해 모집단과의 편향을 가지게 될 가능성이 있다.

3. 응답 수준의 선택 빈도

앞서 우리는 응답의 성실성을 측정하는 방법으로 응답시간을 이용하였다. 조사원의 통제가 없는 상황에서 자기기입식으로 이루어지는 조사의 특성상 응답시간이 지나치게 짧거나 긴 응답자들은 조사에 충실하게 참여하였다고 하기 어렵기 때문이다. 그러나 응답시간의 경우 사람마다 질문을 읽는 속도나 이해하는데 필요한 시간이 다르기 때문에 명확한 기준을 마련하는 일이 쉽지 않다. 즉, 사람에 따라 특정 조사에 소요되는 시간은 매우 유동적이다. 이와 같은 응답시간 기준의 문제점을 보완하기 위하여 응답의 성실성을 측정할 수 있는 다른 방법을 고려해보기로 하자.

Krosnick가 제시한 대략만족적 응답 패턴 중 4번째 패턴은 ‘잘 모름’ 응답 행위이다. 이와 같은 ‘잘 모름’ 응답 행위는 질문에 대한 인지적 판단을 유보하고 본인의 생각을



명확하게 응답하지 않는 행위로, 대표적인 불성실 응답 형태 중 하나이다. 따라서 본 장에서는 리커트 척도(5점 척도)로 조사된 항목을 대상으로 3점(보통)을 선택한 빈도를 확인하고, 이를 응답의 성실성을 구분하는 기준으로 삼을 수 있는지 여부를 확인해보도록 한다.

〈표 1-41〉 각 기준에 따른 응답 수준⁴⁾의 선택 빈도

집단	매우 만족(1)	약간 만족(2)	보통(3)	약간 불만족(4)	매우 불만족(5)
집단1	2.872	5.748	11.246	6.955	8.371
집단2	2.367	5.942	11.132	7.894	8.003
집단3	2.538	6.175	10.570	8.134	8.057
집단4	2.566	6.167	10.270	8.437	8.154
집단5	2.527	6.508	10.307	8.665	7.690
집단6	3.142	6.477	10.216	8.355	7.629
집단7	2.580	6.346	10.375	8.635	7.965
집단8	2.801	6.926	10.208	8.356	7.667
전체	2.674	6.286	10.541	8.178	7.942
F	2.674**	3.339**	2.165*	5.822**	0.997

* $p < .05$, ** $p < .01$

〈표 1-41〉은 사회조사에서 리커트 척도(5점)로 조사된 37개 항목을 대상으로 수준별 평균 응답 횟수를 나타낸 표이다. 전체적으로 보았을 때 보통(3점)을 택한 평균 횟수가 10.541번으로 가장 많았으며, 약간 불만족(4점) 8.178번, 매우 불만족(5점) 7.942번, 그리고 약간 만족(2점) 6.286번 순으로 나타났다. 매우 만족(1점)의 경우에는 평균 응답 횟수가 2.674번으로 다른 응답 수준에 비해 매우 낮게 나타남을 확인할 수 있었다.

평균 3번 응답 횟수에 대한 집단별 차이를 검정한 결과, 집단1과 집단2에서의 평균 3번 응답 횟수가 다른 집단에 비해 유의미하게 크다는 것을 확인할 수 있었다($F=2.165$, $p=0.034$). 또한 전반적으로 불만족에 응답한 평균 응답 횟수가 만족에 응답한 평균 횟수에 비해 크다는 것을 알 수 있으며, 매우 불만족에 응답한 평균 응답 횟수는 집단에 따른 차이 없이 전반적으로 높게 나타남을 알 수 있다. 이를 통하여 응답시간이 짧은 응답자들에게서 대략만족적 응답 패턴과 함께 부정적 응답 패턴이 나타나고 있으며, 인터넷조사 응답자들이 전반적으로 부정적 응답 패턴을 가지고 있음을 알 수 있다.

4) 응답 수준은 문항별 선택 항목에 차이가 있으며, 대표적인 선택 항목인 “매우만족/약간만족/보통/약간불만족/매우불만족”으로 나타남
선택 항목의 수준은 만족(1)에서 불만족(5)의 흐름을 갖도록 각 문항에서 조절함(범죄피해 두려움, 생활시간에 대한 태도)

제7절 결론

1. 요약

인터넷조사는 면접조사나 전화조사와 같은 기존의 전통적인 조사 방식을 대신할 수 있는 대안조사 중 하나로 주목받고 있다. 이와 같은 추세는 IT 관련 기술의 획기적인 발전과 함께 인터넷 이용 인구의 폭발적인 증가와 밀접한 관계를 맺고 있다. 또한 1인 가구가 빠르게 증가하고 맞벌이 부부의 증가에 기인한 주간 재택률의 하락, 그리고 응답자들의 개인보호의식 강화 등은 조사환경을 더욱 열악하게 하고 있다. 열악한 조사환경 속에서 표본조사를 수행하기 위해서는 더 많은 시간과 노력이 필요하며, 여기에 수반하는 비용적인 문제도 간과하기 어렵다. 이와 같은 조사환경의 급속한 변화 속에서 전통적인 방식의 조사방법을 대안할 대체 조사방법에 대한 연구가 필요해졌으며, 그 중 인터넷조사는 큰 주목을 받고 있다.

인터넷조사는 적은 비용으로 많은 사람들을 조사할 수 있다는 큰 장점이 있으며, 조사 진행 시간(응답자들의 응답 가능 시간)이나 면접조사 때 질문하고 답하기 어려운 골고루 문항에 대한 솔직한 답변을 유도하는데 매우 효과적이다. 또한 소프트웨어 등 기술 지원을 통하여 자료 내검을 조사와 동시에 진행할 수 있으며, 자료입력 등에서 발생할 수 있는 처리오차도 크게 줄일 수 있다. 이에 반해 인터넷조사에는 인터넷 이용 가능 인구만을 모집단으로 한다는 포함오차 문제, 확률표본이 아닌 자원자 표본을 이용함으로써 발생할 수 있는 자발성 오차 등 해결하기 어려운 문제점들도 내재되어 있다. 이에 인터넷조사를 연구하는 많은 연구자들은 포함오차 및 자발성에 기인한 자발성 오차를 해결하고 편의가 적은 추정치를 계산하기 위한 추정 방법에 대해서도 지속적인 연구를 계속하고 있다. 그 중 성향가중법을 이용한 추정 관련 연구는 인터넷조사가 적용되는 다양한 상황과 다양한 조사 결과를 바탕으로 진행되고 있으며, 인터넷조사뿐 아니라 기존 조사방법에서의 추정의 정도를 향상시킬 수 있는 방법으로도 그 가능성을 시험받고 있다. 뿐만 아니라 많은 연구자들은 인터넷조사의 품질에 대한 연구를 통하여 인터넷조사의 결과가 기존 조사방법들에 비해 품질면에서 나쁘지 않음을 보였으며, 인터넷조사를 위한 조사방법에 관한 연구도 활발하게 이루어지고 있다. 그러나 그 안에는 인터넷조사에 대한 낙관적인 시각과 비관적인 시각이 여전히 공존하고 있는 상황이다.

본 연구에서는 인터넷조사에 대한 방법론적인 기준이 마련되지 않은 상황에서 대국민을 대상으로 수행된 인터넷을 이용한 사회조사 자료의 품질을 평가함으로써, 국가통계에 인터넷조사를 활용할 수 있는지 여부에 대한 판단을 내리는데 도움을 주고자 한다. 이를 위하여 조사방법에 따른 특성 변수들의 분포를 비교하고, 타당도와 신뢰도를 비교



함으로써 인터넷조사 자료의 품질을 평가하였다. 이어 인터넷조사에서의 정확도 관련 요인을 탐색하였으며, 설문응답 시간과 응답 성실성의 관계에 대해서도 살펴보았다. 다음은 각 분석을 통해 얻은 결과를 요약한 것이다.

- 첫째, 자발적 참여에 의한 인터넷표본은 서울·경기 지역의 20대, 미혼, 고학력, 고소득, 전문·사무직 종사자인 경우가 많았다.
- 둘째, 인터넷조사와 면접조사에서의 만족도 조사 항목을 이용하여 구성타당도와 신뢰도를 비교한 결과, 조사방법에 따른 구성타당도에 차이가 나타났으며, 인터넷조사의 신뢰도는 면접조사와 유사한 수준을 유지하거나 더 높은 것으로 나타났으나, 신뢰도 자체가 높지 않은 것으로 나타났다.
- 셋째, 인구사회학적 특성을 제외한 인터넷조사의 정확도 관련 요인으로는 인터넷조사 참여의향, 계층의식, 신문(일반 및 인터넷) 보는 횟수 등이 있었다.
- 넷째, 인터넷조사의 정확도 관련 요인과 조사의 정확도와의 관계를 확인한 결과, 이들 특성 변수에 따른 조사 결과에 유의미한 차이가 발생함을 확인하였다.
- 다섯째, 응답시간을 전체의 8분위수를 기준으로 분류하고, 각 집단과 인구사회학적 특성과의 관계를 살펴본 결과, 60대 이상, 초졸 이하인 경우에 집단1(평균 응답시간 5분 47초)의 비율이 높았으며, 계층의식의 경우에는 “상하 이상”과 “하하”에서 집단1에 속하는 비율이 높게 나타났다.
- 여섯째, 각 응답집단과 나머지 응답과의 조사 결과 차이를 검정한 결과, 응답시간이 비교적 빠른 집단1~집단3과 응답시간이 비교적 긴 집단6~집단8의 응답 패턴에 차이가 있음을 발견할 수 있었다. 특히, 응답부담이 커지는 테이블형(메트릭스형) 문항에서는 그 차이가 더 크게 나타남을 확인할 수 있었다.
- 일곱째, 응답수준의 선택 빈도와 응답집단을 비교한 결과, 집단1과 집단2의 3번 응답 비율이 가장 높게 나타났으며, 전반적으로 부정적인 응답패턴이 나타남을 알 수 있다. 후자는 응답 시간이 조사 결과에 미치는 영향이 아닌, 인터넷을 이용한 조사의 특성이 미치는 영향인 것으로 보인다.

결과를 종합해보면, 자발적 참여자를 대상으로 하는 인터넷조사 결과를 확률표집방법에 의한 면접조사와 비교한 결과, 인구사회학적 특성과 더불어 응답자의 성향에 따른 특성변수에서도 조사방법에 따른 차이가 발생한다는 것을 알 수 있었다. 또한 인터넷조사에 참여한 응답자들의 응답이 면접조사에 비해 전반적으로 부정적인 것으로 나타났으며, 응답 시간에 따른 응답패턴이 나타나 조사원의 개입이 어려운 자기기입식 설문 방식에 대한 대책이 필요한 것으로 보인다. 인터넷조사 결과의 구성타당도와 신뢰도는 면접조사에 비해 크게 떨어지지 않는 것으로 보이나 신뢰도는 전반적으로 낮게 나타났다.

2. 연구의 한계점 및 향후 과제

조사 비용의 증가, 사생활보호에 대한 의식 강화, 1인 가구 및 맞벌이 부부 증가에 따른 접촉의 어려움 등 조사 환경의 변화는 기존의 표본조사방법을 대안할 수 있는 대안 조사에 대한 관심으로 이어지고 있다. 특히 IT 기술의 발전과 인터넷 이용률의 폭발적인 증가에 기인한 인터넷조사는 여러 대안 조사방법들 중 가장 주목받고 있는 조사방법들 중 하나이다. 그러나 자원자 표본을 이용한 인터넷조사의 경우 통계학적인 한계를 가지고 있으며, 그 적용면에 있어서도 회의적인 시각이 아직까지 팽배한 상황이므로, 인터넷 조사에 관한 연구는 지속적으로 이루어져야 할 필요가 있다.

본 연구는 통계청 사회복지통계과에서 수행한 사회조사를 이용하여 기존의 면접조사와 인터넷조사의 품질을 비교하기 위한 목적으로 수행되었다. 이에 일반 연구를 위한 조사와는 달리 고정된 항목만을 이용할 수 있었으며, 항목 수준의 척도나 배치 등도 조정이 불가하였다. 즉, 이미 조사되어 수집된 자료를 이용하여 분석을 수행하였으므로 실제 분석 가능한 비교 방법은 극히 제한되어 있었다. 또한 자원자 표본을 모집하여 인터넷 조사를 수행함으로써 조사 결과의 차이가 인터넷을 이용한 조사에서 오는 차이인지 자원자 표본의 특성에 의한 차이인지 명확하지 않다. 따라서 인터넷조사의 특성 및 국가통계로써의 활용 가능성을 보다 면밀히 파악하기 위해서는 확률표본추출방법을 통하여 추출된 표본가구 조사 시, 조사방법으로써 면접과 인터넷 중 하나를 선택하도록 하여 조사방법에 따른 차이도 검정해보아야 할 것이다. 이와 같은 조사가 병행된다면 조사방법으로써의 인터넷조사의 효과와 자원자 표본을 통한 인터넷조사의 효과 비교가 가능함으로, 인터넷조사의 활용 가능성을 재고하는데 큰 도움이 될 것이다. 뿐만 아니라 표본조사 결과의 추정 과정을 연구하는데도 크게 기여할 수 있을 것이다.

또한 인터넷조사를 이용하여 표본조사를 수행하는 경우의 가장 큰 장점은 자료 수집 과정에서 조사원들에 의해 발생하는 오차나 자료 입력 오차 등의 비표본오차를 줄일 수 있다는 것이다. 따라서 인터넷조사를 수행하기 위한 시스템을 마련할 때, 보다 정밀하고 유기적으로 구성함으로써 자료의 품질을 크게 상승시킬 수 있다. 예를 들어 한 가구 내용답에서 가구주라는 응답은 전체 가구원 중 한 명만 가능하도록 만든다거나, 해외여행 경험이 없는 경우에는 여행목적 항목에 답할 수 없도록 하는 등 질문 특성을 이용한 시스템적 지원이 가능할 것이다. 또한 인터넷조사의 특성을 살린 질문지 제공이나 화면 구성 등을 개선함으로써 응답자의 편의를 최대한으로 하는 조사를 구성하게 되면, 조사 자료에 대한 품질도 크게 상승할 것으로 예상된다. 따라서 인터넷조사에 대한 기술적인 지원 방법도 함께 연구되어야 할 과제 중 하나이다.



인터넷조사를 기존 표본조사방법들(면접조사, 전화조사, 우편조사 등)의 대안조사로 이용하기 위하여 연구해야 할 관련 과제는 매우 많다. 이 중 인터넷조사의 가장 큰 문제로 지적되고 있는 것이 모집단에 대한 대표성 문제이다. 본문에서도 여러 번 지적하였으나 자원자 표본을 이용한 인터넷조사는 목표모집단에 대한 포함오차를 가지고 있으며, 해당 오차로 인하여 인터넷조사 결과는 큰 편향을 가지게 된다. 많은 연구자들은 이러한 문제를 해결하기 위하여 인터넷조사에 참여한 자원자 표본의 특성을 보정한 성향가중방법을 연구하고 있다. 그러나 추정치의 보정 과정에 이용되는 벤치마킹 통계의 모집단 분포가 알려져 있지 않은 경우가 많으며, 준거 표본 분포를 이용할 경우 해당 표본에 포함된 포함오차가 여전히 남아있다는 문제가 발생하게 된다. 즉, 인터넷조사에 자발적으로 참여한 응답자들의 특성과 관련된 변수의 모집단 분포가 알려져 있거나, 보다 정확한 정보를 포함했을 것으로 여겨지는 준거 표본을 이용할 수 있는 경우가 아니라면, 자원자 표본을 이용한 인터넷조사의 결과는 목표모집단의 참값에서 크게 벗어날 가능성이 있다. 따라서 본 연구에서 다루어진 인터넷조사의 타당도와 신뢰도가 높고, 조사 결과가 가구방문조사와 유사한 경우가 많다고 하더라도, 이는 본 조사에만 해당되는 사항일 뿐 인터넷조사의 일반적인 특성과는 거리가 있다. 이와 같은 결과를 종합해볼 때, 국가 통계로써 자원자 표본을 활용한 인터넷조사를 시행하는 것은 아직은 시기상조인 것으로 여겨지며, 대안조사로써 인터넷조사를 이용해야 할 경우에는 조사방법으로써의 인터넷조사를 고려해보아야 할 것이다.

참고문헌

- 강남준, 백영민 (2004), 대안적 여론조사의 표본편과 문제점과 가중치를 사용한 보정방법: 성향집수 가중과 반복비례 가중을 중심으로, 언론정보연구, 41, pp.43-78.
- 문시연 (2005), 데이터 수집 과정의 품질 향상을 위한 방안, 한국마케팅여론조사협회, 정기세미나 논문집, pp.17-27.
- 이윤석, 이지영, 이경택 (2008), 온라인 조사의 응답오차에 대한 연구: 설문 응답 시간과 응답 성실성의 관계, 조사연구, 9, 2, pp.51-83.
- 이지영, 김진교 (2006), 온라인 패널 조사와 오프라인 대인면접 조사 간 데이터 품질 비교: 신뢰도와 타당도를 중심으로, 마케팅연구, 21, 4, pp.209-231.
- 이홍철, 김원용 (2003), 온라인 조사를 통한 제 16대 대선 예측 및 분석, 한국조사연구학회 춘계학술발표대회 발표논문집, pp.197-215.
- 조성겸, 강남준 (2003), 대안적 사회여론조사방법: 모바일 조사방법의 가능성 검토, 조사연구, 4, 1, pp.1-29.
- 조성겸, 주영수, 조은희 (2005), 인터넷 여론조사의 정확도 관련요인, 조사연구 6, 2, pp.51-74.
- 조성겸, 허명희 (2009), 인터넷 자원자 표본에 대한 추정방법 연구, 통계청 정책연구용역 최종보고서.
- 홍내리, 허명희 (2001), 제 16대 국회의원 선거의 예측조사에 대한 사후적 검증, 조사연구, 2, 1, pp.1-36.
- Battagolia, M., Izrael, D., Hoaglin, D., and Frankel, R. (2004), Tips and Tricks for Raking Survey Data(a,k,a, Sampling Balancing), Unpublished Mimeograph, Cambridge: MA, Abt Associates.
- Braunsberger, Karin, Hans Wybenga and Roger Gates (2007), A Comparison of Reliability between Telephone and Web-Based Surveys, Journal of Business Research, 60, 7, pp.758-764.
- Couper, M. P. (2000), Web Surveys: A Review of Issues and Approaches, Public Opinion Quarterly, 64, 4, pp.464-494.
- D'Agostino, R. and Rubin, D. (2000), Estimating and using propensity scores with partially missing data, Journal of the American Statistical Assosiation, 95, 451, pp.749-759.
- Hoppe, M. and R. Lamp (2001), The Quality of Online Panels: A Methodological Test, Inference and Exhibition Net Effects 4, ESOMAR, Barcelona, Spain, pp.243-262.
- Kiesler, S., J. Sieglar, and T. McGuire (1984), Social Psychological Aspects of Computermediated Communication, American Psychologist, 39, pp.1123-1234.
- Krosnick, J. (1991), Response Strategies for Coping with the Cognitive Demands of Attitude measures in surveys, Applied Cognitive Psychology, 5, pp.213-236.
- McKane and Heisler (2000), Traditional vs. Web: Business-to-business Panel Research in the internet Age: The Bell Atlantic Case, Proceeding, 53rd ESOMAR Congress and Trade Exhibition, ESOMAR, Impact of Networking, Vienna, Austria, pp.323-335.
- Nunnally, J. C. (1978), Psychometric Theory (2nd ed.). McGraw-Hill, New York.
- Rosenbaum, p. and D. Rubin (1983), The Central Role of the Propensity Score in Observational Studies for Causal Effects, Biometrika, 70, 1, pp.41-45.



- Rubin, R. (1997), Estimating causal effects from large data sets using propensity scores, *Annals of Internal Medicine*, 127, pp.757-763.
- Sheehan, K. B. (2002), Online Research Methodology: Reflections and Speculations, *Journal of Interactive Advertising*, 3, 1, <http://jiad.org/vol3/no1/sheehan/index.htm>.
- Schillewart, N., G. Langerak, and T. Duhamel (1998), Non-Probability Sampling for WWW Surveys: A Comparison of Methods, *Journal of the Market Research Society*, 40, 4, pp.307-317.
- Taylor H. (2000), Does internet research work?, *International Journal of Market Research*, 42, 1, pp.51-63.
- Taylor H. Bremer, J., Overmeyer, C., Siegel, J. and Terhanian, G. (2001), The record of internet-based opinion polls in predicting the results of 72 races in the November 2000 U.S. elections, *International Journal of Market Research*, 43, 2, pp.145-153.
- Traugott, M. (2001) Assessing poll performance in the 2000 campaign, *Public Opinion Quarterly*, 65, pp.389-419.
- Tse, A. (1998), Comparing the Response Rate, Response Speed and Response Quality of two Methods of Sending Questionnaires: E-mail vs. Mail, *Journal of the Market Research Society*, 40, 4, pp.353-363.

<부록 1> 조사방법 비교

	면접	인터넷
조사항목	86개	57개
조사기간	15일간	15일간
표본가구수	17,000가구	1,228가구
표본조사구	1,426개	-
조사구당 가구	12가구	-
1인당 조사량	48가구 (4개 조사구)	-
조사원 단가	42,660원	-
조사관리 주체	본청 및 지방청	본청
답례품 지급*	· 1인가구 : 10,000원 · 2~3인 : 15,000원 · 4인이상 : 20,000원	· 1인가구 : 5,000원 · 2인가구 : 10,000원 · 3인이상 : 15,000원
홍보 이벤트	-	· 1등(1명) : 1,000,000원 · 2등(1명) : 300,000원 · 3등(2명) : 100,000원 · 참가상(100명) : 20,000원
매체활용	-	조사도우미(SMS) 이용

* 15세 이상 응답 가구원수 기준

**출처: 2009 사회조사 인터넷조사방법 효과



<부록 2> 인터넷조사 자원자 모집 결과

○ 모집기간 : 5.12~6.10 (30일간)

	전체	1주차	2주차	3주차	4주차	5주차
총응모자	7,619	3,498	1,372	718	904	1,127
참여경로						
신문이나 TV의 뉴스 보도	654	335	85	46	52	136
통계청(지방청포함)사이트 홍보	3,225	1,258	626	383	485	473
경품 사이트 홍보	2,818	1,471	523	209	243	372
기타(주변인의 권유 등)	922	434	138	80	124	146

* 출처: 2009 사회조사 인터넷조사방법 효과

○ 응모자 특성별 현황

<지역별>

전국	서울	부산	대구	인천	광주	대전	울산	경기
7,619	1,873	664	467	404	208	393	188	1,591
	강원	충북	충남	전북	전남	경북	경남	제주
	222	210	189	262	176	305	393	74

* 출처: 2009 사회조사 인터넷조사방법 효과

<취업여부, 교육정도, 주거형태별>

취업여부		교육정도		주거형태		
취업	비취업	고졸이하	대졸이상	단독	아파트	기타
4,065	3,476	2,223	5,318	1,563	4,070	1,986
(53.4)	(45.6)	(29.2)	(69.8)	(20.5)	(53.4)	(26.1)

* 출처: 2009 사회조사 인터넷조사방법 효과

<성별>

가구원*			응모자		
전체	성별		계	성별	
	남	여		남	여
21,277 (100.0)	10,225 (48.1)	11,052 (51.9)	7,541** (100.0)	3,337 (44.3)	4,204 (55.7)

* 만15세 이상 가구원 현황

** 응모는 하되, 특성정보 입력 않은 경우 78건 발생

*** 출처: 2009 사회조사 인터넷조사방법 효과

<연령대별>

	연령대별						계
	10대	20대	30대	40대	50대	>60대	
가구원*	1,189 (5.6)	5,411 (25.4)	5,505 (25.9)	3,331 (15.7)	3,895 (18.3)	1,946 (9.1)	21,277
응모자	166 (2.2)	2,766 (36.7)	2,869 (38.0)	1,225 (16.2)	368 (4.9)	147 (1.9)	7,541** (100.0)

* 만15세 이상 가구원 현황

** 응모는 하되, 특성정보 입력 않은 경우 78건 발생

*** 출처: 2009 사회조사 인터넷조사방법 효과



<부록 3> 학력과 직업 분류

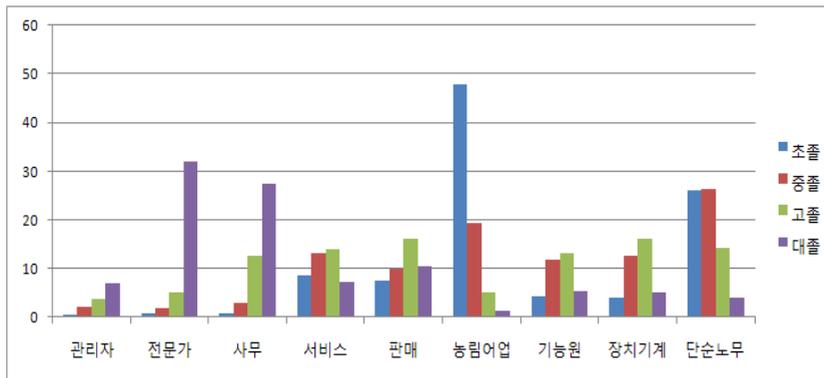
<학력에 따른 직업 분포>

		학력							
		초졸이하		중졸		고졸		대졸이상	
직업*	관리자	16	0.60	42	2.02	284	3.71	636	6.90
	전문가	20	0.75	38	1.83	387	5.05	2959	32.10
	사무	18	0.67	59	2.84	974	12.71	2519	27.33
	서비스	228	8.52	274	13.20	1068	13.94	673	7.30
	판매	199	7.44	207	9.98	1235	16.11	971	10.53
	농림어업	1282	47.91	400	19.28	391	5.10	123	1.33
	기능원	113	4.22	244	11.76	1006	13.13	495	5.37
	장치기계	104	3.89	264	12.72	1234	16.10	466	5.06
	단순노무	696	26.01	547	26.36	1085	14.16	376	4.08

* 무직자 제외

** 면접조사와 인터넷조사에 참여한 모든 응답자 대상

<학력에 따른 직업 분포>



* 무직자 제외

** 면접조사와 인터넷조사에 참여한 모든 응답자 대상