연구논문

혼합식 조사와 웹패널의 (옅은) 빛과 (짙은) 그늘*

Prospects and Problems of the Recent Developments in Social Survey in Korea: Mixed−Mode Survey and Volunteer Web Panel

한신갑a)
Shin−Kap Han

본 논문은 어려워지는 조사환경의 변화에 대한 대응책으로 최근 널리 쓰이고 있는 두 가지 조사방법−혼합식 조사와 (자원자) 웹패널−의 이론적, 방법론적 토대를 검토한다. 우선 한국 사회조사의 생태계라는 현실의 틀 안에서 이들의 자리를 매기고, 각각의 구성과 절차, 한계, 그리고 그 문제점들을 보정하기 위해 개발된 방법들을 소개한다. 특히 '모집단을 얼마나 대표할 수 있는지'라는 가장 기본적인 질문을 기준으로 이들을 평가하고, 이 기준이 제대로 지켜지지 않았을 때 생길 수 있는 편향을 모의실험을 통해 살펴본다. 마지막으로 이 방법들을 도입, 사용하는 데 있어서 취해야 할 구체적이고 실질적인 방안들을 찾아본다. 이 논의를 통해, 사회를 관찰하는 도구로서 기초적인 연구자료와 각종 정책의 수립과 평가에 필요한 근거자료를 제공한다는 사회조사의 원론적 역할을 환기시키고, 자료의 질이 담보되지 않으면 현실적 제약을 이유로 비용만을 낮추는 것은 아무 의미가 없다는 점을 다시 확인한다.

주제어: 자료의 질, 사회조사의 생태계, 혼합식 조사, 웹패널, 보정

* 이 논문은 2012년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구이다(NRF−2010−413−B00031). 2011년 후기사회학 대회 특별기획세션("사회조사의 양과 질")과 2012년 한국사회과학자료원 워크숍("사회조사의 질 개선방안")에서 "폭탄제조: 무응답(편향)을 줄이기 위한 한 방법으로서의 표집틀 섞기와 그 효과"라는 제목으로 초고를 발표했었다. 문제의식을 공감하고 함께 특별기획세션을 조직한 구혜란 박사와 김석호, 박원호, 송유진, 이윤석, 정재기 교수, 소장 자료에 대한 메타데이터를 제공해 준 한국사회과학자료원(KOSSDA)의 김대훈 자료개발실장과 한준 원장, 그리고 이런 작업을 격려해 주신 석현호 선생님께 감사드린다. 구혜란 박사와 이윤석 교수가 원고에 대한 조언을 해 주었고, 김명자, 김영진, 김이선, 이병규, 이상직, 정인관이 작업과정과 원고정리를 도왔다. 세 분 심사위원의 논평이 마지막 단계에서 방향을 잡아주었다.

a) 서울대학교 사회학과 부교수 한신갑.
E−mail: shinkaphan@snu.ac.kr

Mixed-mode survey and (volunteer) web panels are the alternative methods introduced of late as ways to cope with the changing environment. Their formal and substantive bases are reviewed in this paper. After placing them in the ecology of social survey in contemporary Korea, I examine their characteristics and limitations, along with the adjustment methods developed to deal with the problems they incur. Of central concern in the assessment is the fundamental question of whether, or how well, they represent the population. With computer simulations, the biases produced when these requirements are not met are illustrated. Lastly, some concrete and practical steps to take in using these new methods are discussed. The paper once again reminds the role expected of the social survey: It is an instrument to observe the social reality, and it is to provide references for making and evaluating policy as well as basic data for research. In doing so, it reconfirms that the lowering of cost merely for the sake of expediencies and without regard to data quality amounts to nothing useful or meaningful.

Key words: data quality, ecology of social survey, mixed-mode survey, volunteer web panel, adjustment

사회구성원들이 무엇을 생각하고 어떻게 행동하는지를 관찰하고 정리할 수 있게 하는 도구로서의 사회조사는 전후(戰後) 사회과학의 근간 중 하나다. 사회과학 내에서의 다양한 방법론적 입장에도 불구하고, 사회조사자료가 갖는 기초적 역할에 대해서는 넓은 공감대가 형성되어 있고, 그 공감대의 바탕에는 조사자료가 갖는 '대표성', '체계성', 그리고 '효율성'에 대한 신뢰가 깔려 있다.[1] 뒤집어 말하면 조사자료는 이러한 기초적 역할을 감당할 수 있도록 대표성, 체계성, 효율성을 갖춰야 한다는 당위적인 기대가 있는 것이다. 하지만 현재 우리나라의 조사현실은 이런 기대와는 상당한 거리가 있는 것도 사실이다. 지난 수십 년 간의 양적 성장에도 불구하고 그에 상응하는 질적 성숙은 아직 이루어지지 않고 있다는 것이 한국의 사회조사자료를 다루는 많은 사람들

1) 여기서의 '효율성'은 '싸게, 빨리, 많이'의 의미가 아니라, 일억 명에 달하는 미국의 유권자들을 대상으로 사백 명의 표본을 가지고 오차범위 5% 안에서 선거결과를 예측한다는 의미에서의 효율성이다.

에게 공통된 의견이다(차종천 2005; 김상욱 외 2011; 김석호 2011). 사회환경이나 기술환경 등이 바뀜에 따라 자연스럽게 나타나는 여러 가지 변화와 그에 따른 문제점들이 조사환경에 대한 심각한 위협으로 다가오는 것은 바로 이런 양적 성장과 질적 성숙 간의 불균형이 더 악화될 수도 있다는 우려 때문이다(Dillman, Smyth & Christian 2009). 사회조사가 학술연구뿐만 아니라 정책수립 및 집행과정에 기초자료를 제공하는 공공재(公共財), 중간재(中間財), 사회간접재(社會間接財)로서의 역할을 한다는 점을 진지하게 받아들인다면, '대표성, 체계성, 효율성'이라는 사회조사의 필요조건에 대한 이런 논의는 단순히 기술적, 방법론적 문제에 그치지 않고, 보다 큰 사회적 함의를 갖는다(Anderson & Fienberg 1999; National Research Council 2007).

이 글에서는 이런 변화들 중 최근 사용빈도가 높아지고 있고, 특히 선거관련 여론조사와 시장조사에서 조사업계의 표준으로 급속히 자리잡아 가고 있는 두 가지 기법에 대해 그 문제점들을 표본추출과 표집과정을 중심으로 논의하고자 한다. 하나는 '혼합식 조사(MMS: Mixed-Mode Survey, 또는 Multi-Mode Survey)'다. 혼합식 조사는 조사대상에 대한 접근성의 악화 등 사회적 환경의 변화로 생겨나는 문제점들에 효과적으로 대처할 수 있는 대안을 찾아야 한다는 필요의 측면과 기술적 환경의 변화와 함께 사용할 수 있는 자료수집의 도구들이 다양해졌다는 가능의 측면이 맞물리면서 점차 사용이 늘어나고 있는데, 특히 무응답 문제에 대한 '현실적'이고 '경제적'인 해결책의 하나로 도입되고 있다. 같은 맥락에서, 때로는 독자적으로 때로는 혼합식 조사의 한 부분으로 자리잡아 가고 있는 것이 '(자원자) 웹패널(volunteer web panel)'이다.

구체적인 내용은 다르지만 각각에게서 나름의 존재이유를 찾을 수 있는 만큼, 이들에 대한 논의는 현재 한국의 조사환경이라는 맥락 속에서 어떻게 이들이 구현되고 있는가에 초점을 두고 이루어져야 할 것이다. 다음 절에서 다루는 것이 이런 전체적 맥락에 대한 논의이다. 사회조사의 생태계라는 관점에서 현재의 상황을 살펴보고, 가장 쉽게 눈에 띄는 비용에 대한 근시안적 대증요법이 대표성, 체계성, 효율성이라는 사회조사의 기본적 요건과는 상반되는 결과를 낼 수 있다는 점을 지적한다. 본론에 해당하는 2절에서는 현재 쓰이고 있는 혼합식 조사와 웹패널을 비롯한 온라인/인터넷 조사의 구성과 절차, 그리고 문제점 등을 다룬다. 특히 이런 문제점들을 보정하기 위해 개발된 방법들이 얼마나, 또 어떻게, '모집단을 대표할 수 있는 표본의 추출이 가능한가?' 하는 가장 기본적인 질문에 답할 수 있는지를 논의하고, 이 문제가 제대로 다루어지지 않았을 때 발생할 수 있는 편향을 모의실험을 통해 살펴본다. 마지막 절에서는 앞서의

논의를 정리하고, 이 글에서 살펴본 새로운 방법들을 도입, 사용하는 데 있어서 취해야 할 구체적이고 실질적인 방안들을 선발주자의 경험에서 찾아본다. 그리고 그 실천을 위해 재정립이 요구되는 사고의 기본적 틀에 대한 제언으로 글을 맺는다.

Ⅰ. 진단: 증상 · 병인

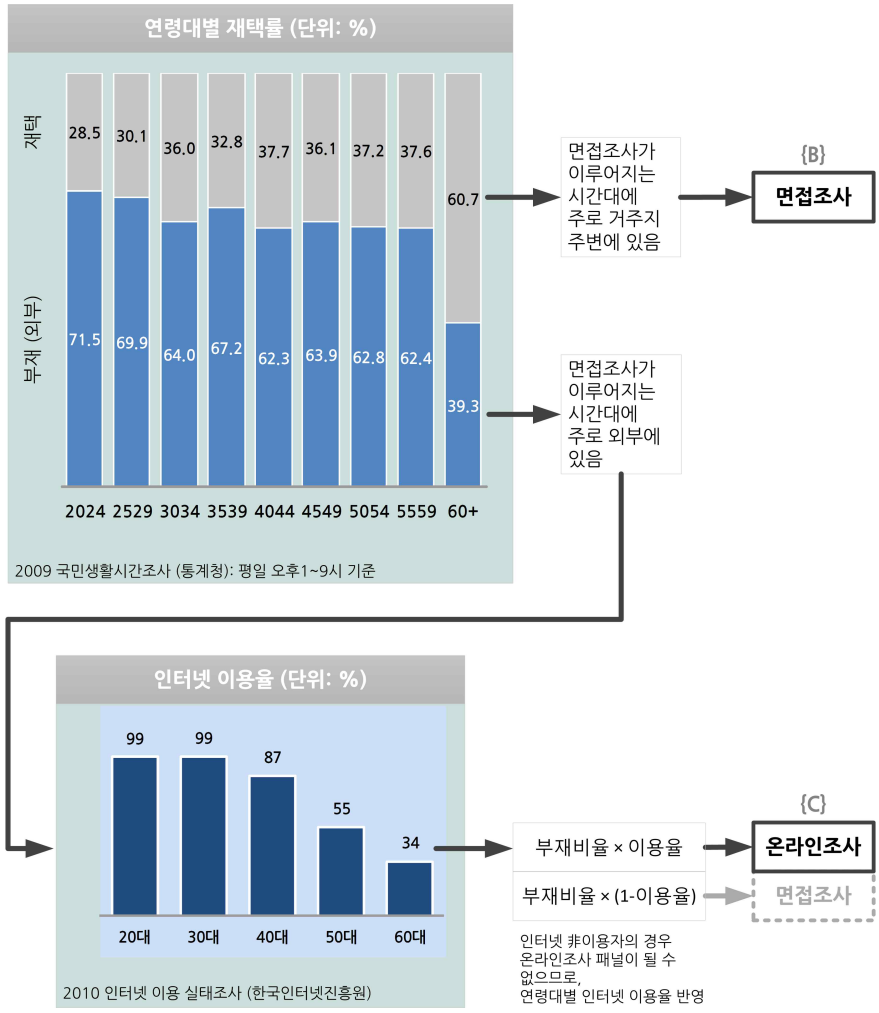
1. 서울, 2011년, ○사

지난 해 한 중견 조사회사(이하 '○사'로 지칭)의 제안설명을 들으면서, 상당히 혼란스러웠던 경험이 있다. 제안의 내용 자체는 낯익은 것이었지만, 듣는 사람에게 혼란을 준 것은 '어차피 제대로는 못 한다', '원칙은 원칙일 뿐이다', '이 정도면 된다', '다들 이렇게 한다'는 식의 제안의 기조였다. 기존의 방법들로는 '어차피' 표본대표성을 얻을 수 없으니까, 온라인을 포함한 자기 회사 방식의 혼합식 조사로 하자는 것이었고, 그 혼합식 조사에서도 사전, 사후의 절차들을 '제대로' 챙긴다는 것은 시간과 비용 면에서 비현실적이므로, 온라인의 경우 자기 회사의 웹패널을 쓰자는 것이었다. 현실적이어야 한다는 것에 기본적인 수긍을 할 수는 있지만, 그렇게 그려낸 그림이 '실제'와 다르다면 어떻게 해야 하나? 더 나아가 그 그림이 실제와 맞는지 아닌지를 확인조차 할 수 없다면 어떻게 해야 하나?

○사가 제안한 표집의 방법을 정리한 것이 〈그림 1〉이다. 다소 복잡해 보이지만 결국은 '재택', 즉 방문이 가능한 사람들은 면접으로, 그렇지 않은 경우는 온라인으로 조사하는 혼합식 방법이다.[2] 특히 결과적으로 표본의 과반수 이상을 만들어 내는 온라인 조사의 경우, ○사가 구성해 유지하고 있는 웹패널을 사용한다는 것이었다. 〈그림 1〉의 표집설계에서 상정하는 틀은 전체 모집단을 재택여부와 인터넷 이용여부로 나눈 〈표 1〉의 2×2 교차표로 정리해 볼 수 있다. 물론 표집의 대상이 되는 이론적 모집단은 {a, B, C, d}이다. 그러나 이 중 ○사가 실제로 표집하는 대상이 무엇인지는 분명치 않다. 원칙적으로는 두 방법 중 어느 것으로도 잡아낼 수 없는 {d}를 제외한 {a, B, C}이어야 한다. 하지만 현재 일반적으로 알려진 인터넷 사용자의 특성 분포를 고려하면서(임경은 2009; 이윤석 · 이지영 · 이경택 2008), 〈그림 1〉의 화살표를 짚어 가면, 실

2) ○사는 이 방법을 '이중표집틀 혼합식 조사(dual-frame MMS)'라고 부른다.

제적으로 다다르는 곳은 면접조사를 통해 표집되는 {B}와 온라인조사를 통해 표집되는 {C}이다. 이 둘을 합친 {B, C}가 어떤 문제를 갖는지는 〈표 1〉의 하단에서처럼 구분 기준 중 하나를 보다 실질적인 예로 바꿔놓고 보면 분명히 드러난다. 이 경우 {B, C}가 잡아내는 것은 인터넷을 이용하는 남성들({C})과 그렇지 않은 여성들({B})이고, 이 둘만으로 전체 인구를 대표하게 된다. 이런 방법이 대표성을 가질 수 있는 유일한 경우는, 두 구분의 축 사이에 아무런 상관관계가 없을 때뿐이다.



〈그림 1〉 ○사의 표집 방법

〈표 1〉 재택 및 인터넷 이용여부에 의한 표집대상 구분

		인터넷 이용	
		예	아니오
재택	예	a	B
	아니오	C	d

		인터넷 이용	
		예	아니오
성별	여성	a	B
	남성	C	d

더군다나 이 중 {C}에 해당하는 부분은 인터넷 이용이 가능한 실제의 모집단을 대상으로 하는 것이 아니라, ○사의 자원자 웹패널({C′})로 그 대상을 대체해 표집하는 것이다. 결국 표집대상은 {a, B, C}가 아닌 {B, C′}가 된다. 이 둘이 같아지려면, 엄밀하게는, {a}=∅, {C}={C′}라는 두 조건이 충족되어야 한다. 둘 모두 아무런 이론적, 경험적 근거 없이 쉽게 받아들일 수 있는 가정들이 아니다. 최대한 현실적인 고려를 하더라도 두 가지는 검토되어야 한다. 하나는 전체 인구, 즉 모집단이 재택여부와 인터넷 이용여부라는 두 축을 기준으로 어떻게 분포되어 있는지를 알아야만 {a}의 크기를 추정할 수 있다는 점이고, 또 하나는, {C′}를 {C}의 확률표본으로 볼 수 없는 이상, {C}와 대비한 {C′}의 분포도 알아야 한다는 점이다.

〈표 2〉 ○사의 선거결과 예측

(1) □□도 도지사 선거

	후보 가(%)	후보 나(%)	오차(%)
실제	47.7	52.3	
K방송사	57.0	43.0	−9.3
A신문사	61.6	38.4	−13.9
E사(MMS)	52.2	47.8	−4.5

(2) △△△선거구 국회의원 선거

	후보 다(%)	후보 라(%)	오차(%)
실제	48.6	51.4	
K방송사	49.4	50.6	−0.8
A신문사	57.2	42.8	−8.6
J신문사	47.9	52.1	0.7
E사(MMS)	48.8	51.2	−0.2

* 2011년 4월 27일 재보궐선거

그날의 제안설명에서 이런 내용들은 언급되지 않았다.[3] 대신 자신들이 제안한 방법의 효용성을 보여주는 근거로 제시한 것은 〈표 2〉에 정리한 두 개의 선거결과였다.[4] 어떻게 보면 현재 우리나라 조사현실의 총체적 위기상황에 대한 '예시1'처럼 보이는 이 자료를 통해 ○사가 보여주려고 했던 것은 자신들의 조사가 다른 회사의 결과에 비해 상대적으로 오차가 작았다는 점이다. 하지만 과연 이 결과가 조사과정상의 여러 문제점들을 상쇄할 수 있을 만큼 튼튼한 것일까? 두 선거 중 한 군데에서만 예측에 성공했다는 것, 오차의 크기나 방향성이 불안정하다는 것 등을 차치한다 하더라도, 앞에서 언급한 문제점들을 고려할 때 이 결과만을 가지고 ○사의 조사방법과 표집틀이 현실을 제대로 반영하는지를 판단하는 근거로 삼을 수 있을까? 이런 우려에 더욱 무게가 실리는 것은 이것이 ○사에서만 발견되는 제한된 현상이 아니고 현재 국내에서 이루어지고 있는 조사들 전반에 나타나는 일반적인 것이기 때문이다(이흥철 2002).

2. 사회조사의 생태계

사회조사의 수요자 쪽에서 원론적으로 기대하는 것은 분석하고자 하는 '모집단 전체를 대상으로 한 무작위 확률표본 추출'이라고 쉽게 규정할 수 있다. 반면에 실제로 공급되는 자료의 성격을 규정하기는 쉽지 않다. 조사자료의 생산과정에 여러 가지 요인들이 복합적으로 연관되어 있기 때문이다. 이 복합적 연관성을 잘 나타내는 개념이 '사회조사의 생태계'다. 통신기술을 중심으로 하는 기술환경, 각종 조사의 수요를 결정하는 시장환경, 관련법과 제도를 포함하는 사회환경 등 다양한 외재적 요인들과 조사의 이론/방법론의 개발과 적용이라는 내재적 요인들이 합쳐져 사회조사의 생태계라는 틀이 만들어지고, 사회조사를 생산하는 공급자와 소비하는 수요자는 그 틀 안에서 각자 자신들의 위치를 찾는다. 이 생태계는 끊임없이 변하기 마련이고, 그런 변화들은 당연히 사회조사의 양과 질에 영향을 미친다(de Leeuw 2005; de Leeuw & Hox 2011).

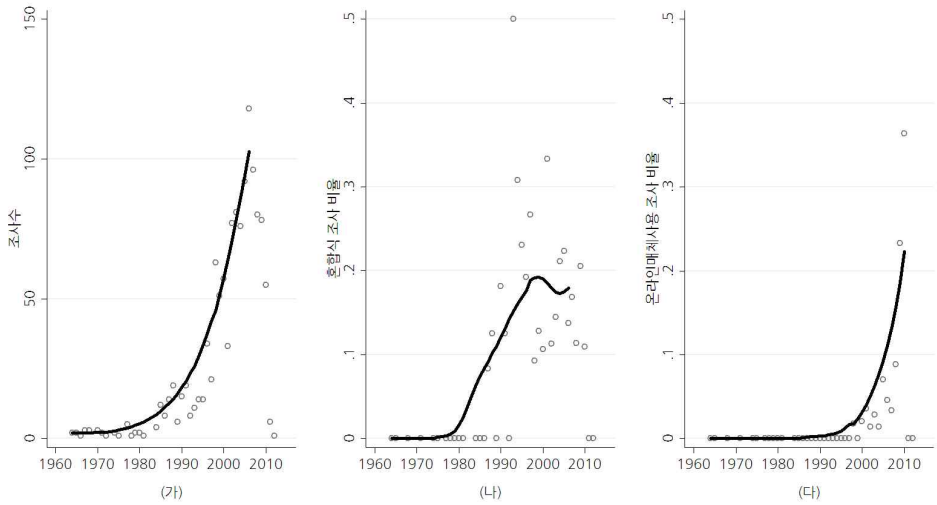
이 관점에서 한국의 사회조사 생태계를 볼 때 가장 먼저 눈에 띄는 것은 양적인 면에서 엄청난 성장을 해 왔다는 것이다. 한국사회과학자료원(KOSSDA)이 2012년 3월

---

3) 또 〈그림 1〉에 의하면 ○사가 웹패널을 표집에 사용하는 틀은 확률표집이 아닌 할당표집인 것으로 보인다. 할당표집의 문제점에 대해서는 구혜란 · 김석호(2011)를 참조하라. 이와 관련하여 또 하나 언급되지 않은 것은 그림의 오른 쪽 하단에 나오는 [부재비율 × (1 − 이용율) − ▸면접조사]라는 부분으로, 〈표 1〉에서는 {d}에 해당하는 부분이다.

4) 다른 비교자료들은 모두 RDD 전화조사를 통해 이루어졌다.

현재 아카이브하고 있는 자료들의 조사년도 분포(<그림 2-가>)를 통해 이 추세를 살펴보면, 1960년대부터 현재까지 50년에 걸친 기간 동안 조사의 수는 기하급수적인 증가세를 보인다는 것, 특히 1980년대와 1990년대를 거치면서 급속하게 그 생산규모가 커졌다는 것을 확인할 수 있다.[5] 이 성장의 외연과 함께 또 하나 주목해야 할 것은 조사방법에 있어서의 변화다. <그림 2-나>와 <그림 2-다>에서 보듯, 1980년대 중반부터 우편, 전화 등의 매체를 면접과 함께 사용하는 혼합식 조사방법의 비중이 크게 높아졌고, 2000년대에 들어오면서는 이메일이나 인터넷 등 온라인 매체를 단독으로 또는 혼합식 방법의 일환으로 사용하는 조사의 수가 빠르게 늘어가고 있음을 볼 수 있다.



* 자료 : 2012년 3월 현재 한국사회과학자료원(KOSSDA)이 아카이브하고 있는 총 1,194건의 조사자료를 대상으로 함.

<그림 2> 한국의 사회조사, 1960~2010

5) <그림 2>를 한국 사회조사의 전반에 나타나는 추세로 읽기 위해서는 두 가지 사항을 고려해야 한다. 그 하나는 2006년 한국사회과학자료원(또는 1983년 그 전신인 한국사회조사데이터아카이브)이 설립되기 이전의 자료들을 아카이브하는 과정에서의 시간경과로 인한 제약(예, 자료유실)과 그로 인한 편향이다. 즉 초기의 자료들일수록 수집가능성이 낮다는 점이다. 또 하나는 이 아카이브에 포함된 자료들은 주로 사회과학과 그 연관분야의 것들로 그 외 분야의 조사자료들은 포함하지 않는다는 점이다. 이 두 가지를 고려하더라도 <그림 2>의 대체적 추세는 변함이 없을 것으로 본다.

〈표 3〉 응답자의 조사과정 참여와 통제의 시대적 변화

조사의 측면	1960년대까지	1970년대와 1980년대	1990년대부터 현재까지
접촉의 직접성 ①	높음: 응답자 가구를 직접 방문해 대면접촉	중간: 전화연결을 통한 원격 접촉	낮음: 기기와 프로그램을 통한 접촉
조사에 대한 신뢰기반 ②	높음: 조사자의 현장 실재에 의한 다양한 강화작용	중간: 음성의 현재감과 대응성에 의한 강화작용	낮음: 허위나 잠재적 유해성 등에 대한 우려
개별 접촉시간 ③	높음: 조사자가 응답자를 방문, 1:1로 정보 수집	중간: 1:1이지만 접촉에 드는 노력은 최소화	낮음: 개별 접촉시간은 최소 또는 전무
개별 접촉밀도 ④	높음: 응답자 소재 확인과 면접을 통해 유지	중간: 순차적으로 통화 시도	낮음: 대량 이메일
응답자의 접근통제 ⑤	낮음: 가구 접근 용이	중간: 미등재 번호, 자동응답기, 수신 모니터링 등	높음: 발신자 확인, 수신차단, 이메일 필터 등
응답자의 참여여부선택 ⑥	낮음: 대면접촉 중단 필요	중간: 전화통화 종료 용이	높음: 사전 동의과정 필요, 참여거부 사회적 인정

* 출처: Dillman, Smyth & Christian (2009: 2). Figure 1.1.

〈표 3〉은 지난 75년간 서구에서 경험한 사회조사 생태계의 변화들을 정리해 보여준다(Dillman, Smyth & Christian 2009:2). 기간이 상대적으로 짧았고, 따라서 변화의 속도가 조금 더 빠르기는 했지만, 그 흐름의 방향은 한국에서도 대체로 같았음을 볼 수 있다(차종천 2005; 김상욱 외 2011). 이 기간 중 저자들이 특히 주목하는 것은 '격동의 시기(turbulent times)'라고 부르는 1990년대 이후의 상황이다. ①에서 ⑥까지의 어느 면을 보든 현재의 상황이 낙관적이지 않은 것은 분명하다. 핵가족화와 가족해체, 여성취업률 증가, 주거형태의 변화와 같은 사회적, 인구학적 추세에서부터 휴대전화(mobile phone), 도어락(doorlock) 등 일상생활에 쓰이는 기기의 보급에 이르기까지 다양한 요인들이 가져온 조사환경의 변화가 응답자들에 대한 접근을 점점 더 어렵게 만들었고(⑤,⑥),[6] 따라서 실제 조사과정에서 접촉의 깊이를 잃게 되었으며(③,④), 그에

6) 이 같은 추세들이 사회조사에 얼마나 큰 영향을 주는지의 한 예로 통계청의 "장래가구추계:

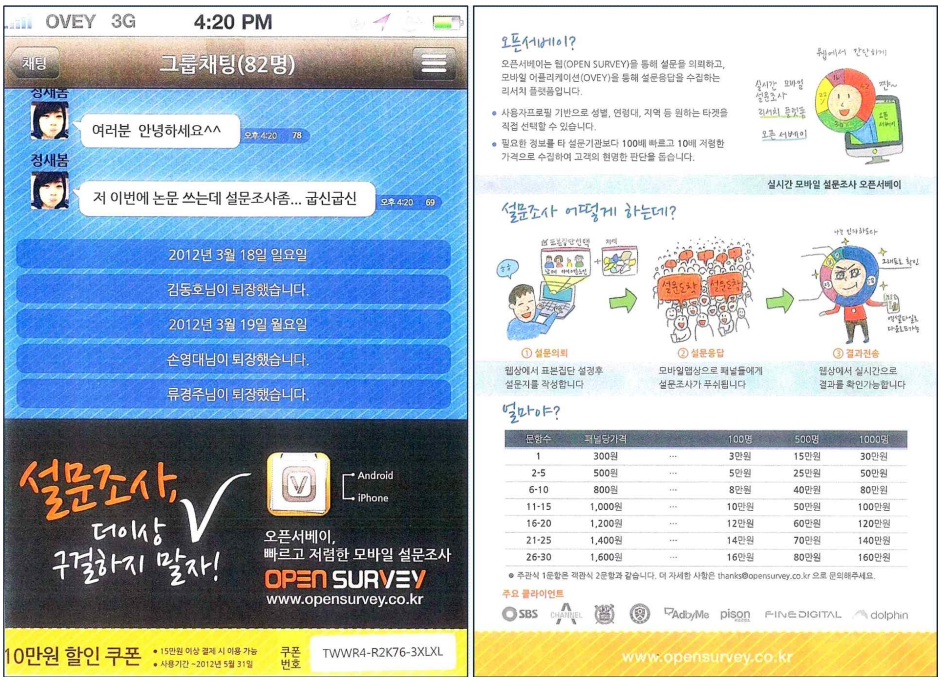
대한 대책('궁여지책')으로, 문제점이 있음에도 불구하고 새로운 매체와 기법을 사용하고 있다는 것이(①,②) 이 표에서 내리는 진단이다(Groves & Couper 1998; Smith 2010).

이런 관점에서 <표 3>에서 구분한 세 시기를 비교해 보면 가장 중요한 변화는 조사의 비용과 효율의 변화이다. 즉, 이 둘 사이의 긴장관계와 그 해소방법의 탐색이 지난 75년간 서구에서 관찰된, 그리고 우리가 지난 50년간 경험한, 사회조사 생태계의 변화를 추동(推動)해 온 주된 내적 요인이었던 것으로 보인다. 초기의 저비용-고효율 환경이 점차 고비용-저효율로 바뀜에 따라 비용은 낮추고 효율은 높이기 위한 새로운 도구와 기법들을 끊임없이 찾아온 과정으로 이 표의 내용을 이해할 수 있기 때문이다(de Leeuw & de Heer 2002; Groves et al. 2009).

그 과정에서 개발되어 현재 쓰이고 있는 다양한 조사방법들을 비용을 기준으로 살펴보면 '온라인 < 우편 < 전화 < 면접'의 순서로 정리할 수 있다. 국가와 상황에 따라, 또 조사의 유형에 따라, 구체적 액수에는 차이가 있지만, 대체로 온라인조사에 드는 비용은 면접조사의 십분의 일, 또는 그보다 더 적은 것으로 알려져 있다 (de Leeuw 2005). 이 비용구조가 사회조사 생태계에 어떤 영향을 줄 것인지는 자명하다. 생태계의 먹이사슬에서 비용은 바로 '먹이'에 해당하기 때문이다.

이렇게 한쪽으로 기운 비용구조로 인한 '쏠림현상'을 막아내기에는 역부족이지만, 마지막 남은 유일한 견제의 방편은 효율의 문제를 다시 제기하는 것이다. 효율은 비용과 산출의 대비이다. 따라서 산출(output), 즉 조사의 양과 질이 유지되지 않으면 비용을 낮추는 것만으로 효율을 높일 수는 없다는 것을 환기시키는 것이다. ○사의 경우에서 보는 것처럼, 최근에 도입되고 있는 방법들은 표본의 크기나 응답률 등 양적인 측면의 산출을 확보하는 데 일차적 주안점을 두고 있고, 그 부분에서 상당한 성과를 거두고 있는 것으로 보인다. 심지어 어떤 조사회사는, <그림 3>에서 보는 것처럼, '100배 빠르고 10배 저렴한 가격으로' 필요한 정보를 수집할 수 있다고까지 한다.

2010년~2035년"(2012)를 살펴보자. 이 자료에 따르면, 2010년 1인 가구수는 415만 3천 가구로, 전체가구수의 23.9%에 해당한다. 이 비율은 2012년 현재 25.3%로 가구원수별 가구구성비에서 가장 큰 비중을 차지하고 있을 것으로 추정되고, 2035년에는 전체의 34.3%로 삼분의 일 선을 넘을 것이라는 전망이다. 이런 증가추세는 사회조사에 큰 문제점을 던진다. 이들이 가구방문이라는 전통적인 방법으로 가장 포착하기 힘들었던 집단이기 때문이고, 또 웹패널로 대체하기에도 어려움이 있기 때문이다.



〈그림 3〉 한 조사회사의 광고

문제의 핵심은 자료의 질(data quality)이다. “자료의 질에 대한 고려 없이 비용만을 논하는 것은 아무런 의미도 없다”던 Deming(1960)의 경고가 아니더라도, 〈그림 3〉에서처럼 100배 빠르고 10배 싸게 만든 자료가 조사대상에 대하여 제대로 된 정보를 주지 못한다면, 그 자료를 만드는 과정에서의 효율은 무슨 의미가 있을까? 이런 관점에서 볼 때, 현재의 추세에서 가장 심각한 문제는 보통 ‘표본’이라고 할 때 기본적으로 가정하는 ‘모집단 전체를 대상으로 한 무작위 확률표본 추출’이라는 원칙과 이상에서 아무런 고민이나 방향감각 없이 이탈해 가고 있다는 것이다. ‘어차피 제대로는 못 한다’, ‘원칙은 원칙일 뿐이다’, ‘이 정도면 된다’, ‘다들 이렇게 한다’는 식의 ‘현실논리’는 결국 만들어 내는 자료의 질에 관계없이 싸게, 빨리, 많이만 만들면 된다는 논리로 이어지고, 그것을 효율성이라고 부른다면, 그 효율성은 대표성과 체계성을 갖추지 못한 ‘부실공사’의 효율성이다. 사회조사에 거는 당위적인 기대, 사회조사의 기초적 역할에 대한 공감대는 이 기본적인 가정에 근거하고 있다. 그리고 그 기초가, 또 그 기초에 대한 신뢰가 흔들리면 조사 생태계 전체가 위협받게 된다.

Ⅱ. 처방: 효능·효과

1. 혼합식 조사

현재 이런 조사환경의 변화에 대한 대응은 주로 기술적인 방법을 통한 것이고, 그 중 가장 대표적인 것이 혼합식 조사다. 앞에서 다룬 사회조사 생태계의 변화가 전통적인 방식의 조사를 어렵게 만들고 있다는 문제의식에서 출발하여, 한편으로는 '온라인 < 우편 < 전화 < 면접'의 비용구조, 또 한편으로는 표집범위오차(coverage error), 표집오차(sampling error), 무응답오차(nonresponse error), 측정오차(measurement error) 등 자료의 질을 결정하는 다양한 요인들을 고려하면서, 다양한 조사방법들의 장단점과 상대적 효율성에 대한 검토에 바탕을 두고 이들을 상호 보완할 수 있도록 합친다는 전략이다. 원론대로라면 상당한 매력을 가지는 전략이고(Revilla 2010; Roberts 2007), 특히 표집범위나 무응답과 관련된 문제들의 상당 부분을 혼합을 통해 '경제적으로' 해소할 수 있을 것이라는 기대가 이 방법이 널리, 그리고 빠르게 보급된 중요한 이유가 되었다(Martin 2011; de Leeuw 2005; Dillman, Smyth & Christian 2009).[7]

미국의 통계청에서 실시하고 있는 American Community Survey(ACS)가 혼합식 조사의 가장 성공적인 예의 하나로 꼽히는데, 매년 실시되는 표본크기 이백만 규모의 전국조사에서 우편, 전화, 면접의 세 가지 방법을 순차적으로 적용하여 〈표 4〉에서 보듯이 95%가 넘는 응답률을 얻어내고 있다.[8] ACS에서는 우선 가장 비용이 저렴한 우편을 통해 표집대상의 절반이 넘는 응답자에 대한 자료를 얻고, 그 다음 전화접촉을 통해 응답률을 추가로 끌어올린 후, 마지막으로 가장 비싼 방법인 방문면접을 통해 95% 이상의 최종응답률을 확보한다(Griffin & Obenski 2002). 비용의 측면을 감안하면서도 자료의 질을 수준 높게 유지할 수 있는 최적의 배합을 찾은 경우다. 하지만 많은 경우, 이 균형은 '제대로 하기'보다는 '싸게 하기'쪽으로 기운다. 또 자료의 질을 고려할 때도 위에서 언급한 자료의 질을 결정하는 네 가지 요인 모두에 대해 고르게 관심이 주어지기보다는 기술적 해결이 쉬운 무응답오차를 줄이는 데 집중적으로, 때로는 독점적으로 주어진다.

---

7) 혼합식 조사가 보급, 확산되면서 국내에서도 이에 대한 관심이 기대 半, 우려 半의 형태로 나타나고 있다. 2011년 한국조사연구학회 춘계학술대회에서 "다매체시대 여론조사 발전의 가능성"이란 제목으로, 또 추계학술대회에서 "뉴미디어를 이용한 여론조사 방법의 패러다임 변화"라는 제목으로 조직된 기획세션들이 이런 관심을 반영한다.

8) www.census.gov/acs/www/. 특히 방법론 부분의 '표본크기와 자료의 질'에 대한 항목을 보라.

〈표 4〉 ACS의 표본 크기와 자료의 질

	Housing Units			Group Quarters Population		
년도	Sample Size*	Response Rates(%)	Coverage Rates(%)	Sample Size*	Response Rates(%)	Coverage Rates(%)
2010	1,917,799	97.5	99.1	144,948	97.6	81.0
2009	1,917,748	98.0	98.9	146,716	98.0	79.1
2008	1,931,955	97.9	98.7	145,974	98.0	80.8
2007	1,937,659	97.7	98.5	142,468	97.8	79.6
2006	1,968,362	97.5	98.7	145,311	97.4	76.2
2005	1,924,527	97.3	98.5	N/A	N/A	N/A
2004	568,966	93.1	97.9	N/A	N/A	N/A
2003	572,447	96.7	97.9	N/A	N/A	N/A
2002	512,768	97.7	97.2	N/A	N/A	N/A
2001	601,875	96.7	97.6	N/A	N/A	N/A
2000	587,519	95.1	99.6	N/A	N/A	N/A

* Number of final interviews.

이런 균형과 비중의 문제가 아니더라도 혼합식 조사는 혼합의 과정 자체에서 오는 자료의 온전성, 또는 일체성의 문제를 갖는다. 두 가지 이상의 조사방식을 통해 만들어진 자료를 하나로 묶어 쓰기 위해서는 합쳐지는 자료들이 등가(equivalent), 또는 비교가능(comparable)하다는 것을 전제할 수 있어야 한다. 즉 조사방식의 차이가 그 방식들을 통해 만들어진 자료에도 영향을 주어 편향으로 남아있어서는 안 된다는 것이다. 이런 차이를 '방식효과(mode effect)'라고 부르는데, 방식효과의 생성기제는 크게 두 가지로 정리된다. 하나는 조사방식에 따라 표집대상의 정의와 범위가 달라져 응답자들이 선별적으로 뽑히는 데서 생기는 차이이다. 예를 들면, 모바일 폰을 통한 조사와 유선전화를 통한 조사에서 노년층과 청년층의 접촉가능성은 매체에 따라 현격한 차이가 나는데, 이 둘을 하나의 자료로 묶어낼 때 생기는 문제점에 대한 것이다(Link et al. 2007). 또 하나는 조사방식에 따라 응답내용이 달라지는 데서 생기는 차이로, 잘 알려진 예는 민감한 사생활이나 개인정보에 대한 설문문항에서 면접과 인터넷 조사 간에 뚜렷한 차이가 있다는 것이다(Martin & Lynn 2011; Revilla 2010; Vannieuwenhuyze, Loosveldt & Molenberghs 2010).

조사방식이 조사의 과정과 그 결과로서의 자료에 영향을 주는 만큼, 각기 다른 방식에 의해 만들어진 자료들이 차이를 보인다는 것 자체는 예상할 수 있는 일이다. 이렇게 차이가 나는 자료들을 혼합했을 때 방식효과로 인해 생기는 차이와 실질적인 효과로 인해 나타나는 차이가 서로 섞여 구분해 낼 수 없게 된다는 데서 문제가 생긴다. 아래에서 더 다루겠지만 이 방식효과의 문제에 대해서는 현재 많은 연구가 진행되고 있고, 다양한 기술적 대안들이 제시되고 있다.

2. 온라인/인터넷 조사와 웹패널

2000년 한국에서 조사연구학회가 만들어지면서 창간된 ≪조사연구≫의 제1권 제1호에 권두논문으로 실린 글은 이계오의 '통계 조사의 현황과 전망'이었다. 그 글에서 저자는 앞으로 인터넷 조사가 전화조사를 '대체'할 것이라고 전망하면서, 그 이유로 낮은 조사비용과 짧은 소요기간 등 인터넷 조사의 월등한 효율성을 들고 있다. 이 '효율성'에 인터넷이라는 새로운 기술에 대한 막연한 호기심과 과장된 기대가 섞이면서 인터넷 조사는 급속히 확산되었고, 앞으로도 그 활용은 더욱 늘어날 것이 분명하다(Couper 2000; Dillman, Smyth & Christian 2009). 하지만, "일반적인 모집단을 대표할 수 있는 표본추출의 틀을 얻을 수 없는 경우가 많다"는 이계오의 지적에서처럼 (2000:9), 이러한 비용 면에서의 낙관적 전망은 그것이 갖는 방법상의, 특히 자료의 질 측면에서의, 문제점과 함께 고려되어야 한다(de Leeuw 2005; Sparrow & Curtice 2004).

문제의 핵심은 인터넷, 온라인이라는 매체의 선택에 따라 표집의 기본적인 틀이 바뀐다는 것이다.[9] 우선 표집의 목표 모집단 자체가 이 매체를 통해 접촉이 가능한가의 여부로 규정되어 '편의표집'의 성격을 갖게 되고, 이런 접촉대상 중 누가, 어떻게 참여하느냐에 따라 '자발표본'이 되기도 한다(Taylor 2000; 조동기 2000). 그 결과 표본의 대표성에 문제가 생기는데, 이렇게 표집확률이 알려져 있지 않거나 통제되지 않은 비확률표집의 경우 확률에 근거한 통계이론을 적용할 수 없게 되고, 따라서 유의확률이나 오차범위 등 통상적인 검증절차를 사용할 수 없게 된다. 인터넷 조사의 질에 대해 긍정적인 평가를 내리는 경우에도, 아직은 마케팅 관련 조사나 제한된 특정 집단을 목표로 하는 조사 등에서 보인 효과가 인터넷 사용의 일반화와 더불어 확대 적용될 수

9) 앞에서 언급한 방식효과의 두 번째 측면에서도 인터넷을 통한 조사는 인터페이스(interface)의 특성에서 기인하는 여러 가지 차이점들을 보인다(Krug 2005).

있을 것이라는(박무익 2000), 또 대표성 문제에 대한 방법론적 해결책이 조만간 개발될 것이라는 '전망'에 기초하고 있는 것이 대부분이다 (조성겸 · 강남준 2003; Roster et al. 2004). 하지만 이런 전망의 실현을 위해서는 여러 가지 검토와 사전 정지작업이 필요한 것으로 보인다(예, 이윤석 · 이지영 · 이경택 2008; 조성겸 · 주영수 · 조은희 2005).

특히 이 조사방법을 최근 많은 조사회사들이 확보해 운영하고 있는 자원자 패널과 결합시켜 사용할 때, 이런 문제점들은 더욱 첨예하게 드러난다(이흥철 2002; 허명회 · 조성겸 2010). 예를 들면, ○사의 경우 50만 명이 넘는 큰 규모의 패널을 확보해 운영하고 있는데, 이 자원자들의 속성과 그 분포, 또 그것이 목표 모집단과 어떤 식으로 얼마나 상응하는지 등을 결정하는 요인과 절차들에 대해 상세하게 공개하지 않고 있다.[10]

인터넷, 또는 온라인 조사라는 틀 안에서도 확률표집을 포함한 여러 조사형태가 가능하다. Couper(2000)는 〈표 5〉와 같이 이들을 크게 확률표집에 근거한 것과 그렇지 않은 것으로 나누는데, 최근 도입되고 있는 방법들은 대부분 후자(왼쪽 열)에 속하는 것들로 대표성의 측면에서 문제가 가장 심각한 것들이기도 하다. 가령, 주어진 정보만으로는 정확히 판단할 수 없지만, 〈그림 3〉에서 본 방법은 이 중 2번에 가까운 것으로 보인다. ○사가 쓰고 있는 방법은 7번에 가까운 것으로 보이지만, 〈그림 4〉에서 보는 자원자패널의 구성절차가 제대로 알려지지 않는다면 3번으로도 볼 수 있다. 당연한 것이지만, 2번이나 3번처럼 확률에 근거하지 않은 방법에 의해 얻어진 표본은 확률이론을 이용하여 모수를 추정할 수 없고, 따라서 표본을 근거로 모집단의 속성을 일반화하거나 추론하는 것이 불가능하다. 그리고 이런 대표성의 문제는 표본의 크기를 늘린다고 해결되는 것이 아니다. 즉, 패널의 규모와 자료의 대표성 간의 상관관계가 단순선형이 아닌 만큼, 패널의 규모를 두 배로 늘린다고 자료의 질이 두 배로 좋아지지는 않는다. 오히려 같은 방법, 같은 절차로 패널의 규모를 늘리면, 편향은 그대로 유지되고, 심지어 더 강화되어 나타날 가능성도 있다.

---

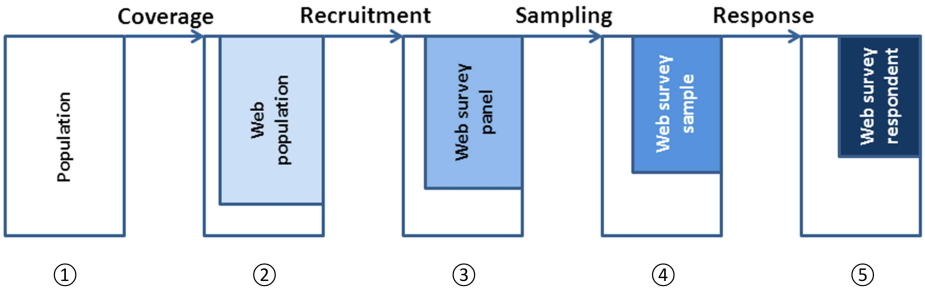
10) 2011년 5월 기준, ○사는 제안설명 자료에서 자신들의 패널현황을 소개하면서 "철저한 패널관리를 통한 높은 반응률을 기반으로 어떠한 조사라도 신속/정확히 수행할" 수 있다고 한다. 보다 구체적으로, "패널 전원에 대해 본인/실명 확인 절차를 거쳐 신뢰성을 확보한 패널을 구성"하고, "자신들만의 지속적이고 다양한 패널 관리기법 개발을 통해 패널들의 로열티를 확보"하고 있으며, "다년간 쌓은 관리기법과 패널 DB를 통해" 조사결과의 신뢰성을 확보하고 있다고 말하고 있다. 이 분야에서 세계적으로 가장 널리 알려진 Harris Poll Online의 경우에도 2000년 8월 기준으로 육백 오십만 명이라는 엄청난 규모의 패널을 확보하고 있지만 그 구성에 대해서는 상세한 정보를 공개하지 않고 있다(www.harrispollonline.com).

〈표 5〉 온라인 조사의 형태

확률에 근거하지 않은 방법	확률에 근거한 방법
1. Polls as entertainment 2. Unrestricted self-selected surveys 3. Volunteer opt-in panels	4. Intercept surveys 5. List-based samples 6. Web option in mixed-mode surveys 7. Pre-recruited panels of Internet users 8. Pre-recruited panels of full population

* 출처: Couper (2000: 477). Table 2.

자원자패널은 〈그림 4〉와 같은 절차를 거쳐 만들어진다(Lee 2006). 각각의 단계도 중요하지만, 이 모든 절차들이 하나의 사슬로 엮여 있어서 이 중 하나의 고리라도 약해지면 사슬 전체가 그만큼 약해지게 된다. 특히 ③의 웹패널(web survey panel)이 구성되기까지의 두 초기 단계가 결정적인데, 이 과정들에 대한 정확한 정보와 이해 없이 ⑤가 ①을 대표할 것이라고 가정하는 것은 위험한 일이다. ①에서 ②로 넘어가는 단계에서 확인되어야 할 것은 표집범위의 문제이다. 즉, 웹을 사용하는, 온라인을 통해 접촉이 가능한 인구가 전체 목표 모집단에서 어느 정도를 차지하는지, 그리고 그 인구가 모집단 안에서 특별한 편향성 없이 분포되어 있어서 이 방법을 사용하더라도 대표성을 갖는 표집이 가능한지의 문제이다. 인터넷의 보급이 늘면서 일반적 적용의 가능성이 점차 확대되기는 하겠지만(박무익 2000), 이 매체를 통한 접촉 가능여부가 조사자료에 아무런 편향을 미칠 수 없게 되는 완전한 보편화의 단계에 이르기 전까지는 대표성이라는 근본적인 문제가 남아 있게 된다.



* 출처: Lee (2006: 330). Figure 1.

〈그림 4〉 자원자 웹패널의 구성과 절차

그 다음 문제는 ②의 대상 중 누가 어떻게 참여해 ③을 구성하느냐 하는 것이다. 일반적인 자발표본의 문제에 더해 패널 모집과정에서의 다른 잠재적 편향까지 겹쳐지면(예, 'Do surveys, get free money.', 'money4surveys.com'에서 보는 것처럼 금전적 요인이 응답자 유치의 수단이 되면), 표본 대표성의 문제는 더욱 심각해진다.[11] 따라서 이런 과정과 절차에 대한 정보는 수요자 측에서는 반드시 요구해야 하고, 또 공급자 측에서는 당연히 제공해야 한다(Baker et al. 2010). 이 투명성의 문제는 그저 도덕적 차원의 것이 아니라, 확률표집이 아닌 방법으로 만들어진 표본에서 얻은 결과들을 이해하고 해석, 적용할 수 있도록 하기 위해 남아있는 유일한 수단이기 때문이다.

3. 기술적 보정의 가능성과 한계

이런 방법들이 현재 국내외 조사업계에 빠른 속도로 전파되고 있는 만큼 위에서 제기한 여러 문제점들에 대한, 특히 '모집단을 대표할 수 있는 표본의 추출이 가능한가?' 하는 가장 기본적인 질문에 대한 검토가 더욱 중요해진다. 지금까지의 연구들은 주로 이런 방법들을 사용했을 때 생길 수 있는 편향에 대한 방법론적, 기술적 보정에 집중하고 있는데, 크게 두 방향으로 이루어지고 있다. 하나는 결합추정방식이다. 두 가지 이상의 방법으로 조사를 진행한 후, 이 결과들을 결합하여 서로를 보완하고 수정하는 것이다.[12] 또 하나는 성향가중치(propensity score)를 이용한 보정방법이다. 특정한 조사방식에 의해 조사될 성향을 조정변수들로 모형화해 가중치를 구하는 것이다. 중요한 것은 이 중 어느 방법을 쓰더라도 보정을 위해서는 준거가 있어야 하고, 그 보정의 근거는 모집단의 분포가 돼야 한다는 것이다. 따라서 이런 준거가 부실하면 사후조정은 어느 방법을 쓰든 효과적일 수 없다(Couper 2000; Lee 2006; 이계오 · 장덕현 2009; 조성겸 · 박아현 · 허명회 2011).

---

11) 한 심사위원의 지적처럼 또 하나 문제되는 부분이 ④에서 ⑤로 넘어가는 단계이다. 패널참여자 중 조사에 적극적으로 참여하는 사람들('활성조사자')과 거의 참여하지 않는 사람들('비활성조사자')은 뚜렷하게 구분된다. 하지만 패널참여자들에 대한 개인정보를 가지고도 왜 누구는 조사에 자주 참여하고, 누구는 참여하지 않는지, 또 그러한 차이가 어떤 문제를 야기하는지에 대한 설명은 내놓지 못하고 있다.

12) 결합 후의 보정, 특히 방식효과로 인한 편향의 보정을 위해서는 마찬가지로 준거가 있어야 하고, 그 준거를 얻기 위해서는 작은 규모라 하더라도 체계적, 확률적으로 표집된 조사가 주어져야 한다. 이를 방식실험(mode experiment)이라고 부른다(de Leeuw 2005; de Leeuw & Hox 2011).

특히 후자의 성향가중치(propensity score)를 이용한 보정방법은 최근에 많은 주목을 받고 있는데, 나이와 성별, 거주지역 등 기본적인 인구통계학적 속성의 분포를 기준으로 이루어지던 종래의 가중치 부여방법에 대한 문제제기에 따라 제안된 방법이다(Rosenbaum & Rubin 1983; Rubin 1997; D'Agostino & Rubin 2000). 조사방법이 무작위확률에 근거해 결정된다면 성향가중에 사용되는 변수들과 조사방법은 조건부 독립의 관계일 것이다. 하지만 앞에서 다룬 혼합식 조사나 온라인/인터넷 조사, 특히 자원자 웹패널의 경우, 그런 원칙적인 관계가 성립할 가능성은 크지 않다. 그 결과 생겨나는 편향을 보정해 전체적인 균형을 맞추기 위해서는 그 편향이 어떻게 만들어지는가에 대한 이해가 필요한데, 성향가중모형을 이용한 보정에서는 이를 관찰된 자료로부터 추론하는 방법을 쓴다(Rosenbaum & Rubin 1983; Valliant & Dever 2011).

대부분 기초적인 인구통계학적 변수들과 더불어 응답자의 특성을 설명할 수 있고 예측변수와의 공변량이 높은 변수들을 성향가중을 위한 변수로 선택하는데, 이 보정방법의 기본적인 틀은 일련의 조정변수들이 주어졌을 때 응답자가 특정한 한 방식에 의해 조사될 가능성, 즉 조건부 확률을 모형화해 가중치를 구하는 것이다. 성향가중치를 이용한 보정이 조사자료의 편향을 완전히 제거하지는 못하지만 어느 정도 줄이는 효과가 있다는 데는 의견이 모아지고 있지만, 그에 따르는 문제점에 대한 논의도 있다.[13]

앞에서 언급했듯이 이 경우에도 준거표본은 필수적이다. 경우에 따라 이론적 표집틀을 같이하는 외부 자료에서 (유사)준거치를 구해 준거표본을 대체하기도 하지만,[14] 어떤 경우에도 보정을 위한 가중치의 값은 실제표본과 준거표본의 상대적 비율을 통해서만 얻을 수 있다(Lee 2006:334; 허명회 · 조성겸 2010; 조성겸 · 박아현 · 허명회 2011). 따라서 이 방법을 쓰는 경우, 보정의 성공여부는 일차적으로 준거자료의 질에 의해 결정된다. 조정을 위한 변수들의 선정, 즉 성향가중모형의 타당성과 적합성 등은 이 일차적 요건이 충족된 다음에 고려해야 할 사항들이다(Couper 2000).

---

13) 예를 들면, Lee(2006)는 추정치의 분산 증가에 대해 주의를 요구하고 있다. 또 하나는 성향가중치를 사용해 조정된 변수들을 실제 분석에서, 특히 평균값이나 비율 등의 단순추정치를 넘어서는 분석모형에서 어떻게 사용할 것인가 하는 것이다(Gelman 2007a, 2007b).

14) 이 경우에도 통계청 등의 자료로부터 전체적인 분포를 원용해 올 수는 있지만 해당 조사의 표집틀과 일치하도록 조정하는 추가적 작업이 필요하다.

다시 정리하자면, 조사비용도 줄이면서 표집범위오차와 무응답오차까지 줄일 수 있다는 기대가 이런 다양한 시도들을 추동하고 있고, 또 어느 정도의 성과를 내고 있는 것도 사실이다. 하지만 어떤 방법을 쓰든 그 방법의 적부를 고려하는 데 가장 중요한 것은 모집단과 표본이 연결되는 틀의 구조, 즉 표집의 범위, 기제와 과정이다. 이 틀의 기본구조가 제대로 짜여 있지 않으면, 또 표집의 범위, 기제와 과정에 대해 제대로 파악하고 있지 않으면, 보정방법을 어떻게 쓸 것인가와 보정결과를 어떻게 읽을 것인가 하는 문제에 대한 혼란을 피할 수 없다.

4. 모의실험 결과

현재 이런 방법들을 사용하는 조사는 〈표 2〉에서 본 ○사의 선거조사처럼 평균값이나 비율 등의 단순추정치를 얻기 위한 것들인 경우가 대부분이다. 여기서는 위에서 다룬 문제점들이 구체적으로 어떻게 드러나는지를 보기 위해, 선거조사의 경우를 사례로 삼아 간단한 모의실험(simulation)을 해 보았다.

우선 통계청의 전국인구추계에 따라 2011년 시점의 연령별 분포에 맞춘 인구 십만 명 규모의 가상선거구를 만들었다.[15] 현행법상 투표권의 기준인 만19세를 고려하여 이 중 가상유권자 77,432명을 추출하고, 이들에게 ○사가 사용한 자료(〈그림 1〉)의 연령별 확률에 따라 재택여부와 인터넷 이용여부 값을 부여했다.[16] 이렇게 구성된 가상유권자 인구의 접근가능성에 따른 분포는 〈표 6〉과 같다. 문제는 이 중 어느 부분이 표집의 대상이 되는가 하는 것이다. 표집대상이 달라짐에 따라 각각의 표본이 모집단의 분포를 어떻게 반영하는지를 보기 위해 표의 하단에 보이는 것과 같이 (가)에서 (라)까지 네 가지의 가능성을 상정하고, 각각에서 규모를 1,000으로 하는 확률표집을 100회씩 독립적으로 실행했다. 관찰기준변수는 정규분포를 따르되, ○사가 자신들의 조사설계에서 가장 핵심적인 기준으로 삼은 연령과 상관계수 0.6 정도의 관계를 갖도록 조작되었다.

---

15) http://sgis.nso.go.kr/pyramid/view_country.asp

16) 예를 들면, 20세부터 24세까지의 인구는 이항분포(binomial distribution)를 따라 28.5%의 확률로 '재택'이 되도록 해당 변수의 값이 주어졌다.

〈표 6〉 **재택 및 인터넷 이용여부에 따른 가상유권자 인구의 분포와 표집의 목표 모집단**

		인터넷 이용		합계(명)
		예(명)	아니오(명)	
재택	예	21,288	9,601	30,889
	아니오	36,665	9,878	46,543
합계(명)		57,953	19,479	77,432

(가)

	인터넷 이용	
재택	a	B
	C	d

(나)

	인터넷 이용	
재택	a	B

(다)

	인터넷 이용	
재택	a	
	C	

(라)

	인터넷 이용	
재택	a	B
	C	

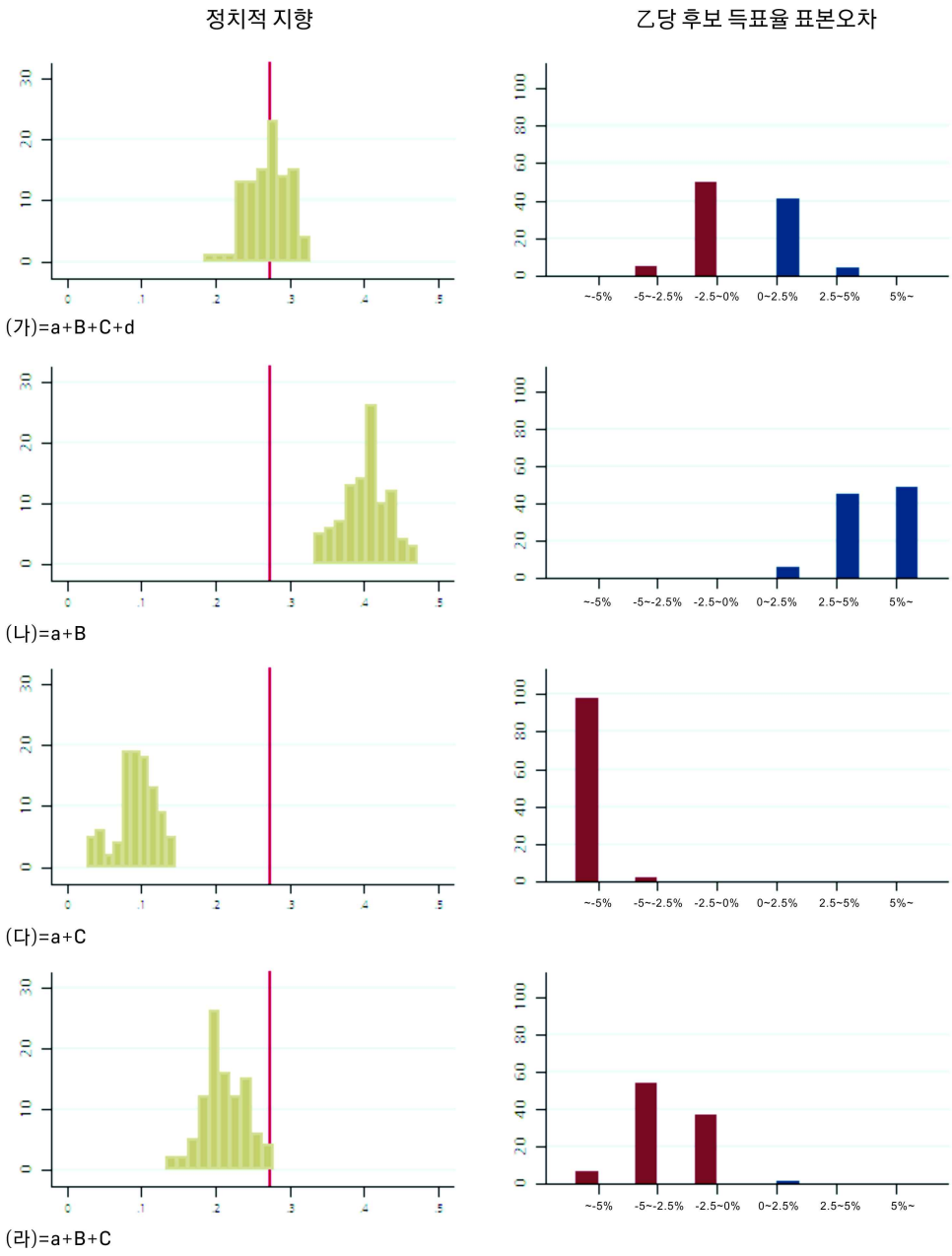
* 하단의 (가)부터 (라)까지 표기된 영문기호는 〈표 1〉에 맞추어 주어짐.

〈그림 5〉는 〈표 6〉과 같이 서로 다른 대상으로부터 표본을 추출했을 때, 관찰기준변수에 대한 추정치가 어떻게 달라지는가를 보여준다. 이 그림의 (가)행에 배치된 그림들은 제대로 된 표집이 가져야 할 모든 속성들을 보여주고 있다. 특히 편향과 관련하여 이론적으로 기대할 수 있는 분포를 보인다.[17] 하지만 (나)와 (다)행의 경우 분포에 있어 상반되는 방향으로 편향을 보이고 있다. 예를 들어, 관찰기준변수가 정치적 지향(진보 : 보수)을 모형화한다고 상정하면, (나)의 재택을 통해 표집된 인구가 보수 쪽으로 더 기우는데 반해, (다)의 인터넷을 통해서 표집된 인구는 진보 쪽으로 확연히 더 기울고 있다(이계오 · 장덕현 2009).

전체 모집단에서 얻는 관찰기준변수(정치적 지향)의 최소값과 최대값이 각각 −3.018과 4.972로 넓은 폭의 분포를 보이고 있는 만큼, 0.2 정도의 차이는 용납할 수 있는 것 아니냐는 주장도 있을 수 있다. 하지만 표본오차를 구해보면 이 추정치가 가지는 문제는 더욱 선명하게 드러난다. 정치적 지향에 따라 후보 선택과 투표가 이루어진다는 가정하에, 각각의 표본에서 얻어진 정치적 지향의 값을 투표결과로 환산해 볼 수 있다. 즉, 관찰기준변수의 값이 0보다 작으면 甲(왼쪽)당 후보에, 0보다 크면 乙(오른쪽)당 후보에 투표하는 것으로 이분처리(dichotomize)해 乙당 후보의 득표율을 계산하고, 그 득표율의 오차를 구해 각 행들을 비교해 보자.

---

17) 물론 면접으로도 온라인으로도 접촉이 되지 않는 오른쪽 하단의 인구는 실제 표집범위에 포함되지 않겠지만 그 비중을 최소화시킨다는 전제하에 (가)행에서 보는 것과 같은 이상적인 준거표본을 상정해 볼 수 있다.



〈그림 5〉 표집틀에 따른 관찰기준변수의 분포(%)

〈그림 5〉의 오른쪽 열(列)에 배치된 그림들은 각 표집에서 얻은 乙당 후보 득표율 추정치의 오차를 선거여론조사에서 통상적으로 사용하는 표본오차인 ±2.5%와 ±5%를 사용해 여섯 개의 구간으로 나누어 보여준다. 이 구간들의 중간에 위치한 −2.5~0%와 0~2.5% 두 구간의 상대빈도가 표본이 모집단을 큰 오차 없이 대표하는 경우의 비율이라면, 왼쪽과 오른쪽 끝에 위치한 두 개의 구간은 표본오차가 5%를 넘게 과소, 또는 과대평가된 경우이다. (나)행도 그렇지만, 특히 (다)행의 표집틀을 사용할 경우 이런 잘못된 추정의 가능성은 위험한 수준에 다다른다. (나)와 (다)행의 합집합에 해당하는 (라)행의 경우 편향은 많이 줄어든다. 그러나 인터넷으로는 접촉이 가능하지만 재택방문을 통한 면접은 할 수 없었던 인구의 비중 때문에 여전히 왼쪽으로, 즉 乙당 후보 득표율의 과소평가 쪽으로 기울 수밖에 없다.[18]

여기서 한 가지 유념해야 할 것은 이 모의실험에서의 표집은 현실과는 상당한 거리가 있는 교과서적인 것이었다는 점이다. 즉, 표집범위가 정해지면 온전한 목표 모집단을 대상으로 확률에 의해 표집하는 이론적인 절차를 그대로 밟은 경우였다. 또 평균값과 득표율 등 단순추정치만을 사용한 경우였다. 사용된 변수들의 모집단 내에서의 분포가 사전에 알려진 이런 경우에는 각각의 표집에 편향이 있더라도 성향가중치 등을 사용한 보정을 통해 정확한 추정치를 얻어낼 수 있다. 하지만 표집방법이 이런 전제를 만족시키지 못할 때, 얻어진 추정치를 얼마나 신뢰해야 하고 또 어떻게 사용해야 하는지에 대해서는 다분히 비체계적이고, 자의적인 기준 이상은 찾아볼 수 없다. 따라서 앞에서 ○사 측이 보여준 선거예측 결과는 유권자의 연령 분포, 연령에 따른 투표경향, 지역구의 성격과 정당, 후보 등이 '운 좋게', '확률에 의해' 맞아 떨어졌던 것이고, 그것을 근거로 다른 경우와 다른 상황으로 확대적용하기에는 어려운 것이었다.[19]

---

18) 〈표 1〉과 〈표 6〉에서 {C}에 해당하는 이 인구는 전체모집단의 47.4%이고, (라)행의 표집대상 인구의 절반이 넘는 54.3%를 차지한다. {C}의 모집단 관찰기준변수 평균값(0.044)과 인터넷 접촉은 되지 않지만 재택인 {B}의 평균값(0.878)은 극명한 대조를 보인다.

19) 2000년 미국 대선에서 Harris Interactive社가 보여주었던 예측결과를 가지고, 확률표본보다 "더 정확한" 비확률표본의 가능성을 말하는 경우도 있다. 이론적으로는 설명하지 못하더라도 "임상결과가 좋으면 시술"할 수 있다는 논리가 설득력을 가질 수도 있지만(조성겸 · 강남준 2003), 그 경우에도 일관된 임상결과가 지속적으로 축적될 때에 생각해 볼 수 있는 설득력이다. Harris Interactive社의 경우도 이후의 선거에서는 그런 결과를 내지 못 하고 있다.

Ⅲ. 예후(豫後): 용법 · 용량

1. 사용상의 주의사항

이처럼 사회조사의 현재 상황에 대한 논의들은 한편으로는 비용을 포함한 현실논리에 대한 수긍이, 또 한편으로는 몇몇 성공사례에 대한 확대추론과 기술적 해결책의 가능성에 대한 기대가 섞인 상태에서 이루어지고 있다. 그리고 대부분은 혼합식 조사나 웹패널 같은 새로운 방법의 도입을 긍정적으로, 또는 최소한 불가피한 것으로 보고 있다. 이 글의 논지 역시 새로운 방법들의 도입과 사용을 무조건 반대하는 것은 아니다. 자료의 질을 평가함에 있어 예산이나 시간 등 조사환경의 현실적 제약들을 무시하자는 것도 아니다(de Leeuw 2005; 구혜란 · 김석호 2011; 박무익 2000). 그보다는 이 방법들의 토대를 이론적, 방법론적으로 다시 한 번 검토하고, 현재의 혼선이나 오류를 바로잡으면서 앞으로는 어떻게 되어야 할지에 대해 생각해보자는 것이다. 즉 지금의 관행에 너무 쉽게, 너무 빨리 익숙해지지 않도록, 현재를 불편하게 느끼도록 하는 데 이 글의 주된 목적이 있다.

그런 맥락에서 이 글에서 제기하고자 했던 한 가지는, 현실적 제약들에 대한 인식에서 출발한다 하더라도, '제대로 된 표집'이 반드시 비현실적인 목표만은 아니라는 것이고, 특히 혼합식 조사나 온라인 조사에 대해 지금과는 다른 의미에서 보다 능동적이고 적극적인 태도를 취해야 할 것이라는 점이다. 첫째로, 혼합식 조사와 관련해서는 이미 좋은 선례들이 나와 있다. 앞에서 다룬 ACS의 경우가 그 하나로, 조사의 전 과정에 걸쳐 자세한 내용이 공개되어 있다(Martin et al. 2007; Groves & Vitrano 2011). 또 그 필요성을 인식하고 도입은 하되, 표집에 포함되어야 할 대상들이 조사절차에서 모두 포함될 수 있는 방안을 찾기 위해 각 회원국에서 다양한 실험을 계속하고 있는 유럽사회조사(European Social Survey, ESS)의 경우도 있다(Lynn et al. 2007). 둘째로, 온라인 매체가 가지는 다양한 가능성을 활용하면서도 대표성을 포함한 표집의 원칙들과 접합시킬 수 있는 방법을 꾸준히 모색해야 할 것이다(Couper 2000; 임경은 2009). 웹패널의 사용과 관련해서는 이미 상세하고 체계적인 지침이 나와 있다. 이 중 특히 눈을 끄는 것은 미국 여론조사협회(AAPOR)의 연구백서(Baker et al. 2010)에 부록으로 실린 미국의 조사연구기관협의회(CASRO)의 '인터넷 관련 윤리강령과 기준수칙'과 유럽 여론 및 시장조사협회(ESOMAR)의 '온라인 표본조사 수요자가 물어야 할

스물여섯 가지'이다.

특히 전자에는 인터넷 조사의 패널 구성과 관련하여 반드시 챙겨야 할 최소한의 기본정보로 아래의 사항들을 열거하고, 수요자의 요청이 있을 경우 조사 회사들은 반드시 이를 제공해야 한다고 규정하고 있다.

- 패널 구성에 관한 정보(패널의 규모, 대상 모집단, 그리고 활동 중인 패널 성원의 범위 규정 등 포함)
- 패널 성원의 유치 방법
- 패널 성원의 활동 상황
- 패널 성원의 보상과 관련된 사항
- 패널 성원의 인증 절차
- 패널의 품질 유지 방안
- 패널과 조사표본의 전체적 분포(응답률, 패널 성원의 참여유형, 표집 시간대의 구성 등 포함)
- 기타 조사와 관련된 정보(이메일을 통한 참여 공지, 표본추출에 사용한 선별기준, 현장실사 일자 등 포함)

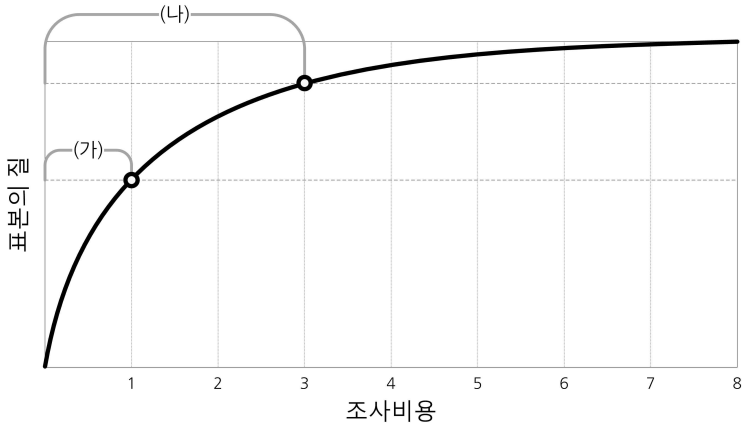
2. 제언

아무런 문제도 없는 순도 100%의 완벽한 조사나 자료가 비현실적인 희망사항일 뿐이라고 해서, '어차피 제대로는 못 한다', '원칙은 원칙일 뿐이다', '이 정도면 된다', '다들 이렇게 한다'는 식으로 곧바로 방향을 트는 것은 마치 '어차피 완벽한 무균환경은 없으니까 차라리 수채통에서 수술을 하자'는 말과 같은 논리적 오류이다(Geertz 1973:30). 이상을 너무 멀고 어렵게 보면, 현실에서 너무 쉽게 많은 것들을 포기하게 된다. 이 글에서 검토하고자 했던 것은 '최대한'이 아니라 '최소한'의 문제들이다. 자료가 만들어지는 과정에 대한 최소한의 근거는 주어져야 그것을 언제 어디서 어떻게 얼마만큼 쓸 수 있는지를 가늠할 수 있고, 문제에 당면했을 때 그 이유를 유추해 해결의 방도를 찾아 볼 수 있다(김석호 2011). 이런 기초적인 근거가 없을 때 자료의 사용자와 소비자에게 주어지는 것은, 믿지 못하면서 마지못해 받아들이거나 안타깝지만 버릴 수밖에 없는 양비(兩非)의 억지 선택이다. 사회를 폭넓고 체계적으로 관찰할 수 있게 해주는 도구로서 기초적인 연구자료와 각종 정책의 수립과 평가에 필요한 기준을 제공

한다는 사회조사의 역할에 대한 기대와 그런 역할을 수행하기 위해 필요한 '대표성', '체계성', '효율성'에 대한 신뢰를 잃으면 사회조사의 존재이유는 없어지고, 사회조사의 생태계는 붕괴된다.

확률표집과 대표성의 중요성에 대한 공감대가 있고, 조사과정의 체계성을 중시해 그것을 추구하고자 하는 의지가 있고, 또 이런 기준과 지침까지 나와 있는 상황에서 제대로 된 조사가 이루어지지 않는 현실의 제약은 도대체 무엇인가? 그 질문에 대해 항상 첫 번째로 나오는 대답은 앞서 사회조사 생태계의 먹이사슬에서 '먹이'에 해당한다고 했던 비용이다. 하지만, 무조건적인 효율과 편의만이 목표가 아니라면, 이 비용의 문제는 상대적인 것이고, 따라서 무엇을 비교의 기준으로 삼느냐에 따라 답이 달라질 수 있다는 점에 유의해야 한다.

즉 연구의 기획단계에서부터 조사의 비용과 자료의 질에 대한 기본적인 시각을 새로 정립해야 한다. 사회조사 생태계의 기본 구조를 규정하는 조사의 비용과 표본의 질 간의 상관관계는 보통 〈그림 6〉과 같은 형태로 상정된다. 이 그림에서 주목할 것은 둘 사이의 관계가 단순선형이 아니라는 점이다. '한계효용체감'의 경우처럼, 비용이 늘어나면 표본의 질이 높아지기는 하지만, 추가비용의 효용은 기하급수적으로 감소한다. 조사 기획단계에서 얼마만큼의 비용을 들여 어느 정도의 질을 확보할지를 결정하기가 어려운 이유가 바로 이것이다. 그리고 이런 상황에서 (나)보다는 (가)를 선택하는 것이 단기적으로는 합리적으로 보일 수 있다(박무익 2000).



〈그림 6〉 조사비용과 표본의 질 간의 상관관계

〈표 7〉 정부출연 연구기관의 패널사업 실시현황

개시연도	주관기관	조사명	2007년 예산 (단위: 백만 원)
1998	한국노동연구원	한국노동패널	850
2001	한국고용정보원	청년패널	779
2002	한국노동연구원	사업체패널	500
2003	한국청소년정책연구원	청소년패널	439
2004	한국직업능력개발원	한국교육고용패널	800
2004	한국교육개발원	한국교육종단연구	250
2004	한국직업능력개발원	인적자본기업패널	450
2005	한국노동연구원	고령화연구패널	769
2005	국민연금관리공단	노후보장패널	700
2006	한국고용정보원	대졸자직업이동패널	1,761
2006	한국여성정책연구원	여성가족패널	770
2006	한국보건사회연구원	복지패널	1,452
2007	한국여성정책연구원	여성인력패널	238
2007	한국장애인고용촉진공단	장애인고용패널	570
2007	육아정책개발센터	아동패널	50
2007	한국보건사회연구원	의료패널	500
2007	한국보건사회연구원	인구패널	280
합계			11,158

* 출처: 이희길(2009: 20, 59).

현재 한국의 현실은 (나)처럼 세 배의 비용을 들여 질 높은 조사 하나를 하는 것보다, (가)처럼 비용이 적게 드는 조사를 3개, 또는 3번 하는 쪽으로 기운다. 그 한 예가 최근 유행처럼 번지고 있는 정부출연 연구기관들의 패널자료 구축이다. 〈표 7〉은 2007년 현재 정부출연 연구기관에 의해 실시되고 있는 대표적인 패널조사들을 정리한 것이다(이희길 2009). 이 자료에 따르면 이들 조사에 사용된 예산은 총 112억 원에 달한다. 이 표에 포함되지 않은 다른 자료들을(예, 지자체출연 연구기관의 패널조사 등) 논외

로 하더라도, 지금 국내에서 이런 조사와 자료 구축에 투입되는 재원은 결코 작은 규모라고 할 수 없다. 현재 우리가 안고 있는 자료의 질 문제가 비용 때문에 생기는 것이라면, 그것은 비용 총액의 문제가 아니라 비용 분배의 문제인 것으로 보인다. 좋은 자료의 구축이라는 논리와는 다른 조사 외적 논리에 의해 분배가 이루어지고 있기 때문에 (나)보다 (가)를 선호하게 되는 것이다. 따라서 단발성 조사를 지양하고, 통합적 자료 구축을 통해 중첩과 남발의 문제만 해결하더라도 재원의 효용을 크게 늘릴 수 있을 것으로 보인다.

비용에 대한 근시안적 관점에서 벗어나, 장기적 관점에서 조사자료의 공공재적 성격을 고려한다면, 또 이 글의 모두에서 논의했던 사회조사의 기초적 역할을 생각한다면, 조사자료를 일회성 소비재가 아닌 중간재로, 또는 사회간접재로 볼 필요가 있고, 비용의 문제를 고려하는 방식도 그런 방향에서 다시 짜야 할 것이다. 이런 발상의 전환은 사회조사의 역할과 그에 대한 기대와 신뢰라는 큰 틀 안에서 보다 적극적인 방식의 운동으로 이어져야 한다. 현재 진행되고 있는 방법론적, 기술적 해법에 대한 다양한 연구들이 사회조사 생태계의 위기라는 '증상'에 대한 제한적 치료에 초점을 맞추고 있다면, 이제는 근본적 '병인(病因)'의 문제로도 눈을 돌려야 한다. 그리고 그 병인을 자료 자체에서 찾게 된다면, 사회조사의 생산과 소비에 관련된 모든 당사자들은 왜 (가)가 아니라 (나)를 선택해야 하는지에 대해 더 고민해야 할 것이고, 그 선택에 대해 더 고집스러워져야 할 것이다.

참고문헌

구혜란 · 김석호. 2011. "표집방법에 따른 자료의 질: 할당표집과 확률표집의 결과 비교." 2011년 후기 사회학대회 발표문.

김상욱 · 김석호 · 김소임 · 김종수 · 김지범. 2011. ≪한국종합사회조사 2010≫. 성균관대학교출판부.

김석호. 2011. "한국 사회조사자료 생산의 문제점과 양질의 자료 구축을 위한 모색." 2011년 후기 조사연구학회 발표문.

박무익. 2000. "우리나라 조사산업의 현황과 전망." ≪조사연구≫ 1(2): 161−177.

이계오. 2000. "통계조사의 현황과 전망." ≪조사연구≫ 1(1): 1−20.

이계오 · 장덕현. 2009 “인터넷 선거조사에서 성향가중모형 적용사례.” ≪조사연구≫ 10(3): 21−36.

이윤석 · 이지영 · 이경택. 2008. “온라인조사의 응답오차에 대한 연구: 설문 응답 시간과 응답 성실성의 관계.” ≪조사연구≫ 9(2): 51−83.

이흥철. 2002. “한국 조사업계의 발달사.” ≪조사연구≫ 3(2): 123−153.

이희길. 2009. “국내 패널조사의 현황분석, 2009.” ≪통계개발원 연구보고서≫.

임경은. 2009. “인터넷조사와 가구방문조사간 데이터 품질비교 −사회조사를 대상으로−.” ≪통계개발원 연구보고서≫.

조동기. 2000. “사이버공간과 사회조사: 온라인 사회조사의 쟁점과 과제.” ≪조사연구≫ 1(1): 73−108.

조성겸 · 강남준. 2003. “대안적 사회여론조사방법: 모바일 조사방법의 가능성 검토.” ≪조사연구≫ 4(1): 1−29.

조성겸 · 박아현 · 허명회. 2011. “생활시간 조사를 이용한 가중치 부여방법: 인터넷 조사에 대한 적용 가능성 검토.” ≪조사연구≫ 12(2): 145−157.

조성겸 · 주영수 · 조은희. 2005. “인터넷 여론조사의 정확도 관련요인.” ≪조사연구≫ 6(2): 51−74.

차종천. 2005. ≪사회조사의 업그레이드: 한국종합사회조사(KGSS) 출발≫. 성균관대학교출판부.

통계청. 2012. “장래가구추계: 2010년~2035년.” 통계청.

허명회 · 조성겸. 2010. “자원자 패널에 의한 인터넷 조사의 성향조정 가중화.” ≪조사연구≫ 11(2): 1−28.

Anderson, Margo and Stephen E. Fienberg. 1999. “To Sample or Not to Sample? The 2000 Census Controversy.” *Journal of Interdisciplinary History* 30(1): 1–36.

Baker, Reg, Stephen J. Blumberg, J. Michael Brick, Mick P. Couper, Melanie Courtright, J. Michael Dennis, Don Dillman, Martin R. Frankel, Philip Garland, Robert M. Groves, Courtney Kennedy, Jon Krosnick, Paul J. Lavrakas, Sunghee Lee, Michael Link, Linda Piekarski, Kumar Rao, Randall K. Thomas, and Dan Zahs. 2010. “Research Synthesis: AAPOR Report on Online Panels.” *Public Opinion Quarterly* 74(4): 711−781.

Couper, Mick. 2000. “Web Survey: A Review of Issues and Approaches.” *Public Opinion Quarterly* 64(4): 464−494.

D’Agostino, Ralph B. Jr. and Donald B. Rubin. 2000. “Estimating and Using Propensity Scores with Partially Missing Data.” *Journal of the American Statistical Association* 95(451): 749−759.

de Leeuw, Edith D. 2005. "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics* 21(2): 233−255.

de Leeuw, Edith D. and Wim de Heer. 2002. "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." pp.41−54 in Robert M. Grouves, Don Dillman, John L. Eltinge, and Roderick J.A. Little(eds.), *Survey Nonresponse*. New York: Wiley.

de Leeuw, Edith D. and Joop J. Hox. 2011. "Internet Surveys as Part of a Mixed−Mode Design." pp. 45−76 in Marcel Das, Peter Ester and Lars Kaczmirek(eds.), *Social and Behavioral Research and the Internet*. New York: Routledge.

Deming, W. Edwards. 1960. *Sample Design in Business Research*. New York: Wiley.

Dillman, Don A., Jolene D. Smyth, and Leah M. Christian. 2009. *Internet, Mail, and Mixed−Mode Surveys: The Tailored Design Method*(3rd Edition). New York: John Wiley & Sons, Inc.

Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books.

Gelman, Andrew. 2007a. "Struggles with Survey Weighting and Regression Modeling." *Statistical Science* 22(2): 153−164.

Gelman, Andrew. 2007b. "Rejoinder: Struggles with Survey Weighting and Regression Modeling." *Statistical Science* 22(2): 184−188.

Griffin, Deborah H. and Sally M. Obenski. 2002. "Meeting 21st Century Demographic Needs −Implementing the American Community Survey: Report 2. Demonstrating Survey Quality." Washington, D.C.: U.S. Census Bureau, Economics and Statistics Administration.

Groves, Robert M. and Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: Wiley.

Groves, Robert M., Floyd J. Fowler, Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology*(2nd ed.). New York: Wiley.

Groves, Robert M. and Frank Vitrano. 2011. "The Decennial Census and the ACS: Looking Back and Looking Ahead." Presentation to Population Reference Bureau, U.S. Census Bureau.

Krug, Steve. 2005. *Don't Make Me Think: A Common Sense Approach to Web Usability* (2nd ed.). New York: New Riders Press.

Lee, Sunghee. 2006. "Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys." *Journal of Official Statistics* 22(2): 329−349.

Link, Michael W., Michael P. Battaglia, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2007. “Reaching the U.S. Cell Phone Generation: Comparison of Cell Phone Survey Results with an Ongoing Landline Telephone Survey.” *Public Opinion Quarterly* 71(5): 814−839.

Lynn, Peter and Sabine Häder. 2007. “How Representative Can a Multi−Nation Survey Be?” pp. 33−52 in Roger Jowell, Caroline Roberts, Rory Fitzgerald, and Gillian Eva(eds.), *Measuring Attitudes Cross−Nationally: Lessons from the European Social Survey*. London: Sage.

Martin, Peter. 2011. “What Makes a Good Mix? Chances and Challenges of Mixed Mode Data Collection in the ESS.” Centre for Comparative Social Surveys Working Paper Series (No. 2). City University London, UK.

Martin, Peter, and Peter Lynn. 2011. “The Effects of Mixed Mode Survey Designs on Simple and Complex Analyses.” Centre for Comparative Social Surveys. Working Paper Series (No. 2011−28). City University London, UK.

Martin, Elizabeth, Jennifer Hunter Childs, Theresa DeMaio, Joan Hill, Courtney Reiser, Eleanor Gerber, Kathleen Styles, and Don Dillman. 2007. “Guidelines for Designing Questionnaires for Administration in Different Modes.” Washington, D.C.: U.S. Census Bureau.(http://www.census.gov/srd/mode−guidelines.pdf)

National Research Council. 2007. “Using the American Community Survey: Benefits and Challenges.” Panel on the Functionality and Usability of Data from the American Community Survey, Constance F. Citro, and Graham Kalton, Editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Revilla, Melanie. 2010. “Quality in Unimode and Mixed−Mode Designs: A Multitrait−Multimethod Approach.” *Survey Research Methods* 4(3): 151−164.

Roberts, Caroline. 2007. “Mixing Modes of Data Collection in Surveys: A Methodological Review.” ESRC National Centre for Research Methods Briefing Paper NCRM/008. ESRC National Centre for Research Methods.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70(1): 41−55.

Roster, Catherine A., Robert D. Rogers, Gerald Albaum, and Darin Klein. 2004. “A Comparison of Response Characteristics from Web and Telephone Surveys.” *International Journal of Market Research* 46(3): 359−373.

Rubin, Donald. 1997. “Estimating Causal Effects from Large Data Sets Using Propensity Scores.” *Annals of Internal Medicine* 127(8): 757−763 .

Smith, Tom W. 2010. “The Globalization of Survey Research.” pp. 477−484 in Harkness, Janet A., Michael Braun, Brad Edwards, Timothy P. Johnson, Lars E. Lyberg, Peter Ph. Mohler, Beth−Ellen Pennell, and Tom W. Smith(eds.), *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*. New York: Wiley.

Sparrow, Nick and John Curtice. 2004. “Measuring the Attitudes of the General Public via Internet Polls: An Evaluation.” *International Journal of Market Research* 46(1): 23−44.

Taylor, Humphrey. 2000. “Does Internet Research Work? Comparing Electronic Survey Results with Telephone Survey.” *International Journal of Market Research* 42(1): 51 - 63.

Valliant, Richard and Jill A. Dever. 2011. “Estimating Propensity Adjustments for Volunteer Web Surveys.” *Sociological Methods & Research* 40(1): 105−137.

Vannieuwenhuyze, Jorre, Geert Loosveldt and Geert Molenberghs 2010. “A Method for Evaluating Mode Effects in Mixed−Mode Surveys.” *Public Opinion Quarterly* 74(5): 1027−1045.

<접수 2012/7/18; 수정 2012/8/19; 게재확정 2012/8/21>