

연구논문

서베이 모드와 사회적 바람직함 편향:

민감한 주제에 대한 전화 면접, 대면 면접, 온라인 서베이 응답 비교\*

The Effects of Survey Mode on Social Desirability Bias: A Comparison of Telephone, Face-to-Face and Online Survey Responses.

송인덕<sup>a)</sup> · 조성겸<sup>b)</sup>

Indeok Song · Sungkyum Cho

본 연구는 사회조사에서 가장 일반적으로 활용되고 있는 서베이 방법의 정확성과 관련하여, 민감한 서베이 주제로 인해 발생하는 응답 상의 사회적 바람직함 편향이 전화 면접, 대면 면접, 그리고 온라인 서베이 모드에 따라 어떻게 달리 나타나는지를 실증적으로 검토하기 위해 수행되었다. 이론적 논의 및 선행연구 결과 검토를 바탕으로, 민감한 질문을 포함한 서베이의 경우, 사회적 바람직함 편향은 전화 면접에서 가장 크게 작용하고 다음으로 대면 면접과 온라인 서베이 순일 것이라는 가설을 제시하였으며, 이를 실험방법을 통해 수집된 자료를 분석함으로써 검증하였다. 연구결과, 본 연구에서 제시한 여러 민감한 주제들 중 게임 중독증과 본인 외모 불만족 정도 측정에서 서베이 모드에 따라 응답의 사회적 바람직함 편향이 달리 나타나는 것이 확인되었다. 또한 서베이 질문의 민감성이 응답자의 잠재적 응답이 무엇인가에 따라 달라지며, 응답자들마다 상이하게 인식된 질문의 민감성에 따라 응답의 왜곡 정도, 즉 사회적 바람직함 편향의 정도가 달라지는 경향이 확인되었다. 이와 같은 연구결과를 바탕으로, 민감한 질문을 포함하는 서베이 수행을 위한 방법론적 제언을 제시하였다.

주제어: 서베이 모드, 모드효과, 서베이 편향

\* 본 연구는 2010년도 정부재원(교육과학기술부 한국사회기반연구사업비)으로 한국연구재단의 지원을 받아 수행되었습니다(NRF-2010-330-B00280).

a) 중부대학교 신문방송학과 교수

b) 교신저자(corresponding author): 충남대학교 언론정보학과 교수 조성겸.  
E-mail: skcho99@gmail.com

The purpose of this study is, with regard to the validity of survey method that have been generally utilized for social research, to empirically examine the effects of various survey modes including telephone interview, face-to-face interview, and online survey on the so-called social desirability bias in survey responses, which is caused by sensitive survey questions. Based upon theoretical discussions and reviews of the previous research, a research hypothesis predicting the degree of social desirability bias will be highest in telephone interview mode and then face-to-face interview and online survey modes in order was proposed. This hypothesis was tested by analysing a set of experiment data. The results show that the degrees of social desirability biases, which were caused by the sensitive questions about respondent's game addiction symptoms as well as dissatisfaction with one's physical appearance, are vary across the three survey modes. Also, it was found that the sensitivity of survey questions may vary depending on what the respondent's potential responses are and in turn, depending on the perceived sensitivity of the questions, the degree of social desirability bias may be determined. Upon consideration of the implications of these results, a methodological suggestion for conducting surveys which include sensitive questions is suggested

**Key words:** survey mode, social desirability bias, survey response

## I. 연구 배경 및 목적

사회조사를 위한 대표적 방법인 서베이(survey)는 사람들이 무엇을 생각하고 어떻게 행동하는가를 탐색, 기술, 설명하기 위해 널리 사용되고 있는 과학적 방법이다. 서베이는 기본적으로 다수의 응답자들로부터 신뢰성 및 타당성 있는 정보를 수집하기 위해 표준화된 질문을 사용한다. 그러나 서베이는 질문 또는 문항 제시 및 응답 방식에 따라 면접원, 우편(설문지), 전화, 온라인(웹) 서베이 등 다양한 모드(mode)로 수행될 수 있다. 최근에는 휴대전화를 비롯한 개인용 모바일 기기의 대중적 확산 및 급속한 기술의 발달을 바탕으로 다양한 유형의 모바일 기기를 활용한 새로운 모드의 서베이가

도 시도되고 있는 상황이다.

서베이는 대규모의 데이터를 비교적 적은 경제적 및 시간적 비용으로 수집하는 것이 가능하며 폭넓은 주제에 관한 정보를 현실적 환경에서 수집할 수 있다는 장점(Babbie 2001; Kays, Gathercoal, & Buhrow 2012; Wimmer & Dominick 2000)을 바탕으로 대규모 사회조사에서 보편적으로 활용되어 왔다. 그러나 사회 변화의 흐름 및 정치·경제·사회·문화 등 다양한 영역에서의 여론 동향을 정확히 진단하고 분석하기 위한 과학적 데이터 제공 및 정책적 반영 목적으로 수행됨에도 불구하고, 사회조사(특히 선거 관련 여론조사)의 예측이 번번이 빗나가는 등의 문제로 인해 서베이 결과의 정확성에 대한 불신 및 전통적 서베이 방법의 한계에 대한 논의가 지속되고 있는 실정이다.

서베이 결과의 타당성은 적절한 방법으로 표집된 대표성 있는 응답자가 자신의 심리적 상태 및 행동 경험을 정확히 인지하고 있으며, 이를 왜곡 없이 정직하게 응답한다는 것을 전제로 한다. 그러나 아무리 정교한 표집을 실시하여 표본의 대표성을 확보함으로써 표본 오차(sampling error)를 줄이더라도 응답자로부터 발생하는 응답 오차(response error)는 연구자가 직접적으로 통제하기 매우 어렵다. 응답 오차란 서베이 응답자가 정확하지 않거나 왜곡된 정보를 제공함으로써 발생하는 오차로서(박용치, 2000; 이윤석·이지영·이경택 2008), 응답자들이 질문을 잘못 이해하거나, 주제와 관련한 정보를 제대로 인출할 수 없거나, 정확하지 못한 추정 및 판단 전략을 사용하거나, 응답을 과장(overreport) 또는 축소(underreport)하거나, 또는 적절한 응답 항목을 찾지 못하는 경우 발생할 수 있다(Krosnick 1991; Tourangeau, Rips, & Rasinski 2000). 따라서 서베이를 디자인하는 과정에서부터 연구자는 표본 오차뿐만 아니라 응답 오차를 최소화하여 데이터의 품질을 높이기 위한 세심한 주의와 노력을 기울여야 한다.

응답 오차의 문제는 서베이가 도박, 술, 마약, 낙태, 심지어 투표행위를 포함하는(Belli, Traugott, & Beckmann 2001) 민감한(sensitive) 주제에 관한 것일 때 더욱 심각해질 수 있다. 이는 앞서 언급한 응답 오차 발생의 일반적 원인들에 더하여 민감한 주제에 대해서 응답자들은 순순히 진실한 정보를 제공하기를 원하지 않기 때문이다(Tourangeau & Yan 2007). 따라서 이 경우 제기되는 중요한 문제는 민감한 질문에 대해 과연 응답자들이 얼마나 진솔한 응답을 하는가의 여부이다. 이와 관련하여 민감한 주제를 다루는 서베이에서의 응답의 정확성에 관한 방법론적 연구들은 서베이 결과로

부터 유추된 추정치 내의 오차의 주요 원인이 응답자로부터의 오응답(misreporting)이라고 주장한다(Belli et al. 2001; Tourangeau & Yan 2007). 오응답이란 사실과 다른 진실하지 못한 응답으로서 의도적으로 정보를 왜곡하거나 실제보다 과장 또는 축소하여 정보를 제공하는 것을 의미한다. 예를 들면, 도박, 술, 마약 관련 경험이 실제로는 있으나 전혀 없다고 응답한다든지 또는 이용량 및 빈도 등을 실제보다 줄여 응답하는 경우가 오응답의 전형적인 사례이다. 한 예로, 우리나라 고등학생들(인문계 남학생 306명, 실업계 여학생 325명)의 흡연 여부를 설문지 및 소변검사를 동시에 적용하여 측정한 연구결과(심수정 등 2003), 설문지에 따른 흡연률은 남자 고등학생 1,2,3학년 각각 14.9%, 20.4%, 22.5%인 반면, 소변검사 결과에 따른 흡연률은 각각 26.6%, 38.0%, 29.6%인 것으로 나타났다. 실업계 여고생 흡연률은 2,3학년이 각각 22.6%와 23.9%인 것으로 설문지를 통해 측정되었으나 소변검사에서는 이보다 훨씬 높은 46.6%와 35.4%가 흡연을 하고 있는 것으로 나타났다. 흡연 여부에 관한 한 자가응답(self-report) 보다 소변검사가 더욱 정확한 측정방법임을 고려하면, 응답자들의 상당수가 오응답을 제공한 것으로 추정된다.

투표행위의 경우에도 실제로는 투표를 하지 않았음에도 불구하고 투표에 참여한 것으로 응답하거나, 투표한(또는 지지하는) 정치후보자를 실제와 다른 전혀 엉뚱한 사람으로 보고하는 등의 오응답이 발생할 수 있다. Belli et al.(2001)은 미국의 전국선거조사(American National Election Studies) 결과와 서베이 응답을 비교하여 실제로는 투표를 하지 않은 사람들 중 20% 이상이 투표를 한 것으로 오응답했다는 사실을 발견하였다. 이를 근거로 연구자들은 미국인들의 투표 여부에 대한 서베이의 추정치에 오차가 발생하는 주요 원인은 표집에서 발생하는 표준 오차(standard error)라기보다는 응답자들의 오응답에서 비롯된 응답 오차라고 결론 내렸다.

그 동안 연구자들은 서베이 결과의 정확성과 관련하여 면접, 전화, 우편, 설문지, 웹 등 다양한 종류의 서베이 모드들의 상대적 효용성을 검토해 왔으나, 일관성 있는 결과가 나타나지 않은 이유로 여전히 논란이 계속되고 있는 상황이다. 이에 서베이 모드를 선택하는 과정에서 연구하고자 하는 주제와의 적합성을 고려하기보다는 단순히 시간적·경제적 비용 효율성만을 고려하는 바람직하지 못한 경향이 있다. 심지어는 동일한 설문을 복수의 서베이 모드(예를 들면, 전화와 면접 조사를 병행하는 경우 등)를 통해 진행하여 데이터를 취합한 후 모집단에 대한 추정치를 산출하는 경우도 있다. 이는 서베이 모드에 따라 응답자들의 응답에 체계적인 편향이 발생할 수 있음을 간과하

고 있는 것이다. 이처럼 서베이 모드와 서베이의 주제, 특히 민감한 주제와의 상호작용 효과에 대한 이해는 여전히 충분하지 못한 상황이다. 이에 본 연구는 서베이에서 응답자 개인의 프라이버시와 관련된 다소 민감한 주제에 대한 설문 응답에 가장 보편적으로 활용되고 있는 서베이 모드들(대면 면접, 전화 면접, 온라인 서베이)이 미치는 영향을 실증적으로 탐색해 보고자 한다.

## II. 선행연구 검토 및 이론적 논의

서베이 질문에 진실하게 답변하지 않는 불량 응답이 많으면 응답 오차가 증가하고 데이터의 품질이 떨어지게 된다는 것은 앞서 언급한 바 있다. 응답 오차의 주요 원인과 관련하여 크로스닉(Krosnick 1991)은 설문에 성실하고 정확한 응답을 제공하는 것은 인지적으로 많은 노력을 요구하기 때문에 응답자들은 ‘희생만족적(satisficing)’ 응답을 하는 경향이 있다고 주장한다. ‘희생만족적’이란 용어는 ‘만족시키다’라는 뜻의 satisfy와 희생이란 의미의 sacrificing을 합성한 것으로, 수고스러운 인지적 과정을 희생시키고 비교적 만족할 만한 응답을 제공하는 경향을 개념화한 것이다. 다시 말해서, 희생만족적 응답이란 질문에 응답하기 위해 필요한 정보 인출 및 정보 통합을 제대로 하지 않거나 전혀 하지 않고 응답하는 행위를 의미한다.

그러나 응답 오차에 대한 이러한 관점은 답변을 의도적으로 왜곡하여 보고하는 행위인 오응답에 대해서는 충분한 설명을 제공해 주지 못하는 것으로 판단된다. 왜냐하면 고의적으로 허위의 사실을 제공하거나 진실을 왜곡하기 위해서는 우선 진실한 정보를 인출하고 통합하는 데 필요한 기본적인 인지적 부담뿐만 아니라 추가적인 인지적 노력이 필요하기 때문이다. 따라서 오응답의 경우는 ‘희생만족’ 개념으로는 충분한 설명이 어렵다. 이러한 문제인식을 바탕으로 본 장에서는 오응답의 주요 원인을 서베이 주제 또는 질문의 민감성에 따른 ‘사회적 바람직함 편향(social desirability bias)’에 있는 것으로 보고, 이러한 요인이 서베이 모드와 상호작용하여 응답의 정확성에 미치는 영향을 논의하고자 한다.

### 1. 민감한 질문이 응답에 미치는 영향

민감한 질문(sensitive question)이란 사회적으로 바람직한 것에 대한 염려뿐 아니

라 응답자들의 입장에서 꺼려지거나 정보 노출에 따른 파급효과에 대한 염려를 유발하는 질문을 포함하는 광범위한 개념이다(Tourangeau & Yan 2007). Tourangeau et al.(2000)은 민감성(sensitivity) 개념을 세 가지 의미로 정의하고 있다. 첫 번째 의미는 주로 개인의 사생활을 침해하는 것으로서 응답자가 응답을 꺼리게끔 하는 질문들이다. 예를 들면, 응답자의 소득, 종교, 학력, 나이 등을 묻는 질문들이 이 부류에 해당될 수 있으며, 이러한 질문들은 응답의 내용과 관계없이 응답자들을 불쾌하게 만들 수 있다.

민감한 질문의 두 번째 의미는 폭로의 위협(threat of disclosure), 즉 진솔한 답변이 제3자에게 알려질 가능성에 대한 우려를 포함한다. 응답자들의 응답이 서베이와 직접적으로 관련이 없는 사람들에게 공개되거나 그럴 가능성에 대한 두려움을 유발하는 질문이라면 이 부류에 해당하는 민감한 질문이라고 할 수 있다. 예를 들어, 서론에서 잠시 언급한 청소년들의 흡연에 관한 질문은 부모 또는 선생님들이 그들의 응답을 알 수 있는 경우라면 민감한 것이 된다. 비록 서베이 연구들은 거의 예외 없이 응답에 대한 비밀보장(confidentiality)을 약속하지만, 응답자들은 이러한 약속을 있는 그대로 받아들이지 않는다. 따라서 정보 노출에 대한 응답자들의 우려는 불법적이거나 사회적으로 적절치 못한 행동 및 태도에 대한 오응답을 야기하는 중요한 요소라고 할 수 있다(Singer & Presser 2007).

민감한 질문의 세 번째 의미는 사회적으로 부적절하거나 바람직하지 못한 응답을 이끌어내는 정도와 관련이 있다(Tourangeau et al. 2000). 즉 질문이 응답자로 하여금 사회적으로 바람직하지 못한 응답을 요구함으로써 응답자가 스스로 사회적 규범을 위반했음을 인정하도록 하는 경우 민감한 질문이 된다는 것이다. 민감성에 대한 이러한 개념은 특정 행동이나 태도와 관련하여 명확한 사회적 규범이 있음을 전제로 하며, 질문의 민감성은 대개 질문에 대한 잠재적 응답에 따라 결정된다. 예를 들어, 투표 참여 여부에 관한 질문은 투표에 실제로 참여한 응답자에게는 전혀 민감하지 않지만 투표에 불참한 응답자에게는 민감한 것일 수 있다.

선행연구들은 민감한 서베이 질문으로부터 초래되는 문제점을 세 가지로 정리하고 있다. 첫째, 민감한 질문을 포함한 서베이는 응답률(response rate)이 대체로 낮은 것으로 알려져 있다(Catania, Gibson, Chitwood, & Coates 1990). Cook, Heath, & Thompson et al.(2000)은 웹 서베이에서의 응답률을 검토한 결과, 서베이의 주제 민감성이 응답률과 부적적으로 관련되어 있음을 보여주었다. 또한 민감한 질문은 항목별

무응답(item nonresponse) 수준을 높이는 부정적 효과를 야기하는 것으로 추정된다 (Tourangeau & Yan 2007). 그러나 이를 실증적으로 검토한 선행연구는 아직까지 없는 실정이다. 다만 Singer, von Thurn, & Miller(1995)은 메타 분석을 통해 비밀보장 약속이 민감한 질문을 포함한 서베이의 응답률 및 질문 항목별 응답률을 높여 준다는 것을 보여주었다. 그리고 소득에 대한 질문에 대해서는 대체적으로 항목별 무응답률이 높다는 것이 선행연구(Juster & Smith 1997; Moore, Stinson, & Welniak 1999)를 통해 일관되게 관찰되어 왔다.

민감한 질문에 나타날 수 있는 가장 심각한 문제점은 응답자의 사회적 바람직함 편향(social desirability bias)을 활성화시킴으로써 응답자가 체계적 오응답(systematic misreporting)을 제공하도록 한다는 것이다. 여기서 사회적 바람직함 편향이란 응답자들이 사회적으로 바람직한 행동이나 태도는 과대응답(overreport)하는 반면 사회적으로 바람직하지 못한 특정 행동이나 태도에 대해서는 과소응답(underreport)하는 경향을 말한다(Tourangeau & Yan 2007). 선행연구들을 통해 과소응답을 유도하는 것으로 알려진 서베이 주제에는 불법 약물(Johnston & O'Malley 1997), 음주(Lemmens, Tan, & Knibbe 1992), 흡연(심수정 등 2003; Patrick et al. 1994), 낙태(Jones & Forrest 1992), 범죄행위(Wyne, 1980), 인종차별적 태도(Krysan 1998) 등이 포함되며, 과대응답을 유도하는 주제로는 투표(Belli et al. 2001), 에너지 절약(Fujii, Hennessy, & Mak 1985) 및 안전벨트 사용(Stulginskis, Verreault, & Pless 1985) 등의 캠페인 참여, 종교행사 참석(Presser & Stinson 1998), 운동(Tourangeau, Smith, & Rasinski 1997) 등이 있다. 과소응답이나 과대응답에 관계없이 이러한 서베이 응답상의 체계적인 편향은 결국 서베이 결과의 품질을 저하시키고 추정치의 정확성에 부정적인 영향을 초래하게 된다.

그러나 Tourangeau et al.(2000)은 특정한 연구 설계 전략을 통해 민감한 질문의 부정적 효과, 즉 사회적 바람직함 편향을 축소시키고 응답의 질을 높일 수도 있다고 주장한다. 이는 비록 실제 서베이에 적용 가능한 구체적 실체로서 '특정한 연구 설계'를 제시하지는 않았으나, 연구자가 서베이 디자인 단계부터 조사과정 전반에 걸쳐 사회적 바람직함 편향을 줄일 수 있도록 세심한 주의를 기울일 필요성이 있음을 제안한 것이다. 따라서 본 연구는 연구자가 자료 수집을 위한 구체적 방법으로서 서베이 모드를 선택하는 과정에서 연구의 주제 및 질문의 민감성에 따른 사회적 바람직함 편향의 정도를 고려하여 가장 적합한 서베이 모드를 신중히 선택하고 활용해야 함을 제시하고자

한다. 이와 관련하여 다음 절에서는 민감한 주제에 대한 응답에 서베이 모드가 미치는 영향을 검토한다.

## 2. 서베이 모드가 응답에 미치는 영향

서베이 과정에서 오응답이 나타나는 것도 문제지만, 특히 그것이 서베이 모드에 따라 정도가 달리 나타나는지 여부도 중요하다. 오응답의 발생은 주로 면접자 및 제3자의 존재 여부가 중요한 영향을 미치는 것으로 나타나고 있다. 민감한 주제를 다루는 서베이에서의 응답 오차에 관한 선행연구들을 종합적으로 검토한 Tourangeau & Yan(2007)은 민감한 질문에 대한 고의적 오응답이 매우 일반적이며 대개는 상황의존적인 것으로 파악한다. 즉 오응답의 정도는 응답자가 응답하기 곤란한 어떤 것이 있는지의 여부와 서베이 디자인의 세부적 특성에 따라 달라진다는 것이다. 또한 민감한 주제에 대한 오응답은 일정 정도 동기화된 과정으로서 응답자들은 면접원에게 창피당하는 것을 피하거나(Holtgraves, Eck, & Lasky 1997; Tourangeau et al. 2000) 제3자들로부터의 간접적인 영향을 회피하기 위해(Aquilino, Wright, & Supple 2000; Rosenfield, Booth-Kewly, & Edwards 1996) 응답을 편집하는 것이라고 주장한다.

Rasinski et al.(1994)은 가상의 서베이 상황을 묘사하고 있는 삽화들을 이용하여 일련의 실험을 수행하였다. 삽화들은 자기기입식 대 면접원 진행방식 그리고 가족 등의 제3자가 존재하거나 존재하지 않는 상황을 보여주었다. 실험참가자들은 삽화 속의 응답자들이 진실을 얘기할지를 평가하였다. 이들 실험의 결과는 응답자들이 진실한 응답을 제공할지를 결정하는 과정에 있어 면접원 및 제3자에 의한 정보의 노출 및 폭로의 위험에 민감하다는 사실을 보여주었다. 이러한 선행연구 결과를 바탕으로 본 연구에서는 민감한 서베이 질문에 대해 응답자가 응답하기 곤란한 어떤 것이 있다는 것을 전제로, 면접원 및 제3자 존재 여부와 관련한 서베이 모드(면접원, 전화, 온라인 서베이)의 효과를 검토하고자 한다.

### 1) 면접원의 존재 여부

서베이는 응답자들로부터 데이터를 수집하기 위해 대면 면접(face-to-face interview), 전화 면접(telephone interview), 우편 설문(mail survey), 컴퓨터 활용 개인 면접(CAPI: Computer-Assisted Personal Interview), 웹 설문(web survey), 설문지(paper-and-pencil questionnaire) 등의 다양한 방법을 활용한다. 비록 이들 다양한

서베이 모드들은 여러 차원에서 서로 상이하지만, 핵심적인 차이점들 중 하나는 설문을 면접원이 진행하는가(interviewer administration) 그렇지 않으면 응답자 본인이 자기기입 방식(self administration)으로 진행하는가의 여부이다.

선행연구들은 면접원 진행 방식에 비해 자기기입 방식이 사회적으로 바람직하지 않은 행동에 대한 솔직한 응답을 증가시킨다는 결과를 보고하고 있다. 다시 말해, 면접원 진행 방식보다 자기기입 방식의 서베이에서 응답자들이 민감한 정보를 더욱 잘 보고하는 경향이 있다는 것이다. 예를 들어, 마리화나 이용에 대한 전화 서베이를 수행하는 과정에서 면접원이 직접 설문을 진행하는 방식과 자기기입식 음성응답방식(IVR: Interactive Voice Response)을 비교한 결과, IVR 방식으로 진행하였을 때 마리화나 이용률이 58% 더 높은 것으로 나타났다(Corkrey & Parkinson 2002). 이와 유사한 대부분의 선행연구들은 많은 수의 응답자들이 면접원 진행 방식에서 보다는 자기기입 방식에서 불법 약물 사용을 보고한다는 것을 보여주고 있다(Aquilino 1994; Gfroerer & Hughes 1992; Schober, Caces, Pergamit, & Branden 1992; Tourangeau & Smith 1996; Turner, Lessler, & Devore 1992). 자기기입 방식은 불법 약물 사용과 마찬가지로 서베이에서 과소응답되는 것으로 알려진 십대 흡연(Brittingham, Tourangeau, & Kay 1998; Currivan, Nyman, Turner, & Biener 2004; Moskowitz 2004) 및 낙태(Lessler & O'Reilly 1997)와 같은 사회적으로 바람직하지 못한 행동들에 대한 응답을 증가시키는 반면, 종교 행사 참석(Presser & Stinson 1998)과 같이 사회적으로 바람직한 행동들에 대한 과대응답을 감소시키는 것으로 보고되고 있다. 또한 자기기입 방식은 우울증이나 불안과 같은 심리적 이상 증세에 대한 응답자의 응답도 증가시키는 것으로 알려져 있다(Epstein, Barker, & Kroutil 2001; Newman et al. 2002).

여기서 중요한 점은 모드 간 차이를 만드는 것은 면접원의 존재 여부가 아니라 응답자들이 다른 사람, 즉 면접원에게 본인들의 응답을 제공하는지의 여부이다. 즉 면접원이 물리적으로 같은 공간 또는 전화 유선 상에 존재하지만 응답자가 응답한 내용을 알지 못하는 경우에는 면접원의 존재 자체가 응답에 그리 영향을 미치지 않는다는 것이다. 예를 들어, Turner et al.(1992)은 마약 이용에 대한 서베이를 수행하는 과정에서 면접원이 질문을 크게 읽되 응답자들은 별도의 응답지에 응답을 표시하는 방법을 사용했다. 면접이 끝나면 응답자는 응답지를 본인이 직접 봉투에 넣고 봉한다. 이러한 방법은 면접원이 응답자가 질문에 어떤 답변을 했는지 전혀 알지 못하게 됨으로써 기존의 자기기입식 설문지 방법만큼 진솔한 응답을 얻는 데 효과적인 동시에 서베이 과

정도 면접원을 통해 효율적으로 진행될 수 있다는 것을 보여준다. 요약하면, 자기기입식 서베이 모드의 핵심은 면접원이 존재하지 않는다는 것이 아니라 응답자의 응답내용이 면접원에게 노출되지 않는다는 것이다.

이러한 선행연구 결과를 바탕으로 본 연구는, 서베이 주제가 민감한 개인의 행동 및 태도를 드러내도록 요구하는 경우, 자기기입식 서베이 모드가 면접원 진행 모드에 비해 사회적 바람직함 편향을 더욱 줄여줄 것으로 예측한다. 보다 구체적으로, 본 연구에서 비교하고자 하는 세 가지 서베이 모드의 경우, 자기기입 방식의 온라인 서베이가 면접원 진행 방식으로 진행되는 대면 및 전화 면접 서베이에 비해 민감한 질문에 대한 응답자들의 사회적 바람직함 편향을 최소화하는 데 더욱 효과적일 것이라는 것이다. 다음 절에서는 전화 면접과 대면 면접 모드 간의 효과 차이를 예측하기 위한 선행연구 검토 및 논의를 진행할 것이다.

## 2) 제3자 존재 여부

서베이가 진행되는 동안 면접원 및 응답자를 제외한 누구도 개입되어서는 안 된다는 것이 서베이 수행의 기본 원칙이다. 그러나 서베이는 종종 응답자의 프라이버시가 완벽히 보장되지 않는 상태에서 진행된다. 다시 말해서, 많은 경우 서베이는 응답자와 면접원 이외에 제3자가 있는 상황에서 진행된다. 문제는 제3자, 특히 부모(Aquilino et al. 2000; Brener et al. 2006; Tourangeau & Yan 2007) 또는 직속 상관(Rosenfield et al. 1996)과 같은 권위 있는 인물의 존재 자체가 민감한 서베이 질문에 대한 응답에 영향을 미치는 것으로 선행연구들이 보고하고 있다는 것이다. 예를 들어, Aquilino et al.(2000)은 부모들의 존재가 알콜 및 마리화나 사용에 대한 응답을 축소시킨다는 것을 발견했다. 이러한 결과를 바탕으로 연구자들은 제3자의 존재가 서베이 응답에 미치는 영향은 응답자가 그들의 응답내용이 제3자에게 알려짐으로써 발생할 수 있는 처벌 등의 파급효과를 염려하는 정도에 따라 달라지는 경향이 있다고 주장한다. 이와 유사하게, 서베이 환경(집 또는 학교)이 청소년들의 위험 행동 응답에 미치는 영향을 실험을 통해 비교한 Brener et al.(2006)의 연구도, 집에서 진행되는 서베이의 경우에는 비록 청소년들이 자기기입 방식으로 응답하더라도 부모의 존재 자체가 위험 행동들에 대한 과소응답을 증가시킨다는 것을 확인하였다.

이러한 제3자 존재의 효과를 서베이 모드와 관련하여 추론해 보면, 앞서 논의한 면접원의 영향과 마찬가지로, 자기기입식 방식이 면접원 진행 방식에 비해 사회적 바람직함 편향에 따른 과소 또는 과대 응답을 감소시키는 데 보다 효율적일 것으로 예상된

다. 이는 자기기입식 방식이 면접원 진행 방식에 비해 제3자 존재를 응답자 본인이 통제할 수 있는 여지를 제공하기 때문이다. 즉 자기기입 방식을 통해 응답자들은 서베이에 참여하는 시간 및 장소 등을 본인의 의지에 따라 선택하거나 응답내용이 제3자에게 누출되지 않도록 자구책을 강구함으로써 프라이버시 수준을 조절할 수 있다. 반면 면접원이 서베이를 진행하는 경우에는 응답자가 자의로 서베이 환경의 프라이버시 수준을 조절할 수 있는 기회가 상대적으로 제약된다. 따라서 제3자 존재 효과와 관련해서도, 본 연구에서 비교하고자 하는 세 가지 서베이 모드 중 자기기입 방식의 온라인 서베이가 면접원 진행 방식인 대면 및 전화 면접 서베이에 비해 민감한 질문에 대한 솔직한 응답을 얻는 데 보다 효율적일 것으로 예측할 수 있다.

그렇다면 면접원 진행 방식에 해당되는 대면 및 전화 면접 서베이에서의 제3자 효과는 어떠한가? 아쉽게도 이 문제와 직접적으로 관련된 선행연구들은 부재한 상황이다. 다만 일부 선행연구들은 민감한 정보를 이끌어내는 데 전화 면접이 대면 면접보다 효율적이라고 주장하는(예: Henson, Cannell, & Roth 1978; Sykes & Collins 1988) 반면, 상반되는 결과를 발견한 연구들도 있다(예: Aquilino 1994; Johnson, Hougland, & Clayton 1989). 또한 전화 면접과 대면 면접이 민감한 질문에 대한 응답에 미치는 효과 차이를 발견하지 못한 연구들도 있다(예: Mangione, Hingson, & Barrett 1982). 그러나 대체적으로는 대면 면접이 전화 면접에 비해 민감한 주제에 대한 솔직한 응답을 유도하는 데 상대적으로 더욱 효율적이라는 증거가 우세하다(de Leeuw & van der Zouwen 1988). 즉 사회적 바람직함 편향은 대면 면접보다 전화 면접에서 더욱 심하게 나타난다는 것이다(Holbrook, Green, & Krosnick 2003). 국내에서 실시된 대규모 도박 연구들을 메타분석한 허태균·김현철·박정열(2009)도 자료수집 방법에 따라 도박 중독 유병률이 2배 이상 차이가 난다는 사실을 발견하였는데, 전화 면접(0.7%)이 대면 면접(1.9%)에 비해 도박 중독 유병률을 축소하는 것으로 나타났다. 이러한 결과를 바탕으로 연구자들은 전 국민을 대상으로 한 대규모 도박 중독 유병률 조사의 경우, 대면 면접 방식을 적용할 것을 권고하고 있다.

비록 선행연구들이 이러한 결과들을 제3자 존재의 효과와 관련하여 설명하고 있지는 않지만, 일반적으로 서베이 상황에서 면접원과 응답자 이외의 제3자의 존재는 정확하고 솔직한 응답을 획득하는 데 있어 부정적인 영향을 미친다. 따라서 전화 면접이든 대면 면접이든 제3자가 존재한다면 면접 장소 또는 시간을 변경하는 등의 방법을 통해 제3자 존재의 영향을 없애거나 가급적 최소화하는 것이 바람직할 것이다. 그런데 이와

같이 면접 환경을 선택함으로써 제3자 존재의 부정적 효과를 통제할 가능성은 전화 면접보다는 면접원과의 대면 면접 상황에서 일반적으로 더 높을 것으로 기대된다. 전화 면접의 경우, 면접원이 응답자 주위에 제3자가 존재하는지 여부를 파악할 수 없을 뿐만 아니라 응답자 스스로가 전화를 받는 곳 주위에 있는 제3자를 피하여 다른 장소에서 전화 면접을 계속하는 것은 쉽지 않다. 이에 비해 대면 면접의 경우에는 면접원이 제3자 존재 여부를 확인하고 응답자의 도움을 얻어 면접 장소를 옮기거나 또는 다른 시간대에 재방문하여 면접을 진행하는 것이 가능하다. 이러한 취지에서 대면 면접이 전화 면접에 비해 제3자 통제 가능성이 높다고 할 수 있으며, 면접 상황에서의 제3자 존재의 부정적 효과는 전화 면접에서 상대적으로 크게 나타날 것으로 예측할 수 있다.

민감한 질문에 대한 응답에 미치는 세 종류의 서베이 모드의 효과와 관련하여 지금까지 논의한 내용을 종합하면, 우선 자기기입 방식으로 진행되는 온라인 서베이가 민감한 질문에 대한 솔직한 응답을 증가시키는 반면 면접원 진행 방식의 대면 및 전화 서베이는 사회적으로 바람직한 방향으로 오응답을 증가시킬 것이다. 그리고 면접원 진행 방식의 서베이의 경우, 상대적으로 제3자의 존재를 통제하기 더욱 용이한 대면 면접이 전화 면접 모드에 비해 민감한 질문에 대한 솔직한 응답을 더욱 많이 이끌어 낼 것이다. 따라서 민감한 서베이 질문에 대한 응답 과정에서 사회적 바람직함 편향을 최소화하고 응답자로부터 보다 진솔한 정보를 수집하는 데는 온라인, 대면 면접, 전화 면접 서베이 순으로 효율적일 것이다. 이상의 논의를 바탕으로 본 연구는 다음과 같은 연구가설을 제시하고 이를 검증하고자 한다.

연구가설: 민감한 질문에 대한 서베이 응답에서 사회적 바람직함 편향은 전화 면접에서 가장 많이 작용할 것이고, 다음으로 대면 면접, 그리고 온라인 서베이 순으로 나타날 것이다.

### Ⅲ. 연구방법

#### 1. 실험 설계 및 실험참가자

민감한 질문에 대한 서베이 응답에 상이한 서베이 모드들이 미치는 영향을 실증적으로 비교·분석하기 위해, 본 연구는 세 종류의 서베이 모드들(대면 면접, 전화 면접, 온라인 서베이)을 기준으로 세 집단 간 반복측정(사전-사후) 실험(3-group between-

subject repeated measures experiment)을 하였다. 집단 간 동질성 검증을 목적으로 실시된 사전조사에는 총남에 소재한 한 대학교에 재학 중인 93명의 학부생들이 참가했으나, 이들 중 사후조사까지 완료한 66명의 응답만을 최종 분석하였다. 총66명 가운데 남자는 38명(57.6%), 여자는 28명(42.4%)이었고, 평균 연령은 만20.61세( $SD=1.97$ )였다.

사전조사(pretest)에서는 모든 실험참가자들이 실험자극물로 제시된 자기기입식 설문지(self-administered paper-and-pencil questionnaire)를 집단적으로 강의실에서 완성하여 제출하였다. 이후 2주일이 경과한 뒤 개별 실험참가자들은 무작위로 대면 면접(21명, 31.8%), 전화 면접(24명, 36.4%), 또는 인터넷 서베이(21명, 31.8%) 조건 중 하나에 할당되어 사전조사에서와 동일한 설문내용에 대한 응답을 제공하였다. 2주일 동안의 경과시간은 동일한 실험자극물의 반복 적용 효과를 최소화하는 동시에 실험자극물로 제시한 설문지 대한 응답자의 응답내용이 시간이 경과함에 따라 실제로 변화할 가능성이 적을 것으로 판단하여 선정하였다.

자기기입식 설문지를 활용한 사전조사의 목적은 이후 무작위 할당(random assignment)을 통해 구성될 세 개의 실험집단의 동질성을 확인 또는 확보하기 위함이다. 이들 세 개 실험집단의 동질성이 사전조사의 응답을 바탕으로 확인 또는 확보되지 않는다면, 상이한 세 종류의 서베이 모드를 적용한 후의 집단 간 응답의 차이가 서베이 모드의 순수한 효과라고 주장하기 어렵다. 그리고 사전조사에서 별도의 서베이 모드(자기기입식 설문지)를 활용한 것은 동일한 모드를 반복 적용함으로써 발생할 수 있는 문제를 최소화하기 위함이다. 만일 세 개의 실험집단(대면, 전화, 온라인) 중 어느 한 집단의 경우, 사전조사에서와 동일한 서베이 내용뿐만 아니라 서베이 모드까지 동일하다면 그 실험집단에서는 반복 측정이 갖는 문제가 더욱 심할 것으로 판단하였다. 결국 본 연구에서는 네 가지 서베이 모드(설문지, 대면, 전화, 온라인)의 영향력이 측정되었으나, 사전조사에서 활용된 설문지(paper-and-pencil questionnaire) 모드는 실험집단들의 동질성을 확보하고, 이후 사후조사의 결과와 비교하여 설문 응답상의 변화의 크기 및 방향을 측정하기 위한 준거(baseline)로 활용되었을 뿐 본 연구의 주요 관심대상인 세 가지 서베이 모드들(대면, 전화, 온라인)과의 비교대상은 아니다.

사후조사(posttest)에서 대면 면접은 실험참가자들과 같은 연령대의 두 명의 훈련받은 여성 면접원들에 의해 진행되었으며, 대면 상황에서 면접원들이 직접 설문내용을 읽어준 후 응답내용을 기록하였다. 그리고 모든 대면 면접은 외부인의 출입이 제한된 면접장소(실험실)에서 진행되거나 또는 응답자가 원하는 장소에서 실시되는 경우에는

면접원의 통제 하에 제3자 존재의 영향을 최소화할 수 있도록 조치되었다. 전화 면접 역시 이들 두 여성 면접원이 실험참가자들에게 직접 전화를 걸어 설문을 완성하였다. 온라인 서베이 조건에 할당된 실험참가자들은 이메일 및 휴대전화 메시지를 통해 공지된 웹 서베이 사이트 주소로 개별적으로 접속하여 자기기입 방식으로 온라인 설문을 완성하였다. 실험참가자들이 사전조사와 사후조사의 응답을 비교하려는 본 실험 설계의 궁극적 목적을 알지 못 하도록 하기 위해 사후조사 참가를 요청할 때에는 사전조사의 데이터가 훼손되었다고 설명하였다. 그러나 사후조사 완료 후 참가자들에게 본 연구의 목적에 대한 설명을 제공하였으며 양해를 구하였다.

## 2. 실험자극물(서베이 내용)

실험자극물로서의 서베이 내용은 실험참가자인 대학생 집단에게 민감한 것으로서 응답 과정에서 사회적 바람직함 편향을 자극할 수 있는 주제를 신중히 선택 또는 고안하였다. 설문 내용은 사회적으로 바람직하지 못한 것으로 일반적으로 인식되고 있는 세 종류의 중독증(도박, 인터넷, 게임) 수준과 본인 외모에 대한 불만족 정도, 성적 가치관(sexual values) 및 성 경험(sexual experience) 행동을 측정할 수 있는 척도 및 질문들을 포함하였다(<표 1> 참조).

<표 1> 실험자극물로서의 민감한 설문 내용 및 측정방법

변인	척도명	항목	척도	N	M(SD)	$\alpha$
도박 중독증	PGSI(Problem Gambling Severity Index)	9	4점	66	9.24(2.54)	.84
인터넷 중독증	성인 인터넷 중독 자가진단 척도 간략형(KS-A)	15	4점	66	29.13(7.24)	.90
게임 중독증	청소년 인터넷 게임중독 척도	20	4점	66	26.80(8.84)	.96
본인 외모 불만족	외모 인식 척도	6	5점	66	14.43(3.71)	.66
성적 가치관	성태도 척도	10	5점	66	27.88(7.14)	.89
첫 성경험 연령	-	1	-	27	18.59(1.67)	-
성관계자 수	-	1	-	26	10.81(3.68)	-
현재 이성과의 성적 행위	-	1	5점	19	3.68(1.42)	-

실험참가자들의 도박 중독증 수준은 일반인들의 문제도박 행동에 대한 최적의 진단 도구로 평가되고 있는(허태균 등 2009) CPGI(Canadian Problem Gambling Index; Ferris & Wynne 2001)에서 추출된 9개 문항을 재구성 및 번역한 PGSI(Problem Gambling Severity Index; 김교현·조성겸·권선중·이동형 2010)를 활용하여 측정하였다. 지난 1년 동안 도박과 관련한 경험들을 질문하는 모든 항목은 1 = “전혀 아니다”로부터 4 = “거의 항상 그렇다”에 이르는 4점 리커트(Likert)형 척도로 측정되었다. 9개 문항으로 구성된 도박 중독증 척도의 사전조사에서의 신뢰도(Cronbach's  $\alpha$ )는 .84로 높은 수준이었으며, 최소 9점에서 최대 36점에 이를 수 있는 개인별 도박 중독증 점수의 평균은 9.24점( $SD = 2.54$ )이었다.

인터넷 중독증 수준은 신광우·김동일·정여주(2011)에 의해 최종 검증되어 현재 국가기관인 한국정보화진흥원(인터넷중독대응센터)에 의해 공식적으로 활용되고 있는 ‘성인 인터넷 중독 자가진단 척도 간략형(KS-A)’을 통해 측정하였다. 이 척도는 총 15개 진술문으로 구성되어 있으며, 각 진술문에 대해 1 = “전혀 그렇지 않다”로부터 4 = “매우 그렇다”에 이르는 리커트형 4점 척도로 응답자의 동의 정도를 파악한다. 사전조사 결과, 15개 문항 전체의 신뢰도는 .90으로 매우 높은 수준이었으며, 최소 15점에서 최대 60점에 사이에 분포 가능한 개인별 인터넷 중독증 점수의 평균은 29.13점( $SD = 7.24$ )이었다.

중독증 가운데 대학생 집단에서 가장 보편적으로 관찰 가능할 것으로 판단되어 측정에 포함된 게임 중독증은 이경옥 등(2006)이 개발한 ‘청소년용 인터넷 게임중독 척도’를 통해 측정하였다. 총 20개의 진술문으로 구성된 이 척도는 원래 게임지향적 생활, 내성과 통제력 상실, 금단과 정서경험의 세 요인으로 구성되어 있으나 분석의 편의를 위해 단일 차원의 개념으로 간주하였다. 사전조사에서의 신뢰도는 .96으로 매우 높은 수준이었다. 각 진술문은 1 = “전혀 그렇지 않다”에서 4 = “항상 그렇다”까지의 4점 리커트형 척도로 측정되었으며, 실험참가자들의 평균은 26.80점( $SD = 8.84$ )으로 게임 중독증 수준이 매우 낮은 것으로 평가되었다.

다음으로 본인 외모에 대한 불만족 정도는 그 질문의 성격상 응답자의 자존감을 드러내는 매우 민감한 주제라는 판단에 따라 포함된 주제이다. 특히 응답자의 외모에 대한 정보는 대면 면접 상황을 제외한 다른 모드에서는 응답자가 고의적으로 왜곡할 가능성이 보다 많을 것으로 예상하여 실험자극물로 제시하였다. 본인 외모에 대한 불만

측 정도는 2007년 민간기관이 수행한 청소년 성문화 연구조사(김백애라·김정아·이목소희 2007)에서 활용된 척도를 통해 측정하였다. 이 척도는 “가슴의 근육이나 유방의 크기 때문에 걱정한 적이 있다”, “성기 형태나 크기 때문에 걱정을 한 적이 있다”, “나는 언젠가는 반드시 성형수술을 받을 것이다”를 포함한 총 6개의 진술문을 1 = “전혀 그렇지 않다”로부터 5 = “매우 그렇다”에 이르는 5점 리커트형 척도로 측정하였다. 사전조사에서의 신뢰도는 .66으로 다소 낮으나 수용 가능한 수준이었으며, 평균은 14.43점( $SD=3.71$ )이었다. 점수는 높을수록 본인 외모에 대한 불만족 정도가 높은 것이다.

실험참가자들의 성적 가치관은 국내외 여러 선행연구들에서 사용된 척도를 재구성한 허은주(2004)의 성태도 척도에서 민감성 수준이 높은 10개의 항목을 사용하여 측정하였다. 포함된 항목에는 “사랑 없이도 성관계를 할 수 있다”, “결혼 전 성경험은 결혼 생활에 도움이 된다고 생각한다”, “돈을 벌기 위해 낮은 사람과 성관계를 할 수 있다” 등이 있다. 이러한 종류의 성적 가치관은 지나치게 진보적 또는 개방적인 것으로서 우리나라의 유교적(보수적) 성문화에 비추어 바람직하지 않은 것으로 인식된다고 연구자들은 보았다. 각 진술문들은 1 = “전혀 동의하지 않는다”로부터 5 = “매우 동의한다”의 5점 척도로 평가되었으며, 사전조사에서의 신뢰도는 .89로 높은 수준이었다. 각 항목의 점수를 합산한 개인의 점수가 낮을수록 ‘보수적’(폐쇄적) 그리고 높을수록 ‘진보적’(개방적) 성적 가치관을 의미하는데, 사전조사의 평균은 27.88점( $SD=7.14$ )으로 중도적 수준이었다.

그리고 본 실험의 목적상, 질문의 민감성 및 서베이 응답의 사회적 바람직함 편향 효과를 극대화하기 위해 실험참가자들의 성적 경험에 관한 질문들을 추가로 포함하였다. 먼저 응답자들의 성교행위(sexual intercourse)와 관련하여 첫 경험 연령과 지금까지 경험한 성교행위 파트너의 수가 몇 명인가를 질문하였다. 그리고 현재 교제 중인 이성관계가 있는지 여부를 확인한 후, 현재의 이성친구와 경험한 성적인 행위를 1 = ‘스킨십’, 2 = ‘포옹’, 3 = ‘키스’, 4 = ‘애무’, 5 = ‘성관계’에 이르는 거트만(Guttman)형 척도를 통해 점수를 부여하였다. 즉 현재의 이성친구와 성관계를 한 것으로 응답한 경우, 하위 수준의 행위를 모두 경험한 것으로 간주하고 5점을 부여하였다. 사전조사에서 성교 경험이 있다고 응답한 27명 응답자들의 평균 첫 경험 연령은 만 18.59세( $SD=1.67$ )였으며, 지금까지 경험한 성교 파트너 수는 10.81명( $SD=3.68$ )으로 연구자들이 예상한 수준을 훨씬 상회하는 것으로 나타났다. 그리고 현재 이성친구와 교제 중인

19명이 경험한 성적 행위 수준은 평균 3.68( $SD=1.42$ )로 키스를 포함한 하위 수준의 행위를 경험한 것으로 나타났다. 마지막으로, 일반적인 서베이의 구성과 마찬가지로 성별, 연령, 가구소득 등의 인구통계학적 질문을 포함하여 설문지 기본 형식을 갖추었다.

### 3. ‘사회적 바람직함 편향’ 측정

실험 데이터에 대한 최종 분석은 개별 실험참가자들의 사전조사에서의 응답과 사후조사에서의 응답을 비교하여 그 변화량을 통계적으로 분석하였다. 구체적으로 설명하면, 민감한 질문들을 포함한 서베이 응답에 세 가지 상이한 모드들이 미치는 영향을 비교·검증하기 위해 본 연구는 각 모드별로 사회적 바람직함 편향이 발생하는 정도를 종속변인으로 설정하였다. 그리고 사회적 바람직함 편향은 사전조사에서의 측정치로부터 사후조사의 측정치를 산술적으로 뺀 값으로 조작적 정의를 하였다. 즉, 사전조사(집단 상황에서 완료된 자기기입식 설문지)의 응답과 비교하여 상이한 서베이 모드들을 통한 사후조사의 응답이 얼마나 다른가를 측정함으로써 각 서베이 모드에서 발생한 사회적 바람직함 편향의 방향과 크기를 측정하는 것이다. 따라서 사전조사의 측정치에서 사후조사의 측정치를 뺀 값이 0에 가까울수록 사회적 바람직함 편향이 거의 발생하지 않은 것이며, 이 값이 양(+의 방향으로 커질수록 사전조사에 비해 사회적 바람직함 편향이 크게 작용한 것이고, 반대로 이 값이 음(-의 방향으로 클수록 사전조사에 비해 사회적 바람직함 편향이 작게 작용한 것으로 볼 수 있다.

이와 같이 각각의 서베이 모드에서 발생한 사회적 바람직함 편향의 정도를 상호 비교함으로써 본 연구에서 제시한 가설을 통계적으로 검증하고자 한다. 그리고 통계 결과의 해석은, 민감한 정보를 얻기 위한 서베이에서 상이한 모드의 효과를 비교·검토한 대부분의 선행연구들과 마찬가지로, 사회적으로 바람직하지 못한 태도 및 행동에 대한 솔직한 응답이 많을수록, 즉 사회적 바람직함 편향이 음(-의 방향으로 클수록 더욱 정확한 자료수집 방법이라는 가정을 바탕으로 이루어졌다. 이는 상식적으로 생각해 볼 때, 서베이 응답자들이 사회적으로 바람직하지 못한 것으로 인식되는 태도나 행동을 고의적으로 왜곡하여 과대응답할 가능성 또는 이유는 거의 없기 때문이다.

〈표 2〉 무작위 할당 후 사전조사 응답의 일원변량분석(ANOVA) 결과

변인	<i>df</i>	<i>F</i>	<i>p</i>
도박 중독증	(2, 63)	2.13	.128
인터넷 중독증	(2, 63)	.09	.913
게임 중독증	(2, 63)	1.29	.281
본인 외모 불만족	(2, 63)	.31	.731
성적 가치관	(2, 63)	.28	.761
첫 경험 연령	(2, 24)	1.80	.187
성관계 파트너 수	(2, 23)	.59	.560
현재 이성과의 성적 행위	(2, 16)	.24	.788

## IV. 연구결과

### 1. 세 실험집단 간 동질성 검증

사전조사에 참가한 93명의 응답자들 중 사후조사에 참여를 거부하거나 연락이 되지 않는 경우를 제외한 66명을 무작위로 대면 면접, 전화 면접, 온라인 서베이 모드 중 하나에 할당하였다. 그리고 무작위 할당의 성공 여부를 확인하기 위해 사전조사에서의 응답을 통계분석하여 세 집단 간 동등성을 확인하였다. 〈표 2〉는 사후조사를 위한 무작위 할당이 성공적으로 이루어졌음을 보여준다. 일원변량분석(ANOVA) 결과, 모든 측정 변인에 대한 사전조사의 응답 수준에 있어 세 집단 간에 통계적으로 유의한 차이는 없는 것으로 확인되었다.

### 2. 가설 검증: 서베이 모드가 민감한 질문에 대한 응답에 미치는 효과 비교

민감한 질문에 대한 응답에 세 종류의 상이한 서베이 모드가 미치는 영향을 검증하기 위해 각 모드별로 발생한 사회적 바람직함 편향 점수(사전조사의 점수 - 사후조사 점수)를 이용하여 세 집단 간 변량분석(3 between-subject ANOVA)을 실시하였다. 앞서 확인하였듯이 민감한 질문들에 대한 사전조사에서의 응답은 무작위 할당을 통해

〈표 3〉 서베이 모드에 따른 사회적 바람직함 편향의 변량분석 결과

종속변인	모드별 사회적 바람직함 편향 점수(SD)*		F- value
인터넷 중독증	전화(n=24)	.71(8.31)	F(2,63)=1.498, p=.231
	대면(n=21)	3.05(6.43)	
	온라인(n=21)	-.48(4.65)	
게임 중독증	전화(n=24)	2.58(6.16)	F(2,63)=3.350, p=.041
	대면(n=21)	3.10(7.42)	
	온라인(n=21)	-2.48(9.57)	
도박 중독증	전화(n=24)	.83(2.01)	F(2,63)=2.289, p=.110
	대면(n=21)	-.57(2.09)	
	온라인(n=21)	.48(2.66)	
본인 외모 불만족	전화(n=24)	2.54(3.16)	F(2,63)=5.579, p=.006
	대면(n=21)	1.19(2.58)	
	온라인(n=21)	-.38(2.97)	
성적 가치관	전화(n=24)	.04(.55)	F(2,61)=.230, p=.796
	대면(n=19)	.00(.94)	
	온라인(n=21)	-.10(.54)	
첫 경험 연령	전화(n=8)	.63(.74)	F(2,19)=.562, p=.579
	대면(n=7)	.29(1.25)	
	온라인(n=7)	1.14(2.27)	
성관계 파트너 수	전화(n=8)	-7.50(25.50)	F(2,19)=.395, p=.679
	대면(n=7)	-4.29(11.34)	
	온라인(n=7)	.43(7.57)	
현재 이성과의 성적 행위	전화(n=6)	-.67(1.63)	F(2,13)=.784, p=.477
	대면(n=4)	.50(1.73)	
	온라인(n=6)	.00(1.10)	

\* 사회적 바람직함 편향 점수 = 사전조사 점수 - 사후조사 점수

구성된 세 집단 간에 통계적으로 유의한 차이가 없었다. 따라서 상이한 서베이 모드에 따라 민감한 질문에 대한 응답이 달라질 것으로 예측한 본 연구의 가설(“민감한 질문에 대한 서베이 응답에서 사회적 바람직함 편향은 전화 면접에서 가장 많이 작용할 것이고, 다음으로 대면 면접, 그리고 온라인 서베이 순으로 나타날 것이다”)은 세 집단 간 사전-사후 조사 응답의 변화량을 비교·분석함으로써 검증할 수 있다. 즉, 사전-사후 조사 간의 응답의 변화량은 자기기입식 설문지를 활용한 사전조사에서의 응답을 기준으로 민감한 질문으로 인한 사회적 바람직함 편향이 세 종류의 서베이 모드들에 따라 상이하게 나타나는 양상을 구체적으로 보여준다. <표 3>은 각 종속변인별로 세 종류의 서베이 모드에 따른 사회적 바람직함 편향의 일원변량분석(one-way ANOVA) 결과를 요약한 것이다.

<표 3>의 변량분석 결과를 보면, 기타 종속변인에서는 세 종류의 서베이 모드에 따른 사회적 바람직함 편향 점수의 차이가 발견되지 않았으나, 게임 중독증과 본인 외모 불만족의 경우에는 사회적 바람직함 편향의 발생 정도, 즉 동일한 질문에 대한 사전조사와 사후조사 사이의 응답의 변화 패턴이 서베이 모드에 따라 상이한 것으로 나타났다.

먼저 게임 중독증의 경우, 대면 면접 모드에서 사회적 바람직함 편향이 가장 높은 수준으로 발견되었으며( $M=3.10$ ,  $SD=7.42$ ), 다음으로 전화 면접 모드에서 응답자들이 사회적으로 바람직한 응답을 제공하는 경향, 즉 응답자 본인의 게임 중독 증상을 축소하여 응답하는 경향을 보였다( $M=2.58$ ,  $SD=6.16$ ). 한편 온라인 서베이 모드에서는 사전조사의 집단적 자기기입식 설문지 작성 모드에 비해 응답상의 사회적 바람직함 편향이 오히려 감소한 것으로 나타났다( $M=-2.48$ ,  $SD=9.57$ ). 이는 응답자들이 사전조사의 자기기입식 설문지에 비해 사후조사의 온라인 서베이를 통해 사회적으로 바람직하지 못한 것으로 인식되는 본인의 게임 중독과 관련한 증상을 더욱 진솔하게 응답한 것으로 해석할 수 있다.

이상의 세 종류의 서베이 모드에서 발생한 사회적 바람직함 편향의 차이는 통계적으로 유의한 것으로 확인되었다:  $F(2,63)=3.35$ ,  $p=.041$ . 결국, 게임 중독증에 관한 서베이의 경우, 자기기입 방식의 온라인 서베이 모드가 면접원 진행 방식의 대면 및 전화 면접 모드에 비해 사회적으로 바람직하지 못한 태도 및 행동(게임 중독증)과 관련한 솔직한 응답을 더욱 증가시킬 것이라는 예측은 확인되었으나, 대면 면접이 전화

면접에 비해 솔직한 응답을 더욱 증가시킬 것이라는 예측은 본 연구에서 확인되지 못했다. 따라서 게임 중독증과 관련한 민감한 질문에 대한 서베이 응답에서 사회적 바람직함 편향은 전화 면접에서 가장 많이 작용할 것이고, 다음으로 대면 면접, 그리고 온라인 서베이 순으로 나타날 것으로 예측한 본 연구의 가설은 부분적으로만 지지되었다.

본인 외모 불만족 정도의 경우에도 서베이 모드에 따라 사전-사후 응답의 변화 패턴이 달라지는 사회적 바람직함 편향의 효과는 통계적으로 유의한 수준에서 확인되었다:  $F(2,63)=5.58, p=.006$ . 전화 면접 모드에서 사회적 바람직함 편향이 작용하여 사후조사에서의 응답 수준(본인 외모에 대한 불만족 정도)이 사전조사에 비해 가장 많이 감소하였으며( $M=2.54, SD=3.16$ ), 다음으로 대면 면접 모드 역시 사전조사의 응답에 비해 본인의 외모에 대한 불만족 정도를 낮게 보고하는 경향이 발견되었다( $M=1.19, SD=2.58$ ). 반면, 온라인 서베이 모드에서는 게임 중독증에 대한 응답과 마찬가지로 사전조사에 비해 사회적 바람직함 편향이 감소하여 본인의 외모에 대한 불만족을 더욱 많이 응답하는 경향이 확인되었다( $M=-.38, SD=2.97$ ). 이상의 결과는 전화 면접, 대면 면접, 온라인 서베이 순으로 본인 외모 불만족에 대한 응답에 영향을 미치는 사회적 바람직함 편향의 수준이 증가할 것으로 예측한 본 연구의 가설이 지지되고 있음을 알려준다.

그러나 본 연구에서 활용한 기타 민감한 질문들(인터넷 중독증, 도박 중독증, 성적 가치관, 첫 경험 연령, 그리고 성관계 파트너 수)에서는 서베이 모드에 따른 사회적 바람직함 편향의 차이가 통계적으로 유의하지 않은 것으로 확인된 바, 본 연구의 가설은 부분적으로만 지지된 것으로 결론내릴 수 있다. 이에 다음 절에서는 이러한 연구결과의 원인 분석을 위해 실시한 추가 분석 결과를 소개하고자 한다.

### 3. 추가 분석: 민감한 질문에 대한 잠재적 응답 수준과 서베이 모드의 상호작용 효과

가설검증 결과, 민감한 서베이 질문에 대한 응답에 세 종류의 서베이 모드가 미치는 효과의 차이는 게임 중독증과 본인 외모 불만족 정도에서만 부분적으로 확인할 수 있었다. 이러한 결과는 본 연구에서 제시한 서베이 질문들의 민감성 수준이 모든 실험 참가자들에게 동일하지 않았기 때문에 발생한 것일 수 있다. 이론적 논의 부분에서 이미 언급하였듯이, 서베이 응답에 반영되는 사회적 바람직함 편향의 정도는 질문의 민

감성에 따라 결정되며, 서베이 질문의 민감성은 응답자의 잠재적 응답에 따라 결정된다. 예를 들어, 미성년자를 대상으로 흡연 여부를 묻는 질문의 민감성은 미성년 응답자들의 잠재적 응답(즉 실제 흡연 여부)에 따라 달라질 것이다. 동일한 질문이라 하더라도 실제 흡연을 하지 않는 미성년자들은 이러한 질문을 전혀 민감한 것으로 받아들이지 않는 반면, 실제 흡연을 하는 미성년자들에게는 민감한 질문이 될 수 있다는 것이다.

이러한 가능성이 본 연구의 상황에도 적용되는 바, 본 연구에서 제시한 여러 가지 민감한 질문들에 대한 사전조사에서의 응답이 사회적으로 바람직한 수준에 있는 응답자들에게는 사후조사가 그리 민감하지 않은 것으로 인식되었을 것이며 따라서 사회적 바람직함 편향이 거의 나타나지 않은 것으로 볼 수 있다. 반면에 사전조사에서 사회적으로 바람직하지 못한 태도나 행동을 높은 수준으로 보고한 응답자들에게는 질문들이 민감한 것으로 인식되고 결국 사후조사에서는 사회적으로 바람직한 방향으로 사전조사의 응답을 왜곡하여 보고할 가능성이 높다. 이러한 논리에 따르면, 민감한 질문에 대한 응답(즉, 사후조사에서의 응답)에 상이한 서베이 모드가 미치는 영향을 보다 정확히 파악하기 위해서는 잠재적 응답(즉, 사전조사에서의 응답)을 기준으로 서베이 질문에 대한 민감성 수준이 높은 응답자들과 낮은 응답자들을 구분할 필요가 있다. 가설 검증을 위해 앞서 실시한 통계방법은 이러한 집단 간 차이를 고려하지 않음으로써 게임 중독증과 본인 외모 불만족 정도를 제외한 나머지 측정변인에 대해서는 서베이 모드의 상이한 효과를 확인할 수 없었던 것으로 판단된다.

이에 추가적 분석으로, 각 주제별로 사전조사에서의 응답 수준과 서베이 모드를 독립변인으로 하는 이원변량분석(two-way ANOVA)을 실시하였다. 사전조사에서의 응답 수준은 중위수(median)를 기준으로 저집단과 고집단으로 구분하였으며, 서베이 모드는 사후조사에 적용된 세 종류(전화 면접, 대면 면접, 온라인 서베이)를 포함하여 2×3 팩토리얼(factorial) 디자인으로 구성하였다. 종속변인은 각 주제별로 사회적 바람직함 편향 점수(사전조사의 응답 점수로부터 사후조사의 응답 점수를 빼서 산출한 응답 점수 변화량)를 활용하였다. 따라서 사회적 바람직함 편향 점수가 양수로 나타나는 것은 사전조사에 비해 사후조사의 솔직한 응답이 감소했음을 의미하는 반면, 음수는 사전조사에 비해 사후조사의 응답이 증가했음을 의미한다. 이원변량분석의 결과는 <표 4>에 요약하였다.

〈표 4〉 잠재적 응답 수준 및 서베이 모드에 따른 사회적 바람직함 편향:  $M(SD)$

		저집단( $n=34$ )	고집단( $n=32$ )	주효과 및 상호작용 효과
인터넷 중독증	대면( $n=21$ )	.50(4.76)	6.44(7.04)	집단: $F(1,60) = 23.15, p = .000$ 모드: $F(2,60) = 2.72, p = .074$ 집단*모드: $F(2,60) = 1.16, p = .322$
	전화( $n=24$ )	-4.17(8.17)	5.58(5.09)	
	온라인( $n=21$ )	-3.00(4.30)	1.82(3.82)	
게임 중독증		저집단( $n=33$ )	고집단( $n=33$ )	
	대면( $n=21$ )	-.44(1.81)	5.75(8.94)	집단: $F(1,60) = 18.54, p = .000$ 모드: $F(2,60) = 5.65, p = .006$ 집단*모드: $F(2,60) = .19, p = .824$
	전화( $n=24$ )	-.20(1.27)	7.22(8.21)	
온라인( $n=21$ )	-7.56(9.82)	1.33(7.73)		
도박 중독증		저집단( $n=43$ )	고집단( $n=23$ )	
	대면( $n=21$ )	-.07(.26)	-1.83(3.82)	집단: $F(1,60) = 1.52, p = .222$ 모드: $F(2,60) = 5.32, p = .007$ 집단*모드: $F(2,60) = 5.06, p = .009$
	전화( $n=24$ )	.00(.00)	2.50(2.93)	
온라인( $n=21$ )	-.08(.29)	1.22(4.06)		
본인 외모 불만족		저집단( $n=35$ )	고집단( $n=31$ )	
	대면( $n=21$ )	1.08(1.83)	1.33(3.46)	집단: $F(1,60) = 7.41, p = .008$ 모드: $F(2,60) = 4.89, p = .011$ 집단*모드: $F(2,60) = 1.51, p = .230$
	전화( $n=24$ )	1.36(1.80)	3.54(3.76)	
온라인( $n=21$ )	-1.75(2.93)	1.44(1.94)		
성적 가치관		저집단( $n=36$ )	고집단( $n=30$ )	
	대면( $n=21$ )	.00(3.77)	1.70(3.20)	집단: $F(1,60) = 7.32, p = .009$ 모드: $F(2,60) = 1.92, p = .155$ 집단*모드: $F(2,60) = 1.50, p = .241$
	전화( $n=24$ )	.29(4.48)	5.80(5.61)	
온라인( $n=21$ )	.00(5.71)	1.60(2.27)		
첫 경험 연령*		저집단( $n=19$ )	고집단( $n=3$ )	
	대면( $n=7$ )	-	.29(1.25)	집단: $F(1,17) = 16.90, p = .001$ 모드: $F(2,17) = 7.56, p = .004$ 집단*모드: $F(1,17) = 19.01, p = .000$
	전화( $n=8$ )	.50(.71)	.67(.82)	
온라인( $n=7$ )	6.00(-)	.33(.82)		
성관계 파트너 수		저집단( $n=10$ )	고집단( $n=12$ )	
	대면( $n=7$ )	.00(.00)	-7.50(15.00)	집단: $F(1,16) = .71, p = .412$ 모드: $F(2,16) = .34, p = .714$ 집단*모드: $F(2,16) = .48, p = .631$
	전화( $n=8$ )	.00(.00)	-15.00(36.97)	
온라인( $n=7$ )	-1.33(2.31)	1.75(10.28)		
현재 이성과의 성적 행위		저집단( $n=8$ )	고집단( $n=8$ )	
	대면( $n=4$ )	-.50(.71)	1.50(2.12)	집단: $F(1,10) = 4.91, p = .051$ 모드: $F(2,10) = .90, p = .437$ 집단*모드: $F(2,10) = .09, p = .917$
	전화( $n=6$ )	-1.33(2.31)	.00(.00)	
온라인( $n=6$ )	-.67(.58)	.67(1.16)		

\* 사전조사 응답 점수(첫 경험 연령)가 낮을수록 잠재적 응답 수준이 높은 고집단으로 분류

인터넷 중독증의 경우, 예상했던 대로 사전조사에서의 응답 수준이 높은 집단과 낮은 집단의 응답 점수 변화량은 통계적으로 유의한 차이가 있는 것으로 나타났다:  $F(1,60)=23.15, p=.000$ . 즉 인터넷 중독에 대한 사후조사 응답에 있어 사전조사에서의 응답 수준이 높았던 고집단( $M=4.62, SE=1.02$ )이 저집단( $M=-2.22, SE=.99$ )에 비해 사전조사의 응답을 더욱 많이 변경한 것으로 나타났다. 고집단은 사전조사에 비해 사후조사에서 인터넷 중독증을 더욱 낮게 보고한 반면, 저집단은 사후조사에서 오히려 인터넷 중독증을 더욱 높게 응답하였다. 인터넷 중독증에 대한 사전-사후 조사 간 응답 변화량에 미치는 서베이 모드의 주효과:  $F(2,60)=2.72, p=.074$  및 집단과 서베이 모드의 상호작용 효과는 없는 것으로 나타났다:  $F(2,60)=1.16, p=.322$ .

게임 중독증에서는 사전조사에서의 응답 수준에 따른 집단:  $F(1,60)=18.54, p=.000$ 과 서베이 모드의 주효과가 각각 확인되었다:  $F(2,60)=5.65, p=.006$ . 사전조사 응답 수준이 높았던 고집단( $M=4.77, SE=1.22$ )은 사후조사에서 게임 중독증을 감소시켜 보고한 반면, 저집단( $M=-2.73, SE=1.24$ )은 사후조사에서 솔직한 응답 수준을 증가시켰다. 그리고 전화 면접( $M=3.51, SE=1.46$ )과 대면 면접( $M=2.65, SE=1.53$ )에서는 사후조사의 응답이 사전조사의 응답에 비해 감소한 반면, 온라인 서베이( $M=-3.11, SE=1.53$ )에서는 증가하였다. 그러나 집단과 서베이 모드 간의 상호작용 효과는 확인되지 않았다:  $F(2,60)=1.16, p=.322$ .

도박 중독증에 대한 사전-사후 응답 점수 변화량에는 사전조사에서의 응답 수준의 주효과가 발견되지 않았다:  $F(1,60)=1.52, p=.222$ . 반면 서베이 모드의 주효과는 발견되었다:  $F(2,60)=5.32, p=.007$ . 전화 설문( $M=1.25, SE=.46$ )에서 사전-사후 조사 간 응답 점수 변화량이 가장 많았으며, 다음으로 면접 설문( $M=-.95, SE=.46$ )과 온라인 서베이( $M=.57, SE=.46$ ) 순이었으나, 면접 설문의 경우에는 사전조사에 비해 사후조사의 응답 점수가 높았다. 사전조사의 응답 수준과 서베이 모드 간의 상호작용 효과도 있는 것으로 나타났다:  $F(2,60)=5.06, p=.009$ . 사전조사에서의 도박 중독증에 대한 응답 수준이 낮은 저집단의 경우에는 서베이 모드와 관계없이 응답 점수의 변화가 거의 없는 반면, 응답 수준이 높았던 고집단에서는 전화 면접( $M=2.50, SE=.76$ ), 대면 면접( $M=-1.83, SE=.86$ ), 온라인 서베이( $M=1.22, SE=.70$ ) 순으로 응답 점수 변화량의 절대값이 많은 것으로 나타났다.

본인 외모 불만족 정도에 있어서도 사전조사에서의 응답 수준의 주효과가 나타났

다:  $F(1,60)=7.41$ ,  $p=.008$ . 사전조사의 응답 수준이 높은 고집단( $M=2.11$ ,  $SE=.50$ )이 저집단( $M=.23$ ,  $SE=.47$ )에 비해 본인의 외모에 대한 불만족 정도가 낮은 것으로 응답을 더욱 많이 변경한 것이다. 서베이 모드에 따라서도 사전-사후 조사 간 응답 변화량에 차이가 있는 것으로 나타났다:  $F(2,60)=4.89$ ,  $p=.011$ . 전화 면접( $M=2.45$ ,  $SE=.57$ )에서 응답 점수 변화량이 가장 많았으며, 다음으로 대면 면접( $M=1.21$ ,  $SE=.61$ ), 온라인 서베이( $M=-.15$ ,  $SE=.61$ ) 순이었다. 사전조사에서의 응답 수준과 서베이 모드 간의 상호작용 효과는 없었다:  $F(2,60)=1.51$ ,  $p=.230$ .

성적 가치관에 대한 사전-사후 조사 간 응답 변화량도 사전조사의 응답 수준에 따라 달라진다는 것이 확인되었다:  $F(1,60)=7.32$ ,  $p=.009$ . 즉 사전조사에서 성적 가치가 비교적 진보적인 것으로 응답한 응답자들( $M=3.03$ ,  $SE=.80$ )이 보수적인 응답자들( $M=.10$ ,  $SE=.74$ )에 비해 사후조사에서 그들의 성적 가치관을 더욱 보수적인 방향으로 변경한 것이다. 그러나 서베이 모드의 주효과:  $F(2,60)=1.92$ ,  $p=.155$  및 사전조사에서의 응답과 서베이 모드의 상호작용 효과는 없는 것으로 나타났다:  $F(2,60)=1.50$ ,  $p=.241$ .

첫 경험 연령에 대한 사후조사 응답도 사전조사에서의 응답 수준에 따라 다르게 변화하는 것으로 나타났다:  $F(1,17)=16.90$ ,  $p=.001$ . 그러나 다른 민감한 주제와 달리, 사전조사의 응답 수준이 낮은 저집단(즉, 첫 경험 연령을 19세를 초과한 것으로 사전조사에서 응답한 집단,  $M=3.25$ ,  $SE=.61$ )이 고집단(첫 경험 연령이 19세 이하인 집단,  $M=.43$ ,  $SE=.23$ )에 비해 사후조사에서 그들의 첫 경험 연령을 더욱 줄여서 응답하는 것으로 나타났다. 서베이 모드에 따라서도 첫 경험 연령에 대한 사전-사후 간 응답 변화량에 차이가 발생하는 것으로 나타났다:  $F(2,17)=7.56$ ,  $p=.004$ . 전체적으로 사전조사에 비해 사후조사에서 응답한 첫 경험 연령이 감소되는 경향을 보이는 가운데, 온라인 서베이( $M=3.17$ ,  $SE=.53$ ), 전화 면접( $M=.58$ ,  $SE=.40$ ), 대면 면접( $M=.29$ ,  $SE=.37$ )의 순으로 응답의 변화량이 많았다. 또한 사전조사의 응답 수준과 서베이 모드 간의 상호작용 효과도 있는 것으로 나타났다:  $F(1,17)=19.01$ ,  $p=.000$ . 특히 사전조사의 응답 수준이 낮은 저집단이 온라인 서베이를 통해 응답하는 경우, 사전조사 응답에 비해 첫 경험 연령이 가장 큰 폭으로 감소하는 것으로 나타났다.

끝으로 성관계 파트너 수와 현재 이성과의 성적행위에 대한 사전-사후 간 응답 점수 변화량에서는 통계적으로 유의한 수준의 주효과 또는 상호작용 효과를 발견할 수 없었다. 이는 질문과 관련하여 충분한 사례 수가 확보되지 않은 가운데 실험참가자들

의 응답들이 지나치게 큰 편차를 보이고 있다는 문제점에 기인한 것으로 판단된다.

이상에서 제시한 추가 분석의 결과를 요약하면, 성관계 파트너 수와 현재 이성과의 성적 행위를 제외한 대부분의 민감한 서베이 주제들(인터넷 중독증, 게임 중독증, 도박 중독증, 본인 외모 불만족 정도, 성적 가치관)의 경우, 서베이 모드와 관계없이 또는 서베이 모드와 상호작용하여 응답자들의 사전조사에서의 응답이 사회적으로 바람직하지 못한 것일수록 사후조사에서는 응답을 변경(대부분 사회적으로 바람직한 방향으로 왜곡)하는 정도가 커진다는 것이 확인되었다. 그러나 첫 경험 연령에 대한 응답은, 연구자들의 기대와 달리, 사전조사에서 응답자들의 첫 경험 연령이 높을수록 사후조사에서는 오히려 낮춰서 응답하는 정도가 심한 것으로 나타났다. 이러한 결과는 서베이 질문의 민감성이 응답자의 잠재적 응답이 무엇인가에 따라 달라지며, 응답자마다 상이하게 인식된 질문의 민감성에 따라 응답의 왜곡 정도(즉, 사회적 바람직함 편향의 정도)가 달라진다는 추론을 뒷받침해 준다.

## V. 논의 및 제언

본 연구는 사회조사에서 가장 일반적으로 활용되고 있는 서베이 방법의 정확성과 관련하여, 그 동안 이론적 및 실무적 차원에서 표본 오차(sampling error)에 비해 상대적으로 충분한 논의 및 고려가 이루어지지 않았던 응답 오차(response error)에 초점을 두고 있다. 구체적으로는, 민감한 서베이 주제로부터 발생하는 응답 상의 사회적 바람직함 편향(social desirability biases) 및 그에 따른 응답의 의도적 왜곡 현상이 다양한 종류의 서베이 모드(survey mode)에 따라 어떻게 달리 나타나는지를 실증적으로 분석함으로써 향후 서베이 조사 결과의 타당성을 제고할 수 있는 방법론적 대안을 제시하고자 하였다.

이를 위해 본 연구에서는 선행연구 결과의 검토 및 이론적 논의를 바탕으로, 서베이 질문의 민감성에 의해 활성화되는 사회적 바람직함 편향의 정도가 면접원의 존재 여부 그리고 제3자의 존재 여부를 달리하는 세 종류의 서베이 모드에 따라 달리 나타날 것으로 예측하였다. 우선 면접원의 존재 여부와 관련하여, 자기기입 방식의 온라인 서베이가 면접원 진행 방식의 대면 및 전화 면접 모드에 비해 사회적 바람직함 편향이 덜 발생하고, 다음으로 제3자의 존재에 대한 통제 가능성을 기준으로 대면 면접이 전

화 면접에 비해 사회적 바람직함 편향을 줄이는 데 더욱 효율적일 것이라고 예측하였다. 요약하면, 개인의 프라이버시 보호 및 응답내용의 비밀보장(confidentiality) 수준과 관련하여, 민감한 서베이 질문이 주어지는 경우 온라인 서베이, 대면 면접, 전화 면접 모드의 순으로 사회적 바람직함 편향의 영향이 축소되어 왜곡되지 않은 진솔한 응답을 이끌어내는 데 더욱 효율적일 것이라는 가설을 제시하였다.

이러한 연구 가설을 검증하기 위해 실시한 실험의 데이터를 분석한 결과, 본 연구에서 제시한 여러 민감한 주제들 중 게임 중독증과 본인 외모 불만족 정도에서만 서베이 모드에 따라 응답의 사회적 바람직함 편향이 달리 나타나는 것으로 확인되었으며, 연구가설로 제시한 서베이 모드별 사회적 바람직함 편향의 정도는 본인 외모 불만족 정도에 대한 응답에서만 확인되었다.

이러한 결과의 원인은 본 연구의 실험 자극물로 제시한 질문들의 민감성이 대학생들로 구성된 실험참가자 집단에서는 연구자들이 기대했던 수준만큼 이르지 못했기 때문일 가능성이 있다. 예를 들어, 인터넷 중독증의 경우, 비록 사회적으로 바람직하지 못한 일종의 병리적 현상으로 일반적으로 인식되고 있지만, 인터넷은 또한 효율적 정보탐색과 업무처리 그리고 원활한 대인관계 및 커뮤니케이션 등을 포함한 긍정적 기능을 위한 수단으로도 그 이용이 사회적으로 권장되고 있는 분위기이다. 따라서 다른 어느 세대집단에 비해 인터넷의 도구적(instrumental) 이용이 불가피하게 많은 대학생집단의 특성을 고려해 볼 때, 인터넷 중독증은 실제로 그 수준이 그리 높지 않거나 민감한 사안으로 받아들여지지 않을 가능성이 있다. 특히 도박 중독증의 경우에는 대학생 집단에서 거의 관찰되지 않는 일반적이지 않은 주제라는 사실이 본 연구에서도 확인되었다(사전조사 결과, 실험참가자들의 도박 중독증 수준은 9~36점에 이르는 척도 상에서  $M=9.24$ ,  $SD=2.54$ 였다.)

이에 반해, 주로 오락적인 목적과 의식적(ritualistic)인 이용으로부터 비롯되는 게임 중독증은 학업 부진, 부모와의 갈등, 대인관계에서의 고립 등의 여러 부정적인 결과들이, 성인으로서 인정받고 있는 대학생집단에서 오히려 더욱 정당화하기 어려움으로 인해, 민감한 주제로 인식되었을 수 있다. 그리고 본인의 외모에 대한 불만족 정도는 대학생집단의 세대적 특성과 외모를 중시하는 사회적 풍토, 그리고 본인의 외모에 대한 평가와 관련된 자존감 등이 복합적으로 연관되어 매우 민감한 주제로 인식되었을 것으로 판단된다.

성과 관련된 태도 및 행동들에 대한 질문 역시 본 실험에 참가한 대학생집단에게는 그리 민감한 주제로 인식되지 않은 것으로 판단된다. 개방적 또는 보수적 태도로 구분되는 성적 가치관의 경우, 보수적인 태도가 사회적으로 바람직한 방향의 것일 것이라는 연구자들의 기대와 달리 대학생집단에서는 그 방향성이 개인에 따라 다를 수도 있다. 즉 어떤 대학생들은 개방적인 성적 가치관 및 행동들을 오히려 사회적으로 바람직한 것으로 인식하고 있을 수도 있다는 것이다. 이는 첫 경험 연령과 성관계 파트너 수, 그리고 현재 이성과의 성적 행위에 대한 응답이 비록 모두 통계적으로 유의하지는 않았으나 사전조사에 비해 사후조사의 응답이 개방적으로 변화하는 경향으로 확인할 수 있다. 특히 전화 면접 모드를 통해 사후조사를 완료한 대학생들은 가설로 제시한 방향과는 반대로 사전조사에 비해 그들의 첫 경험 연령을 낮추고, 지금까지 경험한 성관계 파트너의 수는 늘리고, 현재 이성과의 성적 행위는 오히려 증가시켜 보고하는 경향을 보였다.

이와 같은 이유로 세 종류의 상이한 서베이 모드에 따라 응답의 사회적 바람직함 편향을 감소시키고 보다 진솔한 응답을 이끌어내는 데 있어 차이가 있을 것이라는 연구가설은 본 연구에서 활용한 일부의 서베이 주제와 관련해서만 부분적으로 지지되었다. 그러나 지금까지 논의한 서베이 주제의 민감성 수준을 고려하여 추가적으로 실시한 분석의 결과, 대부분의 주제에 있어 민감성 수준이 높은 집단이 낮은 집단에 비해 사회적 바람직함 편향의 작용으로 인해 발생하는 사전-사후 조사 간의 응답의 변화량이 많다는 것을 확인하였다. 이러한 결과는, 비록 본 연구에서는 표본 수의 제한으로 인해 각 서베이 질문을 민감한 것으로 인식한 집단만을 대상으로 서베이 모드가 미치는 효과를 통계적으로 분석하지 못한 한계가 있으나, 질문의 민감성 수준이 높아질수록 응답을 사회적으로 바람직한 방향으로 왜곡하는 정도가 커지는 경향이 있다는 것을 확인한 것이다. 이러한 연구결과와 관련하여, 향후 관련 연구 및 후속 연구에서는 무엇보다도 서베이 주제 및 질문들의 민감성 수준을 실험참가자 집단과 동일한 특성을 가진 표본을 통해 사전조사하고 그 결과를 바탕으로 신중히 실험자극물을 디자인할 필요성이 있다.

또한 향후 관련 연구에서는 실험참가자를 선정함에 있어서도 실험자극물에 대한 민감성 수준이 높은 대상자를 선별하고 각 실험집단별로 충분한 표본 수를 확보하는 것이 바람직할 것이다. 본 연구에서는 실험참가자들의 탈락률을 사전에 충분히 고려하지 못한 측면이 있었으며, 제한된 규모의 실험참가자들 중에서도 성적 경험 등 본 연

구에서 민감한 주제로 다른 내용을 실제로 경험한 응답자는 더욱 제한적이었다. 이처럼 표본의 크기가 연구결과의 신뢰성을 담보하기 어려운 수준임에도 불구하고, 본 연구에서는 서베이 모드에 따라 민감한 질문에 대한 응답의 변화 패턴을 이해하는 데 도움이 될 수 있는 일종의 탐색적 성격의 자료로 제시하였으나, 향후 후속연구에서는 개선되어야 할 부분이다.

아울러 본 연구의 실험 설계상의 제한점으로 지적될 수 있는 문제점으로서, 실험 조건을 집단별로 체계적으로 조작하고 엄격한 통제 상황에서 가외변인들을 통제하기 보다는 실험조건, 즉 서베이 모드에 따라 독립변인들의 수준(면접원 및 제3자 존재 여부)이 자연스럽게 조작화될 것이라는 전제를 바탕으로 수행된 일종의 의사실험(quasi experiment)이라는 것이다. 문제가 되는 부분은 면접원 존재 여부의 조작은 자기기입식 모드인 온라인 서베이와 면접원 진행방식인 대면 또는 전화 면접 모드를 실험조건으로 제시함으로써 자연스럽게 조작된 것으로 볼 수 있으나, 제3자 존재 여부는 본 연구에서 체계적으로 조작하지 못했다는 것이다. 그러나 일반적인 실제 서베이 상황과 마찬가지로 본 연구의 실험 상황에서도 제3자 존재 여부 및 그에 따른 프라이버시 수준에 대한 통제력의 정도는 무작위로, 그러나 서베이 모드 간에는 일정 정도 체계적으로 발생할 것을 전제로 연구를 진행하였다. 비록 주요 변인에 대한 체계적 조작 및 통제가 결여된 의사실험 연구의 전형적인 한계가 있으나, 대신 인위적 변인 조작을 통한 연구결과의 외적 타당성 감소의 문제는 일정 정도 해소된 것으로 볼 수도 있다. 즉 본 연구에서 확인한 세 가지 서베이 모드들 간의 응답 차이가 논리적 근거로 제시된 제3자 존재 여부에 따라 발생한 것이라고 단정할 수 있는 근거는 희박하지만, 제3자 존재 여부를 조작하고 엄격히 통제하는 것이 현실적으로 거의 불가능한 조건에서 오히려 실제 서베이 상황에서 발생 가능한 서베이 모드 간의 차이의 정도를 간접적으로 확인할 수 있었다고 본다.

또한 서베이 모드에 따른 응답의 변화량 및 방향, 즉 응답의 왜곡 정도를 실증적으로 비교하기 위해 사전-사후 실험결과를 비교하는 경우, 비교의 준거가 되는 사전조사의 응답을 수집하기 위해 활용된 서베이 모드의 효과에 대한 검증작업도 이루어져야 할 것으로 제안한다. 본 연구에서는 집단적 상황에서 설문지(paper-and-pencil questionnaire)를 통해 수집한 응답을 기준으로, 이후 대면 면접, 전화 면접, 그리고 온라인 서베이 방식으로 재차 수집한 응답을 비교함으로써 각 서베이 모드별 응답의 정확성을 추론하였다. 그러나 실제로는 설문지를 통한 사전조사의 응답이 사후조사에서

대면 및 전화 면접 그리고 온라인 서베이를 통해 수집된 응답보다 더욱 진솔한 것인지 또는 오히려 더욱 왜곡된 것인지는 확인이 불가능하다. 즉 민감한 정보를 수집하는 데 있어 상이한 서베이 모드의 효과를 검토한 대부분의 선행연구와 마찬가지로, 본 연구도 민감한 행동에 대한 응답이 많으면 많을수록 더욱 정확한 방법이라는 가정을 전제로 했으나, 이것은 여전히 하나의 확인되지 않은 가정일 뿐이다. 이러한 문제를 해결하기 위해, 향후 연구에서는 적어도 사후조사에서도 설문지를 통해 응답하는 통제집단을 포함하여 과연 동일한 설문지 방식으로 수집한 응답의 경우 사전-사후 조사 기간 동안의 변화가 없는지를 확인하고, 또는 변화가 있다면 그 변화된 정도를 다른 서베이 모드를 통해 수집한 응답에도 동일하게 적용함으로써 모드 간 비교를 객관적으로 할 수 있을 것이다.

비록 본 연구에서 검증하고자 했던 가설은 부분적으로만 지지되었으나, 개인의 사생활과 관련된 민감한 서베이 질문에 대한 응답은 서베이 모드에 의해 영향을 받는 것이 확인되었다. 그리고 사생활 침해 우려가 상대적으로 적은 자기기입식 온라인 서베이가 응답자들로부터 보다 솔직한 응답을 유도하는 반면, 기타 집단 설문지 방식이나 면접원이 개입하는 대면 및 전화 면접 방식은 응답자들로 하여금 프라이버시에 대한 우려 및 그로 인한 왜곡된 응답을 유도할 수 있다. 따라서 사회조사에 있어 서베이가 주제가 민감한 것일 경우, 온라인 서베이 모드가 가장 바람직한 것으로 제시될 수 있으나, 온라인 서베이는 대표성 있는 표본을 확보하는 데 있어 상대적으로 취약한 측면이 있다.

이에 본 연구는 민감한 주제와 관련한 서베이 수행을 위해 전화와 온라인 서베이 방식을 병행하는 듀얼 모드 서베이(dual mode survey)를 대안으로 제안한다. 그러나 이중 매체를 활용하여 조사를 진행하는 혼합 모드 서베이(mixed mode survey)는 표본 간 동질성 및 대표성 확보 그리고 이질적 방법을 통해 수집한 자료의 통합 및 분석 방법상의 문제 등을 신중히 검토해야 한다(장윤재 등 2011). 단순히 실무적인 차원에서 비용효율성만을 고려하여 이중의 서베이 모드를 병행하는 것은 바람직하지 않다. 따라서 본 연구에서 제안하는 전화와 온라인의 듀얼 모드 서베이는 대표성 있는 표본을 확보하기 위한 표집 단계와 데이터 수집 단계를 구분하여 각 단계에서 가장 효율적인 방법을 적용하는 것이다. 즉, 기존 RDD 방식의 전화 면접을 통해 대표성 있는 표본을 선정하되, 조사에 참가하는 응답자들이 보다 진솔하고 정확한 응답을 제공할 수 있도록 설문은 온라인 방식으로 제공하는 것이다. 아무쪼록 본 연구가 서베이 응답의

정확성을 제고하기 위한 이러한 새로운 시도 및 노력에 조금이라도 일조하는 계기가 되기를 바란다.

## 참고 문헌

- 김교헌·조성겸·권선중·이동형. 2010. 《사행산업 이용실태 조사》. 사행산업통합감독위원회.
- 김백애라·김정아·이목소희. 2007. 《십대 청소년의 성격형 및 성태도에 관한 실태조사》. YMCA 아하! 청소년문화센터.
- 박용치. 2000. “자료수집 양식의 응답효과.” 《조사연구》 1(1): 1-17.
- 신광우·김동일·정여주. 2011. 《인터넷중독 진단척도 고도화(3차) 연구》. 한국정보화진흥원.
- 심수정·서홍관·김철환·이성희·김윤진·박민선·김현주. 2003. “요 중 니코틴 검출을 이용한 청소년 흡연 자가보고의 정확도.” 《가정의학회지》 24: 375-383.
- 이경옥·김민화·김승옥·김혜수. 2006. 《유아동 및 청소년의 인터넷 게임중독 척도 개발 연구》. 한국정보문화진흥원.
- 이윤석·이지영·이경택. 2008. “온라인 조사의 응답오차에 대한 연구: 설문 응답 시간과 응답 성실성의 관계.” 《조사연구》 9(2): 51-83.
- 장운재·조성겸·송종현·김옥태·박민규. 2011. “웹기반 조사에서 척도의 크기가 응답의 타당도, 신뢰도, 응답 편이성에 미치는 영향.” 《조사연구》 12(3): 1-23.
- 허은주. 2004. 《대학생의 심리 사회적 성숙 변인과 성태도 및 성지식과의 관계연구》. 인하대학교 박사학위논문.
- 허태균·김현철·박정열. 2009. 《‘전국민대상 대규모 도박이용실태 조사’ 연구용역 위한 사전조사 기획 연구 용역》. 고려대학교 심리학과.
- Aquilino, W.S. 1994. “Interview Mode Effects in Surveys of Drug and Alcohol Use: A Field Experiment.” *Public Opinion Quarterly* 58(2): 210-240.
- Aquilino, W.S., D.L. Wright, and A.J. Supple. 2000. “Response Effects Due to Bystander Presence in CASI and Paper-and-pencil Surveys of Drug Use and Alcohol Use.” *Substance Use and Misuse* 35(6-8): 845-867.
- Babbie, E. 2001. *The Practice of Social Research*(9th ed.). Belmont, CA.: Wadsworth.
- Belli, R.F., M.W. Traugott, and M.N. Beckmann. 2001. “What Leads to Voting Overreports? Contrasts of Overreporters to Validated Voters and Admitted Nonvoters in the American National Election Studies.” *Journal of Official Statistics* 17(4): 479-498.

- Brener, N.D., D.K. Eaton, L. Kann, J.A. Grunbaum, L.A. Gross, T.M. Kyle, and J.G. Ross. 2006. "The Association of Survey Setting and Mode with Self-reported Health Risk Behaviors among High School Students." *Public Opinion Quarterly* 70(3): 354-374.
- Brittingham, A., R. Tourangeau, and W. Kay. 1998. "Reports of Smoking in a National Survey: Self and Proxy Reports in Self- and Interviewer-Administered Questionnaires." *Annals of Epidemiology* 8(6): 393-401.
- Catania, J.A., D.R. Gibson, D.D. Chitwood, and T.J. Coates. 1990. "Methodological Problems in AIDS Behavioral Research: Influences on Measurement Error and Participation Bias in Studies of Sexual Behavior." *Psychological Bulletin* 108(3): 339-362.
- Cook, C., F. Heath, and R.L. Thompson. 2000. "A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys." *Educational and Psychological Measurement* 60(6): 821-836.
- Corkrey, R. and Parkinson, L. 2002. "A Comparison of Four Computer-Based Telephone Interviewing Methods: Getting Answers to Sensitive Questions." *Behavioral Research Methods, Instruments, & Computers* 34(3): 354-363.
- Currivan, D.B., A.L. Nyman, C.F. Turner, & L. Biener. 2004. "Does Telephone Audio Computer-Assisted Self-Interviewing Improve the Accuracy of Prevalence Estimates of Youth Smoking? Evidence from the UMASS Tobacco Study." *Public Opinion Quarterly* 68(4): 542-564.
- de Leeuw, E. and J. van der Zouwen. 1988. "Data Quality in Telephone and Face to Face Surveys: A Comparative Meta-Analysis." in R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg(eds.), *Telephone Survey Methodology* (pp. 283-299). New York: Wiley.
- Epstein, J.F., P.R. Barker, and L.A. Kroutil. 2001. "Mode Effects in Self-Reported Mental Health Data." *Public Opinion Quarterly* 65(4): 529-549.
- Ferris, J. and H. Wynne. 2001. *The Canadian Problem Gambling Index: Final Report*. Toronto, ON: Vanadian Centre on Substance Abuse.
- Fujii, E.T., M. Hennessy, and J. Mak. 1985. "An Evaluation of the Validity and Reliability of Survey Response Data on Household Electricity Conservation." *Evaluation Review* 9(1): 93-104.
- Gfroerer, J. and A. Hughes. 1992. "Collecting Data on Illicit Drug Use by Phone." in C. Turner, J. Lessler and J. Gfroerer (eds.), *Survey Measurement of Drug Use:*

- Methodological Studies*(pp. 277–295). Washington D.C.: U.S. Government Printing Office.
- Henson, R., C.F. Cannell, and A. Roth. 1978. “Effects of Interview Mode on Reporting of Moods, Symptoms, and Need for Social Approval.” *Journal of Social Psychology* 105(1): 123–129.
- Holbrook, A.L., M.C. Green, and J.A. Krosnick. 2003 . “Telephone Versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias.” *Public Opinion Quarterly* 67: 79–125.
- Holtgraves, T., J. Eck, and B. Lasky. 1997. “Face Management, Question Wording, and Social Desirability.” *Journal of Applied Social Psychology* 27(18): 1650–1671.
- Johnson, T.P., J.G. Hougland Jr., and R.R. Clayton. 1989. “Obtaining Reports of Sensitive Behaviors: A Comparison of Substance Use Reports from Telephone and Face-to-Face Interviews.” *Social Science Quarterly* 70(1): 174–183.
- Johnston, L.D. and P.M. O'Malley. 1997. “The Recanting of Earlier Reported Drug Use by Young Adults.” in L. Harrison and A. Hughes(eds.), *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*(pp. 59–80). Rockville, MD: National Institute on Drug Abuse.
- Jones, E.F. and J.D. Forrest. 1992 . “Underreporting of Abortion in Surveys of U.S. Women: 1976 to 1988.” *Demography* 29(1): 113–126.
- Juster, F.T. and P.J. Smith. 1997. “Improving the Quality of Economic Data: Lessons from the HRS and AHEAD.” *Journal of the American Statistical Association* 92(440): 1268–1278.
- Kays, K., K. Gathercoal, and W. Buhrow. 2012. “Does Survey Format Influence Self-Disclosure on Sensitive Question Items?” *Computers in Human Behaviors* 28(1): 251–256.
- Krosnick, J.A. 199 . “Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys.” *Applied Cognitive Psychology* 5(3): 213–236.
- Krysan, M. 1998 . “Privacy and the Expression of White Racial Attitudes: A Comparison across Three Contexts.” *Public Opinion Quarterly* 62(4): 506–544.
- Lemmens, P., E.S. Tan, and R.A. Knibbe. 1992. “Measuring Quantity and Frequency of Drinking in a General Population Survey: A Comparison of Five Indices.” *Journal of Studies on Alcohol and Drugs* 53(5): 476–486.

- Lessler, J.T. and J.M. O'Reilly. 1997. "Mode of Interview and Reporting of Sensitive Issues: Design and Implementation of Audio Computer-Assisted Self-Interviewing." in L. Harrison and A. Hughes(eds.), *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*(pp. 366-382). Rockville, MD: National Institute on Drug Abuse.
- Mangione, T.W., R. Hingson, and J. Barrett. 1982. "Collecting Sensitive Data: A Comparison of Three Survey Strategies." *Sociological Methods & Research* 10(3): 337-346.
- Moore, J.C., L.L. Stinson, and J.E.J. Welniak. 1999. "Income Reporting in Surveys: Cognitive Issues and Measurement Error." in M.G. Sirken, D.J. Herrmann, S. Schechter, N. Schwarz, J.M. Tanur, & R. Tourangeau(eds.). *Cognition and Survey Research*(pp. 155-173). New York: Wiley.
- Moskowitz, J.M. 2004. "Assessment of Cigarette Smoking and Smoking Susceptibility among Youth: Telephone Computer-Assisted Self-Interviews Versus Computer-Assisted Telephone Interviews." *Public Opinion Quarterly* 68(4): 565-587.
- Newman, J.C., D.C. Des Jarlais, C.F. Turner, J. Gribble, P. Cooley, and D. Paone. 2002. "The Differential Effects of Face-to-Face and Computer Interview Modes." *American Journal of Public Health* 92(2): 294-297.
- Patrick, D.L., A. Cheadle, D.C. Thompson, P. Diehr, T. Koepsell, and S. Kinne. 1994. "The Validity of Self-Reported Smoking: A Review and Meta-Analysis." *American Journal of Public Health* 84(7): 1086-1093.
- Presser, S. and L. Stinson. 1998. "Data Collection Mode and Social Desirability Bias in Self-Reported Religious Attendance." *American Sociological Review* 63(1): 137-145.
- Rasinski, K.A., A.K. Baldwin, G.B. Willis, and J.B. Jobe. 1994. "Risk and Loss Perceptions Associated with Survey Reporting of Sensitive Behaviors." The Annual Meeting of the American Statistical Association, Toronto, Canada.
- Rosenfield, P., S. Booth-Kewly, and J.E. Edwards. 1996. "Response on Computer Surveys: Impression Management, Social Desirability, and the Big Brother Syndrome." *Computers in Human Behavior* 12(2): 263-274.
- Schober, S., M.F. Caces, M. Pergamit, and L. Branden. 1992. "Effects of Mode of Administration on Reporting of Drug Use in the National Longitudinal Survey." in C. Turner, J. Lessler, and J. Gfroerer(eds.), *Survey Measurement of Drug Use: Methodological Studies*(pp. 267-276). Rockville, MD: National Institute on Drug Abuse.
- Singer, E. and S. Presser. 2007. "Privacy, Confidentiality, and Respondent Burden as Factors

- in Telephone Survey Nonresponse.” in J.M. Lepkowski, C. Tucker, J.M. Brick, E.D.d. Leeuw, L. Japec, P.J. Lavrakas, M.W. Link, and R.L. Sangster(eds.), *Advances in Telephone Survey Methodology*(pp. 447–470): John Wiley & Sons, Inc.
- Singer, E., von D.R. Thurn, and E.R. Miller. 1995. “Confidentiality Assurances and Response: A Quantitative Review of the Experimental Literature.” *Public Opinion Quarterly* 59(1): 66–77.
- Stulginskas, J.V., R. Nerreault, and I.B. Pless. 1985. “A Comparison of Observed and Reported Restraint Use by Children and Adults.” *Accident Analysis & Prevention* 17(5): 381–386.
- Sykes, W. and M. Collins. 1988. “Effects of Mode of Interview: Experiments in the UK.” in R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls and J. Waksberg (eds.), *Telephone Survey Methodology*(pp. 301–320). New York: Wiley.
- Tourangeau, R., L.J. Rips, and K. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge, England: Cambridge University Press.
- Tourangeau, R., and T.W. Smith. 1996. “Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context.” *Public Opinion Quarterly* 60(2): 275–304.
- Tourangeau, R., T.W. Smith, and K.A. Rasinski. 1997. “Motivation to Report Sensitive Behaviors in Surveys: Evidence from a Bogus Pipeline Experiment.” *Journal of Applied Social Psychology* 27(3): 209–222.
- Tourangeau, R. and T. Yan. 2007. “Sensitive Questions in Surveys.” *Psychological Bulletin* 133(5): 859–883.
- Turner, C.F., J.T. Lessler, and J. Devore. 1992. “Effects of Mode of Administration and Wording on Reporting of Drug Use.” in C. Turner, J. Lessler and J. Gfroerer(eds.), *Survey Measurement of Drug Use: Methodological Studies*(pp. 177–220). Rockville, MD: National Institute on Drug Abuse.
- Wimmer, R.D. and J.R. Dominick. 2000. *Mass Media Research: An Introduction* (6th ed.). Belmont, CA: Wadsworth Publishing Company.
- Wyner, G.A. 1980. “Response Errors in Self–Reported Number of Arrests.” *Sociological Methods & Research* 9(2): 161–177.

