

청소년 조사연구의 재검토

— 조사영역, 표집방법, 보고양식의 문제 —

박 광 배*

- I. 서 론
- II. 조사연구의 개관
- III. 편표집의 문제
- IV. 일차원적 자료분석의 문제
- V. 민감한 영역에 대한 반응의 문제

I. 서 론

청소년기는 아동기와 성인기를 연결하는 발달적 전환기로서 육체적, 인지적, 그리고 행동적으로 독특하고도 격심한 변화를 경험하는 시기이다. 청소년들은 이러한 개인적, 신체적 변화와 그들이 경험하는 사회적 역할의 변화가 서로 부조화함으로써 자신과 사회에 대해서 매우 불안정한 상태에 놓여있게 된다(Coleman, 1992). 예를 들면 청소년들은 국가가 위협에 처해있을 때는 목숨을 바쳐 국가를 위해 희생하는 것이 숭고한 애국정신이라는 가치를 주입받으면서도, 국정에 참여할

수 있는 참정권 및 투표권은 부여받지 못하고 있다(Warren-Adamson, 1992). 청소년들이 경험하는 이러한 사회심리적 부조화와 불안정은 성충동의 억제 및 성역할에 대한 인식, 자아정체감 *egoidentity*의 확립, 가족관계, 친구관계, 학업 등에서 많은 어려움과 혼란을 유발할 수 있고, 이러한 혼란은 청소년들에게 잠재적 위험요인으로 상존하게 된다. 실제로 미국의 경우 청소년들의 죽음을 초래하는 요인들 중에 가장 빈번하게 나타나는 것들이 사고, 타살, 그리고 자살이라고 알려져 있다(Singer, Singer, and Anglin, 1993). 그런데 청소년들의 죽음을 유발하는 이러한 중요한 요인들은 대부분 부주의, 비행, 스트레스 등의 사회심리적, 그리고 발달적 문제들과 연관되어 있고, 따라서 어느 정도 예방가능한 것들이라는 사실에 주목할 필요가 있다.¹⁾ 그런데 Coleman (1992)에 의하면 청소년 정책이나 법규 등은 청

*충북대학교 심리학과 교수

1) 모든 부주의, 비행, 스트레스가 사회심리적 요인들에 의해 생기는 것이라는 주장을 하는 것은 아니다. 예를 들어 주의력 결핍증 *Attention Deficit Hyperactivity Disorder*은 생리적인 이상에 의해 유발된다는 견해도 광범위하게 받아들여지고 있다(Cooley and Morris, 1990 ; Shaywitz and Shaywitz, 1991). 여기에서는 일반적인 경향을 기술하고 있다.

소년 혹은 10대 아동들의 행동이나 도덕관에 대한 근거없는 속단이나 편견에 기초하여 결정되고 제정되는 경우가 많다고 한다. Coleman은 예를 들어 대다수의 청소년들은 부모들을 존경하고 그들의 가치를 수용하며 부모들과 긍정적이고 조화로운 관계를 성공적으로 발전시켜 나가고 있음에도 불구하고, 십대들과 그들의 부모들은 소위 세대차에 의해 서로 갈등하고 상충하고 있다는 편견이 만연하고 있다고 주장한다. 한국에서도 청소년들이 부모들과 원만한 관계를 가지고 있다고 응답하는 비율이 87%, 학교를 좋아한다는 응답이 77%에 이르고 있다는 보고가 있다('91 MBC 청소년 백서, 1991).²⁾ 따라서 청소년들의 사회적응과 인격형성과정에서 발생할 수 있는 제문제들을 파악하고, 그들의 균형된 성장과 성숙을 도와주기 위해서는 예방차원의 기초자료를 비롯하여, 각 문제들을 유발하거나 억제하는 원인들을 규명하는 체계적인 연구노력이 필요시된다.

II. 조사연구의 개관

청소년정책을 담당하는 국가기관이나 청소년의 복지증진을 도모하고자 설립된 사회기관들은 보통 조사자료survey data를 이용하여 청소년의 실태와 시급한 문제들을 파악하고, 예방과 치유를 위한 정책을 수립, 시행하며, 그 시행된 정책이 거둔 성과를 다시 조사방법을 이용하여 평가하는 경우가 많다. 조사연구라고 하면 특정한 모집단의 성향을 추론할 수 있는 통계치를 산출할 목적으로 일정한 절차를 거쳐서 선발된 사람들에게 사

전에 계획된 질문을 하여 정보를 수집하는 것을 말한다(Fowler, Jr., 1988). 이러한 조사연구 이외에도 정부기관이나 공공기관이 보유하고 있는 공식적 자료를 이용하거나, 혹은 통제된 실험연구 등을 통하여 문제파악이나 정책평가를 하는 경우도 많다. 그런데 특별히 조사연구방법을 통하여 현상을 파악하고 규명하는 경우에는 일반적으로 다음의 네가지 이유를 충족시키기 위해서이다. 첫째, 확률표집probability sampling으로부터 구해진 자료는 모집단에 대하여 보다 신뢰성있는, 혹은 모집단에 대하여 보다 일반화할 수 있는 추론을 가능하게 해준다. 둘째, 표준화된 측정도구를 사용하므로써 모든 응답자에게 동일한 질문을 동일한 방식으로 실시할 수 있어서 일관된 자료해석이 가능하다. 셋째, 기본적으로 조사는 다른 수단에 의해서 정보를 수집할 방법이 없는 경우 활용할 수 있는 연구방법이다. 넷째, 연구자가 특정한 분석기법(예를 들면 상관계수 등)을 이용해야 하는 필요가 있는 경우 조사방법에 의해 자료수집을 하는 것이 편리하거나 필요한 경우가 있다.

조사연구는 일반적으로 세가지의 요소로 구성되고 있다. 첫째는 표집sampling으로서 어떤 시간, 장소, 그리고 대상을 선발하여 자료를 수집할 것인지의 여부를 결정하는 과정이다. 이때 가장 핵심적으로 고려되어야 할 것은 여건이 허락하는 범위내에서 연구의 관심이 되는 모집단의 성향을 정확히 추론하게 해주는 표집방법이 무엇인지를 파악하는 일이다. 조사연구를 구성하는 두번째 요소는 설문도구를 설계, 제작하는 것이다. 설문항목들은 조사원이 응답자에게 쓰여진데로 낭독을

2) 그러나 한국청소년의 가정생활에서의 만족도를 다른 나라의 청소년들과 비교한 자료(한국개발조사연구소, 1991)에 의하면 '만족한다'는 응답을 한 사람이 70%를 넘는데도 불구하고 비교된 11개 나라의 청소년들 중에는 최하위를 차지하고 있다. 또한 같은 자료에 의하면 학교생활에 만족한다고 응답한 한국의 청소년은 63.5%로서 과반수를 상회하지만 비교된 11개 나라중에는 10위를 차지하여 중국의 청소년들 다음으로 학교생활에 불만족하는 청소년이 많은 것으로 나타났다.

하여도 어색하거나 왜곡되어 전달되지 않고, 응답자가 그 의미를 자의로 해석하거나 확대, 혹은 축소 해석하지 않도록 명확하게 기술되어야 한다. 세 번째 요소는 면담 *interviewing*이다. 물론 모든 조사연구에서 면담을 행하는 것은 아니지만 면담 방식이 이용되는 경우 면담자의 영향이 자료에 개입하지 않도록 각별한 주의를 기울여야 한다.

그 밖의 조사연구의 핵심적인 방법과 일반적으로 주의해야 할 점들에 대해서는 Fowler, Jr. (1988), Belson(1981), Dillman(1978), Guenzel, Berkman, and Cannell(1983), Turner and Martin (1984) 등을 참고하면 좋다. 본 논문에서는 청소년 연구를 위한 조사연구에서 흔히 간과되기 쉬운 문제들을 논하고 그 해결책을 모색해보고자 한다. 특히 편의표집 *convenience sample*의 문제³⁾, 자료의 일차원적 분석에서 고려되어야 할 점, 그리고 성경험 혹은 비행 등과 같이 민감하거나 사회적으로 바람직하지 않은 영역에 대한 조사에서 중요하게 고려되어야 할 점을 중점적으로 다루어보고자 한다.

III. 편의표집의 문제

과학적인 연구는 최소한 세가지의 기준을 어느 정도 충족시켜야 한다. 그 기준들은 연구의 내적 타당성 *internal validity*, 외적타당성 *external validity*, 그리고 유용성 *usefulness*이다(Cook and Campbell, 1979; Brinberg and McGrath; 1985; Dominowski, 1980).⁴⁾ 내적타당성이란 연

구의 결과에 대한 일관된 해석의 가능성을 말하는 것으로서, 일반적으로 결과해석에 오염 *confounding*을 유발하는 가외변인 *extraneous variable*이 존재하는지의 여부를 의미한다. 그러한 가외변인이 존재하면 얻어진 연구결과가 어떻게 도출되었는지가 불분명해지고, 따라서 연구주제나 가설에 대하여 명확한 결론을 내릴 수 없게 된다. 둘째로, 외적타당성은 특정한 연구로부터 얻어진 자료, 결론, 해석 등이 다른 때, 다른 장소, 다른 대상에게도 적용되고 해당되는 것인지의 여부를 말한다. 이 외적타당성을 '일반화 가능성 *generalizability*'이라고 부르기도 한다. 물론 어떤 연구를 막론하고 그 결과가 모든 시간과 공간을 초월하여 일반화되지는 않는다. 그러나 연구목적과 궁극적 목표에 부합하는 한계 내에서 과학적 연구는 적어도 일정한 기간동안, 합리적인 범위의 다른 장소와 대상에 대하여 그 결과와 결론이 유지되고 적용될 수 있어야 한다. 과학적 연구가 충족해야 하는 세번째 기준인 유용성은 학문적 연구의 결과가 실제 인간의 생활과 복지증진에 얼마나 기여할 수 있는가의 문제이다. 다시 말하면 과학에 대한 실용주의적인 기준인 셈이다.

조사연구는 특히 일반화 가능성 (외적타당성) 과 유용성에 많은 비중을 주는 연구방법이다. 일반화 가능성과 유용성은 동일한 개념은 아니지만 서로 밀접한 관련을 가진다. 어떤 연구의 일반화 가능성이 높으면 그 연구의 결과를 적용할 수 있는 영역의 범위가 넓다는 것을 의미하고, 따라서 유용성이 증가하게 마련이다. 조사연구가 가지는

3) 때로는 이것을 '임의표집'이라고 번안하기도 한다. 그러나 임의라는 용어는 무작위성 *randomness*의 의미도 함께 가지고 있는 까닭에 다소 혼동을 유발할 우려가 있다고 사료된다.

4) 여기에서의 연구타당성이라는 개념을 검사 *test* 혹은 척도 *scale*의 타당도 개념과 혼동해서는 안된다. 연구타당성에 관한 이 세가지 기준 이외에도 학자들에 따라서 '통계적 결론의 타당성 *statistical conclusion validity*' 과 '구성개념 타당성 *construct validity*'을 기준으로서 첨가하기도 한다. 이러한 부가적 연구타당성 개념들에 대해서는 Cook and Campbell (1979)을 참고할 것.

일반화 가능성의 최대한계는 어떻게 결정할 수 있을까? 일반적으로는 연구자가 연구를 수행하기 전에 설정해 놓은 모집단 혹은 전집 *population*이 그 연구의 일반화 가능성의 최대한계라고 말할 수 있다. 예를 들어 연구자가 1993년의 한국청소년을 모집단으로 설정하고, 그들의 성경험을 파악하기 위하여 1500명을 1993년도에 표집하였다면, 이 표본으로부터 구해진 통계치가 일반화될 수 있는 최대범위는 1993년도(시간)의 한국(공간)의 청소년들(대상)이다.

그러나 모든 조사연구의 결과가 주어진 최대범위로 충실히 일반화될 수 있는 것은 아니다. 과연 1500명의 표본으로부터 구해진 통계치를 모집단의 성향으로 간주해도 좋은지, 다시 말하면 주어진 일반화의 최대한계 내에서 얼마나 일반화할 수 있는지의 여부는 또한 그 표본이 모집단에 대한 ‘대표성 *representative*’을 가지는지의 여부에 따라 좌우된다. 대표성의 개념은 통계적 확률 이론에 근거한 개념이다. 만약 표본이 단순 무작위 표집 *simple random sampling* 혹은 다른 확률적 표집방법에 의해 추출되었다면 그 표본은 모집단에 대해 대표성을 가지고 있을 확률이 높다.⁵⁾ 따라서 일반화 가능성의 개념도 확률적 개념이다. 대표성과 일반화 가능성이 확률적 개념인 까닭에 특정한 조사연구에 대해서 일반화 가능성 혹은 대표성이 전무하다거나 완벽하다고 말하는 것은 지나치게 극단적인 평가라고 생각되고, 모든 조사연구는 일반화 가능성의 우월에 있어서 정도의 차이를 가진다고 생각된다.

그러나 모든 조사연구가 확률적 표집방법에 의해 자료를 수집할 수 있는 것은 아니다. 확률적 표집방법은 자금, 인력, 시간 및 활용할 수 있는

설비가 충분히 확보되어야 가능한 경우가 많다. 이러한 여건이 충분하지 못할 때는 소위 ‘편의표집 *convenience sampling*’에 의해 자료수집이 이루어진다. 편의표집이란 연구자가 손쉽게 접근할 수 있는 대상을 상황에 따라 표집하는 것을 말한다. 이러한 편의표집은 확률표집이 아니기 때문에 통계적인 확률개념인 ‘대표성’을 가질 수 없고, 따라서 일반화 가능성도 심각한 제한을 받을 수밖에 없다. 편의표집에 의해 이루어진 조사연구들은 대표성을 가질 수 없으므로 그들을 대표성에 있어서 우월을 가리는 것은 무의미한 일이다. 또한 확률표집의 가정하게 존재하는 여러가지 통계적 지표들, 예를 들면 다음 절에서 설명될 모집단 경향성에 대한 신뢰구간, 을 활용할 수 없다는 단점을 가진다. 따라서 편의표집에 의해 이루어진 조사연구의 일반화 가능성은 다른 기준들에 의해 평가되어야 한다. ‘표본연관성 *sample relevance*’과 ‘표본전형성 *sample prototypicality*’이 그 기준들이다(Sackett and Larson, 1990).

1. 표본연관성

표본연관성이란 의도된 모집단의 핵심적인 구성원과 환경의 특징적 속성이 표본에 반영되고 있는 정도를 말한다. 이 속성을 ‘경계조건 *boundary conditions*’이라고 부른다(Campbell, 1990). 예를 들어, 한국청소년의 성경험에 관한 조사연구를 수행한다면 ‘한국청소년’의 개념을 우선 명확히 정의할 필요가 있다. 만약 한국청소년의 개념을 한국에 거주하는 10대 후반의 사람들로 정의한 후, 서울의 강남지역에서 고등학생들을 편의표집하였다면 이들은 비록 한국청소년을 대표하는 표본은 아니지만 나이에 있어서 정의된 ‘한국청소년

5) 여러가지 확률적 표집방법들에 대해서는 Kalton(1983) 혹은 Yates(1981)를 참고하면 좋다.

년'과 연관성을 가지는 표본이다. 그러나 만약에 표집된 학생들이 중학교 학생들이라면 이들은 연관성을 가지지 않는 표본이다. 말하자면 표본연관성이란 일반화 가능성을 위한 필요조건이지만 충분조건은 아니라고 볼 수 있고, 일반화 가능성의 한계를 보다 좁게 세밀히 규정하는 개념이다. 따라서 편의표집에 의해 수행된 조사연구에서는 그 결과를 일반화할 수 있는 경계조건을 보다 세밀하고 명확하게 기술해야하고, 그 결과의 해석이나 시사점에 대한 논의도 이 경계조건의 한계내에서 이루어져야 한다.

2. 표본전형성

표본전형성이란 표본이 정의된 모집단의 핵심적인 특질을 공유하고(표본연관성), 그와 더불어 핵심적인 모집단의 특질은 아니지만 그 모집단내에서 광범위하게 존재하는 다른 비핵심적 특질들을 공유하는 정도를 말한다. 따라서 표본전형성은 표본연관성을 함축하는 개념이다. 예를 들어서 10대 후반의 한국청소년에 관한 두개의 연구가 있는데 한개는 강남지역의 고등학생들을 표집하였고, 다른 한개는 공장이나 중소기업에서 근무하는 10대 후반의 근로자들을 표집하였다면 두개의 표집은 모두 표본연관성을 가지지만 고등학생 표본이 근로자 표본보다 더 전형성을 가진 표본이다. 왜냐하면 이 연구들에서는 한국청소년의 범위를 정하기 위하여 나이를 핵심적 특질로 삼기는 하였지만 대다수의 한국청소년들이 고등학교에 재학중인 학생들로 구성되어 있기 때문이다. 말하자면 고등학교 교육이 모집단을 규정하는 비핵심적 특질이기는 하지만 그 모집단내에서 광범위하게 존재하는 특질이다. 모집단에 분포된 여러가지 특질들을 언제나 명확히 알 수 없기 때문에 어떤 조사연구가 얼마나 표본전형성을 가지는지의 여부

를 언제나 쉽게 파악할 수 있는 것은 아니다.

3. 삼각도법

삼각도법(triangulation)이란 측량기술에서 도입된 용어로서 흔히 행동과학연구에서 얻어진 결과가 특정한 연구방법을 사용하였기 때문에 발생한 것인지의 여부를 판단하기 위하여 동일한 연구를 다른 방법(e.g., 면접방식, 집단자기보고방식, 우편자기보고방식, 전화면접 등)으로 재실시한 후 처음의 연구와 동일한 결과가 얻어지는지를 검증하는 것을 일컫는다. 만약 그것들이 서로 동일하지 않으면 각기 다른 방법들을 이용하여 얻어진 각기 다른 결과들의 기하학적 평균(geometric mean)을 산출하여 그것을 실제의 상태라고 간주한다. 말하자면 연구결과로부터 방법에 의한 변산(variability)을 통제, 제거하는 방법이다. 또한 이것은 마치 화성과 목성 사이의 거리를 측정하기 위하여 지구와 화성의 거리를 측정하고, 지구와 목성의 거리를 측정한 후, 그 측정치들에 삼각함수를 적용하여 화성과 목성 사이의 거리를 알아내는 것과 흡사한 방법이다.

이 삼각도법의 원리를 편의표집에 의한 조사연구에 적용하므로써 결과의 일반화가능성을 어느 정도 증가시킬 수 있다(Cook and Campbell, 1979). 방법에 의한 변산을 제거하기 위해서 쓰이는 삼각도법을 '방법간 삼각도법(between-methods triangulation)'이라고 부른다면, 이것은 일종의 '방법내 삼각도법(within-method triangulation)'이라고 부를 수 있는 방법이다(Denzin, 1978). 예를 들어서 10대 후반의 한국청소년에 관한 조사연구를 처음에는 고등학생을 대상으로 한 편의표집에 의해 수행하였다면, 그 다음에는 근로청소년을 대상으로한 편의표집을 하여 재실시하는 경우를 말한다. 만약에 두개의 연구결과가

동일하다면 우리는 적어도 교육수준에는 무관하게 우리의 조사결과가 일반화될 수 있을 것이라는 신뢰를 가질 수 있다. 이러한 삼각도법을 써서 얻어진 여러개의 결과들을 취합하고, 그것으로부터 일반화가능성에 대한 판단을 하기 위해서는 흔히 종합분석 *meta-analysis*과 일반화가능도이론 *generalizability theory*이 많이 활용된다. 종합분석에 대해서는 Glass, McGaw, and Smith(1981)과 Hunter, Schmidt, and Jackson(1982) 등을 참고하면 좋고, 일반화가능도이론은 Shavelson and Webb(1991), Cronbach et al. (1972), 그리고 이종성(1988) 등을 참고하면 좋다.

그러나 삼각도법을 활용한 연구에 의해 증가하는 일반화 가능성은 그것이 여러개의 집단에 대해서 여러번 실시되었다고 하더라도 통계적인 확률표집에 근거하여 수행된 연구가 가지는 일반화 가능성과는 성질이 다르다는 사실을 유념하여야 한다. 삼각도법에 의한 일반화는 연구대상이 된 하부집단들 *subpopulations*에 대한 일반화이며, 전체 모집단에 대한 일반화는 아니다. 또한 삼각도법이 적용되는 각 하부집단은 위에서 기술한 표본연관성 및 표본전형성을 가진다는 근거의 제시가 반드시 수반되어야 한다.

IV. 일차원적 자료분석의 문제

청소년 연구를 위한 조사는 단지 호기심에 의해 수행되지는 않는다. 많은 비용과 시간, 그리고 노력을 투입하고 치밀한 계획과 세밀한 검토를 통하여 수집되는 조사자료는 청소년의 복지향상과 정책결정을 위하여 활용될 수 있어야 한다. 이러한 자료의 유용성을 위하여는 우선 실태 혹은 현황파악이 되어야하고, 그러한 실태가 존재하는 이유가 규명되어야 한다.

실태파악을 위한 조사연구는 현상을 간접적으로 관찰하는 것이다. 이러한 조사연구는 흔히 수집된 자료를 각 변인 *variable*에 대하여 일차원적으로 정리하여 제시한다. 예를 들면 만 13~18세 청소년 1500명에게 “자위행위(자신의 성기를 자극하는 행위)를 하고 있는가?” 라는 질문을 한 결과 16.2%가 ‘하고 있다’, 83.8%가 ‘하지 않는다’라는 응답을 하였고, “이제까지 성관계를 가져 본 경험이 있습니까?” 라는 질문을 한 결과 6.3%가 ‘있다’, 93.7%가 ‘없다’라는 응답을 하였다. 이는 자료를 제시하는 경우이다(‘91 MBC 청소년 백서, 1991). 이러한 실태파악을 위한 조사연구를 수행하는 경우 최소한 세가지의 고려해야 할 사항들이 있다. 첫째는 신뢰구간 혹은 오차한계를 설정해야 할 필요이고, 둘째는 무응답 혹은 결측치 *missing value*의 처리문제이며, 셋째는 교차분석 혹은 상관분석의 필요성이다.

1. 신뢰구간의 필요성과 유용성

첫째는 조사연구를 통하여 수집된 자료수치는 그 자체가 유용성을 가지기 보다는 모집단의 성향을 파악하기 위한 수단이며 도구로서의 유용성을 가진다. 즉, 조사연구의 궁극적 목적은 모집단의 성향을 파악하는 것이다. 예를 들어 위에서 인용한 청소년의 성경험 실태에 관한 자료는 표집된 1500명의 성경험 실태를 알아보기 위한 것이 아니라 한국의 청소년 일반의 성경험 실태를 파악하기 위한 것이다. 따라서 앞서 강조되었듯이 조사연구에서는 표집절차와 방식이 매우 중요시된다. 왜냐하면 표본이 모집단을 대표할 수 있다는 확신이 있어야 표본자료로부터 모집단의 성향을 추론할 수 있기 때문이다. 그런데 문제는 표본이 모집단을 대표한다는 것의 의미를 정확히 이해할 필요가 있다. 이때의 ‘대표성’이란 이론적인 대표

성이다. 일반적으로 모집단으로부터 비교적 많은 수의 표본을 무작위로 표집(random sampling)하면, 그때의 표본은 모집단을 대표하는 것으로 간주된다. 그러나 모집단의 성향을 알 수 없으므로 무작위로 추출된 표본이 모집단과 중심경향성(central tendency)에 있어서 동일한지를 검증할 수 없고, 따라서 이에 관한 확신을 가질 수는 없다. 따라서 설혹 표본이 무작위절차에 의해 추출되었다고 하더라도 모집단의 중심경향성에 대한 신뢰구간을 제시하는 것이 매우 중요하다. 신뢰구간이란 표본의 중심경향성 수치를 중심으로 모집단의 중심경향성이 존재할 것으로 추정되는 특정한 범위를 확률적으로 산출해낸 것을 말한다. 예를 들어 자위행위를 한 경험이 있다고 대답한 청소년들이 1500명중에 16.2%였다고 하면, 이 1500명이 한국의 모든 청소년들중에서 무작위로 표집된 사람들이라는 가정하에 95% 신뢰구간을 다음과 같이 구할 수 있다.⁶⁾

$$P \pm 1.96 \frac{\sqrt{P(100-P)}}{\sqrt{N}} = 16.2 \pm 1.96 \frac{\sqrt{1357.56}}{\sqrt{1500}} \\ = 16.2 \pm 1.86$$

이 신뢰구간의 해석은, 모집단 (한국청소년) 전체를 조사한다고 가정하면 그들 중에서 자위행위의 경험이 있는 청소년의 실제 비율은 $16.2\% - 1.86\% = 14.34\%$ 에서 $16.2\% + 1.86\% = 18.06\%$ 사이일 확률이 95%라는 것이다. 반면에 성경험이 있다고 대답한 청소년은 1500명중에 6.3%였는데 이에 대한 95% 신뢰구간을 구하면 $6.3\% \pm 1.23\%$ 로서 모집단에서의 그 비율은 5.07%에서 7.53% 사이에 존재할 확률이 95%이다.

이 신뢰구간이 크면 모집단의 중심경향성에 대한 불확실성이 커진다. 왜냐하면 그 구간내에서 어느 지점에 모집단의 실제 중심경향성이 존재하는지가 애매해지기 때문이다. 위의 예에 의해 알 수 있듯이 추론하고자하는 중심경향성이 %인 경우 50% 근처의 비율지수들은 0% 혹은 100% 근방의 비율지수들보다 신뢰구간이 커지는 경향이 있다. 또한 표본의 크기가 작을수록 이 신뢰구간은 커진다. 예를 들어 자위경험이 있는 청소년의 비율을 산출하고자 1500명이 아닌 150명을 무작위로 표집하여 그중에서 16.2%가 '있다'라고 응답한 자료를 구하였다고 가정하자. 이 경우 모집단의 비율을 위한 95% 신뢰구간은 위의 공식에 의하여 $16.2\% \pm 5.90\%$ 가 되고 이것은 한국의 전체 청소년중 자위경험이 있는 청소년의 비율은 10.3%에서 22.1% 사이의 어딘가에 위치한다는 것을 의미한다. 이와 같이 표본의 비율이 50%에 가까워져나, 표본의 크기가 작으면 신뢰구간이 증가하고, 따라서 모집단에 대한 추론이 불확실해진다. 모집단에 대한 추론이 불확실해진다는 것은 표본자료부터 산출된 비율이 모집단에서의 그 비율에 근사할 것이라는 확신이 감소한다는 것을 의미한다. 모집단의 성향에 대한 확신이 감소한다면 실태파악을 위한 그 조사연구의 유용성도 감소할 수 밖에 없다. 따라서 실태파악을 위한 조사자료의 유용성을 높이기 위해서는 신뢰구간이 작아지도록 주의를 기울이는 것이 대단히 중요하다.⁷⁾ 그런데 표본의 비율과 표본의 크기 중에서 비율은 연구자가 통제할 수 없는 것이므로 모집단에 대한 추론의 확실성과 유용성을 증가시키기

6) 수식에서 P는 표본으로부터 구해진 비율이고, N은 전체 표본사레수 혹은 표본의 크기이다. 이 수식에서 $\frac{\sqrt{P(100-P)}}{\sqrt{N}}$ 부분을 "표준오차(standard error)"라고 부른다.

7) 여기에서의 모든 논의는 물론 표본이 모집단을 대표할 수 있는 무작위 표본이라는 것을 가정한다. 이 대표성이 훼손되었을 때는

위해서 조작하는 것이 불가능하지만, 표본의 크기는 연구자가 어느 정도 통제할 수 있고 그것을 크게 할수록 모집단에 대하여 보다 확신할 수 있는 추론이 가능해진다.

표본의 크기를 통제하여 신뢰구간을 조정하는 것이 가능하다는 것은 연구자가 연구수행을 위하여 자료를 수집하기 이전에 얼마나 많은 대상들에 대하여 자료를 수집해야하는지를 결정할 수 있게 해주는 결정적인 역할을 한다. 이것을 위해서는 우선 기존 연구나 그밖의 다른 이론적인 추론을 이용하여 모집단의 중심경향성에 대해서 대강의 예측을 한다. 둘째로 모집단에 대한 추론을 위한 표본연구가 실제의 정책결정과 시행을 위하여 의미있는 유용성을 가지기 위해서는 신뢰구간이 어느정도 되어야 할지를 결정한다. 이 두가지가 결정되면 그것들을 위의 신뢰구간을 위한 공식에 대입하여 필요한 N을 산출하면 된다. 예를 들어서 10년전에 수행된 조사연구에서 한국청소년중 10%가 성경험을 가지고 있다는 자료를 발표한 적이 있다고 가정하자. 그러면 10년이 지난 지금은 성관념이 과거보다 더 개방되었고 청소년들의 성충동을 자극하는 환경도 더 풍부해졌으므로 지금은 아마도 그 비율이 최소한 15%는 될 것으로 추론할 수도 있다. 이러한 추론을 검증하고, 정확한 비율을 산출하여 청소년을 위한 성교육의 투자범위를 정하는데 활용하기 위해서 정부로서는 신뢰구간 혹은 오차한계의 범위가 최대한 $\pm 1.5\%$ 를 넘지않는 조사연구를 수행하고자 한다고 가정하자. 이러한 추정과 요구를 충족시키는 표본크기 (N)을 정하기 위해서는 위의 신뢰구간 공식에서 범위를 15

$\% \pm 1.5\%$ 로 정한 후 N을 산출하면 된다.

2. 무응답의 처리

만약 구해진 자료에 '무응답'이 존재하는 경우에는 무응답한 사람들까지 모두 포함시켜서 N을 산출한후 신뢰구간을 구하는 것이 그들을 제외하고 N을 산출한 후 신뢰구간을 구하는 것보다 더 작은 신뢰구간이 구해진다.⁸⁾ 따라서 무응답을 N 산출에 포함하는 것이 그것을 N 산출에서 제외하는 것보다 모집단 추론의 통계적 확실성이라는 점에서는 유리하다. 그러나 이러한 통계적 확실성에도 불구하고, 전자의 방식이 후자의 방식보다 언제나 바람직한 것은 아니다. 성경험의 여부를 예로 들면, 전자(무응답을 N 산출에 포함)는 무응답을 성경험이 없다는 응답과 마치 동격인 것으로 간주하는 것이고, 후자(무응답을 N 산출에서 제외)는 무응답한 피험자가 마치 표집되지 않은 것처럼 취급하는 것이다. 두가지중 어떤 방식을 취하느냐하는 문제는 각 연구의 특성과 목적에 따라 판단하여야할 문제이다. 성경험의 여부를 묻는 질문의 경우에는 무응답이 과연 성경험이 없다는 응답과 동일한 것인지에 대하여 심각한 의문이 제기될 수 있다. 따라서 이런 경우 만약 무응답의 빈도가 크지 않다면, 그래서 무응답을 N 산출에서 제외하여도 신뢰구간이 크게 증가하지 않을 때는 무응답을 N 산출에서 제외한 후 신뢰구간을 산출하는 것이 바람직하다고 사료된다. 다른 경우에는 두가지의 신뢰구간(무응답을 N 산출에 포함시킨 후 구해진 것과 그것을 제외시킨 후 구해진 것)을 모두 제시하는 것이 좋다.

그밖의 모든 유용성 논의가 무의미하다. 또한 모든 응답자가 최선을 다하여 진솔한 응답을 한다는 가정하에 신뢰구간, 추론의 확실성, 연구의 유용성 등을 논하고 있다. 응답자가 의도적으로 왜곡된 응답을 하거나 부정직한 응답을 할 가능성이 높은 조사 연구에서 중요하게 고려되어야할 점들은 뒤에서 자세히 논의하고자 한다.

8) 위의 신뢰구간 공식을 검토하면 이 점은 쉽게 증명될 수 있다.

무응답을 N 산출에서 제외하는 경우에 한가지 고려해야할 사항은 무응답을 제외한 나머지 자료만을 고려대상으로 삼으면 비록 표집이 확률적으로 무선절차에 의해 이루어졌다고 하더라도 무응답을 제외한 자료는 더 이상 모집단을 대표하지 않는 특수한 집단의 자료로 변모해 있을 가능성이 있다. 이것은 특정한 성향을 가진 사람들이 무응답을 많이 했는지, 아니면 응답자 성향에 관계없이 무응답이 무작위로 발생했는지에 따라 좌우되는 문제이다. 따라서 만약 A 문항에 대한 무응답이 5% 내지 10% 이상 되는 경우에는 이 문항에 대한 자료를 분석하기 전에 조사에 포함된 다른 문항들에 대한 응답을 이용하여 이 A 문항에 무응답한 사람들과 응답한 사람들이 여러가지 성향에서 서로 구별되는 사람들인지의 여부를 판가름할 필요가 있다. 만약 이 두집단의 성향이 모든 면에서 다르지 않다면 A 문항에 대한 무응답이 무작위적(random)으로 발생한 것으로 간주하고 분석에서 제외한 후 나머지 응답된 자료를 기초로 이 문항에 대한 양상을 파악해도 무방하다. 그러나 만약 이 A 문항에 응답한 집단과 무응답한 집단이 특정한 속성들에서 서로 구별되는 경향이 있으면 이 문항에 대한 분석과 결과해석에는 위에서 설명한 경계조건이 주어져야한다.

3. 교차분할

실태파악을 위한 조사연구에서 분명히 인식되어야할 한가지 목적은 그 조사연구의 결과를 어떻게 활용할 것인가, 즉 어떤 새로운 사실 혹은 지식을 습득하기 위하여 조사연구를 수행하는가이다. 13~18세 사이의 한국청소년 중에서 16.2%가 자위경험이 있다는 것은 많은 것인가, 적은 것인가? 청소년들의 자위행위는 사회문제인가? 16.2%라는 수치는 어떤 의미를 가지는 지식일

까? 이 새로운 지식을 청소년 정책에 반영한다면 어떻게 그것을 해석해야 할까? 이러한 중요한 질문들에 답하기 위해서는 조사연구에 포함된 각 변인들에 대한 반응을 독립적으로 일차원적으로 *unidimensional* 제시해서는 부족한 경우가 많다. 예를 들어서 남자청소년들중에는 30.2%가 자위경험이 있고, 여자청소년들중에는 2.2%가 자위경험이 있다면, 연구자는 ‘남자청소년들의 자위경험이 여자청소년들의 자위경험보다 월등히 많고, 여자청소년들은 남자청소년들에 비하여 자위경험을 거의 가지고 있지 않다’라는 해석을 의미있게 도출해낼 수 있다. 그런데 성별과는 독립적으로 그들 모두 통틀어서 산출된 16.2%라는 수치는 어떠한 의미부여도 가능하게 하지 않는다. 다시 말하면 의미있는 비교대상이 없는 것이다.

일차원적인 자료의 분석과 제시는 경우에 따라서는 매우 왜곡된 정보를 전달할 수도 있다. 다음과 같은 가상적인 경우를 상정해보자. “자위행위를 하는 것이 자연스러운 일이라고 생각하십니까? 해서는 안되는 일이라고 생각하십니까?” 라는 질문을 1750명의 청소년들에게 하고 다음과 같은 자료를 구하였다고 가정하자.

		자위행위는		
		자연스러운 일이다	해서는 안되는 일이다	
남자	700	350		1050
여자	50	650		700
	750	1000		1750

만약 자위행위의 자연스러움에 대한 응답을 성별로 나누지 않고 제시하였다면 한국 청소년의 과반수 이상(1750명중 1000명)이 자위행위를 해서는 안되는 일로 생각한다는 결론을 내리게

되고, 이 결론을 일반화하여 한국청소년들은 일반적으로 자위행위를 부정적으로 생각한다는 편견을 가지게 될 수 있다. 이 견해가 편견인 이유는 이 견해는 남자청소년들에 대해서는 명백하게 잘못된 견해이기 때문이다. 남학생들은 자위행위가 자연스러운 것이라는 태도를 가진 사람들이 그렇지 않은 사람들보다 두배(700 vs. 350)나 더 많다.

위의 가상적인 자료를 정리한 표에서 표의 외곽에 있는 빈도들을 편의상 ‘외곽빈도’, 표의 내부에 있는 빈도들을 편의상 ‘내부빈도’라고 부르면, 일차원적인 자료분석과 제시는 말하자면 외곽빈도만을 파악하는 것이다. 자위행위에 대한 태도가 다른 변인과 상관관계를 가지는 경우에 만약 외곽빈도들만을 파악하면 언제나 방금 설명한 것과 같은 편파된 해석을 도출하게될 위험이 있다. 그렇다고 해서 자위행위에 대한 태도와 상관관계를 가지는 모든 변인을 다 알아내어 내부빈도를 파악하기 위한 교차분할표를 작성하고 분석할 수는 없을 뿐만 아니라 그럴 필요도 없는 것이지만, 성별이나 연령 등과 같이 자위행위와 관련될 것으로 기대되는 중요한 몇가지의 변인들은 반드시 고려되어야 한다. 특히 이러한 조사연구의 자료가 청소년 정책의 수립과 시행을 위한 기초자료로 활용되기 위해서는 성별, 연령, 교육정도, 사회경제적 지위, 등의 소위 구조적 변인(*structural variable*)들에 의한 교차분할분석은 거의 필수적이라고 사료된다.⁹⁾ 왜냐하면 그럼으로서 한정된 자원과 에너지를 효율적으로 분배, 투입할 수 있기 때문이다.

V. 민감한 영역에 대한 반응의 문제

청소년을 대상으로한 조사연구에서 성경험 혹

은 비행에 관한 응답들은 신뢰롭지 못한 경우가 많다. 이러한 연구에서의 비신뢰로운 응답은 일종의 추정오차로서 표집오차(*sampling error*)와는 구별되는 ‘편파(*bias*)’에 의한 오차이다. 표집오차는 확률적인 표집과정에서 개입하는 오차로서 무작위적으로 발생하고, 따라서 가산되는 오차와 감산되는 오차가 궁극적으로는 서로 상쇄된다고 이론화되는 오차이다. 반면에 편파에 의한 오차는 체계적으로 발생하고, 가산되거나 감산되는 어느 한 쪽 방향으로 치우치는 오차이기 때문에 궁극적으로 서로 상쇄되는 것이 아니라 오히려 점차로 누적되는 오차이다. 예를 들어 성경험 혹은 비행에 관한 질문들에 대하여 솔직하지 않은 응답은 거의 언제나 사회적으로 바람직하다고 생각되는 방향으로 왜곡되는 양상을 보이게 된다.

Locander et al. (1976)은 1년동안 음주운전으로 경찰에 적발된 249명의 20대 전후의 젊은이들에게 “지난 1년동안 음주운전을 한 사실이 있는가?” 라는 질문을 일대일 면담, 전화면담, 설문조사 등의 다각적인 방법으로 실시해본 결과 솔직한 응답의 비율은 모든 방법에 걸쳐서 균일하게 46%~65% 정도임을 발견하였다. 특히 음주운전으로 적발된 적이 한번도 없는 사람이 음주운전을 한 사실이 있다고 대답한 경우는 전혀 없었고, 모든 왜곡된 응답은 음주운전으로 적발된 적이 있으면서 음주운전을 한 사실이 없다고 응답한 경우들이었다. Sobell et al. (1974)은 유사한 방법으로 술주정에 의해 경찰에 연행된 사실 여부를 조사하였는데 이때에도 솔직한 응답은 최저 39%, 최고 63.6% 정도에 불과하였다.

청소년들의 비행에 관한 조사질문에 대하여 응답자들이 얼마나 솔직하게 응답하는지를 알아본

9) 빈도의 교차분할분석을 위한 여러가지 통계방법들은 박광배 (1993)을 참고할 것.

연구들도 비교적 많이 존재한다. Ball(1967)은 마약사용에 관한 질문들에 대하여 질문의 유형에 따라서 솔직한 응답의 비율이 32.8%~92%의 광범위한 분포를 보이고 있음을 보고하였다. Hardt and Peterson-Hardt(1977)은 9학년(한국의 중학교 3학년) 914명에게 경찰에 체포된 적이 있는지를 묻는 간단한 질문에 대하여 89%의 솔직한 응답을 확인하였다. Clark and Tifft(1966)는 가출, 폭력, 절도, 등의 청소년비행 전반에 걸쳐서 다양한 조사를 하고 솔직한 응답의 비율이 최저 32.5%에서 최고 100%에 분포하고 있음을 보고하였다. 이들은 솔직성 여부를 판단하기 위하여 거짓말 탐지기를 이용하였다. 이들의 연구에서 한가지 흥미로운 사실은 심각한 비행에

관한 질문일수록 솔직한 응답의 비율이 높아지는 경향이 있다는 것이다. 심각한 비행일수록 그것을 범하는 청소년의 숫자가 적고, 경미한 비행일수록 그 숫자가 많다는 것에 착안하여 각 비행을 범하는 사람의 숫자와 솔직한 응답률 사이의 상관계수를 산출해본 결과 -0.87 이 산출되었다. 즉, 비행이 심각해서 그것을 범하는 사람이 적을수록 솔직한 응답률이 높아진다는 것이다.

성경험에 관한 조사질문에 대하여 솔직한 응답률을 알아본 연구는 Clark and Tifft(1966)와 Udry and Morris(1967)이 대표적이다. 이들의 연구결과가 Wentland and Smith(1993)에 요약되어 있는데 그것의 일부를 옮긴 것이 아래 표에 제시되어 있다.

연구	표본	응답률	질문	솔직한 응답률
Clark & Tifft	45명의 남자대학생	89%	여자를 임신시킨 경험	95%
			매춘부와 접촉경험	95%
			이성교제여부	90%
			동성과 성관계 경험	80%
			이성과 성관계 경험	67.5%
			자위행위의 여부	65%
			포르노물의 소지여부	50%
Udry & Morris	15명의 여성 혹인 근로청소년	100%	지난 24시간 동안 성관계를 가졌는가?	80%

주: Clark and Tifft 연구에서 솔직성 판단은 거짓말 탐지기를 이용하였고, Udry and Morris 연구에서는 소변검사가 이용되었다.

성경험에 관한 Clark and Tifft(1966)의 연구에서도 앞서 비행에 관한 조사에서와 마찬가지로 각 질문에 대하여 긍정적인 응답을 한 사람의 비율과 솔직한 응답률 사이에 부적인 상관관계가 발견되었다($r = -0.89$).

민감한 영역에 관한 조사에서 응답자들이 제공

하는 정보의 진실성에 관한 위의 자료는 이러한 조사연구의 결과를 어떻게 해석해야하는지에 대하여 매우 중요한 시사점을 제공한다. 앞서 우리는 신뢰구간의 필요성과 유용성을 살펴보았다. 신뢰구간은 표집오차의 존재를 가정하고, 그러한 표집오차의 범위가 얼마나 될지에 관한 확률적 추

론이라고 볼 수 있다. 편파에 의한 오차가 심각하게 내재해있는 경우에는 신뢰구간의 의미는 무엇일까? 앞서 인용된 조사에 의하면 13~18세 한국청소년 1500명의 표본중에 16.2%가 자위행위경험이 있고, 6.3%가 성관계경험이 있다고 응답하였다('91 MBC 청소년 백서, 1991). 이 자료들을 기초로 모집단에서의 비율에 대한 95% 신뢰구간을 산출하면 자위행위경험은 최저 14.34%와 최고 18.06%의 범위가 산출되고, 성관계경험은 5.07%와 7.53%의 범위가 산출된다. 그러면 우리는 한국청소년 전체중에 대략 이 범위의 학생들이 자위행위경험과 성관계경험을 가지고 있을 것으로 믿어도 좋은가? Clark and Tiff (1966)의 연구에 의하면 성관계경험에 관한 솔직한 응답은 67.5%, 자위행위에 관한 솔직한 응답은 65%에 불과하다고 한다. 이 연구결과에 기초하여 한국에서의 조사연구에서 모든 응답자가 만약 솔직하게 응답한다고 가정하면 실제로 자위행위경험이 있는 응답자는 16.2%가 아니라 약 $16.2\% + 35\% = 51.2\%$ 가 될 것이다.¹⁰⁾ 마찬가지로 한국의 청소년들중 성관계경험자도 실제로는 6.3%보다 훨씬 많을 가능성이 높다. 만약 한국청소년들중 자위행위경험이 있는 사람이 실제로 50%에 달한다면 앞서 산출된 신뢰구간 14.34%~18.06%는 실제와는 엄청난 괴리가 있는 왜곡된 정보가 될 것이다.

이것은 신뢰구간이라는 것이 언제나 '진실의 상태'를 의미하는 것은 아니라는 것을 자명하게 보여주는 예이다. 특히 자료속에 편파에 의한 오차가 개입되어있을 때는 신뢰구간 혹은 오차한계라

는 것이 그 의미를 상실할 뿐만 아니라 왜곡된 편견을 유도할 위험이 높다. 신뢰구간에 의해서 추정되는 오차한계는 모든 응답자가 최선을 다하여 진술한 응답을 한다는 가정하에 그 의미를 가진다. 물론 기억력의 결함이나 응답기록에서의 실수 등에 의해 응답이 부정확하게 제시될 수는 있지만 이러한 오류도 비체계적인 무선오류이고, 따라서 궁극적으로 서로 상쇄되는 오류이므로 신뢰구간과 오차한계의 해석적 의미에 영향을 주지 않는다. 그러나 응답자의 특정한 동기에 의해 체계적으로 발생하는 부정직한 오류는 신뢰구간과 오차한계의 해석적 의미를 결정적으로 훼손하는 결과를 초래한다.

민감한 영역에 대한 조사연구에서 응답의 솔직성 문제는 매우 해결하기 어려운 문제이고, 아직까지 부정직한 응답을 정확히 가려낼 수 있는 참신한 방법이 모색되지 않고 있다. 그러나 그것이 해결하기 어려운 문제라고 해서 그 문제를 전적으로 무시하고 마치 그러한 문제가 존재하지 않는듯이 조사를 수행하고 그 결과를 발표, 보고하는 것은 많은 자금과, 시간, 노력의 낭비에 불과하다. 민감한 영역에 대한 조사연구에서는 어떠한 형태로든 부정직한 응답의 정도를 가늠하는 노력이 필요하다. 일반적인 행동과학연구에서 솔직하지 않은 응답경향성을 측정해내는 방법들에 대해서는 Rogers (1988), Hall et al. (1991), Pankratz (1983), McClain et al. (1993) 등이 도움이 된다. 그러나 보다 간단한 방법으로는 중고등학생들의 비행에 관한 조사연구의 경우 객관적인 자료가 존재하는 사소한 비행을 이용하여 솔직성의

10) Clark and Tiff의 연구는 대학생을 대상으로 한 것이고 미국과 한국의 성문화차이가 매우 크기 때문에 '성관계경험'에 대한 이러한 수치적 추론은 무의미하다고 사료된다. 그러나 '자위행위'는 타인의 눈에 띄지않는 개인적인 상황에서 행해지므로 성관계보다는 덜 문화적인 영향을 받을 것이고, 대학생들보다 중고등학생들이 성관계의 기회가 더 희박한 까닭에 성충동을 자위행위에 의해 해소할 확률이 더 높으므로 이러한 추론이 가능하다고 생각된다.

정도를 추정하는 방법을 이용할 수 있다. 예를 들면 학생들에게 지난 한학기 동안 결석한 적이 있는지의 여부를 묻고 이 질문에 대한 응답을 출석부와 비교하여 부정확한 응답의 정도를 산출하는 것이다. 앞서 Clark and Tifft (1966)의 연구에 의하면 심각한 비행일수록 솔직한 응답률이 증가한다고 하므로 이러한 사소한 비행에 대한 부정확한 응답률은 보다 심각한 다른 비행에 대한 부정확한 응답률의 상한선을 제공할 것이다. 또한 특정한 학생 본인의 응답과 그 학생에 대한 다른

동료학생의 응답을 함께 수집하여 솔직한 응답의 정도를 대략은 짐작할 수 있는 경우도 있을 것이다. 그밖에 사회적 바람직성을 측정하는 척도 (Edward, 1957) 혹은 다면적 인성검사(MMPI)의 솔직성 척도들을 함께 실시하여 사회적 바람직성 혹은 비솔직성 점수가 지나치게 높은 학생들의 분포를 조사자료와 함께 제시하므로서 그 조사자료를 활용하는 사람들이 주관적으로 부정확한 응답의 정도를 가늠하게 해주는 것도 좋은 연구관행이 되리라고 생각된다.

참 고 문 헌

- '91 MBC 청소년 백서. (1991). 서울: 문화방송.
- 박광배(1993). 빈도분석. 서울: 성원사.
- 이종성(1988). 일반화가능도이론. 서울: 연세대학교.
- 한국갤럽조사연구소. (1991). 한국과 세계청소년의 의식: 우리나라 청소년을 세계 10개국과 비교한 제 2차 국제 조사보고서. 서울: 한국갤럽조사연구소.
- Ball, J. C. (1967). "The reliability and validity of interview data obtained from 59 narcotic drug addicts." *American Journal of Sociology*, 72, 650~654.
- Belson, W. A. (1981). *The Design and Understanding of Survey Questions*. England: Gower.
- Brinberg, D. and McGrath, J. E. (1985). *Validity and the Research Process*. Beverly Hills, CA: Sage.
- Campbell, D. T. (1990). "The role of theory in industrial and organizational psychology." In M. D. Dunnette and L. M. Hough(Eds.), *Handbook of Industrial and Organizational Psychology*, Vol. 1. Palo Alto: Consulting Psychologists Press.
- Clark, J. P. and Tifft, L. L. (1966). "Polygraph and interview validation of self-reported deviant behavior". *American Sociological Review*, 31, 516~523.
- Coleman, J. C. (1992). "The nature of adolescence. In J. C. Coleman and C. Warren-Adamson (Eds.), *Youth Policy in the 1990s: The Way Forward*. London: Routledge.
- Cook, T. D. and Campbell, D. T. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.
- Cooley, E. L. and Morris, R. D. (1990). "Attention in children: A neuropsychologically based model for assessment." *Develop-*

- mental Neuropsychology*, 6, 239~274.
- Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N.(1972). *The Dependability of Behavioral Measurement: Theory of Generalizability of Scores and Profiles*. New York : Wiley.
- Denzin, N. K. (1978). *The Research Act*. New York: McGraw-Hill.
- Dillman, D. A. (1978). *Mail and Telephone Surveys : The Total Design Method*. New York : John Wiley.
- Dominowski, R. (1980). *Research Methods*. Englewood Cliffs, NJ : Prentice-Hall.
- Edward, A. E. (1957). *The Social Desirability Variable in Personality Assessment and Research*. New York : Dryden Press.
- Glass, G. V., McGaw, B., and Smith, M. L. (1981). *Meta-analysis in Social Research*. Beverly Hills : Sage.
- Guenzel, P. J., Berkman, T. R., and Cannell, C. F. (1983). *General Interviewing Techniques*. Ann Arbor : Institute for Social Research.
- Hall, H., Shooter, E. A., Craine, J., and Paulsen, S. (1991). "Explicit alternative testing : A trilogy of studies on faked memory deficits." *Forensic Reports*, 4, 259~279.
- Hardt, R. H. and Peterson-Hardt, S. (1977). "On determining the quality of the delinquency self-report method." *Journal of Research in Crime and Delinquency*, 14, 247~257.
- Hunter, J. E., Schmidt, F. L., and Jackson, G. B.(1982). *Meta-analysis: Cumulating Research Findings Across Studies*. Beverly Hills : Sage.
- Kalton, G. (1983). *Introduction to Survey Sampling*. Beverly Hills : Sage.
- Locander, W. B., Sudman, S., and Bradburn, N. (1976). "An investigation of interview method, threat, and response distortion." *Journal of the American Statistical Association*, 71, 269~275.
- McClain, V. R., Hall, H., Leong, G. A. (1993). *Quick determination of deception using forced choice methodology*. A paper presented at the 1993 annual convention of American Psychological Association. Toronto, Canada.
- Pankratz, L. (1983). "A new technique for the assessment and modification of feigned memory deficits." *Perceptual and Motor Skills*, 57, 367~372.
- Rogers, R. (1988). *Clinical Assessment of Malingering and Deception*. New York : Guilford Press.
- Sackett, P. R. and Larson, J. R.(1990). "Research strategies and tactics in industrial and organizational psychology." In M. D. Dunnette and L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology*, Vol. 1. Palo Alto : Consulting Psychologists Press.
- Shavelson, R. J. and Webb, N. M.(1991). *Generalizability Theory : A Primer*. Newbury Park : Sage.
- Shaywitz, S. E. and Shaywitz, B. A.(1991).

- "Introduction to the special series on attention deficit disorder." *Journal of Learning Disabilities*, 24, 68~72.
- Singer, M. I., Singer, L. T., and Anglin, T. M. (1993). *Handbook for Screening Adolescents at Psychosocial Risk*. New York: Lexington.
- Sobell, M., Sobell, L., and Samuels, F. (1974). "Validity of self-reports of alcohol-related arrests by alcoholics." *Quarterly Journal of Studies on Alcohol*, 35, 276~280.
- Turner, C. S. and Martin, E. (Eds.). (1984). *Surveying Subjective Phenomena*. New York : Russell Sage Foundation.
- Udry, J. R. and Morris, N. M. (1967). "A method for validation of reported sexual data." *Journal of Marriage and the Family*, 29, 442~446.
- Warren-Adamson, C. (1992). Dear minister. In J. C. Coleman and C. Warren-Adamson (Eds.), *Youth Policy in the 1990s: The Way Forward*. London : Routledge.
- Wentland, E. J. and Smith, K. W. (1993). *Survey Responses : An Evaluation of Their Validity*. San Diego : Academic Press.
- Yates, F. (1981). *Sampling Methods for Census-es and Surveys*. New York : McMillan.