

## 연구논문

## 다양한 평가방법을 적용한 조사표 설계 사례연구

Case Study Research: A Multiple-Method Approach to Questionnaire Redesign

박영실<sup>a)</sup>

Young Shil Park

조사표를 평가할 때 한 가지 방법으로는 충분하지 않다. 인지면접, 포커스 그룹면접 등의 조사표 평가방법은 각기 장단점을 갖고 있으므로 다양한 방법을 복합적으로 활용할 경우 조사표 설계의 효과를 높일 수 있다. 이 연구에서는 2라운드의 인지면접, 시험조사, 그리고 조사원보고 방법을 적용하여 2015년 인구주택총조사 표본조사표를 설계한 사례를 보여주고 있다. 세 가지 방법은 응답자, 조사원, 조사표 설계 전문가 등의 다양한 관점을 반영하여 응답오차를 줄이기 위한 목적으로 선택되었다. 이 연구는 첫째, 반복적이고 다양한 조사표 평가방법을 적용해 봄으로써 어느 한 방법을 통해 조사표를 설계할 때 나타날 수 있는 위험성을 줄일 수 있다는 점, 둘째, 조사표 평가 방법에 대한 비교분석이 가능하며 이를 통해 조사표 방법론에 대한 이론적인 지식을 발전시킬 수 있다는 점, 셋째, 조사표 설계의 전 과정을 보여줌으로써 향후 조사표 설계 실무자에게 지침을 줄 수 있다는 점 등의 의의를 갖는다.

**주제어:** 조사표, 인지면접, 시험조사, 조사원보고, 통제집단, 실험집단

To redesign the 2015 Population and Housing Census, the Statistical Research Institute has conducted a pretest of the Census questionnaire using diverse methods. First, we diagnosed problems from two rounds of cognitive interviewing and revised the questionnaire based on the results. Next, in the field, we collected data from split ballot survey and interviewer debriefing with an original and a revised version of questionnaires. These three methods were chosen to reduce the response error from various perspectives including

a) 통계청 통계개발원 동향분석실 주무관 박영실.

E-mail: youngshil@korea.kr

respondents, interviewers, and questionnaire design specialists. One benefit of this multiple-method is its triangulation to decide whether specific items should be revised or not. The second benefit is the theoretical development of questionnaire pretesting methods from comparative analysis. Finally, this study sets guide lines for survey practitioners in Korea by demonstrating the whole process of questionnaire design.

**Key words:** questionnaire, cognitive interviewing, split-ballot survey, interviewer debriefing, control group, test group

## I . 연구배경

그동안 상세화 오차(specification error)는 비표집 오차(nonsampling error) 중에서 간과되어 왔다. 이 오차는 실제 측정되어야 하는 개념과 조사 질문에서 뜻하고 있는 개념 사이에 간극이 있을 경우에 발생한다. 개념 구성물(construct)이 잘못 측정되면 이로 인해 모수가 잘못 추정되고 결과적으로 당초의 이론에 근거한 추론은 타당성을 잃기 마련이다(Biemer 2010: 822). 따라서 완성도 높은 조사표를 만드는 일은 아무리 강조해도 지나침이 없으며, 그 과정에서 필수적인 것이 바로 사전조사(pretest)를 통한 조사표 평가이다. 이는 의사(=조사표 설계 전문가)가 정기검진(=조사표 평가방법)을 통해 숨겨져 있는 질병(=문제점)을 찾아내는 것으로 비유될 수 있다.

조사표 평가방법에 대한 논의가 점화된 이래 지금까지 인지면접(cognitive interviewing)을 중심으로 한 개별적인 평가방법들에 관한 연구가 축적되어 왔으며, 조사표 개선에도 효과적이라는 평가가 있어왔다(Beatty & Willis 2007). 하지만 인지면접과 함께 포커스 그룹면접(focus group interviewing), 조사원보고(interviewer debriefing), 행동분류(behavior coding) 등 다양한 방법을 복합적으로 적용한 연구는 드문 편이다.<sup>1)</sup> 조사표 설계 단계나 목적에 따라 여러 가지 방법을 혼합하여 적용한다면, 어느 한 방법이 갖고 있는 단점을 다른 방법이 갖고 있는 장점으로 극복할 수 있기 때문에 그 효과는 배가될 수 있다. 소수의 실증연구들에서는 꽤 일관적으로 이런 방법을 적용할 경우,

1) 구체적인 조사표 평가방법 소개는 박영실 외(2013 a)을 참고할 것.

다음과 같은 점에서 조사표 설계에 만족할 만한 성과가 있었음을 보여주고 있다 (Esposito 2004; Kaplowitz et al. 2004; Schaeffer & Dykema 2004; Tuttle et al. 2010). 첫째로, 응답자, 조사원, 조사표 설계 전문가 등의 다양한 관점에서 조사 측정 과정을 살펴볼 수 있다는 점과 둘째로, 반복적인 평가를 통해서 이전에 사용했던 방법에서 도출된 한계를 보완하여 재평가 할 수 있다는 점, 그리고 무엇보다 어느 한 방법을 통해서 내린 결정이 초래할 수 있는 위험성을 인지할 수 있다는 점을 긍정적으로 평가하였다.

Kaplowitz et al.(2004)은 환경 평가를 위한 조사항목 개발과정에서 포커스 그룹면접과 인지면접을 반복적으로 실시했다. 잠재적 응답자뿐 아니라 관련 분야의 전문가를 대상으로 한 포커스 그룹면접을 통해서 해당 분야에서의 주요 쟁점과 측정도구로서의 어려움 등을 진단해 내고, 이를 바탕으로 설계된 조사항목을 평가하기 위해서 여러 차례의 인지면접을 실시한 바 있다. Esposito(2004)는 미국 실직자 조사(displaced worker survey) 항목 개선을 위해 조사원보고, 행동분류, 응답자보고(respondent debriefing)의 3가지 방법을 사용하였는데, 1996년부터 2000년까지 매 2년 간격으로 동일한 방법들을 적용하여 그 결과를 비교분석하고 있다는 점이 특징적이다. Tuttle et al.(2010)은 전형적인 조사표 평가 과정을 통해 미국 경제분석국(Bureau of Economic Analysis)에서 수행하고 있는 사업체 조사표를 재설계하였다. 자료 분석가들을 대상으로 한 포커스 그룹면접과 응답자관찰 및 응답자보고 방법을 통해 오차의 원천을 진단하였으며 총 5라운드에 걸친 인지면접을 통해 조사표를 재설계한 후 현장 시험조사를 통해 해당 조사표가 실제 조사환경에서 적합한지를 검증하였다. Schaeffer & Dykema (2004) 또한 부모조사(parent survey)에서 사용되고 있는 특정 개념(예: custody)을 명확히 하기 위해 포커스 그룹면접과 3라운드의 인지면접, 시험조사, 행동분류 방법을 적용하였다. 그리고 추가적으로 시험조사 자료와 행정자료와의 매칭을 통해 응답의 정확성을 진단해 내었는데, 이는 앞의 다른 연구와 구별되는 점으로 조사표 평가의 성공 여부를 가늠할 수 있다는 기준을 제공했다는 점에서 매우 의의가 높다.

2015년은 인구주택총조사의 전환점이 되는 해이다. 인구주택총조사는 전수조사(short form)와 표본조사(long form)로 구분되는데, 향후 전수조사는 행정자료를 이용한 등록센서스로 대체되며 표본조사만 현장조사로 진행될 예정이다. 이에 표본조사에 대한 여러 가지 실험을 진행 중이며, 본 연구에서는 그 중 다양한 방법을 적용하여 진행한 조사표 평가의 일부 결과를 보여주고자 한다. 우리나라에서 조사 과정 중 조사표

설계 부분은 아직까지 블랙박스와의 같다는 점에서 이 연구는 실무적 차원과 이론적 차원 모두에서 주목해 볼 만하다. 먼저 다음 장에서는 조사표 평가 진행 과정 및 적용된 주요 방법을 소개하고, 3장에서는 조사항목이 재설계되는 과정을 기술한 후 4장에서 본 연구의 의의를 적시함으로써 결론을 맺고자 한다.

## II. 연구방법론

### 1. 조사표 재설계 배경

1925년 간이국세조사라는 이름으로 최초의 인구총조사가 시작된 이래로 인구주택총조사는 우리나라의 근간이 되어 온 조사이다. 이에 따라 사회의 변화상을 적확하게 포착하고자 많은 노력을 경주해 왔다. 최신 기술이 접목된 자료수집방법 및 입력방식 등을 도입하였으며, 변화된 사회모습을 반영하기 위해 새로운 조사항목을 개발하기도 하였다. 그러나 여타의 국가통계와 마찬가지로 인구주택총조사 또한 시계열 유지를 목적으로 보수적인 입장에서 기존의 형태를 유지해 온 측면이 있는데 혼인상태를 포함한 몇몇 핵심항목이 그러하다. 하지만 급격한 사회 환경의 변화로 인해 조사표 상에서 사용되는 용어가 시대에 따라 응답자 간 그 의미가 다르게 해석되는 경우가 있으며 혹은 기존의 용어로는 새롭게 발현된 현상들을 포착하기 어려운 경우도 발견되고 있다. 이에 따라 조사표 재설계에 대한 요구가 증가하여, 통계개발원에서는 2012년부터 ‘2015년 인구주택총조사’ 조사표 설계를 위한 연구를 단계적으로 수행하고 있다.<sup>2)</sup> 그 내용은 크게 조사항목 연구와 조사표 디자인 연구로 구분되는데 이 연구에서는 2012년도에 진행된 조사항목 평가 과정을 보여 주고자 한다.

### 2. 조사표 평가 절차

조사항목 연구는 인구주택총조사 핵심항목에 대한 문제점 진단과 이를 기반으로 한 수정안 제안을 목적으로 하였다. <표 1>은 조사표 평가 단계를 제시한 것이다. 먼저 최근 3개년도(2000, 2005, 2010)에 걸쳐 지속적으로 물어진 항목을 조사대상으로 선정

2) 조사표 평가 연구는 현재 진행 중이며, 본 연구는 통계청의 공식 입장이 아님을 밝혀 둔다.

〈표 1〉 조사표 평가 단계

단계	내용	기간	관심 관점
준비	항목 선정 및 검토	2012년 1월~2월	조사표 설계자
평가	인지면접 (2라운드)	2012년 3월~6월	응답자
재설계	조사표 재설계	2012년 7월~10월	-
평가	시험조사, 조사원보고	2012년 11월	응답자, 조사원 현장

하였고 이후 조사표 평가를 위한 기초 작업으로써 항목 검토를 실시하였다. UN 센서스 항목 권고안 및 통계청 조사 지침서를 검토하여 해당 질문이 조사 정의를 잘 준수하고 있는지를 조사하였다. 질문이 불명확하지 않은지, 개념이 모호하지 않은지, 문제가 될 만한 표현은 없는지, 조사표 설계 원칙에서 어긋난 것은 없는지 등을 조사표 설계자의 입장에서 검토하였다.

조사표 평가를 위해 적용한 방법은 크게 인지면접과 시험조사, 조사원보고이다. 응답자의 이해도를 파악하기 위해 두 차례의 인지면접을 수행하였으며 그 결과를 바탕으로 수정된 조사표 안을 제안하였다. 그리고 2010년 조사표인 기존안과 수정안에 대한 시험조사를 실시하였고, 조사 완료 직후 조사원으로부터 표준화된 양식을 이용하여 각각의 조사표에 대한 평가를 받았다. 이 방법들은 응답자와 조사원, 조사표 설계자 등 다각적인 관점에서 조사표를 설계하기 위한 목적으로 선정되었다.

### 3. 적용된 조사표 평가방법

#### 1) 인지면접

조사항목 평가를 위해 제일 먼저 적용한 조사표 평가방법은 인지면접이다. 인지면접은 보통 실험실 상황에서 참가자와의 일대일 면접으로 이루어진다. 면접원이 인지면접 참가자에게 응답에 이르기까지의 전 과정을 묻는 질문을 하면 응답자가 그에 대한 정보를 제공하는 방식으로 면접이 이루어진다. 소리 내어 생각하기(think-aloud), 캐어묻기(probing), 상황대입(vignette), 바꾸어 말하기(paraphrasing) 기법 등이 사용되며 표준화된 면접을 하기 위해 스크립트를 활용하기도 한다.

2012년 4월과 6월에 걸쳐 2라운드의 인지면접이 실시되었는데 1라운드에서는 조사

항목의 문제점을 진단하고, 2라운드에서는 1라운드에서 진단된 문제를 확인하고 진단된 문제점을 수정한 조사항목의 효과 등을 평가하는 데 주안점을 두었다. 통계청 홈페이지와 생활정보신문을 통해 참가자를 모집하였으며 1라운드에 15명, 2라운드에 14명을 선정하였다. 1라운드에서 대졸 이상이면서 자가 소유 및 아파트 거주자가 상대적으로 많이 지원했던 점을 보완하기 위해 2라운드에서는 조사 취약계층에 방점을 두어 참가자를 모집하였으며, 그 결과 인구사회학적 특성을 비교적 고르게 반영하여 참가자를 선정할 수 있었다.<sup>3)</sup>

인지면접은 이 연구의 목적 및 조사내용을 숙지하고 있는 연구진과 실무부서 담당자가 직접 수행하였다. 응답자에게 먼저 자기기입식으로 종이조사표를 작성토록 한 후에 주요 질문에 대해서 묻는 방법으로 진행하였으며, 모든 면접 과정이 녹음되었다.<sup>4)</sup> 이후 녹음자료를 이용하여 항목별 그리고 응답자 특성별 자료를 구축하여 내용분석(content analysis)을 실시하였다.

## 2) 시험조사

인지면접의 결과에 기반하여 수정된 조사표는 2015년 인구주택총조사 제1차 시험조사 틀 안에서 다시 평가되었다. 시험조사는 2012년 11월 1일 기준으로 실시되었다. 전주와 남해 시·군·구 내의 5개 읍·면·동을 시험조사 지역으로 선정하였으며, 이 5개 읍·면·동 총 조사구(415개) 중 20%를 계통추출 한 후 추출된 조사구 내의 모든 가구(4,653)를 조사대상으로 하였다(통계청 2012). 시험조사의 목적이 인지면접을 통해 수정한 조사표의 효과를 기존 조사표(2010년 안)와 비교하는 것이므로 이를 충족시키기 위해 표본을 하위 두 집단으로 분할하였다. 임의로 조사원을 통제집단(2,331가구)과 실험집단(2,322가구)으로 나누었으며 전자에는 기존 조사표를, 후자에는 수정된 조사표를 배부하여 조사하였다(〈표 2〉 참고). 총 조사대상인 4,653가구 중 4,514가구가 응답하여 전체 응답률은 97.0%이었으며, 그 중 통제집단은 2,266가구(97.2%) 실험집단은 2,248가구(96.8%)가 응답하였고 두 집단 간 가구 응답 여부에는 차이가 없었다( $\chi^2=1.752$ ,  $p=0.186$ ).<sup>5)</sup>

3) 구체적인 인지면접 방법론 소개는 박영실 외(2013b)를 참고할 것.

4) 조사원이 없는 상황에서 응답자가 해당 조사표를 어느 정도 이해하는지를 파악하기 위해 자기기입식으로 먼저 조사표를 작성하도록 하였다.

5) 집단별 가구원 수는 통제집단은 6,286명, 실험집단은 6,248명이며, 주택 수는 통제집단은 1,982

〈표 2〉 집단별 조사대상 가구 수 및 담당 조사원 수, 응답률

구분	통제집단	실험집단
조사대상 가구 수	2,331가구	2,322가구
담당 조사원 수	19명	18명
응답 가구 수	2,266가구	2,248가구
응답률	97.21%	96.81%

조사결과의 차이가 조사표에 의한 차이임을 보여주기 위해 다른 모든 절차는 동일하게 진행하였다. 자료수집은 인터넷조사와 방문면접조사의 2단계를 통해 이루어졌다. 인터넷조사는 자발적 참여가구를 대상으로 조사 초반 약 일주일 가량 실시하였으며(10월 28일~11월 4일), 인터넷조사에 참여하지 않은 가구에 한해 조사원이 직접 방문하여 면접을 하였다(11월 5일~19일). 조사완료 직후 조사원에 의한 현지 내용 검토와 자료 입력이 11월 20일~26일까지 실시되었다(통계청 2012). 한편, 전체 응답가구 중 각각의 자료수집 방법이 차지하는 비중은 인터넷조사가 12.1%, 방문면접조사가 87.9%였다.

통제집단과 실험집단 간 동질성 분석의 결과, 성별, 가구구성, 혼인상태, 경제활동상태 등에서는 차이가 없었으나, 연령과 거주 유형에서는 차이가 있었다. 거주 유형의 경우 통제집단에는 아파트가 실험집단에는 아파트 이외의 거처가 많은 것으로 나타났는데, 거주 유형은 주택항목 응답에 영향을 미칠 수 있다는 점에서 이후의 분석 과정에서 조사표 유형과 함께 거주 유형별 분석을 할 필요성이 제기되었다.<sup>6)</sup> 연령은 분석 과정에서 고려하지 않았는데 그 이유는 해당 변수가 응답자의 연령이 아닌 해당 집단 전체의 평균 연령이며 이는 향후 살펴보게 될 응답분포나 무응답에 영향을 미쳤을 것으로 예상하기는 힘들다는 점에 있다.<sup>7)</sup>

호, 실험집단은 1,830호이다.

- 6) 예를 들면 여러 가구가 함께 살고 있는 주택에 거주하는 응답자가 주택에 살고 있는 모든 가구를 포함해서 응답해야 하는 항목에서 본인의 가구만을 대상으로 응답하는 경향이 인지면접 과정에서 드러났다.
- 7) 또한, 연령의 경우 실험 목적상 통제집단에서는 실제 태어난 날짜를 기준으로 실험집단에서는 주민등록상 날짜를 기준으로 조사하였으므로, 조사표 상에서 이미 두 집단 간 차이가 발생할 수 있는 개연성이 내재되어 있었다.

수집된 자료를 이용하여 두 집단 간 항목별 무응답률, 오차율, 그리고 응답분포를 비교·분석함으로써 조사표 효과를 검토하였다. 무응답률은 해당 항목 응답 대상 사례 중 응답하지 않은 사례가 몇 사례인지, 마찬가지로 오차율은 해당 항목 응답 대상 사례 중 오차로 분류된 사례가 몇 사례인지로 측정하였다. 오차는 조사표 내 항목연관성 검토를 통해 논리적이지 않은 응답으로 정의하였다. 이를 통해 무응답률과 오차율이 높을수록 해당 항목에 대한 응답이 어려운 것으로 해석하였다. 응답분포의 경우 편향된 분포를 보이지는 않는지 혹은 제3의 자료와 비교했을 때 응답분포의 차이가 크지 않은지 등을 살펴보았는데, 응답분포는 무응답과 오차 사례 수를 제외한 후 계산하였다.

### 3) 조사원보고

조사원보고 방식에는 여러 가지가 있다. 조사를 수행한 조사원들이 모여서 토의를 진행하기도 하고, 별도의 양식을 통해서 조사항목 각각을 평가하기도 한다. 이번 연구에서는 시험조사 완료 직후 조사를 수행한 조사원들에게 자신이 수행한 조사표에 대한 평가양식을 배포하는 방법을 택하였다. 모든 조사항목에 대해 질문읽기가 얼마나 어려웠는지, 조사원이 보기에 응답자들이 응답을 얼마나 어려워하는 것으로 보이는지, 그리고 응답자들 간에 개념을 다르게 이해하고 있는 것은 없는지를 3점 척도(예: 전혀 어려움이 없었음-약간 어려움이 있었음-매우 큰 어려움이 있었음)로 평가하도록 하였다. 총 33명(89.2%)의 조사원이 응답을 하였으며 이 중 통제집단은 18명, 실험집단은 15명이었다.

## Ⅲ. 조사표 재설계 과정

인구주택총조사 조사표 재설계를 위한 2012년도 연구의 가장 중요한 목적은 사회 및 조사환경의 변화에 따른 기존 조사표의 문제점을 식별해 보고 이를 수정하여 좀 더 정확한 응답을 이끌어내는 데 있다. 이 장에서는 인지면접을 통해서 응답오차 지점을 파악하고, 그 오차를 줄이기 위해 항목을 수정한 후 시험조사 및 조사원보고를 통해 재평가한 과정을 구체적인 사례를 통해 살펴보고자 한다.

### 1. 응답과정과 응답오차

일반적으로, 응답은 질문에 대한 이해(comprehension)→정보인출(retrieval)→판단(judgement)→응답보고(reporting)의 네 단계를 거쳐 이루어지며 응답오차는 이 모든 단계에서 발생할 수 있다. 예컨대 질문을 잘못 해석할 수 있으며, 중요한 정보를 잊어버릴 수도 있고, 자신이 인출한 정보에 근거하여 잘못된 추론을 할 수도 있다. 혹은 자신의 응답을 응답범주에 적절하게 맞추어 응답하지 못 할 수도 있다(Tourangeau et al. 2000)(〈표 3〉참고).<sup>8)</sup>

〈표 3〉 응답과정과 오차 원인

구분	구체적 과정	오차 원인	사례 항목
질문이해	<ul style="list-style-type: none"> <li>- 질문과 지시문에 대한 주의</li> <li>- 질문의 초점 확인</li> <li>- 주요 용어와 개념의 연결</li> </ul>	<ul style="list-style-type: none"> <li>- 어렵거나 모호한 단어나 문장</li> <li>- 이중 부정문</li> <li>- 긴 문장</li> <li>- 질문과 응답범주의 불일치</li> </ul>	가구구분
정보인출	<ul style="list-style-type: none"> <li>- 기억인출 전략과 단서의 생성</li> <li>- 일반적, 구체적 기억 인출</li> <li>- 빠진 세부사항 채우기</li> </ul>	<ul style="list-style-type: none"> <li>- 정확한 정보인출의 어려움</li> <li>- 정보의 망각</li> <li>- 즉시 떠오르지 않는 기억</li> <li>- 구체적 질문에 대한 불확실한 기억</li> </ul>	1년/5년 전 거주지
판단	<ul style="list-style-type: none"> <li>- 기억의 연관성과 완전성 평가</li> <li>- 접근성에 근거한 결론 도출</li> <li>- 인출된 자료들의 통합</li> <li>- 부분인출에 근거한 추정 시도</li> </ul>	<ul style="list-style-type: none"> <li>- 답을 산출하는 데 복잡한 과정을 수행해야 하는 질문</li> </ul>	주거용 연면적
응답보고	<ul style="list-style-type: none"> <li>- 판단과 응답범주의 일치</li> <li>- 응답 수정</li> </ul>	<ul style="list-style-type: none"> <li>- 중첩되는 질문</li> <li>- 응답범주의 모호성</li> <li>- 응답범주가 없는 경우</li> </ul>	혼인상태

자료: Tourangeau et al.(2000) 및 이흥철 외(2012)를 보완하였음.

8) 물론 모든 단계에서 오차가 발생할 수도, 혹은 어느 한 단계에서 오차가 발생할 수도 있다. 하지만 오차가 어느 과정에서 발생했는지 명확히 구별하는 일이 쉽지 않다. 특히, 정보인출과 판단과정의 경계는 모호한 측면이 있어 Conard & Blair(1996)는 인지면접의 결과분석 과정에서 이 두 단계를 합쳐서 작업수행단계로 명명하기도 했다.

## 2. 질문이해: 가구구분 항목

질문이해에는 질문과 지시문에 주의를 기울인 후 질문에 의미를 부여하고 질문의 핵심을 추론하는 과정들이 포함되는데, 질문과 관련한 문제의 상당수가 응답자들이 질문을 잘못 이해했기 때문으로 알려져 있다. 쉬운 예로 한 질문에 동시에 둘 이상의 질문을 하거나 혹은 어렵고 모호한 단어들을 사용할 때 오차가 발생할 수 있다 (Tourangeau et al. 2000). 인지면접을 통해 가장 많이 진단되는 문제 유형 또한 질문 이해와 관련된 것인데(Willis 2005), 가족, 시·군·구, 집, 주택 등 평소 친숙하게 사용하는 용어들임에도 불구하고 조사자와 응답자 간 이해의 차이가 발생하는 것으로 진단되었다(박영실 외 2013b). 특히, 가구구분 항목에 포함되어 있는 ‘가족’이라는 용어는 사회변화에 따라서 응답자들이 이해하고 있는 내용이 달라졌음을 보여주는 대표적인 사례이다. 가구구분 조사항목은 본래 함께 살고 있는 사람의 수와 그들이 가족인지 여부에 따라 응답하게 되어 있다.

‘가족’이란 부부를 중심으로 친족관계에 있는 사람들의 집단이며 혼인, 혈연, 입양 등으로 이루어진 것이라는 조사 정의와 다르게 인지면접 참가자들이 생각하는 가족의 범위는 협소하였다. 1라운드와 2라운드 인지면접 참가자의 69% 가량이 가족의 범위를 부모-자녀 관계로 한정지어 이해하고 있었을 뿐 아니라, 친척의 경우 가족에 친척을 포함시키지는 않았으나, 그렇다고 해서 친척을 남으로 보는 것에도 주저하였다. 이에 따라 친척을 가족과 남의 중간범주로 해석하여 응답범주 ②번의 ‘가족 이외의 사람’으로 분류시키는 경향이 있었는데, 실제 이 조사표에서 가족 이외의 사람은 응답범주 ④번과 ⑤번에 있는 ‘남’과 같은 의미이다(박영실 외 2013b). 이러한 개념 차이는 우리나라 가족구조의 변화와 맥을 같이한다. 조부모-부모-자녀 이외에도 부모의 형제들과 함께 거주하는 대가족 형태가 보편적이었던 과거와 달리 최근에는 부모-자녀의 핵가족 형태가 일반적인데, 가족에 대한 기존의 사전적 정의는 그대로 유지된 채 가족을 구성하는 실체가 달라지면서 차이가 발생한 것으로 볼 수 있다.

가족 개념에 대한 이해의 차이로 인한 응답오차를 줄이기 위해 <표 4>와 같이 조사항목을 수정하였다. ①번의 ‘가족으로 이루어진 가구’를 ‘친인척을 포함한 가족으로 이루어진 가구’로, ②번의 ‘가족과 가족 이외의 사람이 함께 사는 가구’를 ‘친인척을 포함한 가족과 남이 함께 사는 가구’로, 그리고 ④번과 ⑤번의 ‘가족이 아닌 남남끼리’는

〈표 4〉 가구구분 항목: 통제집단과 실험집단

통제 집단	<p>이 가구는 아래에서 어디에 해당합니까?</p> <p>① 가족으로 이루어진 가구                  ② 가족과 가족 이외의 사람이 함께 사는 가구                  ③ 1인가구                  ④ 가족이 아닌 남남끼리 함께 사는 5인 이하의 가구                  ⑤ 가족이 아닌 남남끼리 함께 사는 6인 이상의 가구</p>
실험 집단	<p>이 가구는 아래에서 어디에 해당합니까?</p> <p>① 친인척을 포함한 가족으로 이루어진 가구                  ② 친인척을 포함한 가족과 남이 함께 사는 가구                  ③ 1인가구                  ④ 남남끼리 함께 사는 5인 이하의 가구                  ⑤ 남남끼리 함께 사는 6인 이상의 가구</p>

‘남남끼리’로 수정하였다.<sup>9)</sup>

인지면접을 근거로 수정한 조사항목이 실제 현장에서도 응답오차를 줄이는 데 기여했을까? 통제집단과 실험집단 간 가구구분 항목 응답분포는 두 집단 간에 차이가 없는 것으로 나타났으며( $\chi^2=2.90$ ,  $p=0.575$ ), 이는 항목 무응답률에서도 마찬가지였다( $\chi^2=0.24$ ,  $p=0.625$ ). 그러나 가구주와의 관계 항목을 통해 검토한 오차율은 통제집단보다 실험집단에서 1.6%p 감소하였다( $\chi^2=12.80$ ,  $p<0.001$ ). 특히, 그 오차 내용을 보면 통제집단에서는 응답범주 ①번(48가구)과 ②번(15가구)에서 전체 오차의 약 87% 가량이 발생한 반면 실험집단에서는 응답범주 ①번에서 절반 가까운 오차가 발생했으며 ②번에서는 오차가 거의 없어 응답범주 수정이 효과적이었던 것으로 나타났다.

9) 수정된 조사항목에서 응답혼란의 가능성이 완전히 사라졌는지에 대해서는 논쟁의 여지가 있다. 익명의 심사위원께서 수정된 항목의 가족에 여전히 친인척을 가족으로 보지 않는다는 의미가 내포되어 응답자들의 혼란이 예상된다는 지적을 하였으며, 저자 또한 이 부분을 인정한다. 다만, 향후 추가적인 작업을 통해 조사항목의 수정보완이 이루어질 것임을 알려두고자 한다.

〈표 5〉 가구구분 항목 응답분포 및 무응답률, 오차율: 통제집단과 실험집단 (단위: 빈도, %)

구분	통제집단		실험집단		실험-통제 (%p)
가족 가구	1,809	82.75	1,807	82.02	-0.73
가족+남 가구	2	0.09	4	0.18	0.09
1인가구	373	17.06	388	17.61	0.55
5인 이하 남남 가구	2	0.09	2	0.09	0.00
6인 이상 남남 가구	0	0.00	2	0.09	0.09
전체	2,186	100.00	2,203	100.00	-
무응답	8	0.35	10	0.44	0.09
오차	72	3.18	35	1.56	-1.62

〈표 6〉 가구구분 항목에 대한 조사원보고 결과: 통제집단과 실험집단

구분	통제집단	실험집단
질문읽기의 어려움	1.00	1.00
응답하기의 어려움	1.00	1.07
응답자 간 개념이해의 불일치도	1.00	1.13

주: 3점 척도로 측정되었으며, 점수가 높을수록 어려움 및 불일치도가 높은 것으로 해석됨.

그렇다면 조사원들은 변경 전후의 이 항목을 어떻게 평가했을까? 〈표 6〉을 보면 통제집단과 실험집단 조사원 모두 질문을 읽는 데에는 전혀 어려움이 없다고 평가했음을 알 수 있다. 그러나 조사원들이 평가한 응답 난이도는 비록 그 차이가 미미하기는 하지만 차이가 있었다. 통제집단보다 실험집단에서 응답을 하는 데 약간의 어려움을 더 겪었으며, 개념이해에 대한 불일치도 또한 실험집단에서 높았다.

의심 없이 일상적으로 사용되는 가족이라는 용어가 응답자 간에 다르게 이해되고 이로 인해 오차가 발생할 수 있음을 인지면접을 통해서 파악한 것은 이 연구의 성과 중 하나라고 할 수 있다. 또한 이를 수정한 후 실제 현장에서 조사를 실시해 본 결과, 응답분포에는 영향을 미치지 않는 가운데 유의미하게 응답오차를 줄여주었다는 점은

조사항목 수정의 근거가 될 수 있다. 물론, 가구구분 항목은 인구주택총조사항목 중 가구주와의 관계 항목을 통해서 보완이 가능한 항목이므로 재설계 중요성의 비중이 여타 항목에 비해 낮을 수도 있다. 하지만, 조사표 내에서 가족의 범위를 좀 더 명확하게 해 줌으로써 응답자의 혼란을 줄여줄 수 있고, 동시에 자료처리 과정에서 오차수정에 대한 수고를 줄일 수 있다는 점에서 의미 있는 변경이라고 할 수 있다.

## 2. 정보인출: 1년 전 및 5년 전 거주지 항목

질문에 응답하기 위해서는 기억으로부터 연관된 정보를 인출해야 한다. 이 과정에서 인출전략을 적용하고, 인출을 위해 구체적인 단서(cue)를 발생시키며, 개인적인 기억을 다시 모으는 등의 작업이 진행된다. 이것은 해당 사건이 얼마나 다른 사건들과 명확히 잘 구분될 수 있는지, 질문에서 사용된 용어들 사이에 연관성이 얼마나 높은지, 질문에서 제공되고 있는 단서의 수와 품질이 어떠한지, 기억의 원천이 직접적인 것인지 간접적인 것인지, 그리고 사건이 발생한 시간이 얼마나 오래 전인지 등에 의해서 영향을 받는다(Tourangeau et al. 2000).

인구주택총조사항목 중 5년 전 및 1년 전 거주지 항목은 정보인출 과정에서 오차가 발생할 수 있음을 보여주는 항목이다. <표 7>은 5년 전 거주지 항목이며, 현재 거주지와 비교해 5년 전 거주지가 다른지 여부와 시·군·구를 넘어 이동했을 경우 5년 전 주소를 기입하도록 되어 있다. 1년 전 거주지 항목 또한 동일한 형태를 띠고 있다.

예상과 달리, 인지면접 1라운드에서 참가자들은 자신이 5년 전에 어디에서 살았는지에 대해서는 비교적 어려움 없이 응답하는 것으로 나타났다. 이는 우리나라의 전세 및 월세 계약이 연 단위로 이루어지는 것과 관계있는 것으로 추정할 수 있다. 그러나 인지면접 참가자 15명 중에서 11명이 자가 소유자였으며, 이 중 5년 전과 변동이 없었던 경우가 9명이었던 점을 감안할 때, 자가 소유 여부의 영향이 더 컸던 것으로 볼 수 있다. 따라서 2라운드에서는 전·월세 거주자에 방점을 두고 참가자를 모집해 정보인출의 용이성을 한 번 더 평가해 보았다. 그 과정에서 ‘구’보다는 ‘동’이 5년 전 주소를 인출하는 데 효율적인 단서임이 진단되었다. 5년 전 거주지를 작성하는 데 시간을 지체하고 있는 응답자에게 그 이유를 물었을 때 “동은 알겠는데 구가 잘 안 떠오른다.”는 응답이 있었다. 이는 일상생활에서 주소를 물었을 경우 응답자들이 일반적으로 ‘구’를 응답하기보다 ‘동’을 응답하는 것에 더 익숙한 것과 관련이 깊다(박영실 외 2013).

〈표 7〉 5년 전 거주지 항목: 통제집단과 실험집단

통제 집단	<p>5년 전(2007년 11월 1일)에는 어디에서 살았습니까?</p> <p>① 태어나지 않았음 ② 현재 살고 있는 집 ③ 같은 시군구내 다른 집 ④ 다른 시군구 ⑤ 북한 또는 외국</p> <p>다른 시군구일 때는 해당하는 시도와 시군구의 행정구역명을 기입하여 주십시오. - 북한인 경우에는 북한이라고 기입하고, 외국인 경우에는 국가명만 기입합니다. _____ 특별시·광역시·도·국가명 _____ 시·군·구</p>
실험 집단	<p>5년 전(2007년 11월 1일)에는 어디에서 살았습니까?</p> <p>① 태어나지 않았음 ② 현재 살고 있는 집 ③ 같은 읍면동내 다른 집 ④ 다른 읍면동 ⑤ 북한 또는 외국</p>
	<p>5년 전 살던 한국내 지역 혹은 국가명을 적어 주십시오. - 북한인 경우에는 북한이라고 기입하고, 외국인 경우에는 국가명만 기입합니다. _____ 특별시·광역시·도 _____ 시·군·구 _____ 읍·면·동 북한 또는 외국 국가명: _____</p>

이에 따라, 〈표 7〉과 같이 5년 전 거주지 항목을 시·군·구 수준보다 낮은 읍·면·동 수준으로 묻도록 수정하였다. 이는 첫째, ‘구’는 기억하지 못하되 ‘동’은 기억하는 응답자에게 응답의 용이성을 제공해 줄 수 있으며, 둘째, 좀 더 작은 단위에서의 인구이동 자료를 제공해주어 심층 분석이 가능하다는 점에서 자료이용자에게 긍정적이다.<sup>10)</sup> 그러나 동시에 읍·면·동을 기억하지 못하는 응답자에게는 응답부담이 될 수 있다는 측면이 있어 과연 얼마나 많은 응답자들이 읍·면·동을 응답할 수 있는지 구

10) 통계청은 인구이동자의 개념을 시·군·구 경계를 넘어서 거주지를 이동한 인구로 정의함으로써 읍·면·동 단위의 자료생산이 불가능하다는 한계가 있다. 미국과 영국의 경우, 이동을 현재의 집이 아닌 다른 집에 거주했던 사람으로 정의함으로써 소지역 단위의 인구이동 측정이 가능하다는 것과 비교해 볼 때 우리나라의 경우 다양한 공간단위별로 인구이동 자료를 제공하지 못한다는 한계가 있다(최은영 외 2008).

체적인 데이터를 이용해 검증해 볼 필요가 있다.

주소를 작성하는 주관식 항목의 무응답률은 통제집단과 실험집단 중 어느 집단에서 더 높을까, 낮을까? 실험집단에서의 읍·면·동 항목 무응답률은 시도 및 시·군·구 항목 무응답률과 비교할 때 더 높을까, 낮을까? <표 8>을 보면, 1년 전과 5년 전 거주지의 읍·면·동 항목 무응답률은 실험집단 내 시·도 및 시·군·구 항목 무응답률에 비해서는 높으나 통제집단의 시·도 및 시·군·구 항목 무응답률에 비해 유의미하게 높은 수치는 아님을 알 수 있다. 먼저 1년 전 거주지 항목을 살펴보면, 시·도 및 시·군·구 항목 무응답률은 통제집단과 실험집단 간 차이가 없는 가운데( $\chi^2=0.79$  ( $p=0.375$ ),  $\chi^2=1.61$  ( $p=0.204$ )), 실험집단 내 읍·면·동 무응답률은 동일한 실험집단 내 시·도나 시·군·구보다 약간 높았으나 통제집단과 비교할 때는 차이가 크지 않음을 알 수 있다. 5년 전 거주지 항목에서도 유사한 경향이 발견되었다. 실험집단의 읍·면·동 항목 무응답률은 같은 집단 내 시·도나 시·군·구 항목 무응답률보다는 높은 수치이나, 통제집단의 시·도 및 시·군·구와 비교해 보면 그 차이는 줄어들음을 알 수 있다.

실제 조사 과정에서 조사원들은 새롭게 수정한 조사표를 어떻게 평가했을까? <표 9>를 보면 1년 전 거주지 항목에 대해 응답하기의 어려움을 제외하고 나머지 모든 항목에서는 통제집단이 실험집단에 비해 해당 조사표를 더 어려워하는 것으로 평가하였다. 그러나 가구구분 항목과 마찬가지로 두 집단 간 차이는 미미한 수준임을 알 수 있다.

<표 8> 1년 전 및 5년 전 거주지 주관식 항목 무응답률: 통제집단과 실험집단 (단위: 빈도, %)

구분		통제집단		실험집단		실험-통제 (%p)
1년 전 거주지	시·도	5	4.20	6	2.49	-1.71
	시·군·구	6	5.04	6	2.49	-2.55
	읍·면·동	-	-	13	5.39	-
5년 전 거주지	시·도	12	2.87	9	0.95	-1.92
	시·군·구	9	2.15	9	0.95	-1.20
	읍·면·동	-	-	29	3.06	-

〈표 9〉 5년 전 및 1년전 거주지 항목에 대한 조사원보고 결과: 통제집단과 실험집단

구분	1년 전 거주지		5년 전 거주지	
	통제집단	실험집단	통제집단	실험집단
질문읽기의 어려움	1.11	1.07	1.11	1.00
응답하기의 어려움	1.17	1.21	1.28	1.25
응답자 간 개념이해의 불일치도	1.18	1.11	1.50	1.21

주: 3점 척도로 측정되었으며, 점수가 높을수록 어려움 및 불일치도가 높은 것으로 해석됨.

응답자들이 시·군·구보다 읍·면·동과 같은 세부적인 정보를 기억해 내는 것을 더 어려워 할 것으로 예상하였으나 실제 조사 결과에서 통제집단의 시·도나 시·군·구와 비교할 때 무응답률이 큰 차이가 없는 것으로 나타난 점은 정보를 인출해 내는 과정에서 단서의 역할이 중요함을 보여준다. 하지만, 여전히 단일의 조사표 내에서는 시·도나 시·군·구에 비해 읍·면·동 인출수준이 낮은 것으로 나타났다. 인구주택 총조사에는 태도(attitude)보다 사실(fact)을 묻는 질문이 훨씬 많으며, 이러한 질문 유형에서는 정보인출 관련 오차가 발생할 개연성이 매우 높으므로 조사표 설계에서 기억 인출을 높일 수 있는 전략을 개발하는 것이 중요하다.

한편, 1년 전 및 5년 전 항목에 대한 인지면접 과정에서 인지면접 참가자가 누구인지에 따라서 진단될 수 있는 문제가 달라질 수 있음을 알 수 있었다. 자가 소유자가 많았던 1라운드에서는 발견하지 못했던 문제를 2라운드 전·월세 거주자 인지면접 과정을 통해서 발견한 점은 자발적인 참가자를 대상으로 한 인지면접의 한계와 동시에 반복적인 인지면접의 중요성을 보여주는 것이다.

### 3) 판단: 주거용 연면적 항목

정보인출만으로 질문에 대한 정확한 답을 산출하기 어렵기 때문에 응답자들은 그들이 인출한 정보를 결합하거나 혹은 보충하는 과정을 거친다. 특히, 사실과 관련된 질문에서 응답자들은 인출된 정보의 완전성 혹은 정확성에 대한 판단을 하고, 인출과정에 기반한 추론을 하며, 회상해 낸 것과의 차이를 메꾸기 위해 추론 등을 한다 (Tourangeau et al. 2000). 주거용 연면적 항목은 인출된 정보를 활용하여 추론하는 과정에서 응답자가 근거로 삼고 있는 기준을 잘못 적용하여 오차가 발생한 사례이다.

〈표 10〉 주거용 연면적 항목: 통제집단과 실험집단

통제 집단	이 주택의 주거용 연면적(건물 연면적)은 몇 m <sup>2</sup> 입니까? - 주거에 이용하는 부분만 기입하되, 소수점 이하는 반올림합니다. - 공동주택의 경우 분양면적이 아닌 전용면적을 기준으로 기입합니다. □□□m <sup>2</sup> *1평은 3.3m <sup>2</sup> 입니다.
실험 집단	이 주택의 주거용 연면적(건물 연면적)은 총 몇 제곱미터(m <sup>2</sup> )입니까? - 연면적은 건물 각 층의 바닥면적을 합한 전체 면적을 말함 - 공동주택의 경우 분양면적이 아닌 전용면적을 기준 - 소수점 이하는 반올림 □□□제곱미터(m <sup>2</sup> ) *1평은 3.3m <sup>2</sup>

주거용 연면적은 지금까지 여러 해 동안 조사를 해오면서 응답자들이 어려워하는 항목 중 하나로 꼽혔으나, 인지면접 결과는 기존의 예상과 차이가 있었다. 1라운드 인지면접에서는 주거용 연면적에 응답한 경우가 12명 중 7명으로 절반 이상이었다. 이러한 결과는 앞의 5년 전 거주지와 마찬가지로 참가자 특성과 무관하지 않다. 1라운드 참가자의 상당수가 아파트를 소유한 자(8명)이며, 또한 해당 주택에 입주한 지 얼마 되지 않는 응답자가 일부 있었는데, 이들의 경우 좀 더 쉽게 응답하는 경향이 있었다. 그러나 2라운드에서는 단순히 해당 면적의 정보를 알고 있는지 여부와는 다른 차원의 오차가 진단되었다. 단독 및 연립주택 등 아파트 이외의 거처에 거주하는 참가자 상당수가 층별 연면적을 모두 합산하지 않고 본인이 거주하고 있는 공간의 연면적만을 응답, 즉 응답범위를 잘못 판단함으로써 오차를 발생시켰다. 실제 조사표 작성 과정에서 몇몇 참가자들은 층별 주거용 연면적을 모두 합산해야 하는지를 물어온 경우가 있었다. 따라서, 〈표 10〉과 같이 주거용 연면적 항목에서는 질문과 지시문 두 가지를 수정하였다. ‘이 주택의 주거용 연면적(건물 연면적)은 몇 m<sup>2</sup>입니까?’로 물었던 조사표 원안에 주택 전체의 면적을 합산하라는 의미를 포함시키기 위해 m<sup>2</sup> 앞에 총(total)이라는 수식어를 넣었으며, 지시문에 ‘연면적은 건물 각 층의 바닥면적을 합한 전체 면적을 말함’을 새롭게 추가하였다(박영실 외 2013b).

연면적은 건물 각 층의 바닥면적을 합한 전체 면적을 뜻한다는 지시문을 추가한 실험집단에서 평균면적이 더 넓게 나타났는지를 보기 위해 집단별 평균면적의 차이를 비교해 보았다. 〈표 11〉을 보면 실험집단(101m<sup>2</sup>)이 통제집단(82.5m<sup>2</sup>)에 비해 18.5m<sup>2</sup> 더 넓

은 것으로 분석되었으며(M-W U=126,896 (p<0.001)), 이는 완산구와 남해군의 2010년 인구주택총조사 결과(70.8m<sup>2</sup>)보다도 큰 수치였다. 그런데 아파트의 경우, 독립된 주거형태로 인정해서 인구주택총조사의 평균 주거용 연면적 계산 시에는 본인이 실제로 거주하고 있는 면적만을 응답하도록 하였다. 따라서 건물 각층의 모든 면적을 합해서 응답하라는 지시문의 효과는 아파트보다는 아파트 이외의 거처 유형에서 면적을 유의하게 증가시킨다는 것이 좀 더 합리적인 가설이다. 이를 확인해 보기 위해 거처 유형별로 나누어 면적을 살펴보았다. 단독주택 등 아파트 이외의 거처에서는 그 면적 차이가 유의미하지 않은 것과 달리(M-W U=168,230 (p=0.680)), 아파트에서는 조사표 유형 간에 면적 차이가 있었다(M-W U=448,715 (p<0.001)). 아파트에 거주하는 응답자 집단에서 주거용 연면적이 증가한 것은 지시문의 효과가 아닌 통제집단과 실험집단의 아파트 면적 차이에서 기인했을 가능성을 배제할 수 없다. 여기에서 제시하고 있지는 않으나 사용방 수를 측정한 항목을 보면 통제집단에 비해 실험집단에서 평균 사용방 수(3.8개 vs 4.1개)가 더 많은 것으로 나타나 이러한 추정을 어느 정도 지지해 주고 있다.

한편, <표 12>의 항목 무응답률과 오차율을 보면 항목 무응답률은 집단 간 차이가 없었으나( $\chi^2=0.24$ , p=0.625), 거처 유형 및 총 방 수 항목과 연관성을 검토하여 오차를 살펴본 결과 통제집단이 실험집단에 비해 오차율이 높은 것으로 나타났다.

<표 11> 거처 유형에 따른 주거용 연면적 비교: 통제집단과 실험집단 (단위:m<sup>2</sup>)

구분	전체		아파트		아파트 외	
	통제집단	실험집단	통제집단	실험집단	통제집단	실험집단
평균면적	82.52	100.99	72.39	92.94	113.40	112.95

<표 12> 주거용 연면적 무응답률 및 오차율: 통제집단과 실험집단 (빈도,%)

구분	통제집단		실험집단		실험-통제 (%p)
무응답	7	0.35	44	2.40	2.05
오차	12	0.61	0	0.00	-0.61

〈표 13〉 주거용 연면적 항목에 대한 조사원보고: 통제집단과 실험집단

구분	통제집단	실험집단
질문읽기의 어려움	1.18	1.27
응답하기의 어려움	1.47	1.67
응답자 간 개념이해의 불일치도	1.25	1.40

주: 3점 척도로 측정되었으며, 점수가 높을수록 어려움 및 불일치도가 높은 것으로 해석됨.

〈표 13〉의 조사원보고 결과에 따르면, 조사원이나 응답자 모두 실험집단의 조사표를 더 어려워하는 것으로 나타났다. 앞서 살펴본 가구구분이나 1년 전 및 5년 전 항목과 달리 조사원의 질문읽기에서도 실험집단의 조사표가 더 어렵다고 응답하였다. 주거용 연면적 항목에 대한 응답자들의 응답 어려움이나 응답자 간 개념이해의 불일치도 또한 실험집단에서 높았다. 추가된 지시문으로 인해 질문이 복잡해 보였으며, 조사내용의 정확성을 전달하기 위한 지시문이긴 하나 응답자에게 건물의 각 층의 면적을 모두 합하라는 것이 부담으로 작용했을 가능성이 있다. 한편, 이 항목의 난이도를 평가한 조사원보고 점수가 앞의 가구구분이나 1년 전 및 5년 전 거주지 항목에 비해 높은 것을 통해 전반적으로 주거용 연면적 항목을 응답자들이 더 어려워함을 짐작할 수 있다.

주거용 연면적 항목에 대한 인지면접, 시험조사, 그리고 조사원보고 결과를 종합적으로 판단해 볼 때 주거용 연면적 항목의 수정에 대한 추가적인 검토가 필요할 것으로 판단된다. 연면적은 건물 각 층의 바닥면적을 모두 합해서 계산해 달라는 지시문은 질문의 내용을 정확하게 전달한다는 측면에서 필요한 것이나, 조사항목이 수정된 실험집단에서의 무응답률이 더 높았으며 조사원들이 보기에 응답을 더 어려워하는 것으로 평가하였다는 점을 간과해서는 안 될 것이다. 증가한 면적 차이 또한 조사표 내 지시문의 효과보다는 통제집단과 실험집단으로 추출된 지역적 특성 차이에서 연유했을 가능성이 높기 때문이다. 조사표 효과를 검증하기 위해 표본추출 과정에서 통제집단과 실험집단의 동질성이 무엇보다 중요함을 보여주는 결과이다.

#### 4) 응답보고: 혼인상태

응답의 마지막 과정은 선택(selecting)과 보고(reporting)이다. 선택과 보고는 응답을 적절한 응답범주에 일치시키는 것과 응답기준의 일관성을 위해 수정하는 과정을 말한다. 응답자들은 꽤 명확한 답을 갖고 있을지라도 어떻게 응답해야 할지 막막해 할 때가 있으며 이 과정에서 자의든 타의든 오차가 발생할 수 있다(Tourangeau et al. 2000). <표 14>는 혼인상태 항목으로 통제집단의 질문에서는 실제 혼인상태를 묻고 있으나 응답범주가 법적인 기준에 따라 분류되어 있음을 알 수 있다. 질문과 응답범주 간에 차이가 있는 것은 사회변화에 따라 법적인 테두리 안에서 분류되지 않는 혼인상태가 나타났기 때문이다.

이러한 차이는 응답자들이 어느 것을 기준으로 응답을 해야 할지 혼란을 유발하는 것으로 나타났다. 과연, 법적인 상태와 실제 상태 간 혼인형태가 다른 경우 응답자들은 어느 것을 기준으로 응답을 할까? 동거는 법적으로는 미혼이지만 사실혼 형태로 볼 수 있으며, 별거는 법적으로는 배우자가 있는 상태이나 이혼 등의 의사를 갖고 배우자와 떨어져 사는 형태이다. 1라운드에서는 상황대입 기법을 적용하여 참가자들에게 다음과 같은 질문을 던졌다. “혼인신고를 하지는 않았지만 함께 살고 있는 파트너가 있다면 어디에 응답해야 한다고 생각하십니까? 별거상태라면, 어디에 응답을 해야 한다고 생각하십니까?” 동거의 경우 절반의 참가자가 ‘미혼’에, 그리고 나머지 절반은 ‘배우자 있음’에 응답해야 한다고 한 반면, 별거의 경우 대다수가 ‘배우자 있음’에 응답해야 한다고 하였다. 별거에 대해서는 엄격하게 법적인 상태를 판단근거로 삼고 있어 이혼을 전제로 한 별거의 경우 이혼으로 응답하라는 질문의도와 다른 응답을 하고 있음을 알 수 있었다(박영실 외 2013b). UN 권고안에서 별거 중인 부부가 재혼이 자유롭지 않다는 이유로 배우자가 있는 상태로 간주되는 것을 경계해야 함을 지적한 것은 이러한 상황을 말한 것이다.

동거 및 별거 범주의 추가는 해당 내용의 민감성 때문에 지금껏 논쟁적이었다. 이에 2라운드에서는 두 가지 범주의 민감성 정도에 대해서 물었는데 예상 외로 동거 및 별거 범주가 있더라도 ‘상관없다’는 응답이 절반 이상이었다. 특히 한 50대 여성은 “여기에 만일 이게 없는데 별거를 하고 있다면 어디다가 표시를 하겠어요?”라고 반문해 질문과 응답 범주가 일치하지 않을 경우 응답오차가 발생할 수 있는 지점을 정확하게 짚어 주었다(박영실 외 2013b).



〈표 15〉 혼인상태 항목 응답분포 및 무응답률, 오차율: 통제집단과 실험집단 (단위: 빈도, %)

구분	통제집단		실험집단		실험-통제(%p)	2010 총조사
미혼	1,263	25.19	1,271	24.83	-0.36	29.73
배우자있음	3,119	62.22	3,211	62.74	0.52	57.68
사별	440	8.78	441	8.62	-0.16	8.94
이혼	191	3.81	166	3.24	-0.57	3.64
별거	-	-	29	0.57	-	-
전체	5,013	100.00	5,118	100.00	-	-
무응답	225	4.28	170	3.21	-1.07	
오차	15	0.59	10	0.19	-0.10	

주) 2010년 인구주택총조사 결과는 완산구와 남해군의 결과임.

〈표 16〉 혼인상태 항목에 대한 조사원 평가 결과: 통제집단과 실험집단

구분	통제집단	실험집단
질문읽기의 어려움	1.83	1.67
응답하기의 어려움	1.89	1.73
응답자 간 개념이해의 불일치도	1.53	1.33

주: 3점 척도로 측정되었으며, 점수가 높을수록 어려움 및 불일치도가 높은 것으로 해석됨.

현장 조사에서 조사원들은 통제집단에서 사용한 조사표를 더 어려워하는 것으로 나타났다. 〈표 16〉을 보면 조사원은 질문읽기나 응답하기의 어려움, 응답자 간 개념이해의 불일치도 등 세 가지 항목 모두 통제집단에서 난도가 더 높은 것으로 평가하였다. 앞의 가구구분 항목이나, 1년 전 및 5년 전 거주지, 주거용 연면적 항목과 비교해 볼 때에도 혼인상태 항목에 대한 조사원보고의 난도 평가 점수가 가장 높다.

혼인상태 항목에 대한 인지면접과 시험조사, 그리고 조사원보고를 통해 동거 및 별거를 응답범주에 추가하는 것에 대한 긍정적인 평가를 내릴 수 있다. 특히 대부분의 응답자가 별거에 대해서 사실적인 상태보다는 법적인 상태를 기준으로 응답하며, 동

거나 별거를 응답하는 것이 민감하지 않다는 것을 보여준 인지면접의 결과는 흥미롭다. 이러한 결과를 소수의 자발적인 참여자를 대상으로 한 실험실의 효과로 볼 수 있다는 한계가 있으므로 해당 항목의 현장 적용 가능성을 검토해 보았으며, 그 결과 이 항목에 대한 무응답률이 통제집단에 비해서 낮다는 점과 별거가 0.6% 포착된 것은 조사항목의 개선 가능성을 보여준 것이다. 단, 아직까지 우리 사회에서 동거나 별거에 대한 명확한 합의가 이루어지지 않았기 때문에 이에 대한 추가적인 검토가 필요할 것이다. 별거의 경우 법적으로 배우자가 있는 상태이나 사실혼에서는 이혼과 크게 다르지 않을 수 있다는 측면 등에서는 현재의 5개 응답범주가 상호 배타적이지 않다는 한계가 있다.<sup>13)</sup>

## VI. 결론

조사표 설계 과정은, 조사표 설계 → 평가 → 재설계 → 평가 → ... 로 피비우스의 띠와 같이 끝이 없는 작업을 요구하지만, 그것이 자료의 품질에 미치는 영향을 생각한다면 결코 소홀히 할 수 없는 과정이다. 이번 연구에서는 인지면접, 시험조사, 조사원보고의 세 가지 조사표 평가방법을 적용하여 인구주택총조사 표본조사표 재설계 과정의 일부를 사례로 보여주었다. 현재 진행 중인 프로젝트로 조사항목을 최종적으로 확정지은 단계가 아니나, 연구의 진행과정에서 이 연구가 조사표 설계에서 갖는 다음과 같은 이론적 그리고 실무적 의의를 발견할 수 있었다.

첫째, 기존의 연구(Esposito 2004 등)와 마찬가지로 이 연구에서도 다양한 평가방법을 반복적으로 적용해 봄으로써 조사표 설계의 발전적 과정을 보여줄 수 있었다. 인지면접을 통해서 문제를 진단하고, 진단된 문제를 수정한 후 시험조사를 통해서 다시 평가하는 과정을 거쳤다. 이를 통해 어느 항목의 경우에는 항목 수정의 타당성을 확보할 수 있었으며 또 다른 항목의 경우에는 추가적인 검토가 더 필요하다는 결론에 이른 항목도 있었다. 이는 어느 하나의 방법을 통해서 조사항목을 수정하는 일은 매우 위험할 수도 있음을 역설적으로 말해준다.

13) 익명의 심사위원 또한 이 부분을 지적하였으며, 향후 추가적인 검토를 통해서 수정·보완할 예정이다.

둘째, 이러한 결과는 다양한 관점의 조사표 평가방법을 융합함으로써 얻을 수 있었다. 인지면접은 응답자 관점에서, 조사원보고는 조사원 관점에서 조사표를 평가할 수 있으며, 시험조사는 실제와 유사한 현장에서 조사원과 응답자 관점을 반영한다는 특징을 갖는다. 응답자, 조사원, 그리고 조사표 설계자 모두 조사 과정에서 응답오차가 발생할 수 있는 원천임에도 불구하고 지금까지 우리나라에서 주로 사용되어 오던 조사표 평가방법은 시험조사 후의 조사원 간 집단 토론으로 제한적이었다. 그런데 조사원에 의한 평가는 어느 한 두 항목 혹은 한 두 가구에 대한 인상이 조사표 전체에 대한 평가로 확대해석 될 가능성이 높다. 이러한 한계를 보완하기 위해 인지면접을 통해 응답자 관점에서 조사표를 진단했으며, 시험조사에서는 수정된 조사표의 효과를 살펴보기 위해 기존 조사표와 동시에 자료를 수집하였다. 또한, 집단 간 토론 대신 평가양식을 통한 조사원보고 방식을 취해 모든 조사원의 의견이 고루 반영되도록 하였다. 이러한 방법은 조사원이나 연구자의 어느 한 관점에서 평가되던 조사표 평가방법에서 벗어나 다각적인 관점에서 균형 잡힌 시각을 갖고자 노력하였다는 의의를 갖는다.

셋째, 각각의 방법에 대한 비교·분석을 통해 이론적 지식을 발전시킬 수 있었다. 본 연구의 결과 인지면접과 시험조사는 조사표를 평가하는 데 매우 효과적인 방법으로 보인다. 인지면접은 조사항목 이해에 대한 ‘숨겨진(hidden)’ 문제를 잘 진단해 내었으며 이를 근거로 수정된 조사표의 효과를 검증하기 위해 선택한 시험조사 방법은 적절한 것으로 보인다. 하지만, 수정 전후의 조사항목 평가를 함에 있어서 조사원보고 방법의 효과성에는 의문이 든다. 통제집단과 실험집단 간 평가 점수의 차이가 크지 않아 어느 것이 좋다는 명확한 결론을 내릴 수 없었다. 그렇다면 조사원보고 방법은 조사표를 평가하는 데 쓸모없는 방법일까? 동일 항목에 대한 두 가지 조사표 유형을 비교하기보다는 단일의 조사표 내에서 항목간 비교를 하는 데 더 효과적일 것이다. 예컨대 가구구분 항목에 대한 통제집단과 실험집단 간 차이는 0.07점으로 미미하지만(예: 응답하기 어려움 항목, 통제집단 1.00 vs 실험집단 1.07) 가구구분 항목과 혼인상태 항목에 대한 조사원 평가 결과의 차이는 0.89점으로 매우 뚜렷하다(예: 응답하기 어려움 항목 통제집단 1.00 vs 통제집단 1.89). 이를 통해 조사원이나 응답자가 해당 조사표에서 어느 항목을 특히 어려워하는지, 혹은 쉬워하는지 등을 식별해 낼 수 있다.

넷째, 이번 연구에서는 인지면접을 2라운드에 걸쳐서 실시하였는데 이러한 반복적인 평가는 이전 라운드 인지면접의 한계를 즉각적으로 보완할 수 있다는 장점이 있다. 예컨대 지원자 특성이 특정 경향으로 편향된 것을 보완하기 위해 2라운드에서는 다양

한 층을 포괄하고자 하였으며 그 결과 1라운드에서 진단되지 않은 다른 측면의 문제점이 진단될 수 있었다. 따라서 인지면접의 경우 유사한 특성을 가진 다수에 대한 1라운드 면접보다는 비교적 소수일지라도 여러 라운드에 걸쳐서 실시하는 것이 효과적일 수 있다.

다섯째, 이 연구는 지금까지 잘 알려져 있지 않은 조사표 설계 과정을 가감 없이 보여줌으로써 자료의 품질 향상에서 조사표 역할의 중요성을 보여주었다는 점에서 의의가 있다. 외국과 달리 우리나라에서는 조사표 설계 과정을 문서화하는 데 많은 노력을 기울이지 않는 편이다. 한걸음 더 나아가 반복적이고 다양한 조사표 평가방법이 조사표 설계에 어떻게 적용될 수 있는지를 단계적으로 보여주었다는 점에서 향후 조사표 설계자에게 지침을 제공해 줄 수 있을 것으로 기대한다.

이번 연구 역시 다음과 같은 몇몇 한계는 있다. 응답자나 조사원 관점뿐 아니라 자료 이용자 측면에서 조사표를 평가하는 것 또한 중요하다. 해당 조사항목이 연구 목적을 분명히 전달하는지, 질문이나 응답범주에 이론적인 오류 등은 없는지를 체계적으로 검토해야 한다. 이번 연구 대상인 인구주택총조사의 경우 이용자 측면에서 오랫동안 검토되어 왔다고 볼 수 있으나, 돌다리도 두드려보는 자세로 수정된 항목이 기존의 작성 목적을 동일하게 수행할 수 있는지 등에 대한 세심한 검토가 필요하다. 인지면접이나 시험조사가 갖는 한계 또한 분명히 인식할 필요가 있다. 두 방법 모두 그 대상이 소수의 자발적, 그리고 특정 지역이라는 한계를 갖고 있으므로 결과를 해석하는 데 주의를 기울여야 한다. 향후 이런 점들을 보완한 추가 실험을 진행할 예정이다. 2차, 3차 인구주택총조사 시험조사에서는 그 지역을 다변화해 지역적 특성을 고루 반영할 것이며, 시험조사 과정에서 직접 응답자들의 의견을 듣는 응답자보고 방식을 통해서 실험실 효과 등을 배제하고자 노력할 것이다.

## 참고문헌

- 박영실·박현정·윤연옥. 2013a. “조사표 평가방법에 관한 최근 동향.” 《조사연구》 14(2): 201-227.
- 박영실·박현정·정남수. 2013b. “2015 인구주택총조사 조사표 설계를 위한 인지면접 사례 연구.” 《통계연구》 18(1):34-55.

- 이흥철 · 김영진 · 장윤희 · 박영실 · 박현정 · 박주언. 2012. 《자료수집방법에 따른 조사표설계 가이드라인 작성》 통계개발원 보고서.
- 최은영 · 조순기 · 박영실. 2008. 《인구이동통계 데이터베이스 구축 및 우리나라의 인구이동 특성 분석》 통계개발원 보고서.
- 통계청. 2012. “2015 인구주택총조사 제1차 시험조사 실시계획(안).” 내부자료.
- Beaty, P.C. and G.B. Willis. 2007. “Research Synthesis: The Practice of Cognitive Interviewing.” *Public Opinion Quarterly* 71(2):287–311.
- Biemer, P.P. 2010. “Total Survey Error: Design, Implementation, and Evaluation.” *Public Opinion Quarterly* 74(5):817–848.
- Conard, F. and J. Blair. 1996. “From Impressions to Data: Increasing the Objectivity of Cognitive Interviews.” [www.amstat.org/sections/srms/proceedings/\(2013.9.2\)](http://www.amstat.org/sections/srms/proceedings/(2013.9.2)).
- Esposito, J.L. 2004. “Iterative, Multiple–Method Questionnaire Evaluation Research: A Case Study.” *Journal of Official Statistics* 20(2):143–183.
- Kaplowitz, M.D., F. Lupi, and J.P. Hoehn. 2004. “Multiple Methods for Developing and Evaluating a Stated–Choice Questionnaire to Value Wetlands.” in S. Presser, J.M. Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, J. Martin, E. Singer(eds.). *Methods for Testing and Evaluating Survey Questionnaires*. pp.503–524. John Wiley and Sons, Inc.
- Schaeffer, N.C. and J. Dykema. 2004. “A Multiple–Method Approach to Improving the Clarity of Closely Related Concepts: Distinguishing Legal and Physical Custody of Children.” in S. Presser, J.M. Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, J. Martin and E. Singer(eds.). *Methods for Testing and Evaluating Survey Questionnaires*. pp. 475–502. John Wiley and Sons, Inc.
- Tourangeau, R., L.J. Rips., and K. Rasinski. 2000. *The Psychology of Survey Response*. U.K.: Cambridge University Press.
- Turtle, A.D., R.L. Morrison, and D.K. Willimack. 2010. “From Start to Pilot: A Multiple–Method Approach to the Comprehensive Redesign of an Economic Survey Questionnaire.” *Journal of Official Statistics* 26(1):87–103.

<접수 2013/9/3 , 수정 2013/11/5 , 게재확정 2013/12/16>