〔보 문〕

Journal of the Korean Society
of Clothing and Textiles
Vol. 26, No. 6 (2002) p.779~789

한국의류학회지에 게재된 설문조사 연구의 방법론적 특성 분석

A critical review of survey methods in Journal of the Korean Society of Clothing and Textiles

울산대학교 생활과학부 의류학전공, 고려대학교 행정학과
이 미 영 · 김 태 일

Dept. of Clothing and Textiles, University of Ulsan
Dept. of Public Administration, Korea University
Mi Young Lee · Tae Il Kim

(2001. 10. 22 접수)

Abstract

The purpose of this study is to examine methodological issues of empirical research using survey methods in Journal of the Korean Society of Clothing and Textiles. We reviewed 95 articles from Journal of the Korean Society of Clothing and Textiles in 1996, 1998, and 2000. Methodological issues were categorized into four areas: questionnaire and scale development, sampling and collecting data, variable construction, and external validity(generalization). In each area, a check list was made and 95 articles were reviewed based on this list. Major findings were discussed and some suggestions were made for future research.

Key words: survey method, measurement, sampling, external validity;
설문조사연구, 측정, 표본추출, 외적타당성

I. 서 론

본 연구의 목적은 1996년도, 1998년도, 2000년도 한국의류학회지 게재 논문들 중에서 주관적 인식에 대한 설문조사 연구들을 대상으로 측정도구 구성, 표본추출, 변인값 도출, 결과의 일반화 과정을 분석하고 문제점들을 논의하는 데 있다.

의류학 분야의 실증 연구는 실험(준실험) 연구와 비실험 연구로 구분할 수 있으며 비실험 연구는 다시 계측 등 객관적 자료에 의한 연구와 주관적 인식에 대한 설문조사 연구로 나눌 수 있다. 예를 들어 피복 특성을 분석하는 연구는 대부분 실험 연구에 속하며, 의복 착용과 관련한 연구는 계측에 의한 객관적 자료의 연구로 분류할 수 있다. 그리고 의류 소비자 특성에 대한 연구들은 거의가 주관적 인식에 대한 설문조사 연구에 해당한다.

실증 연구의 결과가 타당하려면 연구의 각 단계가 올바른 방법에 의하여 수행되어야 한다는 것은 모든 실증 연구에서 마찬가지이다. 그런데 주관적 인식에 대한 설문조사 연구는 '비실험 연구'로서 연구대상의 '주관적 인식'을 '설문조사'한다는 특성으로 인하여 다른 실증 연구와 구별되는 방법론적 특성을 갖고 있다. 예를 들면 비실험 연구이기 때문에 실험 연구와는

달리 각종 교란요인의 통제가 매우 중요하게 된다. 그리고 분석을 위한 변인값이 추상적 개념이며, 개개인의 주관적 인식을 통하여 파악되기 때문에 객관적 자료에 의한 연구와는 달리 측정의 타당성이 매우 중요하게 된다.

뒤에서 다시 논의하겠지만 한국의류학회지 게재 논문 중에서 대략 30% 정도는 주관적 인식에 대한 설문조사 연구에 해당하므로 의류학 분야의 연구 방법론 중에서 제법 큰 비중을 차지한다. 따라서 한국의류학회지 게재 논문 중에서 주관적 인식에 대한 설문조사 연구들을 대상으로 '비실험 연구로서 주관적 인식의 설문조사' 라는 특성에 기인한 방법론의 타당성을 평가하고 문제점들에 대하여 논의하는 것은 이 분야의 기존 연구결과를 활용하거나 향후 새로운 연구를 수행하는 데 도움을 줄 수 있을 것으로 판단된다.

설문조사에 의한 연구의 일반적인 절차는 먼저 연구자가 연구문제와 가설을 세우고 연구대상을 정한 다음, ①연구자가 얻으려는 자료 획득에 적합한 설문지 내용을 구성하고, ②표본을 추출하여 설문에 대한 응답으로부터 분석에 필요한 변수값들을 구한 후, ③이를 이용하여 필요한 통계 분석을 행하며, ④분석 결과를 해석하고 일반화하는 단계로 구성된다고 할 수 있다. 그런데 이 과정에서 세 번째 단계와 네 번째 단계의 결과 해석은 설문 조사 연구뿐만 아니라 통계 분석을 사용하는 모든 경험적 연구에서 공통된 것이다.

본 연구에서는 이와 같이 모든 경험적 연구에서 공통되는 과정에 대한 논의는 제외하고, 설문조사 연구라는 특성과 보다 밀접하게 관련되어 있는 설문구성, 표본추출 · 조사, 변인값 도출, 결과의 일반화 과정에서 발생하는 방법론적 특성에 대하여 분석한다.[1]

이를 위하여 먼저 각 과정에서 논점이 되는 사항들을 제시한 후, 이들이 기존 연구들에서 어떻게 반영되고 있는가를 검토한다.[2]

1) 물론 분석 결과의 일반화 역시 설문조사연구 뿐만 아니라 대부분의 경험적 연구에서 해당되는 문제이다. 그러나 뒤에서 논의하듯이 이는 설문조사 연구에서 특히 문제가 되므로 본 연구의 논의에 포함하기로 한다.

II. 분석대상

본 연구가 분석대상으로 삼은 논문은 1996년도, 1998년도, 2000년도 한국의류학회지 게재논문들 중에서 개인의 주관적 인식에 대한 설문조사로 수행된 것들이다. 이에 해당하는 연구가 전체 논문 중에서 차지하는 비중은 어느 정도이며 1996-2000년 사이에 이 비중이 어떻게 변하였는가를 보기 위하여 이 기간의 한국의류학회지 게재논문들을 다음과 같은 네 가지 유형으로 분류하였다.

①질적 연구: 질적인 방법에 의한 연구를 지칭한다. 이론 연구 이외에 사례 연구나 면접 조사 연구도 이에 포함한다.

②실험 연구: 실험 또는 준실험설계에 의한 연구를 지칭한다.

③측정 연구: 객관적 자료의 조사에 의한 연구를 지칭한다. 따라서 설문조사에 의한 연구라도 응답자의 신체치수, 의복지출비, 즐겨 입는 옷 등과 같이 응답자의 객관적 특성 조사가 주목적인 연구는 이 범주에 포함된다.

④설문조사 연구: 비실험설계에 의한 설문조사를 실시한 연구로서 개인의 주관적 인식 조사가 주목적인 연구를 지칭한다. 따라서 개인의 주관적 인식을 조사한 연구라도 실험설계에 의한 것은 이 범주에서 제외하며 실험 연구의 범주에 포함한다.

논문에 따라서는 복수의 범주에 포함되는 연구를 실시한 경우도 있는데 이런 경우는 주된 연구가 무엇인가에 따라서 분류하였다. 예를 들어 어떤 연구는 응답자의 신체치수와 의복유형 등의 객관적 자료의 조사와 함께 주관적 인식(예. 착용감)을 묻기도 하였다. 이런 경우는 연구의 주목적이 객관적 특성의 측정과

2) 본 연구와 같이 어떤 특정 학문 분야의 실증 연구를 대상으로 방법론적 특성을 분석한 국내 연구의 예로는 교육학 및 영양학 분야의 최종후 · 이재창(1990), 경영학 분야의 윤기중 · 안윤기 · 김병수(1987), 행정학 분야의 김태일 · 서주현(1998) 등이 있다. 그리고 의류학 분야의 외국 연구의 예에는 Kang-Park & Sieben(1993), Lennon, Johnson & Park(2001) 등이 있다.

주관적 인식 조사 중에 어디에 있는가에 의하여 범주를 구분하였다.[3]

각 유형에 따라 게재논문들을 분류한 결과는 〈표 1〉에 제시되어 있다.

먼저 3년간 게재논문의 합계에 대한 통계를 보면 설문조사 논문은 전체의 28%정도로서 실험 논문 다음으로 많음을 알 수 있다. 그런데 각 연도별 설문조사 논문의 비중을 보면 1996년도 18.8%, 1998년도 27.7%, 2000년도 37.2%로서 최근에 이를수록 증가하였음을 알 수 있다. 물론 특정 학술지를 대상으로 한 3개년만의 관찰 결과이므로 이것이 의류학 분야 연구 경향의 일반적 추세를 보여주는 것인가는 판단하기 어렵다.

〈표 1〉 연구유형별 게재 논문 수

(단위: 편, %)

	이론	실험	측정	설문	계
1996년도	17(16.8%)	41(40.6%)	24(23.8%)	19(18.8%)	101(100%)
1998년도	17(15.2%)	46(41.1%)	18(16.1%)	31(27.7%)	112(100%)
2000년도	16(13.2%)	38(31.4%)	22(18.2%)	45(37.2%)	121(100%)
계	50(15.0%)	125(37.4%)	64(19.2%)	95(28.4%)	334(100%)

이하에서는 설문조사에 의한 연구논문 95편을 대상으로 설문조사 연구의 각 단계별로 연구가 어떻게 수행되었는가를 분석한다.

Ⅲ. 연구의 각 단계별 분석 결과[4]

1. 측정도구 구성 단계

측정도구 구성 단계에서 검토하는 사항들은 다음과 같다.

· 문항 구성의 근거	· 문항 내용의 제시
· 척도의 수준	· 척도의 범주 수

개인의 인식에 대한 설문조사를 통하여 알아내고자 하는 내용들은 대부분 직접적인 관찰이나 측정이 불가능한 추상적 · 주관적 개념에 해당한다. 따라서 이러한 개념을 측정하려면 먼저 조작적 정의를 통하여 해당 개념을 변수로 전환한 다음 이를 측정하기 위하여 적절한 측정도구(척도)를 구성해야 한다. 측정도구의 구성은 측정 개념의 속성을 나타내 주는 문항내용 구성 단계와 각 문항 척도에 대한 범주(category)를 부여하는 단계로 구분할 수 있다.

(1) 문항 내용 구성 단계

이 단계에서는 문항 내용 구성의 근거 및 문항 내용의 제시 여부를 검토한다. 추상적 · 주관적 개념의 속성을 나타내는 문항을 구성할 때는 동일한 개념을 측정한 선행연구에서 개발된 문항을 사용할 수 있다. 그리고 선행연구가 존재하지 않거나 적절하지 않다고 판단될 때는 연구자가 스스로 문항을 개발하여 사용할 수 있다. 그런데 연구자가 직접 문항을 개발하여 사용할 때는 그 문항의 내용이 무엇이며, 왜 그 문항들이 측정하고자 하는 개념의 속성을 반영한다고 생각하는지에 대한 근거가 제시되어야 할 것이다. 그리고 선행연구의 문항을 사용하는 경우에도 해당 문항의 타당성이 이미 입증되었으며 독자들에게 잘 알려져 있어서 그에 대한 추가적인 설명이 불필요한 경우가 아니라면 문항 내용과 함께 왜 그 문항이 타당한가에 대한 설명이 제시되는 것이 바람직하다. 즉 해당 문항의 측정 도구로서의 적합성을 독자들이 납득할 수 있도록 (또는 판단할 수 있도록) 문항 구성의 근거와 내용이 제시되는 것이 바람직하다.[5]

문항 구성의 근거 제시에 대한 통계는 〈표 2〉에 나타나 있다. 문항 구성 근거 제시 여부는 세 유형으로 나누어 파악하였다. '선행연구'는 선행연구의 문항을 사용

3) 이런 경우에 대부분은 연구의 주목적이 뚜렷하여 범주가 명확하게 분류되었다. 그러나 소수의 논문 중에는 범주 구분이 다소 애매한 경우도 있었다. 따라서 몇몇 논문의 경우는 다른 연구자가 분류하였다면 범주 구분이 달라질 수도 있었을 것이다. 그러나 이런 경우는 본 연구자의 판단으로는 5~6편 이내라고 생각된다.

4) 측정도구 구성 단계와 표본 추출 단계의 방법론적 논점은 김태일 · 서주현 (1998)의 논문에 제시된 체크리스트를 참조하였다. 이 부분에서는 별도의 인용주를 붙이지 않는다.

5) 물론 문항구성의 근거와 문항내용을 제시하였다는 것이 해당 문항의 타당성을 보장하지는 못한다. 하지만 적어도 독자들이 문항의 타당성 여부를 판단하는 데 도움을 줄 것이다.

하거나 이를 수정 보완하여 사용한 경우이며, '스스로 개발' 은 선행연구 문항을 이용하지 않고 직접 개발한 경우이다.[6]

〈표 2〉 문항의 근거

선행연구	스스로 개발	제시 안함	합계
89	4	2	95

주: 분석대상논문의 수가 95편으로 논문 수와 백분비가 거의 같다. 따라서 본문이나 표에서 논문 수를 나타내면서 백분비를 별도로 표시하지는 않기로 한다.

거의 대부분의 논문이 선행연구에서 개발된 문항을 사용하거나 이를 수정 보완하여 사용하였다. 그런데 이 논문들은 대부분 선행연구 문항을 직접 또는 수정 사용하였다는 것 이외에 추가적인 문항 근거에 대한 설명은 하지 않았다. 전술하였듯이 선행연구의 문항을 사용하더라도 해당 문항의 타당성이 이미 입증되었으며 널리 알려진 것이 아니라면 그 문항이 왜 타당한가에 대한 설명을 제시하는 것이 바람직하다. 단지 '선행연구(특히 외국의 연구)에서 사용' 하였다는 것이 그 문항의 타당성을 보장하는 것은 아니기 때문이다. 한편 연구자가 직접 문항을 개발한 경우에도 어떤 근거로 그런 문항을 개발하였는지에 대한 설명은 불충분하였다.

문항 내용의 제시 수준은 〈표 3〉에 나타나 있다. '문항 대부분 제시' 는 분석에 사용된 문항 내용의 거의 전부를 제시한 경우를 의미하며, '문항 일부 제시' 는 분석에 사용된 문항 중 일부에 대해서만 내용을 제시한 경우를 의미한다.[7]

〈표 3〉 문항의 내용

대부분 제시	일부 제시	제시 안함	계
52	18	25	95

6) 선행연구 문항과 스스로 개발한 문항을 모두 사용한 경우는 '선행연구' 로 분류하였다.

7) '대부분' 과 '일부' 의 분류 기준은 절반 이상이냐 미만이냐로 설정하였다. 그런데 실제로는 거의 모든 논문이 절반이 훨씬 넘게 제시한 경우 또는 절반에 훨씬 못미치게 제시한 경우 중 어느 하나에 속하였으며, 절반 내외에 머문 경우는 없었다.

분석대상 논문 95편 중 52편만이 분석에 사용된 문항 내용의 대부분을 제시하였으며 나머지는 일부만을 제시하거나 아예 제시하지 않았다. 독자들이 문항의 타당성 여부를 판단하려면 최소한 분석에 사용된 문항 내용이 무엇인지는 알 수 있어야 한다. 물론 문항의 일부만을 제시하거나 혹은 제시하지 않은 연구의 많은 경우는 선행연구의 문항을 사용한 경우이며, 그렇기 때문에 내용을 제시하지 않더라도 독자들이 문항의 타당성 여부를 알 수 있을 것으로 가정하여 제시하지 않은 것일 수 있다. 그러나 비록 선행연구의 문항을 그대로 사용한 경우라도 그 내용이 누구에게나 널리 알려진 것이 아니라면 독자들이 타당성 여부를 바로 알기는 어려울 것이다. 더구나 외국의 선행연구 문항을 사용한 경우라면 사회문화 환경의 차이와 번역의 한계로 인하여 해당 외국의 연구에서는 타당했던 것이라도 우리 나라의 연구에서는 타당하지 않을 수도 있다.

문항 내용의 타당성은 설문조사 연구의 타당성에 가장 핵심이 된다. 독자들이 분석에 사용된 문항 내용의 타당성을 판단할 수 있도록 문항 내용과 근거를 제시하는 데 좀 더 충실해야 할 것으로 생각된다.

(2) 범주 설정 단계

범주 설정 단계에서는 각 문항에 대한 응답을 몇 개의 범주로 구성하였으며, 각 범주를 구성하는 척도 수준을 무엇으로 간주하였는가를 검토한다.

척도의 수준은 명목(nominal), 서열(ordinal), 등간(interval), 비율(ratio) 척도로 구분되는데 척도의 수준에 따라 적용할 수 있는 분석 기법이 제한된다. 추상적 · 주관적 개념에 대한 개인의 응답은 본질적으로 명목 또는 서열 척도이며 등간 또는 비율 척도가 될 수 없다. 그러나 명목 또는 서열 척도는 사칙연산의 적용이 불가능하므로 분석기법이 크게 제약된다. 이에 따라 설문조사 연구에서는 대부분 본질적으로 서열 척도인 것을 등간 척도로 간주하고 수량적 분석을 실시한다.[8]

8) 서열척도의 등간 척도 간주에 대한 문제점에 대한 논의는 Borgatta & Bohrnstedt(1980), Mayer(1971) 등을 참고할 수 있다.

〈표 4〉 척도의 수준과 범주

척도	명목척도	서열척도						
		제시없음	4점	5점	6점	7점	서열척도간주	등간척도간주
논문 수	2	11	2	62	1	30	2	93

주: 여러 종류의 척도를 함께 사용한 논문은 각각의 척도에 중복 계산되었다. '명목척도'는 명목척도만 사용한 논문을 의미하며, 명목척도와 서열척도를 모두 사용한 논문은 '서열척도'로 분류하였다.

서열 척도를 등간 척도로 간주한다는 것은 서열 척도의 각 간격이 동일한 크기를 갖는 것으로 가정함을 의미한다. 예를 들어 [매우 그렇다, 그렇다, 그저 그렇다, 아니다, 매우 아니다]와 같은 5점 척도를 사용한다는 것은 매우 그렇다와 그렇다 사이의 거리는 그저 그렇다와 아니다 사이의 거리와 동일하며 그렇다와 아니다 사이의 거리의 절반이라고 가정함을 의미한다. 이러한 가정이 실제와 부합하는가의 여부는 분석의 타당성에 매우 중요하다. 척도의 수준과 범주 수는 〈표 4〉에 제시되어 있다.

분석 대상논문 가운데 명목척도만 사용된 경우는 2편이며, 나머지 논문에서는 모두 서열척도가 사용되었다.[9]

서열척도 중에서는 5점 척도가 가장 많이 사용되었다. 그리고 서열척도가 사용된 논문의 거의 대부분은 실제 분석에서는 등간척도를 간주하고 수량적 분석을 하였다. 서열척도로 가능한 분석(예. χ^2검정 등)만을 실시한 논문은 2편에 불과했다.

전술하였듯이 본질적으로는 서열척도인 주관적 응답결과를 등간척도로 간주하여 분석하는 것이 잘못된 것은 아니며, 일반적인 경향이기도 하다.[10]

그러나 이러한 경우에는 '서열척도의 등간척도로의 간주'에 따른 오차가 존재함을 고려하여 분석 결과의 해석에 좀 더 주의해야 할 것이다. 그러나, 서열척도를 등간척도로 간주하고 분석한 논문 중에서 이러한 점을 언급한 논문은 발견할 수 없었다.

2. 표본 추출 단계

이 단계에서 검토하는 논점들은 다음과 같다.

- 확률/비확률추출 여부
- 구체적인 표본추출방법의 제시
- 응답자의 일반적 특성 기술
- 설문의 응답률

(1) 확률/비확률추출

이 단계에서는 먼저 표본이 확률추출 되었는지 아니면 편의추출 되었는지를 검토한다. 표본이 모집단으로부터 확률추출 되지 않았다면 모집단 특성에 대한 통계적 추론은 원칙적으로 의미를 갖지 못한다. 물론 설문조사 연구에서는 확률추출이 불가능하거나 비용·시간 등의 이유로 비확률 표본추출에 의존해야 하는 경우가 자주 발생한다. 그러나 비확률추출에 의하여 연구가 수행된 경우에는 그로 인한 문제점을 인식하고 분석 결과의 해석에서 이를 반영해야 할 것이다. 확률추출 여부에 대한 통계는 〈표 5〉에 제시되어 있다.

〈표 5〉 표본추출 유형

확률추출	편의추출	알 수 없음	계
4	74	17	95

분석대상 논문 중 2편은 층화추출을 사용하였다고 밝히고 있으나 내용을 보면 편의추출을 하였다. 층화추출은 확률추출의 일종으로서 단순히 어떤 특성을 기준으로 모집단 비율대로 표본을 분리 추출하는 것뿐만 아니라 각기 분리된 그룹 내에서는 무작위 추출

9) 참고로 서열 척도를 사용한 논문은 거의 전부가 리커트(Likert) 척도를 사용하였다고 밝히고 있는데, 그 중의 일부 논문에서는 하나의 개념을 단일 문항으로 측정하고 있었다. 리커트 척도는 원래는 서열 척도이지만 등간 척도로 간주하여 사용되는 대표적인 척도법인데 이의 핵심은 하나의 개념을 복수의 문항으로 측정하는 것이다. 하나의 개념을 복수의 문항으로 측정하는 것은 측정 오차를 줄이기 위함이다. 리커트 척도에 대한 설명은 채서일(1992) 등을 참고할 수 있다.

10) 서열 변수를 등간 변수로 가정하고 분석하는 것의 타당성에 대한 논의는 Borgatta & Bohrnstedt(1980), Mayer (1971), Laborvitz(1970) 등을 참고할 수 있다.

이 이루어진 것을 지칭한다. 그러나 이 2편의 연구는 각기 분리된 그룹 내에서 편의추출을 하였으므로 이는 층화추출이 아닌 편의추출에 해당한다. 따라서 이 논문들은 편의추출로 분류하였다. 한편 확률(층화)추출로 분류된 4편의 연구도 표본추출 방법은 확률(층화)추출을 사용하였다고 밝히고 있을 뿐 구체적으로 어떻게 확률(층화)추출을 하였는지는 제시하지 않았다. 따라서 이 논문들 중에도 앞의 2편처럼 실제로는 편의추출을 한 것이 포함되어 있을 수 있다. '알 수 없음' 으로 분류된 17편은 추출방법을 제시하지 않은 경우인데 이들 중 거의 대부분은 편의추출이었을 것으로 짐작된다.

거의 대부분의 연구가 편의추출을 한 것은 의류학 분야 설문 연구의 특성 상 당연하다고도 할 수 있다. 즉 선거 여론 조사와 같이 일반 국민을 대상으로 하는 것이 아니라 대부분이 '특정 의류 소비자' 와 같이 한정된 집단을 대상으로 하는 것이므로 무작위 추출은 힘들 수 밖에 없다. 따라서 '편의추출을 하였다' 는 것 자체가 문제라고 할 수는 없다. 그러나 '편의추출' 은 모집단에 대한 무작위추출이 아니라는 점 때문에 분석에서 많은 한계를 지닌다. 따라서 '편의추출' 로 인한 한계점들을 충분히 고려한 분석이 이루어져야만 한다. '편의추출' 로 인한 한계점들을 보완하려면 후술하는 구체적인 표본추출방법의 제시, 응답자의 일반적 특성 기술, 결과의 일반화 한계 지적 등이 이루어져야 한다.

(2) 구체적인 표본추출 방법의 제시

표본추출 방법의 제시는 확률추출된 경우에도 실제로 모집단에 대한 무작위추출이 이루어졌는가를 판단하기 위하여 필요하지만, 편의추출의 경우에는 더욱 요구된다. 확률추출된 경우는 표본 특성과 모집단 특성의 차이를 표본오차 추정으로 판단할 수 있다. 그러나 편의추출된 경우는 표본오차로부터 이를 알 수 없다. 따라서 이 경우에는 표본이 추출된 구체적인 과정으로부터 표본의 대표성 정도를 '짐작' 할 수밖에 없다. 그러므로 편의추출된 경우에는 특히 표본추출의 구체적인 방법을 제시하는 것이 필요하다. 〈표 6〉은 표본추출과정 제시 여부에 대한 통계를 보여 준다.

〈표 6〉 구체적인 표본추출방법 제시

제시함	제시 안 함	계
35	60	95

표본추출 과정을 제시한 경우보다 안 한 경우가 훨씬 더 많은데 이는 연구자들이 '편의추출' 된 표본이 갖는 문제점에 대하여 충분히 고려하지 않은 경우가 많음을 보여 준다.

(3) 표본의 일반적 특성에 대한 기술

표본의 일반적 특성 기술 역시 확률추출된 경우에도 필요하지만, 편의추출의 경우에는 더욱 요구된다. 편의추출된 경우는 표본의 대표성 정도를 '짐작' 할 수 밖에 없는데 이를 위해서는 구체적인 표본추출 과정의 제시와 함께 추출된 표본의 일반적 특성을 기술하는 것도 필요하다. 표본의 기술 통계치를 살펴봄으로써 표본의 대표성 정도를 어느 정도 판단할 수 있기 때문이다. 그리고 만일 모집단의 기술 통계치를 구할 수 있다면 이를 함께 제시하는 것이 이러한 '판단' 에 더욱 도움을 줄 것이다. 〈표 7〉은 표본의 일반적 특성의 제시 여부에 대한 통계를 보여 준다.

해당없음이란 표본 자체가 매우 제한적(예. 남자 고등학생)이어서 특별히 표본의 일반적 특성을 제시할 실익이 없었던 경우이다.

총 95편 중에서 68편이 표본의 일반적 특성을 제시

〈표 7〉 표본의 일반적 특성 제시

제시	제시 안함	해당 없음	계
68	25	2	95

하고 있어서 다수의 논문에서 이 조건을 충족하고 있었다. 그런데 일반적 특성을 제시한 논문 중에는 이의 제시가 간략하여 표본의 대표성 수준을 파악하기에 충분하지 못한 경우도 많았다.

(4) 설문의 회수율

설문의 회수율이 낮은 경우에는 비록 확률추출이 이루어졌더라도 편의추출과 마찬가지의 효과를 가져온다. 이는 설문의 회수율이 낮은 경우에는 응답자의 자기선택(self selection)으로 인한 편의(bias)가 발생할 수 있기 때문이다. 따라서 설문조사 연구에서는 설문의 회수율을 밝혀야 하며, 회수율이 낮다면 이로 인한 대표성 저하 문제를 고려해야 한다. 〈표 8〉에는 설문지

〈표 8〉 설문지 회수율

	회수율을 알 수 있는 경우의 회수율					알 수 없음	계
	0%~40%	41%~70%	71%~90%	91%~100%	소계		
논문수	3	5	10	28	46	49	95

회수율의 제시 여부와 회수율에 대한 통계가 제시되어 있다.

배부한 설문지 수와 회수된 설문지 수를 명확하게 밝히지 않은 경우가 많았기 때문에 설문지의 회수율을 알 수 있었던 연구는 대략 절반 정도였다.[11] 회수율을 알 수 있었던 논문들의 회수율을 보면 60% 이상의 연구에서 90%가 넘는 회수율을 보였다. 회수율이 90% 이상이라는 것은 설문조사 연구로서는 이례적으로 높은 것이다.[12]

이와 같이 대부분의 연구에서 매우 높은 회수율을 보인 것은 편의추출이며, 면접조사인 경우도 많았기 때문인 것으로 해석된다.

3. 변인값 도출 단계

변인값 도출 단계에서 검토하는 사항은 다음과 같다.

- 신뢰도 검사 여부
- 크론바하 알파 계수값
- 구성 개념 타당도 검사를 위한 요인분석 실시 여부
- 요인분석과정(추출방법, 회전방법, 요인수 선택 기준, 부하량) 제시 여부
- 요인계수값의 크기
- 변인값 계산 방법

설문조사를 마치고 각 문항의 점수를 구한 후에는 이로부터 변인값을 도출하게 되는데 일반적으로 변인값 도출 전에 측정의 신뢰도와 구성 개념 타당도(construct validity)에 대한 문항분석을 한다.[13] 하나의 변인에 대하여 복수의 문항으로 측정도구를 구성한 경우에는 크론바하(Chronbach) 알파 계수를 이용하여 각 문항들이 동일한 개념을 측정하는가를 의미하는 신뢰도를 검사하는 경우가 많다. 그리고 많은 경우 탐색적(exploratory) 요인분석을 통하여 동일한 개념을 측정하는 문항들끼리 수렴하고 다른 개념을 측정하는 문항들과는 서로 구분되는가를 의미하는 구성 개념 타당도를 분석한다.[14]

측정의 신뢰도와 타당도는 설문의 구성 단계에서 우선 고려되어야 한다. 즉 문항을 만들 때 미리 각 변인 개념을 적절하게 측정한다고 생각되는 문항들을 설정해야 한다. 그러나 사전에 구성한 문항들이 실제로 신뢰도와 타당도를 갖추고 있는지는 알 수 없으므로 실제 문항 점수를 구한 후에는 이를 검사해 보고, 적절하지 않은 문항은 제거하는 과정을 거치는 것이 바람직하다고 할 수 있다.[15]

(1) 신뢰도 검사

먼저 신뢰도 검사를 하였는가를 검토한다. 그리고 신뢰도 검사를 한 경우는 크론바하 알파 계수값이 얼

11) 회수율은 $\frac{\text{회수된 설문지 매수}}{\text{배부된 설문지 매수}}$로 계산하였다. 이에 따라 회수된 설문지 매수를 밝히지 않고 배부된 설문지 수와 실제 분석에 사용된 설문지 매수만을 밝힌 경우는 회수율을 계산하지 않았다. 그러나 이 경우는 사실 상 회수율을 밝힌 것과 다름이 없다. 비록 그런 경우가 많지는 않았지만 〈표 8〉에 제시된 회수율 제시 여부 통계의 해석에서는 이를 고려해야 할 것이다.

12) 응답률은 설문의 내용과 대상에 따라 달라지는데 통상 무작위 추출에 의한 우편조사인 경우에는 25% 미만인 경우가 흔하다. 설문조사 방법에 따른 응답률에 대해서는 오인환(1992), 허명회(1998)를 참고할 수 있다.

13) 일반적으로 측정의 타당도는 내용 타당도, 기준 타당도, 구성 개념 타당도로 구분되는데 이 중에서 요인분석으로 판단할 수 있는 것은 구성 개념 타당도이다. 이에 대한 상세한 내용은 채서일(1992)을 참고할 수 있다.

14) 구성 개념 타당도는 동일한 개념을 측정하는 문항들끼리 수렴하는가를 보여주는 수렴 타당도, 다른 개념을 측정하는 문항들과는 서로 구분되는가를 보여주는 판별 타당도라 구분된다.

15) 물론 이 과정을 반드시 거쳐야 한다는 것은 아니다. 이론적으로 충분히 고려하여 문항을 구성하였고, 예비조사 등을 통하여 타당도를 확인하였다면 이러한 과정을 굳이 거칠 필요는 없을 것이다.

〈표 9〉 신뢰도 검사 여부와 크론바하 알파값

신뢰도 검사 실시한 경우의 크론바하 알파값							실시안함	계
0.4	0.5	0.6	0.7	0.8	0.9	제시없음		
2	10	26	20	9	1	4	23	95

마인가를 검토한다. 신뢰도 검사를 한 경우는 모두 크론바하 알파 계수를 측정하는 방법을 사용하였으므로 이 방법만을 고려한다. 신뢰도 검사 여부와 크론바하 알파값은 〈표 9〉에 제시되어 있다. 여기에 제시된 크론바하 알파값은 복수의 알파값 중에서 최소값만을 제시한 것이다.

80% 가까운 논문이 신뢰도 검사를 하였으며, 그 중에서 70% 이상은 최소 크론바하 알파값이 0.6 이상이었다. 크론바하 알파값이 0.6 이상이면 대체로 문항의 신뢰도에 문제가 없는 것으로 간주하는 관례에 비추어 보면 신뢰도 검사를 실시한 연구들의 신뢰도는 비교적 양호한 것으로 판단된다.

(2) 구성 개념 타당도 검사와 변인값 계산

먼저 구성 개념 타당도를 검사하기 위하여 요인분석을 실시하였는가를 검토한다.[16]

그리고 이를 실시한 경우에는 요인분석과정(추출방법, 회전방법, 요인 수 선택기준)을 제시하였는지의 여부와 요인계수(factor coefficient) 값은 얼마인가를 검토한다.

요인분석의 결과는 어떤 추출방법과 회전방법을 사용하였는지, 어떤 기준으로 요인 개수를 선택하였는가에 따라 달라질 수 있다. 이에 따라 요인분석 결과를 제시할 때는 요인의 추출 및 회전방법과 요인수의 선택 기준을 제시하는 것이 관례이다.

요인분석을 실시하면 주어진 문항들이 몇 개의 요인으로 분류되며 요인 개념과의 관련성 정도에 따라 각 문항에 요인계수값이 부여된다. 이 때 특정 요인에 속하는 것으로 분류된 문항의 요인계수의 절대값이 그 요인에 대해서는 비교적 크며, 다른 요인들에 대해서는 작아서 서로 뚜렷하게 구분되면 그 문항은 구성 개념 타당성을 갖춘 것으로 판단할 수 있다. 따라서 구성 개념 타당성을 파악하려면 요인계수값을 분석해야 한다.

주어진 문항들이 각 요인으로 분류되고 요인계수값이 부여되면 이 결과를 이용하여 변인값을 도출하게 된다. 변인값으로는 일반적으로 각 요인에 속한 문항들의 평균점수나 요인점수를 사용한다. 두 방법 중에서 어느 것이 더 타당하다고는 할 수 없으며 이는 각 문항과 요인과의 관계 및 요인분석 과정에서 발생하는 오차를 어떻게 가정하는가에 따라 달라진다. 요인분석 방법론을 전공하는 학자들 사이에서도 서로 선호하는 방법이 다르다(Kim & Mueller, 1978). 따라서 둘 중에 어떤 방법을 사용했는가 자체는 문제가 되지 않지만 어떤 방법을 적용하였는가는 밝히는 것이 바람직하다.

구성 개념 타당도를 검사하기 위한 요인분석을 실시한 연구는 전체 95편 중 55편이었다. 요인분석을 실시한 논문들 중에서 요인분석과정의 제시 여부와 변인값 도출 방법에 대한 통계는 〈표 10〉에 제시되어 있다.

요인분석을 실시한 논문들 중에서 대략 40% 정도는 추출방법과 회전방법을 밝히지 않은데 비하여 요인수 선택기준과 요인계수를 밝히지 않은 논문들은 상대적으로 적었다. 그리고 각각의 요인분석과정을 제시한 연구들을 보면 대부분이 요인추출은 주성분(Principal Component)분석, 요인회전은 Varimax에 의한 직교회전, 요인 수 선택기준으로는 Eigen값(1 이상)을 적용하였다.

각 문항마다 요인계수값은 요인의 수만큼 도출되는데, 같은 요인에 속한 문항 계수값들은 절대값이 크면

16) 요인분석은 탐색적(exploratory) 요인분석과 확인적(confirmatory) 요인분석으로 구분된다. 그런데 구성 개념 타당도 검사를 위하여 요인분석을 할 때는 일반적으로 탐색적 요인분석을 실시한다. 본 연구의 분석대상 논문의 경우에도 1편을 제외하고는 모두 탐색적 요인분석에 의하여 구성 개념 타당도를 검사하였다. 따라서 이하에서 요인분석이라고 할 때는 탐색적 요인분석을 지칭한다.

〈표 10〉 요인분석과정과 계수값, 변인값 도출 방법

추출방법		회전방법		요인 수 선택기준		최소계수값		변인값	
주성분분석	31	Varimax	27	Eigen값	40	0.3	2	요인점수	14
주축인자분석	2	Promax	1	Scree test	1	0.4	25	문항평균	5
		직교회전	3	모두 사용	8	0.5	13		
						0.6	5		
						0.7	1		
계	33	계	31	계	49	계	46	계	19

요인분석을 실시한 논문 55편과 각 항목의 해당 논문 편수(계)와의 차이는 해당 항목의 내용을 밝히지 않은 경우이다.
'직교회전' 은 varimax와 같은 구체적인 회전방법의 제시 없이 직교회전이라고만 제시한 경우이다. Eigen값과 Scree test를 모두 적용한 경우는 '모두 사용' 으로 분류하였다.

수렴 타당성이 있는 것으로 판단한다. 그리고 수렴 타당성이 있으면서 동시에 다른 요인에 속한 문항 계수값들의 절대값은 상대적으로 작은 경우 판별 타당성이 있는 것으로 판단한다. 〈표 10〉에 제시된 계수값은 각 문항이 속한 요인에 대한 계수값 중에서 최소값에 대한 것이므로 수렴 타당성 정도를 보여 준다. 최소계수값은 대부분이 0.4 이상으로서 수렴 타당성은 대체로 양호한 것으로 판단된다. 한편 판별 타당성을 판단하려면 다른 요인에 대한 계수값이 함께 제시되어야 하는데 대부분이 해당 요인에 대한 계수값만을 제시함으로써 수렴 타당성만을 판단할 수 있었으며, 판별 타당성은 판단할 수 없었다. 구성 개념 타당도 파악을 위하여 요인 분석을 실시한 것이라면 다른 요인에 대한 계수값도 함께 제시하는 것이 바람직하다.

그런데 이와 같이 요인분석에 의하여 구성 개념 타당도를 검사한 연구 중에는 요인분석 결과에 대한 해석이 적절하지 못한 경우도 있었다. 예를 들면 어떤 논문에서는 탐색적 요인분석을 통하여 내용 타당도를 검정하였다고 기술하였다. 그리고 다른 논문에서는 요인분석을 통하여 어떤 추상적 개념이 몇 개의 하위 개념으로 구성되어 있음이 밝혀졌다고 기술하였다. 그런데 이들은 모두 부적절한 해석이다.

일반적으로 요인분석을 적용할 때는 미리 연구자가 개념들 간의 구조를 이론적으로 설정한 뒤 각 개념들을 반영한다고 생각하는 문항들을 설정한다. 그리고 요인분석을 통하여 연구자가 같은 개념을 반영한다고 설정한 문항들끼리 동일한 요인으로 분류되며, 다른 개념을 반영하는 것으로 설정한 문항들은 다른 요인으로 분류되는가를 확인한다. 그런데 이는 전술한대로 구성 개념 타당도를 갖추었는가를 보여주는 것이지 내용 타당도를 갖추었는가를 보여주는 것은 아니다. 내용 타당도는 a라는 요인이 A라는 개념을 반영한다고 할 수 있는가에 대한 것으로서 이는 요인분석뿐만 아니라 어떤 통계 기법으로도 판단할 수 없다. 이는 측정도구인 설문문항이 측정대상 개념을 올바르게 측정할 수 있는가에 대한 것으로서 이론에 의하여 혹은 해당 전문가의 판단에 의하여 적정성을 가정할 수 있을 뿐이다. 그리고 이와 같은 구성 개념 타당도의 판단은 미지의 개념 구조를 규명하는 것과도 관련이 없다. 요인분석을 통하여 개념 구조가 밝혀졌다고 하는 것은 일종의 순환론적 오류를 범하는 것에 해당한다.

한편 요인분석을 실시한 연구 중 70% 정도는 어떤 방법으로 변인값을 도출하였는지 알 수 없었다. 변인값 도출 방법을 알 수 있었던 18편 중에서는 15편이 요인점수를 변인값으로 사용하였으며 문항평균점수를 사용한 경우는 5편이었다. 그러나 요인분석을 실시하지 않은 연구에서는 변인값으로 문항평균을 사용할 수밖에 없으므로 분석대상 논문 95편 전체 중에서는 요인점수와 문항평균 중에서 어느 방식이 더 많이 사용되었다고 하기는 어렵다.

4. 결과의 일반화 단계: 외적 타당성의 한계 기술

일반적으로 조사연구의 타당성은 내적 타당성과 외적 타당성으로 구분된다. 내적 타당성은 조사 자체가 적정하게 수행되었는가에 대한 것이며, 외적 타당성

은 조사 결과를 조사 대상의 범위를 넘어서서 일반화 시킬 수 있는가에 대한 것이다.

앞에서 기술한 측정도구 구성 단계, 측정도구 구성 단계, 변인값 도출 단계에 대한 논의는 모두 내적 타당성을 확보하기 위한 조건에 해당하는 것들이라고 할 수 있다. 물론 내적 타당성을 확보하려면 이와 같은 변인값의 획득까지의 단계뿐만 아니라 이를 이용한 통계 분석이 올바르게 이루어져야 한다. 그러나 통계 분석의 타당성에 대한 논의는 본 연구에서는 다루지 않으며 추후의 연구 과제로 남겨 두기로 한다.

일반적으로 응용 연구(applied research)에서는 내적 타당성에 못지 않게 외적 타당성이 매우 중요하다. 아무리 내적 타당성이 높더라도 외적 타당성을 갖추지 못하였다면 그 결과는 현실에 적용되기 어렵기 때문이다. 의류학 분야의 설문 연구도 다른 응용 연구와 마찬가지로 많은 경우 현실에의 적용을 염두에 두고 있기 때문에 외적 타당성의 확보는 매우 중요하다.[17]

외적 타당성이 문제가 되는 경우는 두 가지로 나눌 수 있다. 첫 번째는 표본추출이 무작위로 이루어지지 않은 경우이며, 두 번째는 표본추출의 모집단과 연구목적의 모집단이 다를 경우이다.

표본이 무작위로 추출되지 않은 경우, 즉 편의추출된 경우에는 해당 표본에 대한 통계분석으로부터 도출된 P값, 유의수준, 신뢰구간 등은 부정확하므로 이에 의한 모집단 특성의 추정 또는 검정 결과 역시 부정확하다. 그러므로 편의추출된 표본에 통계분석을 적용하여 모집단 특성을 추론할 때는 '편의추출' 로 인한 일반화의 한계를 논의하는 것이 바람직하다.[18]

무작위 표본에 의한 연구라도 연구목적의 모집단과 표본추출의 모집단이 일치하지 않는 경우에는 일반화의 문제점이 발생한다. 연구목적의 모집단이란 연구자가 연구결과를 통하여 일반화시키려는 대상을 의미하며 표본추출의 모집단이란 직접 표본추출의 대상이 되는 모집단을 의미한다. 표본에 대한 통계적 추론은 표본이 추출된 모집단에 대해서만 유효하므로 이 두 모집단은 일치해야 하는 것이 원칙이다. 그러나 설문조사 연구에서는 이 두 모집단이 일치하지 않는 경우가 흔히 발생한다. 가령 여대생의 의류 소비 행태를 분석하는 연구에서 서울 소재 2개 여자대학교에서 무작위로 500명을 추출하여 분석한 후, 분석결과는 우리 나라 여대생의 일반적인 의류 소비 행태인 것으로 해석하였다고 하자. 이 연구에서 연구목적의 모집단은 우리 나라 여대생 전체이지만 실제 표본추출의 모집단은 서울 소재 2개 여자대학교이다. 따라서 이 연구의 분석결과로부터 표본추출의 대상이 된 서울 소재 2개 여자대학교 학생들의 의류 소비 행태를 추론하는 것은 가능하지만 우리 나라 여대생 전체에 대한 특성을 추론하는 것은 통계학적 정당성을 결여한 것이 된다.

물론 설문조사 연구에서는 연구의 성격 상 또는 비용 · 시간 등의 이유로 연구목적 모집단의 일부만을 표본추출 모집단으로 선택해야만 하는 경우가 많이 발생한다. 이 경우 분석 결과의 해석을 표본추출 모집단으로만 제약하고, 그 범위를 넘는 일반화를 허용하지 않는다면 연구 결과의 활용성이 매우 제약된다. 따라서 표본추출 모집단의 범위를 넘어서는 일반화가 필요할 수 있다. 그러나 이러한 경우에는 두 모집단의 불일치에 따른 일반화의 한계를 인식하고 분석 결과의 해석에서 이를 논의하는 것이 바람직하다.[19]

이상의 논의에 따른 두 유형의 일반화 한계 제시 여부에 대한 통계는 〈표 11〉에 제시되어 있다. '편의추출' 은 편의추출에 따른 일반화 한계만을 지적한 경우이고 '모집단불일치' 는 연구목적 모집단과 표본추출 모집단의 불일치에 따른 일반화 한계만을 지적한 경우이다. 그리고 '모두 제시' 는 두 경우의 한계를 모두

〈표 11〉 외적 타당성(일반화)의 한계 제시 여부

편의추출	모집단불일치	모두 제시	제시 없음	계
2	37	8	48	95

17) 응용 연구에서 외적 타당성 확보의 중요성에 대해서는 McGrath & Brinberg (1983), Wells (2001), Winer (1999)을 참고 할 수 있다.

18) Wells(2001)는 단순히 한계를 기술하는 것만으로는 불충분하며, 같은 조사를 다른 표본을 이용하여 반복하는 (replicate) 등의 방법을 통하여 좀 더 외적 타당성을 높이기 위한 노력이 필요하다고 주장한다.

19) 연구목적의 모집단과 표본추출의 모집단의 불일치에 따른 문제점에 대한 논의는 Henkel(1976)을 참고 할 수 있다.

제시한 경우이다.

절반 이상의 논문에서는 어느 유형의 외적 타당성의 한계에 대한 논의도 제시되지 않았다. 그리고 이를 제시한 논문에서는 대부분 모집단의 불일치에 따른 한계를 지적하였으며 편의추출로 인한 문제점을 지적한 논문은 많지 않았다. 그리고 외적 타당성의 한계를 지적한 경우에도 대부분 논문의 말미에 한 두 줄 정도로 간략히 언급하는 데 그치고 있어서 이에 대한 강조가 좀 더 필요한 것으로 판단된다.

V. 맺음말

본 연구에서는 설문조사를 이용한 실증 연구의 방법론적 특성에는 어떤 것들이 있으며, 이들이 실제 연구에서 어떻게 반영되고 있는가를 한국의류학회지 게재 논문을 대상으로 분석하였다. '추상적 개념을 주관적 인식을 통하여 측정한다는 것' 으로 인하여 요구되는 방법론적 특성을 설문구성, 표본추출 · 조사, 변인값 도출, 결과의 일반화 단계로 구분하여 분석하였다. 그 결과 각 단계에서 요구되는 조건들이 충분히 충족되지 않은 경우가 발견되었다. 실증 연구의 결과가 타당하려면 연구의 각 단계가 올바른 방법에 의하여 수행되어야 함은 물론이다. 실증 연구의 수행자가 방법론의 타당성에 대하여 좀 더 신중한 고려가 있기를 기대한다.

참 고 문 헌

김태일 · 서주현(1998). 행정학 분야에서 설문조사를 이용한 연구의 방법론적 문제점 분석. 한국행정학보, **32**(3), 199−215.

오인환(1992). 사회조사방법론. 나남출판.

윤기중 · 안윤기 · 김병수(1987). 통계의 오용과 효율적 이용에 관한 연구. 산업과 경영, **24**(2), 3−37.

채서일(1992). 사회과학조사방법론. 서울: 학현사.

최종후 · 이재창(1990). 학술논문과 통계적 기법. 자유아카데미.

허명회(1998). 사회여론조사. 자유아카데미.

Borgatta, E. F. & G. W. Bohrnstedt. (1980). Level of Measurement: Once over again. *Sociological Methods and Research*. **9**, 147−160.

Henkel, R. E. (1976). Tests of Significance. Thousand Oaks, CA: Sage Publications.

Kang−Park, J. & Sieben, W. (1993). Research methods and statistical techniques in articles on social psychological aspects of clothing, 1970−1985. In S. Lennon & L. D. Burns (Eds.), Social Science Aspects of Dress: New Directions (pp. 185−196). Monument, CO: ITAA.

Kim, J., & Mueller, C. W. (1978). Factor Analysis. Thousand Oaks, CA: Sage Publications.

Laborvitz, S. (1970). The Assignment of Numbers to Rank Order Categories. *American Sociological Review*, **35**, 515−524.

Lennon, S., Johnson, K. & Park, J. (2001). Research Trends in Textile and Clothing: An Analysis of Three Journals, 1980−1999. *Family and Consumer Science Research Journal*, **30**(2), 117−139.

Mayer, L. S. (1971). A Note on Treating Ordinal Data as Interval Data. *American Sociological Review*. **36**, 519−520.

McGrath, J. E. & Brinberg, D. (1983). External validity and the research process: A comment on the Calder/Lynch dialogue. *Journal of Consumer Research*, **10**, 115−124.

Wells, D. W. (2001). The perils of N=1. *Journal of Consumer Research*, **28**, 494−498.

Winer, R. S. (1999). Experimentation in the twenty−first century: The importance of external validity. *Journal of the Academy of Marketing Science*, **27**(3), 349−358.