

尺度의 項目 數가 信賴度와 妥當度에 미치는 영향

趙 盛 謙

(서울대 新聞學科 助教)

目 次

- | | |
|---|---|
| <p>I. 서 論</p> <p>A. 문제의 제기</p> <p>B. 선행 연구 및 이론적 배경</p> <p>C. 연구문제와 가설</p> <p>II. 研究方法</p> <p>A. 참가자의 選定</p> <p>B. 尺度의 作成</p> | <p>C. 資料의 수집 및 分析 方法</p> <p>III. 研究結果와 論議</p> <p>A. 신뢰도에 미치는 영향</p> <p>B. 반응틀에 미치는 영향</p> <p>C. 적도의 방향성 여부</p> <p>D. 연구결과에 대한 종합적 논의</p> <p>IV. 결 論</p> |
|---|---|

I. 서 論

A. 문제의 제기

측정의 문제는 社會科學 研究의 발전에 중요 장애가 되어 왔다.¹⁾ 이런 문제 중의 하나로 잠재적(latent)으로 연속적인 變因을 제한된 數의 項目(category)의 척도로 측정할 때 발생하는 측정의 비정확성(measurement imprecision)을 들 수 있다.

E.G. Carmines와 J.P. McIver가 주장하듯이 社會科學에서의 중요한 變因들은 대부분 직접 관찰 또는 측정이 어렵다.²⁾ 따라서 이렇게 직접적으로 측정하기 어려운 變因들(unmeasured variables)은 이들을 대표하는 경험적(또는 측정된) 지표(indicators)들을 이용하여 간접적으로 측정할 수 밖에 없게 된다. 그리고 “관찰되기 어려운 이러한 構因(construct)들은 일반적으로 잠재적 차원(latent level)에서 연속성을 지닌 것으로 개념화 되기³⁾ 때문에 構因에 대한 간접적인 측정척도가 얼마나 타당하게(validly) 그리고 신뢰할 수 있게(reliably) 構因을 대표하는가 라는 것이 중요한 문제로 제기된다. 즉 尺度의 신뢰도와 타당도가 측정에서 중요한 위치를 차지한다. 이 점은 측정의 체계적 오차(systematic error) 및 무선적 오차

1) Hubert M. Blalock, “The Presidential Address: Measurement and Conceptualization Problems: The Major Obstacle to Integrating Theory and Research,” *American Sociological Review*, 44 (1979), p. 882.

2) Edward G. Carmines and John P. McIver, “Analyzing Models with Unobserved Variables: Analysis of Covariance Structures,” in G.W. Bohrnstedt and E.F. Borgatta (Eds.), *Social Measurement: Current Issues* (London: Sage, 1981), p. 66.

3) E.F. Borgatta and G.W. Bohrnstedt, “Level of Measurement,” in Bohrnstedt and Borgatta, 앞의 책, p. 28.

(random error)가 통계적 검증에 미치는 영향을 보아도 쉽게 알 수 있다. 측정오차는 變量分析의 검증력을 저하시킬 뿐만 아니라,⁴⁾ 측정치와 측정되지 않은 변인간의 상관정도를 심각하게 감소시킴으로써 假說 검증과 회귀분석의 경우 모수치 추정(parameter estimate)을 왜곡시키기⁵⁾ 때문이다.

커뮤니케이션 분야의 研究에서도 態度 및 태체의 이용 동기등 주요 개념들이 多問項(multi-item)으로 구성된 복합척도(composite scale)로 측정되고 또한 이 척도들이 대부분 Likert型 또는 평가척도라는 점에서, 이러한 척도의 신뢰도 및 타당도에 영향을 주는 要因들에 대한 관심이 요청된다. 태도나 동기를 측정하는 척도는 반응중심의 척도(response-centered scale)에 해당하는 것으로, 자극과 자극간의 관계에 입각한 응답자의 반응 즉 태도가 아닌 판단을 측정하는 자극중심의 척도(stimulus-centered scale)와는 구분된다. Torgerson에 의하면, 자극중심의 척도는, 응답자가 자신의 태도와는 관계없이 자극에 대한 평가를 하도록 요구하기 때문에, 응답자의 편견에 의한 영향이 작아진다고 한다. 그러나 반응중심의 척도에서는 응답자의 태도, 감정 및 능력등이 반응의 주요 인자가 된다고 한다.⁶⁾

이러한 반응중심 척도의 응답에 영향을 미치는 주요 요인으로서, H. Schuman과 S. Presser는 중간항(middle point)의 포함 여부, 설문과 문항의 언어적 표현 및 설문 구성 형태등을 제시하고 있다.⁷⁾ G.D. Jenkins와 J.C. Taber에 의하면, Likert型 척도의 측정오차에 영향을 미치는 變因으로는 ① 응답자의 판단 정확성, ② 척도의 문항 수 및 ③ 문항의 項目(category)數등이 여러 연구에 의해 밝혀 졌지만, 이 중 項目數의 영향에 관한 문제가 가장 많은 연구 관심을 받아왔다고 한다.⁸⁾ 이러한 연구들은 특히 “가장 적절한 項目數는 얼마인가”에 초점을 두고 신뢰도, 타당도, 반응 시간, 중간 항목의 사용 정도 및 項目의 사용 비율등에 미치는 項目數의 영향을 밝히려 노력해 왔다. 그러나 이들 연구결과 제시된 적정 項目數는 최저 2개로부터 20개 이상까지 다양하다. E.P. Cox III는 50년간에 걸친 이러한 項目數에 관한 연구를 검토하고 나서 “비록 적정 項目數에 관한 공식(formula)은 가용하지 않지만, 척도로부터 얻어지는 정보의 質에 대해, 項目數와 함께 작용하는 중요 요인들에 대한 어느 정도의 이해를 얻을 수 있다”고 결론을 내리면서, “그러나 반응오차 및 반응편향에

4) 측정오차가 변량분석에 미치는 영향은 T. Anne Cleary, Robert L. Linn and G. William Walster, “Effects of Reliability and Validity on Power of Statistical Test,” in E.F. Borgatta and G.W. Bohrnstedt (Eds.), *Sociological Methodology* (London: Jossey-Brass, 1977), pp. 130~138 참조.

5) D.R. Johnson and Janes C. Creech, “Ordinal Measures in Multiple Indicator Models, A Simulation Study of Categorization Error,” *American Sociological Review* 48 (June 1983), p. 398에서 再引.

6) Warren S. Torgerson, *Theory and Methods of Scaling* (New York: John Wiley and Sons, 1958), p. 48.

7) Howard Schuman and Stanley Presser, *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording and Context* (Orlando: Academic Press, 1981).

8) G. Douglas Jenkins and Thomas D. Taber, “A Monte Carlo Study of Factors Affecting Three Indices of Composite Reliability,” *Journal of Applied Psychology*, 62: 4 (1977), p. 392.

관한 연구가 무엇보다 우선적으로 요청된다”고 지적하고 있다.⁹⁾ 여기에서 반응오차 및 편차란 곧 신뢰도와 타당도를 의미한다고 볼 수 있다. 따라서 적정 項目 數도 결국은 項目 數와 척도의 신뢰도 및 타당도의 관계를 통해 결정될 수 있을 것이다.

이와같이 척도의 적정 項目 數 즉 項目 數의 신뢰도 및 타당도에 대한 영향은 그 중요성에도 불구하고 아직 명백히 밝혀지지 않았다고 볼 수 있다. 또한 이에 대한 연구들 조차 주로 미국 대학생들을 대상으로 이루어져 왔기 때문에, 이들 연구결과들을 다른 부류의 응답자 특히 한국의 응답자들을 대상으로 한 척도에 그대로 일반화시키는 것은 무리가 있다고 본다. 따라서 본 연구는 복합척도의 項目 數가 그 척도의 신뢰도 및 타당도에 미치는 영향을 실증적 연구를 통해 살펴봄으로써, 척도의 적정 항목 수에 대한 결정에 도움이 되고자 한다.

B. 先行 研究 및 이론적 배경

項目의 數는 척도의 측정오차에 어떤 영향을 미치는가? 따라서 가장 적절한 項目 數는 얼마인가? 이에 대한 지금까지의 研究들은 그 적정 項目 數에 있어서는 물론, 측정오차와 項目 數의 관계에 있어서는 다양하고 때로는 상반되기도하는 결과를 보여주고 있다. 즉 첫째로 項目 數의 증가는 측정의 신뢰도 및 타당도를 높인다는 연구결과와, 둘째 項目 數의 증가는 오히려 타당도를 저해한다는 연구결과, 그리고 셋째로는 項目 數와 측정 오차는 아무런 관련이 없다는 연구결과가 있다.

1. 項目 數의 증가가 信賴度와 妥當度를 높인다는 研究結果들

潜在的으로 연속성이 假定되는 變因들을 제한된 數의 項目의 尺度로 測定할 때 여기에서 발생하는 測定誤差와 項目 數의 관계는 무엇인가? 즉 尺度의 測定值 X 와 潜在的 變因의 “정말 點數” t 의 $X=t+e$ 의 관계에서 項目 數의 증가는 測定誤差 e 에 어떤 영향을 미치는가?

이런 물음에 대한 답을 Monte Carlo 技法을 이용하여 구해 보려는 시도가 이루어져 왔다.

W.S. Martin은 정규분포를 이루고 상관정도가 각기 다른 두개의 變因들을 2개에서 20개까지의 類目으로 나누어 측정된 측정치를 모형화하였다. 이때 原變因간의 상관정도가 높을 수록 그리고 測定類目(category) 數가 적을수록, 測定值간의 상관계수 r 은 原變因간의 상관계수와 많은 차이를 보였다. 원변인간의 상관계수 $r=0.800$ 인 경우 각각 4개의 類目으로 측정했을 때의 測定值간의 상관도는 $r=0.604$, 10개의 유목으로 측정했을 때는 $r=0.777$ 이었다. 또한 그의 模擬實驗에서는 두개의 變因을 각기 다른 유목으로 측정하는 경우, 그 결과는 적은 유목에 의해 지배된다는 것이 발견됐다.¹⁰⁾

9) Eli P. Cox III, "The Optimal Number of Response Alternatives for a Scale: A Review," *Journal of Marketing Research* 17 (November 1980), p. 419.

10) Warren S. Martin, "Effects of Scaling on the Correlation Coefficient: Additional Considerations," *Journal of Marketing Research*, 15 (May 1978), pp. 304~308.

K.A. Bollen과 K.H. Barb의 模擬實驗에서도 위의 Martin의 연구와 비슷한 결과가 나왔다. 그의 연구결과에 의하면, 유목화된(즉 측정된) 變因간의 상관계수는 원변인간의 상관계수보다 언제나 작으나 類目 數가 증가하면, 原變因간의 상관계수에 접근하였다. 5개 이상의 유목이 사용된 경우 원래의 상관계수와와의 차이는 0.1보다 작게 나타났다.¹¹⁾ 이러한 Bollen과 Barb의 연구결과는, 그러나 앞의 Martin의 연구와 마찬가지로 潛在的 變因이 정규분포를 이룬다는 이상적인 상태를 가정한 實驗에서 나온 것이다.

그러나 유목 數의 증가는, 潛在的 變因의 分布 및 유목에 대한 점수부여 방식에 관계없이, 언제나 原變因과 측정치간의 상관관계를 높인다는 결과가 R.M. O'Brien의 模擬實驗에서 나왔다. 즉 O'Brien의 實驗에서는, 정규분포 및 四角分布(uniform distribution)가 偏布보다 상관관계가 높았으며, 점수부여 방식 및 유목의 한계분포에 의한 영향이 작게 나타났다고, 유목 數의 증가는 언제나 상관관계를 높혀 주었다.¹²⁾

한편, 이러한 測定 項目 數의 영향은 척도의 問項 數의 영향과 관계없다는 것이 Jenkins와 Taber의 模擬實驗에서 밝혀졌다. <표 1>이 보여 주듯이, “정말 값”에 의거한 척도치와 측정척도치의 상관계수 r 은 項目 數가 증가할 때, 또는 문항 數가 증가할 때 커지고 있었다. 그러나 양자간에 상호작용 效果는 없었다. 이러한 실험결과에 의거, 그들은, 2개 문항 14점 척도와 9개 문항 2점 척도로 동일한 신뢰도 수준을 얻을 수 있으며, 한편 5점 이상의 척도는, 증가되는 상관계수가 미미하기 때문에 사용할 필요가 없다고 주장하였다.¹³⁾

<표 1> Jenkins와 Taber의 모의실험 결과—“정말 값” 척도치와 측정척도치의 상관계수(r)의 평균치

		項 目 數						
		2	3	5	7	9	10	14
문 항 수	2	.551	.657	.718	.736	.744	.747	.752
	3	.604	.702	.759	.776	.783	.785	.790
	5	.680	.766	.813	.827	.833	.835	.839
	7	.725	.804	.845	.857	.863	.865	.868
	9	.756	.828	.865	.876	.880	.882	.885
	10	.769	.839	.874	.885	.889	.890	.893
	14	.810	.868	.899	.907	.911	.912	.915

이상의 模擬實驗 결과들에 의하면 항목 數의 증가는 尺度의 測定誤差를 감소시키고 있다는 것이다. 尺度에서 나타나는 測定誤差는, Johnson과 Creech¹⁴⁾에 의하면, 세가지로 분류될 수 있다. 첫째는 변형오차(transformation error)로서, 이는 測定된 項目간의 거리

11) Kenneth A. Bollen and Kenney H. Barb, "Pearson's r and Coarsely Categorized Measures," *American Sociological Review*, 46 (April 1981), pp. 232~239.

12) Robert M. O'Brien, "Using Rank Category Variables to Represent Continuous Variables: Defects of Common Practice," *Social Forces*, 59:4 (June 1981), pp. 1149~1163.

13) Jenkins and Taber, 앞의 논문.

14) Johnson and Creech, 앞의 논문, p. 398.

內在하는 變因값의 직선적 변형이 아닐 때 발생한다. 예를 들어, 잠재적 變因이 偏布를 보일 때 등간 점수를 부여하게 되면 변형오차가 발생하여 상관관계가 작아진다.¹⁵⁾ 둘째는 유목화 오차(grouping error)로서, 제한된 숫자로 연속적인 값을 나타내기 때문에, 때로는 낮게 때로는 높게 표현하게 되며, 여기에서 오차가 발생하게 된다. 셋째는 분류오차(misclassification error)로서, 潛在的 變因의 값이 분류되어야 할 項目이 아닌 다른 項目에 포함될 때 나타난다. 이 오차는 無選의일 수도 있고 또는 測定하고자 하는 變因이 아닌 다른 要因에 의해 체계적으로 나타날 수도 있다. 즉 반응의 고정 틀(response set), 일정한 반응경향 및 반응편차등의 영향에 의한 것일 수도 있다는 것이다.

이 세가지 유형의 오차 중에서, 위 Martin의 연구 및 Bollen과 Barb의 연구는 유목화 및 변형오차만을 그들의 實驗에서 고려하였다. 따라서 유목 數의 증가가 測定誤差를 감소시킨다는 이들의 研究結果는, 유목화 및 변형오차를 감소시킨다는 것으로 제한해서 해석되어야 할 것이다. 판단상의 오류를 고려한 Jenkins와 Taber의 실험의 경우도, 판단의 정확도가 項目 數에 관계없이 일정하다는 그들의 전제가 현실적 妥當性이 없다고 보겠다. 즉 項目 數가 많을 때(예를 들어 9점이나 11점, 13점 척도)가 적을 때(3점 또는 5점 척도)보다 項目 간의 차이 구분이 어려울 것이며 따라서 판단의 정확도도 감소할 것으로 보는 것이 좀 더 타당할 것이다.

한편 C.E. Osgood등은 5점 척도를 사용할 경우, 응답자들이 자신의 태도를 충분히 표현하지 못하게 되어 짜증을 보였다고 말하면서, 7점 척도의 사용시, 각 일곱개 項目에 대한 응답이 거의 비슷하기 때문에, 7점 척도가 사용되어야 한다고 주장하였다. 또한 9점 척도를 사용한 결과, 중간점과 극단치 이외의 項目이 별로 사용되지 않았다고 한다.¹⁶⁾ 이들의 연구 결과는 項目 數가 증가되더라도 어느 정도(7점)까지는 판단상의 오류 또는 특정한 반응들이 나타나지 않으며, 따라서 신뢰도가 증가될 것이라는 주장을 내포하고 있다.

이것은 Murphy와 Likert의 실험에서 부분적으로 입증된 바 있다. 즉 44개 문항으로 된 3점 척도 보다는 36개 문항 5점 척도가 신뢰도가 높았다(Spearman-Brown 계수가 0.88에서 0.94로 증가).¹⁷⁾ 그러나 Murphy와 Likert의 이러한 실험은 3점 척도와 5점 척도만을 비교했다는 제한점이 있다.

F.M. Andrews의 연구에서도, 6개의 대규모 조사 자료를 재분석한 결과, 項目 數가 많을 경우에 (20개 이상) 신뢰도와 타당도가 높게 나왔다.¹⁸⁾ 그러나 분석의 대상이 되었던 문항

15) O'Brien의 연구에서 6점 척도일 경우, 정규분포에서는 $r=0.952$, 자유도 1인 카이스퀘어 분포에서는 $r=0.88$ 이었다. Robert M. O'Brien, "Using Rank-order Measures to Represent Continuous Variables," *Social Forces*, 61:1 (September 1982), pp.144~155 참조.

16) Charles E. Osgood, George J. Suci and Percy H. Tannenbaum, *The Measurement of Meaning* (Urbana: Univ. of Illinois Press, 1958), p. 85.

17) Cox III, 앞의 논문, p. 417에서 再引.

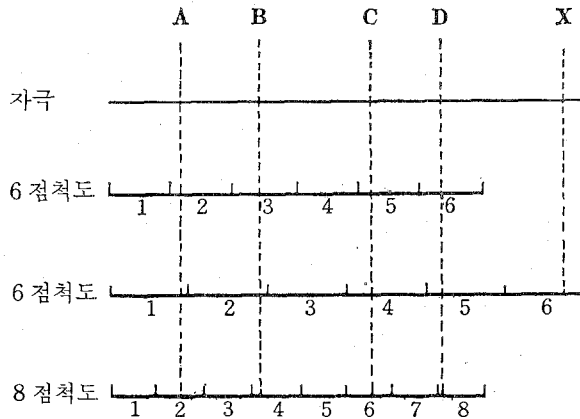
18) Frank M. Andrews, "Construct Validity and Error Components of Survey Measures: A Structural Modelling Approach," *Public Opinion Quarterly*, 48 (1984), pp. 409~442.

들이 태도에 관한 것이 아니라 'TV시청 빈도' 등에 관한 것이었고 또 그 방법도 간접 자료를 이용한 것이었기 때문에, 일반성이 부족하다고 보겠다.

이상에서 본 바와 같이, 項目 數의 증가가 신뢰도를 증가시킨다는 모의실험결과들은, 이상적인 응답자를 가정했다는 점에서, 項目 數와 測定誤差와의 관계를 충분히 설명하지 못하고 있다. 한편 Osgood 등이 비교적 項目 數와 測定誤差의 관계를 현실적 증거를 제시하면서 체계적으로 설명하고 있으나, 이에 대한 경험적 증거들이 부족한 편이다.

2. 項目 數의 증가가 尺度의 妥當度를 저해한다는 研究結果들

項目 數의 증가는 응답자의 반응에 어떤 영향을 주는가? H.S. Upshaw의 시각이론(Perspective Theory)에 의하면, 응답자는 제시된 일련의 자극을 판단할 때, 刺戟群의 범위와 尺度의 범위를 우선 일치시킨 다음, 이에 따라 尺度의 項目를 주관적으로 동등한 간격으로 구분한다고 한다. 즉 아래 <그림 1>과 같이 項目 數가 일정할 때, 자극의 범위가 증가되면 각 項



<그림 1> 자극과 척도와와의 관계에 관한 Upshaw의 그림

目的의 범위가 동등하게 증가한다. 그리고 項目 數만 증가할 경우에는 項目의 범위가 감소된다.¹⁹⁾ 즉 項目 數의 증가는 응답자의 판단 행위에 아무런 영향도 주지 않는다는 것이다.

그러나 일련의 자극에 대한 판단이 아닌 주관적인 태도표현의 경우에는 어떻게 될 것인가? Upshaw의 시각이론은 자극에 대한 판단을 측정하는 즉 자극중심적 척도에 관한 것이지만 태도측정의 경우 즉 반응중심 척도에 대해서도 시사하는 바가 크다. Upshaw에 의하면 응답자의 시각에 따라 尺度項目의 크기가 결정된다.²⁰⁾ 같은 태도를 가진 반응자들이라 하더라도, 尺度의 양 극단치가 표현하는 것에 대한 판단이 다를 경우(즉 시각이 틀린 경우) 尺度에 대한 응답은 다르게 나타난다는 것이다. 또한 여러 문항으로 된 尺度의 경

19) Harry S. Upshaw, "Social Influence on Attitudes and on Anchoring of Congeneric Attitude Scales," *Journal of Experimental Social Psychology*, 14 (1978), p. 329.

20) Harry S. Upshaw, "The Personal Reference Scale: An Approach to Social Judgement," in L. Berkowitz (Eds.), *Advances in Experimental Social Psychology*, 4 (New York: Academic Press, 1969), p. 369.

우에는, 문항간의 동질성에 대한 판단도 다르게 되어 尺度の 측정오차를 유발할 것으로 추측될 수 있다. 척도의 표현문이 뜻하는 바에 대한 해석이 개인마다 달라진다는 점은 강현두와 박은희의 연구에서도 밝혀졌다.²¹⁾ 이와같이, 개인의 태도가 아닌 표현문의 해석등과 같은 다른 요인에 의해 응답자의 반응이 달라지게 되면 척도의 타당성이 저해될 것이다.

J. Cronbach도 반응틀에 관한 연구에서, 태도의 측정에 세분화된 尺度(5점, 7점, 9점등)를 사용한다면, 개인적 해석에 따른 차이가 유발된다고 주장하고 있다.²²⁾ 그에 의하면, 項目數가 증가함으로써 얻게 되는 신뢰도의 증가는 단지 개인마다 특유한 반응틀(Response Set)²³⁾에 의한 것이며, 측정하고자 하는 신념 또는 태도와는 관계가 없는 즉 타당도가 없는 것이라고 한다. 이러한 반응틀은 특히 문항이 어렵거나 또는 애매할 때 더 강하게 나타난다고 한다.²⁴⁾ 또한 반응틀은 개인마다 다를 뿐 아니라 같은 질문에 응답하는 중에도 달라진다는 것이 C.H. Hui와 H.C. Triandis의 연구에서 밝혀졌다.²⁵⁾

S.S. Komorita와 W.K. Graham의 연구에서는 260명의 대학생들이 4개 집단으로 나뉘어 24개 문항의 사회성 척도와 26개 문항의 의미분별 척도에 응답했다. 분석 결과 이질적 문항으로 된 척도(사회성 척도)에서는 2점에서 6점으로 項目數가 증가했을 때 신뢰도가 증가한 반면에(α 계수가 0.62에서 0.74로), 동질적 척도에서는 변하지 않았다($\alpha=0.920, 0.916$). 그들은, 이질적 척도에서 신뢰도가 증가한 것은 반응틀 때문이라고 설명하면서, 따라서 妥當度가 저해된 것이기 때문에 2점 척도를 사용해야 한다고 주장하였다.²⁶⁾ 그러나 그들의 이러한 해석은 사후적 설명의 성격이 강하며, 그들의 연구 또한 반응틀에 대한 직접 검증을 하지 않았다는 점에서, 이들의 주장이 경험적 검증을 바탕으로 했다고는 볼 수 없다.

이상의 연구들에 의하면, 세분화된 척도를 사용할 경우 개인적 시각의 차이 및 반응틀의 영향으로 타당도가 저해된다고 볼 수 있다. 그렇다면 어느 정도 세분화된 척도에서 이런 현상이 나타나는가? Komorita와 Graham은 2점과 6점 척도를 비교하고, 6점 척도에서 나타났다고 말하고 있으나 그들의 연구는 반응틀을 직접 분석하지 않았기 때문에, 명백한 결론을 제시했다고 볼 수 없다. 7점 척도에서는 각 항목에 대한 응답이 비슷한 분포를 보

21) 강현두와 박은희, "척도의 항목에 대한 사람간의 차이에 대한 비교 연구," 신문학보 19집(1984), pp. 397~416.
 22) Lee J. Cronbach, "Further Evidence on Response Sets and Test Design," *Educational and Psychological Measurement*, 10 (1950), p. 22.
 23) Cronbach는 이러한 반응틀의 유형으로 ① 동조성향(무조건 찬성, 또는 반대) ② 회피성향 및 ③ 특정 항목 선호경향의 3가지를 들고 있다. (위의 논문, p. 3. 참조)
 24) 위의 논문, p. 22.
 25) C. Harry Hui and Harry C. Triandis, "The Instability of Response Sets," *Public Opinion Quarterly*, 49 (1985), pp. 253~260. 이 연구에서는 2백개의 문항이 사용되었다(9점 척도).
 26) S.S. Komorita and William K. Graham, "Number of Scale Points and the Reliability of Scales," *Educational and Psychological Measurement*, 25:4 (1965), pp. 987~995.

였으나 9점 척도에서는 그렇지 않았다는 Osgood등의 주장을 근거로 하여 볼 때는 9점 또는 그 이상의 척도에서 반응틀이 나타난다고 볼 수 있다. 물론 Osgood등의 연구도 반응틀을 직접적으로 검증한 것은 아니다. 그러나 Osgood등의 연구결과는 문항이 모호하거나 어려울 때 특히 반응틀이 나타난다는 J.P. Guilford²⁷⁾의 주장에 의해서도 지지된다고 볼 수 있다.

3. 측정 項目 數와 측정오차는 관계없다는 研究들

이와는 달리, 반응중심 척도 특히 Likert型 척도는 방향성만을 측정하는 것이며 따라서 項目 數와 測定誤差는 어떤 체계적인 관계를 보이지 않는다는 연구결과들이 있다. 例컨대, D. Peabody는 3개 척도 총 266개 문항(6점 척도)에 대한 대학생들의 응답을 방향성(direction) 부분과 강도(degree of extremeness)로 나누어 분석해 보았다. 그 결과, 척도 변량의 70~80%가 방향성에 의한 것이었고 단지 10%만이 태도강도에 의한 것이라는 점이 발견되었다. 즉 응답자가 문항에 대해 찬성 또는 반대하는가 라는 부분이, 얼마나 찬성 또는 반대하는가의 부분보다 점수에 대한 영향이 훨씬 크다는 것이다. 이것은 방향성 점수만으로 척도를 재구성했을 때 원 척도와 상관계수가 $r=0.90$ 이상이었다는 점에서도 알 수 있다.²⁸⁾ 또한 그의 연구에서는 태도의 강도를 나타내는 부분은 반응틀에 의한 것이었다.²⁹⁾

Peabody의 연구결과에 따르면, Likert형 척도의 점수는 한 사람이 주어진 태도 문항에 대해 찬성 또는 반대하는 정도를 표현하는 것이 아니라 단지 그 여부만을 반영하는 것이 된다.

M.S. Matell과 J. Jacoby는 각각 60개 문항으로 된 6개의 척도를 項目 數만을 달리해서 (2점부터 19점까지) 실험한 결과, 項目 數와 재측정 신뢰도 및 예언 타당도는 무관하다는 결과를 얻었다. 또한 각 척도 문항의 점수를 2점(項目 數가 짝수) 또는 3점(項目 數가 홀수)으로 축소시킨 척도의 신뢰도를 원척도의 신뢰도와 비교했을 때에도 통계적으로 有意的인 차이 ($p<0.05$ 에서)가 없었다. 따라서 Likert형 척도는 Peabody의 주장과 같이 태도의 방향성만을 나타내는 것이라고 결론을 내리고 있다.³⁰⁾ 그들은 또 다른 연구에서 응답자들이 사용하는 項目 數의 비율이 4점 이상의 척도에서는 일정하다는 것을 발견하고, 이 점에 비추어 보더라도 項目 數의 선정이 중요한 일이 아니라고 주장하였다.³¹⁾

27) J.P. Guilford, "Response Biases and Response Sets," in Martin Fishbein (Eds.), *Readings in Attitude theory and Measurement* (New York: John Wiley and Sons, 1967), p. 279.

28) Dean Peabody, "Two Components in Bipolar Scales; Directions and Extremeness," *Psychological Review*, 69:2 (March 1962), pp. 655~673.

29) 이것은 태도의 강도를 나타내는 점수가 모든 문항에서 정상관을 보였다는 점에서 알 수 있다.

30) Michael S. Matell and Jacob Jacoby, "Is there an Optimal Number of Alternatives for Likert Scale Items? Study I: Reliability and Validity," *Educational and Psychological Measurement*, 31 (1971), pp. 657~674.

31) Michael S. Matell and Jacob Jacoby, "Is there an Optimal Number of Likert Scale Items? Effects of Testing Time and Scale Properties," *Journal of Applied Psychology*, 56:6 (1972), pp. 506~509.

측정 項目 數와 測定誤差는 체계적인 관계가 없다는 이런 연구들은 그러나 이러한 현상에 대한 뚜렷한 설명을 제시하지 못하고 그들의 연구결과에 입각한 일반화만을 시도하고 있다. 그러나 이러한 현상은 이들 연구에서 사용된 문항 수가 많았다는 점에서 (Peabody의 경우 308個, Matell과 Jacoby의 경우 60個) 반응들의 영향 때문인 것으로도 생각될 수 있다. 6점 척도에서 반응들이 나타났다고 단정하기는 어렵지만 Hui와 Triandis의 연구결과에서 보듯이 측정 문항 數가 많을 경우, 같은 질문의 앞 부분과 뒷 부분은 반응률 지수가 다르게 나타났으며, 이것은 문항 수에 의한 영향 때문인 것으로 볼 수 있기 때문이다.

C. 研究問題와 假說

척도의 項目 數와 測定誤差에 관한 지금까지의 研究들은 위해서 보았듯이 명확한 說明을 제시하지 못하고 있다. 項目 數가 신뢰도를 증가시킨다는 Osgood등의 주장은 모의실험 研究 및 Murphy와 Likert등의 研究에 의해 부분적으로 입증되었으나 엄밀한 경험적 검증이 이루어지지 않았고, 또한 그 주장이 근거로 하는 연구들이 특정 부류의 사람만을(미국 대학생) 대상으로 했다는 점에서, 그 一般化에 제약이 있다.

유목 數가 증가하면 반응들의 영향 때문에 척도의 妥當度가 저해된다는 Cronbach등의 주장은 시각理論등에 의해서도 지지되고 있지만, 어느 정도 세분화된 척도에서 반응들이 妥當度를 저해하는가에 대해서는 명확한 답을 주지 못하고 있다. 한편 Likert형 척도는 방향성만을 측정하기 때문에 측정오차와는 관계 없고, 따라서 어느 項目 數의 척도나 연구의 편의에 의해 사용할 수 있다는 Peabody 및 Matell과 Jacoby의 연구결과와는 그런 현상이 나타나는 이유 및 다른 연구결과(즉 신뢰도가 증가한다는 연구결과들)에 대한 체계적인 설명을 제시하지 않고 있다.

1. 研究問題

本 研究는, 따라서, 다음과 같은 研究問題를 설정하고 이에 대한 답을 구함으로써 項目 數와 測定誤差간의 관계에 대한 설명을 제시해 보고자 한다.

- 〈研究問題 I〉 項目 數가 증가하면 그 측정의 신뢰도도 증가하는가?
- 〈研究問題 II〉 반응들의 영향은 어느 정도로 세분화된 척도에서 나타나는가?
- 〈研究問題 III〉 尺度는 태도의 강도를 나타내지 못하는가?

2. 假說과 그 設定根據

위와 같은 研究問題에 대하여 본 연구에서는 다음과 같은 假說들을 설정하고 그것들의 검증을 통해서 주어진 研究問題에 대한 해답을 얻어 보고자 하는데, 우선 그 假說들을 제시하면 아래와 같다.

- 〈가설 I〉 ① 項目 數가 증가하면 그 측정의 신뢰도도 증가할 것이다.

② 이때 인지수준이 높은 被檢者 集團이 낮은 집단보다 신뢰도가 더 높을 것이다.

〈가설 II〉 ① 반응틀은 9점 이상의 척도에서 가장 많이 나타날 것이다.

② 이때 인지수준이 낮은 피험자 집단에서 반응틀이 좀 더 많이, 그리고 더 적은 項目 數의 척도에서 나타날 것이다.

〈가설 III〉 척도는 태도의 강도를 반영할 것이며, 그 정도는 인지수준이 높은 집단이 더 클 것이다.

이상과 같은 假說을 설정하게 된 이유를 간단히 제시하면, 우선 尺度의 신뢰도가 項目 數의 증가에 따라 증가할 것이라는 예측은, 반응틀이 나타나기 전까지는 각 項目에 대해 응답자들이 판별적으로 응답할 것이라는 Osgood 등의 주장과, 반응틀이 나타나더라도 비록 타당도는 저해되지만 신뢰도는 증가할 것이라는 Cronbach의 주장에 근거한 것이다. 반응틀이 나타나는 項目 數에 대해서는 지금까지 研究들이 각기 다른 입장을 보이고 있기 때문에 정확한 예측은 힘들지만 項目간 판별이 어려울 것으로 보이는 9점 내지 7점 정도로 가정할 수 있다. 이러한, 項目간의 판별 여부가 반응틀 및 신뢰도를 결정한다는 본 연구의 입장은 인지수준이 각기 다른 집단의 비교를 통해서 밝혀질 수 있을 것이다. 즉 인지수준이 높은 집단은 낮은 집단 보다 項目간의 판별을 더 쉽게 할 것으로 기대되므로, 신뢰도도 더 높고, 반응틀도 더 늦게 나타날 것으로 가정될 수 있기 때문이다. 또한, 項目에 대해 판별적으로 대하기 때문에, 척도치는 강도를 표현하게 되고, 따라서, 축소 척도와의 상관관계는 작게 나타날 것으로 보였기 때문이다.

II. 研究方法

上記와 같은 假說을 검증하기 위하여, 본 연구에서는 중·고등학교 2학년 386명을 선정, 이들을 각 6개 총 12개 집단으로 나누고 이들로 하여금 TV시청동기와 TV에 대한 태도를 측정하는 4개 척도 18개 문항에 응답케 하였다. 각 집단에 배부된 설문은 응답 項目 數만 3점, 5점, 7점, 9점, 11점 및 13점으로 각각 달랐고 문항은 동일했다. 이렇게 하여 수집된 자료의 분석은 ① 신뢰도 계수 α , ② 반응틀 지수 및 ③ 원척도와 축소 척도와의 상관계수 r 에 대해 변량분석을 이용하여 실시했다. 자료수집 및 분석 방법을 상술해 보면 다음과 같다.

A. 참가자의 선정

참가자는 고등학교 2학년 2백명, 중학교 2학년 186명으로 모두 남학생이었다. 고등학교 2학년생과 중학교 2학년생을 참가자로 선정한 것은, 본 연구의 목적상 인지수준의 상대적

차이를 알아보기 위해서였다. 그러나 본 연구에서 사용된 척도가 TV시청에 관한 것이기 때문에, TV를 전혀 시청하지 않는 학생들이 있을 것으로 예측되는 고등학생 집단을 더 많이 포함시킨 것이다. 또한 지역적 특성 및 성별등 척도項目數의 요인에 의한 영향을 최소화하기 위해 중·고등학교 각 1개교에서 참가자 전원을 선정하였다.

B. 尺度の 작성

본 연구에서는 정보추구동기(4개 문항), 현실도피동기(4개 문항), TV친근도(5개 문항) 및 TV현실감(5개 문항)에 대한 尺度를 사용하였다. 정보추구 및 현실도피동기는 박은희의 연구³²⁾와 서울대학교 사회과학연구소의 시청자 조사³³⁾에서 사용된 문항을 이용하였다. TV친근도 척도는 TV매체에 대해 느끼는 친근감 정도를, 그리고 TV현실감 척도는 TV내용을 어느 정도 현실적으로 받아들이는가를 측정하는 것으로서 박은희의 연구³⁴⁾에서 사용된 것을 이용하였다. 본 연구의 척도 문항들은 앞의 연구들에서 이미 그 타당성이 검토되어 사용되었기 때문에, 실험재료로서의 적합성이 인정된다고 판단되었다.

응답項目은 중간항 有無의 영향³⁵⁾을 배제하기 위해 홀수로만 3점, 5점, 7점, 9점, 11점 및 13점의 6가지 평점척도로 하였다. 각 응답項目은 숫자로 구분하였고, 중간항 및 양 극단치에 “잘 모르겠다” “아주 찬성한다” “전혀 반대한다”를 표시하였다.

C. 자료의 수집 및 분석방법

자료는 1986년 4월 25일(고등학교)과 29일(중학교)의 양일간에 수집하였다. 수업시간을 이용, 척도를 배부하고, 응답요령에 대한 설명을 한 다음 일정시간 후에 회수하였다.

수집된 자료 중에서, TV시청시간이 하루 평균 30분 이내인 응답자 26명의 것은 제외하고 나머지 159명 분을 대상으로 자료 분석을 실시하였다. 최종 분석에 이용된 응답자 수는 다음과 같다.

〈표 2〉 분석에 사용된 응답자 數

인지수준	項目數	3	5	7	9	11	13	계
중 학교		34	33	26	26	34	28	181
고 등 학교		30	27	26	33	31	31	178
계		64	60	52	59	65	59	359

32) 朴恩希, “텔레비전 시청동기와 시청패턴과의 상관관계에 관한 연구,” 석사논문(서강대학교, 1982).

33) 서울대학교 사회과학연구소, TV시청형태 및 이용과 만족의 조사연구: 85년 조사 보고서.

34) 박은희, 앞의 논문.

35) 중간항이 미치는 영향에 대해서는 李義鎭, “커뮤니케이션 연구에 있어서 Likert 尺度의 利用에 관한 소고,” 高大 新聞放送研究所, 커뮤니케이션과학 7집(1985, 8월), pp. 51~86. 참조.

1. 신뢰도 계수의 產出 및 分析(가설 I의 검증방법)

신뢰도 계수의 계산 방법에는 재측정방법, 반분법, Cronbach의 α , θ , 및 Ω 등이 있으나 Cronbach α 가 가장 널리 쓰이고 있다.³⁶⁾ F.M. Nord와 M.R. Novick에 의하면 α 는 복합척도의 신뢰도에 대한 정확한 측정이 아니라 신뢰도의 하한선을 나타낸다고 한다.³⁷⁾ 따라서 척도 문항의 등가성(parallel)을 가정할 수 있을 때 α 는 정확한 신뢰도를 측정하게 되며, 이 가정이 어긋날 경우에는 요인분석에 기초하여 산출된 Ω 나 θ 보다 작은 값을 갖게 된다.⁴⁸⁾ 그러나 본 연구에서는 척도의 각 문항에 대한 가중치를 1로 하였기 때문에(이유는 뒷장에서 상술) 신뢰도 계수로서 α 를 산출하였다.

4개 척도가 사용되었으므로 각 집단별로 4개 총 48개의 α 값을 얻을 수 있었다.

신뢰도 계수 α 에 미치는 인지수준 및 項目 數의 영향은 다음과 같은 변량분석 방안을 이용해 분석하였다.

〈표 3〉 신뢰도 계수의 변량분석 방안 (N=48)

인지수준	항목수	3	5	7	9	11	13
중 학 교		4	4	4	4	4	4
고 등 학 교		4	4	4	4	4	4

2. 반응틀지수의 산출 및 분석(가설 II의 검증방법)

Hui와 Triandis에 의하면 반응틀은 다음 세가지 즉 ① 방향성 : 각 문항에 대한 응답 평균치, ② 반응영역 : 표준편차, ③ 극단치 선호 : 전체 문항에서 극단치를 선택한 회수로 측정할 수 있다고 한다.⁴⁸⁾ 그러나 본 연구에서는 項目 數와 방향성과는 관계가 있을 것이라고 예측되지 않기 때문에 방향성 부분을 제외한 반응영역 및 극단치 선호도와, Cronbach가 제시한 회피성향(중간項 선택도)으로 반응틀을 측정하였는데, 그 구체적인 측정방법은 아래와 같았다.

- ① 극단치의 선호도 : 각 개인의 18개 문항에 대한 응답 중 양 극단치에 응답한 數
- ② 중간항 선호도 : 각 개인의 18개 문항에 대한 응답 중 중간항에 응답한 數
- ③ 반응영역 : 각 개인의 18개 문항에 대한 응답의 표준편차

위 방법들은 Hui와 Triandis가 사용한 방법을 변형한 것이며 이렇게 개인별로 산출된 점수를 가지고, 변량분석과 일원변량분석을 하였다.

36) Edward G. Carmines and Richard A. Zeller, *Reliability and Validity Assessment* (London: Sage, 1979), p. 45.

37) Frederic M. Lord and Melvin R. Novick, *Statistical Theories of Mental Test Score* (London: Addison-Wesley Publishing Co., 1968), p. 94.

38) α , Ω 및 θ 의 관계에 대해서는 D.R. Heise and George W. Bohrnstedt, "Validity, Invalidity and Reliability," in Borgatta and Bohrnstedt (Eds.), 앞의 책, pp. 104-129. 참조.

38) Hui and Triandis, 앞의 논문, p. 254.

3. 축소 척도와 원 척도의 상관관계 분석

(1) 척도치 계산 방법

각 척도의 점수는 척도내 문항점수들의 단순산술평균으로 하였다. 각 문항이 동등하게 내재적 변인을 측정한다고 가정할 수 없을 때에는 통상 요인부하치를 이용한 가중 산술 평균(요인점수)을 구하게 되나, 본 연구에서는 다음과 같은 이유에서 가중치를 1로 하였다.

첫째는 각 척도별로 요인분석한 결과 각 집단마다 요인부하치가 다르게 나타났으며, 일정한 패턴을 찾기가 어려웠다. 이런 요인부하치의 차이를 각 집단별 차이(즉 인지수준에 따른 Factor 구조의 차이, 아니면 항목數의 변화에 따른 Factor 구조의 차이)라고도 볼 수 있으나 본 연구에서의 표본의 크기를 고려해 볼 때 표집오차의 영향 아니면 기타 측정되지 않은 어떤 요인들에 의한 차이도 상당 부분 반영된 것으로 해석되었다. 따라서 구체적인 요인부하치는 무시하는 것이 타당할 것으로 판단되었다. 둘째로는 본 연구에서 사용된 문항들이 다른 연구에서 타당성이 검토된 것이었기 때문이다. 이상의 이유에서 요인분석에 기반하지 않고 단순산술평균으로 척도점수를 계산하였다.³⁹⁾

(2) 축소 척도치의 계산 및 상관관계 분석

원 응답을 중간항은 0, 찬성은 1, 반대는 -1로 치환한 다음 척도치를 계산하였다. 이 축소 척도치와 원척도치의 상관관계 r을 구하여(각 집단별 4개, 총 48개) 이에 대해 앞의 <표 3>에서 보는 바와 같은 방법으로 변량분석을 하였다.

Ⅲ. 研究結果와 論議

A. 신뢰도에 미치는 영향(<가설 I>의 검증 결과)

項目數가 증가하면 신뢰도가 높아지는가? 그리고 인지수준이 높은 집단이 낮은 집단보다 신뢰도가 높은가? 이와 같은 연구문제에 대한 <가설 I-1> 및 <가설 I-2>를 검증해 본 결과, 다음 <표 4>와 <표 5>의 변량분석 결과가 보여주듯이 項目數 증가에 따른 신뢰도의 변화는 나타나지 않았다($p > .05$). 즉 “項目數가 증가하면 신뢰도가 높아질 것이다”는 본 연구의 <가설 I-1>은 부정되었다. 그 반면 “인지수준이 높은 집단이 낮은 집단보다 신뢰도가 더 높을 것이다”는 <가설 I-2>는 긍정되었다.

<가설 I>은 항목數가 증가하면, 어느 정도까지는 응답자들이 각 항목간의 차이에 판별적으로 대할 것이라는 즉 판단상의 오차가 발생하지 않을 것이라는 예측에 입각한 것이었다. 따라서 項目數 증가에 따른 유목화오차 및 변형오차의 감소로 인해 자연히 신뢰도가 증

39) 요인점수, 단순산술평균 및 기타 척도치 계산 방법의 비교에 대해서는 Jae-On Kim and Charles W. Mueller, *Factor Analysis: Statistical Methods and Practical Issues*, (London: Sage, 1978), pp. 60-73 참조.

〈표 4〉 4개 척도의 신뢰도 계수 α 의 평균치 (N=48)

인지수준	유 목 수	3	5	7	9	11	13	평균
중 학 교		.698	.499	.529	.621	.602	.661	.602
고 등 학 교		.750	.772	.667	.737	.769	.618	.719
평균		.724	.635	.598	.679	.685	.639	.66

〈표 5〉 변량분석 결과 (N=48)

변 량 원	자 승 화	df	변량추정치	F	유 의 도
인 지 수 준	.164	1	.164	12.519	.001
항 목 수	.080	5	.016	1.220	.320
인지수준×항목수	.114	5	.023	1.738	.151
오 차	.472	36			
전 체	.830	47			

가할 것으로 예측되었다. 그러므로 이러한 <가설 I-1>이 부정된 것을 응답자들이 항목 증가에 따라 판별적으로 응답하지 않았기 때문이라고 해석할 수도 있다. 하지만 이러한 해석은 반응틀이 나타나더라도 신뢰도는 증가할 것이라는 본 연구 및 Cronbach의 예상과 어긋난 것이다.

이 결과는 다음과 같이 해석해 볼 수 있다. 즉 각 개개인은 개별문항에 대해서는 항목 수가 증가되면서 판별적으로 응답한다 하더라도, 만약 각 응답자가 양 극단치의 판단에 대한 시각이 틀리고 또 문항간 유사성의 정도에 대한 판단이 다르다면, 전체적으로 척도의 신뢰도는 감소하게 될 것이다. 그러나 이러한 개인차로 인한 신뢰도의 저하가 항목 수가 증가하면서 얻게 되는 측정오차의 감소보다 더 크지 않았기 때문에 신뢰도가 항목 수 변화에 따른 증가를 보이지 않았다고 해석할 수 있다.

이렇게 볼 때, <가설 I-2>가 긍정된 것도 두가지 해석이 가능해진다. 즉 첫째는 본 연구에서 가정한 것으로서, 높은 인지수준 집단이 판별력이 높기 때문에 각 項目에 좀 더 판별적으로 대했다는 해석이다. 둘째는 인지수준이 높기 때문에 각 문항을 좀 더 유사한 것으로 인식했기 때문이라는 해석이다. 만약 <가설 I-1>의 부정에 대한 본 연구의 해석이 맞다면 <가설 I-2>의 긍정에 대해서도 두번째 해석이 좀 더 타당할 것이다. 그러나 이 절은 <가설 II>와 <가설 III>의 검증을 통해 좀 더 명확하게 되었는데, 그 결과를 살펴보면 아래와 같다.

B. 항목 수가 반응틀에 미치는 영향(<가설 II>의 검증 결과)

항목 수가 증가하면 반응틀은 언제 나타나는가? 이에 대해 본 연구는 그 동안의 연구

<표 6> 극단치 선호도의 평균치 (N=359)

인지수준	유 목 수	3	5	7	9	11	13	평 균
중 학 교		10.9	5.7	5.6	4.7	5.5	3.2	6.11
고 등 학 교		10.8	6.4	6.0	4.4	5.3	4.4	6.16
평 균		10.9	6.0	5.8	4.6	5.4	3.8	6.13

※ 18개 문항에서 극단치를 선택한 회수.

<표 7> 변량분석 결과 (N=359)

변 량 원	자 승 화	자유도	변량추정치	F	유 의 도
인 지 수 준	5.460	1	5.460	.357	.551
항 목 수	1,943.390	5	388.752	25.422	.000
인지수준×항목수	23.259	5	4.652	.304	.910
오 차	5,306.263	347	15.292		
전 체	7,278.372	358			

<표 8> 극단치 선호도의 Scheffe Test 결과

(N=359)

항목수	13	11	9	7	5	3
13						
11						
9						
7						
5						
3	*	*	*	*	*	

* : 두 집단의 평균치가 p=0.05수준에서 차이 있다는 것을 의미.

결과를 토대로 9점 이상의 척도에서 나타날 것으로 예측하였고<가설 II-1> 그러나 인지 수준이 낮은 집단은 이보다 더 많이, 그리고 더 적은 항목의 척도(7점 정도의)에서 나타날 것으로 가정했었다<가설 II-2>.

이러한 假說들에 대한 검증결과로서, 먼저 <표 6>과 <표 7>에서 극단치 선호도를 보면 인지 수준에 의한 차이는 나타나지 않았으나 항목 수에 의한 차이는 나타났다. 항목 數에서의 차이는 그러나 <표 8>의 Scheffe Test 결과에서 보듯이 3점 척도 집단에 의한 차이였다. 즉 극단치 선호도는 3점에서 5점 척도로 항목 數가 증가할 때는 감소하나 5점 이상의 항목 數 변화에는 어떤 차이를 보이지 않았다. 따라서 극단치 선호 반응들은 7점 척도부터 나타나고 있으며 이것은 인지수준과 무관하다고 볼 수 있다.

중간항 선호의 경우는 <표 9>와 <표 10>의 변량분석 결과가 보여 주듯이 인지수준에 의한 차이는 나타나지 않았다. 반면 항목 수에 의한 차이는 나타났으며, 이것은 <표 11>의 Scheffe Test 결과에서 알 수 있듯이 7점 척도까지는 선호도가 감소하나 9점 이상의 척도에서

〈표 9〉

중간항선택 평균치

(N=359)

인지수준	항목수	3	5	7	9	11	13	평균
중학교		7.1	4.0	2.7	2.8	1.7	2.01	3.48
고등학교		7.2	3.8	2.8	1.8	2.1	1.9	3.22
평균		7.1	3.9	2.8	2.2	1.9	2.0	3.36

※ 18개 문항 중 중간항 선택 수.

〈표 10〉

변량분석 결과

(N=359)

변량원	자승화	자유도	변량추정치	F	유의도
인지수준	1.472	1	1.472	.336	.563
항목수	1,270.725	5	254.145	57.987	.000
인지수준×유목수	20.131	5	4.026	.919	.469
오차	1,520.836	347	4.383		
전체	2,813.164	458			

〈표 11〉 중간항선호도 Scheffe-Test 결과

항목수	13	11	9	7	5	3
13						
11						
9						
7						
5	*	*	*			
3	*	*	*	*	*	

* : 두 집단의 평균치가 $p=0.05$ 수준에서 차이 있다는 것을 의미.

는 차이가 없었다.

즉 중간항 선호 반응틀도 인지수준에 따른 차이는 없었으나, 9점 이상의 척도에서 나타났다.

응답영역의 경우 〈표 12〉와 〈표 13〉의 변량분석 결과, 인지수준에 따른 차이는 없었다. 그리고 Scheffe Test 결과(〈표 14〉), 항목 수에서의 차이는 각 항목마다 나타나고 있는 것으로 해석할 수 있다. 즉 반응영역은 인지수준에 따른 차이는 없었으나 항목 수가 증가할 때 이에 따라 같이 증가했다. 따라서 응답 영역에서의 반응틀은 나타나지 않았다고 볼 수 있다.

이상의 결과로 볼 때, 〈가설 II-1〉은 부분적으로 입증되었다고 볼 수 있다. 즉 항목 수가 3점에서 5점으로 증가할 때는 극단치 선호, 중간항 선호 및 반응영역에서의 반응틀이 나타나지 않았다. 타당도가 저해되지 않은 것이다. 그러나 7점 이상의 척도에서는 극단치 및 중간항 선호 반응틀이 나타나기 시작했다. 즉 9점 척도 이상에서 반응틀이 나타날 것으로 예측했던 본 연구의 가설은, 비록 구체적인 항목 수에서는 정확하지 않았지만, 5점 척도까지 반응틀이 나타나지 않음으로써 어느 정도 입증되었다고 본다. 한편 인지수준에서

〈표 12〉 집단별 응답영역 평균치 (N=359)

인지수준	항목수	3	5	7	9	11	13	평균
중학교		.746	1.29	1.99	2.46	3.53	4.72	2.41
고등학교		.76	1.31	1.97	2.52	3.16	3.63	2.33
평균		.753	1.30	1.98	2.50	3.25	4.15	2.37

※ 18개 문항에 대한 개인별 응답의 표준편차.

〈표 13〉 변량분석 결과

변량원	차승화	자유도	변량추정치	F	유의도
인지수준	4.864	1	4.864	3.135	.249
항목수	497.981	5	99.596	27.344	.000
인지수준×항목수	14.873	5	2.975	.817	.538
오차	1,263.873	347	3.642		
전체	1,781.591	358			

〈표 14〉 응답영역 Scheffe Test 결과

집단	3	5	7	9	11	13
3						
5						
7	*					
9	*	*				
11	*	*	*			
13	*	*	*	*		

* : 두 집단의 평균치가 $\alpha=0.05$ 수준에서 차이가 있다는 것을 의미.

는 어떠한 반응틀도 차이가 나타나지 않음으로써 〈가설 II-2〉는 부정되었다.

〈가설 II〉는 항목 수가 어느 정도 증가할 때 까지는 판별적으로 각 항목을 대하나, 항목 수가 많아지게 되면 항목간의 차이 구분에 노력하지 않고, 특정한 반응틀을 보일 것이라는 가정에서 설정된 것이었다. 그렇기 때문에 인지수준이 높은 집단은 항목간의 차이를 좀 더 잘 판별할 것으로 기대되어 반응틀이 적게 나타날 것으로 예측되었다.

그러나 〈가설 II〉의 검증 결과, 예측과는 달리 인지수준이 높다고 해석 항목간의 차이를 좀 더 잘 판별하지 않는다는 것으로 나타났다. 또한 항목간의 차이를 판별할 수 있는 항목수는 5개까지였으며, 이는 Matell과 Jacoby 및 Peabody 등의 연구결과와도 일치한다고 볼 수 있다.

C. 척도의 방향성 반영 여부(〈가설 III〉에 관한 검증 결과)

척도는 과연 태도의 강도를 표현하지 않는가? 이에 대한 〈가설 III〉의 검증결과로서 〈표 15〉와 〈표 16〉의 원척도와 축소 척도의 상관관계를 보면 Matell과 Jacoby의 연구결과와 비슷하게 나타났다(평균 $r=0.92$). 따라서 방향성을 주로 반영한다고 볼 수 있다. 한편, 변

〈표 15〉 원척도와 축소 척도의 상관관계 r의 평균치 (N=48)

인지수준	항목 수	3	5	7	9	11	13	평균
중 학 교		1.00	.949	.939	.928	.909	.628	.892
고 등 학 교		1.00	.957	.946	.943	.947	.888	.947
평 균		1.00	.952	.943	.936	.928	.758	.92

〈표 16〉 변량 분석 결과 (N=48)

변 량 원	자 승 화	자유도	변량추정치	F	유 의 도
인 지 수 준	.036	1	.036	32.684	0.000
항 목 수	.227	5	.055	50.337	0.000
인지수준×항목수	0.102	5	0.020	18.594	0.000
오 차	0.040	36	0.001		
전 체	.405	47			

량분석 결과(〈표 16〉) 방향성 반영정도는 인지수준 및 항목 수에 따른 차이가 있는 것으로 나타났으나($p=0.000$) 이것은 13점 척도 집단 때문이었다. 13점 척도 및 3점 척도 집단을 제외시키고 5점~11점 척도 집단만을 대상으로 변량분석을 다시 실시한 결과 인지수준($F=2.499$, $df=1$, $p=0.127$) 및 항목 수($F=0.926$, $df=3$, $p=0.443$)에 따른 차이가 나타나지 않았다. 따라서 13점 척도 집단을 제외하고는 전체적으로 방향성 반영 정도는 인지수준 및 항목 수와 무관하다고 볼 수 있다. 즉 본 연구의 〈가설 III〉은 부정되었다.

본 연구의 〈가설 III〉은 응답자들이 각 項目에 관별적으로 응답할 것으로 예측되었기 때문에 척도는 방향성보다는 태도의 정도를 많이 포함할 것으로 생각되어서 설정된 것이며, 따라서 인지수준이 높을수록 판별력이 높으므로 태도 반영정도가 더 클 것으로 기대되어 설정한 것이었다. 그러나 가설검증 결과, 인지수준이 높은 집단과 낮은 집단은 방향성 반영정도에서 차이가 없었다. 이것은 앞의 〈가설 I〉과 〈가설 II〉의 검증결과에서와 마찬가지로 인지수준이 높다고 해서 項目간의 판별을 더 뚜렷하게 하지는 않는다는 것을 보여주고 있다. 방향성의 반영정도가 본 연구의 예측과 달리 높게 나타난 점은 다음의 종합논의에서 〈가설 I〉 〈가설 II〉의 결과와 함께 논의하도록 하겠다.

D. 연구결과에 대한 종합적 논의

본 연구의 결과, 위에서 보듯이 인지수준은 반응틀에 있어서 어떤 차이도 보이지 않았다. 즉 인지수준이 높은 집단이 각 항목간의 차이를 잘 판별하여 응답하지는 않는 것으로 나타났다. 그렇다면 인지수준이 높은 집단에서 신뢰도가 증가한 것은 무엇때문인가?

그것은 〈가설 I〉의 검증결과에서 지적했듯이 인지수준에서의 차이는 문항내의 항목간

의 차이를 판별하는데서 나타나는 것이 아니라 문항과 문항의 동질성 파악에서 나타나기 때문인 것으로 해석된다. 즉 인지수준이 높은 집단은 한 척도내의 문항이 거의 비슷하다는 것을 더 강하게 느낀다는 것이다. 이런 설명은, <가설 I-1>의 검증결과의 해석에도 적용된다고 하겠다. 즉 반응들이 나타나기 전인 3점에서 5점으로 항목 수가 증가한 경우, 왜 신뢰도가 증가하지 않았는가에 대해서도 “문항간 동질성 파악”의 개념이 적용될 수 있다. 즉 상호 약간 다르다고 생각되는 문항일지라도 3점 척도에서는 그 차이를 표현할 수가 없다. 따라서 5점 척도에서는 그런 억제되었던 차이 표현이 나타난 것이다. 그러나 7점 이상의 척도에서는 문항간의 차이를 반영하는 측면보다는 어떤 고정적인 반응틀—본 연구의 경우 극단치 선호—에 의거하여 응답하게 된다고 볼 수 있다.

따라서 7점 이상의 척도는 고정 반응틀인 극단치 선호 경향과 9점 이상의 척도의 경우에는 중간항 선호 경향까지 나타나 척도의 타당도가 저해된다고 볼 수 있다.

<가설 III>의 검증 결과 역시 극단치 선호 경향 개념으로 설명될 수 있다. 극단치가 제일 높은 점수를 부여받게 되어 있으므로, 척도치는 극단치에 의해 지배된다. 그러므로 극단치 선택이 많을 경우 조그만 점수 차로 태도의 차이를 표현하는 부분이 상대적으로 위축된다는 것이다. 따라서 극단치 선호가 강하게 나타난 본 연구의 경우 자연적으로 원척도의 축소 척도는 높은 상관도를 보인 것이다. 한편, 항목 수가 많은 척도에서 상관도가 적어진 것(13점 척도)은 극단치 선호가 아닌 다른 반응틀(예를 들어 13점 척도의 경우, 9점이나 4점 항목 선호도)이 나타났기 때문이 아닌가 생각되나 이 점은 본 연구의 결과만으로는 명확하게 설명될 수는 없다. 그러나 본 연구의 결과는 Likert형 척도가 본질적으로 방향성만을 측정하는 것이 아니라, 단지 극단치 선호가 강하게 나타났기 때문에 방향성이 높게 반영된다는 점을 보여주고 있다.

그러면 이런 극단치 선호 경향은 보편적인 것인가? 미국 대학생을 대상으로 한 Peabody나 Matell과 Jacoby의 연구가 본 연구의 결과와 비슷한 결과를 보였다는 점에서는 상당히 보편성이 있을 것으로 판단되나 좀 더 경험적 검증이 필요한 것으로 보인다.

본 연구의 결과로 볼 때, 2점 척도나 3점 척도가 적절하다는 주장은 설득력이 없다고 보겠다. 즉 3점 척도에서는 응답의 자유도가 적기 때문에 이질적인 문항에 동일한 대답을 강요하는 경우가 될 수도 있기 때문이다. 따라서 반응들이 나타나기 전인 5점 척도의 사용이 적절하다고 판단된다. 그 이상의 항목은 태도의 차이 보다는 응답 패턴에 지배되기 때문에 타당도가 저해된다고 보겠다.

IV. 結 論

그러나 연구문제에 대한 결론에 앞서 몇 가지 본 연구의 제한점을 지적해 보면, 첫째 대상자 선정의 문제이다. 인지수준의 차이에 대한 본 연구의 결과는 좀 더 다양한 대상자에 의한 검증이 요구된다고 본다.

둘째, 평가 척도만을 사용했다는 점이다. 중간항과 극단치는 언어적 표현이 제시됐지만 다른 항목들은 숫자로 표시되었다. 따라서 본 연구의 결과를 다른 유형의 반응척도에 일반화시키는 데에는 제한이 있다고 본다.

셋째, 응답 문항 數 즉 척도가 실제로 행하여지는 경우와 비교하면 너무 적다는 점이다. 따라서 실제 조사의 경우와 동일한 여건이었는지에 대해서는 의문이 있다.

위와 같은 몇가지 제약점을 감안하여, 본 연구의 가설 검증과 논의를 토대로 연구문제에 대한 답을 찾아 본다면 다음과 같다.

첫째, 유목 數의 증가는 신뢰도를 증가시키지 않는다고 하겠다. 그것은 개인간의 문항들에 대한 동질성 여부의 판단 차이가 반영되기 때문이며, 항목이 많아지게 되면 반응틀의 영향이 나타나기 때문이라고 보겠다.

둘째, 반응틀은 극단치와 중간항 선택을 선호하는 경향으로 나타났으며 7점 및 9점 척도에서 부터 나타났다. 따라서 7점 이상의 척도는 타당도가 저해된다고 하겠다.

셋째, 척도치의 상당부분이 방향성을 나타내고 있었으나, 이것은 극단치 선호도의 영향 때문이 아닌가 추측된다. 이 문제에 관해서는 보다 엄밀한 검증이 요구된다.

넷째, 인지수준은 신뢰도는 높이나 타당도와는 관계가 없는 것 같다. 즉 항목간의 판별 정도는 인지수준에 무관하다고 보겠다. 따라서 이상의 결론을 토대로 적정 항목 수를 찾아 본다면 그것은 5점 척도라고 하겠다.

그러나 극단치 선호도가 타당도 및 신뢰도에 미치는 영향을 생각해 볼 때 이런 극단치 선호 경향이 나타나는 원인에 대한 연구가 요청된다고 본다.