

지역사회영양학 연구에서의 리커트 척도 이용 현황: 대한지역사회영양학회지 게재 논문을 중심으로

류 시 현 · 윤 지 현^{1)†}

배재대학교 외식경영학과, ¹⁾서울대학교 식품영양학과 및 생활과학연구소

The Use of Likert Scale in Community Nutrition Research: Analysis of the Articles Published in Korean Journal of Community Nutrition

Si-Hyun Ryu, Jihyun Yoon^{1)†}

Department of Nutrition and FoodService Management, Paichai University, Daejeon, Korea

¹⁾Department of Food and Nutrition & Research Institute of Human Ecology, Seoul National University, Seoul, Korea

Abstract

The purpose of this study was to examine if Likert scales had been properly utilized in community nutrition research. A total of 527 research articles published in the 32 issues of Korean Journal of Community Nutrition from the volume 5, issue 1 in 2000 to the volume 10, issue 2 in 2005 were screened and 55 articles were found to have utilized one or more Likert scales for the studies. Therefore, 109 Likert scales used in the 55 studies were reviewed regarding the name, statement and response items, reliability and validity check, and analysis method. The scales were mostly referred as Likert scales (60%) or Likert-type scales (27%). Some scales were found to be referred as Likert scales although they were Likert-type scales when judged based on the information given in the respective articles. However some scales couldn't be judged for the rightness of the names because the information given for the scales in the articles was not enough. About 23% of the scales consisted of items less than 6 or more than 30, and therefore found to be inappropriate. The percentage of the scales listing all the statement items in the articles was only 25%. Most of the scales (85%) included 5 response items, and the rest included 4 (7%), 7 (6%), or 3 (2%). The percentages of the scales including appropriate center and end items were only 2% and 22%, respectively. Less than half of the scales (41%) were found to have been checked for reliability and only one scale was reported to have been checked for validity. In some scales (6%), the responses were scored improperly for analysis. The responses to the scales were frequently found to have been analyzed by parametric statistics such as mean, ANOVA, t-test, and Pearson's correlation, which might be a problem depending on the size and distribution of study samples. In conclusion, there is much room for improvement in the use of Likert scales in community nutrition research. (*Korean J Community Nutrition* 14(5) : 600~607, 2009)

KEY WORDS : likert scale · likert-type scale · community nutrition · research method

서 론

지역사회영양학은 급변하는 사회에 대응하는 종합학문이

며 응용학문으로 국내에서도 이 분야의 연구가 본격적으로 이루어진지 10년이 넘어섰다. 지역사회영양학은 개인과 집단의 건강유지 및 증진, 개선을 도모하기 위하여 공동체가 지닌 사회, 환경, 경제, 농업, 보건, 교육 문화적 특성을 학제적으로 접근하여야 하는 학문이기(Mo 등 2005) 이 분야의 연구에서 다루어지는 개념들은 복합적인 성격을 띠는 경우가 많다. 이에 지역사회영양학 분야의 연구에서는 이러한 개념들의 측정에 다문항 척도(Multi-item scale), 즉 여러 문항으로 이루어진 측도가 빈번히 사용되고 있다. 이는 다문항 척도가 단일 문항으로는 측정하기 어려운 복합적인 개념을 측정할 수 있으며, 측정의 신뢰성과 정확성을 높일 수 있는

접수일: 2009년 9월 9일 접수

채택일: 2009년 10월 16일 채택

*This study was supported by the 2006 research fund of Paichai University

†Corresponding author: Jihyun Yoon, Department of Food and Nutrition, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea

Tel: (02) 880-8750, Fax: (02) 884-0305

E-mail: hoonyoon@snu.ac.kr

장점을 가지고 있기 때문이다(Hong 2000).

태도를 측정하는 대표적인 다문항 척도로는 리커트 척도(Likert scale), 서스톤 척도(Thurston scale), 거트만 척도(Guttman scale) 등이 알려져 있다(Kim 2001). 이 중 지역사회영양학을 포함한 다양한 분야의 연구에서 가장 빈번히 사용되고 있는 것이 리커트 척도이다. Rensis Likert가 개발하여 1932년에 처음 발표한 리커트 척도는 측정하고자 하는 개념의 속성과 관련된 다문항 서술문에 대한 찬성과 반대의 정도를 나타내는 각 응답 항목에 문항들의 방향을 고려하여 1에서 5까지 점수를 부여하고 응답자가 선택한 개별 문항에 대한 점수를 합산한 총점으로 응답자의 태도와 상대적인 강도를 측정하는 방법이다(Sung 등 2005). 리커트 척도는 다른 척도들에 비하여 수량화가 비교적 용이하고 간편하여 태도, 의견, 속성 및 가치 등을 측정하기 위한 척도로 다양한 분야의 연구에서 널리 사용되어 왔다. 그러나 리커트 척도의 개발 및 사용과 관련하여 애매하고 불확실한 여러 문제들에 대한 논쟁 또한 계속되어 왔다(Jacoby & Matell 1971; Lee 1985; Albaum 1997; Jamieson 2004; Uebersax 2006).

지역사회영양학 연구의 양적 성장과 함께 이 분야에서의 리커트 척도의 사용 빈도 또한 상당히 높아져 왔다. 그러나 개발 및 이용 방법에 대한 충분한 이해 없이 리커트 척도가 사용되어 측정의 타당성, 나아가 연구 결과의 타당성 확보가 어려운 경우가 발견되고 있다. 이에 본 연구에서는 대한지역사회영양학회지에 게재된 연구 논문에서 사용된 리커트 척도의 현황을 그 적절한 활용법에 중점을 두어 분석하고 논의하였다.

조사대상 및 방법

1. 자료 수집

대한지역사회영양학회지 2000년 5권 1호부터 2005년 10권 2호에 게재된 총 527편의 논문 중 논문의 본문 또는 표에 리커트 또는 Likert란 용어가 명시되어 있는 55편의 논문을 수집한 후, 이 논문에서 이용된 총 109개의 리커트 척도를 분석 자료로 선정하였다.

2. 자료 분석

리커트 척도의 명칭, 문항 및 답항 구성의 적절성, 신뢰도와 타당도의 검증 여부, 응답의 분석 방법 등을 검토하였다.

1) 척도의 명칭 검토

논문에서 분석 대상 척도가 어떤 명칭으로 명명되었는지

를 검토하였다. 여러 연구에서 사용되는 리커트 척도는 Likert(1932)가 처음 개발하여 제시한 척도와는 차이가 존재한다. 그래서 여러 문헌(Nitko 1983; Albaum 1997; Jamieson 2004; Edmondson 2005; Uebersax 2006; Vagias 2006; Marsh-Richard 등 2009)에서 일반적으로 권고되거나 허용되고 있는 형태의 리커트 척도는 이러한 최초의 리커트 척도와는 차이의 정도에 따라 ‘리커트 척도(Likert scale)’ 또는 ‘리커트형/식 척도(Likert-type Scale)’로 불리기도 한다. 이에 분석 대상 척도들이 구체적으로 어떠한 명칭으로 사용되었는지를 검토하였다.

2) 문항 구성의 적절성 검토

척도의 문항 구성의 적절성을 평가하기 위해 첫째, 문항의 개수가 논문에 제시되었는지를 검토하였다. 둘째, 문항이 일반적인 다문항 척도의 구성 기준인 최소 6개 이상인지(Ok 2001), 신뢰도를 확보하고 응답자의 지루함을 방지하기 위해 이상적인 약 20~25개(Kwon 2008) 또는 20~30개(Park 등 2007) 정도의 문항으로 구성되었는지를(Kwon 2008) 검토하였다. 셋째, 문항들이 논문에 모두 제시되었는지를 검토하였다.

3) 답항 구성의 적절성 검토

답항 구성의 적절성을 검토하기 위해 첫째, 응답 항목의 수가 적절한 지를 검토하였다. 둘째, 응답 항목이 논문에 모두 제시되었는지를 검토하였고, 제시된 경우, 이러한 항목의 배열 순서를 검토하였다. 셋째, 응답 항목의 표현들의 적절성을 중간 응답 항목과 양 극단 항목을 중심으로 검토하였다. 응답 항목의 표현이 본문에는 한글로, 표에는 영문으로 기술되었으나 우리말 표현에 대한 영어 번역이 오류인 것으로 판단되는 경우에는 연구에 사용된 설문지가 우리말로 작성되었음을 감안하여 우리말로 기술된 표현을 택하여 코딩하였다.

4) 신뢰도 및 타당도의 검증 여부 검토

먼저 척도를 연구에 사용하기 위해 직접 개발하였거나 이전 연구에서 개발된 표준화된 척도를 그대로 또는 수정 보완하여 활용하였는지에 대해 검토하였다. 다음으로 연구 논문에 척도의 신뢰도 및 타당도에 대한 검증에 관한 언급이 있는지를 검토하였다.

5) 응답의 분석 방법 검토

리커트 척도로 측정된 자료에 대한 점수화 방법 및 통계분

척 방법이 자료의 특성에 적합한 지에 대해 검토하였다. 첫째, 응답 항목에 점수를 부여하는 방식에 대해 검토하였다. 둘째, 측정된 자료에 사용된 통계분석 방법을 모수 통계인지 비모수 통계인지를 중심으로 검토하였다.

결 과

1. 척도의 명칭

해당 논문에서 언급된 척도의 명칭을 검토한 결과, 60%가 리커트 척도로, 27%가 리커트식 또는 리커트형 척도로 명명되어 있었다. 그러나 실제로 리커트형 또는 리커트식 척도라 명명되는 것이 적절함에도 불구하고 리커트 척도로 잘못 명명된 경우도 있었고, 그와 반대의 오류도 발견되었다. 단, 해당 논문에 척도 명칭의 적절성을 판단하기 위한 정보, 즉 척도의 문항 및 응답 항목에 대한 정보 등이 충분히 제시되어 있지 않은 경우가 있어 전체 척도 중 어느 정도가 이러한 오류를 범하고 있는지를 분석할 수는 없었다. 일부 논문에서는 리커트형 항목(11%) 또는 리커트 방법(3%)을 사용한 것으로 명명되어 있었다(Table 1).

2. 문항 구성의 적절성

약 7%의 척도는 그 문항 수에 대한 정보가 해당 논문 중에 전혀 제시되어 있지 않았다. 약 28%의 척도의 경우, 문항 수에 대한 정보가 해당 논문의 본문에는 제시되어 있지 않았으나 표에 제시된 문항 항목을 통해 유추할 수 있었다. 다 문항 척도의 가장 적절한 문항 수의 범위에 드는 약 20~30개 정도의 문항으로 구성된 척도는 16%에 불과했으며, 6개 미만 또는 30개를 초과하는 문항으로 구성되어 부적절한 것으로 판단되는 척도가 23%에 달했다. 구체적인 표현이 논문에 제시된 척도는 약 25%에 불과하였으며, 나머지 척도의 경우 대부분 논문 본문에 문장의 몇몇 중요 항목만 서술되었거나 결과 표에 간략하게 요약되어 제시되어 있었고(60%), 문항의 표현에 대한 어떠한 정보도 제공되어 있지 않은 척도도 약 16%에 달하였다(Table 2).

3. 답항 구성의 적절성

응답 항목의 구성을 검토한 결과를 Table 3에 제시하였다. 응답 항목 수가 5개인 척도를 사용한 경우가 85%로 대부분이었다. 그러나 응답 항목 수를 4개(7%)로 줄이거나 7개(6%)로 세분화하여 변형한 경우도 있었으며, 단지 3개(2%)의 응답 항목으로만 구성된 척도도 있었다. 약 11%의 척도는 논문에 응답 항목의 표현을 전혀 제시하지 않는 오류를 범하였다. '매우 동의함'과 같은 긍정 항목이 먼저 제

Table 1. Name of the scales

Name	n (%)	
5 point Likert scale	51 (46.8)	
4 point Likert scale	6 (5.5)	
Likert scale	4 (3.7)	
7 point Likert scale	2 (1.8)	
3 point Likert scale	1 (0.9)	
5 step Likert scale	1 (0.9)	
	65 (59.6)	
5 point Likert type scale	26 (23.9)	
Likert-type scale	Likert type scale	3 (2.8)
	29 (26.6)	
5 step Likert type item	8 (7.3)	
Likert-type item	7 step Likert type item	4 (3.7)
	12 (11.0)	
5 step Likert method	2 (1.8)	
Likert method	4 step Likert method	1 (0.9)
	3 (2.7)	
Total	109 (100.0)	

Table 2. Statement items of the scales

Criteria Classification	n (%)	
Mentioned in method section	70 (64.2)	
If the number of items was mentioned	Not mentioned, but total items were shown in table	31 (28.4)
	Neither mentioned in method section nor shown in table	8 (7.3)
Number of items	1 - 5	18 (16.5)
	6 - 19	59 (54.1)
	20 - 25	15 (13.8)
	26 - 30	2 (1.8)
	31 - 40	5 (4.6)
	41 or more	2 (1.8)
	Not mentioned	8 (7.3)
All items fully listed	27 (24.8)	
If items were listed	All items listed in summary or selected items listed	65 (59.6)
	Not listed at all	17 (15.6)
Total	109 (100.0)	

시된 경우(52%)가 부정 항목부터 제시된 경우(37%) 보다 많았다.

Table 3. Response items of the scales

Criteria Classification		n (%)
Number of items	3	2 (1.8)
	4	8 (7.3)
	5	93 (85.3)
	7	6 (5.5)
If items were listed	All items listed	54 (49.5)
	Part of the items (end items and/or center items) listed	43 (39.4)
	Not listed at all	12 (11.0)
Order of items	Positive items first	57 (52.3)
	Negative items first	40 (36.7)
	Not known	12 (11.0)
Appropriateness of center items	Appropriate (e.g. neutral)	2 (1.8)
	Inappropriate (e.g. moderate, common, don't know)	48 (44.0)
	Unable to judge because center items were not listed	51 (46.8)
Appropriateness of end items	Center items don't exist due to even number of response items	8 (7.3)
	Appropriate (e.g. very/very, strongly/strongly)	24 (22.0)
	Inappropriate (e.g. most/not at all, always/not at all)	73 (67.0)
	Unable to judge because end items were not listed	12 (11.0)
Total		109 (100.0)

응답 항목 중 중간점에 해당하는 용어를 제시하지 않은 경우가 47%나 되었으며, 응답 항목의 수가 짝수이어서 중간 응답이 존재하지 않는 경우는 7%이었다. 단 2개의 척도를 제외한 모든 분석 가능 척도에서 중간점에 해당하는 용어로 ‘보통이다’, ‘그저 그렇다’, ‘잘 모르겠다’ 등의 부적절한 항목을 사용하고 있었다. 중간 항목을 적절히 포함하고 있는 것으로 평가되는 두 척도에서는 ‘neutral’, 즉 ‘찬성도 반대도 아니다’를 중간 항목으로 포함하고 있었다. 응답 항목의 양 극단에 사용된 수식어가 적절한 것으로 판단되는 척도는 약 22%에 불과하였으며, 나머지는 ‘가장’, ‘전혀’, ‘항상’ 등과 같은 극단적인 표현을 쓰거나 이러한 표현을 비대칭적으로 사용하고 있었다.

4. 신뢰도 및 타당도의 검증 여부

분석된 척도 중 선행연구에서 개발된 척도를 수정 보완하여 사용한 경우(41%)가 가장 많았고, 연구목적에 적합하게

Table 4. Reliability and validity check of the scales

Criteria Classification		n (%)
Types of scales	Pre-developed scale with modification	45 (41.3)
	New scale	29 (26.6)
	Pre-developed scale without modification	5 (4.6)
Reliability check	Not mentioned	27 (24.8)
	Mentioned with concrete data	43 (39.4)
	Mentioned without concrete data	2 (1.8)
Validity check	Not mentioned	64 (58.7)
	Mentioned without concrete data	1 (0.9)
	Not mentioned	108 (99.1)
Total		109 (100.0)

척도를 새로이 개발한 경우(27%)도 있었으며, 표준화된 척도를 수정없이 그대로 사용(5%)한 경우도 있었다. 그러나 이러한 척도의 개발 여부 등에 대한 정보가 전혀 제시되어 있지 않은 경우도 약 1/4에 달하였다. 신뢰도가 검증된 것으로 보이는 척도는 41%이었으며, 대부분의 경우, Cronbach α 계수를 통해 신뢰도 검증 결과를 제시하였으나, 이러한 구체적인 자료는 제시하지 않고 신뢰도를 검증했음을 보고한 경우도 있었다. 타당도를 검증한 것으로 보고된 척도는 단 한 개 밖에 없었다(Table 4).

5. 응답의 분석 방법

응답에 대한 점수부여 방법 및 이렇게 산출된 자료의 통계 분석 방법에 대한 검토 결과를 Table 5에 제시하였다. 약 77%의 척도에서는 응답항목 수에 관계없이 가장 비우호적인 응답에 가장 낮은 점수인 1점을, 가장 우호적인 응답에 가장 높은 점수를 부여하고 있었고, 1점이 아닌 0점부터 점수를 부여하는 변형된 방식을 취한 척도(12%)도 있었다. 그러나 중간응답에 0점을 부여한 후 이를 기준으로 하여 최우 응답항목에 대해 대칭적으로 음과 양의 점수를 부여한 척도(6%)도 일부 있었다.

본 연구에서 분석된 척도로 측정된 자료는 주로 모수 통계로 분석된 것으로 나타났다. 중심경향치는 주로 평균값으로 기술되었고(93%), 추론 통계치로 분산분석(50%), t-검정(48%), 피어슨 상관 분석(39%), 다중 회귀분석(19%) 등이 빈번히 사용된 것으로 분석되었다. 분석된 척도 중 소수의 경우만 서열 척도로 간주되어 빈도(10%), 카이제곱 검정(7%) 등이 이용된 것으로 나타났다.

Table 5. Analysis method used for the responses to the scales

Criteria Classification		n (%)	
Methods for assigning points to responses	1 point assigned to the negative end item and the highest point assigned to the positive end item	84 (77.1)	
	0 point assigned to the negative end item and the highest point assigned to the positive end item	13 (11.9)	
	0 point assigned to the center item and both - and + points assigned to the rest items	7 (6.4)	
	Not mentioned	5 (4.6)	
Statistics used for analysis ¹⁾	Mean	101 (92.7)	
	ANOVA	54 (49.5)	
	Parametric statistics	t-test, independent or paired	52 (47.7)
	Pearson's correlation coefficient	44 (38.8)	
	Multiple regression analysis	21 (19.3)	
	Factor analysis	5 (4.6)	
	Non-parametric statistics	Frequency	11 (10.1)
	Chi-square test	8 (7.3)	
	Wilcoxon signed rank test	7 (6.4)	
	Spearman's correlation coefficient	1 (0.9)	

1) the total number and percentage don't equal to 109 and 100%, respectively, due to multiple use of statistics in analysis of responses.

고 찰

지역사회영양학 연구에서 측정도구로서 가장 일반적으로 사용되는 것 중 하나가 다문항척도인 리커트 척도이다. 그러나 본 연구에서 지역사회영양학 연구에서 사용된 109개의 리커트 척도를 검토한 결과 척도를 활용함에 있어 다양한 오류들이 발견되었다.

척도의 명칭에 대해서 살펴본 결과, 리커트형 또는 리커트식 척도라 명명되었어야 함에도 불구하고 리커트 척도라 명명된 경우 또는 이와 반대의 오류를 범한 경우가 발견되었다. 어떠한 척도가 리커트 척도라 명명되기 위해서는 일반적으로 가져야 할 특성이 있다. 즉, 리커트 척도는 다문항 서술문으로 이루어져야 하며, 응답 항목들은 속성에 대한 빈도나 만족 정도를 나타내는 표현도 허용되기도 하나 원칙적으로 동의의 수준을 나타내는 표현으로 구성되어야 하며 일정한 간격 변화를 내포하는 용어들로 중간점을 기준으로 대칭을 이루어야 한다. 또한 응답 항목의 수는 5개를 원칙으로 하여 7, 9, 11개 등의 홀수로만 변경 가능하다(Uebersax 2006). 다른 모든 측면에서 전형적인 리커트 척도와 동일하게 구성되었다더라도 응답 항목의 수가 홀수가 아닌 짝수로 이루어졌거나, 답항이 중간점을 기준으로 대칭을 이루고 있지 않은 척도의 경우는 리커트 척도가 아닌 리커트형 또는 리커트식 척도라 명명되어야 한다(Uebersax 2006; Marsh-Richard 등 2009). 따라서 향후 지역사회영양학 연구에서는 그 특성에 따라 척도의 이름을 보다 정확하게 명명할 필요성이 있다.

척도를 구성하는 문항의 개수를 분석하였을 때, 리커트 척도는 적어도 6개의 문항으로 구성되어야 한다는 일반적인 기준(Ok 2001)을 벗어난 척도가 약 17%에 달하였다. 또한 소수이긴 하지만(6%) 적절한 문항 수의 상한선인 30개를 넘는 문항을 포함하고 있는 척도도 발견되었다. 문항 수가 증가할수록 척도의 신뢰도는 높아지지만(Nunnally 1978), 문항 수를 너무 늘리면 응답자로 하여금 지루한 느낌을 주게 되어 불성실한 응답을 하게 만들거나 뒤쪽으로 갈수록 무응답율이 높아질 수 있다. 리커트 척도의 문항구성 시 몇 개의 문항이 가장 적절한지에 대한 절대적인 기준은 없으나, 문항 분석을 통해 개별문항 점수와 총점간의 상관계수가 높아 변별력이 높은 약 20개(Nitko 1983) 또는 20~25개(Kwon 2008)나 20~30개 정도(Park 등 2007)의 서술문으로 구성하도록 권고되고 있다. 따라서 향후 지역사회영양학 연구에서 리커트 척도의 문항 구성 시 가장 적절한 문항 수에 대해 일괄적으로 규정할 수는 없겠으나, 6개 미만 또는 30개를 초과하는 문항 수는 지양하고 6~30개 범위 내에서도 측정하고자 하는 개념에 따라 가장 적절한 문항 수가 적용되어야 하겠다.

리커트가 처음 개발했을 당시의 리커트 척도는 5개의 답항(strongly approve—approve—undecided—disapprove—strongly disapprove)으로 구성되었지만(Likert 1932), 일반적인 사회과학 분야에서도 마찬가지로 본 연구에서 분석된 척도의 경우에서도 일부 응답 항목의 수에 변화가 있음이 발견되었다. 응답 항목의 수가 증가할수록 척도의 신뢰도는 증가하지만(Oster 1989; Preston & Colman 2000; Weng 2004), 5개를 초과할 경우 신뢰도는 완만하게 증가

한 후 평행을 유지하며 (Lissitz & Green 1975; Jenkins & Taber 1977; Aiken 1983), 응답자의 부담을 증가시킬 수 있다 (Cox 1980). 따라서 향후 지역사회영양학 연구에서 척도의 신뢰도 증가를 위해 응답 항목의 수를 늘리기 보다는 응답자의 성별, 지적수준, 문항에 대한 친숙도, 개인의 속성에 따른 판별력 등을 고려하여 (Cha 1990; Kronsnick 1999) 가장 적절한 응답 항목의 수를 결정하는 것이 필요하겠다.

이렇게 변형된 응답항목의 수가 짝수여서 중간 응답이 존재하지 않은 경우도 있었는데, 중간 응답을 제외하여도 평균 점수에 유의적인 영향을 주지 않으며 (Guy & Norvell 1977; Lee 1985), 응답자의 태도방향에도 변화가 없기에 (Ryan 1980; Ron 1991) 극단응답을 꺼리거나 ‘잘 모르겠다’ 또는 ‘의견 없다’는 태도를 가진 응답자가 중간응답을 선택할 가능성으로 인한 분산측정 훼손 방지를 위해 포함시키지 말아야 한다는 의견도 있다 (Dumas 1999). 반면 중간응답 부재 시 중립 의견을 갖는 응답자에게 강제적으로 부정 또는 긍정을 선택하도록 하여 분포오류가 발생할 가능성이 커지므로 중간응답을 명백하게 제시해야 하며 (Press & Schuman 1980), 응답자가 논리적으로 중간점을 선택할 수 있다면 응답항목 수가 홀수인 것이 더 바람직하다는 반박도 있다 (Cox 1980).

처음 개발되었던 리커트 척도에서는 중간 답항으로 ‘undecided’가 사용되었는데, ‘neither agree nor disagree’ (Albaum 1997)나 ‘uncertain or neutral’ (Loewenthal 1996)로 변형 가능하며, 우리말로써 중간점을 잘 반영할 수 있는 표현인 ‘찬성도 반대도 아니다’ 또는 ‘어느 쪽도 아니다’로 기술하는 것이 적절하다 (Hong 2000). 그러나 본 연구에서는 매우 소수의 척도만이 이러한 적절한 중간 답항을 포함하고 있는 것으로 분석되었다.

분석된 척도 중에는 중간 응답으로 ‘모르겠다’를 제시하는 오류를 범한 경우도 있었다. 조사대상자가 교육수준이 낮은 집단일 때 ‘모르겠다’ 항목을 중간 응답 이외에 추가로 포함 시킴으로써 거짓응답을 유도할 가능성 (Oh 1995)을 배제할 수 있는 장점이 있기는 하나 중간 응답으로 ‘모르겠다’를 제시하는 것은 바람직하지 못하다. 분석된 척도에서 중간 응답으로 가장 빈번히 사용되고 있었던 표현은 ‘보통이다’이었고, ‘그저 그렇다’가 그 다음이었는데 이러한 표현들은 그 의미의 모호성으로 리커트 척도의 답항으로 적절하지 않기에 향후 지역사회영양학 연구에서 활용되는 리커트 척도의 중간 응답으로 사용하는 것은 지양되어야 하겠다.

Albaum (1997)은 중간응답이 포함된 리커트 척도 사용 시 발생할 수 있는 응답의 중앙집중오류를 방지하기 위한 대

안으로 변형된 2단계 리커트 척도를 제안하였다. 이 변형된 척도는 첫 단계에서 응답항목으로 ‘agree/neither agree nor disagree/disagree/no opinion’을 제시하여 문항에 대한 ‘동의 또는 비동의 방향’ 또는 ‘의견 없음’을 물어본 후, 두 번째 단계에서는 ‘very strong/not very strong’의 응답항목을 제시하여 첫 단계에서 응답한 의견에 대한 강도를 물어보는 형식이다. 리커트 척도의 응답항목에 단순히 ‘no opinion’을 추가한 1단계 리커트 척도를 사용한 경우 보다 변형된 2단계 리커트 척도 사용 시 극단부분을 선택한 응답자의 비율이 높게 나타나 응답자의 진실한 의견 파악에 유용하게 사용될 수 있음을 보여주었다.

분석된 척도의 양극단에는 ‘항상’, ‘전혀’, ‘가장’과 같은 절대적인 의미를 지닌 수식어가 자주 사용되고 있었는데, 이런 경우 응답이 중간점에 집중될 수 있기에 (Wyatt & Meyers 1987) 이러한 수식어는 바람직하지 못하다. 또한 리커트 척도의 응답 구성 시 측정하고자 하는 속성에 대한 부정 또는 긍정의 강도를 분명하고, 간결하게 표현할 수 있는 용어를 사용하는 것이 중요하다. 그런데 본 연구에서 분석된 척도에서는 응답 항목 수가 동일한 척도에서조차도 각 응답 항목에 다양한 표현을 사용하고 있는 등 여러 가지 오류가 발견되어 응답항목 수에 따른 응답용어 정립의 필요성이 제기되었다. 단, 분석된 척도 중 응답 항목 전체를 제시한 경우는 약 50%밖에 되지 않았고, 표 하단에 영어로 제시한 경우에는 실제 사용된 설문지에 우리말로 제시된 점을 감안하여 분석하다보니 사용된 응답 용어를 파악하는데 한계가 있었다. 그러므로 응답 항목에 대한 상기의 분석 결과는 이러한 연구의 한계를 감안하여 해석되어야 할 것이다.

본 연구에서 분석된 척도들의 경우, 단순히 ‘신뢰도를 검증한 후 수정 보완하여 사용하였다’고만 설명하는 등 그 신뢰도와 타당도의 검증 또는 그에 대한 언급이 매우 미흡하였다. 특히 Cronbach α 를 제시하여 척도의 신뢰성을 언급한 논문은 다수 있었으나 타당도 검증을 언급한 논문은 단 한 편 밖에 없었다. 문항분석 결과 기술적으로 내적 일관성이 확보되었더라도 그것이 반드시 이론적 타당성을 보장해 주지는 못하며 타당도가 결여될 수도 있다 (Seong 1996). 따라서 향후 지역사회영양학 연구에서는 사용되는 리커트 척도의 신뢰도와 타당도를 둘 다 확보하기 위한 노력이 필요하며 이러한 노력과 검증 결과를 논문에 제시해 주어야만 하겠다.

한편, 선행연구를 통해 표준화된 척도나 외국에서 개발된 척도를 그대로 번역하여 사용한 경우도 있었는데, 조사 대상자나 측정하고자 하는 속성의 시간 경과, 그리고 문화적 차이가 척도의 타당도와 신뢰도에 영향을 줄 수 있기에 측정 도구의 양호도에 관한 수정과 보완이 필요하다고 본다. 또한 외

국 연구에서 타당도 및 신뢰도 검증절차를 거쳐 개발된 리커트 척도라 할지라도 국내 연구에 활용할 경우에는 번안된 응답 용어들이 원래의 척도가 함의하는 의미를 적절히 전달하는지에 대한 검토가 선행되어야 한다.

리커트는 척도 구성 시 제시된 문항에 대한 응답자의 찬반 태도의 정도에 따라 응답항목을 5개로 분류한 후, 각각에 가중치를 부여한 점수와 가중치 대신 단순히 1~5까지의 숫자를 부여해 합산한 점수간의 상관성이 매우 높은 것에 근거하여 5개의 응답항목에 단순히 숫자를 부여하는 방법을 적용하였다(Likert 1932). 응답에 대한 점수 부여 시 문항에 대해 가장 비우호적인 응답항목에 0점 또는 1점부터 시작하여 양의 방향으로 1점 단위로 균등하게 점수를 부여할 수 있는데(Sung 2005), 이 때 문항의 방향을 고려하면서 일관된 방식으로 점수를 부여(즉, 긍정적인 문항에 대해 매우 찬성하거나, 부정적인 문항에 대해 매우 반대하면 가장 높은 점수 부여)함에 유의해야 하겠다.

리커트 척도의 수준을 어떻게 보느냐에 따라 측정된 자료를 분석하는 기술통계 및 추정통계 방법이 달라지기 때문에 리커트 척도가 서열척도인지 등간척도인지에 대한 많은 논쟁이 계속되고 있다. 리커트 척도의 각 응답항목 간 간격이 반드시 동일하다는 것을 증명할 수 없기에 엄밀한 의미에서 서열척도로 보아야 한다는 견해가 있는 반면, 개별문항에 대한 응답점수를 합산한 총점을 등간척도로 간주할 수 있다는 견해도 있다(Albaum 1997). 일부에서는 가감은 가능하지만 승제는 못하는 등간척도의 난점 때문에 모든 수리분석이 가능한 비율척도로까지 리커트 척도를 간주하여 로지트(logit) 점수로 변환시켜 자료를 분석하기도 한다(Kim & Kim 2007).

리커트 척도를 서열척도로 단정할 경우, 측정된 자료의 중심경향치로는 중앙값이, 분산도로는 사분편차가 제시되어야 하고, 상관관계 분석에는 스피어만의 서열상관분석을, 추정 통계에는 만휘트니 U Test, Wilcoxon, Friedman's 2-way ANOVA 등과 같은 비모수적 방법을 사용하는 것이 타당하다(Becker 1999). 그러나 각종 사회과학 연구에서는 리커트 척도를 등간척도로 간주하여 연속형 자료에 적용할 수 있는 통계방법인 피어슨 상관분석, t 검정, ANOVA, 회귀분석, 요인분석까지 실시하고 있으며(Jamieson 2004), 리커트 척도를 활용한 마케팅 분야의 논문들을 분석한 연구(Edmondson 2005)에서도 분석된 논문의 50% 이상에서 중심경향치로 산술평균이 제시되고 있는 것으로 나타났다. 본 연구에서 분석한 다수의 지역사회영양학 연구에서도 이렇듯 리커트 척도를 등간척도로 간주하여 통계분석을 실시하고 있었다.

측정척도의 수준보다는 표본 수와 분포가 모수 통계기법을 적용할지를 결정하는데 있어 더 중요하므로(Knapp 1990) 리커트 척도를 지역사회영양학 연구에 활용하여 수집된 자료를 분석할 경우, 표본 수가 크고 모집단의 분포가 정규분포를 따를 때는 모수통계분석을 적용할 수 있겠다. 무엇보다도 연구에 사용할 측정도구를 작성하기 전에 연구문제에 적합한 통계방법을 고려하여 측정도구를 구성하는 것이 측정도구와 적용한 통계 방법 간에 조화를 이루지 못해 발생할 수 있는 논리적 오류를 방지할 수 있겠다.

요약 및 결론

지역사회영양학 연구에 이용된 109개의 리커트 척도를 검토한 결과를 요약하면 다음과 같다.

1) 척도의 명칭을 검토한 결과, 60%가 리커트 척도(Likert scale)로, 27%가 리커트식 또는 리커트형 척도(Likert-type scale)로 명명되어 있었다. 리커트식 또는 리커트형 척도로 명명되었어야 함에도 불구하고 리커트 척도라 명명되거나 이와 반대인 오류가 발견되었다.

2) 척도의 문항 구성을 검토한 결과, 다문항 척도에서 일반적으로 가장 적절한 20~30개의 문항으로 구성된 척도는 16%에 불과했으며, 6개 미만 또는 30개를 초과하는 문항으로 구성되어 부적절한 것으로 판단되는 척도가 23%에 달하였다. 또한 구성 문항이 논문에 모두 제시된 척도는 25%에 불과하였다.

3) 척도의 답항 구성을 검토한 결과, 척도의 응답항목 수가 5개(85%)인 경우가 대부분이었다. 약 11%의 척도의 경우, 응답 항목의 표현을 전혀 제시하지 않는 오류를 범하였다. 단 두 개를 제외한 모든 분석 대상 척도에서 중간점에 해당하는 용어로 '보통이다', '그리 그렇다', '잘 모르겠다' 등의 부적절한 항목을 사용하고 있었다. 양 극단 응답의 수식을 적절하게 사용한 척도는 약 22%에 불과하였으며, 나머지는 '가장', '전혀', '항상'과 같은 극단적인 표현을 쓰거나 이러한 표현을 비대칭적으로 사용하는 오류를 범하고 있었다.

4) 척도의 신뢰도 및 타당도의 검증 여부를 검토한 결과, 신뢰도가 검증된 것으로 보고된 척도는 41% 수준이었으며, 단 한 개의 척도만이 타당도의 검증을 보고하였다.

5) 척도에 대한 응답 분석 방법을 검토한 결과, 응답에 대한 점수부여 시 가장 비우호적인 응답에 0점부터(12%) 또는 1점부터(77%) 점수를 부여하여 가장 우호적인 응답에는 가장 높은 점수를 부여하고 있었다. 검토된 척도들로 측정된 자료는 대부분 모수 통계로 분석된 것으로 나타났다.

결론적으로, 그 동안 지역사회영양학의 연구에서의 리커트 척도의 이용에 있어 다수의 오류가 있었던 것으로 나타났다. 특히 척도의 문항 및 답항의 구성에 있어 적절치 못한 사례가 많았고, 척도의 신뢰도와 타당도를 확보하기 위한 노력 또한 충분치 않은 것으로 나타났다. 연구에 이용된 척도에 대한 충분한 정보가 논문에 제시되지 않은 것 또한 문제점으로 나타났다.

타학문 분야에서도 리커트 척도의 이용과 관련한 논란은 지속되고 있는 것이 사실이나 일반적으로 합의가 이루어진 부분도 적지 않다. 따라서 지역사회영양학분야의 연구에서도 리커트 척도를 이용할 때 이러한 일반적인 원칙에 근거하여 적절히 이용하기 위한 노력이 필요한 시점이라 생각된다.

참 고 문 헌

- Aiken LR (1983): Number of response categories and statistics on a teacher rating scale. *Educational and Psychological Measurement* 43(2): 397-401
- Albaum G (1997): The Likert scale revisited: an alterate version. *J Marketing Research Society* 5(1): 29-47
- Becker L (1999) : Scales of Measurement. Retrieved April 4, 2006, from <http://web.uccs.edu/lbecker/SPSS/scalemeas.htm>
- Cha BK (1990): 등간척도에서의 등간의 수에 관한 비교문화권적 조사연구. *한국언론학보* 25: 199-223
- Cox EP (1980): The optimal number of response alternatives for a scale: a review. *J Marketing Research* 17(4): 407-422
- Dumas J (1999): Usability testing methods: subjective measures, Part II - measuring attitudes and opinions, pp. 153-162, American Institute for Research. Usability Professional's Association
- Edmondson DR (2005): Likert scales: a history. *Conference Proceedings on Historical Analysis & Research in Marketing* 12: 127-133
- Guy RF, Norvell M (1977): The neutral point on a Likert scale. *J Psychology* 95(2): 199-204
- Hong DS (2000): 사회조사분석(3rd ed.), Dasan Books Co., Seoul
- Jacoby J, Martell MS (1971): Three points Likert scales are good enough. *J Marketing Research* 8(4): 495-500
- Jamieson S (2004): Likert scales: how to (ab)use them. *Medical Education* 38(12): 1217-1218
- Jenkins GD, Taber TD (1977): A monte carlo study of factors affecting three indices of composite scale reliability. *J Applied Psychology* 62(4): 392-398
- Kim ER (2001): Social Research. Korea University Press, Seoul
- Kim HJ, Kim JH (2007): Ordered logit model: 설문조사에 적용되는 척도의 종류. *윌간국토* 310: 94-102
- Kwon DH (2008): Educational Evaluation (2nd ed.), Hakjisa Co., Seoul
- Knapp TR (1990): Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research* 39(2): 121-123
- Kronsnick JA (1999): Survey research. *Annu Rev Psychol* 50(1): 537-567
- Lee EJ (1985): 커뮤니케이션 연구에 있어서 Likert 척도의 이용에 관한 소고. *커뮤니케이션과학* 9: 51-85
- Likert RA (1932): A Technique for the measurement of attitudes. *Archives of Psychology* 140
- Lisstiz RW, Green SB (1975): Effect of the number of scale points on reliability a monte carlo approach. *J Applied Psychology* 60(1): 10-13
- Loewenthal KM (1996): An Introduction to Psychological Tests and Scales, UCL Press Limited, London
- Marsh-Richard DM, Hatzis ES, Mathias CW, Venditti N, Dougherty DM (2009): Adaptive visual analog scales(AVAS): a modifiable software program for the creation, administration, and scoring of visual analog scales. *Behav Res Methods* 41(1): 99-106
- Mo SM, Koo JO, Park YJ, Park YS, Soon SM, Seo JS (2005): Community Nutrition, Kyomunsa Co., Seoul
- Nitko AJ (1983): Educational Tests and Measurement: An Introduction, Harcourt Brace Jovanovich Inc., New York
- Nunnally JC (1978): Psychometric Theory (2nd ed.), McGraw-Hill, New York
- Oh IH (1995): Social Survey Method: Non-sampling Errors, Nanam Publishing House., Seoul
- Ok CS (2001): 연구방법론과 논문작성법, Jigu Publishing Co., Seoul
- Oster TRF (1989): Number of alternatives per choice point and stability of Likert-type scales. *Perceptual and Motor Skills* 68(2): 549-550
- Park (2007): Educational Evaluation, Kyoyookbook Co., Seoul
- Presser S, Schuman H(1980): The measurement of a middle position in attitude survey. *Public Opinion Quarterly* 44(1):70-85
- Preston CC, Colman AM (2000): Optimal number of response categories in rating scales -Reliability, validity, discriminating power, and respondent preferences-. *Acta Psychologica* 104(1): 1-15
- Ron G (1991): The mid-point on s rating scale -Is it desirable?-. *Marketing Bulletin* 91(2): 66-70
- Ryan M (1980): The Likert scale's midpoint in communications research. *Journalism Quarterly* 57(2): 305-313
- Seong TJ (1996): Validity and Reliability, Yangseowon Co., Seoul
- Sung SJ, Yoo TK, Lee SW, Lee GY (2005): Research Methods for Social Work (4th ed.) (Originally written by Rubin A, Babbie ER in 2000). Sigmapress Co., Seoul
- Uebersax JS (2006): Likert Scales: Dispelling the Confusion. *Statistical Methods for Rater Agreement* website. Retrived January 10, 2008, from <http://ourworld.compuserve.com/homepages/jsuebersax/likert.htm>.
- Vagias WM (2006): Likert-type Scale Response Anchors. Clemson International Institute for Tourism. & Research Development, Department of Parks, Recreation and Tourism Management, Clemson University
- Weng LJ (2004): Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement* 64(6): 956-972
- Wyatt RC, Meyers LS (1987): Psychometric properties of four 5-point Likert type response scales. *Educational and Psychological Measurement* 47(1): 27-35