

리커트 척도에서 각 척도점(anchor)에 사용하기 적합한 표현들의 탐색 및 제안*

하 창 현^{1,2)} 박 상 희¹⁾

¹⁾충북대학교 ²⁾Toronto Metropolitan University

리커트 척도는 심리학을 포함한 사회과학 등의 분야에서 자주 사용되는 측정 도구이다. 리커트 척도는 각 점수들 간의 간격이 동등하다는 가정 하에 사용되고 있으나 척도점에 사용되는 구체적인 표현들(예: '약간', '매우')이 무엇인지에 따라 응답자들은 각 점수 간의 간격을 동등하게 지각하지 않을 수 있다. 이는 응답값에 체계적 오류를 발생시켜, 그 결과 분석과 추정에 편향을 일으킬 수 있다. 이러한 문제점을 개선하고자 본 연구에서는 '동의 또는 해당'과 '반대'를 측정하기 위해 사용하는 리커트 척도에 어떤 표현들이 주로 사용되는지를 조사하고 각 표현이 나타내는 의미를 수량화하여, 각 척도점에 사용하기 적합한 표현들을 선정하였다. 먼저 예비 조사를 통해 최근(2020년 1월 - 2022년 12월) 출판된 연구들에서 사용된 척도점의 표현들을 파악하였다. 그 결과, 최근 연구들에서 자주 사용하는 동의 표현이 32개, 반대 표현이 26개임을 확인하였다. 다음으로 본 조사에서는 성인 남녀 205명에게 예비 조사에서 살펴본 각 표현이 의미하는 정도를 슬라이드 척도에 평정하게 하였다. 참가자들이 각 표현이 나타내는 의미가 얼마인지 응답한 값을 기반으로 4점, 5점, 7점 척도의 각 척도점에 사용하기 적합한 표현들이 무엇인지를 살펴보았다. 본 연구에서 선정한 각 척도점 표현들을 연구자들이 맥락에 맞게 수정하여 사용할 수 있도록 제안하며, 각 척도점에 부적합한 표현을 사용하여 발생하는 응답의 체계적 오류를 줄이길 기대한다.

주요어 : 리커트 척도, 동의 또는 해당, 반대, 척도점

* 이 논문 또는 저서는 2023년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2023S1A5B5A17085242).

† 교신저자: 하창현, 충북대학교 심리학과, 충북 청주시 흥덕구 중대로 1,

E-mail: changhyunha@chungbuk.ac.kr



© 2024, Korean Social and Personality Psychological Association.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>)

which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

다양한 사회과학 분야의 연구자들은 응답자로부터 응답 및 반응을 얻기 위하여 여러 측정 도구를 사용한다. 그중 리커트 척도(Likert scale)는 심리학, 교육학 등의 분야에서 자주 사용하는 측정 도구이다. 리커트 척도는 응답자가 제시된 진술문에 얼마나 동의하는지 또는 자신이 얼마나 그에 해당하는지 등을 주어진 숫자 범위 내에서 선택하여 답하는 형태의 척도이며 각 점수 간의 간격이 동등하다는 전제를 가지고 있다(Likert, 1932). 리커트 척도는 높은 신뢰성과 함께 주관적인 생각이나 느낌 등을 수치로 표현할 수 있다는 장점을 가진다(Croasmun & Ostrom, 2011; Joshi et al., 2015; Symonds, 1924). 리커트 척도는 특성상 각 항목 간의 간격이 일정하지 않기에 엄격한 의미에서 등간 척도로 취급해서는 안되고 서열 척도로 간주해야 한다는 주장이 많다(Bishop & Herron, 2015; Jamieson, 2004; McCollin & Ramalhoto, 2007). 이 경우에는 비모수 통계분석 방법을 사용해야 하지만 편의적인 이유로 리커트 척도가 등간척도로 널리 사용되고 있는 것이 현실이다(Knapp, 1990; Stevens, 1946; Tsui, 1997). 주어진 척도를 등간 척도로 간주해도 되는지에 대해서는 척도의 범위를 늘려서 사용하거나(Wu & Leung, 2017) 다분 문항 반응(polytomous item response theory) 모형의 하나인 일반화 부분 모형(generalized partial credit model)을 이용하여 검정이 가능하지만(Dai et al., 2001), 본 연구에서는 이와 반대의 접근을 시도하였다. 즉, 여러 표현들 중 어떤 표현을 사용하였을 때 등간 척도로 간주할 수 있는지에 대하여 보다 직관적이고 간단한 방법으로 알아보았다.

연구자들은 다양한 형태의 리커트 척도를 사용하여 설문을 진행한다. 예를 들어, 각 점수가 나타내는 의미를 적지 않고 숫자로만 구성된 것, 양쪽 끝과 중간에만 부분적으로 척도점의 의미를 적은 것 또는 모든 척도점에 의미를 적은 것이 있다. 그렇다면 이들 중에 어떤 척도를 사용하는 것이 측정에서의 신뢰도를 증가시킬 수 있을까? 또한 리커트 척도가 가정하고 있는

각 점수들 간의 간격에 대한 동등성은 어떤 형태의 척도를 사용하느냐에 따라 변화할까?

응답자들에게 제시하는 선택지가 숫자로만 구성되어 있다면 점수들 사이의 간격은 일정하게 인식될 수 있다. 예를 들어 한 응답자가 어떤 진술문에 동의하는 정도를 5점 척도 상에 답해야 하는 경우, 숫자만 제시되어 있으면 각 척도점에 사용된 표현이 나타내는 의미로부터의 영향을 안 받고 응답할 수 있을 것이다. 그러나 선택지를 숫자로만 제시하는 것은 응답자들에게 혼동을 가져올 수 있다. 응답자들이 얼마나 동의하는지를 측정하는 척도에 답할 때, 예를 들어, 척도에서의 1점이 의미하는 바가 '전혀 동의하지 않음'인지 아니면 '매우 조금 동의함'인지 혼동이 생길 수 있다. 또한 숫자만 제시하는 경우에는 점수의 의미가 추상적일 수도 있으며(Maitland, 2009), 11점 척도의 범위를 -5점부터 +5점까지로 사용하는 것과 0점부터 10점까지로 사용하는 것은 응답자들이 그 의미를 다르게 해석하여 반응에 차이가 나타날 수 있다(Schwarz et al., 1991). 이처럼 숫자로만 척도를 제시하는 것보다 각 척도점에 표현을 제시하는 것은 응답자들에게 척도의 각 점수가 가지는 의미에 대해 더 많은 정보를 제공할 수 있다는 장점이 있다.

각 척도점에 모든 표현을 제시하지 않고 부분적으로 제시하는 경우에도 문제가 존재한다. Alwin과 Krosnick(1991)은 부분적으로 척도점에 표현을 제시한 것보다 각 척도점에 모두 표현을 제시한 경우가 응답에서 더 높은 신뢰도를 보인다고 밝혔다. 이러한 경향은 Alwin(2007)이 종단 연구의 신뢰도를 비교하였을 때에도 나타났으며, 모든 척도점에 표현을 제시한 경우의 신뢰도(Cronbach's $\alpha = .719$)가 부분적으로 표현을 제시한 경우의 신뢰도(Cronbach's $\alpha = .506$)보다 높았다. Saris와 Gallhofer(2007)가 1023개의 설문 문항들을 가지고 수행한 메타 분석에서도 모든 척도점에 표현을 제시하는 것이 신뢰도를 높인다는 결과가 나타났다. 또한 Moors 등(2014)의 연구에서는 척도의 모든 점수에 표현을 제시할 때

응답자들에게서 나타나는 극단응답방식(extreme response styles: ERS)이 낮아지는 것으로 나타났다. 즉, 모든 척도점에 표현을 제시하는 것이 부분적으로 표현을 제시하거나 숫자로만 척도를 제시하는 것보다 체계적 오류를 줄일 수 있는 것이다.

그러나 모든 척도점에 언어 표현이 제시된 척도의 경우에도 문제가 발생할 수 있다. 척도에서 각 점수 간의 간격이 각 척도점에 사용된 표현의 의미에 따라 응답자들에게 다르게 지각될 가능성이 존재한다. 한 응답자가 다섯 개의 언어 표현으로 구성된 척도에 답할 경우, 예를 들어 ‘매우 동의하지 않음’, ‘동의하지 않음’, ‘중간’, ‘동의함’, ‘매우 동의함’의 표현에서 ‘매우 동의함’이라는 표현과 ‘동의함’이라는 표현의 의미상의 간격은 5점과 4점 간의 간격에 해당하는 25%의 의미상의 간격과 차이가 있을 수 있다. 구체적으로, 5점 척도에서 ‘매우 동의함’라는 표현이 의미하는 정도는 100%이고 ‘중간’이 의미하는 정도는 50%이지만, ‘동의함’이 의미하는 정도는 정확히 ‘매우’와 ‘중간’의 중앙인 75%가 아닌 한 쪽에 더 가까운 값(예: ‘중간’이라는 표현에 더 가까운 정도인 60%)일 수 있다. 즉 척도점에 언어 표현을 사용하는 경우에 언어 표현이 나타내는 의미가 숫자의 의미와 불일치할 수 있는 것이다.

동일한 간격을 갖는 척도의 표현들을 사용하는 것은 정확한 측정에 있어서 중요한 요건이다. 동일하지 않은 간격을 가지는 표현들을 사용하는 것은 체계적인 오류를 증가시켜 올바른 추정과 분석에 영향을 미칠 수 있다(Spector, 1976). 척도에 제시되는 표현의 간격이 일정하지 않음으로 인해 각 점수들이 의미하는 바가 달라져 측정의 타당도가 낮아질 수 있으며, 척도로 인한 응답의 편향이 발생하여 문항들 간의 상관관계가 잘못 추정될 수 있다.

한편 응답자들이 문항을 읽은 뒤 척도에 사용된 표현을 인지하고 답을 하기까지는 10초 이내의 짧은 시간이 걸린다(Bassili & Fletcher, 1991).

응답자들은 짧은 시간 동안 척도에 사용된 표현들 사이의 간격을 얼마나 정확하게 파악할 수 있을까? 예를 들어 5점으로 구성된 척도에서 5점은 ‘전적으로 동의함’으로 표현되고 4점에는 ‘꽤 동의함’, ‘약간 동의함’, ‘다소 동의함’, ‘조금 동의함’의 표현 중 하나가 사용되었을 때에 ‘전적으로 동의함’과 ‘꽤 동의함’ 간의 간격과 ‘전적으로 동의함’과 ‘약간 동의함’ 또는 ‘다소 동의함’ 간의 간격을 모두 25%에 해당하는 간격으로 동일하게 지각할까? 짧은 시간 동안 응답자들은 표현된 단어에 의존하여 응답 가능성이 높다. 이러한 이유로 각 척도에 사용되는 표현들은 최대한 응답자의 선택에 체계적 오류를 발생하지 않는 것들을 사용해야 한다. 그럼에도 불구하고 대부분의 연구자들은 척도점에 사용하는 표현에 대한 별다른 고려없이 같은 척도를 사용한 선행 연구가 썼던 척도점의 표현을 그대로 사용하는 경향이 있다. 또한 선행 연구에서 개발된 척도의 척도점 표현에 대한 정보가 제시되지 않은 경우에는 연구자들이 각 척도점의 표현을 임의로 선택하여 제시하기도 한다.

이처럼 척도에 사용되는 표현들에 따라 발생할 수 있는 부정적인 영향을 고려하여 Casper 등(2020)의 연구에서는 실제 문제들을 확인하고 경험적 방법을 통하여 각 표현들이 나타내는 정도(100% 중에서 얼마만큼 의미하는지)를 살펴보았다. 이 연구에서 연구자들은 여러 학회에서 사용된 표현들을 조사한 뒤에, 가장 많이 사용되는 표현과 척도들을 분류하였다. 연구자들은 응답자에게 여러 표현을 제시하여 각 표현(strongly disagree, very much disagree, disagree, do not agree, moderately disagree 등)이 얼마의 정도를 나타내는지 0-100점에 평가하도록 요청하였다. 이렇게 수집된 자료를 바탕으로, Casper 등(2020)은 각 표현들이 나타내는 의미의 범위가 서로 겹치지 않고 각 간격이 일정하도록 하는 표현들을 사용할 것을 제안하였다. 한국에서도 각 점수 간의 간격이 동등하다고 가정하는 리커트 척도를 주로 사용하고 있지만, 리커트 척도의 각

점수에 사용하는 척도점 표현들이 일정한 간격을 의미하고 있는지에 대해 살펴본 한국의 연구는 본 저자들이 조사해 보았으나 찾지 못했다. 본 연구에서는 Casper 등(2020)의 연구를 모델로 삼아 한국인들을 대상으로 한국어 표현을 사용하여 각 척도에 사용하기 적절한 단어들을 경험적 연구방법을 통하여 살펴보았다. 본 연구는 리커트 척도나 등간 척도, 서열 척도 등에 대한 이론적인 측면보다는 척도점에 어떤 표현을 사용하는 것이 적절한지를 탐색하고 제안하는 실용적 측면에 초점을 둔다. 또한 한국에서 연구를 진행할 때 척도점에 다양한 표현을 사용하여 발생할 수 있는 측정의 오류들을 줄이는 것을 목표로 한다. 구체적으로 본 조사에서는 예비 조사를 통해 리커트 척도에서 사용되는 한국어 표현들이 무엇이 있는지를 탐색하였고 본 조사를 통해 척도의 각 점수에 사용되는 한국어 표현들 중에서 4점 척도, 5점 척도, 7점 척도로 사용하기에 가장 적합한 표현들을 판별하여 제안하였다.

예비 조사

예비 조사는 본 조사의 수행을 위해 필요한 기초 자료를 마련하는 단계로 1) '동의 또는 해당'과 '빈도'를 측정하는 척도들에서 자주 사용하는 척도점의 표현들이 무엇인지를 알아봄으로써, 2) 본 조사에서 참가자들에게 제시할 '동의 또는 해당' 표현과 '빈도' 표현의 목록을 만들기 위한 목적으로 수행하였다. '동의', '유사', '빈도', '판단', '선호', '만족' 등을 측정하는 여러 척도 유형들 중 심리학을 포함한 대다수의 사회과학 연구에서는 '동의 또는 해당' 표현과 '빈도' 표현을 널리 사용하여 설문을 진행하기에, 예비 조사에서는 이 두 유형에 초점을 두었다. 구체적으로, 예비 조사에서는 최근 출판된 한국심리학회지의 논문들에 사용된 '동의 또는 해당'과 '빈도'를 측정하는 척도들의 표현을 조사하

여 정리하였다. 한편 예비 조사에서 '동의' 표현과 '유사' 표현을 하나의 유형인 '동의 또는 해당'으로 구분한 이유는 한국어에서 '동의'에 사용되는 척도점의 표현이 '유사'에도 동일하게 적용될 수 있으며, '해당'의 경우에는 한국어 표현상 질문이나 문항에 자신이 얼마나 해당하는지에 '그렇다' 또는 '그렇지 않다'로 답하는 경우들이 있기 때문이었다. 예를 들어 '약간 동의한다'의 '동의' 표현은 '약간 그렇다' 또는 '약간 유사하다' 등의 '해당' 또는 '유사' 표현에도 적용이 가능하다. 또한 '선호(예: 조금 선호한다)'나 '만족(예: 매우 만족한다)'에 대해서도 '동의 또는 해당' 표현을 적용하여 사용할 수 있다.

방 법

조사 대상

예비 조사에서는 응시 및 반응 시간과 같은 측정 방법보다 리커트 척도를 주로 사용하여 참가자에게 질문하는 경우가 많을 것으로 예상되는 학회지들 중에서 3개를 선정하여 조사하였으며, 2020년 1월 1일부터 2022년 12월 1일까지 [한국심리학회지: 사회 및 성격], [한국심리학회지: 산업 및 조직], [한국심리학회지: 임상심리 연구와 실제]에 출판된 논문 180편 중 한국어 척도가 사용되지 않은 15편의 논문들(문헌 조사 논문 11편, 메타분석 논문 1편, 질적 연구 논문 3편, 다차원 척도법 논문 2편, 영어로 작성된 논문 2편)을 제외한 165편의 논문에 사용된 척도점의 표현을 수집하였다.

자료 범주화 및 코딩

예비 조사를 위해 수집된 모든 척도들을 심리학 박사 1명과 심리학 박사수료생 1명이 회의를 통해 다섯 가지 유형('동의 또는 해당', '빈도', '기타', '척도점 표현 없음', '척도점 표현

및 척도 점수 없음)으로 분류하였다. (1) '동의 또는 해당' 유형은 주어지는 진술문에 자신이 얼마나 동의하는지 또는 얼마나 해당하는지를 묻는 척도들이며, 참가자들은 주어지는 진술문들에 대해 동의하거나 자신이 해당하는 정도를 '전혀 동의하지 않음', '전적으로 동의함', '전혀 그렇지 않다', '매우 그렇다' 등에서 선택하여 답한다. (2) '빈도' 유형 척도들은 어떤 행동이나 사건 등을 얼마나 자주 경험하는지 또는 어떠한 행위를 얼마나 자주 하는지를 묻는 표현들로 구성되며, 참가자들은 주어지는 문항에 대해 '전혀 하지 않음', '항상 함', '거의 없음', '자주 있음' 등으로 답한다. (3) '기타' 유형은 사용된 척도가 양극형 척도, 슬라이드 척도 또는 '<', '<<', '<<<'와 같은 기호를 사용한 경우 등에 해당한다. (4) '척도점 표현 없음' 유형은 연구에서 사용한 척도가 몇 점으로 구성되는지를 명시하였으나 척도의 표현을 제시하지 않은 경우이다. (5) '척도점 표현 및 척도 점수 없음' 유형은 사용한 척도의 척도점 표현뿐 아니라 사용한 척도가 몇 점으로 구성되었는지도 제시하지 않은 경우이다. 예를 들어, 척도를 사용한 저자들이 논문에 선행 연구의 척도를 사용하였다고 언급만 한 것이다.

추가로 각 유형들이 한 편의 논문에서 여러 차례 발생할 수 있기에 한 편의 논문에서 발생하는 각 유형의 횟수를 기록하였다. 해당 연구에서 사용한 모든 척도에 대해서 동일한 척도점 표현을 사용하였다고 언급한 경우는 해당 연구

에서 사용한 척도의 수만큼 개별 횟수로 기록하였다.

결 과

다섯 가지 유형에 대한 빈도 분석

165편의 논문에서 사용된 척도 총 722개 중 (표 1) '동의 또는 해당' 유형은 371개로 전체 사용된 척도 중에서 51.4%의 높은 비율을 차지하였다. '빈도' 유형은 54개로 전체에서 7.5%를 차지하였다. '기타' 유형은 37개로 5.1%를 차지하는 것으로 나타났다. '척도점 표현 없음' 유형은 169개로 23.4%를, '척도점 표현 및 척도 점수 없음' 유형은 91개로 12.6%를 차지하는 것으로 나타나 척도점 표현을 제시하지 않은 경우가 전체의 36%의 높은 비율을 차지하는 것으로 나타났다.

'동의 또는 해당'과 '빈도' 유형에 대한 빈도 분석

다음으로 예비 조사의 주 관심 대상인 두 유형('동의 또는 해당', '빈도')에서 사용된 척도점의 표현이 모두 보고되었는지 아니면 부분적으로 보고되었는지를 파악하였다. 모두 보고하는 경우는 예를 들어 5점 척도를 사용하였을 때 1점부터 5점에 사용한 모든 표현에 대해서 논문에 언급한 것이고, 부분적으로 보고한 경우는 5

표 1. 다섯 가지 유형의 빈도 분석 결과

유형	빈도	%	누적 %
동의 또는 해당	371	51.4	51.4
빈도	54	7.5	58.9
기타	37	5.1	64.0
척도점 표현 없음	169	23.4	87.4
척도점 표현 및 척도 점수 없음	91	12.6	100.0
전체	722	100.0	-

점 척도의 예에서 극단점인 1점과 5점(예: 1 = 전혀 동의하지 않음, 5 = 매우 동의함) 등 일부의 표현만을 보고하는 경우이다.

‘동의 또는 해당’ 유형에 속하는 371개 중 사용한 척도의 모든 표현을 논문에 보고한 경우는 14개(3.8%)에 불과하였으며, 대부분(357개, 96.2%)은 척도의 최소 점수와 최대 점수에 사용된 표현만을 보고하였다. 마찬가지로 총 54개의 ‘빈도’ 유형 중 10개(18.5%)만이 논문에 사용한 척도의 모든 척도점 표현을 보고하였으며, 나머지 44개(81.5%)는 척도의 최소 점수와 최대 점수의 척도점의 표현만을 보고하였다. 이러한 결과는 대부분의 저자들이 척도의 각 점수에 사용된 표현들을 논문에 완전하게 보고하지 않음을 의미한다.

추가로, ‘동의 또는 해당’ 유형에 속하는 371개의 척도에서 사용된 표현들이 무엇인지를 파악하였다. 구체적으로 371개의 척도에서 논문에 보고된 척도점의 표현은 중복 표현들을 모두 제외하고 총 30개였다. 개별 척도점에 사용된 표현들 중 가장 많은 빈도로 사용된 표현은 척도의 점수 중 최솟값에 ‘전혀 동의하지 않음/전혀 아니다’(84.6%)와 최댓값에 ‘매우 동의함/매우 그렇다’(87.8%)로 나타났다. 마찬가지로 ‘빈도’ 유형에 해당하는 54개의 척도에서 사용된 표현을 확인하였으며, 논문에 보고된 척도점 표현은 중복 표현을 제외하고 총 24개였다. 척도점에 숫자를 제시한 경우(예: ‘1 = 1~2회’)와 척도점 표현과 함께 명확한 숫자를 제시한 경우(예: ‘가끔 1~2일’)를 목록에서 제외한 결과, 빈도 표현들 중에서 가장 자주 등장하는 표현은 최솟값에 ‘전혀 없다’(62.9%), 최댓값에 ‘항상 그렇다’ (29.6%)였다.

종합하면, 예비 조사에서는 ‘동의 또는 해당’ 유형과 ‘빈도’ 유형이 가장 많이 사용된다는 것을 확인하였다. 또한 조사한 척도들에서 ‘동의 또는 해당’과 ‘빈도’에 무슨 표현들이 사용되었는지를 파악하여 본 조사에서 참가자들에게 제시하기 위한 표현들의 목록을 구성하였다.

본 조사

본 조사에서는 예비 조사에서 수집한 척도점의 표현 목록들을 사용하여 4점, 5점, 7점 척도들의 각 척도점 표현으로 사용하기 가장 적합한 것이 무엇인지를 조사하였다. 조사를 이 세 가지 척도에 한정된 이유는 예비 조사에서 살펴본 722개의 척도들에서 5점 척도가 315번(43.6%)으로 가장 많이 사용되었고 7점 척도가 168번(23.3%)으로 두 번째로, 4점 척도가 52번(7.2%)으로 세 번째로 많이 사용되어 대부분의 경우 이 세 척도 중 하나를 사용한 것으로 나타났기 때문이다.

본 조사에서 참가자들은 ‘동의 또는 해당’과 ‘빈도’ 척도의 척도점에 사용되는 표현들을 제시받고 각 표현들이 얼마 정도의 의미를 가지는지에 대해 연속선(101점 슬라이드 척도) 상에서 응답하였다. 참가자들의 답변을 기반으로 ‘동의 또는 해당’과 ‘빈도’ 유형의 척도점에 사용되는 표현들 중에서 등간격 기준값(5점 척도의 예: 1 = 0, 2 = 25, 3 = 50, 4 = 75, 5 = 100)과 의미상 가장 가까운 표현이 무엇인지를 알아보고, 4점, 5점, 7점 척도의 척도점으로 사용하기 적합한 표현을 선정하였다.

방 법

본 조사에서는 리서치 회사를 통해 205명(여성 104명)의 참가자를 모집하였으며, 참가자들의 평균 나이는 44.59세(SD = 13.74)였다. 참가자들의 연령과 성별에 따라 척도점의 표현에 대한 개념적인 거리에 차이가 나타날 가능성을 고려하여 20세부터 69세의 참가자를 모집하였으며 참가자의 성비 또한 균등하게 모집하였다. 참가자들의 인구통계학적 특성은 표 2에 제시하였다.

표 2. 참가자들의 인구통계적 정보 빈도

	연령대					합계
	20-29세 (n = 40)	30-39세 (n = 40)	40-49세 (n = 41)	50-59세 (n = 41)	60-69세 (n = 43)	
성별						
남성	20	20	20	20	21	101
여성	20	20	21	21	22	104
최종 학력						
초등(국민)학교 졸업	-	-	-	-	1	1
중학교 졸업	-	-	1	1	2	4
고등학교 졸업	5	5	5	8	10	33
전문대학(2-3년제) 재학 또는 졸업	9	7	6	8	4	34
대학교(4년제) 재학 또는 졸업	22	25	23	17	16	103
대학원 재학 또는 졸업	4	3	6	7	10	30

연구 절차

본 조사는 모두 온라인으로 진행되었으며, 연구의 본래 목적을 노출하지 않기 위하여 ‘여러 가지 한국어 표현에 대한 개인의 의견을 살펴보는 조사’라는 가짜 목적이 적힌 연구 대상자 설명서와 연구 참가 동의서를 참가자들에게 제시하였다. 조사 참가에 동의한 참가자들은 성별과 나이, 학력을 묻는 질문에 답한 뒤 연습 시행을 수행하였다. 참가자들은 연습 시행에서 동의 표현(“약간 동의”)과 빈도 표현(“아주 가끔”) 한 문

항씩을 제시받고 이에 응답하였다.

본 시행에서는 참가자들은 먼저 동의 표현에 해당하는 목록 30개를 이미지(그림 1)로 제시받고, 제시된 표현들의 의미를 생각하면서 각 표현을 주의깊게 읽도록 요청받았다. 다음으로 참가자들은 여러 표현이 의미하는 바가 서로 같거나 비슷할 수도 있다는 안내를 받은 뒤에, 제시된 30개의 표현이 어느 정도의 의미를 가지는지(“각 표현은 어느 정도 동의한다는 것을 의미할까요?”)를 0점(전혀 동의하지 않음)부터 100점(전적으로 동의함) 사이의 점수로 슬라이드 척

정말 동의함	비교적 동의함	조금 동의함	꽤 동의함	대체로 동의하지 않음
결코 동의하지 않음	대체로 동의함	다소 동의함	동의하지 않는 편임	꽤 동의하지 않음
매우 동의하지 않음	동의하는 편임	동의함	별로 동의하지 않음	조금 동의하지 않음
매우 강하게 동의함	약간 동의하지 않음	매우 동의하는 편임	매우 동의함	상당히 동의하지 않음
정말 동의하지 않음	매우 많이 동의함	아주 동의함	아주 많이 동의함	다소 동의하지 않음
어느 정도 동의함	약간 동의함	상당히 동의함	동의하지 않음	거의 동의하지 않음

그림 1. 동의 표현 이미지

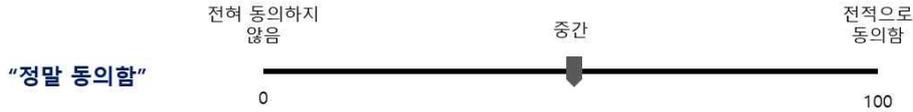


그림 2. 슬라이드 척도 예시

꽤 함	매우 많이 함	그저 그렇게 함	매우 함
거의 안 함	때때로 함	언제나 함	별로 안 함
보통	종종 함	약간 안 함	극히 드물게 함
가끔 함	간혹 함	매우 자주 함	거의 드물게 함
대부분 함	거의 대부분 함	대부분 안 함	많이 함
어쩌다 함	비교적 자주 함	약간 함	자주 함

그림 3. 빈도 표현 이미지

도에 표시하여 답하였다(그림 2). 또한 참가자들은 이전에 응답한 표현에 대해서 조정이 필요하다고 판단한 경우 이전의 응답값을 수정할 수 있도록 허용되었다.

동의 표현에 대한 응답이 끝나고 약 30초 정도 휴식을 가진 뒤, 참가자들은 빈도 표현에 해당하는 목록 24개를 이미지(그림 3)로 제시받고 각 표현을 읽었다. 동의 표현에서와 마찬가지로 참가자들은 각 표현이 나타내는 의미(‘각 표현은 얼마나 자주(빈도) 어떤 행위를 한다는 것을 의미할까요?’)를 0점(‘전혀 하지 않음’)부터 100점(‘항상 함’) 사이의 점수로 슬라이드 척도에 표시하여 답하였다. 참가자들이 각 표현에 대한 응답을 마치면 사후 설명서가 제시되고 연구는 종료되었다. 본 조사의 절차와 측정 도구는 연구자들이 속한 기관의 생명윤리심의위원회(IRB)로부터 승인(승인번호: CBNU-202211-HR-0253)을 받았다.

측정 도구(슬라이드 척도)

슬라이드 척도에서 0부터 100까지의 점수가

나타내는 의미가 무엇인지에 대한 기준을 제공하기 위하여 예비 조사에서 조사한 동의 표현 32개 중에 2개의 표현은 참가자들의 답변을 구하지 않고 슬라이드 척도의 양쪽 끝 위에 제시하였다. 구체적으로, 예비 조사에서 척도의 최솟값의 표현으로 가장 많이 사용되는 것으로 나타난 ‘전혀 동의하지 않음’을 슬라이드 척도의 0점 위에 제시하였고, ‘전적으로 동의함’을 슬라이드 척도의 100점 위에 제시하였으며, 50점 위에는 ‘중간’이라는 표현을 제시하였다. 예비 조사에서는 척도의 최댓값으로 ‘매우 동의함’이 가장 많이 사용되었으나 ‘매우’라는 표현은 국립국어원의 표준국어대사전(2008)에 따르면 “보통 정도보다 훨씬 더”를 의미하기에 최댓값인 100을 나타내기에 부적절하다고 판단하였다. 대신 “하나도 남김없이 모두 다인 것”을 뜻하는 ‘전적(全的)’이라는 표현을 사용하여 ‘전적으로 동의함’을 100점 위에 제시하였다.

빈도 표현에서도 이와 마찬가지로 예비 조사에서 수집된 26개의 빈도 표현 중에서 최솟값에 가장 많이 사용되는 ‘전혀 없다’를 맥락에 맞게 수정하여 ‘전혀 하지 않음’으로 0점 위에, ‘중간’

을 50점 위에, 최댓값에 가장 많이 사용되는 '항상 그렇다'를 '항상 함'으로 수정하여 100점 위에 제시하였다. 추가로 응답자들은 슬라이더의 처음 위치에 의해 영향을 받을 가능성이 있으므로 답변에서 한쪽으로 편향이 발생하는 것을 방지하기 위해 슬라이더의 위치를 50점에 위치시켜 응답자들에게 제시하였다.

결 과

기술통계치

본 조사에서 사용한 '동의 또는 해당' 표현 30개와 '빈도' 표현 24개에 대한 참가자들의 평정 점수의 평균과 표준편차는 부록 1에 제시하였다. 또한 '동의 또는 해당' 표현과 '빈도' 표현을 평균의 크기순으로 정렬하여 부록 2와 3에 그래프로 각각 제시하였다.

선정 기준 및 결과

본 조사에서는 척도점에 사용되기 적절한 표현을 선정하기 위하여, 참가자들이 평정한 표현에 대한 평균과 표준편차를 사용하여 두 가지 기준을 설정하였다. 첫째, 참가자들이 평정한 점수의 분포가 기준값(5점 척도의 예: 1 = 0, 2 = 25, 3 = 50, 4 = 75, 5 = 100)과 가까운지 여부를 판단하였다. 이를 위해 Casper 등(2020)이 사용한 절차를 따라 각 표현들에 대한 Tilton값을 계산하고 각 표현과 기준값의 공유되는 정도가 70% 이하(Tilton값 = 0.77 이상)인 표현들은 후보에서 제외하였다. Tilton값은 Tilton(1937)이 각 점수들이 공유하는 부분(overlap)을 계산하기 위해 제안한 방법이며, 구체적인 계산 방법은 수식 (1)과 같다.

$$\text{수식 (1)} \quad \frac{(M_1 - M_2)}{[(\sigma_1 + \sigma_2)/2]}$$

M1은 계산하려는 두 척도점 표현 중에서 높은 평균값, M2는 낮은 평균값이다. σ_1 은 M1에 대한 개별 점수들의 표준편차, σ_2 은 M2에 대한 개별 점수들의 표준편차이다. 수식(1)을 이용하여 계산된 값을 백분율로 환산한 값은 두 척도점 표현이 겹치는 정도(%)를 의미한다. 본 조사에서는 척도의 등간격 표현을 찾는 것이 목적이므로 4점, 5점과 7점 척도에 대한 기준값과 참가자들이 척도점 표현에 응답한 값의 평균과 표준편차를 사용하였다. 기준값에 대한 표준편차는 존재하지 않으므로, 임의적으로 본 조사에서 측정된 두 유형의 표현에 대한 표준편차 평균값(동의 또는 해당 = 7.88; 빈도 = 9.26)을 사용하였다. 예를 들어, 본 조사에서 참가자들이 응답한 '다소 동의함($M = 70.98$, $SD = 8.43$)', '대체로 동의함($M = 76.91$, $SD = 7.98$)'을 5점 척도에서 4점에 해당하는 기준값($M = 75.00$, $SD = 7.88$)과 수식(1)을 이용하여 Tilton값을 계산하면 '다소 동의함'은 0.49으로, '대체로 동의함'은 0.24으로 나타났으며, 이를 백분율로 환산하면 각각 80%와 90%이다(Tilton, 1937, 표 2 참조). 이는 '대체로 동의함'이 기준값과 90%를 공유하고 있으므로 5점 척도에서 4점의 척도점 표현으로 사용하기에 '다소 동의함'보다 더 적절함을 뜻한다. 이러한 방식으로 각 표현들과 기준값간의 Tilton값을 계산하였다.

둘째, 참가자들이 각 표현들에 응답한 값의 표준편차가 10을 넘는 경우는 그 표현을 후보에서 제외하였다. 즉, 기준값을 50점이라고 하였을 때 어떤 표현의 평균이 50점일지라도 표준편차가 15라면 후보에서 제외된 것이다. 그 이유는 표준편차가 클수록 한 표현의 의미가 참가자마다 다르게 해석되어 그 표현이 가지는 의미의 신뢰성이 낮아지기 때문이다. 이상의 두 가지 기준을 적용하여 4점, 5점, 7점 척도에 사용하기 적절한 표현들을 계산하고 선정하였으며, 동의 표현은 표 3에, 빈도 표현은 표 4에 제시하였다.

본 연구 결과를 토대로 각 척도점에 동의 표

표 3. 기준에 부합하는 동의 표현 목록

	4점 척도	5점 척도	7점 척도
	전혀 동의하지 않음(-)	전혀 동의하지 않음(-)	전혀 동의하지 않음(-)
1점	결코 동의하지 않음(0.39; 84%)	결코 동의하지 않음(0.39; 84%)	결코 동의하지 않음(0.39; 84%)
	정말 동의하지 않음(0.46; 81%)	정말 동의하지 않음(0.46; 81%)	정말 동의하지 않음(0.46; 81%)
	매우 동의하지 않음(0.74; 71%)	매우 동의하지 않음(0.74; 71%)	매우 동의하지 않음(0.74; 71%)
2점	조금 동의하지 않음(0.29; 88%)	별로 동의하지 않음(0.12; 95%)	꽤 동의하지 않음(0.32; 87%)
	약간 동의하지 않음(0.41; 83%)	다소 동의하지 않음(0.21; 91%)	대체로 동의하지 않음(0.41; 83%)
		동의하지 않는 편임(0.30; 88%)	동의하지 않는 편임(0.66; 74%)
		대체로 동의하지 않음(0.52; 79%)	거의 동의하지 않음(0.70; 72%)
3점	비교적 동의함(0.003; 99%)		조금 동의하지 않음(0.28; 88%)
	어느 정도 동의함(0.34; 86%)	중간(-)	
	조금 동의함(0.67; 74%)		약간 동의하지 않음(0.41; 83%)
	약간 동의함(0.73; 71%)		
4점	전적으로 동의함(-)	대체로 동의함(0.24; 90%)	
	매우 강하게 동의함(0.52; 79%)	다소 동의함(0.49; 80%)	중간(-)
	매우 동의함(0.70; 72%)	어느 정도 동의함(0.72; 71%)	
5점		전적으로 동의함(-)	비교적 동의함(0.004; 99%)
		매우 강하게 동의함(0.52; 79%)	어느 정도 동의함(0.33; 86%)
		매우 동의함(0.70; 72%)	다소 동의함(0.53; 79%)
			조금 동의함(0.68; 73%)
			약간 동의(0.74; 71%)
6점			꽤 동의함(0.22; 91%)
			상당히 동의함(0.66; 74%)
			아주 동의함(0.70; 72%)
7점			전적으로 동의함(-)
			매우 강하게 동의함(0.52; 79%)
			매우 동의함(0.70; 72%)

주. 괄호 안에 제시된 값은 기준값과 응답값 간의 공유되는 정도를 Tilton값으로 계산하고 그 값과 백분율로 환산한 값을 제시하였음.

현을 사용할 경우, 5점 척도를 예로 어떤 표현을 사용하는 것이 적절한지는 다음과 같다. Tilton값을 백분율로 환산한 값을 기준으로, 5점 척도에서 1점(기준값 0점)에 '전혀 동의하지 않음', '결코 동의하지 않음', '정말 동의하지 않음', '매우 동의하지 않음' 순으로, 2점(기준값 25점)

표 4. 기준에 부합하는 빈도 표현 목록

	4점 척도	5점 척도	7점 척도
1점	전혀 하지 않음(-)	전혀 하지 않음(-)	전혀 하지 않음(-)
2점	약간 안 함(0.41; 83%)	별로 안 함(0.06; 97%) 거의 드물게 함(0.39; 84%)	거의 안 함(0.63; 74%)
3점	종종 함(0.10; 95%) 때때로 함(0.59; 76%)	중간(-) 그저 그렇게 함(0.13; 94%) 보통(0.23; 90%)	약간 안 함(0.41; 83%)
4점	항상 함(-)	비교적 자주 함(0.11; 95%) 자주 함(0.63; 75%) 꽤 함(0.74; 71%)	중간(-) 그저 그렇게 함(0.13; 94%) 보통(0.23; 90%)
5점	-	항상 함(-)	종종 함(0.11; 95%) 때때로 함(0.59; 76%) 많이 함(0.02; 99%) 대부분 함(0.12; 95%) 꽤 함(0.17; 93%)
6점	-	-	자주 함(0.36; 85%) 매우 자주 함(0.49; 80%) 거의 대부분 함(0.60; 76%) 매우 함(0.65; 74%)
7점	-	-	항상 함(-)

주. 괄호 안에 제시된 값은 기준값과 응답값 간의 공유되는 정도를 Tilton값으로 계산하고 그 값과 백분율로 환산한 값을 제시하였음.

에는 ‘별로 동의하지 않음’, ‘다소 동의하지 않음’, ‘동의하지 않는 편임’, ‘대체로 동의하지 않음’ 순으로 동의 표현을 사용하는 것이 적절한 것으로 해석된다. 3점(기준값 50점)에는 ‘중간’을 사용하는 것이 적절하고 4점(기준값 75점)에는 ‘대체로 동의함’, ‘다소 동의함’, ‘어느 정도 동의함’ 순으로, 5점(기준값 100점)에는 ‘전적으로 동의함’, ‘매우 강하게 동의함’, ‘매우 동의함’ 순으로 동의 표현을 사용하는 것이 적절하다.

이와 마찬가지로 7점 척도를 사용한 경우 각 척도점에 빈도 표현을 사용한다면, 1점(기준값 0

점)에는 ‘전혀 하지 않음’, 2점(기준값 16.67점)에는 ‘거의 안 함’, 3점(기준값 33.33점)에는 ‘약간 안 함’을 사용하는 것이 적절하다. 4점(기준값 50점)에는 ‘중간’, ‘그저 그렇게 함’, ‘보통’ 순으로, 5점(기준값 66.67점)에는 ‘종종 함’, ‘때때로 함’ 순으로, 6점(기준값 83.33점)에는 ‘많이 함’, ‘대부분 함’, ‘꽤 함’, ‘자주 함’, ‘매우 자주 함’, ‘거의 대부분 함’, ‘매우 함’ 순으로 사용하는 것이 적절하며, 7점(기준값 100점)에는 ‘항상 함’을 사용하는 것이 적절하다. 이처럼 연구자들은 표 3과 표 4를 참고하여 각 점수에 선정된 여러 표

현 중에서 연구의 맥락이나 질문의 상황에 맞게 적절한 표현을 사용할 수 있다.

논 의

본 연구는 최근 심리학회지에서 사용된 ‘동의 또는 해당’ 표현과 ‘빈도’ 표현에 대해서 조사하고 참가자들이 생각하는 각 표현들의 의미를 수치로 평정하도록 하여, 4점과 5점, 7점 척도에 사용하기 적절한 표현들의 목록을 구성하여 제시하였다. 리커트 척도에서는 각 점수들 간의 간격을 동등하다고 가정하며 연구자들은 척도의 선행 연구에서 사용된 척도점 표현들 또는 자신들이 임의로 정한 표현들을 사용한다. 많은 연구자들은 리커트 척도의 각 점수들이 등간격으로 구성된다는 가정을 따르기에 각 표현들의 개념적인 거리가 실제로 등간격인지에 대해 별다른 의문을 갖지 않고 사용한다. 그러나 선행 연구(Casper et al., 2020)에 따르면 한 변인을 측정할 때 개념적으로 동등하지 않은 간격의 표현들을 사용하는 경우에는 개념적으로 동등한 거리의 표현들을 사용한 경우보다 신뢰도와 타당도가 더 낮은 것으로 나타났다. 이렇게 개념적으로 간격이 동등하지 않은 표현을 사용하는 것은 참가자들의 응답과 결과의 정확성을 떨어뜨릴 수 있다. 그러나 한국어 설문지에 사용되는 표현들이 개념적으로 동등한 거리를 가지는지를 실증적으로 확인한 연구는 본 저자들이 아는 한 거의 없다. 이 문제를 해결하기 위해 본 연구의 예비 조사에서는 설문에 가장 빈번하게 사용되는 ‘동의 또는 해당’ 유형과 ‘빈도’ 유형의 척도에 사용되는 표현들을 조사하고, 본 조사에서는 각 점수에 사용하기 적절한 표현의 목록을 구성하여 제공하였다.

본 연구에서 확인된 결과의 기대효과 및 의미는 다음과 같다. 첫째, 본 연구에서 선정한 표현들을 사용할 경우, 리커트 척도의 등간격 가정이 위반될 가능성이 줄어 연구 결과에 대한

신뢰도와 타당도가 증가할 것이다. 많은 연구자들은 리커트 척도의 각 점수들이 등간격으로 구성된다는 가정을 따르기에 각 표현들의 개념적인 거리가 실제로 등간격인지에 대해 별다른 의문을 가지지 않고 사용한다. 그러나 선행 연구(Casper et al., 2020)에서 살펴본 것처럼 척도에 사용된 표현에 따라 참가자의 응답에 편향이 발생할 수 있다. 본 연구에서 선정한 표현들을 사용하는 것은 연구자들이 다양한 설문에서 리커트 척도를 사용할 때 발생하는 응답의 편향을 감소시키도록 도울 것이다. 또한, 척도 개발 및 타당화에서 영문으로 구성된 척도점 표현들을 한국어로 번역할 때 연구자들이 척도점 표현을 임의로 선택하거나 단순 번역을 하여 발생할 수 있는 문제들을 본 연구에서 제안하는 표현들 중 가장 적합한 표현을 사용하여 줄일 수 있을 것이다.

둘째, 본 연구에서는 연구자들이 각 점수에 사용하기 적합한 표현들을 제안하였다. 동의를 측정하기 위한 5점 척도에서 2점의 척도점 표현으로 주로 사용되는 ‘조금 동의하지 않음’과 ‘약간 동의하지 않음’, 4점의 척도점 표현으로 자주 사용되는 ‘조금 동의함’, ‘약간 동의함’과 같은 표현이 본 연구에서 설정한 두 가지 기준에 부합하지 않았다. 반면 4점 척도에서는 이러한 표현들(즉 2점에 ‘조금 동의하지 않음’, ‘약간 동의하지 않음’과 3점에 ‘조금 동의함’, ‘약간 동의함’)을 사용하는 것이 적합한 것으로 나타났다. 이는 같은 표현일지라도 사용하는 척도가 몇 점 척도로 구성되었는지에 따라 그 적절성이 달라질 수 있음을 시사한다. 한 가지 가능한 설명으로는 4점 척도와 달리 5점 척도는 중간점을 포함하고 있기에 척도점 표현에 차이가 나타났을 가능성이 존재한다. 그러나 중간점을 포함하고 있는 7점 척도와 중간점을 포함하지 않는 4점 척도의 표현들을 비교해 보았을 때, 4점 척도의 2점(기준점 33.3점에 해당)과 7점 척도의 3점(기준점 33.3점에 해당) 그리고 4점 척도의 3점(기준점 66.6점에 해당)과 7점 척도의 5점(기준점

66.6점에 해당)에 사용하기 적합한 표현들이 비슷한 것으로 나타났다. 그러므로 중간점의 유무에 따른 차이보다는 다른 이유들에 의해서 4점과 5점의 적합한 표현 간의 차이가 나타났을 수 있다. 몇 점 척도인지에 따라 발생하는 표현의 차이가 무엇으로 인해 발생하는지 추후 연구에서 살펴볼 필요가 있다. 또한 지금까지 사용해왔던 각 점수의 표현들이 의미하는 바가 실제 점수와 차이가 있을 가능성이 높으며 결과에 편향을 가져왔을 가능성이 있다(Casper et al., 2020). 이러한 이유로 본 연구에서 여러 기준을 통해 확인한 표현들의 사용을 제안하며, 이는 참가자들이 각 척도에 응답할 때 발생하는 편향을 줄이는 데에 도움이 될 것이다.

셋째, 예비 조사의 결과에서 점수만 보고하고 척도점 표현을 보고하지 않은 비율(23.4%)과 척도점 표현과 점수 모두 보고하지 않은 비율(12.6%)의 합이 36%로 전체의 1/3을 차지하는 것으로 나타났다. 몇 점 척도를 사용했는지 또는 척도점에 어떤 표현을 썼는지에 대해 보고하지 않는 경향은 논문에 제시된 결과를 해석하는 데에 방해가 될 수 있으며, 참가자의 응답값의 범위가 몇 점부터 몇 점까지인지를 파악하지 못하게 되어 표나 그래프에 제시되는 통계치에 대한 해석에 오해가 생길 가능성이 있음을 시사한다. 또한 ‘동의 또는 해당’ 유형과 ‘빈도’ 유형에서도 사용된 척도의 모든 척도점 표현을 제시하지 않고 부분적으로 보고하는 경우가 많은 것으로 나타났다. ‘동의 또는 해당’ 유형의 척도에서 사용된 모든 척도점 표현을 보고한 경우는 3.8%로 매우 낮은 비율을 차지하였고 ‘빈도’ 유형에서도 18.5%로 낮은 비율을 차지하였다. 그러나 실제 연구에서 양쪽 끝과 중간을 제외한 척도점에 표현을 사용하지 않은 것인지, 모든 척도점에 표현을 사용하였지만 보고만 하지 않은 것인지 알기가 어렵다. 이처럼 척도의 각 점수에 대한 척도점 표현들을 보고하지 않는 경우 해당 연구를 후속 연구에서 반복검증하는 데에 방해가 될 수 있다. 후속 연구에서 척도점의 표현을 선행

연구와 다르게 사용하는 것은 참가자들의 답변이 한쪽으로 편향되게 할 가능성이 있기 때문이다. 따라서 후속 연구를 위해 척도에서 사용한 척도점 표현들을 모두 보고하는 것을 제안하며, 척도점의 표현을 모두 제시하지 않고 부분적으로 보고하였을 때, 보고되지 않은 척도점의 표현에 본 연구에서 제안하는 척도점들을 사용하여 척도 표현으로 인해 발생할 수 있는 체계적 오류를 감소시킬 수 있길 기대한다.

끝으로, 본 연구에서는 척도에 사용되는 척도점 표현들이 무엇인지를 파악하고 목록을 구성하는 것에 그치지 않고 다양한 참가자들을 모집하여 각 유형의 표현들이 나타내는 의미를 수량화하여 실증적으로 살펴보았음에 의의가 있다. 본 조사에서는 20대부터 60대까지의 다양한 연령층에 속하는 참가자들을 모집하였으며, 남성과 여성 참가자의 비율도 1:1로 모집하였다. 다양한 참가자들을 모집하여 얻은 결과는 본 연구에서 제안하는 표현들이 여러 대상(성인, 중년, 노년 등)에 대한 ‘동의 또는 해당’과 ‘빈도’를 측정하는 연구에서 활용될 수 있음을 뜻한다.

본 연구의 한계는 다음과 같다. 첫째, ‘동의 또는 해당’ 유형과 ‘빈도’ 유형에서 사용된 표현들이 4점, 5점, 7점 척도의 각 점수(5점 척도의 경우: 1점, 2점, 3점, 4점, 5점)에서 얼마나 사용되었는지를 계산하려고 하였으나 본 연구에서 조사한 많은 논문들은 척도에 사용된 최솟값과 최댓값의 척도점 표현만을 보고하고 있었다. 이러한 이유로 각 점수에서 주로 사용된 척도점 표현들의 빈도를 분석하지 못하였다는 한계점이 존재한다. 추후 연구에서는 척도의 모든 표현을 보고하지 않았을 경우, 논문의 저자에게 요청하여 각 점수들에 사용한 표현들이 무엇이었는지를 조사하여 정리할 필요가 있다.

둘째, 본 조사에서는 척도점의 표현(예: 매우, 약간, 자주, 가끔 등)에 대해서는 살펴보았으나 척도점에 사용될 수 있는 모든 표현들(예: 매우 바쁘다, 자주 먹는다 등)에 따라 척도점 표현이 의미하는 바가 변화할 수 있음을 고려하지 못하

였다. 다시 말해, '먹는다'라는 표현을 사용할 때, '자주 먹는다'와 '가끔 먹는다' 사이의 간격과 '느낀다'라는 표현을 사용할 때 '자주 느낀다'와 '가끔 느낀다' 사이의 간격이 같은지에 대한 여부를 확인하지 않은 것이다. 그러나 '먹는다'나 '느낀다' 등의 척도에 사용할 수 있는 모든 표현을 확인하는 것은 현실적으로 불가능하기에 현재로서는 척도점 구성의 일반적인 원칙이 제시될 때까지는 본 연구에서 제안하는 방법을 사용하는 것이 현실적으로 타당할 것이라 생각한다.

셋째, 본 조사에서 여러 표현들이 의미하는 바를 수량화하기 위하여 0점부터 100점까지의 슬라이드 척도를 사용하여 살펴보았으나 0점과 50점, 100점 위에 제시했던 각 표현이 의미하는 바는 알아보지 않았다. 구체적으로 동의/빈도 표현의 슬라이드 척도에서 0점 위에는 '전혀 동의하지 않음/전혀 하지 않음'을, 100점 위에는 '전적으로 동의함/항상 함'을 제시하였기에 각 표현들이 의미하는 바를 측정하지는 않았다. 이와 마찬가지로 슬라이드 척도의 50점 위에는 '중간'이라는 표현을 적어놓았으나 '중간'이라는 표현이 의미하는 바가 50에 해당하는지를 살펴보지 않았다. 본 조사에서 '전혀'와 '전적으로', '항상', '중간'이라는 표현들에 대한 사전적 정의를 고려하여 0점과 100점, 50점으로 간주하였으나 참가자들이 실제로 각 표현들을 0점이나 100점, 50점으로 지각하는지를 살펴보지는 않았다. 이러한 이유로 참가자들이 생각하는 각 표현의 평균과 표준편차를 계산할 수 없어서 Tilton값 계산에 기준값(예: 0, 25, 50 등)과 각 표현들의 전체 표준편차를 사용하였다. 추후 연구에서는 최솟값과 최댓값에 사용되는 표현들이 의미하는 바를 사전에 확인하거나 각 범위의 끝에 표현을 제시하지 않고 다른 방법으로 측정할 필요가 있다.

넷째, 본 연구에서는 척도점 표현들 간의 등간격 표현을 탐색하기 위하여 101점 슬라이드 척도를 사용하였다. 4점, 5점, 7점 척도와 마찬가지로

101점 슬라이드 척도는 측정학에서 서열 척도로 분류된다. 그러나 Wu와 Leung(2017)에 따르면 척도 점수가 11점인 경우 등간 척도로 가정이 가능하며, 본 연구에서는 101점 척도를 사용함으로써 보다 정밀하게 각 표현의 의미를 확인하였다. 또한 슬라이드 척도는 움직임이나 시야 등으로 인해 응답에 편향을 가질 수 있으나 (Kvam et al., 2023) 본 연구에서는 슬라이드 척도의 타당성을 고려하지 않았다. 추후 연구에서는 다양하고 정밀한 도구를 통해 각 척도점에 적절한 표현이 무엇인지 재확인할 필요가 있다.

끝으로 본 연구에서 동의와 빈도 유형의 4점, 5점, 7점 척도에 사용하기 적합한 표현들을 제안하였으나 9점 척도에 대한 제안은 하지 않았다. 그러한 이유는 본 연구에서 정한 두 가지 기준에 부합하는 표현들이 9점 척도를 구성하기에는 적절하지 않았기 때문이다. 예를 들어 동의 표현에서 평균 점수가 40점 이상 60점 이하에 해당하는 점수들이 없었다. 이처럼 9점 척도를 구성하기 위해서는 본 연구에서 조사했던 동의 또는 빈도 표현들보다 더 많은 표현들에 대한 측정치가 필요할 것으로 예상된다.

저자소개

하창현은 충북대학교 심리학과에서 박사 과정을 수료하였으며, 현재 Toronto Metropolitan University에서 방문 연구원으로 연수 중이다. 사회경제적 지위에 따른 차별, 인상 특징과 언어 유창성에 따른 차별, 타인의 고통 추정 등에 관심을 가지고 있다.

박상희는 충북대학교 심리학과 교수로 재직 중이다. 마음 지각, 집단 간 관계, 도덕적 판단, 빈부 문제, 사회심리학의 정책에의 적용 등에 관심을 가지고 있다.

참고문헌

- 국립국어원 (2008). 매우. 표준국어대사전. 2022년 12월 23일 검색,
<https://stdict.korean.go.kr/search/searchResult.do?pageSize=10&searchKeyword=%EB%A7%A4%EC%9A%B0>
- 국립국어원 (2008). 전적. 표준국어대사전. 2022년 12월 23일 검색,
<https://stdict.korean.go.kr/search/searchResult.do?pageSize=10&searchKeyword=%EC%A0%84%EC%A0%81>
- Alwin, D. F. (2007). *Margins of error: A study of reliability in survey measurement*. New York, NY: John Wiley & Sons.
<https://doi.org/10.1002/9780470146316>
- Alwin, D. F., & Krosnick, J. A. (1991). The reliability of survey attitude measurement: The influence of question and respondent attributes. *Sociological Methods & Research*, 20(1), 139-181.
<https://doi.org/10.1177/0049124191020001005>
- Bassili, J. N., & Fletcher, J. F. (1991). Response-time measurement in survey research a method for CATI and a new look at nonattitudes. *Public Opinion Quarterly*, 55(3), 331-346.
<https://doi.org/10.1086/269265>
- Bishop, P. A., & Herron, R. L. (2015). Use and misuse of the Likert item responses and other ordinal measures. *International Journal of Exercise Science*, 8(3), 297-302.
- Casper, W. C., Edwards, B. D., Wallace, J. C., Landis, R. S., & Fife, D. A. (2020). Selecting response anchors with equal intervals for summated rating scales. *Journal of Applied Psychology*, 105(4), 390-409.
<https://doi.org/10.1037/apl0000444>
- Croasmun, J. T., & Ostrom, L. (2011). Using likert-type scales in the social sciences. *Journal of Adult Education*, 40(1), 19-22.
- Dai, S., Vo, T. T., Kehinde, O. J., He, H., Xue, Y., Demir, C., & Wang, X. (2021). Performance of Polytomous IRT Models With Rating Scale Data: An Investigation Over Sample Size, Instrument Length, and Missing Data. *Frontiers in Education*, 6, 1-18.
<https://doi.org/10.3389/educ.2021.721963>
- Göb, R., McCollin, C., & Ramalhoto, M. F. (2007). Ordinal methodology in the analysis of Likert scales. *Quality & Quantity*, 41, 601-626.
<https://doi.org/10.1007/s11135-007-9089-z>
- Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38, 1212-1218.
<https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, 7(4), 396-403.
<https://doi.org/10.9734/BJAST/2015/14975>
- Knapp, T. R. (1990). Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research*, 39, 121-123.
<https://doi.org/10.1097/00006199-199003000-00019>
- Kvam, P. D., Marley, A. A. J., & Heathcote, A. (2023). A unified theory of discrete and continuous responding. *Psychological Review*, 130(2), 368-400.
<https://doi.org/10.1037/rev0000378>
- Likert, R. (1932). *A technique for the measurement of attitudes*. New York, NY: Columbia University Press.
- Maitland, A. (2009). Should I label all scale points or just the end points for attitudinal questions?. *Survey Practice*, 2(4), 1-5.
- Moors, G., Kieruj, N. D., & Vermunt, J. K. (2014). The effect of labeling and numbering of response scales on the likelihood of response

- bias. *Sociological Methodology*, 44(1), 369-399.
<https://doi.org/10.1177/0081175013516114>
- Saris, W. E., & Gallhofer, I. N. (2007). *Design, evaluation, and analysis of questionnaires for survey research*. Hoboken, NJ: Wiley-Interscience.
<https://doi.org/10.1002/9780470165195>
- Schwarz, N., Knäuper, B., Hippler, H. J., Noelle-Neumann, E., & Clark, L. (1991). Rating scales numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55(4), 570-582.
- Spector, P. E. (1976). Choosing response categories for summated rating scales. *Journal of Applied Psychology*, 61(3), 374-375.
<https://doi.org/10.1037/0021-9010.61.3.374>
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103, 677-680.
<https://doi.org/10.1126/science.103.2684.677>
- Symonds, P. M. (1924). On the loss of Reliability in Ratings Due to Coarseness of the Scale. *Journal of Experimental Psychology*, 7(6), 456-461.
<https://doi.org/10.1037/h0074469>
- Tilton, J. W. (1937). The measurement of overlapping. *Journal of Educational Psychology*, 28(9), 656-662. <https://doi.org/10.1037/h0053750>
- Tsui, M. S. (1997). Empirical Research on Social Work Supervision: The State of the Art (1970-1995). *Journal of Social Service Research*, 23(2), 39-54.
https://doi.org/10.1300/J079v23n02_03
- Wu, H., & Leung, S. O. (2017). Can Likert scales be treated as interval scales?—A simulation study. *Journal of Social Service Research*, 43(4), 527-532.
<https://doi.org/10.1080/01488376.2017.1329775>

1 차원고접수 : 2024. 06. 19.

수정원고접수 : 2024. 07. 12.

최종게재결정 : 2024. 07. 14.

Exploring and proposing conceptually equal-distanced anchor wording for each scale point of Likert scales to measure agreement, accordance, and frequency

Chang Hyun Ha^{1),2)}

Sang Hee Park¹⁾

¹⁾Chungbuk National University

²⁾Toronto Metropolitan University

The Likert scale is a measurement tool commonly used in various fields including social sciences such as psychology. Although the Likert scale is used with the assumption that the intervals between each scale point are equal, respondents may not perceive the intervals between each scale point equally, depending on the wordings used in scales (e.g., 'slightly', 'very'). This can lead to systematic errors in responses and biases in analysis and estimation of results. In this study, we empirically explored and proposed anchor wordings suitable for use in each point, considering the errors induced by the wordings. Specifically, we investigated what anchor wordings have been commonly used in the Likert scale to measure 'consent or accordance' and 'frequency'. We calculated and proposed what anchor wordings are most suitable for each scale point. In the preliminary survey, we collected and identified wordings of scale points used in recently published studies (January 2020 - December 2022). In the main survey, we recruited participants and had them rate the meaning of each anchor wordings obtained from the preliminary survey using slide scales. The preliminary survey found that 32 anchor wordings of consent and 26 of frequency were frequently used in recent studies. In the main survey, we proposed what anchor wordings are suitable for each scale point on the 4-, 5-, and 7-point scales based on the values that participants responded to what each wordings represent.

Key words : Likert scale, Agreement, Frequency, Anchor