

연구논문

웹기반 조사에서 척도의 크기가 응답의 타당도, 신뢰도, 응답 편이성에 미치는 영향*

Effects of Scale Size on Validity, Reliability and Easiness of Response
in a Web-based Survey

장윤재** · 조성검*** · 송중현**** · 김옥태***** · 박민규*****

Yoon Jae Jang · Sung Kyum Cho · Chong-Hyun Song · Ock-Tae Kim · Mingue Park

최근 유선전화 기반 조사방법에 웹이나 모바일을 결합한 혼합모드조사에 대한 관심이 증가하고 있다. 이와 관련해 본 논문에서는 척도 구성의 측면에 주목했다. 전통적인 조사연구에서 문항 작성 시 주로 다문항의 다점 척도를 사용한 경우가 많았던 데 반해, 이러한 문항 구성 방식이 웹기반 조사에서도 장점을 갖는지 분명하지 않기 때문이다. 따라서 웹기반 조사를 중심으로 척도의 크기에 따라 응답의 타당도와 신뢰도, 그리고 응답 편이성에 차이가 생기는지 검토하고자 했다.

주제어 : 웹기반 조사, 혼합모드조사, 타당도, 신뢰도, 응답 편이성

Regarding the growing concern for the mixture of landline and web-based/mobile survey, researchers are increasingly interested in methodological examination of anticipated benefits and problems of mixed mode surveys. This study focuses on scale construction. Multiple questions with a multi-point scale have been preferred in traditional surveys. However, it is not obvious if those scales have advantages also

* 이 연구는 2010년도 정부재원(교육과학기술부 한국사회기반연구사업비)으로 한국연구재단의 지원을 받아 이루어졌습니다(NRF-2010-330-B00280).

** 교신저자(corresponding author): 서울여자대학교 언론영상학부 전임강사 장윤재.
E-mail: yjang@swu.ac.kr

*** 충남대학교 언론정보학과 교수

**** 선문대학교 언론광고학부 조교수

***** 서울대학교 언론정보연구소 객원연구원

***** 고려대학교 통계학과 부교수

in a web-based survey. We tested if, in a web-based survey, validity, reliability, and easiness of response differ according to the size of scales.

key words : easiness of response, mixed mode survey, reliability, validity, web-based survey

I. 문제 제기

선거예측조사가 번번이 빗나가는 등 여론조사의 정확성에 문제가 제기되면서 전통적 조사방법의 한계에 대한 논의가 활발하게 이루어지고 있다. 이 과정에서 재택률 차이에 따른 표본편포 문제, 낮은 응답률로 인한 표본의 대표성 문제, 개인 투표성향을 조사하는 데 방문 면접이나 가구전화를 사용해 발생하는 문제 등 체계적 편차 요인에 대한 논의와 응답률 제고를 위한 방안이 논의되고 있다. 뿐만 아니라, 인터넷 이용이 확산되고 휴대전화의 사용이 일반화되는 등 디지털 미디어 보급이 가속화되면서 이러한 매체들을 이용해 기존의 조사방법을 보완하거나 대체함으로써 기존 조사방법의 한계를 극복하려는 시도가 이루어지고 있다.

그런데 기존의 조사방법에 웹이나 모바일을 결합한 새로운 조사방법, 예컨대 혼합모드조사(mixed mode survey) 방법을 수립하기 위해서는 연구설계, 표집, 데이터 수집, 조사결과에 대한 분석 및 평가 등 조사의 각 단계별로 기존의 방법론에 대한 면밀한 재검토가 필요하다. 기존의 조사방법론은 유선전화 및 면접조사 방법을 중심으로, 그리고 단일 모드를 중심으로 구축되어 온 만큼, 이중 매체를 활용해 조사 대상자를 모집할 때 표본간의 이질성을 어떻게 처리하고 대표성을 확보할 것인지, 측정을 위한 문항 및 설문지를 어떻게 구성할 것인지, 그리고 수집한 자료를 어떻게 통합적으로 분석할 것인지 등 전반적인 방법을 새로운 조사 환경에 맞게 수정·보완할 필요가 있다.

본 논문에서는 웹, 모바일 등을 활용한 새로운 조사기법과 관련된 방법론적 이슈들 가운데서 척도 구성의 측면에 주목했다. 전통적인 조사연구에서는 문항을 작성할 때 다 문항의 다점 척도를 이용하는 경우가 많다. 즉 하나의 개념을 측정하기 위해 여러 개의 문항을 사용하고, 각각의 문항은 5점에서 11점에 달하는 다점 척도를 사용해 왔다. 그런

데 이러한 방식이 웹이나 모바일을 활용한 응답상황에도 적합한지는 분명하지 않다. 피조사자들의 집중도가 현저하게 떨어지는 온라인 조사에서 길고 복잡한 문항은 탈락률을 높일 가능성이 있고, 이는 표본의 문제와 함께 조사의 타당성과 결과의 신뢰도 및 응답의 편이성을 저해하는 주요인이 될 수 있다.

따라서 웹과모바일 조사에서 문항 및 척도 구성방식이 피조사자의 응답에 어떤 영향을 미치는지를 면밀하게 검토할 필요가 있다고 보고, 본 논문에서는 우선 웹기반 조사를 중심으로 척도 크기에 따라 응답의 타당도와 신뢰도, 그리고 응답의 편이성에 차이가 생기는지 검토하고자 했다. 구체적으로 동일한 질문에 대해 상이한 크기의 척도가 제시됐을 때 주관적 차원(피조사자 스스로가 느끼는 응답의 용이성, 부담감, 짜증)과 객관적 차원(문항에 대한 반응 시간)에서 응답의 편이성에 차이가 있는지, 그리고 측정 문항의 (사후적) 타당도와 신뢰도가 어떻게 달라지는지 비교하고자 했다.

II. 이론적 논의

1. 웹기반 조사의 장점과 한계

전통적으로 사회 조사에는 면접조사나 전화조사 방법이 주를 이루어 왔다. 그런데 이러한 방법들은 기본적으로 가구 방문 또는 가구 단위의 유선전화를 접촉 단위로 하기 때문에 낮 시간 재택인구의 감소나 무선전화의 보급과 같은 생활환경의 변화와 함께 여러 가지 문제점에 부딪히게 되었다. 이에 따라 요일이나 시간대별로 재택률의 차이가 발생하는 데 따른 표본편포 문제, 응답률, 가구기반 조사로 인한 체계적 편파 요인 등이 대두되고 있다(강남준·백영민 2005). 이러한 기존 조사방법의 한계를 극복하기 위해 RDD (random digit dialing)를 사용해 표집 문제를 해결하거나, 조사 시간대를 할당변수로 추가해 인구학적 속성에 따른 재택률 차이가 유발하는 체계적 응답자 선택편향을 해결하는 방법 등이 제시되기도 했다(허명희·황진모 2006).

또 대안적 조사방법으로 인터넷이나 모바일 미디어를 이용한 사회조사 방법이 제시되고 있으며 이와 관련해 활발한 논의가 이루어지고 있다. 특히 웹기반 조사는 유선조사로는 접근하기 어려운 1인 가구나, 20-30대 젊은 층을 대상으로 하는 조사를 수행하는 데 용이하고(이계오·장덕현 2009), 조사원의 일정과 관계 없이 시간대에 구애받지 않고 자

유로운 조사 참여가 가능하며, 응답의 별도 코딩이 불필요하고, 멀티미디어 자료를 활용할 수 있으며 조사원에 의한 오차도 감소하는 등(Tourangeau et al. 1997) 여러 가지 장점을 가지고 있다. 이에 따라 여러 사회조사에서 웹기반 조사가 점차 증가하는 추세이다.

하지만 웹기반 조사방법이 아직 체계적으로 정립되어 있지 않고 그 타당성 또한 아직 검증되지 않은 상태이다. 또 여러 가지 오차 요인도 제시되고 있다. 우선 인터넷 조사는 인터넷에 접근 가능한 응답자만이 참여할 수 있다는 점에서 표본의 포함률 오차(coverage error)가 발생한다(강남준·백영민 2005). 인터넷으로 조사를 시행하는 경우 인터넷 보급률이 상대적으로 낮은 계층, 예컨대 고령층이나 인터넷 접속이 어려운 지역의 거주자 그리고 시각 장애인 등의 조사 참여가 어렵기 때문이다. 또 현재 이루어지고 있는 인터넷 조사는 보통 조사 회사가 구축하고 있는 패널들 가운데 희망자들이 자발적으로 참여하는 방식으로 진행되는데, 이에 따라 자기선택에 의한 오차(self-selection error)가 발생하기 쉽다. 또한 무응답 비율이 면접조사에 비해 더 높은 것으로 알려져 있다(Tourangeau et al. 1997).

이 가운데 포함률 오차는 인터넷 보급률이 증가하면서 점차 감소하고 있지만, 자기선택 오차는 여전히 중요한 이슈이다(허명희·조성겸 2010). 이를 해결하기 위해 표집 단계에서 확률표집 방법을 적용하거나, 통계적 모델을 이용해 오차를 제거하는 방법 등이 제시되었다. 예컨대 인터넷을 통한 조사 참여와 관련된 변수들을 검토해 인터넷 표본이 전화조사 표본과 유사하도록 보정하는 성향지수 가중법(propensity weighting)이 한 예이다. 이러한 방법 외에, 기존의 조사방법과 웹기반 조사를 혼용함으로써 웹기반 조사의 한계를 보완하고 극복하는 방안이 제시되기도 한다.

이러한 한계점들 외에 최근에는 조사 내용의 구성 방식에 대한 재검토가 필요하다는 논의도 이루어지고 있다. 웹기반 조사에서는 다문항의 다점 척도에 대한 응답 협조도가 낮아지는 점을 해결하기 위해 응답 시간을 이용해 불성실 응답자를 판명하거나(이윤석 외 2008), 문항의 작성 단계에서부터 웹기반 조사의 특수성을 고려해야 할 필요성이 제기되고 있다. 웹기반 조사의 환경이 유선전화나 면접조사 등의 경우와는 상이한 만큼, 기존의 조사환경에 근거해 구축되어 온 문항 및 척도구성 방법에 대한 수정이 필요하다는 것이다. 이에 본 논문에서는 웹기반 조사와 관련된 다양한 방법론적 주제들 가운데 척도의 문제에 주목했다.

2. 웹기반 조사에서 척도 크기와 타당도, 신뢰도, 응답 편이성의 문제

어떤 대상에 대한 사람들의 판단은 영속적이고 불변하는 것이라기보다는 여러 가지 상황적 요인으로 인하여 임시적으로 구성되는, 즉 맥락 의존적인 경우가 많다(Wilson & Hodges 1992). 사람들이 평소에 어떤 대상에 대한 평가나 신념을 가지고 있는 것은 사실이지만, 대상에 대한 순간적인 판단, 예컨대 질문이 주어졌을 때 그에 대한 응답은 자신이 처한 상황이나 선행자극(prime) 등에 따라 달라질 수 있다. 사람들이 판단을 내릴 때 많은 인지적 노력을 기울이기보다 최소한의 인지 자원을 활용해 빠른 판단을 내리려는 경향이 있기 때문이다(Fiske 1992; Tversky & Kahneman 1974). 즉, 설문에 대한 응답은 질문이 주어진 순간에 가장 접근 가능한 기억에 근거하여 이루어진 정보 처리의 결과인 경우가 많다(Fazio & Williams 1986; Oskamp & Schultz 2005 참조).

조사연구에서 문항 및 설문의 구성이 중요한 이유가 바로 이 때문이다. 질문이 제시되는 순서나 안내문의 내용, 응답지 등에 따라 피조사자들이 활성화하는 기억 정보가 달라지고, 이에 따라 응답의 결과는 달라질 수 있다. 예컨대 “최근 초·중·고교 무상급식이 사회적 이슈로 대두되었습니다.”라고 안내문이 제시된 경우와, “최근 초·중·고교에서 무상 급식을 실시해야 한다는 의견이 늘어나고 있습니다.”라고 안내문이 제시된 경우, 이에 대한 의견을 물었을 때의 응답의 결과가 다를 수 있다. “무상급식에 대한 한국 사회의 여론은 대체로 어떤 것 같습니까?”라는 질문과 “무상급식에 대한 귀하의 의견은 어떻습니까?”라는 질문 가운데 어떤 것이 먼저 제시되는가에 따라서도 응답이 달라질 수 있다. 또 응답지가 ‘매우 찬성’과 ‘매우 반대’ 사이에서 선택하도록 한 경우와 ‘찬성’과 ‘반대’ 사이에서 선택하도록 한 경우에도 답이 달라질 수 있다.

이처럼 맥락적 요소에 따라 응답이 달라질 수 있는 만큼, 문항과 척도를 어떻게 구성하는가 하는 것은 매우 중요하다. 일반적으로 설문 및 문항구성의 가장 중요한 원칙은 (1) 조사목적에 맞는 자료를 수집할 수 있는 타당한 내용으로 설문과 문항을 구성해야 하고, 이를 통해 (2) 신뢰성 있는 자료를 수집할 수 있어야 한다는 것이다(Fowler 1995; Tourangeau et al. 2000). 따라서 연구자들은 사전에 전문가 및 예비조사 풀을 활용해 문항의 내용 및 응답지를 신중하게 검토해야 한다. 또 피조사자들이 제시된 질문을 연구자가 의도한 대로 정확하게 이해할 수 있도록 질문과 응답지를 명료하게 작성하고, 피조사자들이 지나치게 바람직한(socially desirable) 응답을 하려 하거나, 거부감이나 부담감

을 느끼지 않도록 문항과 응답지를 작성해야 한다(Tourangeau et al. 2000). 즉 측정의 타당성, 응답의 신뢰성과 함께 응답의 편이성을 높이기 위해 노력해야 하는 것이다.

이러한 문항 구성의 문제와 관련해서는 주로 문항 작성 자체에 대한 관심이 높다. 즉, 문항을 명료하게 작성해 오해의 가능성을 줄이고, 하나의 질문에 두 가지 이상의 내용이 포함되지 않도록 작성해야 한다는 등의 지침이 그것이다. 그런데 이러한 문항 작성에 있어 질문의 구성뿐만 아니라 척도의 크기 역시 매우 중요한 요소이다. 척도의 크기에 따라 응답의 용이성이나 응답 결과가 달라질 수 있기 때문이다.

일반적으로, 여건이 허락하는 한 다점의 척도를 사용하는 것이 일반적이다. 척도의 크기가 클수록 응답의 분산이 커져 자료 분석에 용이하기 때문이다. 물론 너무 큰 척도를 사용하는 것도 바람직하지는 않다. 10점 이상의 척도를 사용한다고 해서 더 좋은 결과나 새로운 정보가 제시되는 것은 아니라고 한다(Andrews 1984). 그런데 구체적으로 척도의 크기에 따라 문항의 신뢰도 및 타당도가 어떻게 달라지는지에 대한 연구 결과는 다양하다. 예컨대 Jacoby & Matell(1971)은 다점 척도(2~19점)를 2점 및 3점 척도로 변환해 신뢰도와 타당도를 비교했다. 그 결과 타당도와 신뢰도는 원 척도를 사용한 경우에 비해 다소 낮았지만 통계적으로 유의한 수준은 아니었다. 따라서 이들은 척도의 수를 줄이는 것이 신뢰도나 타당도에 부정적인 영향을 미치지 않는다고 보았다. 그런가 하면 5점까지는 신뢰도가 증가하지만 그 이상의 척도를 사용한다고 해서 신뢰도가 증가하지는 않는다고(Jenkins & Taber 1977; Lissitz & Green 1975, Cicchetti et al. 1985에서 재인용), 7점까지는 척도 크기에 따라 신뢰도가 증가한다는 결과도 있어(예, Cicchetti et al. 1985) 이에 대한 명확한 판단은 어려운 상태이다.

다만 척도 크기가 작을수록 응답의 분포가 좁기 때문에 자료의 분석이 수월하지 않다는 점에 대해서는 여러 연구자들이 공감하고 있다. 예컨대 Cohen(1983)은 척도를 이항화하는 경우 무엇보다도 통계적 검증력(statistical power)의 감소가 문제라고 보았다. 척도를 이항화하는 경우 통계적 검증력이 적게는 38%에서 많게는 60%까지 감소한다는 것이다. 통계 분석 기법은 대부분 자료의 분산과 관련이 크기 때문에 척도 크기가 클수록 연구자가 예상한 바에 가까운 결과를 얻을 수 있고, 이항 척도를 사용하는 경우 적용하기 어려운 분석법도 많다. 따라서 5~7점 정도의 척도를 사용하는 것이 대개의 경우 가장 적절한 것으로 간주되고 있다(Fowler 1995).

이러한 문제는 사회정치적 이슈와 관련한 조사에서 더 중요하게 간주된다. 사회적 질

문에 대한 답을 예/아니오 또는 찬성/반대와 같은 이항으로 응답하는 것이 쉽지 않고, 극단적인 태도를 가진 사람들이 용이하게 응답을 할 수 없으며, 그에 따라 결과적으로 정보의 손실이 발생할 수 있기 때문이다(Stober & Dette 2002). 따라서 적어도 6점 이상의 척도를 사용하는 것이 좋다(Green & Rao 1970)는 견해가 제시되곤 한다.

또한 척도 크기는 응답의 용이성에도 영향을 미치는 것으로 알려져 있다. 보통 전화 조사에서는 척도 크기를 작게, 지필 조사에서는 크게 하는 경우가 많다. 전화 조사의 경우 응답지를 들으며 기억해야 하는데, 많은 응답 보기가 제시되는 경우 참여자들이 이를 일일이 기억하고 응답하기가 어렵다(Fowler 1995). 때문에 조사 회사에서는 전화 조사를 실시할 때 2점이나 4점의 척도를 선호한다고 한다(2011. 4. 28. 현대리서치연구소 이상경 대표이사와의 대화). 그에 비해 지필 조사의 경우에는 사람들이 일일이 응답지를 기억할 필요가 없어 기억의 어려움으로 인한 응답의 용이성 문제가 발생하지 않으므로, 분산을 증가시켜 분석의 용이성을 높이기 위해 보다 큰 척도를 사용하는 경우가 많다.

그렇다면 웹기반 조사에서는 어떤 크기의 척도가 더 적합할 것인가? 설문의 외형만을 본다면 웹기반 조사는 전화 조사보다는 지필 조사와 더 유사하다. 이 점을 생각한다면 다점 척도를 사용해도 무방할 것이다. 하지만 웹기반 조사는 지필 조사에 비해 참여자들의 성실도가 현저하게 낮다는 문제가 있다. 웹기반 조사에서는 무응답 비율이 면접조사에 비해 더 높으며(Tourangeau et al. 1997), 응답에 소요되는 시간도 매우 짧기 때문이다. 이에 따라 다점 척도에 대한 기피율이 증가하게 되어(이윤석 외 2008) 응답의 불성실도가 증가하여 연구의 타당도와 신뢰도에 영향을 미칠 수 있다. 또 보통 인지적 노력이 불충분할수록, 예컨대 교육 수준이 낮은 사람일수록 응답의 정확도가 떨어지며 긍정적인 답(찬성)을 선호하는 경향이 있는데(Converse & Presser 1986), 그렇다면 웹기반 조사에서도 이와 같이 응답의 편향이 나타날 수 있다.

이에 본 논문에서는 웹기반 조사에서 상이한 크기의 척도를 사용했을 때 응답의 사후적 타당도와 문항간 신뢰도, 그리고 피조사자의 주관적, 객관적 응답 편이성을 비교해 웹기반 조사에서 어떤 크기의 척도를 사용하는 것이 더 적절한지에 대한 답을 모색하고자 했다. 본 연구에서 탐색하고자 하는 연구문제는 다음과 같다.

연구문제 1. 웹기반 조사에서 2점, 6점, 10점 척도를 사용할 때 측정 변인의 사후적 타당도에 어떤 차이가 있는가?

연구문제 2. 웹기반 조사에서 2점, 6점, 10점 척도를 사용할 때 문항간 신뢰도에 어떤 차이가 있는가?

연구문제 3. 웹기반 조사에서 2점, 6점, 10점 척도를 사용할 때 피조사자의 주관적 응답 편이성(응답 용이성, 짜증, 부담감), 객관적 응답 편이성(문항 반응 시간)에 어떤 차이가 있는가?

III. 연구방법

서울과 충남에 소재한 4년제 대학에서 커뮤니케이션 관련 강좌를 수강하는 학생 258명을 대상으로 일원요인설계(척도 크기: 2점 vs. 6점 vs. 10점)를 적용한 실험을 실시했다. 실험 참여자들의 평균 연령은 21.38세($SD = 2.32$)였고, 이들 가운데 남학생이 92명(35.7%), 여학생이 166명(64.3%)이었다.

1. 실험 절차

각 대학의 강좌별 담당 교수가 전자우편이나 강의 게시판을 이용해 실험 사이트 주소를 안내하고, 학생들이 직접 웹사이트에 접속해 실험에 응하도록 했다. 실험용 웹사이트는 쉐트릭스(Qualtrics.com)의 서베이 솔루션을 이용해 제작했다. 참여자들이 웹사이트에 접속해 연구에 대한 간략한 안내문을 읽고(인터넷 뉴스 이용 및 평가에 관한 연구로 소개) 연구 참여에 동의하면, 학생들은 척도 크기에 따른 세 개의 실험조건 가운데 한 조건에 무작위로 할당되었다.

실험 자극으로는 인터넷 뉴스와 관련된 질문들을 담고 있는 총 4쪽의 설문을 제시하고 조건에 따라 서로 다른 척도를 사용한 설문에 응답하게 했다. 참여자들이 응답하는 동안 반응에 소요된 시간을 100분의 1초 단위로 기록했다. 실험 자극에 대한 응답을 마친 후 참여자들은 다음 페이지로 이동해 설문을 진행하는 동안 응답하기가 얼마나 수월했는지, 얼마나 부담스러웠는지, 얼마나 짜증을 느꼈는지 응답했다. 마지막으로 성별, 연령, 국적을 입력하고(외국인 학생의 응답은 분석에서 제외) 실험을 종료했다.

2. 실험 처치

실험 처치를 위해 2점, 6점, 또는 10점 척도로 이루어진 설문에 응하게 했다. 이와 같은 척도를 사용한 이유는 중간값의 존재 여부가 응답 편이성에 영향을 미칠 수 있기 때문에 일관되게 짝수 또는 홀수로 된 척도를 사용해야 하리라고 보았기 때문이다. 그런데 중간값은 중립적인 의견을 의미할 수도 있고 의견 없음 또는 이슈에 대한 무지를 의미할 수도 있어(Fowler 1995) 결과의 해석에 복잡성을 초래한다. 본 연구에서는 척도 크기에 따른 차이에 주목하고자 하므로 중간값은 제외하기로 했다.

모든 문항을 세 종류의 척도로 제시해야 하므로 성별, 지역, 학력과 같은 변인을 피하고 2점으로 또는 10점으로 제시해도 어색하지 않을 문장들을 제시하고자 했다. 구체적으로 인터넷 뉴스 이용과 관련된 일반적 질문, 인터넷 뉴스에 대한 평가, 분야별 인터넷 뉴스 이용 정도, 인터넷 뉴스 이용 동기를 묻는 총 33개 문항을 4쪽에 걸쳐 제시했다. 설문이 점점 길어진다는 느낌을 주기 위해 다음 페이지로 이동할수록 문항 수가 증가하게 했다(첫 페이지 4문항에서 2페이지 7문항, 3페이지 9문항, 4페이지 13문항으로 증가).

첫 번째 페이지에는 인터넷 이용 빈도, 종이신문과 비교한 인터넷 신문 선호도, 댓글을 읽고 쓰는 정도를 묻는 4개 문항을 제시했다. 두 번째 페이지에는 인터넷 뉴스에 대한 평가를 묻는 7개 문항을 제시했다. 이 문항들은 매체 및 정보원 공신력에 대한 기존 연구들을 참조해(Flanagin & Metzger 2003; McCroskey & Teven 1999) 구성한 것으로, 최근 온라인 뉴스에 대한 평가를 측정하는 국내의 선행 연구에서도 유사한 문항들을 사용한 바 있다(김은미·선유화 2006; 이은주·장윤재 2009).

다음으로 세 번째 페이지에서는 9개의 뉴스 주제 분야를 제시하고, 분야별 인터넷 뉴스 이용을 물었다. 마지막으로 네 번째 페이지에서는 인터넷 뉴스 이용 동기를 묻는 13개 문항을 제시했다. 박상호(2009)가 제시한 8가지의 인터넷 포털뉴스 이용동기(뉴스 흥미성 추구, 편리성 추구, 뉴스전문성 추구, 뉴스사이트 유용/신뢰성 추구, 사건/정보 추구, 휴식/재미 추구, 의견교환/토론 추구, 일상성 추구) 가운데 편리성 추구 동기, 뉴스전문성 추구 동기, 사건/정보 추구 동기를 측정하는 문항을 사용했다. 구체적인 문항 및 척도 구성 내용은 <표 1>에 제시했다.

〈표 1〉 실험처치를 위한 문항 구성

변인	측정문항		측정척도
1. 인터넷 뉴스 이용 일반	나는 인터넷을 자주 사용한다 나는 인터넷 뉴스보다 종이신문을 더 많이 본다 나는 인터넷 뉴스를 읽을 때 기사에 달린 댓글을 읽는다 나는 온라인 뉴스를 읽고 댓글을 단다		2점 척도: 그렇다(1), 아니다(2) 6점 척도: 전혀 아니다(1) ~ 매우 그렇다(6)
2. 인터넷 뉴스 평가	전문적이다 사실적이다 객관적이다 믿을 만하다 공정하다 신속하다 선정적이다		10점 척도: 전혀 아니다(1) ~ 매우 그렇다(10)
3. 주제별 뉴스 이용	정치 경제 사건/범죄 국제 연예 스포츠 IT/정보기술 환경 패션		2점 척도: 이용안함(1), 이용함(2) 6점 척도: 전혀 이용안함(1) ~ 자주 이용함(6) 10점 척도: 전혀 이용안함(1) ~ 자주 이용함(10)
4. 인터넷 뉴스 이용 동기	편리성 추구	인터넷 뉴스로 보는 것이 편하기 때문에 타 매체를 볼 필요가 없어서 시공간의 제약이 거의 없어서 포털을 통해 모두 한 번에 해결할 수 있어서	2점 척도: 그렇다(1), 아니다(2) 6점 척도: 전혀 아니다(1) ~ 매우 그렇다(6) 10점 척도: 전혀 아니다(1) ~ 매우 그렇다(10)
	뉴스 전문성 추구	뉴스기사 내용이 충실하고 깊이 있기 때문에 뉴스가 믿음이 가고 정확하기 때문에 각 언론사별 뉴스를 비교할 수 있어서 지난 뉴스를 확인할 수 있어서	
	사건 /정보 추구	사회의 다양한 사건/문제 점을 알 수 있어서 원하는 자료나 정보를 찾기 위해 일상생활의 다양한 정보를 얻기 때문에 큰 사건의 발생을 확인하기 위해 다른 매체에서 얻은 정보를 확인하기 위해	

3. 측정

1) 응답 타당도

일반적으로 측정 문항 및 척도의 타당성은 통계적으로 검증하기보다 분야 전문가의 주관적 판단에 근거해 평가하는 경우가 더 많다. 하지만 표준화된 척도나 여타 연구와의 비교, 또는 요인분석 등의 방법을 활용해 사후적으로 타당도를 검증하기도 한다. 특히 요인분석은 잠재적 요인을 중심으로 측정 문항이나 변수들의 선형 결합 가능성을 보여주므로 비록 결과 자체가 내용의 타당성을 입증해 주지는 않지만 측정 문항들의 논리적 일관성을 뒷받침하는 바, 타당도 검증을 위한 도구로도 종종 이용한다(Garson 2011 참조).

본고에서는 타당도를 사후적으로 확인하기 위해 실험 자극으로 사용한 문항들을 이용해 탐색적 요인분석을 실시했다. 실험자극 문항 가운데 인터넷 뉴스 평가를 측정한 7개 문항, 주제 분야별 뉴스 이용을 측정한 9개 문항, 인터넷 뉴스 이용 동기를 묻은 13개 문항을 대상으로 베리맥스 직교회전을 적용한 주성분 분석을 수행했다. 다만 2점 척도를 사용한 응답에 대해 직접 요인분석을 실시하는 경우 상관계수의 왜곡 가능성이 있으므로, 2점 척도 사용 집단의 응답에 대해서는 우선 사분상관행렬(tetrachoric correlation matrix)을 추출한 후 이를 사용해 요인분석을 실시했다(Fleming 2005 참조).

2) 응답 신뢰도

신뢰도 점수로는 가장 일반적으로 사용하는 크론바흐의 알파(Cronbach's α) 값을 사용했다. 유사 연구에서도 크론바흐의 알파 계수를 사용한 경우가 많다(Cicchetti et al. 1985; Jacoby & Matell 1971). 다만 이항 척도에서 신뢰도가 과소하게 평가될 것을 고려해, 크론바흐 알파 값 외에 추가로 상관분석을 실시하여 피어슨 상관계수(Pearson's r)를 사용하기로 했다. Cicchetti et al.(1985)의 연구에서도 크론바흐 알파 계수를 적용하기 어려운 척도에 대해서는 상관계수를 사용해 신뢰도를 분석한 바 있다.

구체적으로, 뉴스 이용 동기 가운데 “뉴스가 믿음이 가고 정확하기 때문에”라는 문항에 대해 긍정적으로 답한 참여자들은 인터넷 뉴스에 대한 평가 항목 가운데 “객관적이다”, “믿을 만하다”, “공정하다”는 문항에 대해서도 긍정적으로 답할 가능성이 높을 것으로 보고 이들 문항 간의 상관을 분석했다. 또, 인터넷 뉴스에 대해 공신력을 높이 평가하기보다 신속성과 흥미성을 높이 평가한 점과 관련하여 종이신문에 비해 인터넷 신문을

더 많이 이용하는 것은 주로 인터넷 뉴스의 편리성과 관련되어 있을 것이라고 예상하고 “나는 인터넷 뉴스보다 종이신문을 더 많이 본다”는 문항과 인터넷 이용 동기 가운데 편리성 추구 동기를 측정하는 4개 문항과의 상관을 분석했다.

3) 응답 편이성

응답 편이성은 주관적 차원과 행위 차원으로 구분해 측정했다. 우선 응답자들이 실험 자극 문항들에 대해 응답하는 동안 문항에 대한 반응 시간을 기록하여 행위 차원의 응답 편이성을 파악하고자 했다. 반응 시간은 100분의 1초 단위로 기록했으며, 페이지별로 첫 번째 클릭, 마지막 클릭, 페이지 이동 버튼 클릭까지 소요된 시간을 기록했다. 주의 분산 및 여타 행동에 쓴 시간 등을 최대한 통제하고자, 마지막 질문에 대한 응답 시각에서 첫 번째 질문에 응답한 시각을 빼고 (페이지당 문항 수 - 1)로 나누어 문항당 평균 응답 소요시간을 구했다¹⁾.

주관적 차원의 응답 편이성은 설문을 진행하며 응답하기가 얼마나 쉬웠는지, 부담스럽거나 짜증스러웠는지를 묻는 문항들을 제시하고 7점 척도상에 응답하게 하여 측정했다. 응답의 용이성 측정을 위해서는 앞서 제시된 질문들에 답을 하기가 쉬웠는지 묻고 ‘즉시 답하기 어려웠다(1)’ ~ ‘즉시 답할 수 있었다(7)’, ‘내 의견을 제대로 표현할 수 없었다(1)’ ~ ‘내 의견을 제대로 표현할 수 있었다(7)’, ‘답하기 불편했다(1)’ ~ ‘답하기 편했다(7)’는 문항에 응답하게 했다. 이들 문항에 대한 응답의 평균을 구해 응답 용이성 점수로 사용했다($M = 5.76$, $SD = 1.18$, Cronbach's $\alpha = .81$).

이어서 질문에 대답하며 어떤 느낌을 받았는지 묻고, 짜증을 측정하기 위해 ‘짜증나지 않았다(1)’ ~ ‘짜증이 났다(7)’, ‘귀찮지 않았다(1)’ ~ ‘귀찮았다(7)’, ‘번거롭지 않았다(1)’ ~ ‘번거로웠다(7)’는 3개 문항을 제시하고, 이들 문항에 대한 응답의 평균을 구했다($M = 2.00$, $SD = 1.24$, Cronbach's $\alpha = .92$), 부담감을 측정하기 위해 ‘부담스럽지 않았다(1)’ ~ ‘부담스러웠다(7)’, ‘질문이 많지 않았다(1)’ ~ ‘질문이 많았다(7)’, ‘힘들지 않았다(1)’ ~ ‘힘이 들었다(7)’는 3개 문항을 제시하고 평균값을 구해 사용했다($M = 1.77$, $SD = .96$, Cronbach's $\alpha = .83$).

1) 예를 들어 어떤 참여자가 문항이 7개인 두 번째 페이지를 로드한 후 5초만에 첫 번째 문항에 응답하고 23초에 마지막 7번째 문항에 응답했다면, 이 참여자의 문항당 평균 반응시간은 $(23-5)/(7-1)$, 즉 3초이다.

IV. 분석 결과

세 집단이 인구학적으로 동질한지 확인하기 위해(equivalence test) 각 집단의 연령과 성비를 비교한 결과, 세 집단 사이에 연령차($F = .35, p = .71$)나 성차($\chi^2 = 1.42, p = .49$)가 존재하지 않음을 확인했다. 실험 자극은 집단별로 상이한 척도를 제시하는 방식으로 이루어졌으므로, 별도의 조작화 검증(manipulation check)은 실시하지 않았다.

본격적인 분석에 앞서 척도 크기별 응답값을 Fisher의 Z점수로 표준화한 값을 아래 <표 2>에 제시했다. 대체로 척도 크기가 클수록 평균 및 표준편차가 일정하게 증가하거나 감소하는 경향을 보이고 있다²⁾.

<표 2> 주요 변인의 표준화된 기술통계치

측정변인		10점척도		6점척도		2점척도	
		M	SD	M	SD	M	SD
인터넷 이용	전반	.94	1.01	-.01	.50	-.92	.11
	뉴스	-.57	1.36	.00	.72	.56	.21
	댓글읽기	.48	1.37	-.12	.86	-.35	.21
	댓글쓰기	-.64	1.30	-.02	.69	.65	.19
인터넷 뉴스 평가		-.62	1.36	.11	.70	.50	.22
인터넷 뉴스 이용동기	편리성	-.53	1.40	.11	.77	.41	.18
	전문성	.52	1.44	-.01	.62	-.51	.19
	정보추구	.79	1.19	-.02	.61	-.76	.16
주제별 인터넷 뉴스 이용	정치	-.16	1.48	.16	.85	.00	.30
	경제	-.39	1.42	.16	.84	.22	.30
	사건/범죄	.42	1.40	-.02	.83	-.40	.20
	국제	.18	1.48	.01	.84	-.18	.25
	연예	.71	1.21	-.04	.78	-.66	.15
	스포츠	-.05	1.50	.12	.85	-.07	.23
	IT/정보기술	.02	1.48	-.11	.88	.08	.26
	환경	-.03	1.49	-.14	.83	.17	.31
패션	.08	1.54	.02	.78	-.10	.24	

주) 10점 척도 집단 n = 85, 6점 척도 집단 n = 87, 2점 척도 집단 n = 86

2) Z 점수는 평균을 0, 표준편차를 1로 표준화한 값이므로 척도 크기가 클수록 표준편차가 증가하리라는 것은 쉽게 짐작할 수 있다. 평균의 경우, 일정하게 증감하기는 하지만 증감의 방향은 문항에 따라 차이를 보인다. 인터넷 이용과 관련된 문항들인 만큼 응답자들의 성향 및 사회적 바람직성(social desirability) 등에 따라 이 결과의 편향성을 분석해 볼 수 있을 것이나 본 연구에서는 관련 자료가 부족하고 연구문제가 응답의 타당도, 신뢰도, 편이성에 초점을 두고 있으므로 관련 분석을 실시하지 않았다.

〈표 3〉 인터넷 뉴스 평가, 이용, 이용 동기 측정 문항 요인분석 결과

		10점 척도			6점 척도			2점 척도		
요인	요인	고유값	설명 분산	요인	고유값	설명 분산	요인	고유값	설명 분산	
인터넷 뉴스 평가										
I	전문, 사실, 객관, 신뢰, 공정	3.77	53.28	전문, 사실, 객관, 신뢰, 공정	3.23	43.68	객관, 신뢰, 공정, 신속	3.49	48.38	
II	신속, 선정	1.44	21.26	신속, 선정	1.10	18.15	전문, 사실, 선정	2.19	32.86	
주제별 인터넷 뉴스 이용										
I	정치, 사건/범죄, 국제	3.40	25.77	정치, 사건/범죄, 국제, 환경, 경제, IT/정보기술	3.33	33.00	정치, 사건/범죄, 국제, 환경	3.02	29.33	
II	연예, 패션, 환경	1.50	22.07	연예, 패션	1.50	16.59	연예, 패션	2.23	25.03	
III	경제, 스포츠, IT/정보기술	1.15	19.42	스포츠	1.00	15.21	경제, 스포츠, IT/정보기술	1.50	20.72	
인터넷 뉴스 이용 동기										
I	뉴스 확인, 다양한 사건 확인, 검색, 일상정보, 대형사건확인, 타매체 보도 확인, 타매체 볼 필요 없음	5.09	23.62	뉴스 비교, 뉴스 확인, 다양한 사건 확인, 검색	2.25	16.48	믿음가고 정확, 뉴스 확인, 다양한 사건 확인, 검색, 일상정보	6.15	34.29	
II	내용충실, 깊이, 믿음가고 정확	1.30	15.04	내용충실, 깊이, 믿음가고 정확	1.41	15.13	내용충실, 깊이, 타매체 보도 확인	2.11	20.04	
III	편리, 시공간 제약없음, 포털통해 한 번에 해결	2.08	23.06	편리, 시공간 제약없음, 포털통해 한 번에 해결, 타 매체 볼 필요 없음	3.57	18.04	타매체 볼 필요 없음, 시공간 제약없음	1.68	18.29	
IV	뉴스 비교	1.04	11.33	일상정보, 대형사건확인, 타 매체 보도 확인	1.12	14.56	뉴스 비교, 대형사건확인, 편리, 포털통해 한번에 해결	3.63	31.69	

1. 타당도 분석

요인분석 결과, 척도에 따라 요인 추출 결과가 상이하게 나타났다(〈표 3〉 참조). 인터넷 뉴스에 대한 평가는 10점 척도와 6점 척도 집단에서 동일하게 뉴스의 공신력(전문, 사실, 객관, 신뢰, 공정성)과 선정주의(신속, 선정성)을 가리키는 두 개의 요인이 추출되었다. 2점 척도를 사용한 집단의 응답에서도 두 개의 요인이 추출되기는 했지만 나머지 두 집단과는 다르게 객관, 신뢰, 공정, 신속성을 포함하는 요인과 전문, 사실, 선정성을 포함하는 요인이 추출되었다.

주제 분야별 뉴스 이용에 대해서는 10점 척도와 2점 척도가 유사한 결과를 보였다. 10점 척도에서는 정치, 사건/범죄, 국제 vs. 연예, 환경, 패션 vs. 경제, 스포츠, IT/정보기술로 구성되는 세 개 요인이, 2점 척도에서는 정치, 사건/범죄, 국제, 환경 vs. 연예, 패션 vs. 경제, 스포츠, IT/정보기술로 구성되는 세 개 요인이 추출됐다. 이와 달리 6점 척도를 사용한 집단에서는 연예, 패션(요인 II)과 스포츠(요인 III)를 제외한 나머지 모든 주제가 하나의 요인으로 묶였다.

인터넷 뉴스 이용 동기의 경우, 10점 척도를 사용한 경우와 6점 척도를 사용한 경우 일부 문항이 일치하기도 하지만 요인의 성분이 정확하게 일치하는 경우는 없었다. 이 문항을 제시한 박상호(2009)의 연구대로라면 편리성 추구 동기, 뉴스 전문성 추구 동기, 사건/정보 추구 동기의 세 개 요인이 추출될 것으로 기대되지만, 본 실험에서는 모든 집단에서 네 개 요인이 추출되었다. 편리성, 뉴스 전문성 및 정보 추구 동기로 명명할 만한 요인들이 추출되기는 했지만, 각 요인의 소속 변수들이 예상한 바와는 상당한 차이를 보일 뿐 아니라 척도에 따라 상당한 불일치를 보이고 있다. 모든 경우 설명분산은 2점 척도가 가장 크고 다음이 10점 척도, 6점 척도 순이었다.

2. 신뢰도 분석

집단별로 요인 추출 결과가 상이한 바, 유사한 주성분 분석 결과를 보인 변인들을 선별해 문항간 신뢰도를 분석했다. 인터넷 뉴스에 대한 평가 측정 문항 가운데서는 ‘객관적이다’, ‘믿을 만하다’, ‘공정하다’ 세 개 문항을, 주제 분야별로는 정치, 사건/범죄, 국제 뉴스에 대한 이용 응답을 대상으로 신뢰도를 분석했다. 인터넷 뉴스 이용 동기 측정 문항에 대해서는 세 집단의 응답에 대한 요인분석 결과가 상이하므로, 선행 연구(박상호 2009)에서 제시한 대로 편리성 추구 동기, 뉴스 전문성 추구 동기, 사건/정보 추구 동기

로 문항을 묶어 새로운 변수를 구성하는 경우를 가정했다.

〈표 4〉에서 확인할 수 있는 것처럼, 문항간 신뢰도는 대체로 10점 척도가 가장 높고 2점 척도가 가장 낮았다. 그런데 이러한 결과는 다점, 다문항 척도일수록 신뢰도가 높아지는 신뢰도 산출 공식상의 특징으로 인한 것일 수 있다. 따라서 추가적으로 일부 변인을 대상으로 하는 상관분석을 실시해 응답의 일관성을 검토했다.

뉴스가 믿음이 가고 정확하기 때문에 인터넷 뉴스를 이용한다는 문항에 대한 응답과 인터넷 뉴스에 대한 평가(객관성, 신뢰성, 공정성) 사이의 상관을 분석한 결과, 세 집단 모두에서 높은 수준의 상관을 보이고 있었다(〈표 5〉 참조). 하지만 인터넷 뉴스보다 종이

〈표 4〉 문항간 신뢰도 분석 결과*

	10점 척도	6점 척도	2점 척도
인터넷 뉴스 평가(객관적, 믿을 만, 공정)	.91	.80	.45
인터넷 뉴스(정치, 사건/범죄, 국제 뉴스)	.73	.77	.38
인터넷 뉴스 이용 동기: 편리성 추구	.69	.66	.28
인터넷 뉴스 이용 동기: 뉴스 전문성 추구	.84	.64	.57
인터넷 뉴스 이용 동기: 사건/정보 추구	.74	.70	.49

* 표에 제시한 수치는 크론바흐의 알파 값임.

〈표 5〉 척도 크기별 인터넷 뉴스 이용 동기, 평가, 이용정도간 상관분석 결과(Pearson's r)

	10점 척도 (n = 85)	6점 척도 (n = 87)	2점 척도 (n = 86)
이용 동기: 뉴스가 믿음이 가고 정확하기 때문에			
vs. 인터넷 뉴스는 객관적이다	.49 ***	.32 **	.44 ***
vs. 인터넷 뉴스는 믿을 만하다	.26 ***	.56 ***	.37 **
vs. 인터넷 뉴스는 공정하다	.60 ***	.54 ***	.39 ***
나는 인터넷 뉴스보다 종이신문을 더 많이 본다			
vs. 인터넷 뉴스로 보는 것이 편하기 때문에	-.08	-.24 *	-.49 ***
vs. 사-공간의 제약이 거의 없어서	.02	-.07	-.25 *
vs. 포털을 통해 모두 한 번에 해결할 수 있어서	.06	-.27 *	-.25 *

* p < .05, ** p < .01, *** p < .001

신문을 더 많이 본다는 문항과 인터넷 뉴스 이용의 편리성 추구 동기 사이의 상관은 2점 척도와 6점 척도 집단에서만 존재했다. 10점 척도를 사용한 경우에는 이들 문항간에 상관관계가 존재하지 않았다. 편리성 추구 동기보다는 “뉴스기사 내용이 충실하고 깊이있기 때문” 혹은 “일상 생활의 다양한 정보를 얻기 때문에”와 같은 여타 문항과의 상관관계가 더 높았다(모든 $r > .21$, 모든 $p < .05$)

3. 응답 편이성

실험자극으로 제시한 설문 페이지별, 그리고 설문 전체에 대한 문항 당 평균 반응시간을 <표 6>에 제시했다. 분산분석 결과 모든 페이지에 걸쳐($F = 9.26, p < .001, np^2 = .07$), 그리고 각 페이지별로(모든 $F > 3.66$, 모든 $p < .02$) 척도 크기가 클수록 반응 시간은 유의하게 증가했다. 사후검증(Scheffe) 결과 모든 페이지에 걸쳐 2점 척도와 10점 척도 집단이 유의한 차이를 보인 것으로 나타났다. 평균적으로는 6점 척도 집단과 10점 척도 집단 사이에도 유의한 차이가 존재했다.

피험자들이 주관적으로 인식한 응답 편이성, 즉 응답의 용이성과 짜증, 부담감 등에 척도 크기가 영향을 미치는지 분석한 결과 세 변인 모두에 대해 집단 간 차이가 존재하지 않았다(모든 $F < 1.60$). 즉 척도 크기와 무관하게 응답자들은 모두 설문에 대해 대체로 응답하기 쉬우며($M = 5.76, SD = 1.18$), 짜증이 거의 나지 않고($M = 2.00, SD = 1.24$), 부담감도 거의 느끼지 않는다고($M = 1.77, SD = .96$) 응답했다.

<표 6> 페이지별, 전체 문항 당 평균 반응시간

	10점 척도(n = 85)		6점 척도(n = 87)		2점 척도(n = 86)	
	M	SD	M	SD	M	SD
1페이지 (일반질문)	5.20 ^b	3.25	4.32 ^{ab}	2.19	3.82 ^a	4.17
2쪽(인터넷 뉴스 평가)	3.23 ^b	1.80	2.62 ^{ab}	2.09	2.33 ^a	1.58
3쪽(주제 분야별 이용)	2.99 ^b	4.24	2.19 ^{ab}	1.51	1.72 ^a	2.07
4쪽(인터넷 뉴스 이용 동기)	3.29 ^b	1.98	2.86 ^{ab}	1.49	2.46 ^a	1.22
평균	3.68 ^b	2.19	3.00 ^a	1.14	2.58 ^a	1.50

주) 단위: 초. 평균에 제시된 첨자는 사후검증 결과를 가리킴. 즉 같은 첨자가 제시된 집단 사이에는 평균 차이가 없음을, 다른 첨자가 붙은 집단 사이에는 평균 차이가 유의함을 의미한다.

V. 논의 및 결론

웹기반 조사의 활용 가능성에 대한 관심이 증가하는 가운데, 본고에서는 웹기반 조사 환경이 면접원이 굳이 필요하지 않으며 다점 척도 사용이 비교적 용이하다는 등의 측면에서는 지필 조사의 경우와 유사하지만, 피조사자의 집중도나 응답 성실도는 낮아지기 쉽다는 데 주목했다. 따라서 지필 조사의 문항 및 척도 구성 방법을 웹기반 조사에 그대로 적용하기 어려울 수 있을 것이라 가정하고 상이한 척도를 사용할 때 측정 문항의 타당도, 신뢰도 및 응답 편이성에 차이가 있는지 확인하고자 했다.

1. 문항의 타당도

요인분석 결과로 추정해 본 측정 문항의 타당도는 10점 척도를 사용한 집단에서 상대적으로 안정적인 결과를 보이고 있다. 인터넷 뉴스에 대한 평가는 공신력과 속도·선정성으로, 인터넷 뉴스 이용은 전통적 뉴스, 여성 선호형 뉴스, 남성 선호형 뉴스로, 인터넷 뉴스 이용 동기는 정보 추구, 편리성 추구, 전문성 추구로 요인 설정이 가능해 보인다. 즉 전반적으로 설명 가능한 요인이 추출되었다고 하겠다.

이에 비해 2점과 6점 척도를 사용한 경우에는 변인에 따라 결과가 차이를 보인다. 6점 척도 집단에서는 주제별 이용이, 2점 척도 집단에서는 뉴스 이용 동기가 다른 척도를 사용한 집단에 비해 일목요연한 결과를 보이지 않고 있다. 특히 2점 척도를 사용하는 경우 기존 연구에서 제시된 것과 같은 요인분석 결과를 기대하기 어렵다는 것을 알 수 있다.

이 결과를 토대로 판단한다면, 웹기반 조사에서도 다점 척도를 사용하는 것이 통계적 수준에서 타당도를 확보하는 데 유리하다는 잠정적인 결론을 내릴 수 있을 것이다. 하지만 이 결과는 어디까지나 요인분석을 통해 추정된 타당도에 국한되는 것임을 고려해야 한다. 요인분석은 측정에 사용한 문항들이 잠재적 변수와 얼마나 합치성을 보이는가를 판단하기 위한 것으로, 조사에 사용한 문항들을 어떠한 구인들(components)로 분류 가능한지만을 보여준다. 따라서 타당도에 대한 최종적인 판단은 연구자들의 이론적 판단에 달려 있다고 보아야 할 것이다.

이와 관련해, 비교적 오랜 시간에 걸쳐 검증되어 온 인터넷 뉴스 평가 관련 문항들은 2, 6, 10점 척도 집단 모두에서 대체로 비슷한 요인분석 결과를 보이고 있다는 점에 주목

할 필요가 있다. 척도 크기와 관계없이 인터넷 뉴스에 대한 평가에 대한 분석에서는 뉴스 공신력 관련 요인(객관성, 신뢰성, 공정성 등)과 선정성 관련 요인(신속성, 선정성 등)이 추출되었다. 이에 비해 인터넷 뉴스의 주제별 이용 정도나 인터넷 뉴스 이용 동기는 비교적 최근에 개발된 문항들로, 타당성이 충분히 검증되지 않은 문항들이기 때문에 척도 크기에 민감하게 반응했을 가능성이 있다.

그렇다면 이 결과는 이 연구에서 사용한 문항들의 내적 타당도가 반영된 결과일 수 있으며, 이 경우 웹기반 조사에서 어떤 크기의 척도를 사용하는가 하는 것은 응답의 타당도에 큰 영향을 미치지 않는다고도 볼 수 있다. 따라서 웹기반 조사에서 척도 크기에 따른 응답의 타당도와 관련해서는, 이론적 수준에서 타당화가 충분히 이루어진 문항들을 이용해 재검토를 실시할 필요가 있을 것으로 보인다.

2. 문항간 신뢰도

크론바흐의 알파 계수를 사용해 비교한 문항간 신뢰도는 척도 크기와 비례해 증가하고 있어, 다점 척도가 응답의 신뢰도를 높인다는 선행 연구들과 유사한 결과가 도출되었다(Cicchetti et al. 1985 참조). 하지만, 이는 신뢰도를 산출하는 산술적 논리에도 관련성이 큰 것으로 보인다. 크론바흐 알파 계수는 문항의 수와 개별 문항 및 전체 관찰값의 분산에 따라 결정되는 값으로, 동일한 문항을 사용한 경우 개별 문항의 분산은 크고 전체 응답값 사이의 분산은 작을수록 증가하게 된다. 그런데 척도 크기가 작을수록 개별 문항의 분산이 감소하고 응답의 변동(fluctuation)은 크므로 전체 응답의 분산이 증가해 결과적으로 신뢰도가 낮아지기 쉽다. 따라서 이 결과에 대해서도 좀더 신중하게 해석할 필요가 있다.

추가로 실시한 상관분석 결과를 보면 척도 크기가 작다고 해서 변인간 관계의 일관성이 떨어지는 것은 아님을 알 수 있다. ‘뉴스가 믿음이 가고 정확하기 때문에’ 인터넷 뉴스를 사용한다는 문항에 대한 동의 정도와 인터넷 뉴스의 객관성, 신뢰성, 공정성에 대한 응답 간의 상관은 모든 척도에서 매우 유의한 수준인 것으로 나타났다. 이는 이론적으로 타당한 변인들간의 관계를 설명하는 데 2점 척도도 충분히 유용한 역할을 할 수 있음을 보여주는 결과라 하겠다. 즉, 웹기반 조사에서도 전통적인 조사의 경우와 마찬가지로 통계적인 차원에서의 신뢰도를 높이는 데 다점 척도가 더 유리하지만, 상호 연관성이 높은 변인들에 대한 분석에는 2점 척도 역시 사용 가능하리라고 판단된다.

3. 주관적, 객관적 응답 편이성

척도 크기는 피조사자들이 주관적으로 느끼는 응답 편이성에 영향을 미치지 않았다. 어떤 척도를 사용하건 피조사자들이 설문을 진행하면서 느끼는 짜증이나 부담감에 차이가 없다면, 웹기반 조사에서는 연구자들이 연구의 목적에 따라 비교적 자유롭게 척도 크기를 선택할 여지가 있다고 하겠다.

하지만 응답 편이성의 객관적 지표라고 할 수 있는 반응 시간은 척도 크기에 따라 차이를 보였다. 즉, 다점 척도를 사용한 경우 응답에 소요된 시간이 더 길었다. 전반적으로 본 조사 참여자들의 반응 시간은 문항당 평균 3초 정도에 불과해, 이 결과를 응답에 대한 피로도나 부담감을 짐작케 하는 결과로 해석하기는 어렵다. 하지만 (1)웹기반 조사 참여자들이 설문을 매우 빠른 속도로 진행한다든 점, (2)이항 척도를 사용하는 경우 응답 속도가 더욱 빨라진다는 점은 추후 웹기반 조사 설계 시 충분히 고려해야 할 사항이다.

면접 조사나 전화 조사를 실시하는 경우, 면접원의 존재로 인해 응답자들은 조사에 어느 정도 집중하게 된다. 하지만 웹기반 조사에서는 면접 또는 전화 조사의 경우와 같은 수준의 집중도를 기대하기 어렵다. 피조사자는 응답 도중에 다른 일을 하거나 컴퓨터에 실행 중인 다른 프로그램을 이용하고 또는 다른 웹사이트를 방문할 수 있다. 이처럼 방해요인이 더 많음에도 불구하고 평균 응답 시간이 문항당 3초라는 것은 우려할 만한 결과이다.

특히 2점 척도를 사용한 경우에는 반응 시간이 더 짧아져, 10점 척도를 사용한 경우보다 평균 1초 이상 차이를 보였다. 따라서 웹기반 조사에서는 조사 내용에 따라 척도 크기를 신중하게 선택할 필요가 있을 것으로 보인다. 질문에 대해 얼마간이라도 숙고한 후의 응답이 필요한 경우라면 다점 척도를 사용하는 것이 유리할 것이다. 반면 사람들의 즉각적, 자동적인 반응을 측정할 필요가 있는 연구라면 이항 척도를 사용해 웹기반 조사를 실시하는 것이 매우 유용할 것이다.

4. 결론

이상의 결과들을 종합하면 다음과 같은 시사점을 도출할 수 있다. 첫째, 측정 도구의 타당도와 신뢰도는 무엇보다 문항 구성의 토대가 되는 이론적, 내적 타당성과 관련성이

높다고 볼 수 있다. 비교적 오랜 기간 많은 연구를 통해 검증되어 온 척도인 (인터넷) 뉴스에 대한 평가 측정 문항에서, 요인분석 및 상관분석 결과가 척도 크기와 관계없이 비교적 일관된 것을 볼 때 그러하다. 따라서 내적 타당도를 확보한 문항을 사용한다면 굳이 다점 척도를 사용하지 않더라도 충분히 신뢰할 만한 결과를 얻을 수 있을 것이다.

둘째, 피조사자들이 느끼는 주관적 응답 편이성은 척도에 따라 차이를 보이지 않지만, 전반적으로 문항에 대한 반응 시간이 매우 짧으며, 반응 시간은 척도 크기와 비례했다. 면접원과 통화를 하는 경우는 모든 응답자들이 동일한 시간 동안 응답하게 된다. 하지만 웹기반 조사에서는 이 점과 관련해 큰 개인차를 보일 수 있을 뿐 아니라, 전반적으로 응답 시간이 짧다는 사실은 우려할 만한 점이다. 특히 최근 들어 자택이나 직장 등에서 직접 설문조사 사이트에 접속해 진행되는 형태의 조사가 증가하고 있는데, 이 경우 본 조사에서와 같이 피조사자들이 매우 빠른 속도로 응답할 가능성이 높다. 이 점을 생각할 때, 웹기반 조사에서 연구 내용과 목적에 따라 어떤 크기의 척도를 사용할 것인지를 주의 깊게 선택할 필요가 있다고 하겠다.

다만 이 실험의 결과를 해석하는 데 있어 다음의 몇 가지 사항을 고려해야 할 것이다. 첫째, 피조사자들이 모두 대학생으로 비교적 동질한 집단이며 연령과 성별에 있어 집단 간 차이가 없다는 점을 확인하였지만, 그 외에 평소 인터넷 이용 시간, 인터넷 이용 능숙도 등 웹기반 조사 응답에 영향을 미칠 수 있는 요인에 대해서도 확인을 했더라면 좋았을 것이다. 둘째, 본 연구의 피조사자들이 모두 대학생이므로 인터넷 이용 시간이 전반적으로 길고 인터넷 이용에도 별 어려움이 없는 계층일 것으로 예상된다. 그렇다면 중·장년층 및 고령자에게서는 다른 양상의 응답 편향이 나타날 가능성이 있다.

본 연구는 웹기반 조사에 한정해 척도 크기에 따른 응답을 비교한 것으로, 추가적으로 보다 면밀한 후속 연구가 필요하다. 예컨대 일반적으로 2점 척도 사용 시 신뢰도 값이 매우 낮다면, 이러한 척도를 사용할 때 문항간 신뢰도를 추정할 수 있는 보완적인 방법에 어떤 것이 있는지, 그리고 웹기반 조사에서 전반적으로 문항에 대한 평균 반응 시간이 매우 짧은 점을 고려할 때, 어느 정도 속도가 필요한 사안에 대한 설문에 과연 웹기반 조사가 적합하다고 할 수 있는지 등에 대한 탐구가 필요할 것으로 보인다.

참고 문헌

- 강남준·백영민. 2005. “대안적 여론조사의 표본편파 문제점과 가중치를 사용한 보정 방법.” 《언론정보연구》 41(2): 43-78.
- 김은미·선유화. 2006. “댓글에 대한 노출이 뉴스 수용에 미치는 효과.” 《한국언론학보》 50(4): 33-64.
- 박상호. 2009. “포털뉴스 이용동기가 인터넷 자기효능감, 정치적 신뢰, 정치냉소주의와 정치참여에 미치는 영향에 관한 연구.” 《한국언론학보》 53(5): 153-175.
- 이계오·장덕현. 2009. “인터넷 선거조사에서 성향가중모형 적용 사례.” 《조사연구》 10(3): 51-83.
- 이윤석·이지영·이경택. 2008. “온라인조사의 응답오차에 대한 연구.” 《조사연구》 9(2): 51-83.
- 이은주·장윤재. 2009. “인터넷 뉴스 댓글이 여론 및 기사의 사회적 영향력에 대한 지각과 수용자의 의견에 미치는 효과.” 《한국언론학보》 53(4): 50-71.
- 허명희·조성겸. 2010. “자원자 패널에 의한 인터넷 조사의 성향조정 가중화.” 《조사연구》 11(2): 1-28.
- 허명희·황진모. 2006. “전화조사를 위한 시간균형할당표본추출.” 《조사연구》 7(2): 39-52.
- Andrews, F.M. 1984. “Construct Validity and Error Components of Survey Measures: A Structural Modeling Approach.” *Public Opinion Quarterly* 52(1): 53-83.
- Cicchetti, D.V., D. Showalter, and P.J. Tyrer. 1985. “The Effect of Number of Rating Scale Categories on Levels of Interrater Reliability: A Monte Carlo Investigation.” *Applied Psychological Measurement* 9: 31-36.
- Cohen, J. 1983. “The Cost of Dichotomization.” *Applied Psychological Measurement* 7: 249-253.
- Converse, J.M. and S. Presser. 1986. *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills, CA: Sage.
- Fazio, R.H. and C.J. Williams. 1986. “Attitude Accessibility as a Moderator of the Attitude-Perception and Attitude-Behavior Relations: An Investigation of the 1984 Presidential Election.” *Journal of Personality and Social Psychology* 51(3): 505-514.
- Fiske, S.T. 1992. “Thinking is for Doing: Portraits of Social Cognition from Daguerreotype to Laserphoto.” *Journal of Personality and Social Psychology* 63(6): 877-889.

- Flanagin, A.J. and M.J. Metzger. 2003. "The Perceived Credibility of Personal Web Page Information as Influenced by the Sex of the Source." *Computers in Human Behavior* 19(6): 683-701.
- Fleming, J.S. 2005. "TETCORR: A Computer Program to Compute Smoothes Tetrachoric Correlation Matrices." *Behavior Research Methods* 37: 59.
- Fowler Jr., F.J. 1995. *Improving Survey Questions: Design and Evaluation*. Thousand Oaks, CA: Sage.
- Garson, G.D. 2011. *Factor Analysis*. Retrieved November 3, 2011, from <http://faculty.chass.ncsu.edu/garson/PA765/factor.htm>
- Green, P.E. and V.R. Rao. 1970. "Rating Scales and Information Recovery: How Many Scales and Response Categories to Use?" *The Journal of Marketing* 34: 33-39.
- Jacoby, J. and M.S. Martell. 1971. "Three-Point Likert Scales are Good Enough." *Journal of Marketing Research* 8: 495-500.
- McCroskey, J.C. and J.J. Teven. 1999. "Goodwill: A Reexamination of the Construct and its Measurement." *Communication Monographs* 66(1): 90-103.
- Oskamp, S., and P.W. Schultz. 2005. *Attitudes and Opinions*(3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Stober, J. and D.E. Dette. 2002. "Comparing Continuous and Dichotomous Scoring of the Balanced Inventory of Desirable Responding." *Journal of Personality Assessment*, 78, 370-389.
- Tversky, A. and D. Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185(4157): 1124-1131.
- Wilson, T.D. and S. Hodges. 1992. "Attitudes as Temporary Constructions." In L. Martin and A. Tesser (eds.), *The Construction of Social Judgments*(pp. 37-66). New York: Springer-Verlag.
- Tourangeau, R., L.J. Rips, and K.R. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press.
- Tourangeau, R., K. Rasinski, J.B. Jobe, T.W. Smith, and W. Pratt. 1997. "Sources of Error in a Survey of Sexual Behavior." *Journal of Official Statistics* 13: 341-365.