

Causes of Careless Responding in Surveys and Preventions in the Survey Design*

설문조사에서 불성실 응답의 원인과 조사설계 단계에서의 예방

Won-Woo Park(First Author)

Professor, College of Business Administration, Seoul National University
(*wwwpark@snu.ac.kr*)

Yoowoo Lee(Corresponding Author)

Seoul National University
(*youfmc@snu.ac.kr*)

Sunghyuck Mah(Co-Author)

Seoul National University
(*ryanshmah@snu.ac.kr*)

Jayoung Kim(Co-Author)

Seoul National University
(*marsb612@snu.ac.kr*)

Suhyun Bae(Co-Author)

Seoul National University
(*top3043@snu.ac.kr*)

Seonyoung Ji(Co-Author)

Seoul National University
(*ji179@purdue.edu*)

.....

This study investigates the level of Korean researchers' awareness of careless responding (CR) in surveys, the causes of CR, as well as the potential remedies available with regards to survey design. We first reviewed 827 survey studies published in seven major journals in the field of business administration and analyzed the level of recognition of Korean researchers regarding CR. We also investigated the causes of CR through a literature review that particularly focused on the motives and burdens of respondents. We then conducted in-depth interviews with academic and practical experts to investigate the perceptions and practices regarding CR. Finally, based on the literature review and the in-depth interviews, we drew up measures to prevent CR. We suggest that there need to be academic standards that prompt researchers to report specific CR prevention or detection methods used for their research in a more detailed manner. This study, as a cornerstone, contributes to further discussions and the development of preventive measures of CR in the Korean scholarly community, particularly in the field of business administration.

Key Words: careless responding, causes of careless responding, prevention of careless responding, survey design

.....

Submission Date: 02. 26. 2021

Revised Date: (1st: 06. 18. 2021)

Accepted Date: 07. 05. 2021

* This research is supported by the Institute of Industrial Relations of Seoul National University Research Fund (in 2021).

1. 서론

설문조사는 경영학, 심리학, 사회학 등 사회과학 제반 분야에서 폭넓게 사용되는 자료수집 방법이다. 이에, 연구자들은 방법론 측면에서 설문조사의 오차를 줄이고 연구의 내적 타당성을 높이고자 다양한 학술적 논의를 전개해왔다(Deutskens et al., 2004; Galesic and Bosnjak, 2009; Gideon, 2012; Groves, 1989; Park and Lee, 2017). 특히, 대규모 설문조사를 수행하는 사회학 또는 인구통계학자들은 별도의 학술지(예, Sociological Methods and Research, The American Sociologist)와 연구분야(예, Survey Methodology)를 형성할 정도로 설문조사 방법에 대해 오랜 시간 연구해왔다(Cantor et al., 2008; Görtitz, 2004; Saris and Gallhofer, 2014). 그러나 해당 분야의 연구들은 주로 응답의 양적 측면, 즉, 설문 대상자 대비 실제 참여자의 비율을 의미하는 ‘응답률(response rate)’을 제고하기 위한 설문조사의 설계에 치중해왔다(Gideon, 2012; Singer and Ye, 2013). 한편, 응답의 질적 측면, 즉 응답자들이 연구자의 의도에 맞게 응답하였는가에 대한 문제와 응답의 질(response quality)을 높이기 위한 방안에 대한 연구는 측정 심리학을 중심으로 비교적 제한적으로 수행되어왔다. 그러나 SAP사의 쉐트릭스(Qualtrics), 아마존(Amazon)사의 M-Turk과 같은 온라인 설문조사 관련 도구와 서비스의 활용이 용이해지고 인터넷을 활용한 온라인 설문이 보편화 되면서, 최근 학계에서도 응답의 질에 대한 관심이 높아지고 있다. 특히, 응답자가 설문 내용 충분히 주의를 기울이지 않고 부주의하게 응답(careless responding)하거나 의도적으로 무작위하게 응답(random responding)

하는 불성실 응답의 영향과 탐지에 대한 연구(예, Meade and Craig, 2012)가 활발하게 진행되고 있다.

설문조사의 질적 측면에서, 불성실 응답은 오차분산(error variance), 변수 간 상관관계(correlations), 내적 일치도 계수(internal consistency validity coefficient), 요인 적재치(factor loading)와 같은 다양한 통계적 수치에 영향을 미친다(Curran, 2016; DeSimone et al., 2018; Meade and Craig, 2012; Ward and Meade, 2018). 따라서 불성실 응답에 따른 응답 오차가 커질수록, 많은 비용과 노력을 통해 얻은 연구 데이터의 신뢰성과 타당성이 떨어질 수 있으므로 연구자들의 주의가 요구된다(Park et al., 2020). 그러나 이러한 중요성에도 불구하고 아직까지 국내 학술문헌에 있어 불성실 응답에 대한 논의는 활발하지 않은 실정이다. 예외적으로, Park et al.(2020)은 불성실 응답의 개념과 유형을 정리하고, 이를 탐지할 수 있는 다양한 통계적 방법들(예, 가짜문항, 마할라노비스 거리 등)을 소개하였다. 다만, 이는 불성실 응답이 발생한 후에 대처할 수 있는 사후적 조치에 해당하며, Park et al.(2020)은 사전적으로 불성실 응답을 예방하는 방안에 대해서는 후속 연구가 필요함을 촉구하였다.

본 연구는 이러한 요구에 부응하여, 불성실 응답을 예방하는 사전적 조치를 다루고자 한다. 예컨대, Park et al.(2020)의 선행연구가 불성실 응답에 대한 ‘치료제’를 다루었다면, 본 연구에서는 불성실 응답에 대한 ‘백신’을 다룬다고 설명할 수 있다. 백신 개발의 시작점이 진단 및 발생원인 파악에 있듯이, 본 연구 또한 불성실 응답에 관한 현재의 인식수준 및 발생 원인을 고찰하는 데에서 논의를 시작하고자 한다. 구체적으로는 (1) 최근 10년 이내에 경영학 관련 주요 일곱 개 학술지에 등재된 설문조사 연구에

서의 불성실 응답 인식 및 대처 수준을 분석하고, (2) 문헌조사를 통해 불성실 응답의 기저에 있는 발생 원인에 대하여 고찰하고자 한다. 나아가, (3) 학계 및 실무 전문가들을 심층 인터뷰함으로써 불성실 응답에 대한 인식과 예방방안을 조사하며, (4) 선행연구와 인터뷰를 기반으로 불성실 응답 예방방안을 도출할 것이다. 본 연구는 이처럼 정량적 실태분석, 문헌조사 및 인터뷰를 통합한 접근을 통하여 불성실 응답이라는 주제를 보다 심도 있게 살펴며, 이에 대한 지속적인 관심과 논의를 촉구하고, 향후 설문조사를 수행하는 연구자들이 실제로 예방방안을 참고할 수 있도록 도움으로써 학계의 실질적인 변화와 발전을 이끌어내고자 한다.

II. 불성실 응답에 대한 국내연구의 인식 및 대처 현황

불성실 응답이란 설문조사에서 응답자가 설문내용에 충분히 주의를 기울이지 않고 부주의하게 응답하거나 의도적으로 무작위하게 응답하는 것을 의미한다(Beach, 1989; Park et al., 2020; Schmitt and Stults, 1985; Ward and Pond, 2015). 초기연구들은 응답자가 주의를 전혀 기울이지 않은 무작위 응답만을 불성실 응답으로 정의하였으나(Beach, 1989; Berry et al., 1992), 근래의 연구들은 충분한 주의를 기울이지 않은 부주의 응답까지 포함하여 불성실 응답을 정의하고 있다(Ward and Pond, 2015). 이러한 추세에 따라 본 연구에서는 불성실 응답을 의도적으로 무작위 응답을 하는 경우(Meade and Craig, 2012), 부주의하거나 응답 시 충분한 노력을 기울이지 않는 경우(Huang, Bowling et

al., 2015), 그리고 응답의 동기가 낮거나 없어서 진정성 없이 응답한 경우(Huang et al., 2012, Ward and Meade, 2018)까지를 포괄하도록 정의한다. 본 연구는 상기의 개념적 정의에 바탕하여, 2011년에서 2019년 사이에 경영학 분야 학술지에 게재된 논문들을 분석하여 국내 연구자들의 불성실 응답에 대한 인식 및 대처현황을 살펴본다.

2.1 불성실 응답 인식수준

불성실 응답에 대한 국내의 인식수준과 해결을 위한 노력을 살펴보기 위해, 경영학 전반을 다루고 있는 학술지인 *경영학연구*, *인사·조직 분야의 조직과 인사관리연구*, *인사조직연구*, 그리고 마케팅 분야의 *마케팅연구*, *마케팅관리연구*와 국제/전략 분야의 *전략경영연구*, *국제경영연구*로 범위를 한정하여 최근 약 10년(2011~2019년) 이내에 게재된 논문들을 전수조사하였다. 해당 학술 분야를 선택한 이유는, 해당 분야가 경영학의 제 분야 중 설문조사의 활용의 빈도가 높아 불성실 응답 인식 실태의 비율과 추이를 보다 명확하게 확인할 수 있기 때문이다. 전수조사를 통해 (1) 질적연구를 제외한 정량연구를 분류하고, (2) 그 중에서 설문연구의 비중이 얼마나 높은지 확인하였으며, (3) 이러한 설문연구 중 불성실 응답에 대한 인식이나 데이터 처리에 대한 언급 여부를 조사하여 3단계로 분류하였다. 불성실 응답을 인식하였다고 판단되어 마지막 카테고리 분류된 연구의 경우, 그 기준과 예시를 다음 장에 상술하였다. 개괄적 결과는 <Table 1>에 제시된 바와 같다. 먼저, 경영학 전반의 분야를 다루고 있는 *경영학연구*의 경우, 2011년부터 2019년까지 총 565편의 논문이 게재되었다. 정량연구에 해당하는 연구는 524편(92.7%)이었고, 이 중에서 212편(40.5%)이 설문

〈Table 1〉 불성실 응답에 대한 인식수준

단위: 빈도, (%)

학술지	구분	2011	2012	2013	2014	2015	2016	2017	2018	2019	합계
경영학 연구	정량연구	51	53	61	75	59	63	55	52	55	524
	설문연구	19	25	18	31	27	29	23	18	22	212
	불성실 응답 언급 연구	7 (36.8)	11 (44.0)	11 (61.1)	12 (38.7)	13 (48.1)	12 (41.4)	8 (34.8)	7 (38.9)	10 (45.5)	91 (42.9)
조직과 인사관리 연구	정량연구	28	29	28	29	22	29	20	22	23	230
	설문연구	18	24	20	24	18	25	13	15	18	175
	불성실 응답 언급 연구	8 (44.4)	12 (50.0)	10 (50.0)	13 (54.2)	10 (55.6)	13 (52.0)	8 (61.5)	6 (40.0)	9 (50.0)	89 (50.9)
인사조직 연구	정량연구	17	15	20	14	11	15	11	15	12	130
	설문연구	13	15	13	12	9	12	7	13	10	104
	불성실 응답 언급 연구	6 (46.2)	6 (40.0)	4 (30.8)	7 (58.3)	4 (44.4)	4 (33.3)	3 (42.9)	8 (61.5)	5 (50.0)	47 (45.2)
마케팅 연구	정량연구	21	31	49	31	15	19	19	17	19	221
	설문연구	8	16	23	12	6	7	3	7	8	90
	불성실 응답 언급 연구	6 (75.0)	8 (50.0)	12 (52.2)	6 (50.0)	3 (50.0)	2 (28.6)	2 (66.7)	4 (57.1)	3 (37.5)	46 (51.1)
마케팅 관리연구	정량연구	24	22	29	30	23	19	21	21	19	208
	설문연구	17	18	22	22	17	15	17	13	15	156
	불성실 응답 언급 연구	10 (58.8)	12 (66.7)	13 (59.1)	15 (68.2)	12 (70.6)	12 (80.0)	11 (64.7)	8 (61.5)	7 (46.7)	100 (64.1)
국제경영 연구	정량연구	15	16	14	21	20	15	17	15	18	151
	설문연구	6	9	2	11	15	9	5	7	9	73
	불성실 응답 언급 연구	1 (16.7)	1 (11.1)	0 (0.0)	0 (0.0)	4 (26.7)	2 (22.2)	1 (20.0)	2 (28.6)	2 (22.2)	13 (17.8)
전략경영 연구	정량연구	15	8	10	16	11	12	9	12	9	102
	설문연구	1	1	4	5	1	1	1	2	1	17
	불성실 응답 언급 연구	0 (0.0)	0 (0.0)	2 (50.0)	3 (60.0)	0 (0.0)	0 (0.0)	1 (100.0)	0 (0.0)	1 (100.0)	7 (41.2)
합계	정량연구	171	174	211	216	161	172	152	154	155	1,566
	설문연구	82	108	102	117	93	98	69	75	83	827
	불성실 응답 언급 연구	38 (46.3)	50 (46.3)	52 (51.0)	56 (47.9)	46 (49.5)	45 (45.9)	34 (49.3)	35 (46.7)	37 (44.6)	393 (47.5)

주1) 불성실 응답 언급 연구의 백분율(%)은 해당 연도 설문연구 중 불성실 응답 언급 연구의 비율을 나타냄.

주2) 표의 설문연구 빈도는 설문조사 연구 외에도 설문조사를 활용한 실험연구를 포함한 수치임.

연구 방법을 채택하였다. 설문연구 중 연구 과정에서 발생한 불성실 응답에 대해 간략하게라도 언급한 연구는 총 91편으로 설문연구의 42.9%에 해당하였다.

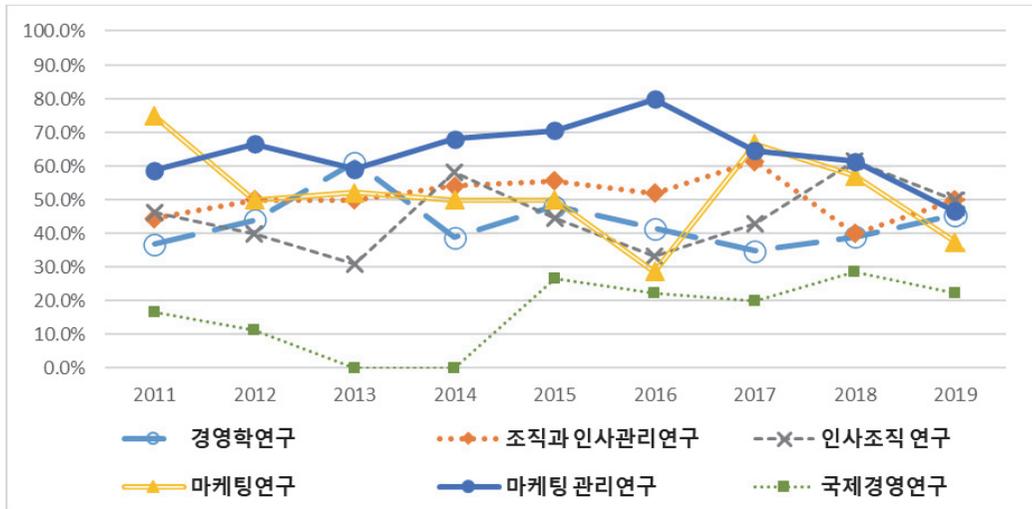
인사·조직 분야 학술지 중 조직과 인사관리연구는 2011년부터 2019년까지 총 245편의 논문이 게재되었다. 정량연구에 해당하는 연구는 총 230편(93.9%)이었고, 이들 정량연구의 76.1%(175편)가 설문연구 방법을 사용하였다. 설문연구 중 불성실 응답에 대해 언급한 연구는 총 89편으로 설문연구의 50.9%에 해당하였다. 한편, 인사조직연구의 경우, 최근 약 10년간 게재된 논문 총 225편 중 130편(57.8%)이 정량연구였고, 이들 정량연구의 80%(104편)가 설문연구 방법을 사용하였다. 설문연구 중 불성실 응답에 대해 언급한 연구는 총 47편으로 설문연구의 45.2%에 해당하였다. 평균적으로 인사조직연구는 조직과 인사관리연구 대비 설문연구의 절대적 수가 적었지만, 불성실 응답 언급 연구의 비율은 비슷하게 나타났다.

마케팅 분야 학술지 중 마케팅연구는 2011년부터 2019년까지 총 240편의 논문 중 정량연구의 수가 총 221편(92.1%)이었고, 이들 정량연구의 40.7%(90편)가 설문연구 방법을 사용하였다. 설문연구 중 불성실 응답에 대해 언급한 연구의 수는 총 46편으로 설문연구의 51.1%에 해당하였다. 연간 추세를 살펴보면, 타 학술지에 비해 설문연구의 수가 적음에도 불구하고 불성실 응답 언급 연구의 비율이 마케팅관리연구에 다음가는 수준으로 높게 나타났다. 마케팅관리연구의 경우, 최근 약 10년간 게재된 총 215편의 논문 중 정량연구가 총 208편(96.7%)이었고, 이들 정량연구의 75.0%(156편)가 설문연구 방법을 사용하였다. 설문연구 중 불성실 응답에 대해 언급한 연구의 수는 총 100편으로 설문연구의 64.1%를 차지하였다. 연도별 불성실 응답 인식현황

을 비교한 결과, 마케팅연구보다 마케팅관리연구에서 대체적으로 불성실 응답을 언급한 연구의 수와 비율이 높았으며, 마케팅관리연구에서 설문연구 중 불성실 응답을 언급한 연구의 비율은 7개 학술지 중에서도 가장 높은 수준에 해당하였다.

국제/경영전략 분야 학술지 중 국제경영연구는 2011년부터 2019년까지 총 161편의 논문 중 정량연구의 수가 총 151편(93.8%)이었고, 이들 정량연구의 48.34%(73편)가 설문연구 방법을 사용하였다. 설문연구 중 불성실 응답에 대해 언급한 연구의 수는 총 13편으로 설문연구의 17.8%에 해당하였다. 비록, 국제경영연구의 설문연구 비율(48.34%)은 마케팅연구의 설문연구 비율(37.5%)보다 높았지만, 불성실 응답을 언급한 비율은 다른 분야의 학술지에 비해 낮은 것으로 나타났다. 전략경영연구의 경우, 2011년부터 2019년까지 게재된 총 129편의 논문 중 정량연구의 수가 총 102편(79.1%)이었고, 이 중 설문연구를 활용한 연구의 비중은 16.7%(17편)로 다른 학술지에 비해 상당히 낮았다. 이 중 불성실 응답에 대해 언급한 연구는 총 7편으로 설문연구의 41.2%를 차지하였다. 전략경영연구의 경우, 설문연구 비율이 확연하게 낮으며 불성실 응답 언급 논문도 연간 4편 이내에 그쳤기에, 불성실 응답 언급 비율의 해석에 있어 주의가 필요할 것이다.

위 결과를 종합하면, 2011년~2019년 조사기간 동안 7개 주요 학술지에 발간된 총 1,566편의 논문 중 827편(52.8%)이 정량연구를 택하였으며, 이들 중 393편(47.5%)이 불성실 응답에 대한 최소한의 언급이나 데이터 처리 과정 기술을 포함하였다. 절대 숫자가 지나치게 작아 추세를 왜곡하여 보여줄 가능성이 있는 전략경영연구를 제외하고, 여섯 개 학술지의 평균적인 불성실 응답 언급 비율을 <Figure 1>에 제시하였다. 이 때, 평균 불성실 응답 언급 비율



〈Figure 1〉 설문조사를 활용한 연구 중 불성실 응답을 언급한 논문 비율

은 거의 절반에 가까운 수준으로 나타났는데, 이는 서론에서 언급한 불성실 응답에 대한 명시적인 관심 부족과 연구의 부재에 비해 비교적 높은 수치로 간주할 수 있을 것이다. 그러나 각 문헌에서 얼마나 상세하고 명확하게 불성실 응답을 인식 및 언급하고 있는지는 보다 상세히 살펴볼 필요가 있다. 또한, 10여 년간 불성실 응답 언급 비율이 우상향하고 있지 않고 정체되어있다는 점을 고려할 때, 해당 이슈에 대한 인식수준이 크게 변화하지 않았다고 볼 수 있다.

2.2 불성실 응답 인식 및 대처유형

위의 결과를 보다 명확하게 이해하고자, 불성실 응답 관련 언급 유형을 다음 〈Table 2〉와 같이 네 가지 유형으로 분류하였다. 첫째, 무언급형은 불성실 응답이 데이터에 포함되었을 것으로 추측되지만, 불성실 응답이나 데이터 클리닝과 관련한 과정에 대한 언급을 제시하지 않은 경우이다. 둘째, 데이터 클리닝 언급형은 ‘불성실 응답’이라는 단어를 사용하지

는 않았지만 여러 가지 방식으로 데이터 클리닝을 하고 이를 본문에 제시한 경우이다. 셋째, 불성실 응답 언급형은 ‘불성실 응답’의 키워드를 직접적으로 본문에 제시한 경우이다. 넷째, 불성실 응답 예방 및 탐지방법 언급형은 불성실한 응답의 예방 및 탐지 기준을 본문에 구체적으로 제시하고 처치한 경우이다. 셋째와 넷째 분류에 해당하는 경우, 위의 〈Table 1〉에서 ‘불성실 응답 언급 연구’로 분류하였다.

〈Table 3〉은 최근 약 10년간 각 학술지에 게재된 설문연구들을 위에 제시한 기준에 따라 네 가지 유형으로 분류하여 제시하였다. 일곱 개 학술지의 설문연구(총 827편)를 네 가지 유형으로 분류한 결과에 따르면, 불성실 응답 언급형이 370편(44.7%)으로 가장 많았으며, 다음으로는 불성실 응답 관련 무언급형이 237편(28.7%), 데이터 클리닝 언급형이 197편(23.8%), 마지막으로 불성실 응답 예방 및 탐지방법 언급형이 23편(2.8%)으로 조사되었다. 또한, 학술지별로 불성실 응답 인식 유형을 살펴본 결과, 경영학연구, 조직과 인사관리연구, 인사조직연구,

<Table 2> 불성실 응답에 대한 인식 및 대처유형

유형	내용
I. 무언급형	불성실 응답이나 데이터 클리닝 관련 언급을 본문에 제시하지 않음
	예시: '항목에 누락이 없었다.', '수거된 모든 설문지가 유효하였다.', '무응답의 오류가 없었다.' 등
II. 데이터 클리닝 언급형	'불성실'에 관한 언급은 없지만 여러 가지 방식으로 데이터 클리닝을 하고 본문에 제시
	예시: '이상치를 보인', '부정확한 응답', '중심화 경향이 강한', '결측치가 있는 설문지를 제거한', '응답이 불충분한', '부적절하다고 판단되는', '신뢰성이 현저히 떨어지는' 등
III. 불성실 응답 언급형	'불성실'에 관한 언급을 본문에 제시
	예시: '불성실하게 응답한', '불성실한 응답', '응답이 불성실한'
IV. 불성실 응답 예방 및 탐지방법 언급형	불성실한 응답의 예방 및 탐지기준을 본문에 구체적으로 제시
	예시: 경고성 지시문의 활용, 지시적 조작 점검 문항을 활용하여 불성실 응답이 의심되는 응답을 제외한 표본을 분석에 사용

<Table 3> 경영학 관련 주요 학술지에 나타난 불성실 응답에 대한 인식 및 대처유형

단위: 빈도, (%)

학술지 \ 유형	I. 무언급형	II. 데이터 클리닝 언급형	불성실 응답 언급 연구			설문연구 총계 (I + II + III + IV)
			III. 불성실 응답 언급형	IV. 불성실 응답 예방 및 탐지방법 언급형	소계 (III + IV)	
경영학연구	63 (29.7)	58 (27.4)	78 (36.8)	13 (6.1)	91 (42.9)	212 (100.0)
조직과 인사관리연구	46 (26.3)	40 (22.9)	89 (50.9)	0 (0.0)	89 (50.9)	175 (100.0)
인사조직연구	38 (36.5)	19 (18.3)	47 (45.2)	0 (0.0)	47 (45.2)	104 (100.0)
마케팅연구	28 (31.1)	16 (17.8)	46 (51.1)	0 (0.0)	46 (51.1)	90 (100.0)
마케팅 관리연구	21 (13.5)	35 (22.4)	91 (58.3)	9 (5.8)	100 (64.1)	156 (100.0)
국제경영연구	33 (45.2)	27 (37.0)	12 (16.4)	1 (1.4)	13 (17.8)	73 (100.0)
전략경영연구	8 (47.1)	2 (11.8)	7 (41.2)	0 (0.0)	7 (41.2)	17 (100.0)
계	237 (28.7)	197 (23.8)	370 (44.7)	23 (2.8)	393 (47.5)	827 (100.0)

주1) 최근 약 10년간(2011~2019년) 일곱 개 학술지에서의 설문연구 및 설문조사를 활용한 실험연구를 대상으로 함

주2) 백분율(%)은 해당 학술지 내 설문조사를 활용한 연구 수 대비 각 유형별 연구 수의 비율임

주3) 표의 설문연구 총계는 설문연구와 설문조사를 활용한 실험연구를 포함한 수치임

마케팅연구, 마케팅관리연구의 다섯 개 학술지에서 불성실 응답 언급형이 가장 큰 비중을 차지하였다. 국제경영연구와 전략경영연구는 무언급형이 가장 많은 비중을 차지하였다. 그러나 불성실 응답 예방 및 탐지방법 언급형은 조직과 인사관리연구, 인사조직연구, 마케팅연구에서는 찾아볼 수 없었으며, 다만 경영학 전반에 관해 다루는 경영학연구에서 6.1% (13편)로 다른 학술지에 비해서 조금 더 불성실 응답에 관해 명확한 조치를 실시한 연구들이 있는 것을 확인할 수 있었다. 따라서 학술지별 불성실 응답에 대한 인식 및 대처유형을 분석한 결과, 대부분의 학술지에서 불성실 응답에 대한 인식을 포함한 예방 및 대처방안이 여전히 미흡한 것으로 나타났다.

가장 바람직한 사례로 볼 수 있는 네 번째 유형인 ‘불성실 응답 예방 및 탐지방법 언급형’에 속한 몇 가지 사례를 소개하면 다음과 같다. Seo and Lee (2019)는 온라인 설문조사 도구인 서베이 몽키 (Survey Monkey)를 이용한 연구에서 ‘이 문항에서는 특정 답을 고르세요’라는 지시적 조작 점검 문항 (instructional manipulation check item)을 활용하여 불성실 응답이 의심되는 응답을 제외한 표본을 분석에 사용하였다고 명시적으로 밝혔다. 또한, Jun et al.(2017)는 현금성 마일리지를 설문 응답에 대한 인센티브로 제공하였으며, 모든 문항에 대해 응답이 이루어지도록, 그리고 특정 페이지에서는 응답시간이 30초 미만인 경우 다음 페이지로 이동할 수 없도록 설문을 구성하였다. 마지막으로, Kang et al.(2015)는 브랜드에 대한 기사 내용이 브랜드 평가에 미치는 영향에 관한 연구에서, 의미론적 (semantic) 불성실 응답 탐지 방식을 활용하여, 부정적 기사를 읽었음에도 브랜드에 대한 평가가 긍정적으로 이루어진 경우를 불성실 응답으로 간주하여 제거하였다고 밝혔다.

III. 불성실 응답의 원인

지금까지 불성실 응답에 대한 국내 연구자들의 인식수준과 대처현황을 살펴보았다. 다음으로는 불성실 응답의 예방방안을 논의하기에 앞서, 이를 유발하는 원인에 대해 살펴보고자 한다. Krosnick(1991)은 설문조사에 참여한 응답자들이 충분한 인지적 노력을 들이지 않고 적당하게 응답하는 경향이 있다고 주장하였는데, 이와 같은 행동을 유발하는 주요 원인으로서는 응답자의 능력, 동기 부족, 응답자 부담, 그리고 이들 간의 상호작용을 제시하였다. 이러한 기저 원인 외에도, 불성실 응답은 응답자들의 설문에 관한 관심 부족(Meade and Craig, 2012), 긴 설문의 길이(Palaniappan and Kum, 2019), 연구자와 참가자 사이의 제한적인 접촉(Ward and Pond, 2015), 주의를 분산시키는 환경(Zwarun and Hall, 2014)과 같은 다양한 요인에 의해 발생할 수 있다. 불성실 응답에 영향을 미치는 이러한 요인과 특성은 <Table 4>와 같이 주로 두 가지 기저 원인(동기와 부담)과 그 밖의 경우로 구분된다.

3.1 동기부여

불성실 응답은 설문응답자의 동기와 관련하여 발생할 수 있다(Huang et al., 2012; Maniaci and Rogge, 2014; Meade and Craig, 2012). 구체적으로, 불성실 응답은 설문 조사 지침을 준수하고 항목 내용을 올바르게 해석하며 정확한 응답을 제공할 동기가 없을 때 발생한다. 불성실 응답과 관련 있는 심리적 동기로는 사회적 동기(Meade and Craig, 2012; Ward and Pond, 2015), 자극추구 동기(Brower, 2018; Meade and Craig, 2012), 성

(Table 4) 불성실 응답에 영향을 미치는 주요 원인과 이유

주요 원인		불성실 응답이 발생하는 이유	관련 연구
동기	참여 동기	설문의 내용에 주의를 기울이지 않아 발생	Anduiza and Galais(2017), Chmielewski and Kucker(2020), Shamon and Berning(2020)
	자극 추구 동기	설문에 대한 흥미, 관심, 재미를 충족하지 못해 발생	Brower(2018), Maniaci and Rogge(2014), Meade and Craig(2012)
	사회적 동기	설문자와 응답자 간 사회적 상호작용의 부족으로 발생	Meade and Craig(2012), Ward and Meade(2018), Ward and Pond(2015)
	공격성 동기	공격성 동기로 인해 능동적 불성실 응답 행동이 발생	DeSimone et al.(2020), McKay et al.(2018)
응답자 부담		장시간의 설문으로 응답자의 피로가 누적되어 발생	Berry et al.(1992), Meade and Craig(2012), Palaniappan and Kum(2019)
기타		성격, 정서, 태도 등과 같은 개인 특성으로 발생	Bowling et al.(2016), DeSimone et al.(2020), Huang, Liu et al.(2015)
		나이, 성별, 교육수준 등과 인구통계학적 특성으로 발생	Maniaci and Rogge(2014), Nichols and Edlund(2020)
		인종, 문화와 같은 문화적 요인으로 발생	Grau et al.(2019)
		산만한 환경과 같은 상황적 요인으로 발생	Zwarun and Hall(2014)

취동기(Kühne and Kroh, 2018; Ward and Meade, 2018), 공격성 동기(DeSimone et al., 2020) 등이 있다. 이 중, 사회적 동기, 자극추구 동기, 성취동기가 부족하거나 공격성 동기가 작용할 경우 불성실 응답이 발생한다. 구체적으로는 (1) 보너스 등에 주의를 빼앗겨 설문 내용에 충분한 주의를 기울이지 않는 경우, (2) 설문에 대한 흥미, 관심, 재미 등이 충족되지 못하는 경우, (3) 설문 설계자와 응답자 간 사회적 상호작용이 부족한 경우, (4) 응답자가 설문 또는 응답자에 대해 공격성을 갖고 능동적으로 불성실 응답 행동을 보이는 경우와 같이 네 가지 이유로 인해 불성실 응답이 발생한다.

첫째, 만일 응답자들이 설문에는 건성으로 응답하

고 설문 후 지급되는 보너스와 같은 인센티브에 더 집중할 경우, 설문의 내용에 충분한 주의를 기울이지 않아 불성실 응답이 발생하게 된다(Anduiza and Galais, 2017; Bowling et al., 2016; Chmielewski and Kucker, 2020; Shamon and Berning, 2020). Akbulut(2015)는 온라인 설문 응답자들이 개인의 인구통계학적 정보를 잘못 체크하는 경우, 응답자들이 충분한 주의를 기울이지 않고 설문조사 후 지급되는 보너스 포인트에 초점을 맞추기 때문에 설문에 건성으로 답변을 제공했을 수 있다고 주장하였다. Meade and Craig(2012)도 응답에 대한 보너스 포인트는 성실한 응답을 이끌어 내지 못할 수도 있다고 설명하였다. 예를 들어, 설문

완료자 중에 보상을 목적으로 한 응답자가 다수 포함될 경우, 이들이 이야기하는 불성실 응답은 조사 데이터 품질에 부정적 영향을 미치기 때문이다(Lee et al., 2008). 또한 Bowling et al.(2016)은 불성실한 응답으로 인해 참여 인센티브를 받지 못할 수 있다고 경고하는 것은 참가자들이 신중하게 대응하도록 동기를 증가시켜 결과적으로 불성실 응답 발생률이 낮아진다고 보고하였다.

둘째, 설문을 받아들이는 개인의 관심과 흥미에 따라서 불성실 응답이 발생할 수 있다(Brower, 2018; Maniaci and Rogge, 2014; Meade and Craig, 2012). 설문에 대한 응답자의 관심과 흥미는 응답자가 설문에 신중하고 정확하게 답변하고자 하는 동기와 관련이 깊다(Meade and Craig, 2012). 예를 들어, Brower(2018)은 설문길이가 길수록 불성실 응답이 증가하는데 특히 응답자의 흥미가 낮을 때 불성실 응답이 더 많이 나타난다고 보고하였다. Meade and Craig(2012)는 채용에 활용되는 설문이나 피드백을 받기 위해 참여하는 피드백 설문과 같이 설문의 결과가 응답자의 생활에 직접적인 영향을 미치거나 응답자가 설문의 내용에 관심이 있는 경우 혹은 응답자들이 설문에 대해 진실한 흥미를 느끼고 있는 경우, 불성실 응답이 발생할 가능성이 낮아진다고 보고하였다. 또한 Maniaci and Rogge (2014)는 연구에 관한 관심과 지원, 자신과 타인 또는 세상을 이해하려는 목적, 재미 추구 등의 내적 동기는 불성실 응답과 부적 관계이며 반면, 보상 또는 대학교 강의 추가점수 획득 등의 외적 동기는 불성실 응답과 정적 관계라고 보고하였다.

셋째, 설문설계자와 응답자 간 사회적 상호작용 부족이나, 응답자의 주의를 분산시키는 응답 환경과 같은 환경적 차원의 특성 또한 불성실 응답을 유발할 수 있다(Meade and Craig, 2012; Ward and

Meade, 2018; Ward and Pond, 2015). 설문 설계자와 응답자 간의 상호작용이나 적절한 설문조사 행동에 대한 사회적 규범 인식이 설문응답자가 신중하고 정확하게 응답해야 한다는 동기에 영향을 미칠 수 있다(Meade and Craig, 2012; Ward and Meade, 2018; Ward and Pond, 2015). 설문 설계자와 응답자 간의 직접적인 사회적 상호작용이 이루어지기 어려운 온라인 설문조사의 경우, 응답 과정을 관찰하거나 감독하기 어려우므로 불성실 응답 빈도가 높을 수 있다. 이에, 물리적 거리가 사라지고 개인이 익명화되는 온라인의 특성으로 인해 응답자들의 책임감이 낮아질 수 있고, 이에 따라 불성실 응답과 같은 바람직하지 않은 응답 패턴이 발생하게 된다(Johnson, 2005; Meade and Craig, 2012).

넷째, 공격성 동기로 인해 응답자가 능동적으로 불성실 응답을 보일 수도 있다(DeSimone et al., 2020). DeSimone et al. (2020)은 잠재공격성이 높은 응답자일수록 설문과 문항에 더욱 좌절감을 느껴, 사회 규범과 관습을 무시하려는 의도를 표출하고 연구를 방해할 목적으로 연구자의 성실 응답 장려 시도를 무시하는 경향이 있을 것이기에 불성실 응답을 보일 가능성이 높다고 주장하였다. 또한, 불성실 응답과 상관관계를 갖는 것으로 기존에 알려진 타 변수를 통제한 후에도, 잠재공격성(implicit aggression)이 불성실 응답과 정적 관계를 보인다는 것을 확인하였다. 한편, McKay et al.(2018)은 응답자의 악의적 특성(malevolent traits)이 불성실 응답과 밀접한 관련이 있다고 보고하였다. 정직성 및 겸손함, 우호성과 같은 호의적인 특성(benevolent traits)보다 정신병질(psychopathy), 사회지배 지향성(social dominance orientation), 윤리적 영역에서의 위험감수(risk taking in ethical

domains) 등의 악의적 특성이 지시적 조작 점검문항과 한 줄 응답의 탐지방법을 사용하여 불성실 응답을 탐지한 결과에 영향을 미치는 것을 확인한 것이다. 그러나 이와 같이 공격성 동기로 인해 응답자가 능동적으로 불성실 응답을 보인다는 주장은 비교적 최근에 제기되었기에, 관점의 정립을 위해서는 보다 많은 실증연구 결과의 축적이 필요할 것으로 보인다.

3.2 응답자 부담

위에서 언급한 동기부여 이외에도, 응답자가 설문 길이, 설문 내용에 대한 인지적 부담으로 피로 혹은 지루함을 느끼면 불성실 응답을 할 수 있다. 구체적으로 응답자는 설문의 길이가 너무 길거나, 설문조사 질문에 답하는 데 필요한 노력이 너무 크거나 설문조사로 인한 스트레스가 너무 클 것 같다고 느껴지거나, 혹은 설문에 참여해달라는 요청을 너무 자주 받을 경우에 쉽게 부담감을 느낀다(Galesic and Bosnjak, 2009; Palaniappan and Kum, 2019). 따라서 응답자가 설문에 부담을 느끼면 설문응답률과 완료율이 낮아지며 데이터의 질에도 부정적인 영향을 미친다(Galesic and Bosnjak, 2009; Gideon, 2012). 설문조사는 응답자에게 유사한 응답 형식 및 항목에 응답할 것을 요구하는데, 사람들은 이러한 반복적이고 변화가 없는 작업에 참여할 때 종종 지루함을 경험한다(Gibson and Bowling, 2020). 또한 응답자는 설문의 질문을 이해하고 응답을 선택하기 위해서 인지적 노력을 기울여야 하므로(Krosnick, 1991) 질문에 응답하는 동안 피로함을 느낄 수 있다(Berry et al., 1992; Galesic and Bosnjak, 2009). 일반적으로 긴 길이의 설문은 응답자의 피로를 증가시키기에 불성실 응답을 유발하는 요인으

로 여겨진다(Gibson and Bowling, 2020; Meade and Craig, 2012; Palaniappan and Kum, 2019). 특히, 설문이 오랜 시간 동안 지속되는 경우, 응답자의 피로가 누적되어 설문의 중반이나 후반부에 응답의 질이 떨어지게 된다(Galesic and Bosnjak, 2009). Berry et al.(1992)의 연구는 대학생, 지역사회 자원봉사자, 직무지원자를 포함한 여러 표본을 대상으로 실시한 설문조사에 나타난 무작위 응답을 연구한 결과, 직무지원자와 같이 설문에 참여하는 동기가 강한 경우에도 설문의 후반부에 무작위 응답이 빈번하게 나타나는 것을 확인하였다.

3.3 기타

동기부여와 응답자 부담 이외의 기타 원인으로는 개인 특성, 인구통계학적 특성, 문화적 요인, 환경적 특성 등이 있다. 첫째, 성격 특성과 불성실 응답 간의 관계에 관한 연구는 특정한 성격 특성이 설문에 성실하게 응답할 동기에 영향을 미친다고 가정한다(Bowling et al., 2016; Ward et al., 2017). 연구자들은 성실성과 우호성이 높은 응답자는 설문에 더욱 주의를 기울이고, 지시사항을 잘 따르며, 연구자를 기꺼이 돕거나 기쁘게 하려는 등 이타적 태도를 보이기에 불성실 응답을 보일 확률이 낮다고 보고하였다(Bowling et al., 2016; DeSimone et al., 2020; Dunn et al., 2018; Grau et al., 2019; Maniaci and Rogge, 2014; Ward et al., 2017). 또한, 불성실 응답을 보이는 응답자들은 그렇지 않은 응답자들에 비해 낮은 수준의 자아 존중감을 가지는 것으로 나타났다(Maniaci and Rogge, 2014; McKay et al., 2018; Nichols and Edlund, 2020). 성격 특성 이외에도 여러 개인 특성이 불성실 응답과 상관관계를 가지고 있는 것으

로 밝혀졌다. 예를 들어, Huang et al.(2015)은 응답자의 정서와 태도가 불성실 응답과 상관관계를 갖는다는 것을 확인하였다.

둘째, 어떠한 특성의 응답자들이 불성실 응답을 보일 경향이 큰지를 파악하고자 나이, 성별, 교육수준 등의 인구통계학적 특성과 불성실 응답 간의 관계를 살핀 선행연구도 다수 존재한다. Maniaci와 Rogge (2014), Nichols와 Edlund(2020)는 응답자의 나이가 어릴수록 불성실 응답 비율이 높으며, 남성이 여성에 비해 불성실 응답의 비율이 높다고 제시하였지만, 반면 Grau et al.(2019)은 응답자의 성별과 나이로 불성실 응답을 예측할 수 없다고 보고하였다. 교육수준과 불성실 응답 간의 관계에 관해서는 대부분의 연구가 교육수준이 낮을수록 불성실 응답 비율이 높다고 보고하였다(Grau et al., 2019; Kim et al., 2019; Maniaci and Rogge, 2014; Nichols and Edlund, 2020).

셋째, 개인이 아닌, 국가와 대륙 등의 집단 단위에 서 나타나는 불성실 응답의 원인을 지리적 위치, 인종과 문화의 특성 등에서 찾고자 하는 연구들도 이루어지고 있다. Nichols and Edlund(2020)는 아마존(Amazon)사의 M-Turk 표본을 이용한 연구에서 아시아 응답자들이 북미 응답자들과 유럽 응답자들에 비해 현저히 높은 비율로 불성실 응답을 보인다는 것을 확인하였다. 또한, 표본의 수는 적었으나 남미 응답자들이 북미 응답자들에 비해 높은 비율의 불성실 응답을 보이는 것 역시 확인하였다. Grau et al.(2019)은 문화의 다양한 특성들과 불성실 응답과의 상관관계를 확인하였다. 국가 수준에서, 인간개발지수와 개인주의, 해방지수는 불성실 응답과 부적 관계에 있는 반면, 성불평등 지수와 권력 거리 지수는 불성실 응답과 정적 관계를 보였다. 특히, 국가의 해방지수가 해당 국가의 불성실 응답을 잘 예

측하였으며, 개인 수준에서도 거주 국가의 해방지수가 불성실 응답을 예측할 수 있는 것으로 나타났다.

넷째, 설문조사에 집중하기 어려운, 산만한 응답 환경과 같이 응답자가 신중한 응답을 제공하는 데 방해가 되는 상황적 요인 또한 불성실 응답을 유발하는 환경적 특성으로 작용할 수 있다. 예를 들어, 온라인 설문조사의 경우에는 응답자가 설문을 진행하는 동안 텔레비전을 시청하거나, 다른 업무를 병행하는 등의 멀티태스킹(multitasking)을 시도할 수도 있다(Meade and Craig, 2012). Zwarun and Hall(2014)은 온라인 설문조사에 참여한 표본의 거의 30%에 해당하는 응답자가 대략 30분의 설문조사 동안 주의를 분산시키는 하나 이상의 활동을 경험하였다고 보고하였다. 이러한 현상은 특히 어린 세대의 응답자들이 온라인 설문조사에 참여할 경우 두드러지는 것으로 나타났다(Carrier et al., 2009; Zwarun and Hall, 2014).

IV. 불성실 응답에 대한 전문가들의 인식과 대처방안

4.1 자료수집 및 분석과정

불성실 응답에 대한 학계 및 관련 업계 내 인식과 대처방안을 살펴보기 위해 각계 전문가를 대상으로 반구조화된 심층면접을 개별적으로 실시하였다. 면접 대상은 설문조사법의 활용빈도가 높은 경영학과 사회과학 분야 연구자와 설문조사 관련 업계 종사자 가운데 총 11명의 전문가들을 선정하여 <Table 5>와 같이 구성하였다. 학계에서는 설문조사 및 방법론 관련 연구수행 경험이 풍부하며, 관련 저서 또는 강

의경력이 있는 저명한 교수를 포함해 8명의 전문가를 선정하였다. 업계의 경우, 컨설팅 업체 또는 설문조사 및 방법론 관련 업체의 대표 또는 임원에 해당하며, 관련 사업수행 경험이 풍부한 전문가 3명을 선정하였다. 이 때, 학계 전문가들의 평균 교수경력은 16년에 달하였고, 업계 전문가들의 평균 업계 경력은 12년에 해당하였다. 체계적이면서도 심층적인

답변을 이끌어내기 위해 반구조화 면접법을 사용하였으며, 면접에 사용한 문항은 불성실 응답의 경험 및 인식, 불성실 응답의 예방으로 나누어 <Table 6> 과 같이 10개 문항으로 구성하였다. 자료 수집은 2020년 10월부터 2개월에 걸쳐 대면면접, 서면면접, 전화면접, 화상면접 방식으로 개별 실시하였다. 면접은 대상자 개인별로 30분에서 60분가량 소요되

<Table 5> 면접 참여자 목록

면접 대상자	영역	분야	교수 혹은 업계경력	면접방법
A	학계	경영학	6년	서면면접
B			3년	
C		교육학	28년	대면면접
D		사회학	41년	
E			19년	
F		13년	6년	대면면접
G				전화면접
H		소비자학	12년	대면면접
I		관련 업계	컨설팅	16년
J	12년			
K	설문조사		8년	

<Table 6> 면접 질문지

분류	문항 내용
불성실 응답의 경험 및 인식	불성실 응답은 무엇이라고 생각하시나요?
	불성실 응답을 경험하신 적이 있나요? 있다면 몇 퍼센트 정도 되는 것 같나요?
	경험하신 불성실 응답에 대해서 자세히 이야기해주시겠습니까? 주로 어떤 유형의 불성실 응답이었나요?
	불성실 응답이 설문조사 결과에 영향을 미친다고 생각하십니까?
	불성실 응답을 인식하신 경우, 어떤 조치를 취하셨나요?
	조치한 결과를 어떻게 보고하시나요?
불성실 응답의 예방	경험상, 어떤 상황에서 불성실 응답이 많았나요?
	불성실 응답을 예방하기 위해 어떤 방법을 활용하셨나요?
	어떤 방식이 불성실 응답을 줄이는 데 효과적/비효과적이었나요?
	향후 설문연구를 진행하실 경우, 보다 성실한 응답을 얻기 위해 어떤 점을 개선하실 예정인가요? 혹은 어떤 방식을 추천하시나요?

었으며, 면접 내용은 면접 참여자에게 사전 동의를 구하여 녹취 및 수기 기록하였다. 면접 후에는 참여자에게 소정의 기념품을 제공하였다. 수집된 자료는 연구진의 반복 검토와 합의를 통해 불성실 응답에 대한 인식, 경험에 의거한 실천적 예방방안, 연구 내 불성실 응답 보고방법, 불성실 응답에 관한 교육 및 연구 활성화의 네 가지 주제로 해석 및 분류되었다.

4.2 분석결과

4.2.1 불성실 응답에 대한 인식

전문가들은 불성실 응답이 데이터의 질을 저하시킬 가능성이 있다는 인식을 공통적으로 가지고 있었다. 대체로 불성실 응답이 의심되는 경우는 설문 응답의 5~10% 범위에 해당하는 것으로 인지하고 있었으며, 일부 전문가들은 온라인 설문의 경우 불성실 응답이 의심되는 경우가 더 높은 비율로 나타나는 것으로 보았다. 전문가들은 일반적으로 불성실 응답을 “의도적 혹은 비의도적으로 설문에 집중하지 못해서 발생한 응답(F)”을 나타낸다고 보았는데, 전문가에 따라 “연구자들의 의도대로 대답하지 않는 응답(G)”과 같은 광의적인 측면이나, “모든 설문을 같은 번호로 체크하는 경우(B)”와 같은 좁은 범위로 접근하는 등의 차이를 보였다. 이러한 차이는 응답에 있어서 데이터의 질을 저하시키는 여러 문제들(예: 무응답, 사회적 바람직성)에 대해 연구자들이 통합적으로 접근하기 때문으로 해석되지만, 그동안 국내 학계에서 불성실 응답에 대한 논의가 부족했다는 점에서도 원인이 있다. 한편, “응답자는 기본적으로 불성실하다고 전제해야 한다(...) 불성실 응답은 설문이 제대로 만들어지지 않았을 때도 발생하기에 조사설계 과정에서 충분한 노력과 시간을 들여서 설

문을 만들어야 한다(F)”, “설문 대상이 적합하지 않거나, 응답자가 제대로 이해하지 못하는 수준의 질문으로 구성되거나, 너무 많은 개방형 질문이 포함된 경우와 같이 설문이 제대로 설계가 되어있지 않다고 판단될 때는 불성실 응답이 아닌 설문의 문제로 본다(K)”와 같이, 불성실 응답을 논의하기 위해서는 연구자의 연구 의도에 맞도록 충분한 시간과 노력을 들여 작성하고 검증된 설문의 설계가 선행되어야 함을 다수의 전문가들이 강조하였다.

4.2.2 경험에 의거한 실천적 예방방안

A가 “조사 설계에서 불성실 응답을 낮추는 방안에 대한 연구가 국내는 전무하다.”라고 지적한 바와 같이, 불성실 응답 예방에 관한 명확한 가이드라인이 부재하기에, 각 참여자들은 각자의 시행착오와 경험에 기반하여 다양한 방식들을 활용하여 왔다.

그 중에서도 인센티브 제공의 중요성을 다수의 응답자들이 공통적으로 언급하였다. 한 전문가는 연구 진행 시 일반적으로 성실한 응답을 가정하지만, 이 자체가 지나치게 긍정적 기대임을 날카롭게 지적하며, “인센티브가 없으면 불성실하지 않을 이유가 없지 않은가? (...) 성실이라는 것 자체가 개인의 에너지 시간이라는 비용을 요구하는 것이다(C)”라고 언급하였다. 여건이 허락하는 한 가급적 큰 금액의 금전적 보상을 주는 것이 좋으며, 선물의 경우, “부과가 큰 품목에 대한 만족도가 높았다는 점(G)”도 언급되었다. 다만, 인센티브도 동기부여 목적의 수단일 뿐, 인센티브의 효과성은 한계 체감하며, 동기부여가 더욱 중요하다는 의견도 존재하였다(A, D).

설문의 내부적 설계 측면에서는 응답자들이 설문에 익숙해져 발생하는 ‘성숙효과(mature effect)’를 예방하는 방식이 권장되었다. 예로 “통으로 동일 문

항을 반복하지 않도록 제시방법을 다변화(D)”하거나, ‘당신은 술을 마십니까?’라는 분류 질문을 던지고, 음주 횟수나 주량 등 관련 세부 질문들이 나열되는 패턴을 반복할 경우, 응답자들이 질문패턴을 학습하여 빠른 종료를 위해 다음 분류 질문에 사실과 다르게 응답할 수 있으므로, 분류 질문들을 모두 한꺼번에 질문한 후, 이에 따른 세부 질문을 추후에 보여주는 방식(I) 등이 언급되었다.

설문 외적으로는 설문의 대상과 시기에 대한 고려의 중요성을 다수 언급하였다. 질문에 대한 답을 명확히 알고 있는 사람들에게 설문을 보내야 하며(J), 기업 설문 시 반드시 산업특성을 고려해야 하고(B), 관심을 가질만한 참여자들을 선정(K)해야 한다는 것이다. 또한, 설문을 실시하는 시간도 중요한데, 직장인을 대상으로 하는 경우, 연말연초 등 인사발령을 전후한 시기는 가급적 지양해야 하며, 아침이나 월요일 시간대에는 확인하지 못할 가능성이 있음을 고려해야 한다고 지적하였다(I). 이외에도 다수의 참여자들이 불성실 응답의 예방과 관련하여 공통적으로 지적한 사항들을 다음과 같이 정리하였다. 이외에도 다수의 참여자들이 불성실 응답의 예방과 관련하여 공통적으로 지적한 사항들을 다음과 같이 정리하였다.

- 오프라인 대비 온라인 설문조사 상황에서 특히 빈번히 발생하므로 주의(A, E)
- 설문 내 가짜문항 또는 논리적 불일치 문항 방식을 적극 활용(D, H)
- 설문의 길이에 주의, 1회 설문의 총량이 최대 30~50문항을 넘지 않도록 설계(E, F, J, I)
- 연구지시문에 설문의 취지와 중요성을 강조하여 작성(D, H)
- 설문에 악의적인 태도를 지닐 수 있는 동기 미리 고려(예, 노조 등) (B)

4.2.3 연구 내 불성실 응답 보고방법

불성실 응답을 이해하고 파악한 후에는 구체적으로 어떻게 불성실 응답을 제거했는지를 연구 본문을 통해 보고하는 것이 필수적이라는 점에 대해 다수의 전문가들이 동의하였다(E, G, H, J, K). 그러나 “전체 표집한 대상자 수 대비 연구에 실제로 사용한 대상자의 수는 보고하지만, 데이터 처치를 통해 걸러낸 불성실 응답의 경우는 보고하지 않는다(E)”와 같이, 전체적으로 제거된 불성실 응답의 수를, ‘~제외하고 분석하였다’처럼 구체적인 제거 기준 없이 보고하는 것이 대부분의 연구에서 사용되는 방식이다(H). 이는 불성실 응답 보고에 대한 명확한 기준 역시 존재하지 않기에, 연구 상에서 보고하는 방법 또한 일정하지 않은 것으로 해석할 수 있다. 유의한 연구결과만을 쫓아, 명확한 기준 없이 불성실 응답으로 판단하여 제거하는 경우, 연구 윤리상의 문제도 발생할 수 있다(G). 따라서 F는 “국내에서는 불성실 응답을 보고하는 윤리적 기준이 명확하지 않아 지금까지 문제가 되지 않았지만, 국외에서 국내의 연구를 바라볼 때 신뢰할만하지 않다고 느끼는 경우가 많다”라고 지적하며, 연구에 사용된 원본 데이터를 공개하여 불성실 응답을 제거하는 과정에 대해 투명하게 제시할 필요가 있음을 강조하였다.

4.2.4 불성실 응답에 관한 교육 및 연구 활성화

위에서 언급한 것과 같은, 조사설계와 분석 단계에서 불성실 응답을 최소화시키는 대처방안 이외에도, 학계 전문가들은 실제적으로 불성실 응답을 줄이기 위해서는 “조사방법론 수업을 들었던 연구자들이 연구를 설계해야 한다(G)”며 최소한의 가이드라인의 필요성을 역설하였다. 또한 “한국 사회에서 데

이터를 수집할 때 기반이 되는 공공성과 시민성과 같은 사회적 가치에 대한 인식 개선이 연구자와 응답자 양측 모두에게 필요하며 이는 대학원의 커리큘럼에서 체계적으로 현장과 이론을 접목한 조사방법론 교육으로 개선될 수 있다(F)”고 언급하였다. 부정확한 척도 사용이 불성실 응답을 높인다는 측면에서도 실제로 많은 연구자들이 “검증되지 않은 부정확한 척도를 활용하여 설문조사를 실시하는 것에 우려(G)”를 나타냈으며 이는 곧 불성실 응답 발생을 높이는 원인이 될 수 있기에 방법론 교육을 통해 개선해야 한다고 지적하였다. 마지막으로 해외에서 조사방법론으로 박사를 취득한 전문가는 “국내 사회과학 분야에서 조사방법론에 대한 이해도가 낮으며, 해외에서는 20년 전에 사용하였던 순차 번역(back translation)과 같은 방법론을 아직도 사용하는 현실(I)”이므로 장기적으로 불성실 응답을 줄여나갈 수 있도록 학계의 지속적인 관심과 연구가 뒷받침되어야 한다고 강조하였다.

V. 불성실 응답을 예방하는 설문조사의 설계

앞서 살펴본 불성실 응답 발생의 주요 기저원인과 전문가 인터뷰 결과에 따라, 설문 참여에 대한 응답자의 동기부여와 부담을 중심으로, 불성실 응답을 예방할 수 있는 설문조사 설계방법에 대하여 논하고자 한다. 따라서 본 장에서는, 응답률 관련 연구에서 풍부한 실험을 통해 검증된 연구설계 원리들을 불성실 응답 예방에 준용하여 논의한다. 물론, 가장 바람직한 참고자료는 Meade and Craig(2012) 등의 논의 이후, 불성실 응답을 탐지하는 다양한 방법들

(예, 마할라노비스 거리 등)을 활용하여 연구설계와 불성실 응답 간의 관계를 구체적으로 탐구한 사례들이다. 이러한 연구에 대해서는 별도로 불성실 응답에 관한 연구임을 명시하고 그 활용 방안에 대해 보다 상세히 논하고자 한다.

5.1 연구 지시문 작성

불성실 응답을 예방하기 위한 지시문이란, 프레임을 사용하여 응답자가 설문에 성실하게 응답할 동기부여가 되도록 설문 시작 부분의 지시사항에 성실 응답을 독려하는 문구 혹은 경고를 삽입하는 것을 의미한다(Gideon, 2012; Gibson, 2019; Huang et al., 2012). 프레임은 대단위 설문조사를 실시하는 연구들에서 응답률을 높이기 위해서 고안되었는데, 이는 응답률을 높이는 것 뿐 아니라 불성실 응답을 예방하는 방법으로도 활용할 수 있다(Gideon, 2012). 설문조사에서 지시문은 일반적으로 정직, 정확성 및 익명성을 강조하는데(Huang et al., 2012), 불성실 응답을 줄이기 위한 프레임의 활용은 설문 응답 시 의사결정이 명확하고 효과적으로 이루어지도록 설문의 내용을 구성하고 설문 주제를 제시하는 역할을 보조할 수 있다(Gideon, 2012; Woodside and Dubelaar, 2003).

구체적으로, 프레임을 활용한 지시문은 긍정적 지시문, 경고형 지시문, 상호작용 지시문의 세 가지 형태로 구분할 수 있다. 첫째, 긍정적 지시문은 동기부여를 통하여 설문응답에 몰입하도록 독려하는 형태로, ‘조직구성원들의 응답이 조직의 중요한 의사결정에 영향을 미친다.’ 혹은 ‘귀하가 응답하신 설문은 향후 이런 주제의 연구 발전에 크게 기여할 것이다.’와 같은 문구를 삽입하는 것이다(Gibson and Bowling, 2020; Gideon, 2012; Kanfer et al., 2017;

Marshall, 2019; Meade and Craig, 2012; Ward and Meade, 2018). Marshall(2019)은 과정기반 접근(process-based approaches)을 활용하여 지시문을 활용한 동기부여가 가능하다고 설명하면서, 동기이론에 기반을 둔 예방방법은 참가자가 설문지를 빨리 완료하려고 건성으로 응답하는 것을 막을 수 있다고 주장하였다. 그러나 동기가 활성화되기 위해서는, 개인이 행동을 수행하기 위해 투입한 노력의 양과 행동에 기여이 투자할 시간이 비례해야 하므로, 단순한 긍정적 지시문만으로 동기부여가 일어나 불성실 응답의 각종 측정값이 줄어든다고 보기에는 한계가 있다(Diefendorff and Chandler, 2011). 즉, 응답자의 동기 요소가 상황적 제약과 상호작용하여 불성실 응답에 영향을 미칠 수 있기에 불성실 응답을 예방하는 측면에서 다른 예방방법과 함께 동기부여를 통한 긍정적 지시문이 활용될 수 있다.

둘째, 경고형 지시문으로는 사회적 파워 이론(social power theory)을 기반으로 단순 경고성 문구와 실제적 경고성 문구를 삽입하는 방식이 가장 많이 사용되고 있다(Gibson, 2019). 단순 경고성 문구는 '응답이 성의 없을 경우 통계처리에서 제외될 수 있다.' 혹은 '불성실 응답이 발생하면 많은 시간과 노력을 들여서 하는 연구 결과가 왜곡될 수 있다.' 등이며 실제적 경고성 문구는 '불성실 응답으로 간주될 경우, 인센티브 지급에서 제외될 수 있다.' 혹은 '불성실 응답으로 간주될 경우, 학점에서 불이익을 받을 수 있다.' 등이다(Fan et al., 2012; Gibson and Bowling, 2020; Huang et al., 2012; Meade and Craig, 2012). Huang et al.(2012)의 연구는 불성실 응답 탐지방법을 명시한 경고 메시지가 존재하고 불성실 응답으로 인해 처벌받게 될 것이라는 경고 지시문이 있을 경우, 네 개의 불성실 응답지수 중 세 가지 경우에서 대조 그룹에 비해 불성실 응

답이 감소함을 제시하였다. 또한 부정적 경고는 응답시간(Ward and Pond, 2015), 의미론적 반의어(Huang et al., 2012) 및 한 줄 응답(Huang, Bowling, et al., 2015)과 같이 간단한 불성실 응답지표에 효과적이다. 그러나 단순한 경고성 지시문은 짝수 홀수 일관성(Ward and Pond, 2015) 또는 마할라노비스 거리(Ward and Pond, 2015)와 같은 보다 정교한 탐지기법에는 유의미한 차이가 없다고 보고되었다. 일반적인 경고형 지시문은 부주의한 반응을 불성실 응답으로 명확하게 탐지하기 어렵고, 응답자의 성실한 응답행동과 무관할 수 있으며, 거친 표현은 오히려 설문응답자의 몰입을 방해할 수 있기에, 부정적 경고형 지시문의 사용은 신중하게 설문문의 맥락 안에서 활용되어야 한다(Gibson, 2019).

셋째, 상호작용 기법은 응답자와 설문자 간의 상호작용이 불성실 응답을 줄인다는 가정 하에 지시문이나 지시사항을 사용하여 설문 프로세스 안에서 혹은 설문이 완료된 후 응답자에게 직·간접적으로 응답의 질에 피드백을 주는 것을 의미한다(Gibson, 2019; Ward and Pond, 2015). 예를 들어 '응답자에게 설문의 내용과 질에 대한 피드백을 제공할 것이다(Ward and Pond, 2015).' 혹은 '설문 후 불성실 응답이 아닌 경우 엑스트라 인센티브를 제공한다(Marshall, 2019).'와 같은 내용의 지시문과 함께 영상 혹은 녹음파일로 지시사항을 제공할 수 있다(Ward and Meade, 2018). 예를 들어, Ward and Pond(2015)는 설문이 끝난 후에 응답자들에게 응답의 질에 대한 피드백을 제공할 것을 지시문에 명시한 경우, 응답자들이 설문을 완료하는데 더 많은 시간을 들인다고 보고하였다. Gibson(2019)은 온라인 설문조사 상황에서 대화형 메시지를 프로그램을 통해 상호작용 경고 지시문(interactive warning message)과 상호작용 독려

지시문(interactive encouraging message)을 사용하여 불성실 응답을 비교하였다. 이 때 상호작용 독려 지시문이 있는 그룹이 상호작용 경고 지시문이 있는 그룹에 비해서 페이지 시간 지표에서 불성실 응답이 통계적으로 유의하게 줄어들었다고 보고하였다. 그러나 상호작용 피드백 지시문도 불성실 응답을 확인하는 모든 지표에서 상호작용 피드백이 불성실 응답에 미치는 효과가 줄어들지 않았다(Ward and Pond, 2015). 구체적으로 상호작용을 위한 지시사항을 영상이 아닌 지시문으로 제공하였을 때에는 오히려 응답자의 흥미가 감소하였고, 불성실 응답의 객관적 지표가 변화하지 않았다는 연구결과도 존재한다(Ward and Meade, 2018). 따라서 상호작용 기법의 활용도 단순한 지시문을 벗어나 영상이나 녹음파일과 같은 다양한 방법을 활용할 필요성이 제기된다.

종합하면, 지시문을 활용한 선행연구들은 하나의 지시문을 사용하여 불성실 응답의 지표를 줄이는 데는 한계가 있다고 보고하였으며, 경고성 지시문을 사용할 때는 응답자가 거부감을 느끼지 않도록 신중하게 설문의 맥락 안에서 활용되어야 한다고 강조하였다. 앞서 제시한 세 가지 형태의 지시문 외에도 인지부조화 또는 위선(hypocrisy)을 유발하는 지시사항을 활용하거나(Ward and Meade, 2018), 온라인 설문 페이지에서 가상 존재를 배치하거나(Ward and Pond, 2015), 지시문에서 미리 설문소요시간을 알려주는(Galesic, 2006) 등의 방법이 활용될 수 있다. 따라서 설문조사에서 지시문을 사용할 때는 설문자의 권위가 전달되면서 응답자의 몰입이 떨어지지 않도록 문맥 안에서 신중하게 사용되어야 하며 최대한 상호작용이 가능하도록 구성하는 방안을 모색해야 한다.

5.2 개별 문항 설계

다음에서는 설문지에서 개별 문항 1개가 지닌 특성이 응답자의 불성실한 응답에 어떻게 영향을 미치는지 살펴본다. 여기서 언급되는 문항의 길이는 설문지 전체의 길이(예, 100~150개 문항 수)가 아니라 설문 문항 1개의 길이임에 유의한다. 설문지 전체 길이의 특성이 불성실 응답에 미치는 영향에 관하여는 다음 장에서 별도로 논의한다.

일반적으로 질문의 길이가 짧을수록 응답자의 인지적 부담을 줄여주어, 성실하고 정확한 응답을 이끌어낼 것이라고 추정할 수 있다. 다만, 기존의 연구들에 따르면, 구체적인 사항이 적시된 개별 문항이 길수록 좋은 응답결과를 얻을 수 있다는 증거 역시 존재한다(Blair et al., 1977; Laurent, 1972). 연구자들은 문항의 길이가 응답의 정확도에 미치는 영향을 확인하기 위해, 동일한 내용을 짧은 문장과 긴 문장으로 질문한 뒤, 객관적으로 비교할 수 있는 자료(예, 실제 의료기록 등)와의 일치 수준을 비교하였다. 그 결과, Blair et al.(1977)의 음주행동 빈도를 묻는 설문과 Laurent(1972)의 환자들을 대상으로 한 청각문제에 대한 설문, 두 설문 모두 문항이 길고 자세하게 설명이 되어 있을 때 응답자가 더욱 정확하게 이해하고 성실하게 답변하는 것으로 나타났다. 그러나 문항 자체의 길이에 관한 연구는 풍부하지 않은데, 설문연구의 경우 신뢰도와 타당도 검증을 엄격히 거친 설문도구를 주로 활용하기 때문이다. 문항의 길이와 불성실 응답 간의 관계를 확인한 연구는 현재까지는 존재하지 않지만, 기존의 발견들과 앞서 논의한 불성실 응답의 주요 원인을 고려할 때, 문항의 길이를 지나치게 줄일 필요는 없을 것으로 보인다. 오히려 길이를 유지하되 번역과정에서 될 수 있으면 다양한 표현을 적용하

는 것(paraphrasing)이 불성실 응답의 주요 원인 중 한 가지인 지루함을 최소화 할 수 있는 방안이다.

설문조사에서 수치화된 값을 응답하도록 하는 평정 척도(rating scale)은 이를 처음 고안한 심리학자의 이름을 따서 리커트 척도(Likert scale)라고도 불리며 가장 많이 활용되는 방식이다(Joshi et al., 2015). 이 리커트 척도를 사용함에 있어, 몇 점 척도를 사용할 것인가는 매우 중요한 이슈이다. 많은 연구자들이 적절한 척도의 수를 확인하고자 노력하였으나, 큰 차이는 발견되지 않았다. 일반적으로 신뢰도는 7점에 가까울수록 증대되며, 이후에는 신뢰도에 차이가 없는 것으로 확인되었다(Givon and Shapira, 1984). 불성실 응답의 원인에 관련하여서는 척도의 숫자가 지나치게 커질 경우, 인지적 부담이 증가할 수 있으므로, 예외적인 상황이 아니면 기존의 연구들이 권고하는 7점 척도로 범주를 제한하는 것이 바람직하다.

불성실 응답을 예방함에 있어 보다 중요한 결정요인은 중간 보기에 대한 고려이다. 7점 또는 5점 척도에서 중간 보기, 또는 '보통이다'라는 항목이 존재할 경우, 응답자들이 인지적 부담을 줄이고, 설문 응답속도를 높이고자 가운데 보기를 무조건 선택할 가능성도 없지 않다. 그러나 문항에 따라 긍정도 부정도 아닌 중간의 태도를 실제로 지닌 응답자들이 존재할 수 있기에, 이들을 불성실한 응답이라 일괄적으로 간주하는 것은 바람직하지 않다(Shuman and Presser, 1981). 반대로, O'Muircheartaigh and Campanelli(1999)은 중간 보기가 아예 주어지지 않는 경우, 응답자들이 불성실하게 아무 보거나 임의로 선택하였음을 보고 한 바 있다. 이러한 경향성을 재확인해보고자, Park et al.(2020)에서 활용된 3,030명의 응답형태를 분석한 결과에서도, 개별 문항에 대하여 중간 보기를 선택한 경우는 10%대

를 넘지 않았다. 적어도 한국적 맥락에서는 응답자들이 자신의 긍정 또는 부정적인 입장을 나타내는 것을 선호하고, 생각보다 중간 보기('보통이다')를 선택하는 경향이 크지 않았다. 또한 척도의 라벨(예, 높다, 자주 한다 등 정도에 상응하는 표현의 병기)의 활용에 있어서, 설문지의 공간 또는 사용하는 설문 시스템이 지원하는 한 가급적 모든 척도에 붙여주는 것이 양극단의 값에만 설명을 붙이는 것보다 바람직하다(Krosnick and Berent, 1993). 이는 응답자들이 개별 척도의 수준을 해석하는 인지적 부담을 줄여줄 수 있다. 기존의 연구에서도 모든 척도에 라벨링을 하는 방식이 신뢰성을 높여주었을 뿐만 아니라, 응답자들의 만족도 또한 높여주었다(Dickinson and Zellinger, 1980).

동일하게 정량적 내용을 묻더라도, 개방형 질문은 보다 풍부한 정보를 확보할 수 있다. 따라서, 나이와 같이 쉽고 정확하게 응답할 수 있는 질문은 선택형(예, 10대, 20대 등)이 아닌 개방형을 질문하는 것이 유리하다(Mondak, 2001). Couper et al.(2001)에 따르면 질문의 난이도에 따라 답변을 입력하는 공란의 크기 또한 응답의 질에 영향을 미치므로, 예상되는 주관 답변의 길이와 내용의 깊이에 따라 설문지 상 공란의 크기를 조정할 필요가 있다. 반면, 재산, 근로소득, 특정 기간 동안 참여횟수 등 응답자들이 정확한 수치를 환산하기 어려워 어렵게산으로 대답할 확률이 높은 항목들에 대해서는 오히려 개방형보다 선택형 질문을 제시하는 것이, 불성실한 응답을 예방하는데 도움이 된다.

역코딩 문항은 종종 설문에 주의를 기울이지 않고 답하는 사람들을 확인하고, 주의를 환기하고자 활용되어 왔으나 결론적으로는 측정 도구의 타당도와 신뢰도만 저하시키고 충분한 역할을 하지 못하고 있다(Vigil-Colet et al., 2020). 역코딩 문항의 사용

은 무효한 응답을 줄이지 못한다는 연구가 존재하며 (Schmitt and Stults, 1985), Curran(2016)은 역코딩 문항을 불성실 응답 이슈를 해결하는 통치약으로 활용하여서는 안 된다고 지적한 바 있다. 결론적으로, 불성실 응답을 탐지 또는 줄이려는 목적으로 역코딩 문항을 일부러 줄이거나 삽입해서는 안 된다. 역코딩 문항이 불성실 응답의 예방에 미치는 영향은 불확실하나, 역코딩 문항의 존재가 불성실 응답 사후적 탐지결과에 영향을 미친다는 점은 확실하기 때문이다(Curran, 2016). Curran(2016)은 역코딩 문항이 없는 설문, 역질문 문항은 있지만 역코딩이 되지 않은 데이터, 역질문 문항이 있고 역코딩이 된 데이터에 있어서 불성실 응답을 어떻게 접근해야 할지에 대하여 가이드라인을 제공하고 있으므로 이를 참고할 만하다.

5.3 설문지 구성

설문 설계과정에서 적정 설문길이, 설문구조, 시각적 요소 삽입 등의 예방적 기법을 사용해 응답자의 피로감을 최소화하고 설문을 끝까지 성실하게 완료하도록 도울 수 있게 설문을 구성하여 불성실 응답을 예방할 수 있다(Bowling et al., 2016; Huang, Liu et al., 2015; Meade and Craig, 2012). 설문항목의 복잡함이나 모호함은 응답자로 하여금 항목의 의미를 작위적으로 해석하게 하여 무작위 응답(random responding)을 증가시킬 수 있다(Podsakoff et al., 2003). 이에, 효과적인 설문구성은 전체 설문길이, 형태 및 내용의 측면에서 발생하는 불성실 응답을 예방하고 이로 인해 발생하는 측정 오차를 줄일 수 있다(Park and Lee, 2017).

첫째, 적정한 설문길이는 응답률을 높이고 불성실 응답을 줄일 수 있다. 불성실 응답 연구가 본격적으

로 논의되기 이전부터 선행연구들은 설문길이와 응답률, 응답의 질의 관계에 대해서 논의하였다(Deutskens et al., 2004; Galesic and Bosnjak, 2009; Groves, 1989). 응답자는 자신이 인식하는 적절한 시간 동안 설문을 완료하려는 성향을 보이기에, 대체적으로 설문길이가 짧을수록 응답률이 높아지며 (Deutskens et al., 2004), 설문조사 소요시간이 길어질수록 응답의 질이 떨어지게 된다(Groves, 1989). 설문길이는 불성실 응답과도 밀접한 관계가 있는데, 지나치게 긴 설문은 응답자에게 피로감과 지겨움을 발생시켜 성실 응답을 방해하게 된다(Bäckström and Björklund, 2019; Galesic, 2006; Galesic and Bosnjak, 2009; Gibson and Bowling, 2020; Liu et al., 2013; Nguyen, 2017). 구체적으로, Reymond et al.(2000)은 불성실 응답과 설문길이에 관련하여 전체 설문시간이 15분이 넘어 가면 불성실 응답이 발생하기 쉽다고 제시하였다. 그러나 최신 연구들은 설문길이에 소요되는 시간만으로 불성실 응답을 줄이는 것은 무리가 있다고 주장하며(Gibson, 2019), Gibson and Bowling(2020)는 설문의 길이와 부정적 경고형 지시문의 상호작용이 불성실 응답을 줄인다고 보고하였다. 따라서 설문시간만으로 불성실 응답을 줄인다고 단언할 수는 없지만, 극단적으로 긴 길이의 설문은 성실한 응답을 이끌어내기에 적절치 않다고 볼 수 있다.

둘째, 문항들이 설문지 전체의 흐름에 맞게 자연스럽게 공통된 주제를 가지도록 배치함으로써, 설문 응답에 필요한 인지적 부담을 줄여 응답의 질을 향상시키고 불성실 응답을 예방할 수 있다(Park and Lee, 2017; Weijters et al., 2014). 설문구조와 응답의 질의 관계에 대한 Couper et al.(2001)와 Tourangeau et al.(2004)의 연구는 인터넷에서 설문 문항들이 한 화면에 한 개씩 제시된 것 보다,

한 화면에 한꺼번에 제시되었을 때, 문항 간 상관계수가 더 높아진다고 보고하였다. 그러나 유사한 주제의 질문을 모아서 제시할 경우, 설문지의 구조에 의한 측정오차인 맥락효과(context effect)가 발생할 수 있다(Park and Lee, 2017; Tourangeau et al., 2003). 맥락효과는 설문구조가 어떻게 배치되어 있는지에 따라 설문의 응답의 질에 영향을 미치며, 특히 설문문항이 모호하거나 설문문항이 응답자에게 익숙하지 않은 경우 더욱 두드러진다(Tourangeau et al., 2003). 문항배치에서는 맥락효과 뿐 아니라 동일한 개념을 측정하는 문항들을 인접하게 배치하였을 때, 인지적 앵커링(anchoring)과 조정(adjusting)을 통한 인지적 휴리스틱(heuristic)의 작용을 일으키는 동화효과(assimilation effect)도 주의해야 한다(Weijters et al., 2014; Gehlbach and Barge 2012).

요약하면, 연구자는 설문 설계단계에서 응답자의 흥미 수준이 유지될 수 있도록 설문지를 구성해야 하며, 설문 소요시간이 길어질 경우, 중요 질문을 설문지의 초반에 배치하여 응답의 질이 떨어지지 않도록 세심한 설계가 필요하다. 또한, 응답자의 인지적 부담을 줄이기 위한 효과적 설문구성이 필요하며, 특히 온라인 설문에서는 시각적 요소를 도입하여 연구자와 응답자의 상호작용을 높이는 설문구성 역시 고려해야 한다.

5.4 인센티브 제공

많은 연구자가 설문조사의 수행에 있어 인센티브 제공의 중요성을 인식하고 있다(Singer, 2002). Couper(2008)는 선행연구를 기반으로 인센티브 제공에 대하여 다음과 같은 네 가지 결론을 도출하였다. (1) 어떤 형태의 인센티브도 인센티브를 제공

하지 않는 것보다 좋으며, (2) 향후 지급을 약속한 인센티브보다 선지급형 인센티브가 좋다, (3) 현금이 비금전적 인센티브보다 좋으며, (4) 추첨을 통해 일부에게 비싼 금액의 선물을 지급하는 것 보다 전원에게 작은 인센티브를 지급하는 것이 좋다. 이러한 인센티브에 관한 연구결과들을 참고할 때 가장 유의해야 할 것은 대다수의 인센티브 연구가 연구의 참여율 또는 응답률을 높이는 것을 목적으로 하고 있다는 점이다(예, Cantor et al., 2008; Göritz, 2004). 인센티브의 제공, 특히 선지급형 인센티브가 설문 참여 거절을 줄이는 것은 확인되었으나, 인센티브가 설문응답의 질에 미치는 영향은 알려진 바가 극히 드물다(Singer and Ye, 2013). 인센티브와 불성실 응답 사이에 어떠한 연관성이 있는지를 연구한 경우 또한 상당히 드물기에, 소수의 연구결과들을 보다 자세히 소개한다.

Göritz(2004)는 아마존(Amazon)사의 M-Turk과 같은 온라인 패널 시스템의 참여자들을 대상으로 현금, 복권, 선물(시계 등) 다양한 형태의 인센티브를 조작화한 결과, 인센티브의 유형이 응답률에는 영향을 미쳤지만, 응답의 질에는 유의미한 영향을 주지 않는다고 보고하였다. Shamon and Berning(2020)는 사전적 탐지문항을 넣은 후 온라인 설문을 실시하였는데, 인센티브를 받은 응답자들이 그렇지 않은 샘플에 비하여 높은 불성실 응답 비율을 나타낸다고 보고하였다. 그러나 이 두 집단은 무선 할당된 집단이 아니라 최초부터 돈을 받기로 약속하고 참여한 집단과 자발적으로 참여한 집단으로써 후자의 경우 내적 동기부여 수준이 더 높기에, 애초에 성실히 참여하고자 하는 동기가 낮은 참가자들이 첫째 집단에 포함되었을 확률이 높다. Gibson(2016)의 연구에서는 동일 표본(대학생)에 대하여, 두 집단으로 무선할당 후, 실험집단에 대해서는 통제집단

보다 다음과 같은 과정을 추가하였다. 설문 진행에 앞서 '고도의 통계적 방법을 활용하여, 정확하고 주의를 충분히 기울인 응답자에 한하여, \$30 상당의 기프트카드 추첨에 포함한다.'라는 내용을 구두 및 설문의 지시문에 담았다. 그 결과, 성실한 응답에 대하여 금전적 인센티브를 보장받은 집단이 그렇지 않은 집단에 비하여 사전적 탐지문항(가짜문항) 방식에 덜 탐지되었으며, 설문시간(지나치게 빨리 넘긴 경우 불성실 응답으로 간주)으로 탐지한 경우에도 덜 탐지된 것으로 나타났다.

최근의 연구들은 응답자들에게 개인화된 피드백(personalized feedback)을 제공하는 것에 관심을 두고 있다(Kühne and Kroh, 2018; Wenz et al., 2020). 이는 금전적 인센티브의 제공이 현실적으로 제한될 수 있음을 고려할 때, 매력적인 대안이 될 수 있다. Kühne and Kroh (2018)의 피드백이 응답행동에 미치는 영향에 관한 연구는 피드백 제공 여부에 따라 문항비율, 응답시간, 응답의 내적 일관성, 사회적으로 바람직한 응답성향 등의 차이가 존재하는지 분석하였다. 그 결과, 개인화된 피드백을 받은 집단이 설문이 더 재미있고, 덜 지루하였고, 스스로에 대해 더 잘 알게 되었다고 응답하였다. 아직 충분한 연구결과가 누적되지 않아 비금전적 인센티브가 불성실 응답에 미치는 영향에 대한 추가적인 확인이 필요하지만, 응답자들의 만족을 위해서 이러한 피드백 제공을 고려하는 것이 권장되는 바이다. '계량화된 나(quantified self)'라는 용어가 등장할 만큼(Swan, 2013), 경영 연구의 대상이 되는 응답자들은 스스로에 대하여 어느 때보다 높은 관심을 가진다. 따라서 이들에게 개인화된 정보를 제공하는 것은 향후 성실한 응답을 유도함에 있어 좋은 동기 부여 방법이 될 수 있다.

5.5 기타

불성실 응답을 유발하는 주요 원인인 동기와 응답자 부담 외에도 상황적 맥락에 의해 온라인 설문에서 불성실 응답이 응답자 의도와 상관없이 발생할 수 있다. 이미 선행연구들은 오프라인 설문조사보다 온라인 설문조사에서 응답자들의 주의가 더 쉽게 분산된다고 보고하였으며 이를 예방하는 방안에 대한 논의가 진행되고 있다(Ward and Pond, 2015; Gibson, 2019). 먼저, 온라인 설문조사의 경우, 사진과 같은 시각적 요소를 사용하여 응답자의 흥미와 집중력을 향상시켜서 불성실 응답을 예방할 수 있다(Ward and Pond, 2015). 예를 들어, Ward and Pond(2015)는 숨을 쉬고 눈을 깜빡이는 가상 인간 형태(virtual human presence)를 온라인 설문에 삽입해 연구자와 응답자 간 상호작용을 실시할 경우, 응답자가 설문에 좀 더 몰입하여 불성실 응답을 예방할 수 있다고 보고하였다. 한편, 나이가 어린 응답자들이 온라인 설문조사에 참여할 경우 불성실 응답빈도가 증가하는 것으로 보고되었기에 특정 대상이나 세대를 고려한 세심한 설계가 요구된다(Zwarun and Hall, 2014). 예를 들어, 온라인 설문조사 시 한 문항 당 응답시간을 다른 표본들과 비교하여 과도하게 짧거나 길게 응답한 데이터는 제거해야 하는 방법을 사용할 수 있다.

이상을 종합하여, 불성실 응답을 예방하는 설문조사의 설계를, (1) 응답자 동기를 고양시키는 연구 지시문의 작성, 설문지 구성, 인센티브와 (2) 응답자 부담을 완화하는 측면에서의 개별 문항의 설계, 설문지 구성과 (3) 기타의 온라인 설문의 상황적 요인으로 구분하여 <Table 7>에 제시하였다.

<Table 7> 주요 원인에 따른 불성실 응답을 예방하는 설계

주요 원인에 따른 대처방안	구체적 설계	관련 연구
동기	연구 지시문의 작성	• 긍정적 지시문(예, '조직구성원들의 응답이 조직의 중요한 의사결정에 영향을 미친다.' 혹은 '귀하가 응답하신 설문은 향후 이런 주제의 연구 발전에 크게 기여할 것이다.') Gibson(2019), Gideon(2012), Kanfer et al.(2017), Marshall(2019), Ward and Meade(2018)
		• 경고형 지시문(예, '응답이 성의 없을 경우 통계처리에서 제외될 수 있다.', '불성실 응답으로 간주될 경우, 인센티브 지급에서 제외될 수 있다.', '불성실 응답으로 간주될 경우, 학점에서 불이익을 받을 수 있다.') Fan et al.(2012), Gibson and Bowling(2020), Huang et al.(2012), Meade and Craig(2012)
		• 상호작용 지시문(예, '응답자에게 설문의 내용과 질에 대한 피드백을 제공할 것이다.', '설문 후 불성실 응답이 아닌 경우 엑스트라 인센티브를 제공한다.') Gibson(2019), Ward and Meade(2018), Ward and Pond(2015)
	설문지 구성	• 흥미와 집중력을 높이는 시각적 요소 사용(예, 사진 혹은 가상의 인간 형태) Deutskens et al.(2004), Ward and Pond(2015)
	인센티브	• 금전적 인센티브 제공(예, 선지급형 인센티브, 현금 지급, 전원에게 인센티브 지급) Gibson(2016), Singer and Ye(2013)
• 비금전적 인센티브 제공(예, 명시적인 피드백) Kühne and Kroh, (2018), Wenz et al.(2020)		
응답자 부담 완화	개별 문항의 설계	• 적절한 개별 문항의 길이 Groves(1989)
		• 과하지 않은 척도 사용(예, 7점 척도) Givon and Shapira(1984)
		• 중간 보기 활용(예, '보통이다.') Krosnick and Berent(1993), O'Muircheartaigh et al.(1999)
		• 개방형 문항 활용(예, 나이) Mondak(2001)
		• 선택형 문항 활용(예, 재산, 근로소득, 특정 기간 동안 참여횟수 등 응답자들이 정확한 수치를 환산하기 어려워 어렵게산으로 대답할 확률이 높은 문항) Huttenlocher et al.(1990)
	• 역코딩 문항 사용 주의 Curran(2016)	
설문지 구성	• 적절한 설문길이(예, 전체 설문시간이 10분이내) Bäckström and Björklund(2019), Gibson and Bowling(2020)	
	• 공통된 주제를 가진 문항 배치 Couper et al.(2001), Tourangeau et al.(2004)	
기타	개인 특성 문화적 요인	• 조사 설계 과정에서 대처방안을 고안할 수 없는 요인
	인구통계학적 특성	• 연구 대상(장애인, 어린이 등)에 따른 세심한 설문 설계 Carrier et al.(2009), Zwarun and Hall(2014)
	온라인 설문 (상황적 요인)	• 사진과 같은 시각적 요소를 온라인 설문에 삽입하여 응답자의 흥미와 집중력을 향상 Breitsohl and Steidelmüller(2018), Meade and Craig(2012), Zwarun and Hall(2014) • 인터넷에서 설문문항들이 한 화면에 한 개씩 제시된 것 보다, 한 화면에 한꺼번에 제시

VI. 논의 및 결론

본 연구는 최근 10년 이내의 경영학 분야 일곱 개 학술지에서 설문조사 방법을 사용한 논문을 대상으로 불성실 응답에 대한 인식 및 대처수준을 분석하고, 문헌연구와 반구조화된 심층면접을 통해 불성실 응답의 주요 기저 원인과 대처방안에 대하여 고찰한 후, 조사설계 수준에서 활용 가능한 예방방안을 소개하였다. 이는 다음과 같은 학술적 의의와 시사점을 지니고 있다.

첫째, 본 연구는 ‘불성실 응답의 예방’이라는 아직은 익숙하지 않은 주제를 세 가지 연구방법론(실태 분석, 문헌조사, 심층 인터뷰)으로 다각도에서 분석하여 새로운 주제에 익숙해질 수 있는 계기를 마련하였다는 점에서 의의가 있다. 독자들은 각자 익숙한 접근 방식을 통해 ‘불성실 응답의 예방’에 대한 관심과 이해를 증진시킬 수 있을 것이다. 특히, 본 연구는 2,400편에 달하는 논문을 분류 및 정량화하여 국내 경영학계의 불성실 응답에 대한 인식 및 대처수준을 직관적으로 파악하는 한편, 전문가 인터뷰를 통해 이를 심도 있게 분석하였다. 이와 같은 다각도적인 접근은 독자들이 하여금 현상을 체계적으로 이해하도록 도움으로써, 불성실 응답의 예방에 대한 관심과 필요성을 독려하는 데 기여한다.

분석 결과에 따르면, 전체 조사 대상 중 불성실 응답 예방 및 탐지방법을 구체적으로 언급한 사례는 2.8%에 불과하며, 반대로 아예 데이터의 정제과정에 대해 언급하지 않은 논문은 28.7%에 달했다. 실제 검토과정을 거쳤으나 논문에 명시하지 않은 경우를 감안하더라도, 불성실 응답을 명확하게 인식하고 대처하며 그 과정을 투명하게 밝히는 과정은 아직까지 익숙하지 않다고 결론지을 수 있다. 연구방법론

에 있어서 현재는 보편화된 동일방법편의(common method bias)의 문제에 관하여 국내 학계에 처음으로 논의를 시도했던 Park et al.(2007)의 연구처럼, 차츰 연구자들 사이에 관심이 확대되고 공감대가 형성되면, 향후에는 대다수의 설문연구에서 불성실 응답을 예방하고 탐지하는 장치들과 그 활용에 대한 학계의 기준점을 찾아나갈 수 있을 것이다.

둘째, 응답의 양(quantity)에서 질(quality)로 연구 관심의 전환을 이끌었다는 점에서 본 연구의 또 다른 주된 의의가 존재한다. 그동안 대단위 설문조사에서 패널 유지를 위한 충분한 응답 수 확보(Singer, 2002), 통계적인 측면에서 검정력 확보를 위한 충분한 응답 확보 필요성 등은 이미 충분히 논의된 바 있으나(Price et al., 2005), 개별 응답의 품질에 대한 논의를 충분하지 않았던 것이 사실이다(Curran, 2016; DeSimone et al., 2018). 연구, 교육 및 비즈니스 환경을 막론하고, 코로나(COVID-19) 시대를 거치며, 대면 설문조사가 온라인 또는 모바일 설문조사로 대체되는 속도는 가속화 될 것이다(Fleischer et al., 2015). 이러한 비대면 설문조사의 확대는 설문의 양을 확보하기는 용이해지나, 응답의 질을 보장받기는 어려워진다는 의미이기도 하다. 예를 들어, 설문연구 참여에 사전동의한 일정 패널들이 소정의 보상을 받고 연구에 참여하는 아마존(Amazon)사의 M-Turk은 활용방법에 대한 별도의 리뷰 연구(Keith et al., 2017)가 나올 만큼 학술분야에서의 활용이 보편화되었다. 하지만 최근 부주의한 응답으로 인한 ‘M-Turk의 위기(M-Turk Crisis)’를 제목으로 한 연구가 발간될 만큼(Chmielewski et al., 2020) 응답의 질적인 측면에 대한 경각심 역시 늘어나고 있다. 따라서, 설문연구 참여자들이 충분히 성실한 응답을 하는지 점검하고, 예방하는 것의 중요성이 점점 더 확대되어 갈

것이기에, 국내에서도 본 연구가 그 초석으로서의 의의를 지닌다.

셋째, 참고하여 활용할 수 있는 불성실 응답 예방 방안을 제시했다는 점에서 의의가 있다. 후속 연구자들이 연구 설계 시 효율적으로 원인과 예방방안을 파악할 수 있도록 <Table 8>에 이를 종합하여 제공하였다. 이는 설문 설계 시의 활용과 더불어, 학술적인 추가연구의 가능성도 열어주고 있다. 특히, 불성실 응답의 주된 요인을 동기별로 분류하여 제공한 것은 다차원적 접근을 가능케 한다는 점에서 장점이 있다. 예를 들어, 자극추구 동기와 사회적 동기, 응답자의 부담 중 어떤 것이 불성실 응답에 미치는 영향이 더 큰지, 그 효과는 개인의 특성에 따라 어떻게 달라지는지 검증해 볼 수 있을 것이다.

동시에 이 부분은 본 연구가 지닌 고유한 의의와 시사점에도 불구하고, 단일 연구로 온전히 명쾌하고 통일된 예방방안을 제시하기 어렵다는 점에서 연구의 한계점으로도 남아있다. 사전 예방방안 대신 다양한 사후 대처방안을 소개한 Curran(2016)의 연구에서도, 특정 한 가지 방법을 제시하기보다 각 연구설계에 사용하여서는 안 되는 방법들을 제시하였다. 이는 연구들의 다양한 환경과 맥락을 고려했을 때, 하나의 최적(best-fit) 방안을 제시하는 것 자체가 적합하지 않을 수 있음을 시사한다. 따라서 예방방안에 있어서 보다 많은 연구결과들이 누적된다면, Curran(2016)의 접근처럼, 연구의 각 상황에 맞춘 권고안을 제시할 수 있을 것이다.

또한, 기대했던 바와 달리, 본 연구에서 수십 년간 실무와 연구를 이어온 전문가들과의 심층 인터뷰에서 통일될 만한 권고안을 도출하지 못한 점은 아쉬움으로 남는다. 이는 주로 각 영역의 전문가들이 서로 상이한 학술분야에 종사하였음에 귀인한다. 불성실 응답을 정의하고 받아들이는 태도 또한 분야에

따라서 크게 상이했는데, 예로 교육학 분야의 경우, 학교환경에서의 설문은 학생들의 인정받으려는 욕구로 인해 불성실 응답이 발생하지 않는다(면접자 C)라고 조언했다. 반면 사회학 분야의 경우, 성실 응답을 위해 인구 전체를 대상으로 하는 정교한 계층 표집이 필요함을 강조하였으며, 개별 응답의 질에 대해서는 많은 의견을 표하지 않았다(면접자 D). 이러한 결과를 바탕으로 보아, 향후 연구에서는 전문가 집단에 대한 의견을 구할 때 조사대상을 설문조사의 대상과 형태, 실시목적이 비교적 유사한 영역(예, 마케팅, 조직행동)으로 한정하면 보다 명확한 결과를 얻을 수 있을 것으로 기대한다.

한편, 본 연구에서 명쾌한 해답을 제시하지 못한 또 다른 원인은 불성실 응답을 직접으로 다룬 실증연구가 충분히 누적되지 않았기 때문으로 볼 수 있다. 예를 들어, 설문의 길이-연구 참여율을 다룬 연구는 누적되어 있으나(Galesic and Bosnjak, 2009; Groves, 1989), 설문의 길이-불성실 응답의 발생을 다룬 연구는 충분치 않다. 마찬가지로, 연구설계에 있어서의 다른 여러 가지 요인들(인센티브의 크기, 설문 환경, 응답 시간)과 불성실 응답에 관한 연구 또한 그 수가 충분하지 않다. 따라서, 인센티브와 불성실 응답 간의 관계를 실험을 통해 밝힌 Shamon and Berning(2020)의 연구처럼, 본 연구에서 소개된 예방방안의 효과성을 입증할 수 있는 다양한 실험 및 실증연구들이 향후에 이루어져야 할 것이다. 즉, 본 연구가 다양한 백신의 존재를 소개하는 ‘백서’와 같다면, 앞으로는 개별 백신의 효과성을 차근차근 검증해나갈 필요성이 있다. 예를 들어, 자기중심성이 강화된 MZ세대를 대상으로 한 설문조사에서 개인맞춤형 피드백의 제공이 불성실 응답의 발생을 예방하는지, 그 효과는 금전적 인센티브 제공을 얼마나 대체할 수 있는지 등을 실험을 통해 확인해볼

<Table 8> 불성실 응답이 발생하는 주요 원인에 따른 대처방안

주요 원인	불성실 응답 발생 이유	대처방안	구체적 설계
동기	참여 동기 설문의 내용에 주의를 기울이지 않아 발생	지시문	• 긍정적 지시문, 완곡한 경고형 지시문, 상호작용 지시문
		인센티브	• 선지급형 인센티브 활용, 현금 지급, 응답자 모두에게 인센티브 지급
	자극 추구 동기 설문에 대한 흥미, 관심, 재미를 충족하지 못해 발생	지시문	• 긍정적 지시문
		설문지 구성	• 사진과 같은 시각적 요소를 사용, 상호작용을 높이기 위한 가상 사람 형태 활용하며 중요 질문을 설문지의 초반에 배치하여 응답의 질이 떨어지지 않도록 세심한 설계
사회적 동기 설문자와 응답자 사회적 상호작용의 부족으로 인해 발생	지시문	• 긍정적 지시문(예, '설문 후 불성실 응답이 아닌 경우 추가 인센티브를 제공한다.') • 경고형 지시문(예, '불성실 응답으로 간주될 경우, 학점에서 불이익을 받을 수 있다.') • 상호작용 지시문(예, '응답자에게 설문지의 내용과 질에 대한 피드백을 제공할 것이다.')	
	설문지 구성	• 사진과 같은 시각적 요소를 사용, 상호작용을 높이기 위한 가상 존재 활용	
공격성 동기 공격성 동기로 인해 능동적 불성실 응답 행동이 발생	설문지 구성	• 중복된 항목이나 불필요한 항목을 넣는 것과 같이 응답자의 공격성 동기가 점화되지 않도록 신중하게 설문 구성	
응답자 부담	응답자의 피로가 누적되어 발생	개별 문항 설계	• 몰입을 높이는 개별 문항의 설계, 적절한 설문길이, 과하지 않은 척도 사용, 중간 보기 활용, 개방형 문항 활용, 선택형 문항 활용, 역코딩 문항 사용 주의
		설문지 구성	• 적절한 설문길이, 공통된 주제를 가진 문항 배치
기타	온라인 설문과 같은 상황적 요인에 의해 발생	온라인 설문 시 몰입감 유도	• 사진과 같은 시각적 요소와 가상의 존재를 온라인 설문에 삽입하여 응답자의 흥미와 집중력을 향상 • 인터넷에서 설문문항들이 한 화면에 한 개씩 제시된 것 보다, 한 화면에 한꺼번에 제시

수 있을 것이다.

마지막으로, 본 연구에서는 불성실 응답의 기존 개념을 활용하고 간략하게 다루었으나, 개념적 또는 이론적 측면에서의 불성실 응답에 대한 논의도 지속적으로 발전시켜나갈 필요가 있다. 본 연구에서는 기존 학자들의 정의에 따라 예방방안을 도출하였으나, 불성실 응답의 정의를 확장할 경우 또 다른 예방방안이 나올 수도 있다. 예를 들어, '사회적 바람직성

(social desirability)'은 설문조사에서 전통적으로 고려되어 온 이슈이다(Nederhof, 1985). '사회적 바람직성'을 고려한 응답과 불성실 응답은 어떠한 개념적 구분과 중첩이 있는지, 과도하게 긍정적인 응답 또한 불성실한 응답으로 간주해야 한다면 이를 예방하는 방안이 무엇일지에 대한 별도의 고민도 필요할 것이다.

국내외를 막론하고 응답의 질과 불성실 응답에 관

한 관심이 대폭 증대되고 있다는 점은 긍정적이다 (Kang et al., 2020; Goldammer et al., 2020). 근래 발행되는 실증연구에서는 불성실 응답을 예방하고 검증하는 도구들을 사전에 설계하고(예, 주의 점검 문항), 이를 통해 탐지된 불성실 응답을 분석 시에 적절히 걸러내고, 그러한 과정을 투명하게 원고에 담은 사례들이 점점 더 자주 관찰되고 있다(예, Fleming et al., 2020). 본 연구가 이러한 변화의 흐름에 마중물이 될 수 있기를 바라며, 사회과학 전반에 있어 명확하고 재현 가능한 사회과학적 발견에 기여할 수 있기를 기대한다.

참고문헌

- Akbulut, Y.(2015), "Predictors of Inconsistent Responding in Web Surveys," *Internet Research*, 25(1), pp.131-147.
- Anduiza, E. and C. Galais(2017), "Answering Without Reading: IMCs and Strong Satisficing in Online Surveys," *International Journal of Public Opinion Research*, 29(3), pp.497-519.
- Bäckström, M. and F. Björklund(2019), "Is Reliability Compromised Towards the End of Long Personality Inventories?," *European Journal of Psychological Assessment*, 35(1), pp.14-23.
- Beach, D. A.(1989), "Identifying the Random Responder," *The Journal of Psychology*, 123(1), pp.101-103.
- Berry, D. T., M. W. Wetter, R. A. Baer, L. Larsen, C. Clark, and K. Monroe(1992), "MMPI-2 Random Responding Indices: Validation Using a Self-Report Methodology," *Psychological Assessment*, 4(3), pp.340-345.
- Blair, E., S. Sudman, N. M. Bradburn, and C. Stocking(1977), "How to Ask Questions about Drinking and Sex: Response Effects in Measuring Consumer Behavior," *Journal of Marketing Research*, 14(3), pp.316-321.
- Bowling, N. A., J. L. Huang, C. B. Bragg, S. Khazon, M. Liu, and C. E. Blackmore(2016), "Who Cares and Who Is Careless? Insufficient Effort Responding as a Reflection of Respondent Personality," *Journal of Personality and Social Psychology*, 111(2), pp.218-229.
- Breitsohl, H. and C. Steidelmüller(2018), "The Impact of Insufficient Effort Responding Detection Methods on Substantive Responses: Results from an Experiment Testing Parameter Invariance," *Applied Psychology*, 67(2), pp. 284-308.
- Brower, C. K.(2018), *Too Long and Too Boring: The Effects of Survey Length and Interest on Careless Responding* [Master's thesis, Wright State University], CORE Scholar. https://corescholar.libraries.wright.edu/etd_all/1918(retrieved February 2021).
- Carrier, L. M., N. A. Cheever, L. D. Rosen, S. Benitez, and J. Chang(2009), "Multitasking Across Generations: Multitasking Choices and Difficulty Ratings in Three Generations of Americans," *Computers in Human Behavior*, 25(2), pp.483-489.
- Camus, K. A.(2015), *Once Careless, Always Careless? Temporal and Situational Stability of Insufficient Effort Responding (IER)* [Master's thesis, Wright State University], CORE Scholar. https://corescholar.libraries.wright.edu/etd_all/1617(retrieved February 2021).

- Cantor, D., B. C. O'Hare, and K. S. O'Connor (2008), "The Use of Monetary Incentives to Reduce Nonresponse in Random Digit Dial Telephone Surveys," *Advances in Telephone Survey Methodology*, pp.471-498.
- Chmielewski, M. and S. C. Kucker(2020), "An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results," *Social Psychological and Personality Science*, 11(4), pp.464-473.
- Couper, M. P.(2008), *Designing Effective Web Surveys*, New York, Cambridge University Press.
- Couper, M. P., M. W. Traugott, and M. J. Lamias (2001), "Web Survey Design and Administration," *Public Opinion Quarterly*, 65(2), pp.230-253.
- Curran, P. G.(2016), "Methods for the Detection of Carelessly Invalid Responses in Survey Data," *Journal of Experimental Social Psychology*, 66, pp.4-19.
- DeSimone, J. A., H. K. Davison, J. L. Schoen, and M. N. Bing(2020), "Insufficient Effort Responding as a Partial Function of Implicit Aggression," *Organizational Research Methods*, 23(1), pp.154-180.
- DeSimone, J. A., A. J. DeSimone, P. D. Harms, and D. Wood(2018), "The Differential Impacts of Two Forms of Insufficient Effort Responding," *Applied Psychology*, 67(2), pp.309-338.
- Deutskens, E., K. De Ruyter, M. Wetzels, and P. Oosterveld(2004), "Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study," *Marketing Letters*, 15(1), pp.21-36.
- Dickinson, T. L. and P. M. Zellinger(1980), "A Comparison of the Behaviorally Anchored Rating and Mixed Standard Scale Formats," *Journal of Applied Psychology*, 65(2), pp. 147-154.
- Diefendorff, J. M. and M. M. Chandler(2011), "Motivating Employees," In S. Zedeck (Ed.), *APA Handbook of Industrial and Organizational Psychology, Vol 3: Maintaining, Expanding, and Contracting the Organization*, Washington, DC, American Psychological Association, pp.65-135.
- Dunn, A. M., E. D. Heggstad, L. R. Shanock, and N. Theilgard(2018), "Intra-Individual Response Variability as an Indicator of Insufficient Effort Responding: Comparison to Other Indicators and Relationships with Individual Differences," *Journal of Business and Psychology*, 33(1), pp.105-121.
- Fan, J., D. Gao, S. A. Carroll, F. J. Lopez, T. S. Tian, and H. Meng(2012), "Testing the Efficacy of a New Procedure for Reducing Faking on Personality Tests Within Selection Contexts," *Journal of Applied Psychology*, 97(4), pp.866-880.
- Fleischer, A., Mead, A. D., and J. Huang(2015), "Inattentive responding in MTurk and other online samples," *Industrial and Organizational Psychology*, 8(2), pp.196-202.
- Fleming, A. C., Hlebasko, H., Adams, S. C., Roach, K. N., and N. D. Christiansen(2020), "Effects of sexism and job - applicant match on leadership candidate evaluations," *Social Behavior and Personality: An International Journal*, 48(9), pp.1-8.
- Galesic, M.(2006), "Dropouts on the Web: Effects of Interest and Burden Experienced During an Online Survey," *Journal of Official Statistics*, 22(2), pp.313-328.

- Galesic, M., and M. Bosnjak(2009), "Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey," *Public Opinion Quarterly*, 73(2), pp.349-360.
- Gehlbach, H. and S. Barge(2012), "Anchoring and Adjusting in Questionnaire Responses," *Basic and Applied Social Psychology*, 34(5), pp.417-433.
- Gibson, A.(2016), *Stop With the Questions Already! The Effects of Questionnaire Length and Monetary Incentives on Insufficient Effort Responding* [Master's thesis, Wright State University], CORE Scholar, https://corescholar.libraries.wright.edu/etd_all/1656(retrieved February 2021).
- Gibson, A.(2019), *Stop What You're Doing, Right Now! Effects of Interactive Messages on Careless Responding* [Doctoral dissertation, Wright State University], CORE Scholar, https://corescholar.libraries.wright.edu/etd_all/2119/(retrieved February 2021).
- Gibson, A. M. and N. A. Bowling(2020), "The Effects of Questionnaire Length and Behavioral Consequences on Careless Responding," *European Journal of Psychological Assessment*, 36(2), pp.410-420.
- Gideon, L.(2012), *Handbook of Survey Methodology for the Social Sciences*, New York, Springer.
- Givon, M. M. and Z. Shapira(1984), "Response to Rating Scales: A Theoretical Model and Its Application to the Number of Categories Problem," *Journal of Marketing Research*, 21(4), pp.410-419.
- Goldammer, P., Annen, H., Stöckli, P. L., and K. Jonas(2020), "Careless responding in questionnaire measures: Detection, impact, and remedies," *The Leadership Quarterly*, 31(4), 101384.
- Göritz, A. S.(2004), "The Impact of Material Incentives on Response Quantity, Response Quality, Sample Composition, Survey Outcome, and Cost in Online Access Panels," *International Journal of Market Research*, 46(3), pp.327-345.
- Grau, I., C. Ebbeler, and R. Banse(2019), "Cultural Differences in Careless Responding," *Journal of Cross-Cultural Psychology*, 50(3), pp. 336-357.
- Groves, R. M.(1989), *Survey Errors and Survey Costs*. New York, John Wiley and Sons.
- Huang, J. L., N. A. Bowling, M. Liu, and Y. Li (2015), "Detecting Insufficient Effort Responding with an Infrequency Scale: Evaluating Validity and Participant Reactions," *Journal of Business and Psychology*, 30(2), pp.299-311.
- Huang, J. L., P. G. Curran, J. Keeney, E. M. Poposki, and R. P. DeShon(2012), "Detecting and Detering Insufficient Effort Responding to Surveys," *Journal of Business and Psychology*, 27(1), pp.99-114.
- Huang, J. L., M. Liu, and N. A. Bowling(2015), "Insufficient Effort Responding: Examining an Insidious Confound in Survey Data," *Journal of Applied Psychology*, 100(3), pp. 828-845.
- Huttenlocher, J., L. V. Hedges, and N. M. Bradburn (1990), "Reports of Elapsed Time: Bounding and Rounding Processes in Estimation," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(2), pp.196-213.
- Johnson, J. A.(2005), "Ascertaining the Validity of Individual Protocols from Web-Based Per-

- sonality Inventories," *Journal of Research in Personality*, 39(1), pp.103-129.
- Joshi, A., S. Kale, S. Chandel, and D. K. Pal(2015), "Likert Scale: Explored and Explained," *Current Journal of Applied Science and Technology*, 7(4), pp.396-403.
- Jun, J., T. Lee, S. Chung, and C. Park(2017), "A Study on Determinants of Intention to Use Car Sharing: Moderating Effect of Consumer Innovativeness," *Journal of Marketing Management Research*, 22(2), pp.49-66.
- Kanfer, R., M. Frese, and R. E. Johnson(2017), "Motivation Related to Work: A Century of Progress," *Journal of Applied Psychology*, 102(3), pp.338-355.
- Kang, C. W., S. B. Choi, and M. S. Kwon(2020), "False Responses Detection of Survey System using Machine Learning," *Journal of The Korean Data Analysis Society*, 22(3), pp. 1013-1021.
- Kang, W., C. Yoo, and D. Kim(2015), "The Effects of Negative Publicity on Brand Evaluation: Moderating Roles of Brand Strategy and Prior Attitude," *Journal of Marketing Management Research*, 20(3), pp.65-99.
- Keith, M. G., Tay, L., and P. D. Harms(2017), "Systems perspective of Amazon Mechanical Turk for organizational research: Review and recommendations," *Frontiers in Psychology*, 8, 1359.
- Kim, Y., J. Dykema, J. Stevenson, P. Black, and D. P. Moberg(2019), "Straightlining: Overview of Measurement, Comparison of Indicators, and Effects in Mail - Web Mixed-Mode Surveys," *Social Science Computer Review*, 37(2), pp. 214-233.
- Krosnick, J. A.(1991), "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys," *Applied Cognitive Psychology*, 5(3), pp.213-236.
- Krosnick, J. A. and M. K. Berent(1993), "Comparisons of Party Identification and Policy Preferences: The Impact of Survey Question Format," *American Journal of Political Science*, 37(3), pp.941-964.
- Kühne, S. and M. Kroh(2018), "Personalized Feedback in Web Surveys: Does it Affect Respondents' Motivation and Data Quality?," *Social Science Computer Review*, 36(6), pp.744-755.
- Laurent, A.(1972), "Effects of Question Length on Reporting Behavior in the Survey Interview," *Journal of the American Statistical Association*, 67(338), pp.298-305.
- Lee, Y. S., J. Y. Lee, and K. T. Lee(2008), "Amounts of Responding Times and Unreliable Responses at Online Surveys," *Survey Research*, 9(2), pp.51-83.
- Liu, M., N. A. Bowling, J. L. Huang, and T. A. Kent(2013), "Insufficient Effort Responding to Surveys as a Threat to Validity: The Perceptions and Practices of SIOP Members," *The Industrial-Organizational Psychologist*, 51(1), pp.32-38.
- Maniaci, M. R., and R. D. Rogge(2014), "Caring about Carelessness: Participant Inattention and Its Effects on Research," *Journal of Research in Personality*, 48(1), pp.61-83.
- Marshall, A.(2019), *Caring More About Careless Responding: Applying the Theory of Planned Behavior to Reduce Careless Responding on Online Surveys* [Doctoral dissertation, Colorado State University], Mountain Scholar, <https://hdl.handle.net/10217/195385>(retrieved February 2021).

- McKay, A. S., D. M. Garcia, J. P. Clapper, and K. S. Shultz(2018), "The Attentive and the Careless: Examining the Relationship Between Benevolent and Malevolent Personality Traits with Careless Responding in Online Surveys," *Computers in Human Behavior*, 84, pp.295-303.
- Meade, A. W. and S. B. Craig(2012), "Identifying Careless Responses in Survey Data," *Psychological Methods*, 17(3), pp.437-455.
- Mondak, J. J.(2001), "Developing Valid Knowledge Scales," *American Journal of Political Science*, 45(1), pp.224-238.
- Nederhof, A. J.(1985), "Methods of coping with social desirability bias: A review," *European Journal of Social Psychology*, 15(3), pp. 263-280.
- Nguyen, H. L. T.(2017), *Tired of Survey Fatigue? Insufficient Effort Responding Due to Survey Fatigue* [Master's thesis, Middle Tennessee State University], JEWLScholar@MTSU, <http://jewlscholar.mtsu.edu/xmlui/handle/mtsu/5302>(retrieved February 2021).
- Nichols, A. L. and J. E. Edlund(2020), "Why Don't We Care More About Carelessness? Understanding the Causes and Consequences of Careless Participants," *International Journal of Social Research Methodology*, 23(6), pp. 625-638.
- O'Muircheartaigh, C. and P. Campanelli(1999), "A Multilevel Exploration of the Role of Interviewers in Survey Non response," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 162(3), pp.437-446.
- Palaniappan, K. and I. Y. S. Kum(2019), "Underlying Causes Behind Research Study Participants' Careless and Biased Responses in the Field of Sciences," *Current Psychology*, 38(6), pp. 1737-1747.
- Park, H-J. and S. H. Lee(2017), *Scientific Survey Method to Change the World*, Seoul, Republic of Korea, Chang-ji Co.
- Park, W-W., M. S. Kim, S. M. Jeong, and K. M. Huh(2007), "Cause and Remedies of Common Method Bias," *Korean Journal of Management*, 15, pp.89-133.
- Park, W-W., S. H. Mah, S. H. Bae, S. Y. Ji, Y. W. Lee, and J. Y. Kim(2020), "Careless Responding in Surveys: Detection Methods and the Impact of Screening on Research," *Korean Management Review*, 49(2), pp.331-364.
- Podsakoff, P. M., S. B. MacKenzie, J. Y. Lee, and N. P. Podsakoff(2003), "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88(5), pp.879-903.
- Price, J. H., Daek, J. A., Murnan, J., Dimmig, J., and S. Akpanudo(2005), "Power analysis in survey research: Importance and use for health educators," *American Journal of Health Education*, 36(4), pp.202-209.
- Reymond, P., H. Weber, M. Damond, and E. E. Farmer(2000), "Differential Gene Expression in Response to Mechanical Wounding and Insect Feeding in Arabidopsis," *The Plant Cell*, 12(5), pp.707-719.
- Rogelberg, S. G. and J. M. Stanton(2007), "Introduction: Understanding and Dealing with Organizational Survey Nonresponse," *Organizational Research Methods*, 10(2), pp.195-209.
- Saris, W. E. and I. N. Gallhofer(2014), *Design, Evaluation, and Analysis of Questionnaires*

- for *Survey Research*, New York, John Wiley and Sons.
- Schmitt, N. and D. M. Stults(1985), "Factors Defined by Negatively Keyed Items: The Result of Careless Respondents?," *Applied Psychological Measurement*, 9, pp.367-373.
- Seo, R. and J. Lee(2019), "How to Manage the Green-Eyed Monster: Money Priming May Modify the Effects of Jealousy on the Preferences for Attention-Grabbing and Luxury Products," *Korean Management Review*, 48 (5), pp.1221-1252.
- Shamon, H. and C. C. Berning(2020), "Attention Check Items and Instructions in Online Surveys with Incentivized and Non-Incentivized Samples: Boon or Bane for Data Quality?," *Survey Research Methods*, 14(1), pp.55-77.
- Shuman, H. and S. Presser(1981), *Questions and Answers in Attitude Surveys Experiments on Questions Forms and Context*. Thousand Oaks, CA, Sage.
- Singer, E.(2002), "The Use of Incentives to Reduce Nonresponse in Household Surveys," In R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little (Eds.), *Survey Nonresponse*, New York, Wiley, pp.163-177.
- Singer, E. and C. Ye(2013), "The Use and Effects of Incentives in Surveys," *The Annals of the American Academy of Political and Social Science*, 645(1), pp.112-141.
- Song, M. R., W. Chu, and M. Im(2018), "Psychological Characteristics and Usage Patterns of Electric Vehicle Owners in South Korea," *Korean Management Review*, 47(2), pp.271-306.
- Swan, M.(2013), "The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery," *Big Data*, 1(2), pp.85-99.
- Tourangeau, R., M. P. Couper, and F. Conrad(2004), "Spacing, Position, and Order: Interpretive Heuristics for Visual Features of Survey Questions," *Public Opinion Quarterly*, 68(3), pp.368-393.
- Tourangeau, R., E. Singer, and S. Presser(2003), "Context Effects in Attitude Surveys: Effects on Remote Items and Impact on Predictive Validity," *Sociological Methods and Research*, 31(4), pp.486-513.
- Vigil-Colet, A., D. Navarro-Gonzalez, and F. Morales-Vives(2020), "To Reverse or to not Reverse Likert-Type Items: That is the Question," *Psicothema*, 32(1), pp.108-114.
- Ward, M. K. and A. W. Meade(2018), "Applying Social Psychology to Prevent Careless Responding During Online Surveys," *Applied Psychology*, 67(2), pp.231-263.
- Ward, M. K., A. W. Meade, C. M. Allred, G. Pappalardo, and J. W. Stoughton(2017), "Careless Response and Attrition as Sources of Bias in Online Survey Assessments of Personality Traits and Performance," *Computers in Human Behavior*, 76, pp.417-430.
- Ward, M. K. and S. B. Pond III(2015), "Using Virtual Presence and Survey Instructions to Minimize Careless Responding on Internet-Based Surveys," *Computers in Human Behavior*, 48, pp.554-568.
- Weijters, B., A. De Beuckelaer, and H. Baumgartner (2014), "Discriminant Validity Where There Should Be None: Positioning Same-Scale Items in Separated Blocks of a Questionnaire," *Applied Psychological Measurement*, 38(6), pp.450-463.
- Wenz, A., A. Jackle, J. Burton, and M. P. Couper (2020), "The Effects of Personalized Feedback

- on Participation and Reporting in Mobile App Data Collection," *Social Science Computer Review*, Advance online publication, <https://doi.org/10.1177/0894439320914261> (retrieved February 2021).
- Woodside, A. G. and C. Dubelaar(2003), "Increasing Quality in Measuring Advertising Effectiveness: A Meta-Analysis of Question Framing in Conversion Studies," *Journal of Advertising Research*, 43(1), pp.78-85.
- Zwarun, L. and A. Hall(2014), "What's Going On? Age, Distraction, and Multitasking During Online Survey Taking," *Computers in Human Behavior*, 41, pp.236-244.

- The author Won-Woo Park received Bachelor and Master of Business Administration from Seoul National University, and Ph.D. (1989) from the University of Pittsburgh. He worked as a faculty member of some private universities of the USA and of South Korea before joining Seoul National University in 1998. He published about 130 research articles and 16 books, received teaching awards from both the college (Business School) and the university (SNU). He was the recipient of the '2020 Emerald Literati Award' from Emerald Publishing.
- The author Yoowoo Lee is a Ph.D. majoring in Vocational Education & Workforce Development at Seoul National University (SNU). She received her MA in majoring in organizational behaviors and human resource management in the School of Business Administration at Chungang University (CAU). Her research interests include age management, mid-career change, and survey research methodology.
- The author Sunghyuck Mah is a Ph.D. candidate majoring in organizational behaviors and human resource management at the School of Business Administration, Seoul National University (SNU). He earned his Bachelor and Master of Business Administration from Yonsei University and SNU, respectively. His research interests include empowering leadership, overqualification among young generations, and survey research methodology.
- The author Jayoung Kim is a Ph.D. candidate majoring in strategy and international management at the Graduate School of Business, Seoul National University. She received her BS in Housing & Interior Design from Yonsei University and MS in Business Administration from Seoul National University. Her current research interests include growth strategies of the platform business, entrepreneurship, and research methodology.
- The author Suhyun Bae is a Ph.D. majoring in Vocational Education & Workforce Development at Seoul National University (SNU). She received her BA in Psychology from the University at Albany - State University of New York, and MA in Education from Seoul National University. Major research areas include HRD, generational conflict, organizational culture, vocational education, and research methodology.
- The author Seonyoung Ji is a student of Organizational Behavior & Human Resources Ph.D. program at Purdue University, West Lafayette. She received her B.B.A. degree from Sungkyunkwan University and M.S. in Business Administration from Seoul National University. Her main research interests include meaningful work, positive organizational psychology, and research methodology.