

## 메타분석 결과에서 통계적 오류의 영향: 헤지스·올킨 방법과 헌터·슈미트 방법의 비교 분석\*

이병관\*\*

한양대학교 광고홍보학과 교수

손영곤\*\*\*

한양대학교 광고홍보학과 강사

본 연구는 메타분석의 대표적인 두 가지 방법인 헤지스와 올킨(Hedges & Olkin)이 제시하는 방식과 헌터와 슈미트(Hunter & Schmidt)의 방식간의 접근방법과 특징에 대해 메타분석 방법론적 측면에서 비교분석하기 위해 진행하였다. 메타분석 방법과 관련하여 두 방식간의 방법론적 기본 가정과 차이점을 주요 쟁점으로 살펴보았으며, 구체적인 산출 과정과 그 결과를 상호 비교하는데 초점을 맞춰 상이한 분석방법과 절차, 그리고 결과 해석의 차이점과 유의점 등을 검증하였다. 구체적으로 국내에서 2015년까지 수행된 인간 행동을 설명하기 위한 대표적인 사회인지이론 가운데 하나인 계획된 행동이론을 적용한 연구논문을 대상으로 헤지스와 올킨 메타분석방법과 헌터와 슈미트 메타분석방법에 기초하여 그 결과를 제시하였다. 표집오차 및 측정오차 보정, 신뢰구간과 확산구간, 피셔(Fisher) z 변환값 사용 여부, 효과크기 비교, 동질성 검사 여부 등 방법론적 차이에 대해서도 논의하였다.

주제어 : 메타분석, 통계적 오류, 표집오차 보정, 측정오차 보정, 계획된 행동이론

\* 이 논문은 2015년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2015S1A5A2A01009538).

\*\* gogreen@hanyang.ac.kr, 주저자

\*\*\* noesis4@naver.com, 교신저자

## 1. 들어가는 말

현대적 의미에서 메타분석은 “연구결과들을 통합할 목적으로 수행되는 개별 연구결과들에 대한 통계적 분석”(Glass, 1976, p.4)이다. 개별적으로 진행되는 연구들은 각기 다른 연구대상 및 연구환경에서 수행되며, 연구결과 또한 일관되게 나타나기보다 서로 다른 결과를 제시하는 경우가 많다. 메타분석은 연구결과가 하나의 방향으로 수렴되지 않고 제각각의 양상을 보일 때, 이를 통합해서 바라볼 필요성에서 시작되었다. 지식이 발전함에 따라 관련 연구물도 증가하게 되고, 데이터베이스를 통해 각종 문헌과 연구결과에 대한 접근과 검색 용이성은 메타분석을 한층 발전시키는데 기여한 바 크다. 대표적인 연구물 검색 데이터베이스인 ISI의 Web of science에 나타난 바에 따르면, 1997년 이전까지 2,000편도 안되던 메타분석 연구는 2012년 말 95,719편으로 최근 5년 동안 연 18.86% 성장하여 양적으로 크게 팽창했음을 알 수 있다(Zhu, Jiang, Cao, Sun, Chen, & Liu, 2015).

이러한 메타분석 연구의 발전은 국내외를 막론하고 의학, 심리학, 교육학, 사회학 등 다양한 학문 영역에서 관심의 대상이 되고 있다. 보건의료 분야의 전세계적인 조직체인 코크렌 데이터베이스 (<http://www.cochranelibrary.com>)에 의하면, 2016년 12월 현재까지 수록된 연구결과물 총 9,604건 가운데 2,234개가 ‘meta-analysis’와 관련된 연구 결과임을 찾을 수 있었다. 커뮤니케이션 연구 분야만을 놓고 보더라도 외국의 경우에는 상당히 많은 메타분석 연구들이 수행된 바 있다. 그러나 국내 커뮤니케이션 연구에 있어 메타분석을 활용한 연구들은 많지 않은 것이 사실이다. 국내 커뮤니케이션 연구에서도 여러 편의 메타분석 시도가 있었지만(예: 김광재, 2010; 김유경·강형구·최일도, 2008; 류준호·윤승금·이준호, 2010; 손영곤, 2014; 손영곤·이병관, 2012; 유형진, 2005; 이병관 외, 2013, 2014; 전근영·이정교, 2015; 황

상재·박석철, 2004; 황용석, 2006), 최근 몇 편의 연구(예: 손영곤, 2014; 손영곤·이병관, 2012; 이병관 외, 2013, 2014; 전근영·이정교, 2015)를 제외하곤 실질적인 접근이나 방법은 진정한 의미의 메타분석이라기보다는 '체계적 문헌 분석(systematic review)'으로 보아야 할 것이다. 메타분석의 가장 큰 특징이자 장점은 효과크기(effect size)를 계산해서 중재의 방향과 크기를 제공해주는데 있다고 할 때(Cooper, 2009), 국내 커뮤니케이션 분야에서 이루어진 메타분석 연구의 상당수는 효과크기를 통한 통합적 분석이라고 말할 수 없기 때문이다. 메타분석이라고 칭하면서 효과크기를 계산하지 않는 연구는 엄밀한 의미의 메타분석과는 구분하여 체계적 문헌 분석과 같이 다른 용어를 사용하는 것이 바람직하다(신인수·박은영, 2011).

오늘날 메타분석의 접근방법은 크게 두 가지로 집약된다. 하나는 헤지스와 올킨(Hedges & Olkin, 1985)이 제시하는 방법(이하 HO방법)이고, 다른 하나는 헌터와 슈미트(Hunter & Schmidt, 2004) 방법(이하 HS방법)이 바로 그것이다<sup>1)</sup>. 이들 메타분석 방법은 그 역사나 접근방법 등 개념적으로 큰 대조를 보이고 있다(Johnson, Mullen, & Salas, 1995). 개별 효과크기를 계산할 때 HO방법은 상관계수를 피셔(Fisher)의  $z$ 값으로 변환하여 분석한 후 통합 효과크기를 산출할 때 이를 다시  $r$ 로 변환하는 과정을 거치는 것을 특징으로 한다. 이에 비해 HS방법은  $z$ 값으로 변환하지 않고 표집오차, 측정오차, 범위제한(range restriction) 등 다양한 통계적 오차를 고려하여 상관관계 값을 보정하고 이를 토대로 통합 효과크기를 산출하는 절차를 거친다. 어떤 방법

1) 이외에도 로젠탈(Rosenthal, 1991) 방법도 포함하여 메타분석의 3대 주류로 설명되기도 하지만(Johnson, Mullen, & Salas, 1995), 로젠탈(Rosenthal)이 제시하는 방법은 HO방법과 기본 가정에서 유사한 양상을 보인다. 이에 따라 최근에는 본 연구에서 다루고자 하는 두 가지 접근방법으로 분류하는 경우가 많다(예: Cooper, 2009, 2010; Lipsey & Wilson, 2001).

이 우수항가에 대해서는 지속적으로 논의가 거듭되고 있지만(Schulze, 2007), 메타분석 결과의 대부분은 이들 두 가지 방식 가운데 어느 하나를 차용하여 사용하고 있다. 국내에서 HO방법은 교육, 경영, 심리 등 연구영역 전반에 두루 적용되고 있으며(예: 김세경·천성문, 2015; 유성경·박은선·김수정·조효진, 2014; 조용래·노상선·조기현·홍세희, 2014), HS방법에 기초한 메타분석은 산업 및 조직심리 분야를 중심으로 다루어지고 있다(예: 박형인·남숙경·양은주, 2011; 오인수 외, 2007; 유태열, 2009; 최한나·권윤주·강귀옥, 2015; 허만용, 2008; 홍세정·장재윤, 2015). 이는 HS방법이 통계적 기본 가정에 충실하므로 복잡한 계산과정을 거칠 뿐만 아니라, 분석과정에서 발생하는 다양한 통계적 오류를 고려해 각 과정별로 보정하는 등 절차상의 어려움이 복합적으로 작용하고 있기 때문인 것으로 풀이된다.

헌터와 슈미트(Hunter & Schmidt, 2004)는 연구결과에서 나타나는 ‘본질적 특성이 아닌 인공적인(artifactual 혹은 man-made) 오차가 연구의 불완전성(imperfections)으로 인해 필연적으로 발생하게 되는데, 이를 ‘통계적 오류(statistical artifacts)’라고 언급하고 있다. 따라서 HS방법에서는 어떤 연구이든 완벽한 연구는 존재하지 않기 때문에 모든 개별 연구들은 다양한 형태의 오차를 포함하고 있으며, 결국 이러한 오차들 즉 통계적 오류가 메타분석의 결과에 영향을 미칠 수 있다고 주장한다(Hunter & Schmidt, 2004). HS방법을 지지하는 학자들은 다양한 통계적 오류가 실제 상관관계를 왜곡시키기 때문에 상관관계 연구들에 대한 메타분석에서 이를 보정하지 않을 경우, 실제 상관관계 분포의 직접적인 추정이 불가능해진다고 주장한다. 그러나 지금까지 여러 연구들(e.g., Callender & Osburn, 1980; Charles, 2005; Field, 2001; Hunter, Schmidt, & Le, 2006; Law, Schmidt, & Hunter, 1994; Le & Schmidt, 2006)이 시뮬레이션을 통해 통계적 오류, 특히 측정오차와 범위제한에 의한 통계적 오류의 보정 절차 및 결과에 대한 논의해왔지만, 그

결과는 연구마다 상이하게 나타났다. 예를 들면, 어떤 연구들(e.g., Callender & Osburn, 1980; Charles, 2005; Field, 2001)은 HS방법이 문제가 있다고 주장하는 반면, 또 다른 연구들(Hunter et al., 2006; Law et al., 1994; Le & Schmidt, 2006)은 HS방법이 메타분석 연구에서 효과크기를 추정하는데 있어 보다 더 정확한 절차라고 주장하기도 한다.

사실, HS방법에서는 이러한 통계적 오류로서 '표집오차(sampling error), 종속변인의 측정오차(measurement error), 독립변인의 측정오차, 연속적인 독립변인의 이분화(dichotomization), 연속적인 종속변인의 이분화, 독립변인의 범위제한(range restriction), 종속변인의 범위제한, 독립변인의 구성개념 타당도(construct validity)의 불완전성, 종속변인의 구성개념 타당도의 불완전성, 보고 및 자료입력 과정에서의 오차, 외적 요인에 의한 변량 등 총 11개의 통계적 오류들을 제시하고 있지만(Hunter & Schmidt, 2004), 이들 가운데 사회과학 분야의 개별 연구 논문에서 얻을 수 있는 정보들은 극히 제한적임을 유념할 필요가 있다. 특히 커뮤니케이션 연구 분야의 논문 중에서 각종 통계적 오류들을 보정하기 위한 기초 정보를 얻을 수 있는 것은 표집오차, 측정오차, 범위제한 등과 관련된 내용이 대부분인데, 그나마 범위 제한을 보정할 수 있는 정보<sup>2)</sup>를 제시한 개별 연구는 거의 없는 편이다. 따라

---

2) 범위제한을 보정하기 위해서는 기본적으로 제한되지 않은 표본 변인의 표준편차(standard deviation of unrestricted sample)를 알아야만 하는데, 현실적인 상황에서 이를 산출하는 것은 불가능하다. 알렉산더와 동료들(Alexander, Alliger, & Hanges, 1984)은 코헨의 비율(Cohen's ratio)을 사용하여 제한되지 않은 표본 변인의 표준편차를 제한된 표본 변인의 표준편차(standard deviation of restricted sample)를 통해서 추정할 수 있는 공식을 제안하였지만, 코헨의 비율을 산출하기 위해서는 측정변인의 범위가 반드시 제공되어야 한다(Cohen, 1959). 그러나 현실적으로 커뮤니케이션 연구 분야의 대부분의 연구들은 등간척도를 사용하고 있고, 이에 따라 연구자들 또한 측정변인의 범위를 보고하지 않는 것으로 보인다. 실제로 본 연구의 분석대상인 계획된 행동이론의 개별 연구에서는 측

서 본 연구는 먼저 메타분석의 기본 개념과 HO방법과 HS방법의 차이, 그리고 HS방법에서 제시하는 통계적 오류의 보정에 대해 살펴보고, 표집오차와 측정오차의 보정을 기초로 계획된 행동이론(theory of planned behavior: 이하 TPB)의 실제 자료(real-world data)를 통해 메타분석을 수행하는 두 가지 방법에 의한 분석 과정과 그 결과의 차이를 살펴보고자 한다. TPB의 경우 커뮤니케이션 연구 분야를 위시하여 다양한 영역에 걸쳐 인간 행동을 설명하기 위해 널리 사용되는 대표적인 사회인지적 이론의 하나이며, 주요 개념의 측정이나 조작적 정의에 대한 기본 가이드가 제시되고 있어 상대적으로 연구자의 자의적 해석은 적을 수 있다. 다시 말해, 본 연구는 시뮬레이션 자료가 아닌 실제 자료를 이용해, 통계적 오류의 보정에 따라 HO방법과 HS방법에 의한 메타분석 결과가 어떤 차이를 보이는지를 비교·분석함으로써 메타분석의 연구방법론적 함의를 논의하고자 한다.

## 2. 메타분석의 기본 개념 및 일반 절차

메타분석을 실시하기 위해서는 분석대상이 되는 기존의 선행연구 결과들이 다양한 형태로 제시하고 있는 통계치를 하나의 통일된 효과크기로 변환하는 것에서부터 시작한다. 효과크기는 각기 다른 통계치로 제시된 결과를 비교가능한 지수로 만들어 메타분석이 가능하게 해주는 기본 단위이다. 메타분석에서 사용되는 지수인 효과크기는 측정 척도에 대한 집단간 표준화된 평균 차이(standardized mean difference: Cohen's  $d$ , Hedges'  $g$  등)를 나타내는 효과크기( $ES_m$ ), 변인간의

---

정변인의 범위를 보고한 연구가 단 한 건도 나타나지 않았다. 따라서 커뮤니케이션 연구 분야의 메타분석에서 범위 제한을 고려한 보정은 현실적으로 불가능한 것으로 판단된다.

관계를 나타내는 상관계수(correlation coefficient)의 효과크기( $ES_c$ ), 그리고 두 개의 이항변인간의 관계를 나타내는 승산비(odds-ratio)의 효과크기( $ES_{or}$ ) 등으로 다양하다. 이밖에도  $R^2$ 나  $\eta^2$ ,  $\omega^2$  등도 효과크기로 사용할 수 있으나, 이들은 메타분석을 위한 효과크기로는 적절하지 않다(Cheung, 2015). 이들 개개의 지수는 개념적으로 관찰된 효과크기를 표준화하였다는 점에서 동일한 선상에서 비교가 가능하며, 서로 다른 효과크기로 변환할 수 있다. 특히 상관관계를 이용한 효과크기 계산 방법은 다른 방법에 비해 복잡하지 않고 두 변인간의 일정한 방향성을 파악하기 용이하다는 점에서, 그리고 해석하기 용이하다는 점에서 많이 이용되고 있다(Ellis, 2010; Rosenthal & DiMatteo, 2001).

메타분석은 대체로 여섯 단계를 거쳐 진행되며 일반적인 연구절차와 유사하다(오성삼, 2002; Cooper, 2010; Field & Gillett, 2010). 첫 번째 단계는 연구문제를 설정하는 단계로 연구주제에 대한 관심 변인의 설정, 변인의 정의 등을 하는 단계이다. 두 번째는 분석 대상 논문의 소재를 파악하고 연구 자료의 범위를 선정하여 자료를 수집하는 단계이다. 세 번째는 분석 자료를 부호화하고 연구 특성을 요약하는 단계이다. 네 번째는 분석이 가능한 변인들에 대한 부호화된 자료를 가지고 통계분석을 실시하는 단계이다. 메타분석에서는 기초 자료로 효과크기를 사용하는데, 이는 개별 연구결과들을 서로 비교할 수 있도록 고안된 표준화된 통계값을 의미한다. 효과크기 병합 방법, 동질성 검증, 출판편향, 조절변인에 의한 연구결과의 차이 분석 등을 하는 단계이다. 다섯 번째는 분석 결과들을 제시하고 해석하는 단계이며, 여섯 번째는 결과를 작성하는 단계이다. 메타분석의 일반 절차 중에서 첫째와 둘째, 그리고 셋째 절차는 HS방식이나 HO방식, 어떤 방식의 메타분석을 하더라도 차이가 없다. 본 연구에서는 넷째 단계, 즉 개별 효과크기를 산출하고 이를 통합한 가중 평균 효과크기 계산, 평균 효과크기의 유의성 검증, 동질성 검증과 같이 실제 메타 통계분석

을 진행하는 과정, 그리고 다섯째 단계인 앞서 메타 통계분석에서 나타난 결과를 제시하고 해석하는 것에 주안점을 두고 살펴보기로 한다. HS방식의 메타분석과 HO메타분석은 관찰된 상관계수 값 그 자체를 사용할 것인지 혹은 이를 변환하여 사용할 것인가 하는 점에서 출발하여 이후 과정에서 차이를 보이게 된다.

### 1) 효과크기 산출

연구방법론으로 메타분석에 대한 논의가 거듭되면서 이를 실제로 수행하는 방법에 대한 연구도 함께 진행되어 왔다. 메타분석을 수행하기 위한 첫 단계로 효과크기를 산출하게 되는데, 개별 효과크기를 계산하고 이를 통합하는 과정에서 전자는 상관계수를 피셔(Fisher)의  $z$  값으로 변환하여 분석한 후 통합 효과크기를 산출할 때 이를 다시  $r$ 로 변환하는 과정을 거치는 것을 특징으로 한다. 하지만 메타분석을 실시하기 위해 상관계수 자체를 사용할 것인지, 아니면 상관계수를 피셔의  $z$ 로 변환한 후 이를 토대로 메타분석을 실시하는 것이 좋은지에 대해서는 끊임없는 논란이 제기되고 있다(e.g., Hall & Brannick, 2002; James, Demaree, & Mulaik, 1986; Johnson et al., 1995; Schmidt et al., 1988; Schmidt & Hunter, 1999a; Schulze, 2007). 울프(Wolf, 1986)는 원 상관계수 값보다 피셔의  $z$ 가 조금 높게 나타나지만 모집단에 비해 표본크기가 작은 경우가 아니면 무시할 수 있는 수준이라고 주장하기도 한다. HO방법의 메타분석에서는 분석의 기초자료로 상관계수를 피셔의  $z$  변환한 값을 이용한 후, 이를 다시  $r$ 값으로 복원하여 최종 결과를 산출한다. 상관계수로 효과크기가 제시되지 않은 경우( $t$ 값,  $F$ 값,  $\chi^2$  등)는 상관계수로 변환하는 공식을 이용하게 된다(예: Wolf, 1986). 이에 비해 HS방법의 메타분석에서는 기본적으로 상관계수  $r$ 값 자체를 분석의 기초로 하고 있다. 피셔의  $z$ 로 변환한 값을 가지고 분



석을 하는 경우에는 평균상관관계를 높이는 방향으로 편향된 결과를 가져올 수 있으며, 전환하지 않은 상태에서 분석한 것보다 정확도가 낮아진다고 주장하고 있다(Hunter & Schmidt, 2004). 특히 상관관계가 높은 경우에는 정적으로 편향도가 크게 높아진다고(positive bias) 한다. 메타분석을 진행하는데 있어 HO방법과 HS방법을 둘러싼 논쟁의 핵심은 상관계수  $r$  자체를 이용할 것인가 혹은 피셔의  $z$ 로 변환한 후 사용할 것인가 하는 문제와 개별 효과크기를 통합할 때 가중치를 어떻게 적용할 것인가 하는 문제로 집약된다(Sshulze, 2007).

## 2) 효과크기 통합 및 동질성 검정

전체 메타분석 결과를 제시하기 위해서는 적절한 통계적 방법을 사용하여 연구결과들을 통합하는 과정을 거쳐야 한다. 메타분석에서 각 연구의 효과 차이를 통합할 때 가중치를 적용하는 것이 일반적인데, 이는 표본크기가 작은 연구는 표본수가 큰 연구에 비해 우연한 차이로 인한 영향을 더 많이 받을 수 있으므로 통합 효과크기를 제시할 때에는 표본수가 큰 연구에 상대적으로 더 큰 가중치를 부여하게 된다(이준영, 2008; Egger, Smith, & Phillips, 1997).

전체 효과크기는 동질성 분석에 따라 두 가지 결과로 나타나게 된다. 동질성 분석의 영가설은 '개별 연구들의 효과크기들이 동일 모집단에서 추출되었다는 것으로 설정된다. 만일 동질성 분석 영가설을 기각하지 못하면 동일 모집단에서 효과크기가 추출된 것으로 판단하여 고정효과모형(fixed-effect model)을 사용하고, 동질성 검정 결과 통계적으로 유의하여 영가설을 기각하게 되면 표집오차만이 아니라 무작위 오차(random error)까지 고려하는 임의효과모형(random-effect model)을 이용하게 된다. 고정효과모형에서는 메타분석에서의 모든 효과크기는 공통의 평균을 공유하며, 개별 효과크기의 차이가 나타나는 것

은 단지 표집오차에 의한 것, 즉 서로 다른 표본수를 사용하기 때문이라고 가정한다. 반면에 임의효과모형에서는 표집오차에 의한 차이(sampling variation)뿐만 아니라, 개별 효과크기는 미지분포(unknown distribution)로부터 추출되기 때문에 차이를 보일 수밖에 없다는 무작위 변이(random variation)를 함께 가정한다(Borenstein et al., 2010). 결국 이들 두 모형의 차이는 각 연구에서 얻은 결과가 어떤 분포를 이루고 있는가에 대한 통계적 관점의 차이라고 할 수 있다(이준영, 2008).

통계적 동질성 혹은 이질성 평가는 코크랜(Cochran)의  $Q$ 나  $I^2$ 를 이용하는 방법이 많이 사용되고 있으며,  $I^2$ 의 경우 75% 이상이면 효과크기가 이질적인 것으로 인정되고 있다(Higgins, Thompson, Deeks, & Altman, 2003). 일반적으로 임의효과모형을 사용해야 할 때 고정효과모형을 사용하면 효과크기를 과대 추정하는 경향이 있다(Borenstein et al., 2009).  $Q$ 통계량은 기본적으로  $\chi^2$  분포를 따르며, 전체 표본에 대한 메타분석 이후 하위집단으로 분석할 수 있는지를 판단하는데 사용된다. HO 메타분석에서는 관찰된 변량(observed variance)이 표집오차로 인해 예상되는 변량보다 큰지 여부를  $\chi^2$  동질성 검사를 한다.  $\chi^2$  검사를 통해 변량이 일정수준 이상이 된다면 동질성이 기각되고 이질적인 것으로 판단하여 임의효과분석을 하게 된다.

한편, 통합 효과크기의 해석과 관련해서는 코헨(Cohen, 1988)이 제시하는 기준을 따르는 것이 일반적이다. 이에 따르면, 상관계수의 영향력의 크기를 .10 이하면 작은 효과크기, .25 정도면 중간 효과크기, .40 이상이면 큰 효과크기로 보고 있다. 분석대상이 되는 연구물들에 대한 평균 효과크기를 구하고 난 다음, 효과크기의 유의성은 신뢰구간과  $z$ 검증을 통해 추론할 수 있다. 평균 효과크기의 신뢰구간을 설정하기 위해서는 표준오차의 값을 알아야 하는데, 이 값은 메타분석 방법에 따라 다양하다.

### 3) 신뢰구간과 확신구간 산정

HO방법의 메타분석에서는 표본크기를 감안한 가중 평균 효과크기를 산출하고 표준오차(standard error)를 통해 신뢰구간(confidence interval, CI)을 산정한다. 여기서 사용하는 신뢰구간은 모수치 자체가 아닌 모수 추정치를 가리킨다. 신뢰구간은 결국 상관관계 평균의 추정치를 나타내는 것으로 표준오차는 표본크기에 의존하고, 이는 결국 표집오차와 깊은 관련이 있다. 이에 비해 HS방법의 메타분석에서는 모수치(parameter values)의 분포를 언급하는 것으로 확신구간(credibility interval, CV)을 사용하고 있다. HS방법의 메타분석에서 일차적으로 얻은 것은 실제 점수(true score) 상관관계( $\bar{\rho}$ )이다. 그러나 이러한 실제 점수 상관관계는 엄격히 말하면 모수치 자체가 아니며, 오차를 교정하여 추정한 모수치를 가리킨다. 따라서 확신구간은 모수 추정치( $\bar{\rho}$ )의 표준오차를 사용하여 구하는 신뢰구간과 달리 표집오차를 포함한 각종 오류들이 제거된 모집단 상관관계( $\rho$ )의 표준편차를 활용하여 얻게 된다. 다시 말해 확신구간은 모수치의 분포를 말하는 것으로 표집오차의 영향을 받지 않는다. 반면 신뢰구간은 표준오차를 반영하여 값을 구하게 되는데, 표준오차는 표집오차에 따라 차이를 받을 수 있으므로 표집오차의 영향을 크게 받는다. 이 때문에 모수치에 관심을 갖는 HS방법의 메타분석에서는 확신구간이 신뢰구간보다 일반적으로 더 중요한 의미를 가지게 된다. 그러나 조절변인 등을 확인하는 데는 신뢰구간이 더 중요한 역할을 담당하기도 한다(허만용, 2008; Hunter & Schmidt, 2004).

기본적으로 HO방법의 메타분석은 측정치 수준(measure or scale level)을 나타내는 평균 효과크기를 추정하는데 초점을 맞추고 있으나, HS방법의 메타분석은 구성개념 수준(construct level)의 관계를 통계적

으로 추정하는데 초점을 맞추고 있다고 할 수 있다. 이 때문에 확신구간은 연구간 모수치 변량을 허용하지 않아 기본 정의상 확신구간이 “0”일 수밖에 없는 고정효과모형과 달리 임의효과모형과 깊은 관련이 있다. 결국 신뢰구간은 평균 실제 점수 상관에 대한 점 추정치에 신뢰수준(일반적으로 95%를 가정하게 됨)에 따른 오차 범위를 결합해서 얻는 구간 추정치로, 95%의 신뢰구간이 “0”을 포함하지 않는다면 추정된 실제 점수 상관이 통계적으로 유의미하다는 것을 의미한다. 반면 확신구간은 실제 점수의 상관의 분포를 가리키는 것으로 0을 포함하지 않고 초과(혹은 미달)하는 범위에 있으면, 개별적으로 교정된 상관계수의 대부분이 0을 기점으로 정적 혹은 부적으로 집중되어 있음을 의미한다. 확신구간의 범위를 헌터와 슈미트(Hunter & Schmidt, 2004)는 80%를 제시하고 있다. 즉 신뢰구간은 추정된 단일값(평균 실제 점수 상관계수)의 정확도에 대한 정보를 제공해주며, 확신구간은 모수값의 분포로 조절변인에 의한 해석 여부의 필요성에 대한 정보를 제공해 준다(Hunter & Schmidt, 2004; Whitener, 1990).

#### 4) 조절변인의 탐색

앞서 기술한 바와 같이, 전체 연구들에 대한 통합 효과크기를 추정 한 후에는 개별 연구들에서 추정된 효과크기가 동질적인지에 대한 동질성 검증을 실시하게 된다. 여기서 나타난 결과가 동질적이지 않은 것으로 판단될 경우, 연구자는 조절변인에 의한 영향이 있었는지를 살펴보아야 한다. 즉, 추정된 효과크기가 동질적인지 혹은 이질적인지의 여부는 고정효과모형이나 임의효과모형 중 어떤 모형을 토대로 결과의 해석을 할 것인지를 결정하는 것과 함께 조절변인에 의한 효과크기 차이를 탐색할 것인가를 결정하는 근거가 된다.

HS 메타분석에서는 하위집단의 조절효과를 판단하기 위한 방법으

로 세 가지 기준을 제시하고 있다. 첫 번째는 조절변인의 기각 규칙으로 “75% 규칙”을 제시하고 있다. 이는 총 변량 중에서 표집오차를 비롯한 각종 통계적 오류로 인한 변량이 75% 이상을 차지한다면 실제적으로는 모든 변량이 오류에 의한 것이며, 나머지 변량은 통제할 수 없는 오류들로 인한 것으로 본다. 즉 메타분석 결과 관찰된 총 변량에서 통계적 오류(예, 표집오차, 측정오차)로 인한 변량이 75%보다 적다면 두 변인간의 관계를 조절하는 변인이 존재한다고 판단하며, 반대로 변량의 75% 이상이 설명되면 실질적인 조절변인이 존재하지 않는다고 본다(Hunter & Schmidt, 2004). 하지만 HS 메타분석을 하는 연구자들 중에도 75% 규칙에 이견을 제시하는 의견도 존재한다(Borenstein et al., 2009).

HS방법에서 조절효과를 검증하는 두 번째 기준은  $Q$ 통계량으로, 이는 실제 상관관계의 동질성에 대한 유의도 검정이다.  $Q$ 통계량은 통합효과크기를 산출하고 난 후 하위집단으로 분류할 수 있는지 판단하는데 유용하게 사용할 수 있다. 모든 실제 상관관계가 동일하다는 영가설 하에서  $Q$ 는 자유도  $k-1$ 인  $\chi^2$  분포를 따른다. 헌터와 슈미트(Hunter & Schmidt, 2004)는 일반적으로 통계적 오류들의 변량 때문에 제1종 오류가 발생하기 때문에  $\chi^2$  검정은 수정된 상관관계를 기초로 계산되어야 하며, 수정되지 않은 상관관계로 계산된  $\chi^2$  검정은 통계적 편향(statistical bias)을 발생시킬 수밖에 없다고 주장한다. 이렇게 볼 때 관찰된 전체 변량 중에서 통계적 오류들에 의해 설명되는 변량이 작고(75% 이하),  $Q$  통계량이 유의미할 때 개별 연구들에서 추정된 효과크기들간의 변동성(variability)을 설명해주는 조절변인이 있다고 판단할 수 있으며, 잠재적 조절변인을 탐색할 필요성이 있다.

조절변인의 영향을 살펴보기 위한 세 번째는 하위집단간 신뢰구간의 중첩 여부를 확인하는 것으로 각 하위집단 상관관계의 95% 신뢰구간을 비교하여 그 구간에서 중첩되는 부분이 있는지 파악하는 것

이다. 신뢰구간 중첩 여부는 하위집단으로 분류하기 위해 기준이 되었던 조절변인의 효과를 확인하는데 결정적인 역할을 하게 된다. 비교 대상이 된 각 하위집단에서 평균 상관관계의 신뢰구간이 중첩되지 않는다면 조절효과가 있는 것으로 판단한다.

### 5) 출판편향과 보정

연구자가 특정 주제에 관한 선행연구 결과들을 통합 분석할 때, 해당 주제에서 이루어진 모든 연구를 총망라하지 못하고 일부 연구결과들만 가지고 종합하게 되면 표집의 대표성이 문제가 될 수 있는데, 로젠탈(Rosenthal)은 이를 책상서랍의 문제(file drawer problem) 혹은 출판편향(publication bias)이라고 한다.

현재까지 출판편향 여부를 확인하기 위한 다양한 방법이 소개되었으나 출판편향의 문제를 확인할 수 있는 하나의 완벽한 방법은 존재하지 않는다(Borenstein, 2005). 여러 가지 출판편향을 보정하기 위한 방법이 개발, 제시되는 것도 바로 이러한 이유 때문이기도 하다. 출판편향의 존재 여부는 가지치기와 채우기(trim and fill: Duval, & Tweedie, 2000), 깔때기 그림(funnel plot: Egger, Smith, Schneider, & Minder, 1997), 에거(Egger)의 회귀분석(Egger's regression analysis: Egger, Smith, & Altman, 2003) 등을 통해 확인할 수 있다. 퍼거슨과 브래닉(Ferguson & Brannick, 2012)의 연구에 따르면, 심리학 관련 연구 데이터베이스인 PsycINFO를 중심으로 2005년에서 2009년까지 5년간 수행된 메타분석 연구에서 출판편향 문제를 가지치기와 채우기(24%), 안전성 계수(22%), 출판된 연구의 효과크기와 미출판 연구의 효과크기 비교(21%), 깔때기 그림(5%), 에거의 회귀분석(5%) 등의 방법을 통해 확인하고 있는 것으로 나타났다.

깔때기 그림은 관찰 상관(Y축)을 표본크기(X축)에 따라 도면상에

나타내는 것으로, 산점도(scatter plot)의 모양이 옆으로 누운 깔때기 모양으로 대칭을 이루면 출판편향이 없는 것으로 본다. 이는 표본크기가 작을 때는 효과크기의 변동성이 크기 때문에 산점도가 퍼지고 표본크기가 클 때는 오차가 줄어들기 때문에 추정된 실제 점수로 수렴한다는 가정에 기초한다(Vevea & Woods, 2005). 안전성 계수는 메타분석을 통해 산출되는 효과크기가 0으로 나오기 위해 필요한 연구 논문의 수를 계산하여 준다. 하지만 이 방법은 누락된 연구의 평균 효과크기에 따라  $N$ 에 대한 추정이 매우 다양하게 산출된다는 취약점이 있으며(Becker, 2005), 안전성 계수는 실질적인 유의성이 아니라 통계적 유의성에 초점을 두고 있으므로 출판편향을 확인하는 방법으로 사용하지 않는 것이 바람직하다는 주장도 있다(Littell, Corcoran, & Pillai, 2008). 가지치기와 채우기 방법은 효과크기와 표준오차가 좌우 대칭 분포를 보이는지 확인하여 좌우 대칭이 아닐 때 대칭을 만들기 위해 보고되지 않은 효과크기를 추정하여 채워 넣음으로써 대칭 분포를 만들고, 관찰 효과크기 값과 비교하여 그 차이가 크면 출판편향의 가능성이 있는 것으로 보게 된다. 회귀분석 방법은 실제 효과크기와 표준오차의 상관관계를 통해 효과크기 분포의 비대칭성을 수치화하는 방법을 제시하고 있는데, 출판편향이 없는 경우 표준화된 효과크기의 정확도를 통해 추정된 회귀절편은 0과 유의한 차이를 보이지 않는다. 한편, 출판편향과 관련하여 2005년까지 수행된 메타분석에 대한 로스타인과 동료들(Rothstein et al., 2005)의 연구에 따르면, 약 50%는 그 영향이 미미한(minimal) 수준이고, 약 45%는 보통(modest), 나머지 약 5%만이 심각한(severe) 수준으로 나타났다.

### 3. HO방법과 HS방법의 메타분석 수행 과정

1970년대 중반 이후 교육심리학에서 시작된 메타분석은 HS방법과 HO방법 등의 두 주류를 이루며 사용되어 왔다. 홀과 브래닉(Hall & Brannick, 2002)은 크게 3가지 측면에서 HS 메타분석 방법과 HO 메타분석 방법의 차이를 설명하고 있다. 첫째는 상관계수  $r$ 값 자체를 이용하는가 아니면 피셔(Fisher)의  $z$ 값으로 변환한 값을 이용하는가 하는 점이다. 둘째는 모집단 상관계수를 추정하는 과정에서 감소 요인을 보정하는가 혹은 보정하지 않는가 하는 점이다. 셋째는 가중치를 어떻게 적용하여 통합 효과크기를 산출할 것인가 하는 점이다. 예를 들면 가중치를 고려하여 통합 효과크기를 계산하는 과정에서 HS방법은 표본크기에 비례하여 더 큰 가중치를 부여하는 반면, HO방법은 모집단 변량이 작을 경우 큰 가중치를 부여하게 된다. HO 메타분석에서는 표본크기를 감안한 가중 평균 효과크기를 산출하고 표준오차를 통해 신뢰구간을 산정하는데 주안점을 두고 있다. 여기서 사용하는 신뢰구간은 모수치 자체가 아닌 모수추정치로서 구체적으로는 상관관계 평균의 추정치를 나타내는 것으로 표준오차는 표본크기에 의존하고, 이는 결국 표집오차와 깊은 관계가 있다. 이에 반해, HS 메타분석에서는 모수치의 분포를 언급하는 것으로서 확신구간을 사용하고 있다. HS 메타분석을 통해 얻는 결과는 실제 점수 상관관계( $\bar{\rho}$ )이다. 물론 이 또한 엄격히 말하면, 모수치 자체는 아니며, 오차를 보정하여 추정한 모수치라고 표현하는 것이 더 정확하다. 이에 따라 확신구간은 모수추정치( $\bar{\rho}$ )의 표준오차를 사용하여 구하는 신뢰구간과 달리, 표집오차를 포함한 각종의 오류들이 제거된 모집단 상관관계( $\rho$ )의 표준편차를 활용하여 얻는다.



### 1) 효과크기 산출

연구대상 논문들의 효과크기를 상관계수(Pearson's  $r$ )로 표현된 연구 결과들을 메타분석하는 가장 전형적인 방법은 상관계수 자체를 분석의 기초자료로 활용하거나, 상관계수를 피셔의  $z$ (Fisher's  $z$ )로 변환한 값을 분석의 기초자료로 활용하는 방법이 있다. 피셔의  $z$ 값을 사용하는 이유는 상관계수 분포는 비대칭 분포를 보이는 것이 보통이므로 상관계수  $r$ 을 사용하지 않고 피셔의  $z$ 값으로 변환하여 정규 분포에 가까운 대칭 분포를 만들기 위해서이다(Borenstein et al., 2009). HO 메타분석에서는 분석의 기초자료로 상관계수를 피셔  $z$  변환값을 이용하는데, 피셔  $z$ 의 산출(공식 1)과 분산(공식 2), 및 표준오차(공식 3)를 산출하는 공식은 다음과 같다(오성삼, 2002; Borenstein et al., 2009).

$$ES_{zr} = 0.5 * \ln\left(\frac{1+r}{1-r}\right) \quad \text{식 1.1}$$

$$V_z = \frac{1}{n-3} \quad \text{식 1.2}$$

$$SE_z = \sqrt{V_z} \quad \text{식 1.3}$$

Fisher  $z$ 로 변환한 효과크기를 이를 다시  $r$ 값으로 복원하는 방법은 다음과 같다(Borenstein et al., 2009).

$$r = \frac{e^{2z} - 1}{e^{2z} + 1} \quad \text{식 1.4}$$

## 2) 통계적 오류의 보정

HS방법은 기본적으로 측정 수준(measure or scale level)이 아닌 구성 개념 수준(construct level)의 관계를 추정한다는 점에서 HO방법과는 큰 차이를 보인다. 이 때문에 각 변인이 갖는 측정오차 등에 의한 편향과 같은 각종 통계적 오류들을 보정하는 것에 초점을 맞추고 있으며, 모든 오차를 보정하고 난 후 상관관계를 통합하는 과정을 거치는 것으로 메타분석을 실시한다. HS방법은 개별 연구결과들이 다양한 형태의 오차를 포함하고 있기 때문에, 이러한 인위적이고 편향된 오차를 보정해야만 실제 값을 제대로 추정할 수 있다고 주장한다.

HS방법에 의한 메타분석에서는 ‘관찰된(observed or unadjusted)’ 효과크기와 ‘진정한(true or adjusted)’ 효과크기를 구분하여 설명한다. 이에 따르면 관찰된 효과크기는 측정의 불안전성으로 인해 실제 효과크기라고 보기 어려우며 이 때문에 ‘감소되거나(attenuated)’, 혹은 ‘감소되지 않은(unattenuated)’ 오차를 보정해야 실제 값을 확보할 수 있다고 가정한다. 이러한 출발선상에서 HS방법의 메타분석은 가정 먼저 발생할 수 있는 오차나 연구자들이 사용한 측정도구의 오류(artifacts or man-made errors)들의 크기를 확인하고, 이를 보정하는 과정을 통해 모집단 효과크기를 확인하는 과정을 거친다(Hunter & Schmidt, 2004). 그러나 앞서 언급한 바와 같이, 헌터와 슈미트가 정리하고 있는 다양한 오류들 가운데 개별 연구 논문에서 얻을 수 있는 정보가 극히 제한적이기 때문에 실제 메타분석 연구에 대한 적용에 있어서는 표집오차와 측정오차 등과 관련된 내용이 대부분이다.

### (1) 표집오차의 보정

HS방법에서는 각 연구에서 제시하고 있는 상관계수(관찰 상관계수)에 표본수를 고려한 가중치를 적용하여 상관계수 평균( $r$ )과 연구간 관

찰 상관관계 변량( $s_r^2$ ), 그리고 모집단 상관관계 변량( $\sigma_\rho^2$ )을 구하게 된다(Hunter & Schmidt, 2004). 여기까지가 헌터와 슈미트가 주장하는 기본적인 메타분석(bare bones meta-analysis)으로 표집오차만을 고려한 메타분석이다. 그러나 HO방법과 다르게 HS방법은 가중상관관계 평균( $r$ )을 그대로 사용하지 않고, 다양한 통계적 오류를 보정한 실제 상관관계 평균( $\rho$  또는  $Ave(\rho)$ )을 기초 자료로 분석한다. 일반적으로 효과크기는 HO 메타분석 결과보다 HS 메타분석 결과에서 더 높게 나타나게 되며, 그 정도는 가능한 통계적 오류를 얼마나 많이 보정하는가에 비례하여 증가한다.

HS방법에서는 표집오차 보정을 메타분석 연구를 진행하기 위한 가장 기본적인 출발점으로 본다. 만일 모집단 상관관계가 모든 연구에 걸쳐 일정하다면, 그 상관관계에 대한 최상의 추정치는 연구들간의 단순 평균이 아니라 가중 평균, 즉 개별 연구에 사용된 표본수에 의해 가중치가 반영된 평균 상관관계가 되어야 한다(Hunter & Schmidt, 2004). 이를 수식으로 표현하면 다음과 같다.

$$r = \frac{\sum [N_i r_i]}{\sum N_i} \quad (\text{식 2.1}) \quad \begin{array}{l} r_i \text{는 연구 } i \text{의 상관관계} \\ N_i \text{는 연구 } i \text{의 표본수} \end{array}$$

여기서 연구들의 변량도 일상적인 표본 변량이 아니라 가중치가 반영된 변량이어야 한다(Hunter & Schmidt, 2004). 이를 수식으로 표현하면 다음과 같다.

$$s_r^2 = \frac{\sum [(r_i - \bar{r})^2]}{\sum N} \quad (\text{식 2.2}) \quad \begin{array}{l} r_i \text{는 연구 } i \text{의 상관관계} \\ N_i \text{는 연구 } i \text{의 표본수} \end{array}$$

위의 (식 2.1)과 (식 2.2)를 통해서 표집오차 변량(식 2.3)과 모집단 상관관계 변량(식 2.4)을 추정할 수 있다.

$$\sigma_e^2 = \frac{(1 - \bar{r}^2)^2}{(N-1)} \quad (\text{식 2.3}) \quad \begin{array}{l} r_i \text{는 연구 } i \text{의 상관관계} \\ N_i \text{는 연구 } i \text{의 표본수} \end{array}$$

$$\sigma_\rho^2 = \sigma_r^2 - \sigma_e^2 \quad (\text{식 2.4}) \quad \begin{array}{l} \sigma_r^2 \text{는 표본 상관관계 변량} \\ \sigma_e^2 \text{는 표집오차 변량} \end{array}$$

## (2) 측정오차의 보정

효과크기를 산출하는데 있어 측정오차를 고려한다는 점이 HO방법과 HS방법의 메타분석 방법론상의 가장 큰 차이점이다. HO방법의 메타분석에서는 측정오차를 포함한 대부분의 오차들에 대한 해결방법을 크게 고려하지 않는다. 하지만 커뮤니케이션 현상과 같은 사회과학을 다루는 모든 연구에서 완벽하게 측정되는 변인은 존재하지 않는다. 이렇게 발생한 체계적 오차(systematic error), 즉 측정오차로 인해 개별 연구에서 상관관계는 모집단 상관관계보다 효과크기가 낮아지게 되며, 이러한 감소모형(attenuation model)을 고려하여 산출하는 것이 측정 오차를 보정한 HS방법의 메타분석이다. 상관관계는 복합감소계수(the compound attenuation factor)를 이용하여 보정할 수 있는데, 복합감소계수는 분석대상 논문에서 제시되는 측정 변인의 측정오차(신뢰도 계수)를 곱하는 방식으로 산출할 수 있으며 평균 복합감소계수는 각각의 측정오차들의 평균치들을 곱하는 방식으로 산출해낼 수 있다(Hunter & Schmidt, 2004).

측정오차에 의한 체계적 감소(systematic attenuation)의 보정은 다음과 같은 전통적인 공식을 따른다.

$$\rho TU = \frac{\rho_{xy}}{\sqrt{s_{xx}} \sqrt{s_{yy}}} \quad (\text{식 2.5})$$

$\rho TU$ 는 모집단 상관관계  
 $\rho_{xy}$ 는 관찰된 상관관계  
 $\sqrt{s_{xx}}$ 는 변인 x의 신뢰도  
 $\sqrt{s_{yy}}$ 는 변인 y의 신뢰도

예를 들어, 두 구성개념, 광고태도와 제품태도 간의 상관관계의 관찰된 값이 .30이라고 할 때, 이들 변인이 완벽하게 측정되었다면 이들 상관관계의 실제 점수, 즉 모집단 상관관계 역시 .30으로 추정될 수 있다. 그러나 만일 광고태도와 제품태도의 신뢰도가 각각 .50, .60 이라면 이들 관찰된 상관관계 값은 다음과 같이 보정되어야 한다. 즉, 위의 (식 2.5)에서 제시된 바와 같이, 측정오차는 모집단 상관관계의 실제 점수를 .55에서 .30으로 감소시키는 역할을 했다고 볼 수 있다.

$$\rho TU = \frac{\rho_{xy}}{\sqrt{s_{xx}} \sqrt{s_{yy}}} = \frac{.30}{\sqrt{.50} \sqrt{.60}} = .55 \quad (\text{식 2.6})$$

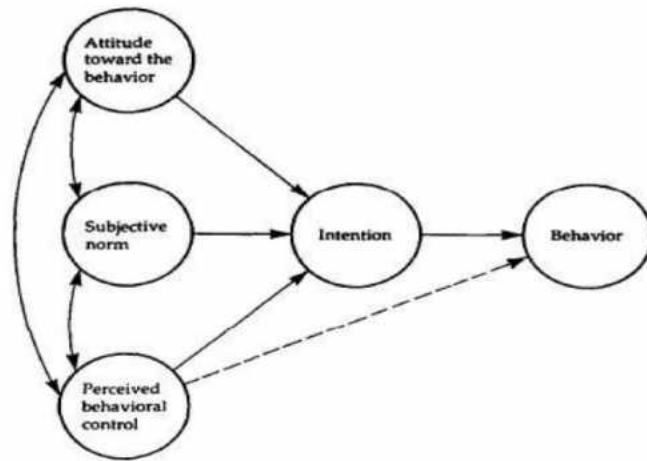
지금까지 메타분석의 기본 개념 및 절차, 그리고 HS방법의 통계적 오류의 보정, 즉 표집오차와 측정오차의 보정을 살펴보았다. 지금부터는 본 연구의 목적인 표집오차와 측정오차의 보정에 따라 HO방법과 HS방법에 의한 메타분석 결과가 어떤 차이를 보이는지를 실제 자료를 통해 비교·분석하고자 한다.

#### 4. 연구방법

본 연구는 TPB를 적용한 연구를 대상으로 표집오차와 측정오차의 보정에 따른 HO방법과 HS방법에 의한 메타분석 결과 차이를 비교하기 위해 진행되었다. TPB는 인간 행동을 설명하는 강력한 이론적 틀을 제공하는 것으로 평가받고 있으며(Courneya, 1995; Hausenblas, Carron, & Mack, 1997), 측정이나 분석방법에 대해서도 에이젠(Ajzen, 1985)이 초기에 제시했던 이론적·방법론적 틀 안에서 국내에서 상당히 많은 연구가 진행되고 있어, 본 연구의 분석 사례로 적절하다고 판단되었다. 본 연구는 지난 2011년까지 국내에서 수행된 TPB를 적용한 연구에 대한 메타분석 연구(손영곤·이병관, 2012)를 2015년까지로 확장하여 연구를 수행하였다.

##### 1) TPB 개관

TPB는 인간의 행동을 이해하고 예측하는데 유용하게 사용되어 온 대표적인 사회심리학 이론인 이성적 행위이론(theory of reasoned action: Ajzen & Fishbein, 1980)의 확장된 이론이다. TPB는 행동, 행동의도, 행동에 대한 태도, 주관적 규범, 지각된 행동통제 등의 개념으로 구성된다. TPB는 사람들이 어떤 행동을 하는 것은 그와 같은 행동을 수행하려는 행동의도(behavioral intention)에 의해 결정되고, 행동의도는 다시 대상에 대한 태도(attitude), 주관적 규범(subjective norm), 지각된 행동통제(perceived behavioral control)에 의해 결정되는 것으로 설명한다(Ajzen, 1985; Conner & Sparks, 2001). <그림 1>에 나타난 바와 같이, ‘개인의 특정 행동에 대한 신념과 그 행동 결과에 대한 평가의 함수로 구성되는 태도’와 ‘특정 준거집단의 기대인 규범적 신념과 그 준거집단의 기대에 따르려는 동기인 순응동기의 함수인 주관적 규범’은



〈그림 1〉 TPB 모형(Ajzen, 1991)

행동의도에 직접적인 예측변인으로 구성된다. 또한 이전의 이성적 행위이론에서는 제시되지 않았던 개념인 지각된 행동통제는 행동 수행의 용이성 및 행동 관련 자원과 장애에 대한 신념을 말하는데, 이 구성개념은 행동의도를 통해 행동에 영향을 미치는 간접효과는 물론 행동의 직접적인 예측변인으로 작용한다. 끝으로 TPB는 행동의도를 행동의 직접적인 예측변인으로 구성하였다.

본 연구는 TPB의 각 구성개념의 효과크기, 즉 태도, 주관적 규범, 지각된 행동통제가 행동의도에 미치는 효과크기와 행동의도 및 지각된 행동통제가 행동에 미치는 효과크기를 HO방법과 HS방법의 메타분석을 수행하여 비교하였다.

## 2) 자료 수집 및 선정

본 연구는 2015년까지 국내에서 발표된 TPB를 적용한 연구를 분

석대상으로 실시하였다. 연구논문은 한국교육학술정보원에서 제공하는 학술연구정보서비스(www.riss4u.net), 한국학술정보에서 제공하는 검색 사이트(http://kiss.kstudy.com), 학술교육원(http://www.earticle.net), 누리미디어 DB Pia(http://www.dbpia.co.kr), 교보문고 스킨라(http://scholar.dkyobobook.co.kr), 뉴논문(http://www.newnonmun.com) 등 국내 학술지를 검색할 수 있는 사이트를 이용하여 연구 목록과 자료를 수집하였다. 결과물은 '계획된 행동(위)이론', '계획적 행동(위)이론', '계획행동(위)이론', 'Theory of Planned Behavior', 'TPB' 등의 검색어를 통해 나타나는 논문을 일차적으로 선정하였다.

검색어 결과 모두 464편을 찾을 수 있었는데, 본 연구에서는 논문 수집과정의 중복 및 편의를 위해 선행 연구(손영곤·이병관, 2012) 이후 발표된 논문을 중심으로 수집하였다. 그 결과 2011년까지 발표된 논문 188편과 2012년에서 2015년 4년간 발표된 276편의 논문을 찾을 수 있었다. 이들 가운데 중복 논문 53편, TPB와 관련이 없는 논문 12편, 한국연구재단 등재지나 등재후보지가 아닌 논문 6편 등 71편의 연구물을 제외한 393편을 대상으로 본 연구에 적합한 논문을 선별하는 과정을 거쳐 최종적으로 202편의 논문을 대상으로 분석을 실시하였다. 이를 위해 HO방법은 보렌스타인 등(Borenstein et al., 2009)이 확대 발전한 CMA(Comprehensive Meta-Analysis) 프로그램을, 통계적 오류의 보정을 통한 HS방법은 R programing language & statistical software(R Development Core Team, 2011)의 psychometric 패키지(Fletcher, 2010)를 이용하여 분석하였다.

### 3) 분석 대상 논문 포함/제외 기준 및 자료처리

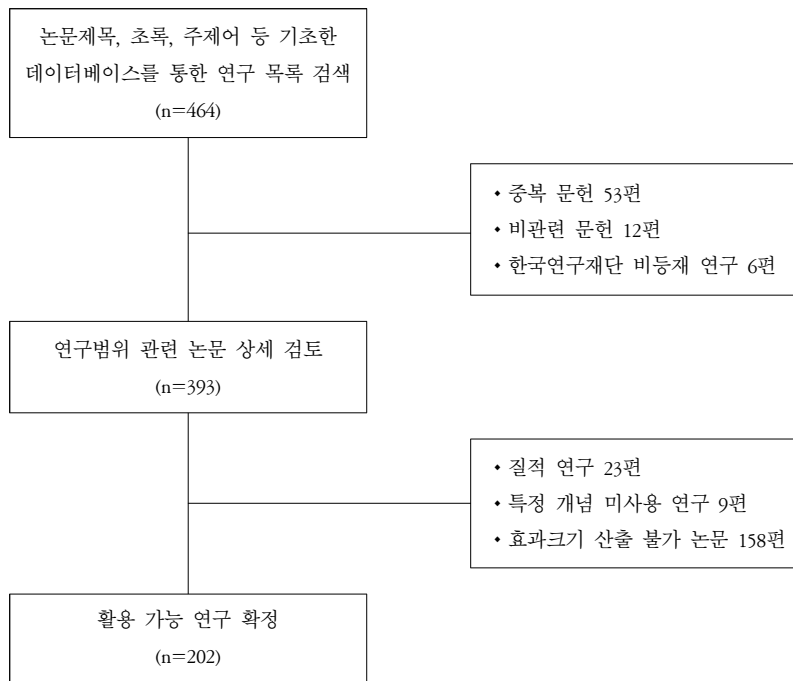
메타분석을 위한 첫 단계는 연구주제와 대상의 선정에서 출발한다. 본 연구는 국내에서 수행된 TPB를 적용한 논문을 대상으로 하였다.



구체적으로 어떠한 원칙에 따라 어떠한 논문을 포함하고, 또 제외하였는가를 상술하면 다음과 같다.

첫째, TPB를 적용한 실증적 연구논문이어야 한다. 문헌연구나 다른 이론적 논의와의 연장선상에서 TPB를 다루고 있는 논문 23편을 제외하였다. 둘째, TPB에서 가정하는 기본 개념 즉, 태도, 주관적 규범, 지각된 행동통제와 행동의도나 행동 개념을 모두 포함한 관계를 다루고 있어야 한다. TPB를 구성하는 태도나 주관적 규범, 지각된 행동 통제 중 어느 하나의 개념이라도 측정하지 않은 연구, 또는 행동의도나 행동을 종속변인으로 하고 있지 않은 연구 9편을 제외하였다. 셋째, 변인들간의 관계를 상관관계로 제시한 논문을 대상으로 하였다. 회귀분석이나 구조방정식모형을 통해 결과를 제시하고 있는 경우는 제외하였다. 회귀계수나 구조방정식을 이용하는 메타분석 방법에 대해서도 논의되고 있기는 하지만(예: Aloe & Becker, 2012; Becker & Wu, 2007; Peterson & Brown, 2005), 아직 방법론적 타당성을 확보하지 못하고 있는 관계로 본 연구에서는 이를 제외하였다. 이 과정에서 158편의 논문을 배제할 수 있었다. 단, 이들 논문에 대해서도 연구자들에게 별도로 연락하여 상관관계 결과를 구할 수 있는 경우 포함하여 자료 손실을 최소화하고자 하였다. 넷째, TPB를 구성하는 개념들을 모두 사용하고 있고, 변인들간의 상관관계를 제시하고 있는 경우라고 해도 구성개념간의 상관관계가 아닌 측정 변인간의 상관관계를 제시하고 있는 연구는 제외하였다. HS방법에 의한 메타분석은 기본적으로 구성개념 단위의 관계에 초점을 맞추고 있기 때문이다. 다섯째, HO방법에 의한 메타분석의 경우 표본크기와 상관계수 등의 통계치만 제시되어 있으면 산출이 가능하다. 하지만 HS방법의 경우 표집 오차만이 아니라 측정 과정에서 발생하는 통계적 오류를 수정하기 위해서 별도의 통계치도 감안하였다. HS방법에 의한 메타분석에서 측정오차는 신뢰도 계수(Cronbach's  $\alpha$ )를 통해 보정하게 된다. 이에 따

라 본 연구에서도 TPB를 구성하는 태도, 주관적 규범, 지각된 행동통제, 행동의도, 행동 개념에 대한 측정의 신뢰도를 확인하였다. 신뢰도가 보고되지 않은 경우에는 전체 평균으로 대체하였다. 본 연구에서 누락되어 있는 연구의 신뢰도는 전체 평균(태도  $\alpha=.858$ , 주관적 규범  $\alpha=.836$ , 지각된 행동통제  $\alpha=.806$ , 행동의도  $\alpha=.880$ , 행동  $\alpha=.889$ )으로 대체하여 계산하였다. 단, 단일 문항으로 측정된 경우와 같이 신뢰도 계수를 확인할 수 없는 변인에 대한 신뢰도는 1로 처리하여 분석을 실시하였다.



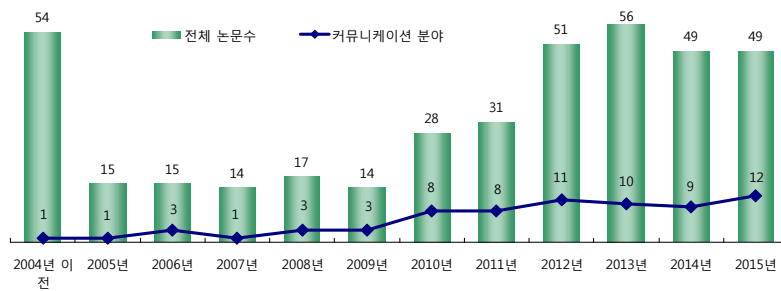
〈그림 2〉 TPB 확장 메타분석 대상 논문 수집 과정

## 5. 연구결과

### 1) TPB 연구 동향 및 연구대상 논문 특성

<그림 3>은 국내에서 TPB를 적용한 연구 동향을 살펴본 결과이다. 그 결과 2011년까지 188편이었으나, 2012년 이후 2015년까지 4년간 205편이 발행된 것으로 나타나 최근 들어 양적으로 크게 증가하고 있는 양상을 보이고 있었다. 특히 2012년 이후 매년 50여편 정도가 지속적으로 발표되고 있었다. 커뮤니케이션 연구분야에 국한하여 보더라도 2010년 이후 크게 증가하여 매년 10여편의 논문이 발표되고 있음을 알 수 있었다.

국내 TPB 연구에 대한 메타분석에 포함된 논문들의 전반적인 특성을 살펴보면 <표 1>과 같다. 총 202편의 논문 중 2010년 이후 발간된 것으로 나타나 최근 들어 활발한 연구가 진행되고 있는 것으로 나타났다. 학문 분야별로는 관광학, 스포츠/체육학에서 가장 많이 다루어지고 있었으며, 그 다음이 광고/커뮤니케이션, 경제/경영학, 의학/간호학/보건학, 심리학 등의 순이었다. 본 메타분석에 포함된 연구논문의 총 표본크기는 78,173명이었다.



<그림 3> 연도별 TPB 연구 동향

〈표 1〉 분석대상논문 특성

		연구수	%
전체		202	100.0
간행 년도	2000년 이전	2	1.0
	2000 ~ 2004년	16	8.1
	2005 ~ 2009년	44	22.0
	2010 ~ 2015년	140	68.9
학술지 구분	등재지	192	94.7
	등재후보지	10	5.3
분야별	언론/광고/커뮤니케이션	33	13.4
	의학/간호학/보건학	15	8.1
	경제/경영학	24	13.4
	체육/스포츠	41	22.0
	심리학	17	8.1
	관광	48	23.0
	기타	24	12.0

## 2) 표집오차의 보정의 결과

앞서 제시된 수식 (식 2.1)과 (식 2.2)를 통해서 표집오차 변량(식 2.3)과 모집단 상관관계 변량(식 2.4)을 적용하여 TPB를 구성하는 주요 개념들간의 효과크기를 HS방법에 따라 표집오차를 보정하였고, 이러한 효과크기를 HO방법(식 1.1~식 1.4)에 의해 산출된 각각의 효과크기와 비교하였다. 이에 대한 결과는 <표 2>에 제시된 바와 같다.

분석결과, 태도와 행동의도 관계(HS방법=.456, HO방법=.500), 주관

〈표 2〉 표집오차 보정에 따른 메타분석 결과 및 비교

관계	메타분석 방법	ES	k	n	$\sigma_{\rho}^2$	$\sigma_{\rho}$	95% CI	80% CV
태도-행동의도	HS방법	.456	240	77,996	.033	.182	.432 ~ .479	.224 ~ .688
	HO방법	.500			-	-	.471 ~ .527	-
주관적 규범-행동의도	HS방법	.402	240	77,114	.036	.191	.377 ~ .426	.157 ~ .646
	HO방법	.438			-	-	.411 ~ .464	-
지각된 행동통제-행동의도	HS방법	.412	240	76,843	.037	.191	.387 ~ .436	.167 ~ .657
	HO방법	.443			-	-	.416 ~ .469	-
지각된 행동통제-행동	HS방법	.340	110	35,475	.034	.185	.304 ~ .375	.103 ~ .577
	HO방법	.392			-	-	.354 ~ .429	-
행동의도-행동	HS방법	.469	109	35,741	.036	.189	.432 ~ .505	.227 ~ .711
	HO방법	.503			-	-	.456 ~ .547	-

ES: 통합 효과크기, k: 효과크기 수, n: 표본크기,  $\sigma_{\rho}^2$ : 추정 평균 상관계수 변량,  $\sigma_{\rho}$ : 추정 평균 상관계수 표준편차, 95CI: 95% 신뢰구간, 80%CV: 80% 확신구간

적 규범과 행동의도(HS방법=.402, HO방법=.438), 지각된 행동통제와 행동의도(HS방법=.412, HO방법=.443), 지각된 행동통제와 행동(HS방법=.340, HO방법=.392), 행동의도와 행동(HS방법=.469, HO방법=.503) 등 TPB에서 가정하고 있는 모든 관계에 대해 HO방법에 의한 메타분석 결과가 표집오차만을 고려한 HS방법보다 약간 높게 나타나고 있었다. 이러한 결과는 HO방법과 HS방법의 차이에 대한 기존의 주장(예: Field, 2001; Johnson et al., 1995)과 일치하는 것으로 나타났다.

## 2) 표집오차 및 측정오차의 통합 보정

표집오차에 이어 구성개념의 측정오차를 보정한 상관관계를 바탕으로 HS방법과 HO방법의 차이를 살펴보았다. 측정오차의 보정은 앞

의 수식(식 2.5)을 적용하여 수행되었으며, 이렇게 산출된 TPB 각 구성변인의 효과크기는 HO방법에 의해 산출된 각각의 효과크기와 비교하였다. 측정오차를 고려할 경우 복합감소계수를 이용하여 상관관계를 보정하는 절차를 거치게 된다. 복합감소계수는 관찰 상관계수를 측정 변인의 측정오차(신뢰도 계수)를 나누는 방식으로 산출할 수 있다. 이에 대한 결과는 <표 3>에 제시된 바와 같다.

그 결과, 태도와 행동의도(HS방법=.524, HO방법=.500), 주관적 규범과 행동의도(HS방법=.472, HO방법=.438), 지각된 행동통제와 행동의도(HS방법=.493, HO방법=.443), 지각된 행동통제와 행동(HS방법=.403, HO방법=.392), 행동의도와 행동(HS방법=.538, HO방법=.503) 등 HS방법에 의한 결과가 HO방법보다 큰 것으로 나타났다. HS방법만을 놓고 보더라도 표집오차만을 고려했을 때에 비해 측정오차까지 보정할 경우 상대적으로 크게 증가하고 있었다. 이러한 결과는 HS방

<표 3> 표집오차와 측정오차 통합 보정에 따른 메타분석 결과 및 비교

관계	메타분석 방법	ES	k	n	$\sigma_{\rho}^2$	$\sigma_{\rho}$	95% CI	80% CV
태도-행동의도	HS방법	.524	220	77,996	.042	.205	.497 ~ .550	.261 ~ .786
	HO방법	.500			-	-	.471 ~ .527	-
주관적 규범-행동의도	HS방법	.472	217	77,114	.074	.272	.437 ~ .507	.124 ~ .820
	HO방법	.438			-	-	.411 ~ .464	-
지각된 행동통제-행동의도	HS방법	.493	219	76,843	.065	.255	.460 ~ .525	.166 ~ .819
	HO방법	.443			-	-	.416 ~ .469	-
지각된 행동통제-행동	HS방법	.403	132	35,475	.044	.209	.363 ~ .443	.136 ~ .670
	HO방법	.392			-	-	.354 ~ .429	-
행동의도-행동	HS방법	.538	104	35,741	.105	.324	.477 ~ .599	.123 ~ .953
	HO방법	.503			-	-	.456 ~ .547	-

ES: 통합 효과크기, k: 개별 효과크기 수, n: 표본크기,  $\sigma_{\rho}^2$ : 추정 평균 상관계수 변량,  $\sigma_{\rho}$ : 추정 평균 상관계수 표준편차, 95CI: 95% 신뢰구간, 80%CV: 80% 확신구간

식을 사용한 기존의 메타분석 연구의 결과(예: 최한나 외, 2015)와 일치하는 결과를 보여주었다.

이와 같은 결과를 HS방법에서 제시하는 통계치에 따라 요약 정리하면 다음과 같다. TPB에서 가정하는 태도와 행동의도의 관계에 대한 관찰 평균 상관계수( $\bar{r}$ )는 .456으로 나타났다. 또한 표집오차와 측정오차를 보정하여 추정된 실제 평균 상관(true mean correlation)은 .524이다. 이는 코헨(Cohen, 1988)이 제시하는 효과크기 해석 기준에 따르면, 큰 효과를 갖는 것으로 해석할 수 있다. 80% 확신구간(.261-.786)에서 0을 포함하고 있지 않으므로 분포의 90% 이상이 정적 상관을 보이고 있는 것으로 짐작할 수 있다. 그러나 메타분석에 포함된 상관관계 값들의 전체 변량이 표집오차와 측정오차에 의해 5.05%만이 설명되고 있었으며, 나머지 약 95%의 변량은 다른 요인에 의해 설명되고 있었다. 따라서 헌터와 슈미트(Hunter & Schmidt, 2004)의 75% 규칙을 적용해 볼 때 추가적으로 조절변인에 의한 분석이 필요하다고 해석할 수 있다.

또한 본 연구는 학위논문 등을 제외하고 출판이 된 논문만을 포함하고 있는데서 발생할 수 있는 표집의 대표성을 확인해보는 절차도 수행하였다. 즉 출판편향이 있었는가를 검증하기 위해 로젠탈(Rosenthal, 1979)의 책상서랍 분석(file drawer analysis)과 깔때기 그림을 이용하였다. 책상서랍 분석을 보면 본 메타분석의 결과를 무효화하기 위해서는 상당히 많은 수의 독립 표본들이 존재해야 하는 것으로 나타났다(fail-safe  $N=231,951 \sim 1,896,667$ ). 즉 가장 적은 안전성 계수를 보인 지각된 행동통제와 행동의도에 대한 실제 점수 상관의 경우 본 연구결과를 무효로 하기 위해서는 효과크기가 0이거나 확인되지 않은 논문(또는 개별 연구수)이 최소한 231,951개가 존재해야만 한다. 따라서 국내에서 TPB 구성개념간의 관계에 대한 효과크기는 책상서랍의 문제가 존재하기 어렵다고 볼 수 있다.

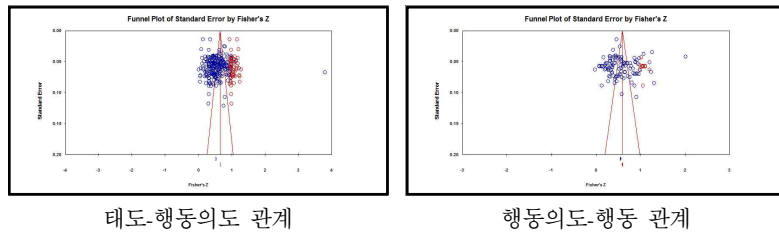
〈표 4〉 HS방법에 의한 TPB 메타분석 결과

관계	<i>k</i>	<i>N</i>	$\bar{\rho}$	$SD_{\rho}$	80%CI	95%CI	%Var	$\bar{r}$	<i>SD<sub>r</sub></i>	<i>f<sub>iN</sub></i>
ATT-INT	240	77,996	.524	.205	.261~.786	.497~.550	5.05	.456	.182	1,896,667
SN-INT	240	77,114	.472	.272	.124~.820	.437~.507	3.45	.402	.191	1,389,344
PBC-INT	240	76,843	.493	.255	.166~.819	.460~.525	3.99	.412	.191	1,389,797
PBC-BEH	110	34,375	.403	.209	.136~.670	.363~.443	5.76	.340	.185	231,951
INT-BEH	109	35,741	.538	.324	.123~.953	.477~.599	2.03	.469	.189	374,935

ATT: 행동에 대한 태도, SN: 주관적 규범, PBC: 지각된 행동통제, INT: 행동의도, BEH: 행동, *k*: 개별 효과크기 수, *N*: 전체 사례수,  $\bar{\rho}$ : 추정된 실제 점수(true score),  $SD_{\rho}$ : 실제 점수 상관계수의 추정된 표준편차, 80%CI: 실제 점수 상관계수의 확신구간(credibility interval), 95%CI: 실제 점수 상관계수의 95% 신뢰구간(confidence interval), %Var: 통계적 인공물에 의해 설명되는 변량의 비율,  $\bar{r}$ : 표본크기를 가중치한 평균 상관계수, *SD<sub>r</sub>*: 표집오차 변량을 제거한 후 평균 상관계수의 표준편차, *f<sub>iN</sub>*: 현 효과크기를 "0"으로 만들기 위해 필요한 개별 연구수(fail-safe N)

깔때기 그림은 출판편향이 있는지 없는지를 육안으로 확인해 볼 수 있는 방법이다. 이 방법은 효과크기와 표준오차의 분포가 좌우 대칭 형태를 보이는가를 바탕으로 좌우 대칭을 이루고 있으면 출판편향이 없거나 심하지 않은 것으로 파악하게 된다. 듀벨과 트위디(Duval & Tweedie, 2000)의 가지치기와 채움 방식을 적용하여 깔때기 그림을 통해 출판편향을 살펴본 결과, TPB에서 가정하고 있는 태도와 행동의도, 주관적 규범과 행동의도, 지각된 행동통제와 행동의도, 지각된 행동통제와 행동, 행동의도와 행동의 모든 관계에서 좌우 대칭 분포를 보이고 있어 출판편향은 발견되지 않음을 확인할 수 있었다. 다만 미처 포함시키지 않은 관련 연구가 축적되거나 좌우 대칭을 위해 임의의 결과가 추가될 경우 실제 효과크기는 현재 나타난 그것보다 다소 높아질 수 있음을 보여주고 있었다. <그림 4>는 CMA 프로그램을 통해 구현한 본 연구의 태도와 행동의도, 행동의도와 행동의 관계에 대한 깔때기 그림이다.





〈그림 4〉 깔때기 그림

## 6. 결론 및 논의

경험적 연구를 수행하는 과정에서 적절한 연구방법론의 사용은 특정 이론을 지지하고 발전시키는데 있어 필수적인 요건이다. 커뮤니케이션을 포함한 모든 연구에서 올바른 분석방법을 선택하고 사용하는 것은 연구문제나 연구가설을 검증하는데 그치지 않고 이론적 추론을 완성하는데 중요한 역할을 담당한다. 커뮤니케이션 현상을 설명하기 위한 연구방법도 다양한 분석기법을 통해 이론을 형성하고 정립하는 도구로 확고하게 자리매김하고 있으며 방법론 자체의 정교화를 통해 지속적으로 발전하고 있다. 이 점에서 메타분석은 커뮤니케이션 연구 분야의 이론을 보완하고 공고히 할 수 있는 연구방법론 가운데 하나가 될 것이다.

메타분석은 체계적 문헌고찰을 수행하는 과정에서 2개 이상의 개별 연구들의 요약 추정치를 통합함으로써 효과에 영향을 미치는 중재 요인들을 모두 고려한 통합된 가중 평균 추정치를 정량적으로 산출하여 전반적인 효과성을 평가하기 위해 사용되는 통계적 기법이다. 메타분석을 수행함으로써 얻을 수 있는 이점으로는 다음과 같은 몇 가지를 들 수 있다(한국보건의료원, 2011). 첫째, 검정력을 증가시킬

수 있다. 표본크기와 관련해서 한계가 있을 수밖에 없는 개별 연구들을 통합하게 되면 보다 명확한 효과를 발견해낼 기회가 더 높아지게 된다. 둘째, 정밀성을 향상시킬 수 있다. 중재 요인에 대한 효과 추정치는 보다 많은 정보가 있을 때 훨씬 엄밀한 분석을 실시할 수 있다. 셋째, 개별 연구들이 파악하지 않은 질문에 답할 수 있다. 서로 다른 특성을 지닌 개별 연구들을 포함하여 메타분석을 수행한 경우 개별 연구 사이에서 중재 효과의 일관성을 조사할 수 있으며, 조사된 효과 추정치에 차이가 있는 경우 이에 대한 이유를 설명할 수 있게 된다. 넷째, 상반된 연구 사이에서 발생하는 논쟁을 조정하거나 새로운 가설을 설정할 수 있다. 결과에 대한 통계적 분석은 상반된 정도를 평가할 수 있으며, 개별 연구에서 서로 다른 결과에 대한 이유를 파악하고 이를 정량화할 수 있다.

이와 같은 메타분석의 의의는 헬스커뮤니케이션 연구에서도 그대로 찾아볼 수 있다. 사람들이 바람직한 건강행동을 할 것인지 혹은 하지 않을 것인지는 다양한 영향 요인을 받아 형성된다. 즉 환경에 의해 건강신념이 바뀔 수 있고, 건강신념에 따라 건강행동이 변화할 수도 있다(정의철, 2013). 일반적으로 흡연이나 감염 등은 사회경제적 지위가 낮은 계층에서, 정기검진과 의사 상담은 사회경제적 지위가 높은 계층에서 많이 이루어지는 등 건강 불평등 현상은 질병의 발생이나 사망률, 건강 관련 행동에서도 나타나는 것으로 알려져 있다. 사회경제적 지위가 높은 사람은 건강관리를 위해 내시경이나 정기검진을 받을 기회도 많다고 한다(Viswanath et al., 2006). 사회경제적 환경 요인을 포함하여, 특정 행동에 대한 지식이나 건강상태에 대한 주관적 인식 등과 조절변인에 따른 건강행동의 차이나 그 역할을 규명하는 것도 메타분석을 통해서만 가능하다. 축적된 연구결과를 바탕으로 사람들의 건강증진을 위한 건강 관련 신념, 태도를 강화하거나 설득 커뮤니케이션 메시지 전략을 어떻게 구성할 것인가 하는 점

도 헬스커뮤니케이션과 메타분석 연구방법이 접목될 수 있는 부분이라고 할 수 있다.

본 연구는 헬스커뮤니케이션 연구 영역에서 메타분석 접근 방법을 통한 공통 분모를 찾기 위한 하나의 시도이다. 이를 위해 국내에서 수행된 TPB를 적용한 연구에 나타난 실제 자료를 이용하여 메타분석의 두 가지 접근방법에 따른 결과 차이를 살펴보기 위해 진행하였다. TPB의 경우 이론을 주창한 에이젠(Ajzen, 1985)이 주요 개념의 조작적 정의와 함께 이들을 어떻게 측정해야 하는가에 대해서도 예시를 제공하고 있으므로, 연구자에 의한 자의적 개념 해석 여지가 상대적으로 적다고 판단하였다. 즉 TPB는 비교적 일정한 가이드에 의거하여 주요 개념이 측정되고 있으므로 측정의 타당도를 일반화하는데 대표적인 이론이 될 수 있으며, 구성개념 단위의 효과크기를 보정하는 HS방법을 검증하는데 적절할 것으로 가정할 수 있었다. 이를 위해 2015년까지 국내 학술지에 게재된 논문 202편을 대상으로 메타분석의 두 가지 방법이 제시하는 절차에 따라 결과를 산출·비교하였다.

본 연구에서 도출된 결과를 보면 먼저, 표집오차만을 보정하였을 경우 TPB의 모든 구성변인의 관계에서 HS방법에 의한 효과크기가 HO방법의 효과크기보다 비교적 작은 것으로 나타났다. 이러한 차이는 HS방법과 HO방법의 대표적 차이 중 하나라고 볼 수 있는데, 즉 HS방법이 변환하지 않은 원래의  $r$ 값(untransformed  $r$  values)을 사용하는데 반해 HO방법은 피셔의  $z$ 변환(Fisher's  $z$ -transformation)을 사용하는데서 기인한다(Field, 2001; Johnson, Mullen, & Salas, 1995). 이에 대해 슈츠(Schulze, 2007)는 효과크기를 산출하는데 있어 상관계수  $r$  자체를 이용하지 않고 피셔의  $z$ 로 변환한 값을 이용하는 것은 변량을 안정화시킬 수 있으므로 더 좋다고 할 수 있으며, 특히 통합 추정치의 변량이 작을 경우 이를 역변량 가중치로 이용하는 것은 통계적으로 이상적인 방법이 될 수 있다고 주장한다. 그러나  $r$ 값 자체에 편차가 있으

므로 표본크기가 크지 않을 때 상관관계 추정치의 표준오차를 가중치로 적용하게 되면 무시할 수 없는 왜곡된 메타분석 결과를 산출할 수 있기 때문에, 연구결과물이 이질적이라고 예상되는 경우  $r$ 이 아닌 피셔의  $z$ 값으로 변환하여 사용하는 것은 바람직하지 않으므로 HS방법을 통해 효과크기를 산출할 것을 권고하고 있다(Schulze, 2007). 필드(Field, 2001) 역시 자신의 몬테카를로 시뮬레이션 연구에서 연구결과물이 동질적인 상황(고정효과)에서는 HO방법의 효과크기가 약간 과대 추정되는 경향을 보인 반면 HS방법에서는 효과크기가 약간 과소 추정되는 경향을 보였지만, 이러한 차이는 반올림 오차(rounding error) 수준의 미세한 차이라고 주장하였다. 그러나 연구결과물이 이질적인 상황(임의효과)에서는 두 방법 간의 효과크기 추정에는 상당한 차이를 발견할 수 있었는데, 이는 본 연구의 메타분석 자료가 임의효과모형이라는 점을 고려할 때 유사한 결과라고 할 수 있다. 필드는 임의효과 상황에서는 오히려 HS방법이 HO방법에 비해 보다 정확한 모집단 효과크기를 추정하고 있다고 주장한다. 만일 우리가 실제 세계의 자료에서는 고정효과모형이 존재하는 경우는 거의 불가능하고 임의효과모형이 보다 현실적이라는 주장(예: Field, 2001; Hunter & Schmidt, 2000; Osburn & Callender, 1992; Schmidt & Hunter, 1999b)을 지지한다면 본 연구의 결과는 메타분석 연구자에게 실제적 함의를 줄 수 있을 것이다.

본 연구의 두 번째 결과는 HS 메타분석과 HO 메타분석이 갖는 또 하나의 본질적인 차이, 즉 모집단 상관계수를 추정하는 과정에서 감소 요인을 보정하는가 혹은 보정하지 않는가와 관련지어 해석할 수 있다. 본 연구에서 표집오차와 함께 구성개념의 측정오차를 보정한 경우 TPB의 모든 구성변인의 관계에서 HS방법에 의한 효과크기가 HO방법의 효과크기보다 상당히 높은 것으로 나타났다. 이러한 결과는 홀과 브래닉(Hall & Brannick, 2002)의 몬테카를로 시뮬레이션 연구

의 결과와 유사한 것으로 보인다. 이들의 연구에 따르면, 통계적 오류에 의한 감소(attenuation)가 없는 조건에서는 HS방법과 HO방법 모두 비교적 덜 편향되게(unbiased) 모집단 상관관계( $\rho$ )를 추정하였지만, 감소가 있는 조건에서는 감소보정(correction for attenuation)을 HS방법의 결과는 표준편차가 없거나( $\sigma_\rho = .00$ ) 작을 때( $\sigma_\rho = .15$ )에는 모집단 상관관계( $\rho$ )가 미미하게 과소 추정되거나 과대 추정되지만, 현저하게 클 때( $\sigma_\rho = .30$ )에는 모집단 상관관계( $\rho$ )의 상당한 과대추정을 보이는 것으로 나타났다. 그러나 감소보정(correction for attenuation)을 하지 않은 HO방법에서는 모든 조건에서 모집단 상관관계( $\rho$ )가 심하게 과소 추정되는 것으로 나타났다. 홀과 브래닉(Hall & Brannick, 2002)은 또한 존슨과 동료들(Johnson, Mullen, & Salas, 1995)의 연구결과<sup>3)</sup>에서 HO방법과 HS방법이 큰 차이를 보이지 않는 것도 통계적 오류에 의한 감소 요인을 고려하지 않았기 때문이라고 반박한다. 통계적 오류와 같은 감소 요인이 있을 경우 HO방법이나 HS방법중 어떤 방법을 이용하여 효과크기를 산출할 것인가는 가중 평균 통합 효과크기에서 비교적 큰 차이를 보이고 있기 때문이라고 주장하였다. 결국, 연구자가 HO방법과 HS방법 중 어떠한 메타분석 방법을 택할 것인가 하는 문제는 이들 방법 중 어느 방법이 더 좋은가를 판단하는 것보다는 연구자가 예상되는 통계적 오류를 수정할 것인가의 여부와 더 밀접하

3) 존슨과 동료들(Johnson, Mullen, & Salas, 1995)은 가상 사례를 바탕으로 HO방법과 HS방법을 비교·분석하였는데, 이들의 연구에 따르면 개별 효과크기를 추정하는데 있어 HO방법이나 HS방법 어떤 방법을 적용하는가에 관계없이 모두 비슷한 추정치를 보이며, 다만 신뢰구간과 같은 변량 측면에서 HO방법에 비해 HS방법이 크게 나타나 대규모 연구결과를 통합하는 과정에서는 HS방법에 비해 HO방법을 사용하는 것이 예측의 정확성 측면에서 볼 때 더 우수하다고 주장하였다. 이에 대해 슈미트와 헌터(Schmidt & Hunter, 1999a)는 반박 논문을 통해 존슨과 동료들의 연구결과는 HS방법의 기본 가정과 효과크기 산출 과정을 잘못 적용해서 나타난 결과라고 주장하고 있다.

게 관련이 있다고 할 수 있다.<sup>4)</sup>

메타분석의 목적이 분석에 포함되는 연구들이 완벽하게 수행되었다는 가정에서 관찰된 관계를 추정하는 것(Hunter & Schmidt, 2004)에 있다면, HS메타분석 방법이 HO메타분석 방법보다 메타분석의 기본 목적에 충실한 방법론이라 할 수 있다(Aguinis, Sturman, & Pierce, 2008; Hall & Brannick, 2002). HO 메타분석은 측정치 수준(measure or scale level)을 나타내는 평균 효과크기를 추정하는데 초점을 맞추고 있으나, HS 메타분석은 구성개념 수준(construct level)의 관계를 통계적으로 추정하는데 초점을 맞추고 있다. 이렇게 볼 때 전반적인 평균 효과크기(overall mean effect size)를 살펴보는데 관심이 있다면 HO방법을, 여기서 한 걸음 더 나아가 다양한 통계적 오류, 즉 표집오차나 측정오차(혹은 범위제한) 등을 고려한 실제 모집단의 진정한 효과(true effect), 그리고 측정의 엄밀성을 확보해야 하는 경우에는 HS방법에 기초한 메타분석을 실시하는 것이 바람직하다고 할 수 있다. 특히 개념을 어떻게 측정할 것인가에 대한 논란이 있는 경우 HS방법에 기초한 메타

4) 측정오차의 보정이 과학적 성과의 정확한 측정과 과학적 이론의 평가에 매우 중요하다는 것이 지속적으로 역설되어 왔음에도 불구하고, 아직까지 측정오차와 이것이 연구결과에 미치는 영향 등에 대한 논의나 실행은 미비하다는 슈미트와 헌터(Schmidt & Hunter, 1999b)의 비판에도 불구하고, 측정오차의 감소보정을 실행하는 것이 옳은 것이냐에 대해서는 상당히 오랫동안 논란(예: Nunnally, 1978; Seymour, 1988)이 계속 되어온 것 또한 사실이다. 1950년대부터 시작된 측정오차의 감소보정에 대한 학자들간의 논란의 핵심은 측정오차의 보정이 실제 상관관계의 강도를 비이상적으로 과대추정하거나(예: Johnson, 1950; Seymour, 1988), 보정의 대상이 되는 적절한 신뢰도 계수의 유형(예: Raju et al., 1993), 보정공식의 오류(예: Muchinsky, 1996)에 관한 것으로 대별될 수 있다. 특히 지금으로부터 약 110년 전에 스피어맨(Spearman, 1904)에 의해 제시된 보정공식은 측정오차의 감소보정의 심각한 비판 대상이 되고 있는데, 즉 특정 조건 하에서 보정된 상관관계 값이 한계치인 1을 넘는 경우가 발생한다는 점이다(이에 대한 자세한 논의는 Muchinsky, 1996; Schmidt & Hunter, 1999b를 참조).

분석의 사용은 개념적 타당도를 한층 높이는데 크게 기여할 수 있다. 그러나 현실적으로 발표되는 연구논문에서 통계적 오류를 교정하는데 필요한 정보를 전부 얻기 어려울 뿐만 아니라, 이런 정보들을 평균치로 간주하여 통계적으로 추론하는 HS 메타분석 또한 방법론적 한계를 갖고 있다. 게다가 최근 들어 R 프로그래밍을 활용한 HS방법의 메타분석이 가능해져 연구자들이 보다 손쉽게 HS방법의 메타분석을 수행할 수 있다 하더라도, R 프로그래밍에 익숙하지 않은 연구자에게 HS방법을 이용한 메타분석은 여전히 쉽지 않은 것이 사실이다. 반면, 메타분석을 위한 전용 프로그램(예: Comprehensive Meta Analysis)을 사용하는 HO방법의 메타분석은 HS방법에 비해 산출과정이 복잡하지 않고 분석결과들을 이용자 입장에서 확인이 용이하며, 특히 출판편향을 확인할 수 있는 다양한 방법들을 시각적으로 그리고 자동 계산해 줌으로써 연구 분석 결과들을 여러 측면에서 살펴보는데 편리함이 적지 않다. 그럼에도 불구하고 아직까지는 두 가지 메타분석 방법 모두 각기 방법론적 장단점을 갖고 있기 때문에 어떤 연구 주제에 어떤 메타분석 방법을 적용하는 것이 좋은 것인가에 대한 판단과 분석방법론의 차용은 여전히 연구자의 몫으로 남겨질 수밖에 없다.

국내 커뮤니케이션 연구 분야에서 메타분석을 이용한 연구는 아직 초보적 단계에 머무르고 있다고 해도 과언이 아니다. 논문 제목이나 초록, 키워드 검색 등을 통해 메타분석이라는 용어가 포함되어 있다고 하더라도 엄밀한 의미의 메타분석, 즉 효과크기 분석이나 선행 연구결과들을 다시 통계적으로 재분석하는 연구가 아닌 체계적 리뷰나 내용분석을 가리키는 경우도 적지 않았다. 또한 몇 편의 메타분석 연구도 관련 연구주제에 대한 잠정적 결론을 도출하고 향후 연구방향을 제시하는 메타분석 연구의 일차적인 목적에 충실하게 진행되고 있었다. 한 걸음 더 나아가 본 연구는 연구방법론으로서 메타분석을

수행하는 과정 그 자체에 관심을 두고 메타분석을 둘러싼 두 가지 큰 갈래에 대한 비교, 분석을 실시하였다. 이를 위해 본 연구는 2015년까지 국내에서 수행된 TPB를 적용한 개별 연구들을 대상으로 이를 통합하는 방법론적 차이점을 다루고 있다는 점에서 그 의의를 찾을 수 있다. TPB를 구성하는 개념은 연구자간 개념화나 조작적 정의, 측정에 있어 에이젠(Ajzen)이 제시한 틀 내에서 사용되고 있으므로 커다란 편차를 보이지 않는다. TPB는 측정의 타당도 일반화에 관심을 갖는 HS방법과 많은 연구영역에서 두루 사용되고 있는 HO방법에 의한 결과치를 비교하는데 적절한 이론이라고 할 수 있다. 더불어 과거 수행된 메타분석(손영곤·이병관, 2012)과의 연장선상에서 전반적인 경향을 살펴보는 데도 유용할 것이라고 판단하였다.

본 연구는 TPB를 구성하는 개념간의 실제 상관관계가 어떻게 나타나는가를 두 가지 메타분석 방법을 통해 비교·분석하여 측정과 관련된 통계적 인공물에 의한 모수 추정의 차이를 살펴보았으나 본 연구 또한 몇 가지 한계가 있다. 첫째, 전체 메타분석 결과 관찰 변량에서 통계적 인공물이 설명하는 비율이 약 2%에서 5%로 나타나 하위집단 분석을 가능하게 하는 여러 조절변인이 있을 수 있으나 본 연구에서는 이를 다루지 못하였다. 고려해 볼 수 있는 조절변인으로 TPB 변인이 몇 개의 하위 차원으로 구성되어 있다는 점을 들 수 있다(Ajzen, 2002, 2006). 즉 행동에 대한 태도는 경험적(experiential) 태도와 도구적(instrumental) 태도, 주관적 규범은 지시적(injunctive) 규범과 기술적(descriptive) 규범, 지각된 행동통제는 자기효능감(self-efficacy)과 통제력(controllability) 등으로 구성되어 있다는 주장은 이를 대변하고 있다. 연구대상, 적용 분야, 연구시기 등과 같은 연구설계와 같은 조절변인에 의한 차이도 살펴볼 필요가 있다. TPB에 대한 국내외 메타분석(손영곤·이병관, 2012; Armitage & Conner, 2001) 결과에서도 개념의 측정, 연구환경변인에 따른 차이가 나타나고 있었다. 특히 손영



곤·이병관(2012)의 연구결과와 비교해 볼 때 TPB 구성개념간의 관계에 대한 효과크기가 증가하고 있었는데, 이는 본 연구결과의 출판편향에서 볼 수 있듯이 포함시키지 않은 연구가 추가될 경우 효과크기가 다소 높아질 수 있음과 맥을 같이 한다고 할 수 있다. 둘째, 국내에서 2015년까지 축적된 연구논문중 메타분석이 요구하는 정보나 통계치를 보고하지 않은 상당수의 논문을 배제할 수밖에 없었다. 메타분석은 선행 연구에 나타난 결과 값을 기초 자료로 하여 통계적으로 재분석하는 연구방법이기 때문에 선행 연구에 필요한 통계치가 제시되지 않을 경우 포함시키지 못하였으며, 이에 따른 출판편향의 문제도 발생할 수 있다. 계량적 연구를 진행하는 경우 추후 후속 연구자에 의한 메타분석의 가능성을 염두에 두고 논문을 작성하는 자세가 필요한 것도 이 때문이다. 이 점에서 섉스톤(Saxton, 2005)이 제시하고 있는 모든 변인의 평균, 표준편차, 표본크기 명시, 구체적인  $p$ 값 표기 등 메타분석 활성화를 위한 조건이나 한국언론학회나 한국심리학회의 학술지 논문 작성 지침은 논문을 작성할 때 유념해야 할 것으로 판단된다. 셋째, 신뢰도를 보고하지 않거나 단일 문항으로 측정하여 측정오차를 완벽하게 통제하지 못하였다. 본 연구에서 총 267개의 개별 효과크기에서 태도 24개, 주관적 규범 37개, 지각된 행동통제 31개, 행동의도 58개, 행동 52개에서 단일 문항으로 측정하여 신뢰도를 1로 간주하거나 신뢰도가 제시되지 않아 평균 신뢰도로 대체한 것에서 발생할 수 있는 실제 모집단 상관계수 추정의 오류도 전혀 무시할 수 없었다. 넷째, 보고된 결과치의 한계로 인해 범위제한과 같은 다양한 통계적 오류를 모두 포괄하지 못하고 표집오차와 측정오차만을 고려할 수밖에 없었다는 점에서 HS메타분석에서 제시하는 측면을 종합적으로 다루지 못했다는 점도 연구의 한계로 지적하고자 한다.

## 참고문헌

- 김성태 (2005). 국내 내용분석 연구의 방법론에 대한 고찰 및 제언. <커뮤니케이션이론>, 1권 2호, 39-67.
- 김세경·천성문 (2015). 자기통제력과 관련된 변인에 대한 메타분석: 가정환경 변인군과 개인 심리 변인군을 중심으로. <한국심리학회지: 상담 및 심리치료>, 27권 1호, 85-107.
- 박형인·남숙경·양은주 (2011). 직무소진과 직무태도 및 이직의도와 의 관계: 메타분석적 문헌 고찰. <한국심리학회지: 산업 및 조직>, 24권 3호, 457-491.
- 손영곤 (2014). 메타분석에 의한 국내 유명인 모델의 광고효과 연구 현황 및 개관. <한국광고홍보학보>, 16권 2호, 183-220.
- 손영곤·이병관 (2012). 계획된 행동이론을 적용한 사회인지적 행동 모델의 유용성에 대한 메타분석. <한국언론학보>, 56권 6호, 127-161.
- 신인수·박은영 (2011). 특수 교육 및 관련 분야의 메타분석논문에 대한 고찰. <중북·지체부자유연구>, 54권 4호, 157-176.
- 오성삼 (2002). 메타분석의 이론과 실제. 서울: 건국대학교 출판부.
- 오인수·김광현·Darnold, T. C.·황중오·유태용·박영아·박량희 (2007). 직무만족, 조직몰입, 성과, 이직의도간의 관련성: 문헌고찰 및 메타분석. <인사조직연구>, 15권 4호, 43-86.
- 우형진 (2005). 국내 인터넷 연구의 발전적 모색: 1995~2005년까지의 커뮤니케이션 학술지에 게재된 인터넷 연구에 대한 메타분석. <커뮤니케이션이론>, 1권 1호, 332-366.
- 유성경·박은선·김수정·조효진 (2014). 삶의 의미와 정신건강과의 관계에 대한 메타분석: 1997년-2013년 국내 연구를 중심으로. <한국심리학회지: 일반>, 33권 2호, 441-465.

- 유태열 (2009). 조직공정성과 직무만족간 관계 및 조절 변수에 관한 메타분석. 서울: 성균관대학교 석사학위논문.
- 이병관·손영곤·서동명·좌보경·홍현호·이진우 (2013). 지난 40년간 공포소구 연구의 통합: 국내 공포소구 연구에 대한 메타분석. <한국광고홍보학보>, 15권 3호, 126-155.
- 이병관·손영곤·이상록·윤문영·김민희·김채린 (2014). 건강 관련 행동의 예측을 위한 사회인지이론의 유용성: 국내 건강신념모델 연구의 메타분석. <홍보학연구>, 18권 2호, 163-206.
- 이준영 (2008). 메타분석. <대한내분비학회지>, 23권 6호, 361-378.
- 전근영·이정교 (2015). 국내 유머 광고에 대한 메타 분석. <한국언론학보>, 59권 6호, 477-504.
- 정의철 (2013). 헬스커뮤니케이션. 서울: 커뮤니케이션북스.
- 조경호 (1998). Hunter의 메타방법론을 적용한 조직몰입과 직무만족간의 상관관계 분석. <한국행정학보>, 32권 2호, 147-164.
- 조용래·노상선·조기현·홍세희 (2014). 우울과 불안증상에 대한 마음 챙김에 기반을 둔 개입의 효과: 메타분석. <한국심리학회지: 일반>, 33권 4호, 903-928.
- 최한나·권윤주·강귀옥 (2015). 작업동맹과 상담성과의 관계에 대한 메타분석(1990-2014). <상담학연구>, 16권 3호, 85-99.
- 한국보건사회연구원 (2013). 한국의 건강불평등 지표와 정책 과제. 서울: 한국보건사회연구원.
- 한국보건의료연구원 (2011). NECA 체계적 문헌고찰 매뉴얼. 서울: 한국보건의료연구원.
- 허만용 (2008). Hunter & Schmidt의 메타분석을 적용한 시장지향과 성과의 관계: 목표와 문화의 조절효과. <한국정책과학학회>, 12권 1호, 45-70.
- 홍세정·장재윤 (2015). 내적 동기와 창의성간의 관계에 대한 메타분

- 석. <한국심리학회지: 일반>, 34권 1호, 57-86.
- 황상재 · 박석철 (2004). 국내 인터넷 연구의 메타분석: 연구 주제와 방법을 중심으로. <한국방송학보>, 18권 2호, 68-91.
- 황용석 (2006). 한국 온라인저널리즘의 연구의 주제와 접근방법에 대한 메타분석. <커뮤니케이션이론>, 2권 1호, 128-169.
- Aguinis, H., Sturman, M. C., & Pierce, C. A. (2008). Comparison of three meta-analytic procedures for estimating moderating effects of categorical variables. *Organizational Research Methods*, 11(1), 9-34.
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior (pp.11-39). Springer Berlin Heidelberg.
- Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, 32, 665-683.
- Ajzen, I. (2006). Constructing a TpB questionnaire: Conceptual and methodological considerations. (n.d.) Retrieved from <http://people.umass.edu/aizen/pdf/tpb.measurement.pdf>.
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Alexander, R. A., Alliger, G. M., & Hanges, P. J. (1984). Correcting for range restriction when the population variance is unknown. *Applied psychological measurement*, 8(4), 431-437.
- Aloe, A. M., & Becker, B. J. (2012). An effect size for regression predictors in meta-analysis. *Journal of Educational and Behavioral Statistics*, 37(2), 278-297.
- Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behavior: A meta-analytic review. *British Journal of Social Psychology*, 40, 471-499.

- Becker, B. J. (2005). Failsafe N or file-drawer number. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Chichester, UK: John Wiley & Sons.
- Borenstein, M. (2005). Software for publication bias. In H. R. Rothstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustment* (pp.194-220). West Sussex, UK: John Wiley & Sons.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: John Wiley & Sons.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed effect and random effects models for meta-analysis. *Research Synthesis Methods*, 1, 97-111.
- Callender, J. C., & Osburn, H. G. (1980). Development and test of a new model for validity generalization. *Journal of Applied Psychology*, 65(5), 543-558.
- Charles, E. P. (2005). The correction for attenuation due to measurement error: Clarifying concepts and creating confidence sets. *Psychological methods*, 10(2), 206-226.
- Cheung, M. W. L. (2015). *Meta-analysis: A Structural equation modeling approach*. West Sussex, UK: John Wiley & Sons, Ltd.
- Cohen, A. C. (1959). Simplified estimators for the normal distribution when samples are singly censored or truncated. *Technometrics*, 1(3), 217-237.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Conner, M., & Sparks, P. (2001). The theory of planned behavior and health behaviors (pp.121-162). In M. Conner, & P. Norman (Eds.).

- Predicting health behavior (pp.1-22). Buckingham: Open University Press.
- Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.
- Cooper, H., & Hedges, L. V. (2009). Research synthesis as a scientific process. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds). *The handbook of research synthesis and meta-analysis* (2nd ed.). NY: Russell Sage Foundation.
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel plot based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455-463.
- Egger, M., Smith, D. W., & Altman, D. G. (2003). *Systematic reviews in health care: Meta-analysis in context* (2nd ed.). London, UK: BMJ Books.
- Egger, M., Smith, D., & Phillips, A. N. (1997). Meta-analysis: Principles and procedures. *BMJ*, 315, 1533-1537.
- Egger, M., Smith, D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629-634.
- Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, 17(1), 120-128.
- Field, A. P. (2001). Meta-analysis of correlation coefficients: a Monte Carlo comparison of fixed-and random-effects methods. *Psychological methods*, 6(2), 161-180.
- Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, 63, 665-694.

- Fletcher, T. D. (2010). *Psychometric: Applied Psychometric Theory*. R package. version 2.2. <http://CAN.R-project.org/package=psychometric>.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational researcher*, 5(10), 3-8.
- Hall, S. M., & Brannick, M. T. (2002). Comparison of two random-effects methods of meta-analysis. *Journal of Applied Psychology*, 87(2), 377-389.
- Hausenblas, H. A., Carron, A. V., & Mack, D. E. (1997). Application of the theories of reasoned action and planned behaviour to exercise behaviour: A meta-analysis. *Journal of Sport & Exercise Psychology*, 19, 36-51.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta analysis*. New York: Academic Press.
- Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, 8(4), 275-292.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2nd ed.). Thousand Oaks, CA: Sage.
- Hunter, J. E., Schmidt, F. L., & Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology*, 91(3), 594-612.
- Johnson, B. T., Mullen, B., & Salas, E. (1995). Comparison of three major meta-analytic approaches. *Journal of Applied Psychology*, 80(1), 94-106.
- Law, K. S., Schmidt, F. L., & Hunter, J. E. (1994). A test of two refinements in procedures for meta-analysis. *Journal of Applied Psychology*, 79(6), 978-986.

- Le, H., & Schmidt, F. L. (2006). Correcting for indirect range restriction in meta-analysis: Testing a new meta-analytic procedure. *Psychological methods*, 11(4), 416-438.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: Sage.
- Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis*. NY: Oxford University Press.
- Muchinsky, P. M. (1996). The correction for attenuation. *Educational and psychological measurement*, 56(1), 63-75.
- Nunnally, J. C. (1978). *Psychometric theory* (2d ed.). NY: McGraw-Hill.
- Osburn, H. G., & Callender, J. (1992). A note on the sampling variance of the mean uncorrected correlation in meta-analysis and validity generalization. *Journal of Applied Psychology*, 77(2), 115-122.
- Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology*, 90(1), 175-181.
- R Development Core Team (2011). *R: A language and environment for statistical computing*. R foundation for statistical computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://project.org/>.
- Raju, N. S., Burke, M. J., Normand, J., & Lezotte, D. V. (1993). What would be if what is wasn't? Rejoinder to Judiesch, Schmidt, and Hunter (1993). *Journal of Applied Psychology*, 78(6), 912-916.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, California: Sage Press.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-Analysis: Recent development in quantitative methods for literature reviews. *Annual*



Review of Psychology, 52, 59-82.

- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005a). Publication Bias in Meta-Analysis. In H. R. Rothstein (Eds.), *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustment*(pp.1-7). West Sussex, UK: John Wiley & Sons.
- Saxton, M. L. (2006). Meta-analysis in library and information science: Method, history, and recommendations for reporting research. *Library Trends*, 55(1), 158-170.
- Schmidt, F. L., & Hunter, J. E. (1999a). Comparison of three meta-analysis methods revisited: An analysis of Johnson, Mullen, and Salas (1995). *Journal of Applied Psychology*, 84(1), 144-148.
- Schmidt, F. L., & Hunter, J. E. (1999b). Theory testing and measurement error. *Intelligence*, 27(3), 183-198.
- Schulze, R. (2007). Current methods for meta-analysis: Approaches, issues, and developments. *Journal of Psychology*, 215(2), 90-103.
- Seymour, R. T. (1988). Why plaintiffs' counsel challenge tests, and how they can successfully challenge the theory of "validity generalization". *Journal of Vocational Behavior*, 33(3), 331-364.
- Spearman, C. (1904). The proof and measurement of association between two things. *The American journal of psychology*, 15(1), 72-101.
- Thorndike, R. L. (1947). *Research problems and techniques* (Rep. No. 3 AAF Aviation Psychology Program Research Reports). Washington, DC: U.S. Government Printing Office.
- Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods*, 10, 428-443.
- Viechtbauer, W. (2010). *Conducting meta-analyses in R with the metafor*

- package. *Journal of Statistical Software*, 36(3), 1-48. URL <http://www.jstatsoft.org/v36/i03/>.
- Viswanath, K., Breen, N., Meissner, H., Moser, R. P., Hesse, B. & Rakowski, W. (2006). Cancer knowledge and disparities in the information age. *Journal of Health Communication*, 11, 1-17.
- Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, 75(3), 315-321.
- Wolf, F. (1986). *Meta-Analysis: Quantitative methods for research synthesis*. Beverly Hills, CA: Sage.
- Zhu, C., Jiang, T., Cao, H., Sun, W., Chen, Z., & Liu, J. (2015). Longitudinal analysis of meta-analysis literatures in the database of ISI Web of Science. *International journal of clinical and experimental medicine*, 8(3), 35-59.

최초 투고일: 2016년 10월 23일  
논문 수정일: 2016년 12월 12일  
게재 확정일: 2016년 12월 14일

**The Effects of Statistical Artifacts on Outcomes in  
Meta-Analysis: The Comparative Analysis of Hedges-Olkin  
and Hunter-Schmidt Meta-Analytic Approaches**

**Lee, Byoungkwan**

Professor, Department of Advertising & Public Relations,  
Hanyang University

**Sohn, Youngkon**

Lecturer, Department of Advertising & Public Relations,  
Hanyang University

This study attempted to review and compare two major meta-analytic approaches, the Hedges and Olkin approach and the Hunter and Schmidt approach, in the aspects of meta-analytic methodology. The basic methodological assumptions and characteristics between the two approaches were also discussed in terms of the analytic techniques, methodological procedures, and interpretations of the results. In addition, the two meta-analyses of the theory of planned behavior, one of the major social cognitive theories, were conducted. Based on the results, Fisher's  $z$ -transformation, effect size estimation, heterogeneity assessment, correction for sampling error and measurement error, and confidence interval and creditability interval were discussed.

*Key words : meta-analysis, statistical artifacts, Hedges-Olkin meta-analysis, Hunter-Schmidt meta-analysis, Theory of Planned Behavior*