



## 지각된 메시지 효과성은 실제 메시지 효과를 얼마나 반영하는가?

설득 캠페인 메시지의 지각된 효과와 실제 효과

**이병현** 성균관대학교 미디어커뮤니케이션학과 연구원

**정성은** 성균관대학교 미디어커뮤니케이션학과 교수

### To What Degree Does Perceived Message Effectiveness Reflect Actual Message Effect?

Perceived vs. Actual Message Effect in Persuasive Campaigns

**Byeong-Hyeon Lee\***

Researcher, Department of Media and Communication, Sungkyunkwan University

**Sungeun Chung\*\***

Professor, Department of Media and Communication, Sungkyunkwan University

To predict the effectiveness of persuasive campaign messages in advance, campaign practitioners utilize perceived message effectiveness (PME). However, it has been debated whether actual message effect (AME) can be predicted with PME. Based on correlational data, some researchers argued that PME is a causal antecedent of AME and can be used in place of AME in formative campaign research. O'Keefe criticized the use of correlations to test the diagnosticity of PME and proposed a new method, comparing the direction of difference in PME and that of AME for message pairs. In a meta-analysis, O'Keefe found that the direction of difference between PME and AME were matched in only 58% of message pairs. To resolve this debate, we critically examined existing testing methods and improved those methods to more rigorously test the PME-AME relationship. Also, applying the method used in the study of perceived polling effect on individuals' attitudes, we suggested to directly compare PME with AME to find the proportion that PME reflects AME. We also classified various types of PME measures into perceived argument quality, estimated message impact on self (or others), and perceived change in attitude of oneself (or others) due to message exposure. We conducted a two-wave panel survey and measured perceived and actual effect of four campaign messages that were created by public agencies to promote the safety of nuclear power plants in South Korea ( $N=475$ ). Initial attitudes of respondents were measured in the first survey, and a week later, four messages were randomly presented to different groups, then various measures of PME and AME were

---

\* zkflldb@gmail.com

\*\* chseun@gmail.com

measured. We found that both across-message and message-specific PME-AME correlations was positive and relatively high. With a modified method of O'Keefe the direction of the difference in PME matched with that of AME in all message pairs (30/30) showing high level of diagnosticity. Respondents with negative initial attitudes underestimated the message effect, reporting their attitude change due to message exposure as 43.2% of the actual attitude change. Those with neutral attitudes reported their attitude change as 76.3% of the actual attitude change. These results suggest that while PME can be used for campaign message design, practitioners need to be aware of the systematic bias it contains.

**Keywords:** perceived message effectiveness, actual message effect, persuasive campaign, estimated message impact, perceived attitude change

# 1. 서론

설득캠페인은 특정 수용자 집단의 신념, 태도, 또는 행동에 영향 끼치기 위한 지속적 활동이다 (Pfau & Parrott, 1993). 설득캠페인은 목표와 수용자를 특정하는 기획단계, 설득 메시지를 제작하여 전달하는 집행단계, 그리고 캠페인의 효과를 평가하는 단계로 구성된다 (Pfau & Parrott, 1993). 캠페인 집행단계에서 실무자들은 제작 중인 메시지를 미리 평가하기 위해 메시지 효과에 대한 조사를 실시한다(Duke et al., 2015; O'Keefe & Reid, 1990). 캠페인 메시지의 실제 효과를 미리 진단하는 이상적인 방법은 목표 수용자의 일부를 대상으로 하여 메시지 노출로 인한 태도나 행동의 변화를 측정하는 것이다(Atkin & Freimuth, 2001; Pfau & Parrott, 1993). 하지만 캠페인 메시지 노출로 인한 실제 태도나 행동의 변화를 사전에 검증하는 방법은 시간적 또는 재정적인 문제로 실행하기 어려운 경우가 많다(Davis & Duke, 2018). 실무자들은 대안적으로 전문가 집단 또는 목표 수용자 일부에게 제작 중인 메시지를 제시하고 메시지가 얼마나 효과적일지를 예측 또는 평가하게 하여 이를 기초로 메시지를 제작한다(Bigsby, Cappella, & Seitz, 2013; Davis & Duke, 2018; O'Keefe, 2018; Pfau & Parrott, 1993; Yzer, LoRusso, & Nagler, 2015). 즉, 메시지 제작자들은 측정하기 어려운 실제 메시지 효과(actual message effect, AME) 대신 지각된 메시지 효과성(perceived message effectiveness, PME)을 측정하여 활용한다(Davis & Duke, 2018). 이 방법은 PME가 AME와 밀접하게 연관되어 있어 AME를 진단할 수 있다는 가정에 기초하고 있다.

딜러드와 동료들(Dillard, Weber, & Vail, 2007; Dillard, Shen, & Vail, 2007)은 메타분석과 상관관계를 이용한 연구들을 통해 이 가정을 경험적으로 검증하였다. 그러나 지각된 메시지 효과성이 실제 메시지 효과와 일치하지 않을 것이라는 주장 또한 제기되었다(O'Keefe, 1993, 2018). 오키프는 1993년 논문에서 PME는 AME와 다른 심리적 과정을 통해 형성되므로 AME와 동일시 될 수 없다는 것을 이론적으로 주장하고, 2018년 논문에서 상관관계를 이용해 PME의 AME에 대한 진단능력 검증하는 방법을 비판하였다. 또한 그는 두 메시지의 지각된 효과성의 순위와 실제 효과의 순위의 일치 여부를 이용하는 방법을 제안하였다. 이 방법을 적용한 메타분석 연구의 결과, 지각된 효과성과 실제 효과의 순위가 일치한 경우의 비율(58%)은 우연적 확률보다 크지 않았다. 오키프는 이를 근거로 지각된 효과성이 실제 효과를 효과적으로 진단할 수 없다고 주장하였다. 많은 학자들이 이 논쟁에 관심을 보여 이 주제에 관해 두 개의 주요한 국제 학술회의 (the Kentucky Conference on Health Communication, 2018; Preconference for ICA's 2019 Annual Conference)가 개최되고, 국제 학술지 특별호

(*Journal of Communication*, Special forum, 2018)가 기획되었다. 하지만 이에 관해 명확한 결론은 내려지지 않았으며 PME 사용에 관한 다양한 주제들(예, PME 측정방식에 따른 차이와 메시지 수용자에 따른 차이)로 확장되고 있다(Nabi, 2018; Noar, Barker, & Yzer, 2018).

PME의 AME에 대한 진단능력을 검증하기 위해 기존 연구들은 두 가지 다른 방법, 즉 지각된 효과성과 실제 효과 간 상관관계를 이용하는 방법(Dillard, Weber, & Vail, 2007; Dillard, Shen, & Vail, 2007) 그리고 두 변수에 대한 순위의 일치여부를 이용하는 방법(O'Keefe, 2018)을 사용하였다. 이 두 가지 방법 이외에 PME와 AME를 서로 비교 가능하게 측정된 후 비교하는 방법이 있다. 선거 여론조사 결과보다 영향력 지각과 관련되어 설득캠페인 메시지 영향력 지각 연구(Chung, Heo, & Moon, 2018)에서 연구자들은 선거 여론조사 결과보다의 영향력 지각의 정확성을 검증하기 위해 보도노출로 인한 실제 태도변화의 정도를 측정하고 이를 응답자가 지각한 태도변화 정도와 직접 비교하였다. 이 방법을 설득캠페인 메시지의 PME의 AME에 대한 진단능력 검증에 적용할 경우, PME가 AME를 정확하게 반영하는지 아니면 어느 정도나 과소 또는 과대하게 반영하는지를 검증할 수 있다. 본 연구는 지각된 효과성과 실제 효과 간의 상관관계를 이용하는 방법, 두 메시지 간 지각된 효과성과 실제 효과 간의 상대적 순위 일치 여부를 이용하는 방법과 더불어 두 변수의 직접적 비교를 통한 검증 방법을 동시에 사용하여 지각된 메시지의 진단능력을 종합적으로 규명하고자 한다.

노어 외(Noar, Bell, Kelly, Barker, & Yzer, 2018)는 기존 연구들이 지각된 효과성을 각기 다른 방식으로 측정된 것이 지각된 효과성의 진단능력을 검증하기 어렵게 만든 이유 중 하나라고 지적하였다. PME의 진단능력을 검증한 기존 연구들은 여러 다른 방식으로 지각된 효과성과 실제 효과를 측정해왔다. 지각된 메시지 효과성은 논증의 품질 등 메시지의 특성(perceived argument quality; Cesario, Grant, & Higgins, 2004; Hale, Mongeau, & Randi, 1991; Hullett & Boster, 2001; Hullett, 2002; Lavine & Snyder, 1996; Mitchell, 2000; Reid, Gunther, & Smith, 2005)에 대한 평가로 측정되거나, 전반적인 메시지의 설득성(perceived persuasiveness; Collins, Taylor, Wood, & Thompson, 1988; Dillard & Peck, 2000; Dillard & Witte, 1993), 또는 자신 또는 다른 사람들에게 얼마나 영향력을 미칠 것인지에 대한 추정(estimated impact on oneself; Cesario et al., 2004; Hullett, 2004; Hullett & Boster, 2001; Mitchell, Brown, Morris-Villagran, & Villagran, 2001; Reid 2005; estimated impact on others; Davison, 1983, Cohen, Mutz, Price, & Gunther, 1988), 그리고 메시지 노출로 발생할 자신 또는 다른 사람들의 태

도 변화 정도(perceived amount of attitude change on oneself or on others; Chung et al., 2018)와 같이 다양한 방식으로 측정되었다. 실제 메시지 효과는 메시지 노출 후 태도(post-message attitude; Lavine & Snyder, 1996), 메시지 노출 사전 사후 태도 변화의 정도(amount of attitude change; Chung, Heo, & Moon et al., 2018; Collins, Taylor, Wood, & Thompson et al., 1986), 또는 행동 의도(behavior intention; Bigsby, Cappella, & Seitz, 2013; Dillard & Witte, 1993; Hullett & Boster, 2002) 등과 같이 다양한 방식을 통해 측정되었다. 본 연구는 PME의 진단능력을 체계적으로 검증하기 위해 PME를 측정유형에 따라 구분하여 각 유형의 진단능력을 검증하고 측정유형 간에 진단능력의 차이가 있는지를 검증한다. AME의 다양한 측정유형을 검토하고 PME의 AME에 대한 진단능력 검증 시 메시지 노출 사전 사후 태도 변화의 정도를 이용한 AME의 측정의 효과성을 논의한다. 본 연구는 PME의 AME에 대한 진단능력 검증에 사용된 세 가지 다른 방법을 여러 다른 PME 측정 유형과 메시지 노출 사전 사후 태도 변화의 정도를 이용하여 종합적으로 검증하고자 한다.

## 2. 이론적 배경

### 1) 지각된 메시지 효과성의 진단능력 검증 방법 1: 지각된 효과성과 실제 효과 간의 상관관계

지각된 메시지 효과성 (PME)는 특정 메시지의 설득성에 대한 지각 (perceived persuasiveness; Dillard & Peck, 2000) 또는 메시지의 수용자의 신념, 태도, 행동 등에 대한 영향 추정 (estimated message impact; Hullett & Boster, 2001)를 가리킨다. 딜러드, 웨버 외(Dillard, Weber et al., 2007)는 상관관계를 이용하여 PME가 실제 메시지 효과 (AME)를 예측할 수 있는지를 검증하였다. 16개 연구의 결과에 대한 메타분석 결과, PME와 AME 간 상관 계수의 평균은 .41로 나타났으며, 95%(38/40)의 경우 상관관계는 정(+)적이었다. 이 결과를 바탕으로 그들은 PME가 AME를 예측할 수 있다고 주장하였다. 딜러드, 웨버 외는 또한 PME의 측정유형을 메시지 특성 측정(attribute measures)과 메시지 영향 측정(impact measures)으로 구분하여 두 측정 유형에 따라 지각된 효과성과 실제 효과 간의 상관관계가 다를 수 있음을 또한 발견했다. 딜러드, 쉐인 외(Dillard, Shen et al., 2007)는 PME와 AME의 간의 인과관계를 검증하기 위해 다섯 개의 실험을 진행하였다. 그들은 메시지 특성 측정과 메시지 영향력 측정을 혼합하여 PME를 측정하고 메시지가 제안한 행동에 대한 메시지 사후 태도

(post-message attitude) 또는 행동의도로 AME를 측정하였다. 그들은 다섯 실험의 결과로 발견된 17개의 PME와 AME 간의 상관관계를 구조방정식으로 분석한 결과를 토대로 PME가 AME의 인과적으로 선행변수임을 주장하였다.

딜러드 외의 주장과 달리 오키프(O'Keefe, 1993; 2018)는 PME가 AME를 유의미하게 예측할 수 없다고 주장하였다. 오키프(1993)는 메시지 수용자들이 설득에 대해 상식적인 신념 (commonsensical beliefs about persuasion)을 가지고 있으며, 이를 근거로 메시지의 설득력을 평가한다고 보았다. 하지만 설득에 관한 수용자들의 상식적인 신념들에는 오류가 많이 있으므로 PME가 메시지의 실제 효과를 제대로 반영하지 못한다고 주장했다. 또한 오키프(2018)는 PME의 진단능력을 검증하기 위해 PME와 AME의 상관관계를 이용하는 것에 의문을 제기하였다. 오키프는 두 메시지에 대한 PME와 AME의 가상의 자료를 제시한 후, 메시지별 PME와 AME의 상관관계가 +1.00임에도 불구하고 PME의 평균이 상대적으로 높은 메시지가 AME의 평균은 상대적으로 낮을 수 있음을 보여주었다. 이 예들을 토대로 오키프는 특정 메시지에 한정된 PME와 AME의 상관관계(message-specific correlation)는 AME에 대한 PME의 진단능력의 증거가 될 수 없다고 주장하였다.

오키프가 사용한 가상 자료에서 각 메시지에 한정된 상관관계는 두 메시지 모두 +1.00이지만, 전체 자료에 대한 PME와 AME의 상관관계(across-message correlation)는 -.96이었다. 상관관계는 전체 표본에서는 부(-)적일지라도 표본의 하위집단에서는 정(+)적일 수 있다(즉, 심슨의 역설, Wagner, 1982). PME-AME의 상관관계가 개별 메시지에서는 정(+)적이더라도 전체 메시지에서 부(-)적이라면 그 메시지들의 실제 효과를 예측하기 위해 PME를 이용할 수 없다. 하지만 오키프의 주장은 한편 PME-AME 상관관계가 전체 표본에서도 정(+)적이고 개별 메시지에 대해서도 정(+)적이라면, AME의 진단을 위해 PME를 사용할 수 있음을 의미한다. 특히, PME-AME 상관관계가 전체 표본과 개별 메시지 모두에서 정(+)적인 경우, 특정 메시지의 PME 평균이 다른 비교 메시지의 평균보다 통계적으로 유의미하게 크다면, 해당 메시지의 AME가 비교 메시지의 AME보다 큼을 합리적으로 예측할 수 있을 것이다. 즉, 전체 표본과 메시지별 표본에서 PME-AME 상관관계가 모두 정(+)적인 조건에서는 해당 상관관계들이 PME의 AME에 대한 진단능력의 정도를 지표하고 있다고 볼 수 있다. 기존 연구에서는 전체 표본과 메시지 표본별로 PME의 AME 간 상관관계를 보고하지 않았다. 본 연구는 PME-AME 상관관계를 이용한 PME의 검증능력 평가와 관련해 다음과 같은 연구 문제를 제시하고 검증하고자 한다.

**RQ1:** 복수의 캠페인 메시지에 대하여 전체 표본에서의 PME와 AME 간의 상관관계와 메시지별 표본에서의 PME와 AME 간의 상관관계는 서로 일치하는가 그리고 상관관계의 정도는 어떠한가?

## 2) 지각된 메시지 효과성의 진단능력 검증 방법 2: 두 메시지 간 PME와 AME 순위 비교

오키프(2018)는 AME에 대한 PME의 진단능력을 검증하기 위해 상관관계가 아니라 경쟁하는 두 메시지의 PME에 관한 순위와 AME에 관한 순위가 일치하는지를 확인해야 한다고 주장하였다. 그는 30개의 연구에서 서로 비교 가능한 151개의 메시지 쌍(message pairs)을 추출하여 두 메시지의 PME의 상대적 순위와 AME의 상대적 순위가 일치하는지 확인하였다. 오키프는 두 메시지 간 PME 순위와 AME 순위의 일치 여부를 판정하기 위해 다음의 기준을 사용하였다. 첫째, 특정 메시지가 비교 메시지보다 PME가 높고 AME도 높은 경우(즉,  $PME_{M1} > PME_{M2} \ \& \ AME_{M1} > AME_{M2}$ ), PME의 진단능력이 있는 경우로 판정하였다. 둘째, 두 메시지의 PME는 서로 다르지만 AME가 같을 경우( $PME_{M1} \neq PME_{M2} \ \& \ AME_{M1} = AME_{M2}$ ), PME가 실제 효과가 더 낮은 메시지를 선택하지 않게 했다는 점에서 PME의 진단능력이 있다고 보았다. 셋째, 두 메시지의 PME는 같지만 AME가 서로 다를 경우 ( $PME_{M1} = PME_{M2} \ \& \ AME_{M1} \neq AME_{M2}$ ), 측정된 PME가 실제 효과가 더 낮은 메시지를 선택하게 할 수도 있으므로 진단능력이 없다고 보았다. 분석결과, PME의 우월성과 AME의 우월성이 일치한 경우는 58%로 우연적 기댓값(50%)과 통계적으로 유의미하게 다르지 않았다. 오키프는 이를 근거로 PME가 AME를 제대로 진단하지 못한다고 주장하였다.

오키프(2018)가 PME의 진단능력을 검증하기 위해 사용한 기준들의 타당성을 검토할 필요가 있다. 첫째, 오키프는 AME와 PME의 순위를 결정하기 위해 단지 수치만을 비교하였으며, 수치 간의 차이가 통계적으로 유의미한가를 고려하지 않았다. 하지만 오키프가 분석에 사용한 자료들은 모두 표본(samples)이기 때문에 통계적으로 유의미한 차이가 아닐 경우, 수치가 다르더라도 실제로 차이가 있다고 단정할 수 없다. 오키프는 PME 간 차이가 통계적으로 유의미한 경우만을 분리해 따로 PME 순위와 AME 순위의 일치 여부를 검증하여 진단확률을 67%라고 보고하였다. 표본을 이용할 경우, 차이에 관한 통계적 유의미성은 반드시 고려되어야 하므로 오키프가 초록에서 보고한 진단확률(58%)보다 통계적 유의미성을 고려하여 측정된 PME의 진단확률(67%)이 더 정확한 수치라고 할 수 있다(Noar, Barker et al., 2018). 오키프는 두 메시지의 AME 간 차이의 유무를 판단하는 과정에서도 통계적 유의미성을 전혀 고려하지 않았다. AME의 차이 또한 통계적 유의미성을 고려하여 판단하는 것이 바람직하다. 둘째, 오키프가 메타

분석에서 사용한 기존 연구들은 PME와 AME를 여러 다른 방식으로 측정하였다. 특히 AME는 연구에 따라 태도, 감정, 행동의도 등 매우 다르게 측정하였다(예, Byrne, Katz, Mathios, & Niederdeppe, 2015). 그러므로 오키프의 메타 분석 연구 결과는 일반화하기 어렵다.

본 연구는 PME의 진단능력을 검증하기 위해 오키프(2018)가 사용한 방법을 수정 및 보완한다. 먼저 PME의 진단능력을 효과적으로 검증하기 위해 특정 사안에 대해 같은 주장을 하는 여러 개의 설득 메시지를 응답자들에게 제시하고 동일한 내용의 문항을 사용하여 PME와 AME를 각각 측정한다. 그리고 본 연구는 PME와 AME의 차이를 판정하기 위한 오키프의 기준을 다음과 같이 수정하여 사용하고자 한다. 첫째, 해당 두 메시지 간 PME가 통계적으로 유의미한 차이가 없고 AME도 통계적으로 유의미한 차이가 없는 경우( $[PME_{M1} \approx PME_{M2}]$  &  $[AME_{M1} \approx AME_{M2}]$ , “ $\approx$ ”는 통계적으로 유의미한 차이가 없는 경우를 표시)에는 PME와 AME의 내용이 일치하기 때문에 AME에 대한 PME의 진단능력을 인정할 수 있다(원칙 1). 오키프는 이 경우에 대해 따로 원칙을 제시하지 않았다. 둘째, 두 메시지 간 PME가 통계적으로 유의미한 차이가 없지만 AME가 통계적으로 유의미한 차이가 나는 경우 ( $[PME_{M1} \approx PME_{M2}]$  &  $[AME_{M1} \gg AME_{M2}]$ , “ $\gg$ ”는 통계적으로 유의미한 차이가 있는 경우를 표시)이다. 이 경우는 두 메시지의 효과가 실제에서는 차이가 있지만 PME가 그러한 차이를 포착하는 데 실패한 경우이다. 이는 두 집단 간 평균의 차이에 관한 가설을 통계적 방법으로 검증할 시 실제로는 차이가 있지만 없음으로 잘못 판정하는 2종 오류와 유사하다. 이 경우, PME가 메시지 간 실제하는 효과의 차이를 포착하는 데 실패하였기 때문에 AME에 대한 PME의 진단능력을 인정하지 않는다(원칙 2).

셋째, 두 메시지 간 PME가 통계적으로 유의미한 차이가 나지만 AME가 통계적으로 유의미한 차이가 없는 경우( $[PME_{M1} \gg PME_{M2}]$  &  $[AME_{M1} \approx AME_{M2}]$ )이다. 이 경우는 실제에서는 두 메시지 간 효과의 차이가 존재하지 않지만, PME의 결과는 차이 나는 것으로 나타난 경우이다. 이 경우는 가설검증 시, 두 집단 간 평균이 실제에서는 차이가 나지 않지만 관찰된 표본의 평균을 토대로 두 집단 간 차이가 있다고 잘못 판정하는 1종 오류와 상응하는 것으로 원칙적으로 오류이다. 이 경우 원칙적으로는 오류이지만 여러 이유로 진단능력을 부정하기에는 무리가 따른다. 실제 캠페인에서는 메시지가 일정 기간 반복적으로 메시지가 제시되지만, 실험실 연구 또는 설문 조사에서 메시지가 대개 한 번만 제시되므로 그런 연구에서 실제 캠페인에서 발생할 메시지 효과가 재현되지 않거나 축소되어 나타날 수 있다. 그리고 지각된 효과성에 비해 실제 효과는 측정이 상대적으로 어려워(Shen, Sun, & Pan, 2018) 측정오차가 더 크게 개입될 수 있다. 그런 이유로 형성평가에서 실제 효과가 부분적으로만 측정되어 메시지 간 차이가 나타나지



않을 수 있다. 오키프 또한 이 경우 PME가 최소한 더 낮은 AME를 갖은 메시지를 선택하도록 이끌지 않으므로 AME에 대한 PME의 진단능력을 인정하였다. 본 연구도 또한 이러한 이유로 PME의 진단능력을 부정하지 않는다(약한 인정, 원칙 3).

넷째, PME에서 유의미한 차이가 있고, AME에서 PME와 같은 방향으로 유의미하게 차이가 있을 경우(예,  $[PME_{M1} \gg PME_{M2}]$  &  $[AME_{M1} \gg AME_{M2}]$ ), PME의 진단능력이 있다고 판정한다. 이 경우 PME의 진단능력은 매우 높은 정도로 인정된다 (강한 인정, 원칙 4). 다섯째, PME에서 차이가 있고, AME에서 PME와 다른 방향으로 유의미한 차이가 나타났을 때 (예,  $[PME_{M1} \gg PME_{M2}]$  &  $[AME_{M1} \ll AME_{M2}]$ ), PME의 진단능력이 없다고 판정한다. 이 경우 PME의 진단능력은 매우 높은 정도로 부정된다 (강한 부정, 원칙 5). 본 연구는 새로운 기준을 사용하여 설득 메시지 간 PME의 상대적 순위와 AME의 상대적 순위의 일치 여부를 검증한다.<sup>1)</sup> 이를 검증하기 위해 실제로 사용된 복수의 메시지를 사용하여 PME와 AME를 측정 한 후, 메시지 간 PME 순위와 AME 순위의 일치 여부를 검증한다. 본 연구는 다음과 같은 연구 문제를 제시한다.

**RQ2.** 통계적 차이를 고려하였을 경우, 설득 메시지 간 PME와 AME의 상대적 순위의 일치 정도는 어떠한가?

### 3) 지각된 메시지 효과성의 진단능력 검증 방법 3: 지각된 효과성과 실제 효과의 직접 비교

딜러드와 동료들(Dillard, Weber et al., 2007)은 PME와 AME의 관계를 규명하기 위해 상관관계를 이용하였다. 특정 메시지에서 PME와 AME가 높은 상관관계를 보이면, PME를 높이 지각한 수용자일수록, 확률적으로 AME도 높을 것이라고 합리적으로 예측할 수 있다. 만일 두 메시지를 비교할 때, 전체 메시지에서의 정(+) 상관관계가 높다면 PME 평균이 높은 메시지가 AME 평균 또한 높을 것이라고 합리적으로 예측할 수 있다. 하지만 PME-AME의 상관관계는 PME가 AME를 과소평가하는지 혹은 과대평가하는지, 즉 어느 정도 반영하는지에 관한 정보는 제공하지 못한다. 오키프(2018)는 PME와 AME를 각각 0 - 100 척도를 이용하여 측정한 5개의 사례에 대

---

1) 오키프가 메타 분석에 사용한 논문 중 하나인 Byrne et al., (2015)의 연구를 오키프가 사용한 기준으로 분석하였을 경우, AME에 대한 PME의 진단능력 인정 비율은 75.0%(24/32)로 나타났다. 진단능력 인정 비율은 실제 효과의 측정 유형(신념, 공포, 행동의도)에 따라 다르게 나타났다. 본 연구에서 제시한 기준을 사용하였을 경우 AME에 대한 PME의 진단능력 인정 비율은 전체적으로 더 높게 측정되었다(84.4%, 27/32). 이 결과는 통계적 유의미성을 고려하였는지에 따라 PME 진단능력평가가 다르게 나타난다는 것을 보여준다.

한 가상 자료를 이용하여 상관관계의 문제점을 보여주었다. 같은 방식으로 다음의 다섯 사례의 자료를 가정하여 상관관계의 한계를 설명해 보고자 한다(PME, AME): (43, 73), (44, 74), (45, 75), (46, 76), (47, 77). 이 경우 PME-AME의 상관관계는 +1.00이다. 이 자료는 PME는 평균적으로 AME의 60%(= 45/75)를 반영하고 있다. 이는 PME가 AME를 과소평가함을 보여준다. 만일 다른 연구에서 다음의 다섯 사례의 자료를 발견했다고 하자(PME, AME): (13, 73), (14, 74), (15, 75), (16, 76), (17, 77). 이 경우도 PME-AME의 상관관계가 +1.00이지만 PME의 반영률은 20%에 불과하다. 만일 발견된 자료가 (73, 43), (74, 44), (75, 45), (76, 46), (77, 47) 라면 역시 상관관계는 +1.00이지만 PME가 AME를 오히려 1.67배 과대평가하고 있다. 캠페인 제작에 있어 실제 효과의 과대평가와 과소평가는 전혀 다른 실천적 의미를 지닌다. 상관관계를 넘어서 PME가 AME를 과소평가하는지 과대평가하는지 즉, 어느 정도 반영하는지를 검증하는 방법은 두 변수를 비교할 수 있게 측정하여 직접 비교하는 것이다.

설득캠페인 메시지효과 지각은 아니지만, 선거 여론조사 결과보도의 영향력 지각과 관련하여 지각된 영향력과 실제 영향력을 직접 비교한 연구가 있다. 정성은 외(Chung et al., 2018)는 선거 여론조사 결과보도의 영향력에 대한 지각에 편향이 있는지를 검증하기 위해 지각된 영향력을 실제 영향력과 직접 비교하였다. 그들은 선거 여론조사 결과보도에 노출되기 전 후보에 대한 태도와 결과보도 노출 후의 후보에 대한 태도의 차이로 여론조사 결과보도의 실제 영향력을 측정하였다. 그리고 연구자들은 응답자들에게 여론조사 결과보도 노출로 인한 응답자 자신의 후보에 대한 태도변화의 정도를 보고하게 하였다. 그들은 실제 태도 변화 측정과 지각된 태도 변화 정도를 동일한 척도(0-100)를 사용하여 측정하였다. 그 결과 연구자들은 지각된 보도 영향력과 실제 보도 영향력을 직접 비교할 수 있었다. 실험 결과, 선거 여론조사 결과 노출 후 긍정적으로 변한 집단과 부정적으로 변한 집단이 나타났으며, 두 집단 모두 실제 태도 변화 정도(AME)보다 과소하게 태도 변화 정도를 보고(PME)하는 것을 발견했다. 인상관리 동기(impression management motivation)와 설득에 관한 연구(McGuire & Millman, 1965; Schlenker, 1980)들은 사람들은 외부메시지에 쉽게 영향받거나 설득당하지 않는 사람으로 보이려는 인상관리 동기가 있으므로 설혹 메시지로 인한 태도의 변화가 있을지라도 인정하지 않음을 주장했다. 미디어 영향력 지각에 관한 연구들 또한 인상관리동기로 인해 자신이 받은 메시지 영향력을 실제보다 과소평가함을 보여주고 있다(Douglas & Sutton, 2004; Tal-Or & Drukman, 2010). 기존 연구들은 그러한 메시지의 자신에 대한 영향력을 실제보다 과소평가하는 경향이나 다른 사람이 받을 영향보다 과소하게 평가하는 경향이 주로 사회적으로 바람직하지 않은 메시지일 경우(Douglas & Sutton, 2004; Perloff, 2009; Schoenbach & Becker, 1995) 또는 자신의 기

존 태도와 일치하지 않는 경우(정성은 이원지, 2012; Meirick, 2004)에 나타난다고 보고하고 있다. 이러한 동기의 작용은 PME의 측정 유형에 상관없이 모든 PME에 영향을 끼칠 수 있다.

이러한 검증 방법을 설득캠페인 PME의 진단능력 검증에 적용할 수 있다. 설득 메시지 노출 전후의 태도 간 차이로써 AME를 측정하고 메시지 노출로 인한 자신의 태도 변화정도에 대한 지각으로 PME를 측정하면 두 변수를 직접 비교할 수 있고 상관관계뿐만 아니라 AME에 대한 PME의 반영률도 측정할 수 있다. 본 연구는 특히 특정 사안에 대해 설득캠페인의 주장과 일치하지 않는 태도를 보이던 수용자들이 캠페인 메시지 노출로 인한 자신의 태도 변화의 정도를 실제 태도 변화를 어느 정도 반영하는지를 살펴보고자 한다.

**RQ3.** 설득캠페인 메시지의 입장과 자신의 기존 태도가 불일치할 경우, 수용자들이 지각하는 메시지 노출로 인한 자신의 태도 변화 정도(즉, 지각된 메시지 효과성)는 메시지 노출로 인한 실제 태도 변화 정도(즉, 실제 메시지 효과)와 비교하여 어느 정도인가?

#### 4) 지각된 메시지 효과성과 실제 메시지 효과의 측정 유형

노어 외(Noar, Bell et al., 2018)는 PME의 진단능력이 논란이 되는 이유로 기존 연구들이 서로 다른 측정 방식을 사용하였다는 점을 지적하였다. 설득 메시지 AME에 대한 PME의 진단능력을 체계적으로 검증하기 위해서는 기존 연구에서 사용된 PME와 AME 측정 방식을 유형별로 구분하고 하나의 연구에서 동시에 사용하여 검증할 필요가 있다. 기존 연구들이 사용한 주요 PME 측정 방식은 첫째, 메시지가 얼마나 논리적인지(logical), 근거를 제시하고 있는지(evidence), 생생한지(vivid), 일관성 있는지(coherent) 등을 측정한 지각된 논증 품질(perceived argument quality; Dillard, Weber et al., 2007)이다. 기존 연구들이 사용한 PME 유형 중 하나는 응답자 자신에게 발생할 메시지 영향력 추정 즉, 메시지 영향력 추정(estimated message impact; Dillard, Weber et al., 2007, Hullett, 2004)이다. 기존 연구에서 사용된 지각된 메시지 영향력은 메시지 대상자를 응답자 또는 다른 사람들로 특정한 경우와 특정하지 않은 경우로 나뉜다(Bigsby et al., 2013; Noar, Barker et al., 2018). 본 연구는 메시지 대상이 응답자인 경우와 다른 목표 수용자들인 경우로 구분한다. 기존 연구들에서 사용된 PME 측정 유형 중 다른 하나는 지각된 태도 변화 정도(perceived attitude change)이다. 정성은 외(Chung et al., 2018)는 실험 참여자들에게 가상의 선거 후보자들에 대한 선거 여론조사 결과를 제시한 후 그들은 여론조사결과에 대한 노출로 인한 자신의 태도 변화 정도를 보고하게 하였다. 이후 연구자들은 지각된 태도 변화 정도와 실제 태도 변화 정도(즉, 선거 여론

조사 결과 노출 전과 후에 측정된 선거 후보자에 대한 태도의 차이 정도)를 직접 비교하였다. PME를 이 방식으로 측정한다면 PME가 AME를 얼마나 정확하게 반영하는지를 확인할 수 있다. 지각된 태도 변화 정도는 PME의 다른 측정 유형인 메시지 영향력 지각과 내용적으로 일치한다. 메시지 영향력 지각은 메시지의 신념, 태도, 행동 등에 대한 영향력을 영향의 크기(적다, 크다)로써 측정하는 데 비해, 지각된 태도 변화 정도는 메시지의 영향력을 태도 측정에서 사용된 척도를 사용하여 그 정도를 표시하게 하는데 차이가 있다. 그러므로 지각된 태도 변화 정도 또한 PME의 한 측정유형이라 할 수 있다. 본 연구는 PME의 진단능력을 체계적으로 검증하기 위해 PME를 측정유형에 따라 구분하여 각 유형의 진단능력을 검증하고 측정 유형 간에 진단능력의 차이가 있는지를 또한 검증한다.

기존 연구들은 실제 메시지 효과 또한 다양한 방식으로 측정하였다. 메시지의 실제 효과를 측정하기 위해 기존 연구에서 많이 사용된 방법의 하나는 사후-메시지 태도이다(post-message attitude; Dillard & Peck, 2000; Dillard & Witte, 1993; Hullett, 2004). 하지만 이러한 방식으로 AME를 측정할 경우, 사안에 대한 기존 태도 등의 가외변수가 PME와 AME 모두에게 영향을 끼칠 수 있다. 이런 경우, PME와 AME 간에 발견된 상관관계는 가외변수의 통제 여부에 따라 다르게 나타날 수 있으므로 PME의 AME에 대한 인과성 또는 PME의 진단능력을 표시하기에 한계를 가진다. 이와 달리 몇 연구들(예, Collins et al., 1988; Douglas & Sutton, 2004; Hullett, 2002)은 AME를 설득 메시지 노출 전과 메시지 노출 후에 태도를 각각 측정된 후 두 태도 간의 차이(즉, 메시지 노출 전후의 태도 변화의 정도)로 AME를 측정하여 PME와 AME의 관계를 검증하였다. 메시지 노출 전후의 차이를 구하여 이를 종속변수로 사용하면 응답자의 초기태도 등 응답자 속성 관련 고정 효과(unit-specific fixed effects)를 통제할 수 있다(Meyer, 1995). 그러므로 메시지 노출 전후의 태도 변화의 정도로 AME를 측정할 경우, PME-AME 관계를 더 정확하게 검증할 수 있다. 또한 메시지 노출 전후의 태도 변화의 정도로 AME를 측정하면 이에 상응하는 PME(즉, 메시지 노출 전후의 태도 변화 정도에 대한 지각)를 이용하여 PME와 직접 비교할 수 있다. 직접 비교를 통해 둘 간의 차이(discrepancy)가 어느 정도인지 검증할 수 있다.

많은 공공캠페인은 궁극적으로 목표 수용자의 행동변화를 목표로 하는 경우가 많으므로 AME를 실제 행동 또는 행동의도의 변화로 측정하는 것이 이상적이다. 그러나 메시지 노출로 인한 실제 행동의 변화를 측정하는 것은 많은 참여자가 필요하고 행동의 변화까지 시간이 소요되기 때문에 현실적으로 어려움이 크다. 한편 수용자의 행동변화를 최종목표로 하지만 해당 단계에서 태도나 신념의 변화를 목표로 하는 설득캠페인도 많다(Miller, 1980; Noar, Bell et al.,

2018; Pfau & Parrott, 1993). 예를 들어, 원자력 발전소의 안전성을 관련 통계를 중심으로 주장하는 캠페인은 원자력 발전소의 안전성에 대한 동의, 즉 태도의 변화를 목표로 한다. 이러한 캠페인의 경우는 태도 변화의 정도로 캠페인의 실제 효과를 측정하는 것이 타당하다. 이러한 이유로 본 연구는 태도 변화를 목표로 하는 캠페인들을 대상으로 하여 메시지 노출 전후 태도 변화의 정도로 AME를 측정하여 사용하고자 한다. 본 연구는 PME를 각각 지각된 메시지 품질, 메시지 영향력 추정(자신 또는 타인), 그리고 지각된 태도 변화 정도로 구분하여 측정하고 AME를 메시지 노출 전후의 태도 변화의 정도로 측정하여 PME와 AME의 관계를 위에서 설명한 세 가지 방법을 적용하여 검증하고자 한다.

**RQ4.** 지각된 메시지 효과성과 실제 메시지 효과 간의 관계는 지각된 메시지 효과성의 측정 유형(지각된 논증 품질, 메시지 영향력 추정, 지각된 태도 변화 정도)에 따라 다르게 나타나는가?

### 3. 연구 방법

#### 1) 표본 구성과 연구 절차

제안된 연구 가설은 한국 원자력 발전소의 안전성을 설득하는 캠페인 메시지들을 이용하여 검증하였다. 한국 원자력 발전소 안전 문제는 오랫동안 사회적 논쟁 사안이었고(Kim, 2017), 한국 수력원자력주식회사 그리고 원자력안전위원회 등이 여러 차례 설득캠페인을 진행했기 때문에 캠페인 메시지의 실제 효과와 지각된 효과 간의 관계를 검증하기에 적합한 사안으로 보였다. 가설의 검증을 위해 여론조사 업체를 통해 두 차례에 걸친 패널조사를 실시하였다(two-wave panel study). 일차 설문 조사에서 원자력 발전소의 안전에 대한 기존 태도를 측정하고, 일주일 후 이차 조사에서 응답자들을 무작위로 네 개의 집단으로 구분하였다. 각 집단에 속한 응답자들은 하나의 캠페인 메시지를 본 후 질문에 응답하였다. 다수의 공공 설득캠페인은 해당 사안에 대해 반대하거나 중립적인 태도를 지닌 사람들의 태도나 행동의 변화를 목표로 한다(Atkins & Rice, 2013; Pfau & Parrott, 1993). 본 연구는 일차조사에서 원자력 발전소 안전에 대해 중립적이거나 반대 태도를 보인 시민들( $N = 1,202$ )의 응답을 측정하였다. 일주일 후 이차 조사 기간 동안 915명의 참가자가 메일에 응답하였으며 최종적으로 580명이 응답을 완료하였다. 주목 검증 문항에 실패한 응답자와 논리적으로 연결된 두 질문에 대해 모순적인 답변을 한 응답자들을 제외한 475명의 응답을 분석하였다. 조사 참가자의 성별은 남성( $n = 240, 50.5%$ ), 여성( $n =$

235, 49.5%)으로 나타났으며 연령대는 20세부터 69세까지( $M = 44.21$ ,  $SD = 13.08$ )로 나타났다. 온라인 설문 일차 조사에서 메시지 주제에 대한 조사 참가자들의 기존 태도를 측정하였다(중립:  $n = 243$ , 반대:  $n = 232$ ). 일주일 후, 이차 조사에서 한국 원자력 발전소는 안전하다는 설득 메시지를 제시하고 응답자들의 태도를 재측정하고 PME를 다양한 방식으로 측정하였다.

## 2) 설득캠페인 메시지

조사에 사용된 설득캠페인 메시지들은 ‘한국 원자력 발전소의 안전성’을 알리기 위해 특정 기관(즉, 대한민국 정부, 원자력안전위원회, 그리고 한국수력원자력)이 유튜브(YouTube)에서 실제로 사용한 메시지들로 선정되었다. 메시지들은 메시지의 표현 형식(실제 vs. 그래픽), 등장인물(사람 등장 vs. 사람이 등장하지 않음), 내용(전문적 vs. 일반적)에서 서로 다른 메시지들을 선정하였다. 선정된 메시지들의 길이는 크게 다르지 않도록 편집하였다. 사용된 설득캠페인 메시지의 주요 특성은 <Table 1>과 같다.

## 3) 주요 변수 측정

### (1) 지각된 논증 품질

지각된 논증의 품질을 측정하기 위해 기존 연구들(예, Dillard, Weber et al., 2007; Hullett, 2002; O’Keefe, 2018)을 참고하여 다음과 같은 문항들을 사용하였다: 1. 제시된 메시지는 주목을 끈다; 2. 완성도가 높다; 3. 논리적이다; 4. 정확한 사실에 기초하고 있다; 5. 주장을 뒷받침하는 근거들 충분히 제시하고 있다(11-point scale, 0 = 전혀 그렇지 않다, 10 = 매우 그렇다); 일곱 개의 문항으로부터 측정된 값들의 평균을 사용하였다(Cronbach’s  $\alpha = .93$ ,  $M = 6.00$ ,  $SD = 1.64$ ).

Table 1. Messages Used in the Study

Message	Title	Time	Source	Speaker	Narrator	Information
1	Our Lives (우리 삶)	2:18	KHNP	CEO of KHNP	CEO of KHNP	Explanation about the issue, specific evidence(e.g., temperature, feedback effect, big data statistic).
2	Soowoni & Jaro (수원아와 자로)	2:00	KHNP	Character (animation)	Voice actor	Explanation about the issue and evidence (e.g., comparison with Japan cases).
3	Is It Safe? (안전한가요?)	2:47	KHNP	Employees of KHNP	Voice actor	Specific and scientific information and evidence.
4	Jang Do-Yeon's Interview (장도연 인터뷰)	3:54	Korean government, NSSC	Comedian and chairperson of NSSC	Comedian and chairperson of NSSC	Explanation about the issue, no specific evidence.

Note. KHNP: Korea Hydro & Nuclear Power Co.,Ltd; NSSC: Nuclear Safety and Security Commission.

Message 1: <https://youtu.be/H9DgVKgtj7g>; Message 2: <https://youtu.be/hyIBGqk-PDU>;

Message 3: <https://youtu.be/Luglw99alFI>; Message 4: <https://youtu.be/IFg0QKHIn0>.

## (2) 메시지 영향력 추정

자신에 대한 메시지 영향력 추정을 측정하기 위해 기존 연구들(예, Dillard, Peck, 2000; Dillard, Weber et al., 2007; Hullett, 2004)을 근거로 다음과 같은 문항들을 사용하였다:

1. 나는 앞에서 제시된 캠페인 메시지를 통해 새로운 것을 배웠다; 2. 나는 앞에서 제시된 캠페인 메시지를 읽은 후 '한국 원자력 발전소의 안전성'에 대한 확신이 증가했다; 3. 앞에서 제시된 캠페인 메시지는 나에게 도움이 되었다; 4. 앞에서 제시된 캠페인 메시지는 '한국 원자력 발전소의 안전성'에 대한 나의 태도에 긍정적으로 영향을 미쳤다(11-point scale; 0 = 전혀 그렇지 않다, 10 = 매우 그렇다), Cronbach's  $\alpha = .95$ ,  $M = 5.96$ ,  $SD = 1.88$ ). 다른 일반 사람들(성인)에 대한 메시지 영향력 추정(Cronbach's  $\alpha = .97$ ,  $M = 6.16$ ,  $SD = 1.69$ ) 또한 같은 질문을 사용하여 측정하였다.

## (3) 이슈에 대한 태도와 자신의 실제 태도 변화 정도

자신의 실제 태도 변화 정도를 측정하기 위해 정성은 외(Chung et al., 2018)의 연구를 참고하여 다음과 같은 문항들을 사용하였다. 귀하께서는 '한국 원자력 발전소는 안전하다'라는 주장에 대해 지금 어느 정도 동의하십니까? 귀하의 현재 동의 정도를 아래에 표시해주시기 바랍니다 (11-point scale; 0% - 100%). 응답자들의 사안에 대한 태도를 반대를 의미하는 왼쪽 부분(0% - 50%)과 찬성을 의미하는 오른쪽 부분(50% - 100%)으로 구분하였으며, 10% 단위로 선택할 수 있도록 제시하였다. 또한, 조사 참가자들이 질문의 내용에 혼란이 없도록 다음과 같은

설명을 추가로 제시하였다: 0%: 한국 원자력 발전소는 안전하다는 주장에 전혀 동의하지 않음; 100% 한국 원자력 발전소는 안전하다는 주장에 전적으로 동의함. 이슈에 대한 실험 참여자 자신의 태도를 일차 조사( $M = 38.55$ ,  $SD = 14.85$ )와 이차 조사( $M = 58.15$ ,  $SD = 18.41$ )에서 각각 측정하였다. 자신의 실제 태도 변화 정도를 측정하기 위해 메시지 노출 전과 후에 측정된 사안에 대한 태도의 차이를 계산하여 사용하였다( $M = 19.60$ ,  $SD = 15.51$ ).

#### (4) 회상된 사안에 대한 자신의 기존 태도

응답자들이 메시지로 인한 자신의 태도 변화 정도를 성실하게 표시하였는지를 확인하기 위해 기존 연구(Douglas & Sutton, 2004)를 참고하여 사안에 대한 자신의 기존 태도를 측정하였다. 이차조사에서 조사 참여자들은 회상한 메시지 노출 전 자신의 태도를 측정하도록 요청받았다: 일주일 전 조사에서 “한국 원자력 발전소는 안전하다”라는 주장에 대해 귀하께서 표기하신 동의 정도를 최대한 기억하여 아래에 표기해 주시기 바랍니다. 조사 참여자들에게 자신이 지각하는 일차 조사에서 측정된 자신의 태도를 빈칸에 직접 기재하도록 요청하였다: 한국 원자력 발전소는 안전하다는 주장에 대한 일주일 전의 나의 동의 정도: [빈 칸] % ( $M = 48.64$ ,  $SD = 18.05$ ).

#### (5) 지각된 자신의 태도 변화 정도와 지각된 다른 사람들의 태도 변화 정도

지각된 자신의 태도 변화 정도를 측정하기 위해 기존 연구(Chung et al., 2018)를 근거로 다음과 같은 문항들을 사용하였다: “일차 조사와 비교해, 이슈에 대한 당신의 동의 정도는 얼마나 달라졌습니까?” 조사 참여자들은 지각된 이슈에 대한 동의 정도의 증감 여부를 제시된 콤보박스를 사용하여 측정하였다(11-point scale: 0% - 100%). 콤보박스의 선택지는 0% 증가(변화 없음)부터 100% 증가까지 총 열한 개의 선택지로 구성되었다( $M = 11.09$ ,  $SD = 14.40$ ). 본 연구는 위와 같은 방법으로 지각된 다른 사람들의 태도 변화 정도( $M = 12.11$ ,  $SD = 12.82$ )를 측정하였다.

#### (6) 사안에 관한 관심 정도

사안에 관한 조사 참가자들의 관심 정도를 측정하기 위해 기존 연구(Chung et al., 2018)를 근거로 다음과 같은 질문들을 사용하였다: 1. 나는 ‘한국 원자력 발전소의 안전성’에 대한 미디어 보도를 주의 깊게 본다; 2. 나는 ‘한국 원자력 발전소의 안전성’에 대해 주위 사람들과 이야기를 나누는 편이다(0 = 전혀 그렇지 않다, 10 = 매우 그렇다); 3. 귀하께서는 ‘한국 원자력 발전소의 안전성’과 관련한 사안에 관해 얼마나 많이 알고 있다고 생각하십니까? (0 = 전혀 모른다, 10 = 매우 많이 알고 있다); 4. 귀하께서는 ‘한국 원자력 발전소의 안전성’과 관련한 사안에 관



해 얼마나 잘 이해하고 계십니까? (0 = 이해도 매우 낮음, 10 = 이해도 매우 높음) 측정된 값들의 평균을 사용하였다(Cronbach's  $\alpha = .87$ ,  $M = 5.75$ ,  $SD = 1.62$ ).

## 4. 연구 결과

1) 전체 메시지의 PME와 AME 간 상관관계와 각 메시지에 한정된 두 변수 간 상관관계 RQ1 (PME-AME 상관관계)와 RQ4 (PME 측정유형별 상관관계 정도)를 동시에 검증하였다. 이를 위해 PME와 AME 간의 상관관계를 네 개 메시지 전체(across-message correlation)에 대해 그리고 메시지별로(message-specific correlation), 그리고 PME의 측정 유형별로 검증하였다(Table 2)). PME의 측정 유형은 지각된 논증 품질(perceived argument quality), 자신(또는 타인)에 대한 메시지 영향력 추정(estimated message impact on oneself/others), 지각된 자신 (또는 타인) 태도 변화 정도(perceived attitude change of oneself/others)로 구분하여 측정하였다.

Table 2. Across-Message and Within-Message PME-AME Correlations for Various PME Measures

	Perceived argument quality	Estimated impact on oneself	Estimated impact on others	Perceived attitude change of oneself	Perceived attitude change of others	Average correlation
All ( $n = 475$ )	.417 <sub>ab</sub> <sup>***</sup>	.457 <sub>a</sub> <sup>***</sup>	.368 <sub>ab</sub> <sup>***</sup>	.431 <sub>ab</sub> <sup>***</sup>	.332 <sub>b</sub> <sup>***</sup>	.401
M1 ( $n = 117$ )	.503 <sub>a</sub> <sup>***</sup>	.522 <sub>a</sub> <sup>***</sup>	.437 <sub>a</sub> <sup>***</sup>	.498 <sub>a</sub> <sup>***</sup>	.402 <sub>a</sub> <sup>***</sup>	.472
M2 ( $n = 128$ )	.344 <sub>ab</sub> <sup>***</sup>	.430 <sub>a</sub> <sup>***</sup>	.304 <sub>ab</sub> <sup>***</sup>	.443 <sub>a</sub> <sup>***</sup>	.190 <sub>b</sub> <sup>*</sup>	.342
M3 ( $n = 112$ )	.420 <sub>a</sub> <sup>***</sup>	.482 <sub>a</sub> <sup>***</sup>	.416 <sub>a</sub> <sup>***</sup>	.539 <sub>a</sub> <sup>***</sup>	.489 <sub>a</sub> <sup>***</sup>	.469
M4 ( $n = 118$ )	.443 <sub>a</sub> <sup>***</sup>	.428 <sub>a</sub> <sup>***</sup>	.342 <sub>a</sub> <sup>***</sup>	.232 <sub>a</sub> <sup>*</sup>	.259 <sub>a</sub> <sup>***</sup>	.341

Note. PME = perceived message effectiveness. AME = actual message effectiveness (amount of actual attitude change). M1 is "Our Lives." M2 is "Soowoni and Jaro." M3 is "Is It Safe?" M4 is "Jang Do-Yeon's Interview." Correlations with common subscripts in the same row are not statistically significantly different from each other. Correlations without common subscripts in the same row are statistically significantly different.

\* $p < .05$ . \*\*\* $p < .001$ .

RQ1에 대한 분석결과, 메시지 전체의 PME-AME 간의 상관관계는 PME 측정 유형과 관계 없이 모두 통계적으로 유의미하게 정(+ )적으로 나타났고(.332  $\leq r \leq$  .457), 개별 메시지에 관한 PME-AME 간의 상관관계 또한 네 개의 메시지 모두에서 통계적으로 유의미하게 정(+ )적으로 나타났다 (M1: .402  $\leq r \leq$  .522; M2: .190  $\leq r \leq$  .443; M3: .416  $\leq r \leq$  .539; M4: .232  $\leq r \leq$  .443). 메시지 전체의 경우와 개별 메시지 경우 모두 PME-AME 간의 상관관계가

정(+ )적으로 나타났기 때문에 이 경우 PME와 AME 간의 상관관계가 PME의 AME에 대한 진단능력을 일정 정도 반영하고 있다고 볼 수 있다. PME 측정 유형 별 상관관계의 평균은 전체 메시지 경우 .401, 개별 메시지의 경우 .341 - .472로 나타났다. 이는 딜러드 외가 메타 분석에서 발견한 .41과 유사한 수치이다. 코헨(Cohen, 1988; 1992)에 의하면 이 수치는 보통수준 상관관계(moderate, .30)와 높은 수준 상관관계(large, .50)의 중간 수준으로 비교적 높은 상관관계라고 할 수 있다. 이러한 결과는 AME에 대한 PME의 진단능력이 비교적 높은 수준임을 제시한다.

RQ4를 PME-AME 상관관계에 적용하여 살펴보았다. 개별 메시지의 경우, 두 번째 메시지를 제외하고 나머지 세 개의 메시지에서 PME-AME 상관관계는 다섯 개의 PME 측정 유형 간에 통계적으로 유의미한 차이를 보이지 않았다. 두 번째 메시지도 지각된 다른 사람들의 태도 변화 정도를 제외하고 서로 차이를 보이지 않았다. 이러한 결과는 기존에 사용된 PME 측정 유형들이 유사한 수준에서 AME와 비교적 높은 수준(약 .40)의 상관관계를 맺는 것을 보여준다. 즉, 여러 PME 측정 유형들이 모두 AME에 대한 비슷한 정도의 진단능력을 가지는 것을 보여준다. 특이한 점은 응답자가 지각한 자신의 태도 변화 정도는 기존 연구에서 많이 사용되지 않았으나 두 개의 메시지에서 수치상으로 가장 높은 상관관계를 보인 점이다.

## 2) 지각된 메시지 효과성과 실제 메시지 효과의 순위 간의 일치 정도 검증

RQ2의 검증에 앞서 네 개의 메시지 간 각 PME 측정유형이 통계적으로 유의미하게 차이가 있는지 그리고 메시지가 AME에서 차이가 있는지를 검증해 보았다. 일원변량분석 (ANOVA)를 통해 메시지 유형 (4개 메시지)에 따라 PME 측정 변수가 다른지를 각 PME 측정 유형별로 검증하였다. 성별, 관심정도를 통제변수로 사용하였다. 지각된 논증 품질의 경우, 메시지 간 차이가 유의미하게 나타났다,  $F(3, 465) = 3.50, p = 0.15$ , 부분  $\eta^2 = .02$ . 자신에 대한 추정된 메시지 영향력의 경우, 메시지 간 차이가 유의미하게 나타나지 않았다,  $F(3, 465) = 2.43, p = 0.06$ , 부분  $\eta^2 = .015$ . 다른 사람들에 대한 추정된 메시지 영향력의 경우, 메시지 간 차이가 유의미하게 나타나지 않았다,  $F(3, 465) = 2.15, p = 0.09$ , 부분  $\eta^2 = .01$ . 지각된 자신의 태도변화 정도의 경우, 메시지 간 차이가 유의미하게 나타나지 않았다,  $F(3, 465) = 2.21, p = 0.09$ , 부분  $\eta^2 = .014$ . 지각된 다른 사람들의 태도변화 정도의 경우, 메시지 간 차이가 유의미하게 나타났다,  $F(3, 465) = 2.95, p = 0.03$ , 부분  $\eta^2 = .019$ . 각 메시지 짝들 간의 평균 비교는 <Table 3>에 자세히 보고하였다.

RQ2(순위 비교)와 RQ4 (PME 측정 유형 간 차이)를 검증하기 위해 메시지 짝을 추출해 PME와 AME 순위를 비교하였다. 총 네 개의 메시지를 이용하여 여섯 개의 메시지 짝을 만들고

PME 순위와 AME 순위의 일치 여부 정도(진단능력 인정률)를 구하였다. 진단능력 인정률은 PME 측정 유형 별로 검증하였고 차이에 관한 통계적 유의미성을 고려한 새로운 기준을 적용하였다. 분석결과는 <Table 3>에 제시하였다.

분석결과, PME 측정 유형 모두에서 100%(30/30) 진단능력 인정률이 나타났다. 구체적으로 살펴보면 AME는 모든 메시지 짝에서 통계적으로 유의미한 차이가 나타나지 않았다. AME와 PME에서 모두 통계적으로 유의미한 차이가 나타나지 않아서 PME의 진단능력이 인정된 경우(원칙 1)는 지각된 논증 품질의 경우(50.0%, 3/6), 자신에 대한 영향력 추정(estimated impact on oneself)의 경우(66.7%, 4/6), 다른 사람들에 대한 영향력 추정(estimated impact on other) 경우(66.7%, 4/6), 지각된 자신의 태도 변화 정도(perceived attitude change of oneself)의 경우(83.3%, 5/6), 지각된 다른 사람들의 태도 변화 정도(perceived attitude change of other)의 경우(66.7%, 4/6)로 나타났다. PME에서 유의미한 차이가 있음에도 AME에서 유의미한 차이가 없었기 때문에 진단능력이 인정된 경우(약한 인정, 원칙 3)는 지각된 논증 품질의 경우(50.0%, 3/6), 자신에게 미칠 영향력에 대한 지각의 경우(33.3%, 2/6), 다른 사람들에게 미칠 영향력에 대한 지각의 경우(33.3%, 2/6), 지각된 자신의 태도 변화 정도의 경우(16.7%, 1/6), 지각된 다른 사람들의 태도 변화 정도의 경우(33.3%, 2/6)로 나타났다.

RQ4를 두 메시지 간 PME와 AME 순위 비교에 적용하여 살펴보았다. PME 진단능력 인정률은 PME 측정 유형 별로 차이가 나타나지 않았다(모두 100.0%). 세부적으로 AME와 PME에서 모두 통계적으로 유의미한 차이가 나타나지 않아서 PME의 진단능력이 인정된 경우(원칙 1)는 지각된 자신의 태도 변화 정도(83.3%, 5/6)가 다른 PME 측정 유형보다 높게 나타났으며, PME에서 유의미한 차이가 있음에도 AME에서 유의미한 차이가 없어서 진단능력이 인정된 경우(원칙 3)는 지각된 논증 품질의 경우(50.5%, 3/6) 다른 PME 측정 유형보다 높게 나타났다.

Table 3. Comparison in PME and AME between Two Messages for Various Measures of PME

Message Pair		Actual attitude change of oneself			Perceived argument quality				Estimated impact on oneself				Estimated impact on others				Perceived attitude change of oneself				Perceived attitude change of others			
A	B	A	B	t	A	B	t	Diag	A	B	t	Diag	A	B	t	Diag	A	B	t	Diag	A	B	t	Diag
M1	M2	19.15 (14.12)	20.47 (15.26)	.70	6.03 (1.68)	6.15 (1.64)	.55	Yes	6.07 (1.81)	6.01 (2.06)	.25	Yes	6.12 (1.75)	6.30 (1.74)	.79	Yes	10.85 (13.10)	13.52 (16.53)	1.39	Yes	13.16 (13.69)	14.06 (12.95)	.53	Yes
M1	M3	19.15	19.11 (17.17)	.02	6.03	6.24 (1.53)	.99	Yes	6.07	6.16 (1.87)	.35	Yes	6.12	6.35 (1.56)	1.04	Yes	10.85	10.89 (14.80)	.02	Yes	13.16	11.52 (12.82)	.94	Yes
M1	M4	19.15	19.58 (15.60)	.22	6.03	5.58 (1.65)	2.09*	Yes	6.07	5.61 (1.73)	1.98*	Yes	6.12	5.86 (1.66)	1.17	Yes	10.85	8.90 (12.39)	1.18	Yes	13.16	9.49 (11.39)	2.24*	Yes
M2	M3	20.47	19.11	.65	6.15	6.24	.45	Yes	6.01	6.16	.58	Yes	6.30	6.35	.24	Yes	13.52	10.89	1.29	Yes	14.06	11.52	1.53	Yes
M2	M4	20.47	19.58	.45	6.15	5.58	2.73**	Yes	6.01	5.61	1.62	Yes	6.30	5.86	2.01*	Yes	13.52	8.90	2.46*	Yes	14.06	9.49	2.93**	Yes
M3	M4	19.11	19.58	.22	6.24	5.58	3.17**	Yes	6.16	5.61	2.29*	Yes	6.35	5.86	2.30*	Yes	10.89	8.90	1.11	Yes	11.52	9.49	1.27	Yes
Diagnosticity rate		100.0% (6/6)				100.0% (6/6)				100.0% (6/6)				100.0% (6/6)				100.0% (6/6)						

Note. PME is perceived message effectiveness. AME is actual message effect and measured as the amount of actual attitude change. Diag. is PME's diagnosticity for AME. Diagnosticity is judged as good if the ranking of two messages in PME is matched with the ranking in AME. M1: Our Lives; M2: Soowoni & Jaro; M3: Is It Safe?; M4: Jang Do-Yeon's Interview. \* $p < .05$ . \*\* $p < .01$ .

### 3) 설득 메시지에 대한 지각된 메시지 효과성의 정확성 검증

RQ3(설득캠페인 메시지의 입장과 자신의 기존 태도가 불일치할 경우, 수용자들이 지각하는 메시지 효과성과 실제 메시지 효과 비교)을 검증하기 위해 공분산 분석(ANCOVA)을 하였다. 태도 변화 정도의 유형(즉, 실제 효과, 지각된 효과성)을 반복측정 변수로 사용하였고 태도 차이 유형(즉, 실제 효과 vs. 지각된 효과성)을 종속변수로 사용하였다. 사안에 대한 기존 태도, 사안에 관한 관심 정도, 그리고 성별을 통제 변수로 사용하였다. 분석결과, 지각된 자신의 태도 변화 정도( $M = 11.09$ ,  $SD = 14.40$ )와 실제 자신의 태도 변화 정도( $M = 19.60$ ,  $SD = 15.51$ ) 간의 차이가 통계적으로 유의미하게 나타났다,  $F(1,462) = 33.27$ ,  $\eta_p^2 = .07$ ,  $p < .001$ . 지각된 태도 변화 정도는 실제 태도 변화 정도를 56.6%(19.60 vs. 11.09) 반영하고 있는 것으로 나타났다(즉, 과소평가). 태도 변화 정도의 차이(지각 vs. 실제)와 사안에 대한 기존 태도(중립 vs. 반대) 간에 통계적으로 유의미한 상호작용이 측정되었다,  $F(1,462) = 43.27$ ,  $\eta_p^2 = .09$ ,  $p < .001$  (〈Figure 1〉 참조). 이를 더 자세히 분석하기 위해 사안에 대한 기존 태도를 두 집단(즉, 중립 집단, 반대 집단)으로 나누어서 반복측정 분산분석(repeated ANOVA)을 시행하였다. 중립 집단의 경우( $n = 243$ ), 실제 자신의 태도 변화 정도( $M = 15.47$ ,  $SD = 12.70$ )와 지각된 자신의 태도 변화 정도( $M = 11.81$ ,  $SD = 14.43$ ) 간에 차이가 통계적으로 유의미하게 나타났다,  $F(1,231) = 4.78$ ,  $\eta_p^2 = .02$ ,  $p = .030$ . 이 경우 지각된 태도 변화 정도는 실제 태도 변화 정도를 76.3%(15.47 vs. 11.81) 반영하는 것으로 나타났다. 반대 집단의 경우( $n = 232$ ), 실제 자신의 태도 변화 정도( $M = 23.92$ ,  $SD = 16.97$ )와 지각된 자신의 태도 변화 정도( $M = 10.34$ ,  $SD = 14.35$ ) 간에 차이가 통계적으로 유의미하게 나타났다,  $F(1,220) = 31.16$ ,  $\eta_p^2 = .12$ ,  $p < .001$ . 이 경우 지각된 태도 변화 정도는 실제 태도 변화 정도를 43.2%(23.92 vs. 10.34)만 반영하는 것으로 나타났다. 즉, 초기 반대 태도일 경우 초기 중립 태도보다 과소평가의 경향이 더 강하게 나타났다.

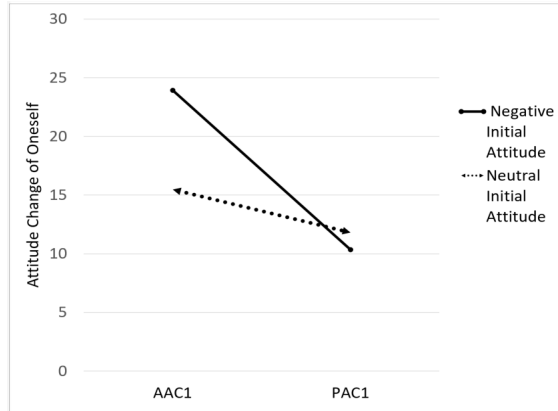


Figure 1. Perceived attitude change of oneself (PAC1) and actual attitude change of oneself (AAC1) by initial attitudes

메시지 별로 실제 태도 변화 정도와 지각된 태도 변화 정도를 비교한 결과도 메시지 전체 결과와 대체로 동일하게 나타났다(Table 4 참조).

Table 4. Perceived Attitude Change of Oneself (PAC1) and Actual Attitude Change of Oneself (AAC1) by Initial Attitudes

	Neutral Initial Attitudes				Negative Initial Attitudes			
	AAC1	PAC1	t	Proportion	AAC1	PAC1	t	Proportion
All (n = 475)	15.47 (12.70)	11.81 (14.43)	3.75**	76.3%	23.92 (16.97)	10.34 (14.35)	13.61***	43.2%
M1 (n = 117)	15.82 (10.83)	11.82 (13.62)	2.33*	74.7%	22.10 (16.01)	10.00 (12.68)	7.08***	45.2%
M2 (n = 128)	16.76 (13.43)	15.44 (17.40)	.626	92.1%	24.67 (16.20)	11.33 (15.35)	7.59***	45.9%
M3 (n = 112)	12.81 (11.76)	8.42 (11.62)	2.56*	65.6%	25.64 (19.41)	13.45 (17.24)	5.33***	52.5%
M4 (n = 118)	16.19 (14.08)	10.95 (13.29)	2.51*	67.6%	23.45 (16.47)	6.55 (10.92)	7.53***	27.9%

Note. Proportion = the proportion of PAC1 to AAC1(i.e., PAC1/AAC1). Numbers in the parentheses below means are standard deviations. \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .01$ .

두 번째 메시지 (수원이와 자료)의 경우 중립적인 초기태도를 가졌던 응답자들은 메시지 영향력 과소평가 경향이 나타나지 않았다. 그 외에는 모두 과소평가 경향이 나타났고 중립적 초기태도 경우보다 부정적 초기태도의 경우 메시지 영향력 과소평가의 정도는 더 크게 나타났다.

## 5. 논의 및 결론

설득캠페인 메시지를 제작하는 과정에서 실무자들은 메시지의 실제 효과(actual message effect, AME)를 예측하기 위해 지각된 효과성(perceived message effect, PME)을 이용한다. 하지만 기존 연구들은 PME의 진단능력에 대해 서로 다른 연구 결과를 제시하고 있다. 일부 연구자들(e.g., Cappella, 2018; Davis & Duke, 2018; Dillard, Weber, & Vail et al., 2007; Noar, Barker et al., 2018)은 지각된 효과성이 실제 효과를 예측한다고 주장하지만, 오키프(O'Keefe, 2018)는 지각된 효과성이 실제 효과를 예측하는 확률은 우연적 확률에 불과하다고 주장하였다. PME의 진단능력을 인정하는 대부분의 연구들(예, Dillard, Weber et al., 2018)은 PME-AME 상관관계를 비교하는 방법을 사용한다. 이에 대해 오키프(1993)는 대립하는 의견을 논리적으로 제시한 바 있고, 상관관계를 이용한 검증방법을 비판하면서 두 메시지 간 PME와 AME 순위 비교 방법을 제시하여 그 결과를 토대로 진단능력을 부정하였다. 본 연구는 기존에 사용된 검증 방법들을 비판적으로 검토하고 한계점을 보완하였다. 그리고 여론조사보도의 영향력 지각 관련 연구(Chung et al., 2018)에서 사용된 지각된 보도의 영향력과 실제 영향력을 비교하는 방법을 설득캠페인 메시지 효과지각 연구에 적용하였다. 본 연구는 수정된 세 가지 검증 방법을 하나의 연구에 동시에 적용하여 AME에 대한 PME의 진단능력을 종합적으로 검증하였다.

연구 가설 및 문제를 검증하기 위해, 두 차례에 걸친 패널 조사(two-wave panel study)를 진행하였다. 메시지는 원자력 관련 공공기관(즉, 대한민국 정부, 원자력안전위원회, 한국수력원자력)이 유튜브(YouTube)에서 원자력 안전을 홍보하기 위해 제작한 동영상 자료를 사용하였다. 메시지 노출로 인한 실제 태도 변화와 응답자들이 지각한 메시지 영향력을 측정하여 분석한 주요 결과는 다음과 같다. 첫째, 메시지 전체의 PME-AME 상관관계는 PME 측정 유형에 상관없이 모두 통계적으로 유의미하게 정(+)적으로 나타났으며 상관관계의 평균은 .40으로 나타났다. 개별 메시지에 관한 PME-AME 간의 상관관계 또한 PME 측정 유형에 상관없이 모두 통계적으로 유의미하게 정(+)적으로 나타났으며 최소 .34에서 최대 .47로 나타났다. 이러한 수치는 비교적 높은 상관관계로(Cohen, 1988), 딜러드, 웨버 외(Dillard, Weber et al., 2007)가 메타 분석에서 발견한 PME-AME 상관관계의 평균 .41과 유사하다. 개별 메시지를 PME 측정 유형별로 분석한 결과, 모든 PME 측정 유형들이 AME와 비교적 높은 수준의 상관관계를 보였으며, 이는 PME가 측정 유형에 상관없이 AME에 대해 비슷한 정도의 진단능력을 가지고 있음을 보여준다. 지각된 태도 변화 정도는 기존 연구에서 많이 사용되지 않았음에도 두 개의 메시지

에서 AME와 가장 높은 상관관계를 나타냈다. 이러한 결과는 PME 진단능력 검증에서 지각된 태도 변화 정도가 유용한 변수로 사용될 수 있다는 것을 나타낸다.

오키프는 두 메시지 간 PME 순위와 AME 순위를 비교하는 방법을 제시하였지만, 분석에 사용된 자료들이 표본임에도 불구하고 통계적 유의미성을 분석에서 고려하지 않았다. 본 연구는 통계적 유의미성을 고려하는 등 오키프의 기준을 보완한 새로운 기준을 사용하여 두 메시지 간 PME와 AME 순위를 비교하였다. 분석결과, PME 진단능력 (즉, PME 순위와 AME 순위 일치)는 모든 메시지에서 PME 측정 유형에 상관없이 모두 인정되었다. 세부적으로 PME와 AME 모두 통계적으로 유의미한 차이가 나타나지 않아서 PME의 진단능력이 인정된 경우는 66.7%(20/30)로 나타났다. PME에서 통계적으로 유의미한 차이가 나타났지만, AME에서 통계적으로 유의미한 차이가 나타나지 않았기 때문에 PME의 진단능력을 약한 정도로 인정한 경우는 33.3%(10/30)로 나타났다. 이러한 결과는 PME의 진단능력을 매우 강하게 입증하는 것은 아니다. 강한 입증은 많은 경우에서 PME와 AME에서 모두 통계적으로 유의미한 차이가 같은 방향으로 발견될 때이다. 하지만 낮은 수준일지라도 AME에 대한 PME의 진단능력은 모든 경우에서 인정되었다. 이런 결과는 상관관계를 이용한 연구 결과와 일치한다.

본 연구는 또한 PME를 지각된 태도 변화 정도 그리고 AME를 실제 태도 변화 정도로 측정 후 두 변수를 직접 비교하였다. 분석결과, 사람들은 메시지의 입장과 자신의 기존 태도가 불일치할 때 실제 태도도 메시지 주장 방향으로 변화하고 자신의 태도가 메시지 주장 방향으로 변했다고 지각한 것으로 나타났다. 이는 지각된 태도 변화가 실제 태도 변화를 반영하고 있음을 의미한다. 그러나 응답자들은 자신들이 경험한 실제 태도 변화를 56.6% 축소하여 보고하였다. 초기 중립태도를 가진 집단의 경우, 지각된 태도 변화 정도는 실제 태도 변화 정도를 76.3% 반영하였고, 초기 반대태도 집단의 경우 지각된 태도 변화 정도는 실제 태도 변화 정도를 43.2%만 반영하고 있었다. 초기태도가 반대일 경우(vs. 초기태도 중립) 실제 태도 변화 정도를 과소평가 하는 경향이 더 강하게 나타났다.

메시지 수용자들이 지각한 메시지 효과성이 실제 메시지 효과를 어느 정도 반영하는지는 실제 캠페인 메시지 제작에 매우 중요한 문제이다. 본 연구의 결과는 메시지 제작에서 PME의 활용에 대해 다음과 같은 점을 시사한다. 첫째, 오키프(O'Keefe, 1993, 2018) 등은 캠페인 메시지 제작과 선택에서 PME를 사용하는 것에 회의적이었지만 본 연구의 결과는 PME가 AME와 비교적 높은 정도의 상관관계를 보이므로 효과적으로 활용될 수 있다는 점을 시사한다. 오키프(2018)의 검증방법 즉, PME 순위와 AME 순위의 일치정도도 본 연구에서는 매우 높게 나타났다. 본 연구는 PME를 지각된 메시지 논증(perceived argument quality), 메시지 영향력



추정(estimated message impact), 지각된 태도 변화 정도(perceived attitude change) 등으로 다양하게 측정하였다. PME-AME 상관관계 검증과 PME, AME 순위일치 여부 검증 등에서 PME 측정 유형 별 유의미한 차이는 발견되지 않았다. 캠페인 제작과정에서는 가능하다면 다양한 측정들을 통합적으로 사용하면 실제 효과를 보다 정확하게 예측할 수 있을 것이다. 본 연구 결과는 PME가 AME를 반영하기는 하지만 축소 반영할 가능성을 제시한다. 특히 캠페인의 목표대상이 메시지 주장에 대해 반대 견해를 가지고 있는 사람들일 경우, 그들은 메시지 노출로 인한 자신의 태도 변화 정도를 43% 정도로 축소하여 보고한다는 점이 발견되었다. 초기 반대 입장의 수용자들이 지각된 메시지 효과를 낮게 평가하더라도 실제에서는 메시지 효과가 있을 수 있다는 점을 보여준다. 초기 반대입장 수용자보다는 낮은 정도지만, 초기 중립입장을 가진 수용자들도 자신의 실제 태도 변화를 일정 정도 (76.3%)로 낮추어서 보고하는 경향이 나타났다. 캠페인 실무자들은 이런 경향성을 고려하여 지각된 메시지 효과성 자료를 활용해야 할 것이다.

본 연구의 결과는 설득캠페인 실무뿐 아니라 PME와 AME가 관련된 이론적 논쟁에도 시사점이 있다. AME에 대한 PME의 진단능력에 대한 학술적 논쟁은 오키프(2018)로 시작되었다. 이에 대해 다양한 반론들이 논문 또는 학술대회 논문으로 발표되었다. 하지만 기존 연구들은 오키프의 핵심 주장들을 체계적으로 검토하고 평가하지 못했다. 본 연구는 상관관계를 이용한 PME 진단능력 검증에 대한 오키프의 주장을 검토하고 오키프의 가상적인 자료 분석의 결과와 주장이 상관관계를 이용한 PME 진단능력 검증을 완전히 부정하지 못한다는 점을 밝혔다. 본 연구는 두 메시지 간 PME 순위와 AME 순위의 일치정도를 이용해 PME의 진단능력을 판단한 오키프(2018)의 기준을 새롭게 보완하였다. 새 기준을 사용하여 검증한 결과 PME의 진단능력은 100%로 나타났다. 본 연구가 제안한 기준들은 오키프의 검증 방법의 발전에 기여하고 향후 더 나은 검증 방법을 고안하는 데 중요한 도움이 될 것으로 기대한다.

본 연구는 AME에 대한 PME의 진단능력을 검증하면서 기존 연구들의 한계를 넘어 PME가 AME를 어느 정도 반영하는지를 측정하였는데 중요한 이론적 의미가 있다. 기존 연구들은 PME-AME 간 상관관계를 주로 검증하였다. 두 변수 간 상관관계가 매우 높더라도 PME는 AME를 과대평가할 수도 있고 과소평가할 수도 있다. 과대평가와 과소평가는 캠페인에 대한 실천적 의미가 크게 다르므로 마땅히 구분되어야 한다. 본 연구는 상관관계를 넘어 지각된 메시지 효과성의 정도를 실제 메시지 효과의 크기와 비교하였다. 비교 결과 응답자들이 실제 메시지 효과를 축소 보고한다는 것을 발견하였다. 그리고 실제 태도 변화에 대한 지각된 태도변화 정도의 반영률을 계산할 수 있었다. 이러한 결과들은 PME-AME 간의 관계를 규명하는 데 중요한 기여가 될 것이다.

본 연구는 AME에 대한 PME의 진단능력을 규명하기 위해 세 가지 방법을 동시에 한 연구에서 사용하였다는 점에 의의가 있다. 세 가지 방법들의 결과는 서로 내용적으로 일치하면서 서로의 검증을 상호보완하고 있었다. 세 가지 방법을 동시 사용하였을 때 PME-AME의 관계를 종합적으로 평가할 수 있었고 더 신뢰할 수 있는 결과를 산출할 수 있었다. 또한, 본 연구는 AME에 대한 PME의 진단능력을 규명하기 위해 공공기관이 실제로 실행한 설득캠페인 메시지를 조사에서 사용하였다. 실제 사용된 메시지들을 이용함으로써 기존 연구들보다 상대적으로 높은 생태학적 타당도(ecological validity)를 확보할 수 있었다.

본 연구는 한계점들은 다음과 같다. 본 연구는 원자력 발전소의 안전성을 홍보하기 위해 실제 기관들이 제작해 사용한 메시지들을 조사를 위해 사용하였다. 형식, 등장인물, 논리적 근거 등에서 서로 다른 메시지들을 비교하였다. 해당 메시지들에 대해 지각된 메시지 효과성은 서로 다르게 나타났지만, 실제 메시지 효과, 즉 메시지 노출로 인한 태도 변화의 정도는 메시지 간 통계적으로 유의미하게 차이 나지 않았다. 연구문제 중 하나는 PME와 AME에서의 순위의 일치 여부를 이용한 진단능력 판단이었다. 이 방법의 적용은 메시지들의 AME가 통계적 유의미하게 차이가 났을 때, 더 확실한 의미가 있을 수 있다. 메시지 간 실제 효과가 차이 나지 않았기 때문에 두 번째 연구 문제에 대한 결과를 실제 효과가 다른 메시지들의 경우로 일반화하는 데 어려움이 있다. 후속연구는 실제 효과에서도 차이가 나는 메시지들을 사용하여 PME의 진단능력을 검증할 필요가 있다.

본 연구는 네 개의 메시지를 이용하여 여섯 개의 메시지 쌍을 비교하였다. 여섯 개의 메시지 쌍을 사용할 경우, PME의 진단확률은 6 수준으로만 표시된다. 표본의 수가 6개로 한정되기 때문에 진단확률을 우연적 확률(50%)와 통계적으로 비교하기 힘들다. 본 연구는 메시지 노출 전과 후의 태도변화정도로 메시지의 실제 효과를 측정하였다. 이 연구설계에서 응답자가 복수의 메시지를 접하고 태도를 표시할 수 없으므로 더 많은 메시지를 사용하기에 어려움이 있었다. 후속 연구들은 보다 많은 메시지를 비교할 필요가 있다. 또한, 메시지 쌍을 비교해 진단확률을 측정하면서 세 개의 메시지 쌍에 동일한 메시지가 반복해서 사용되었다. 사례들이 서로 완전 독립적이지 않으므로 통계적 검증에 어려움이 발생한다. 이 문제는 오키프(2018)의 연구에도 해당된다. 본 연구는 진단확률이 100%로 나타났기에 사례들의 효과를 따로 통제할 필요는 없었다. 하지만 원칙적으로 각 메시지의 독자적인 효과를 가상변수(dummy variable)를 사용하여 통제하는 방법을 사용할 필요가 있다.

본 연구에서 사안에 대한 태도를 먼저 측정한 후 지각된 메시지 영향력을 측정하였기 때문에 태도의 표현 여부가 PME와 AME 간의 상관관계 등에 영향을 끼쳤을 가능성이 있다. 본 연

구 참여자들은 일차조사에서 자신의 태도를 이차 조사에서 정확하게 회상하지 못했다. 그러므로 실제 태도변화 정도를 기억해서 그에 맞추어 지각된 영향력을 보고했다고는 보기 어렵다. 이차 조사에서의 사안에 대한 태도가 더 긍정적일수록 메시지 영향력을 더 크게 보고했을 가능성을 제기할 수도 있다. 하지만 후속적인 자료분석의 결과 두 번째 조사에서의 태도와 지각된 메시지 영향력 간에 부(-)인 관계가 발견되었다 ( $r = -.326$ ). 후속연구에서는 자신의 태도를 표기하는 행동과 메시지 영향력 지각 간의 관계를 검증할 필요가 있다. 이를 검증하기 위해서는 두 가지 조건에서 연구를 진행할 필요가 있다. 한 조건에서는 두 번째 조사에서 메시지를 읽은 후 메시지 영향력 지각을 태도보다 먼저 측정하고, 다른 조건에서는 태도를 먼저 측정하고 지각된 영향력을 측정하는 것이다. 그러면 더 정확하게 PME와 AME 간의 상관관계를 측정할 수 있을 것이다.

지각된 메시지 효과성과 실제 메시지의 효과성을 서로 다른 사람들에게서 측정하고 비교하는 방법을 사용할 수도 있다. 두 변수를 다른 사람들에게 측정하는 것은 조사 참여자들을 두 개의 집단으로 무선할당한 후 한 집단의 참여자들은 모든 메시지에 대해 실제 영향력만을 측정하고 다른 한 집단의 참여자들은 모든 메시지에 대해 지각된 영향력만 측정할 때 가능하다. 만일 기존에 여러 메시지에 대해 메시지별 영향력 지각을 측정한 자료가 있다면 이를 활용해 메시지 실제 효과만을 측정한 후 순위를 비교할 수 있다.

본 연구에서 오키프 (2018)이 제안한 방법, 즉 두 메시지 간 PME와 AME 순위 일치여부를 확인하는 방법을 보완하여 사용하였다. 통계적 유의미성을 고려하여 순위 일치 여부를 판정하기 위해 기준들을 새로 정립하였다. 본 연구는 두 메시지 간 PME는 통계적으로 유의미한 차이가 나지만 AME는 통계적으로 유의미한 차이가 나타나지 않은 경우, 원칙적으로는 PME가 AME를 정확하게 진단하지 못했지만 PME의 진단능력을 부정하지 않았다 (원칙 3). 본 연구는 실제 메시지 효과는 단기간 또는 일회적인 노출로 나타나기 힘들 수 있고, 실제 메시지 효과를 측정하는 것은 상대적으로 어려워(Shen et al., 2018) 측정된 AME에서 차이가 나타나지 않을 수 있다는 이유로 PME의 진단능력을 부정하지 않았다. 오키프 (2018) 또한 이 경우 PME가 최소한 실제 효과가 더 낮은 메시지를 선택하지 않게 했다는 이유로 PME의 진단능력을 부정하지 않았다. 하지만 엄격하게 본다면 이 경우도 PME가 AME를 제대로 진단하지 못한 경우라고 볼 수 있다. 만일, 이 경우 PME의 진단능력을 인정하지 않는다면, 본 연구에서의 PME 진단능력 인정률은 100%에서 70.8%(17/24)로 낮아진다. 후속연구에서는 PME의 진단능력 인정 여부에 대해 어떠한 기준을 사용하는 것이 적절한지에 대해 더 자세한 논의가 필요하다.

지각된 메시지 효과성의 실제 효과 진단능력에 대한 논쟁은 오키프(1993)로부터 비롯되었다. 해당 논문에서 오키프는 지각된 메시지 효과성은 실제 메시지 효과와는 다른 심리적 과정들

을 거쳐 형성된다고 주장하였다. 즉, 수용자들은 메시지의 효과에 관한 상식적인 관념들이 있고 그런 관념들에 근거해서 지각된 메시지 효과성을 형성하지만, 실제 메시지 효과는 수용자가 의식하지 못하는 심리적 과정들에 의해서도 영향받는다고 주장하였다. 지각된 효과성과 실제 효과의 형성과정과 개입 변수가 다르므로 두 변수의 내용과 함의가 다르다고 주장하였다. 본 연구는 지각된 메시지 효과성과 실제 메시지 효과를 비교하였을 뿐, 그 형성과정에 대해 깊게 논의하거나 검증하지 않았다. 지각된 메시지 효과성을 측정하기 위해 지각된 논증 품질, 메시지 영향력 추정, 지각된 태도 변화 정도 등을 측정했지만 그 변수 간의 관계에 대해 논구하지 않았다. 지각된 논증 품질은 메시지 영향력 추정과 실제 태도 변화에 모두 영향을 끼칠 수 있다. 메시지 영향력 추정과 실제 태도 변화 정도의 상관관계의 일부는 공통된 선행요인(예, 지각된 논증 품질)에서 비롯되었을 수 있다. 하지만 수용자가 영향 과정을 알아차리기 힘들거나 인정하기 힘든 설득요인들(예, 정보원의 매력성, 감성 호소) 등은 실제 효과에서는 발견되지만 지각된 메시지 효과성에서는 나타나지 않을 수 있다. 지각된 메시지 효과성과 실제 메시지 효과의 공통된 선행요인들은 무엇인지 각 변수에 특수하게 영향을 미치는 선행요인들은 무엇인지를 이론적으로 설명하고 경험적으로 검증하는 노력이 앞으로 필요하다.

## References

- Atkin, C. K., & Freimuth, V. (2001). Formative evaluation research in campaign design. In R. E. Rice & C. K. Atkin (Eds.), *Public communication campaigns* (pp. 125-145). Thousand Oaks, CA: SAGE. doi:10.4135/9781452233260.n7
- Atkin, C. K., & Rice, R. E. (2013). Advances in public communication campaigns. In E. Scharrer (Ed.), *The international encyclopedia of media studies: Media effects/media psychology* (pp. 526-551). London, UK: Wiley-Blackwell
- Biggsby, E., Cappella, J. N., & Seitz, H. H. (2013). Efficiently and effectively evaluating public service announcements: Additional evidence for the utility of perceived effectiveness. *Communication Monographs, 80*, 1-23. doi:10.1080/03637751.2012.739706
- Byrne, S., Katz, S. J., Mathios, A., & Niederdeppe, J. (2015). Do the ends justify the means? A test of alternatives to the FDA proposed cigarette warning labels. *Health Communication, 30*, 680-693. doi:10.1080/10410236.2014.895282
- Cappella, J. N. (2018). Perceived message effectiveness meets the requirements of a reliable, valid, and efficient measure of persuasiveness. *Journal of Communication, 68*, 994-997. doi:10.1093/joc/jqy044
- Cesario, J., Grant, H., & Higgins, E. T. (2004). Regulatory fit from persuasion: Transfer from “feeling right.” *Journal of Personality and Social Psychology, 86*, 388-404. doi:10.1037/0022-3514.86.3.388
- Chung, S., Heo, Y.-J., & Moon, J.-H. (2018). Perceived versus actual polling effects: Biases in perceptions of election poll effects on candidate evaluations. *International Journal of Public Opinion Research, 30*, 420-442. doi:10.1093/ijpor/edx004
- Chung, S., & Lee, W. (2012). Reconstructing of the third-person perception hypothesis: Focusing on the effect of preexisting attitudes, the level of knowledge, and message strength. *Korean Journal of Journalism & Communication Studies, 56*(5), 322-349.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi:10.1037/0033-2909.112.1.155
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge.
- Cohen, J., Mutz, D., Price, V., & Gunther, A. (1988). Perceived impact of defamation: An experiment on third-person effects. *Public Opinion Quarterly, 52*, 161-173. doi:10.1086/269092
- Collins, R. L., Taylor, S. E., Wood, J. V., & Thompson, S. C. (1988). The vividness effect: Elusive or illusory? *Journal of Experimental Social Psychology, 24*, 1-18. doi:10.1016/0022-1031(88)90041-8

- Davison, W. P. (1983). The third-person effect in communication. *Public Opinion Quarterly*, 47, 1-15. doi: 10.1086/268763
- Davis, K. C., & Duke, J. C. (2018). Evidence of the real-world effectiveness of public health media campaigns reinforces the value of perceived message effectiveness in campaign planning. *Journal of Communication*, 68, 998-1000. doi:10.1093/joc/jqy045
- Dillard, J. P., & Peck, E. (2000). Affect and persuasion: Emotional responses to public service announcements. *Communication Research*, 27, 461-495. doi:10.1177/009365000027004003
- Dillard, J. P., Shen, L., & Vail, R. G. (2007). Does perceived message effectiveness cause persuasion or vice versa? 17 consistent answers. *Human Communication Research*, 33, 467-488. doi:10.1111/j.1468-2958.2007.00308.x
- Dillard, J. P., Weber, K. M., & Vail, R. G. (2007). The relationship between the perceived and actual effectiveness of persuasive messages: A meta-analysis with implications for formative campaign research. *Journal of Communication*, 57, 613-631. doi:10.1111/j.1460-2466.2007.00360.x
- Dillard, J. P., & Witte, K. (1993). Possessions theory of persuasion: An examination of its basis and range. *Communication Studies*, 44, 188-199. doi:10.1080/10510979309368394
- Douglas, K. M., & Sutton, R. M. (2004). Right about others, wrong about ourselves? Actual and perceived self-other differences in resistance to persuasion. *British Journal of Social Psychology*, 43, 585-603. doi:10.1348/0144666042565416
- Duke, J. C., Alexander, T. N., Zhao, X., Delahanty, J. C., Allen, J. A., MacMonegle, A. J., & Farrelly, M. C. (2015). Youth's awareness of and reactions to the real cost national tobacco public education campaign. *PLoS One*, 10. doi:10.1371/journal.pone.0144827
- Hale, J. L., Mongeau, P. A., & Randi, M. T. (1991). Cognitive processing of one- and two-sided persuasive messages. *Western Journal of Speech Communication*, 55, 380-389. doi:10.1080/10570319109374394
- Hullett, C. R. (2002). Charting the process underlying the change of value-expressive attitudes: The importance of value-relevance in predicting the matching effect. *Communication Monographs*, 69, 158-178. doi:10.1080/714041711
- Hullett, C. R. (2004). Using functional theory to promote sexually transmitted disease (STD) testing: The impact of value-expressive messages and guilt. *Communication Research*, 31, 363-396. doi:10.1177/0093650204266103
- Hullett, C. R., & Boster, F. J. (2001). Matching messages to values underlying value-expressive and

- social-adjustive attitudes: Reconciling an old theory with a contemporary measurement approach. *Communication Monographs*, 68, 133-153. doi:10.1080/03637750128057
- Kim, S. H. (2017, 7, 23). The controversy over nuclear power plants, this is the fact. The Hankyoreh. Retrieved from <http://www.hani.co.kr/arti/economy/marketing/803903.html>
- Lavine, H., & Snyder, M. (1996). Cognitive processing and the functional matching effect in persuasion: The mediating role of subjective perceptions of message quality. *Journal of Experimental Social Psychology*, 32, 580-604. doi:10.1006/jesp.1996.0026
- McGuire, W. J., & Millman, S. (1965). Anticipatory belief lowering following forewarning of a persuasive attack. *Journal of Personality and Social Psychology*, 2, 471-479. doi:10.1037/h0022486
- Meirick, P. C. (2004). Topic-relevant reference groups and dimensions of distance: Political advertising and first-and third-person effects. *Communication Research*, 31, 234-255. doi:10.1177/0093650203261514
- Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13, 151-161. doi:10.2307/1392369
- Miller, G. R. (1980). On being persuaded: Some basic distinctions. *Persuasion: New Directions in Theory and Research*, 1, 1-28.
- Mitchell, M. (2000). Able but not motivated? The relative effects of happy and sad mood on persuasive message processing. *Communication Monographs*, 67, 215-226. doi:10.1080/03637750009376505
- Mitchell, M., Brown, K., Morris-Villagran, M., & Villagran, P. (2001). The effects of anger, sadness and happiness on persuasive message processing: a test of the negative state relief Model. *Communication Monographs*, 68, 347-359. doi:10.1080/03637750128070
- Noar, S. M., Barker, J., & Yzer, M. (2018). Measurement and design heterogeneity in perceived message effectiveness studies: A call for research. *Journal of Communication*, 68, 990-993. doi:10.1093/joc/jqy047
- Noar, S. M., Bell, T., Kelley, D., Barker, J., & Yzer, M. (2018). Perceived message effectiveness measures in tobacco education campaigns: A systematic review. *Communication Methods and Measures*, 12, 295-313. doi:10.1080/19312458.2018.1483017
- O'Keefe, D. J. (1993). Understanding social influence: Relations between lay and technical perspectives. *Communication Studies*, 44, 228-238. doi:10.1080/10510979309368397
- O'Keefe, D. J. (2018). Message pretesting using assessments of expected or perceived persuasiveness: Evidence about diagnosticity of relative actual persuasiveness. *Journal of Communication*, 68,

120-142. doi:10.1093/joc/jqx009

- O'Keefe, D. J., & Reid, K. (1990). The uses and effects of public service advertising. In L. A. Grunig & J. E. Grunig (Eds.), *Public relations research annual* (Vol. 2, pp. 67-91). Hillsdale, NJ: Erlbaum.
- Perloff, R. M. (2009). Mass media, social perception and the third-person effect. In J. Bryant & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (3rd ed., pp. 252-268). New York, NY: Routledge.
- Pfau, M., & Parrott, R. (1993). *Persuasive communication campaigns*. Needham Heights, MA: Allyn and Bacon.
- Reid, S. A., Gunter, H. N., & Smith, J. R. (2005). Aboriginal self-determination in Australia. *Human Communication Research*, 31, 189-211. doi:10.1093/hcr/31.2.189
- Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations*. Monterey, CA: Brooks/Cole
- Schoenbach, K., & Becker, L. B. (1995). Origins and consequences of mediated public opinion. In T. L. Glasser & C. T. Salmon (Eds.), *Public opinion and the communication of consent* (pp. 323-347). New York, NY: Guilford Press.
- Shen, L., Sun, Y., & Pan, Z. (2018). Not all perceptual gaps were created equal: Explicating the third-person perception (TPP) as a cognitive fallacy. *Mass Communication and Society*, 21, 399-424. doi:10.1080/15205436.2017.1420194
- Tal-Or, N., & Drukman, D. (2010). Third-person perception as an impression management tactic. *Media Psychology*, 13, 301-322. doi:10.1080/15213269.2010.503516
- Wagner, C. H. (1982). Simpson's paradox in real life. *The American Statistician*, 36, 46-48. doi:10.1080/00031305.1982.10482778
- Yzer, M., LoRusso, S., & Nagler, R. H. (2015). On the conceptual ambiguity surrounding perceived message effectiveness. *Health Communication*, 30, 125-134. doi:10.1080/10410236.2014.974131

최초 투고일 2020년 8월 6일

게재 확정일 2020년 11월 13일

논문 수정일 2020년 11월 30일



## 부록 I. 국내 참고 문헌

김성환 (2017, 7, 23). 탈원전 논란, 이것이 팩트다. <한겨레>. URL:<http://www.hani.co.kr/arti/economy/marketing/803903.html>

정성은·이원지 (2012). 제삼자 지각 가설의 재구성: 메시지 강도, 관련 지식 보유 정도, 기존 태도를 중심으로. <한국언론학보>, 56권 5호, 323-349.

## 지각된 메시지 효과성은 실제 메시지 효과를 얼마나 반영하는가? 설득 캠페인 메시지의 지각된 효과와 실제 효과

이병현

(성균관대학교 미디어커뮤니케이션학과 연구원)

정성은

(성균관대학교 미디어커뮤니케이션학과 교수)

설득캠페인 메시지의 효과를 사전에 예측하기 위해 캠페인 실무자들은 주로 지각된 메시지 효과성(perceived message effectiveness, PME)을 활용한다. 하지만 PME로 실제 효과(actual message effect, AME)를 예측할 수 있는지에 대한 논쟁은 지속적으로 진행되고 있다. 이 논쟁의 해결을 위해 본 연구는 기존 연구들이 사용한 세 가지 방법, 즉 PME와 AME 간의 상관관계를 이용하는 방법, 두 메시지 간 PME와 AME의 순위일치 여부를 이용하는 방법, 그리고 메시지 노출로 인한 실제 태도변화의 정도와 지각된 태도변화의 정도를 직접 비교하는 방법들을 체계적으로 검토 보완하였다. PME와 AME의 다양한 측정 방식을 검토하여 이를 유형화하였다. 본 연구는 각 검증 방법과 관련된 연구 문제들을 제시하고 세 가지 검증 방법을 동시 적용하여 PME와 AME의 관계를 종합적으로 검증하였다. 연구 문제의 검증을 위해, 공공기관이 원자력 발전소 안전성을 홍보하기 위해 제작 배포했던 네 개의 동영상 메시지를 사용하였다. 두 차례에 걸친 패널 조사(N = 475)의 일차 조사에서 응답자들의 초기 태도를 측정하였으며, 일주일 후 이차 조사에서 네 개의 메시지를 서로 다른 집단에게 무선할당하여 제시한 후, PME를 유형 별로 측정하였고, 메시지 노출 전후 사안에 관한 태도의 차이로 AME를 측정하였다. 연구 결과, PME-AME 상관관계는 전체 메시지의 경우와 개별 메시지의 경우 모두 정(+)적으로 나타났으며, 전체 메시지 상관관계는 평균 .40, 개별 메시지에 관한 상관관계는 최소 .34에서 최대 .47로 비교적 높게 나타났다. 오키프(O'Keefe, 2018)의 방법을 보완하여 두 메시지 간 PME와 AME의 순위 일치를 분석한 결과, PME의 진단능력 인정률은 100%(30/30)으로 나타났다. 사안에 관한 초기태도가 메시지의 입장과 반대일 경우, 응답자들은 메시지 노출로 인한 자신의 태도 변화 정도를 실제 태도 변화의 43.2%로 축소하여 보고함을 발견하였다. 초기 태도가 중립일 경우에는 실제 태도의 76.3%를 지각된 태도 변화로 보고하였다. 이러한 결과는 설득캠페인 메시지 제작에서 PME를 활용하여 AME를 진단할 수 있지만, 실무자들은 PME가 포함하고 있는 체계적 편향을 인지할 필요가 있음을 시사한다.

**핵심어:** 지각된 메시지 효과성, 실제 메시지 효과, 설득캠페인, 메시지 영향력 추정, 지각된 태도변화 정도