ARTICLE IN PRESS
JID: AMJ [m5G;March 29, 2018;2:25]
Australasian Marketing Journal 000 (2018) 1–6
Contents lists available at ScienceDirect
Australasian Marketing Journal
journal homepage: www.elsevier.com/locate/ausmj

Invited Paper

Mediation analysis revisited: Practical suggestions for addressing common deficiencies

Jungkeun Kim[a,*], Euejung Hwang[a], Megan Phillips[a], Sungha Jang[b], Jae-Eun Kim[c], Mark T. Spence[d], Jongwon Park[e]

[a] Department of Marketing, Auckland University of Technology, Auckland, New Zealand
[b] Department of Marketing, Kansas State University, Manhattan, KS 66506, USA
[c] Department of Marketing, University of Auckland, Auckland, New Zealand
[d] Bond Business School, Bond University, QLD 4226, Australia
[e] Korea University Business School, Korea University, Seoul, South Korea

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Mediation analysis
Baron and Kenny causal chain approach
Reverse mediation
Experimental-causal-chain design
Moderation-of-process design

ABSTRACT

Four issues that can affect statistical conclusions from mediation analysis are presented here: The implications of omitting mediators; not conducting reverse mediation analysis; using inappropriate measures; and not considering a wider array of experiment-based methods. Suggestions for addressing each of these are advanced. Previous issues of AMJ, JMR and JCR are then examined to gauge the extent to which these suggestions were used. Less than half of the published papers inspected (44.4% of the total) endeavored to address at least three of the four issues raised above. AMJ authors will realize higher statistical as well as theoretical rigor if they consider these suggestions.

1. Introduction

Providing evidence of mediation in empirical experimental studies is critical to theory development and theory testing in marketing research. A mediator is “a variable (M) that transmits the effect of an antecedent variable (X) to an outcome variable (Y) in a casual sequence such that X causes M and M causes Y” (i.e., $X \rightarrow M \rightarrow Y$, MacKinnon et al., 2013, p. 338). In the last 30 years significant advancements have been made on how to identify mediators, both theoretically (e.g., Baron and Kenny, 1986; Spencer et al., 2005) and statistically (e.g., Baron and Kenny, 1986; Hayes, 2009). In this research note, we review some of the main approaches and their considerations when conducting mediation analysis, and then advance practical suggestions for how to address common pitfalls with respect to mediation analysis. Specifically, we discuss the statistical ramifications of omitting mediators, the benefits of reverse mediation analysis, using multi-item measures to reduce measurement error, and combining experiment-based methods. We conclude by perusing recent issues of *Journal of Marketing Research, Journal of Consumer Research* as well as *Australasian Marketing Journal* to assess the extent to which authors embrace one or more of the suggestions advanced.

2. The causal chain approach to simple mediation analysis

Efforts to unearth mediating variables have a long and rich history (consider, for example, Festinger, 1957). Early studies might have inferred mediators based on theoretical arguments or provided evidence through correlational analyses (Spencer et al., 2005). Baron and Kenny’s seminal article (1986) changed the landscape. They distinguished between mediators and moderators, and suggested a way of empirically testing for a mediation relationship based on a series of regression models, the causal chain approach.[1] They proposed that three conditions can demonstrate simple mediation. First, X should significantly influence M (i.e., path a in Fig. 1). Second, X should significantly affect Y (path c). Finally, when both X and M are included in the model to predict Y, M should significantly influence Y (path b) and the effect of X on Y (path c') should be no longer significant or reduced significantly compared to the direct effect of X on Y (path c). Thus, ab captures the indirect effect of X on Y via the mediator, and c' is the direct effect of X on Y. If path c' is zero, it indicates full mediation ($c = ab$); when path c' is not zero, it suggests partial mediation. Since its explication

* Corresponding author.
E-mail addresses: jungkeun.kim@aut.ac.nz (J. Kim), euejung.hwang@aut.ac.nz (E. Hwang), megan.phillips@aut.ac.nz (M. Phillips), sjang@ksu.edu (S. Jang), bestjek@gmail.com (J.-E. Kim), mspence@bond.edu.au (M.T. Spence), amadeus@korea.ac.kr (J. Park).

[1] It is also referred to as a ‘measurement-of-mediation design’ (Spencer et al., 2005).

https://doi.org/10.1016/j.ausmj.2018.03.002
1441-3582/

Please cite this article as: J. Kim et al., Mediation analysis revisited: Practical suggestions for addressing common deficiencies, Australasian Marketing Journal (2018), https://doi.org/10.1016/j.ausmj.2018.03.002

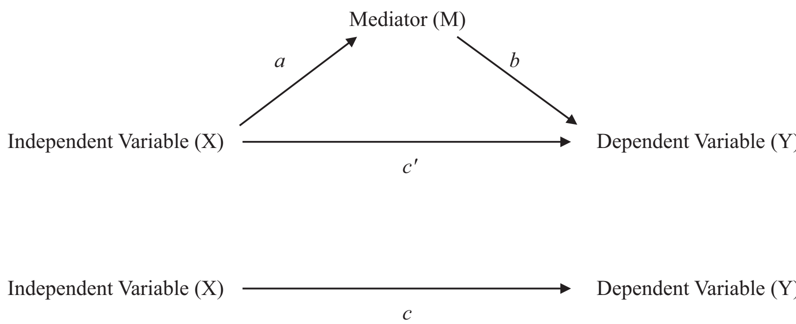

Fig. 1. The Baron and Kenny's model of mediation.

three decades ago, the Baron and Kenny causal chain approach has proven quite popular in psychology as well as in marketing, cited over 73,500 times based on Google scholar as of January 2018.

However, weaknesses have been identified with this approach (Jacoby and Sassenberg, 2011; MacKinnon and Pirlott, 2015; Zhao et al., 2010). First, except in the case of full mediation (i.e., when the direct effect of X on Y becomes zero taking a mediator into account),[2] there is a possibility of omitted mediator(s) that might bias the effect of X on Y (Bullock et al., 2010; Zhao et al., 2010). Second, unobserved variables could have an influence on both X and M, or M and Y, which could also lead to biased statistical conclusions (MacKinnon and Pirlott, 2015). Third, there can be a measurement order effect. Measuring the mediator first can change Y, or conversely, measuring Y can influence the measurement of the mediator. Researchers often measure all constructs in one setting and it is common to measure the mediator(s) after measuring the dependent variable (Fiedler et al. 2011; Iacobucci et al., 2007; Jacoby and Sassenberg, 2011). Fourth, although Baron and Kenny (1986) suggest using the Sobel (1982) test for confirming significance of the indirect effect of X on Y via M, the Sobel test assumes that the sampling distribution of the indirect effect is normal. Hayes (2009) notes that this may not be the case. A fifth issue to consider is that raised by Zhao et al. (2010) and Shrout and Bolger (2002) who argue that there do not need to be a significant relationship between X on Y (path c in Fig. 1) for there to be an indirect effect of X on Y through M. In such a situation, if a researcher were to strictly follow the methodology advanced by Baron and Kenny (1986), they would discontinue their investigation after the first $X \rightarrow Y$ regression model failed.

Scholars have therefore advanced different statistical approaches to conducting mediation analysis, most notably bootstrapping (Shrout and Bolger, 2002; Preacher and Hayes, 2004; Zhao et al., 2010). Simulations have shown bootstrapping to be a more powerful approach to identifying mediators relative to the causal chain approach (Hayes, 2009). In addition, researchers often use latent measurement models such as structural equation models because they consider the measurement error of the mediator(s) and/or dependent variable (Iacobucci et al., 2007); it is also easier to run more complex models that consider a wider array of constructs/construct relationships (MacKinnon and Pirlott, 2015). Nevertheless, Baron and Kenny deserve much praise for greatly advancing the statistical approach to identifying mediators, and the underpinnings of their approach are largely intact.

Methodological advancements aside, within the next section, we discuss four issues that can adversely affect statistical conclusions reached from mediation analyses and suggest remedies. We start with the case where erroneous conclusions are reached regarding the residual direct effect of X on Y (path c') due to the omission of one or more additional mediating variables.

3. Common issues and possible remedies

3.1. Consider the implications of omitted mediators

Recall that if $c = ab$, there is no direct effect of X on Y; instead, M fully mediates the relationship between X and Y (Baron and Kenny, 1986; Shrout and Bolger, 2002; Zhao et al., 2010). In these situations researchers commonly assume that there is no omitted mediator(s). However, we demonstrate that there is a possibility that unaccounted mediators can exist, even when $ab = c$ (or $c' = 0$). Through a simulation, it is shown that significantly different conclusions can be reached (e.g., an insignificant direct effect becomes negative) due to incomplete model specification; this is a separate concern to that of under-defining the psychological processes at play, what Spencer et al. (2005) call the theoretical analysis.

Consider Fig. 2. Assume in an omniscient world that there are three mediators in the model. M_1 (i.e., $a_1 \times b_1 = 1 \times 1 = 1$) and M_2 (i.e., $a_2 \times b_2 = 1 \times 1 = 1$) have positive influences on Y, whereas M_3 (i.e., $a_3 \times b_3 = 1 \times -1 = -1$) has a negative influence on Y. If a researcher theorized a parsimonious $X \rightarrow M \rightarrow Y$ relationship, hence only measured M_1, the results would indicate full mediation and the researcher could erroneously assume that M_1 is the only mediating mechanism for X on Y. In fact, in this hypothetical case two opposite mediators (i.e., M_2 and M_3) exist. Claiming *full* mediation based on the single mediator is therefore risky.[3]

Endeavoring to eliminate extraneous variables is a criterion for demonstrating causality. Here, we explore the statistical ramifications of the obverse: omitting relevant variables that can lead to confounder bias (MacKinnon and Pirlott, 2015). Referring again to Fig. 2, if a researcher measured M_1 and M_2 in her empirical setting, the direct effect of X on Y (i.e., c') could be significantly negative. In contrast, if the researcher measured only M_1 and M_3, the direct effect of X on Y (i.e., c') could be significantly positive. We demonstrate this using a simulation.

Consider the following general model:

$$\begin{aligned} M_1 &= a_1 X + u_1 \\ M_2 &= a_2 X + u_2 \\ M_3 &= a_3 X + u_3 \\ Y &= b_1 M_1 + b_2 M_2 + b_3 M_3 + c' X + \varepsilon \end{aligned}$$

To be consistent with the settings in Fig. 2, we set a_1, a_2, and a_3 to 1, b_1 and b_2 to 1, b_3 to -1, and $c' = 0$. Let X be drawn from a standard uniform distribution and assume that each error term

[2] In section 3.1 we present a simulation that demonstrates mediating variables could be omitted even in the case of full mediation.

[3] Rucker et al. (2011) made a similar warning regarding full mediation; however, their argument is based on measurement error in the case of two mediators.