Omnibus Tests of the Indirect Effect in Statistical Mediation Analysis with a Multicategorical Independent Variable

Patrick S. Creedon[1], Andrew F. Hayes[1], and Kristopher J. Preacher[2]

[1]Department of Psychology, The Ohio State University [2]Department of Psychology and Human Development, Vanderbilt University

Abstract

Hayes and Preacher (2014, *British J of Math & Statistical Psych*) present an approach to statistical mediation analysis when the independent variable *X* is multicategorical. However, whether *M* can be deemed a mediator of the effect of *X* on *Y* depends on how the groups are coded. We evaluated an omnibus test of mediation without this limitation based on adjusted R^2 when estimating *M* from *X* multiplied by the effect of *M* on *Y*. We also evaluated the test of joint significance and the causal steps approach popularized by Baron and Kenny (1986, JPSP). The test of joint significance performed adequately if not better than methods based on adjusted R^2. The causal steps method performed worst.

Mediation with a Multicategorical X

- Mediation analysis is used to answer questions of "how" or "by what process" *X* affects *Y*.
- The *indirect effect* of *X* on *Y* quantifies the sequence of causal steps by which *X* affects *Y* through a mediator variable *M*.
- When *X* represents *k* distinct groups, the effect of group on *M* and *Y* can be estimated using *k* – 1 dummy variables D_i or some other group coding system.
- Using OLS regression, the indirect effect of *X* on *Y* can be estimated using two linear models:

$$(1)\quad M = i_M + \sum_1^{k-1} a_i X + e_M$$
$$(2)\quad Y = i_Y + bM + \sum_1^{k-1} c'_i X + e_Y$$

- The *k* – 1 *relative indirect effects* of *X* on *Y* are the products of the a_i and *b* paths from equations 1 and 2 (see Figure 1).
- Each relative indirect effect estimates how much a particular group differs from a reference group on *Y* as a result of *X*'s influence on *Y* through *M*.

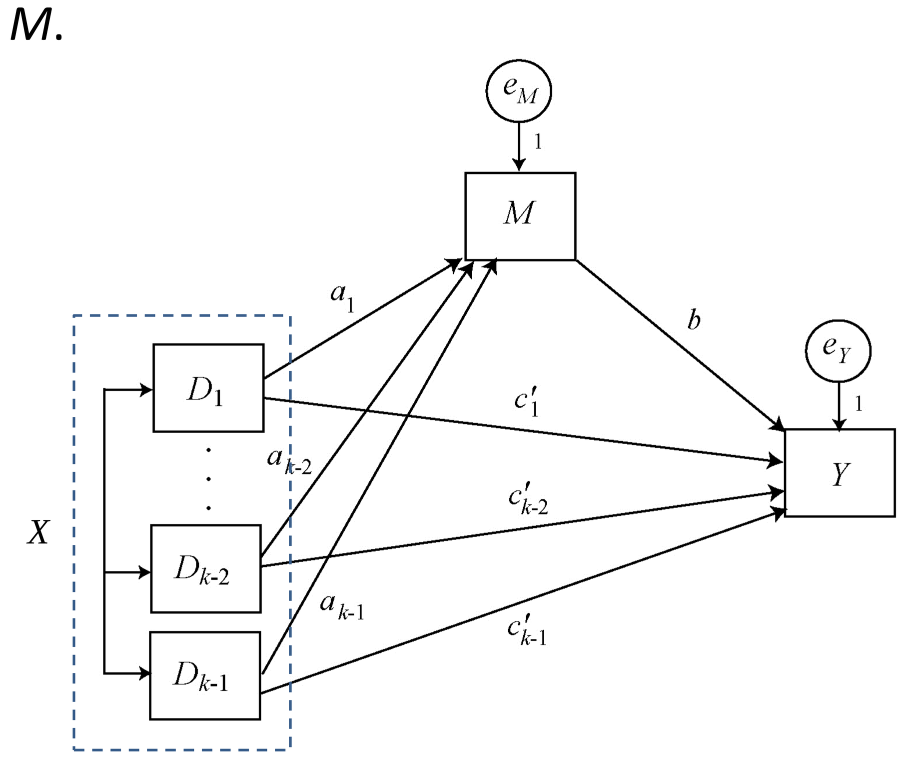


Figure 1. A mediation path diagram with *X* representing *k* groups

Problems with Inference

- Inference about relative indirect effects can be assessed using a bootstrap confidence interval. However, omnibus inference about mediation of the effect of *X* can depend on how the groups are coded.
- An omnibus test of mediation based on a confidence interval (CI) for R^2_{MX} multiplied by *b* avoids this problem, as it is invariant across methods of coding groups. A CI that doesn't include zero = mediation.
- However, the squared sample multiple correlation coefficient is known be positively biased, which likely would produce an invalid omnibus test of mediation without some kind of adjustment.
- A number of "adjusted R^2" measures exist that are less biased. We explore their relative performance in this omnibus test of mediation.

The Monte Carlo Simulation

- We conducted a Monte Carlo simulation, generating samples from a population mediation process. We varied:
  - Number (*k*) of groups (*k* = 3, 4, or 5)
  - Pattern of spread of the *k* group means on *M*
  - Size of the *k* – 1 a_i paths to manipulate the % of variance in *M* explained by group (R^2_{MX} = .00, .02, .13, .26)
  - Size of the *b* path (*b* = 0, .14, .39 .59)
  - Sample size per group (*n* = 25, 50, 100)
- *k* X *n* values of *M* were generated from equation (1), where e_M was a normal deviate with variance determined by R^2_{MX}.
- *Y* was generated from equation (2) where e_Y was a standard normal deviate. Relative direct effects (c'_i) were fixed at zero.
- The omnibus indirect effect was tested using the test of joint-significance (both R^2_{MX} and *b* statistically significant), the causal steps method (joint significance plus a statistically significant total effect of *X*), and bootstrap CIs (1000 bootstrap samples) for $R^2_{MX}b$ using 6 measures of adjusted R^2 (Smith, Wherry-1, Wherry-2, Olkin-Pratt, Pratt, and Claudy-3) as well as unadjusted R^2.
- Type I error rate was calculated for each combination of conditions corresponding to no indirect effect of *X*, across 1000 replications per condition (i.e., in conditions in which population $R^2_{MX}b$ = 0).

Figure 2. $b = 0$

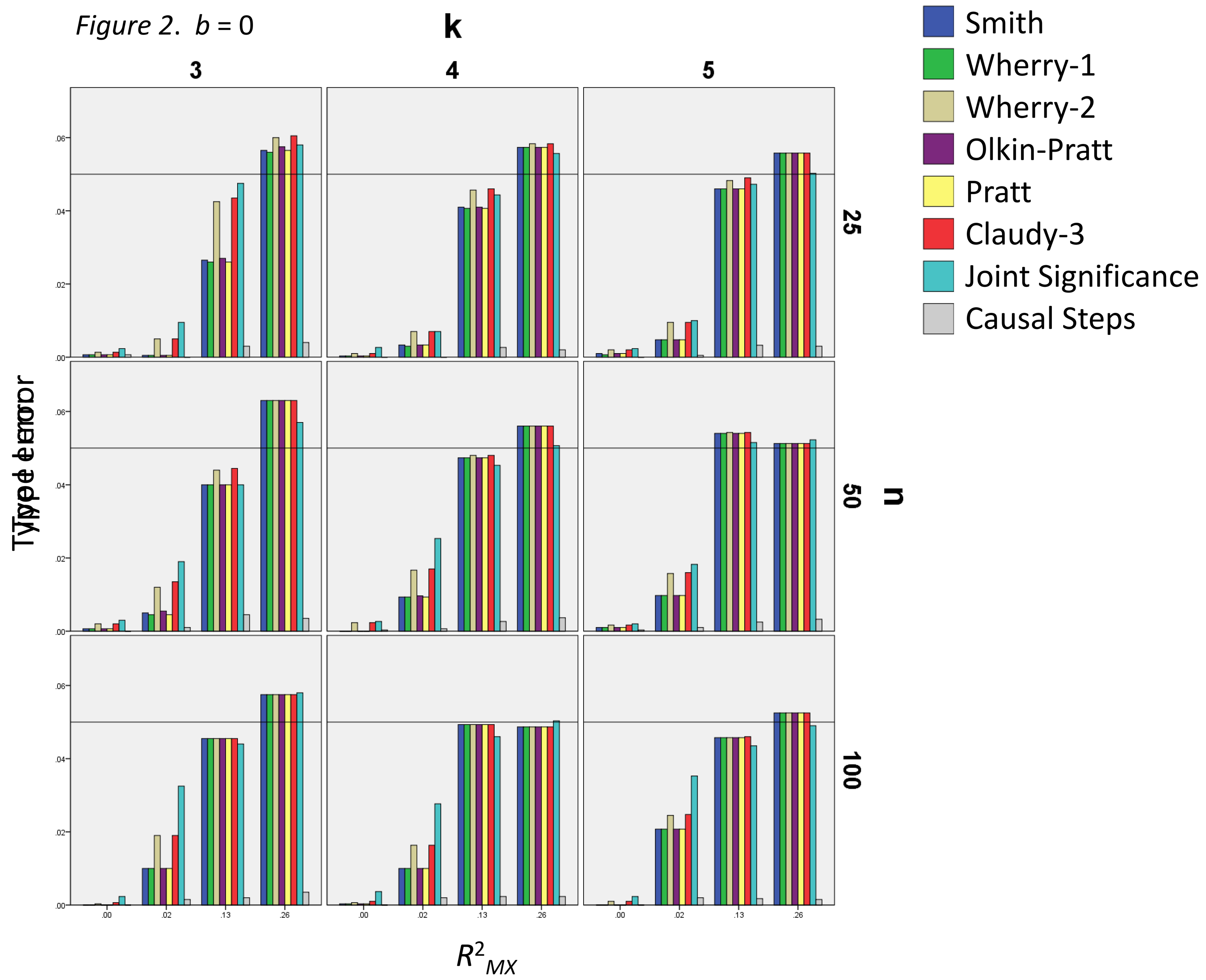
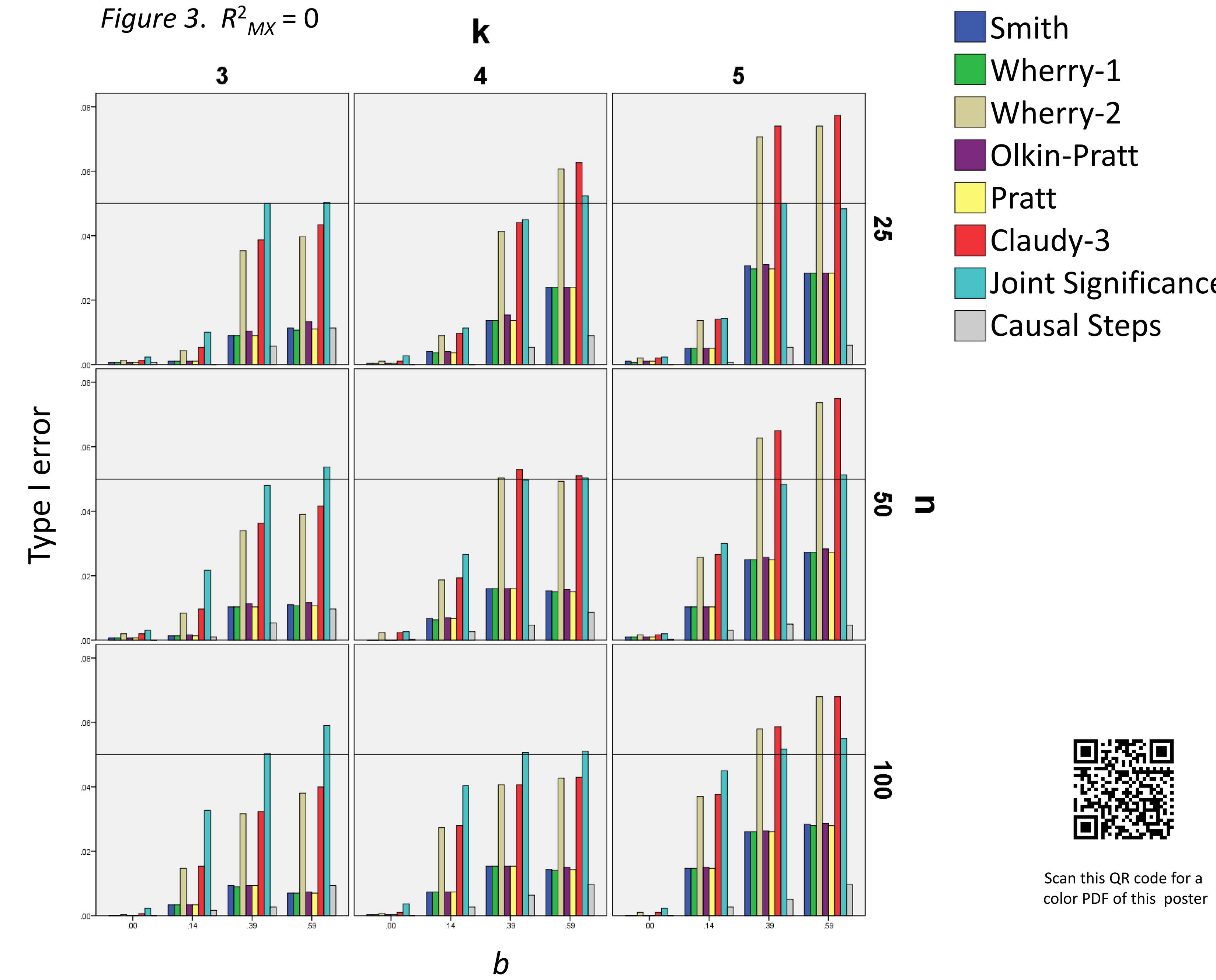


Figure 3. $R^2_{MX} = 0$



Results and Discussion

- Type I error rates can be found in Figures 2 and 3. Results collapse across the spread of the group means on *M*, as this had little effect on the results.
- As expected, using an unadjusted R^2 *in* $R^2_{MX}b$ performed terribly, with Type I error rates approaching one as *M*'s effect on *Y* or *n* grows larger. These results are not displayed is the figures below.
- When *b* = 0, each test performed about the same, with conservatism that declines as R^2_{XM} increases. The exception is the causal steps ("Baron and Kenny") method, which was always very conservative.
- When R^2_{MX} = 0, most of the methods performed similarly, with conservatism that declined as *b* increased. But using the Wherry-2 and Claudy-3 adjusted R^2 measures in the index resulted in a liberal test when *b* was larger. The causal steps method was very conservative.
- The Type I error rate of the test of joint significance (where R^2_{MX} and *b* are both significant) was generally closer to the .05 level than the other methods.

Recommendation:

- The test of joint significance is adequate as an omnibus test of the indirect effect. It is generally less conservative than other tests we examined. It doesn't yield an interval estimate, but that is ok because $R^2_{MX}b$ has no meaningful interpretation anyway.
- Avoid the causal steps approach. Requiring a total effect of *X* prior to testing the indirect effect results in a *very* conservative test, much more so than other methods. Its Type I error rate is consistently very small even when one of the components of the indirect effect is large.

Scan this QR code for a color PDF of this poster