The problem is on the radar in psychology; for example, Smith (2012) in a recent editorial in *Journal of Personality and Social Psychology* stated:

"if the independent variable (X) is manipulated and the mediator (M) and dependent variable (Y) are measured, the usual analysis will be biased if unobserved causes of M are correlated with unobserved causes of Y" (p. 2).

$$x \rightarrow m \rightarrow y$$

Why? There could be a third unmeasured variable, *q*, that predicts both *m* and *y*. Or *m* could be endogenous for other reasons (i.e., measurement errors)

With randomization there is no correlation between *x* and *q* (Antonakis, et al., 2010). Thus measuring or not *q* and including it in a regression model will make no difference to the effect of *x* on *m* or any other DV. That is:

$$m = \gamma_0 + \gamma_1 x + v \qquad \text{Eq. 1}$$

$$m = \delta_0 + \delta_1 x + \delta_2 q + w \qquad \text{Eq. 1.1}$$

Asymptotically: $\delta_1 = \gamma_1$

Although in Eq. 1, *q* is in *v*, *v* is orthogonal to *x*. Even if $\delta_2 \neq 0$, δ_1 is unaffected.

Thus, the OLS (ordinary least squares) or ML (maximum likelihood) estimate of ξ_1 in Eq. 6 will be biased because (from Eq. 5 and 10):

$$\xi_1 = \frac{cov(y,m)}{var(m)} = \eta_1 + \eta_2\kappa_1 \neq \eta_1 \qquad \text{Eq. 11}$$

Recall, the true model is:

$$m = \lambda_0 + \lambda_1 x + \lambda_2 q + \varepsilon \qquad \text{Eq. 3}$$

$$y = \eta_0 + \eta_1 m + \eta_2 q + \theta \qquad \text{Eq. 4}$$

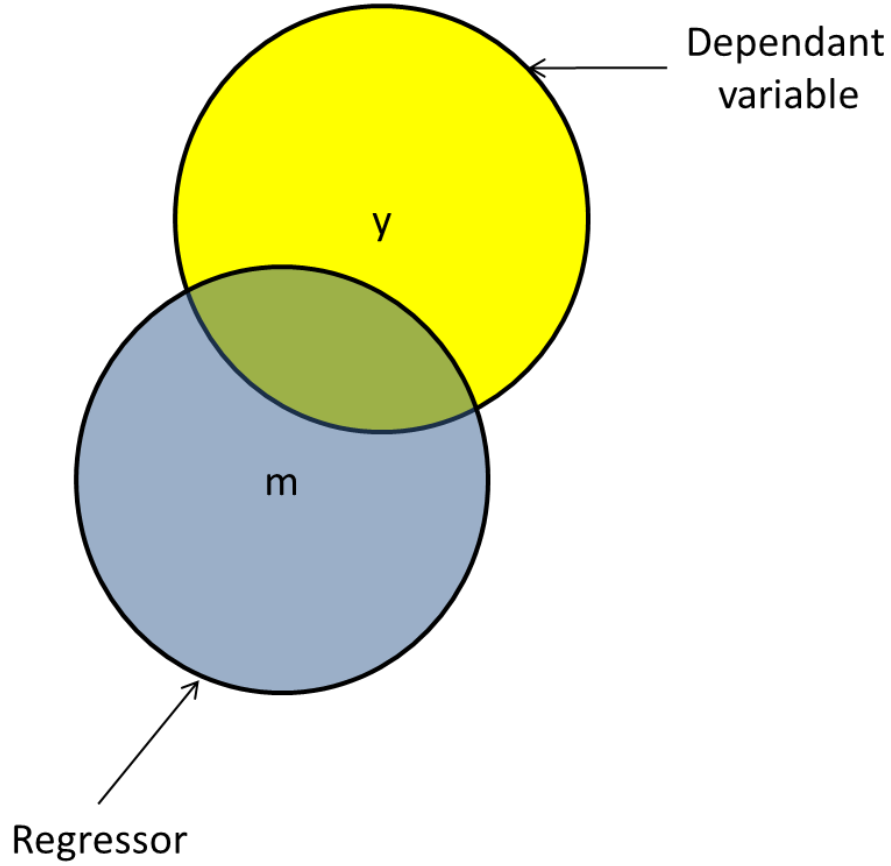
And we estimate:

$$m = \omega_0 + \omega_1 x + u \qquad \text{Eq. 5}$$

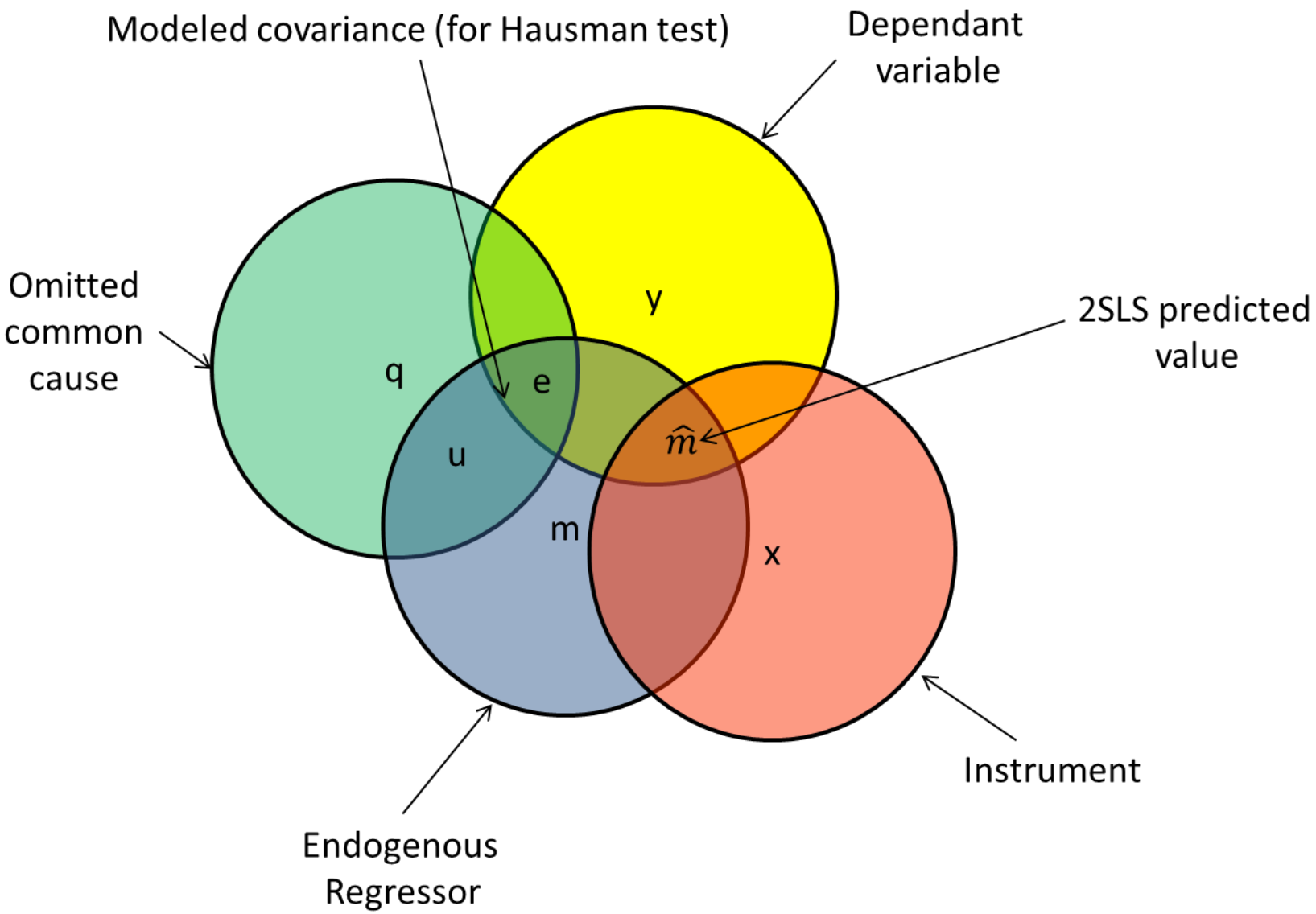
$$y = \xi_0 + \xi_1 m + o \qquad \text{Eq. 6}$$

ξ_1 will only equal η_1 if either η_2 or κ_1 are zero (i.e., if q is unrelated to y or to m).

The solution to the endogeneity problem in mediation.



The overlap in ballentines is the variance that the two variables share.



Voilà: This is the general formula for an instrumental-variable estimate (e.g., see Bollen, 2012):

$$\hat{\eta}_1 = \frac{Cov(y, x)}{Cov(m, x)}$$

Thus, *x*, the instrument, *must* correlate with *y* and *m* if *m* is a true cause of *y* (and if *m* is endogenous)*;* else, the estimate of Eq. 15 is either zero or undefined, which shows up another fallacy in mediation—and a widely believed one too—that *x* need not correlate with *y* to establish mediation (Shrout & Bolger, 2002; Zhao, Lynch, & Chen, 2010); this belief is only true if *m* is exogenous!