Minji Kim, Joseph Cappella

Efficient Versus Accurate Message Testing: Choosing an Optimal Sample Size to Evaluate Message Characteristics

Abstract

Message testing, where raters evaluate messages on various structural and substantive features, is a crucial part of many persuasion studies. In testing, researchers often face a trade-off between accuracy and efficiency specifically with regard to the number of raters per message. In this light, the current study examines how the number of raters affects the accuracy of overall ratings relative to the population. Five data sets evaluating anti-smoking messages were utilized. Bootstrapped samples and populations are compared using various techniques. Results suggest that after 23 to 25 raters per message, increasing the number of raters does not enhance evaluation accuracy appreciably. Nonetheless, variation across messages and tolerance level for misclassification may require more or fewer raters per message.