

EMPIRICAL INVESTIGATION OF DATA QUALITY AND VALIDITY OF COGNITIVE VALENCE AND CONTENT OF THOUGHTS: TRAINED INDEPENDENT RATERS VERSUS TRAINED ORIGINAL RESPONDENTS

Karin Braunsberger, University of South Florida St. Petersburg, United States
Brian Buckler, Avila University, United States
David J. Ortinau, University of South Florida, United States

ABSTRACT

The present study empirically investigates data quality of cognitive responses. Cognitive response researchers using qualitative techniques generally ask selected respondents to write down their cognitions, feelings about, or behaviors toward particular experimental stimuli in an unstructured format (see Wright, 1973). Once these cognitive responses have been produced, a decision has to be made whether the original respondents who generated the thoughts or independent raters are asked to use coding schemas to classify the cognitive output into more interpretive data. Even though the use of independent raters seems to be a general research convention (Coulter, 1998; Meirick, 2002), many studies caution that using independent raters to classify cognitive responses is problematic (Cacioppo and Petty, 1981; Perreault and Leigh, 1989; Wright, 1980).

The present study is the first known attempt to dissect cognitive intent congruence into a cognitive valence component and a cognitive content component to understand differences and similarities between the thought classifications provided by independent raters and original respondents based on post-facto classification schemas. It thus takes the “cognitive intent congruence” issue one step further than previous research because it empirically investigates whether the degree of congruence between the trained respondents who generated the original cognitions and trained independent raters is higher depending on whether the comparisons are restricted to the valence or the content of the cognitions.

A review of the literature suggests that data quality and reliability issues surrounding the use of post-facto coding schemas have been studied by primarily using “inter-rater reliability” measures to determine the consistency of the classification coding results between independent raters. While this reliability measure is very useful in assessing the overall consistency of the coding (classification) results between specific independent raters, reliance on inter-rater reliability measures can lead to a number of problems relating to data quality, reliability, and validity (Braunsberger, Buckler and Ortinau, 2005). Further, inter-rater reliability measures might mask problems with data validity, which is commonly defined as “the degree to which what the researcher was trying to measure was actually measured” (McDaniel and Gates, 2005, 276). That is, measuring whether a group of independent raters agrees on the classification of cognitive responses does not tell us if these raters have correctly interpreted the meaning of the cognitions as intended by the respondents who generated those thoughts.

The study offers 12 research hypotheses, and the two main components of overall cognitive intent validity (valence and content congruencies) between the original respondents and different sets of independent raters (undergraduate, graduate and Ph.D.) are tested under varying conditions of familiarity with the original experimental stimuli (familiar with ad versus unfamiliar with ad) and type of coding (one step versus two step).

The findings reveal wide ranges of cognitive valence congruence among all types of independent raters, providing strong evidence that independent raters have difficulty reproducing the cognitive thought coding patterns of the original respondents. The results also strongly suggest that less complex schemas are better. Our analysis reveals that providing independent raters with the actual experimental ad used by the original respondents does not significantly improve these raters’ ability in matching original respondents’ coding of cognitive content of thoughts or the valence of those thoughts. In contrast, the between-rater average congruence results suggest that familiarity with the experimental ad will be more beneficial to independent Ph.D. raters for coding of the valence and content of the original cognitions than for the other rater groups. Furthermore, familiarity of the experimental ad only marginally improves graduate raters’ cognitive valence congruence more so than undergraduate raters.

In summary, the findings strongly support that independent raters perform much better in coding the original respondents’ cognitive valence of thoughts as compared to the cognitive content of those thoughts; however, the congruence levels between independent raters and original subjects are surprisingly low on both. Thus, whenever possible, original respondents should be trained and allowed to code their own cognitive responses.

References Available on Request