Calibrating Personality Self-Report Scores to Acquaintance Ratings

John A. Johnson
Pennsylvania State University, DuBois

Abstract

By convention in individual personality assessment, scores on self-report questionnaires within ±5 standard deviation of the mean score for that trait are considered "average," whereas scores outside that range are reported as "high" or "low" levels of the trait. To date, no one has examined how well this convention corresponds to perceptions of low, average, or high trait levels by acquaintances. The present study compares the accuracy of the conventional ±5 SD cutoffs for low, average, and high trait levels with cutoffs derived from Optimal Data Analysis (ODA; Yarnold & Soltysik, 2005) for 160 participants who completed the IPIP-NEO (Goldberg, 1999) and were rated by acquaintances on the 30 facets and 5 domains of that instrument. In 32 of 35 comparisons, the ±5 convention was found to be less accurate than the ODA method for deriving cutoffs.

What About Acquaintance Perceptions?

Although defining self-report inventory scores as "low," "average," or "high" according to whether they fall within ±5 SD of the mean makes good statistical sense for labeling a score with respect to others' self-report scores, this does not mean that someone receiving an "average" score on a self-report measure will be perceived as average for that trait by acquaintances who know the person well. It is an empirical question how accurately the ±5 SD cutoffs for a particular self-report measure will correspond to judgments of personality by acquaintances. The current study tested the accuracy of the ±5 SD convention against a method called Optimal Data Analysis (ODA; Yarnold & Soltysik, 2005) for deriving cutoffs for maximizing valid classification.

METHOD

196 undergraduate students agreed to complete Goldberg's (1999) 300-item International Personality Item Pool (IPIP) representation of Costa and McCrae's (1992) NEO-PI R (hereafter IPIP-NEO). Participants provided the names of three knowledgeable acquaintances, who were sent to a Web site containing descriptions of the five domains and 30 facets measured by the IPIP-NEO. Acquaintances were asked to rate participants on 35 scales with the following percentile anchor points: 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 99. From the 196 persons recruited, 160 (59 males, 101 females) actually completed the IPIP-NEO and had at least one acquaintance complete the personality ratings (116 had three raters, 39 had two raters, and 5 had one rater). Acquaintance ratings were averaged for participants with two or more raters to increase reliability. While future analyses may examine five categories of ratings, the current student classified averaged ratings between 30-70 percentile as "average" and ratings outside that range as "low" or "high." Below is a sample portion of the Web rating form.

DISCUSSION

Cutoff scores from Optimal Data Analysis (ODA; Yarnold & Soltysik, 2005) reliably classify acquaintance perceptions of low, average and high levels of traits more accurately than the convention of classifying scores within ±5 SD of the mean as "average" and scores outside that range as "low" or "high." Optimal cutoffs varied considerably across scales, averaging more than 1 SD from the mean. Further analyses are needed to clarify the relative sensitivity, predictive value, effect strengths, and model efficiency for the two methods.

References


