Text-Based Similarity Measures and Attitude Scale Development

John Ford

U.S. Merit Systems Protection Board

Abstract: This research investigates the utility of text-based similarity measures in developing agree/disagree attitude scale questions in three ways. First, text mining of the narrative responses to an open-ended survey question about employee performance identified characteristic issues raised by highly engaged versus less engaged employees. These issues were used to develop new items that increased the reliability of the engagement scale. Second, latent semantic analysis can predict interrelatedness of attitude scale questions before scale administration. This information can inform decisions about which items to retain in the scale for field testing. Finally, inter-item semantic similarity can be a more appropriate, non-zero baseline for comparing correlations between items after scale administration. This allows more sensitive detection of item correlations which differ from text-based item similarity rather than from zero. These techniques have been used in development of attitude scales that measure employee engagement, work motivation, and adherence to merit system principles.