Concreteness, Concretely: A Case Study for Validation in Natural Language

Michael Yeomans

Abstract. Concreteness is central to psychological theories of learning and thinking, in psychology and organizational behavior. However, the literature provides diffuse and competing approaches to measuring concreteness in natural language. In this paper, we develop a concrete definition of concreteness. We first gather datasets from several domains, including written advice, and plan-making (total N = 9,780). We find that many dictionary- and category-based measures fail to consistently detect concreteness in new domains, while in-domain supervised models are quite robust. We generate simple guidelines for automated concreteness detection within and across domains, which are provided in a corresponding R package, *doc2concrete.* Our approach also offers a model for how open science can support more co-operation and theoretical integration in organizational behavior research.