

Determinants of learning difficulty and boundary uncertainty in unidimensional category learning

Gyslain Giguère

The University of Texas at Austin

Bradley C. Love

The University of Texas at Austin

Abstract: Many real-world learning and classification problems involve mastering subtle and noisy distinctions. For example, predicting the winner of a baseball game is challenging due to numerous interactions and inherent noise in outcomes. One common approach in machine learning to reduce overfitting and boost performance (i.e., generalization) on novel test items is to use regularization methods that effectively smooth the training data. One research question is whether human performance can be boosted by effectively regularizing (i.e., idealizing) the training set. In the first two studies, participants learned and were tested on perceptual categories. Idealization was manipulated across conditions in terms of distribution variance and feedback type. In the third study, participants predicted the outcome of actual baseball games after training on either actual or idealized game results. In all three studies, participants performed better on generalization tests (involving actual items) when trained on idealized items than on actual items.