Measuring Category Learning One Step at a Time

Jonathan M. Kwasny (jkwasny@binghamton.edu)
Kenneth J. Kurtz (kkurtz@binghamton.edu)
Department of Psychology, Binghamton University,
PO Box 6000, Binghamton, NY 13902 USA

Traditionally, researchers have studied category acquisition by evaluating whether and how quickly learners reach performance criterion on a classification task and by conducting subsequent transfer tests. With some exceptions, research carried out within this paradigm makes little use of rich sources of trial-by-trial information about the time course of acquisition and response latency. In order to track the acquisition process, we present an experimental paradigm based on the inclusion of additional on-line behavioral measures within each classification learning trial. We believe the approach of investigating the microstructure of category learning will inform theoretical discussion of the representations and processes underlying categorization.

In order to study category acquisition 'as it happens' a set of novel categories was designed for participants to learn. The stimuli were three types of grid patterns derived by introducing a level of distortion to source prototypes never seen by participants. This approach follows from Posner and Keele's (1968) use of dot pattern stimuli. Each stimulus item was a grid pattern consisting of a 4 x 4 matrix composed of individual cells colored either black or white. Category instances were created by applying a level of distortion to the prototype by reversing the color of a number of cells. The prototypes themselves are much like everyday tile patterns with diagonal symmetry and a 50-50 ratio of black-to-white cells.

Several design features bear further emphasis. The studies use a three-way classification task as opposed to the commonly used binary choice to promote the learning of a positively defined set of categories rather than merely learning members versus non-members of a single category. The studies have been designed to maximize naturalism by using large category sets (16 items per category) rather than small sets that easily become a 'toy problem'. In addition the stimuli are designed to be convenient for experimental control, but not overly artificial as clear instantiations of a small set of binary feature values. The grid patterns invite individual interpretation and construction of higher order features.

While recent research has explored alternatives to classification as the basic task during category learning (e.g., Ross, 1997; Yamauchi & Markman, 1998), the present investigation uses standard supervised classification with the addition of a "follow-up" task on each trial and measurement of classification response latency. Goodness ratings provide information about the development of systematic graded structure over the time course of learning. Confidence ratings supplement the classification accuracy data by providing insight into participants' awareness of their own learning progress and emerging category organization.

The overall design of the studies includes a learning phase and a test phase. During learning, in addition to the classification task, participants were also asked to complete either goodness-of-exemplar judgments for correct items in which they rated how good a member they considered each item to be of the selected category or confidence judgments before and after classifying each item about their ability to correctly classify the item. After a fixed number of learning trials, participants were asked to make a set of similarity judgments for sets of same-category and different-category pairs. Participants also completed a written questionnaire including an open-ended category description task and a prototype generation task in which they filled in empty grids to graphically represent their idea of each prototype.

Experiment 1 was designed primarily to evaluate the novel paradigm using grid patterns with minimal distortion from the prototypes (one cell reversed) and a goodness rating for the follow-up task. Experiment 2 used 2-cell distortions, confidence ratings for the follow-up task and added the transfer tests described above. Experiment 3 used 1, 2, and 3-cell distortions with the goodness rating follow-up task. To briefly summarize the findings, classification accuracy increases and response latency decreases across the time course of learning. For initial learning only, confidence is higher after the learning trial. Classification accuracy increases mainly in the beginning portion of each learning pass while goodness and confidence ratings increase steadily. Items with greater distortion were more difficult to classify, but no overall effect on goodness ratings was observed. These findings combined with the results of the similarity judgment, category description, and prototype generation tasks provide useful contributions to our understanding of category learning and representation.

References

Posner, M.I., & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353-363.

Ross, B.H. (1997). The use of categories affects classification. *Journal of Memory and Language, 37,* 240-267.

Yamauchi, T., & Markman, A.B. (1998). Category-learning by inference and classification. *Journal of Memory and Language, 39,* 124-148.