Title
Generalization in Category Learning: One and Two Category Problems

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 19(0)

Authors
Wills, A. J.
McLaren, I. P. L.

Publication Date
1997

Peer reviewed

Generalization in Category Learning: One and Two Category Problems

A. J. Wills (AJW43@CUS.CAM.AC.UK)
I. P. L. McLaren (IPLM2@CUS.CAM.AC.UK)
Psychological Laboratories, University Of Cambridge
Downing St., Cambridge CB2 3EB. United Kingdom

Generalization is one of the most pervasive concepts in the study of cognition. For a very wide range of situations and species it can be shown that if a response is trained to a stimulus then the likelihood of another stimulus also evoking that response is a function of the similarity between them. One of the established facts about generalization gradients, at least within animal learning studies, is that the presence of a second, differentially reinforced, stimulus in training leads to decreased generalization to the target stimulus at test. Prior to our work, no direct study of this effect had been done with normal adults. The purpose of our research was to investigate whether a similar result could be found in human category learning and, if so, to produce an appropriate connectionist account.

Experiments

Our first experiment compared gradients of generalization following one and two category training. Such a comparison raises two methodological problems. First, how may meaningful category membership information be provided in a one category situation? The typical "guess-and-correct" method of two category training is inappropriate for the one category condition as the answer will always be the same. We resolved this by presenting the category label alongside the stimulus in both conditions. The second problem was that of the number of the training trials that should be given in the non-discriminative condition, i.e. does one control for the total number of stimuli seen, or for the number from the relevant category? We address this problem by including both controls in the experiment.

The stimuli used were novel to the subjects, prototype-based, and constructed from 12 symbols positioned on an invisible grid. Subjects studied a number of different, sequentially presented, examples. They were then were asked to classify, as quickly as possible, other examples which ranged from highly typical of one category to highly typical of the other (sometimes unseen) category. Two category training led to a sharper generalization gradient than one category training, and the gradients in the one-category conditions did not differ from each other. The pattern of reaction times was also affected by the presence of a contrast category.

The typical explanation of this result in the animal learning literature is that the presentation of the contrast stimulus somehow neutralizes the effect "incidental" components of the stimulus or the context. Incidental components are those that do not change systematically across the range in which generalization is tested. It is shown that a simple feedforward error-correcting network (of the same type as used in Gluck & Bower, 1988) can correctly predict the results of the first experiment if one assumes that the stimulus representation does indeed include incidental components.

This type of explanation of the result leads to a further prediction to the extent that there is a difference between one and two category training there should be a comparable difference between blocked and intermixed two category training. Specifically, if all examples of one category are presented before any examples of the other, the resulting generalization gradient should be shallower than if there is no systematic ordering. The second experiment directly compared one category, blocked two category and intermixed two category training. Although a one vs. two category difference is again seen, no difference is observed between blocked and intermixed training. The incidental stimulus hypothesis was therefore disconfirmed, and we forwarded an alternative explanation.

Modeling

The basis of our model is that the decision processes engaged following two category training are different to those engaged following one category training. After two category training the decision is relative whilst after one category training it is not. The model takes Gluck & Bower (1988) as its starting point, but adds a threshold unit and a *winner-take-all* network which together implement the hypothesized decisional processes. It predicts, in some detail, the choice probabilities and the reaction times found in our experiments.

Acknowledgements

This research was supported by a grant from the BBSRC (UK) to I.P.L. McLaren. The authors would like to thank N. J. Mackintosh and J. K. Kruschke for their helpful comments.

Reference

Gluck, M. A. & Bower, G. H. (1988) From Conditioning to Category Learning: An Adaptive Network Model. *Journal of Experimental Psychology: General*, 117, 3, 227 247