Title
Transfer Along a Continuum: Differentiation or Association?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

Authors
McLaren, I.P.L.
Suret, M.

Publication Date
2000

Peer reviewed

Transfer Along a Continuum: Differentiation or Association?

I.P.L. McLaren (iplm2@cus.cam.ac.uk)
Department of Experimental Psychology; Downing Street
Cambridge, CB2 3EB UK

M. Suret (mbs22@cam.ac.uk)
Department of Experimental Psychology; Downing Street
Cambridge, CB2 3EB UK

Abstract

There has been a revival of interest in the question of the optimal training schedule for a difficult discrimination. McClelland (personal communication) argues that the optimal schedule is one which starts with a much easier discrimination on the same dimension as the difficult one, arranged so that the easy problem can be made to gradually converge on the more difficult one. He further argues, in agreement with Saksida (1999), that the reason for this is that representations are more easily formed during acquisition of the easy problem, which can then be put to use in solving the difficult discrimination. As associative learning theorists we are more familiar with another account - that initiated by Lawrence (1952) - which agrees that the optimal schedule is one which employs a strategy of transfer along a continuum. Where the accounts differ is in the mechanism for transfer; rather than appealing to representation formation, this account explains the benefits of training on the easier problem in terms of the dimensional features / elements that acquire associative strength and their ability to generalize appropriately to the hard problem. In this paper we report experiments that attempt to distinguish between these two accounts by manipulating / deconfounding stimulus exposure and training. We demonstrate the basic effect, and show that pre-exposure to the stimuli that comprise the easy problem is less effective than pre-exposure to the stimuli that make up the more difficult discrimination. Our conclusion is that this latter result is not what one would predict from the non-associative account given above, but that it fits well with McLaren, Kaye, and Mackintosh's model of perceptual and associative learning.

Introduction

Lawrence (1952) demonstrated that it was possible for training on an easy perceptual discrimination to transfer to a more difficult problem on the same dimension. An example of the type of problem that he studied would be a brightness discrimination between two rectangles in similar shades of gray (hard version) or black and white rectangles (easy version). Groups of pigeons trained on these problems for the same number of trials can reach a point where the group trained on the easy problem have solved it, i.e. they have learned to peck at one rectangle for grain and to ignore the other, whereas the group trained on the hard problem have made little progress. If both groups are now trained on the hard problem for a further number of trials, i.e. the group previously trained on the easy version of the problem are now switched to the harder version, then the result of interest is that the group switched from easy to hard acquires the hard problem much more rapidly than the group trained on the hard problem from the outset. This result holds despite the fact that the total amount of training is the same for both groups, and that the group that acquires the problem more slowly is the one that has received more training on that specific problem. This is the phenomenon of transfer along a continuum (TAC), and is the subject of the research reported in this paper.

The standard associative account of this phenomenon appeals to the notion of generalization. The stimuli for the easy problem become associated with reward and non-reward respectively, and then generalize to the stimuli for the hard problem (e.g. see Mackintosh, 1983). This is more effective than training on the hard problem itself because it is so difficult to learn, which is taken to be because the stimuli are so similar to one another. Figure 1 can be used to illustrate one possible instantiation of this explanation. On this approach, a stimulus is represented by a set of activated elements or units, a distributed representation. Variation along a stimulus dimension such as brightness will, for the most part, be represented by different elements corresponding to different values on the dimension, rather than the activation level of an individual element being the primary indicator of value on the dimension (c.f. chapter by Thompson in Mostofsky, 1965). Each element has a 'tuning curve' such that it responds most strongly to a certain value on the dimension and this response drops off fairly rapidly with 'distance' from this optimal value. Note that many elements will be active when any stimulus with value on that dimension is present, the coding is via a pattern of activation. Learning will proceed via association between the elements activated by a stimulus and other units representing reward. We are now in a position to explain Lawrence's results. In the case of the easy problem (shown top in the figure) the stimuli are well separated on the dimension and there is relatively little overlap between the patterns of activation that represent them. Learning proceeds rapidly, favoring those elements which are most active on a trial, and there is little generalization between stimuli to slow acquisition of the problem. In the case of the hard problem (shown bottom in the figure)

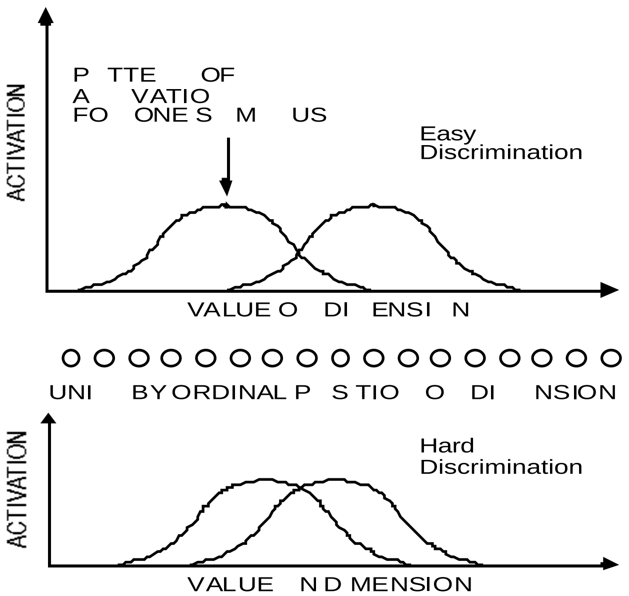

Figure 1: Stimulus representation on a dimension for easy (top) and hard (bottom) discriminations.

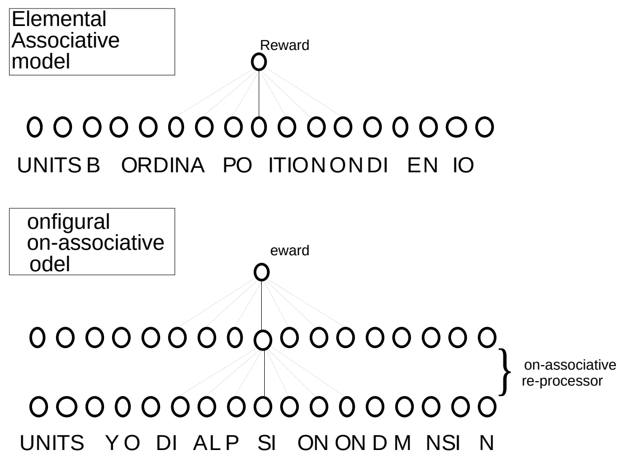

Figure 2: Architecture for associative and non-associative models of TAC.

the situation is somewhat different, in that the large degree of overlap between the stimulus representations results in considerable generalization between the stimuli, and this is what makes the discrimination difficult. The elements that are most active, and so dominate learning, are not those that best discriminate between the stimuli. As a result acquisition is slow. Now consider the case where the easy problem is first acquired, and then the subjects are switched on to the hard problem. The training on the easy problem will result in exactly the elements that are the most predictive of reward or non-reward in the hard problem gaining considerable associative strength because they are highly activated by the easy stimuli. Thus the learning will transfer well to the hard problem, and will be more than an equivalent amount of training on the hard problem would have provided (because of the relatively large activations of the elements concerned).

There is another tradition in psychology, however, that appeals to quite different, non-associative processes to explain the phenomenon of transfer along a continuum. It can be traced back at least as far as the work of Eleanor Gibson (1969), who conjectured that a process of differentiation, contingent on exposure to the stimuli in question, resulted in representations of the stimuli that better enabled discrimination between them. Gibson's thesis is perhaps most naturally captured in terms of competitive learning coding schemes that require no explicit instruction to develop representations that capture the structure of a stimulus set that they are exposed to. Our example of such a system is that due to Saksida (1999), which is explicitly designed to deal with phenomena of the kind under consideration here. Figure 2 allows us to contrast Saksida's model with the standard associative account. Instead of stimulus elements being directly associated with reward representations (shown top), there is a non-associative pre-processor prior to association to reward representations (shown bottom in the figure). The model develops a representation of the input at this intermediate, competitive layer, and, in Saksida's model, it does so in a way that drives initially overlapping representations for two stimuli apart so that they become more discriminable (i.e. differentiation). The interesting possibility raised by such a model is that the explanation for TAC might be quite different to that generated by the associative account given earlier. Instead of appealing to generalization of associative strength from the easy problem to the hard version, it could be that training on the relatively easy problem could develop a coding at the competitive layer that meant that the hard problem was no longer as difficult as would have been the case. Whereas before the hard problem would have (initially at least) given rise to highly overlapping patterns of activation at the competitive layer, now the patterns of activation are better separated because of the coding scheme developed for the dimension whilst solving the easier problem. In a sense what happens is that the process of developing discriminable representations for the easier problem drags apart the representations for the harder problem as well. Saksida herself is quite definite on this..."One clear prediction of the current model is that exposure to a pair of similar stimuli will facilitate discrimination of stimuli that are even more similar along the same dimension." and..."pre-exposure to two stimuli will facilitate discrimination of other stimuli whose representations fall between them on the competitive layer" Saksida (1999). The only provisos being that the easier discrimination should itself employ relatively similar stimuli, and should be studied long enough for the stimuli to become discriminable (i.e. for the competitive layer to develop the necessary representations.

The strategy adopted in this paper is to contrast these two accounts with specific reference to the issue of whether or not TAC is best characterized as due to elementally-based generalization or rather to perceptual

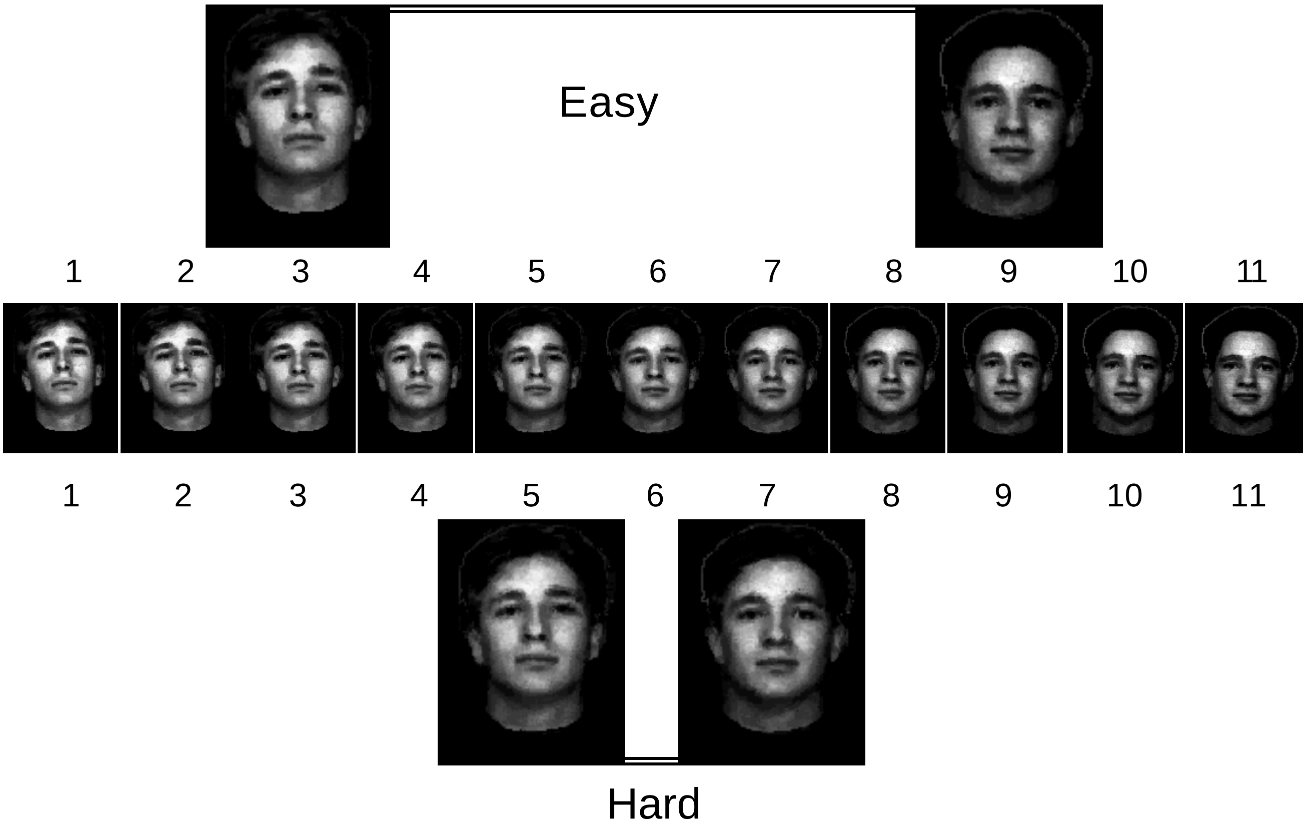

Figure 3: One of the four morphed face dimensions used in these experiments.

differentiation as a result of representation formation. Experiment 1 demonstrates TAC using stimuli that meet Saksida's criteria for her models application. Experiment 2 then assesses whether the effect can be explained predominantly in terms of representation development by looking at the effects of pre-exposure to the stimuli used in the easy and hard problems. The logic here is that if TAC is mainly due to representation development, then the training phase is effectively equivalent to a pre-exposure phase, and so explicit pre-exposure should generate the same pattern of results.

Experiment 1

Stimuli and Apparatus

In all phases of the experiment, pictures of faces were shown in the form of gray-scale images. These had been created from standard passport photographs of university undergraduates, which had been scanned into the computer. These stimuli were presented on an Apple Macintosh computer running Microsoft Basic. They were 3.5 cm by 4.5 cm and subjects sat approximately 50 cm from the screen. The face stimuli for this experiment were constructed by taking pairs of faces and morphing from one to the other in 10 equal steps, giving a dimension with 11 values in all. The faces in a given pair are chosen to be similar (which aids the morphing process in keeping the transitions smooth) so that neighboring stimuli on the dimension are very similar indeed. Figure 3 illustrates the morphed face dimension for one pair of faces, there were four pairs of faces in total and the faces at 3 and 9 on the dimension always constituted the easy problem and those at 5 and 7 the hard version. All four dimensions were used concurrently for every subject, with the assignment of the face dimensions to the conditions of the experiment counterbalanced appropriately. Pilot testing revealed that the discriminations were difficult (even for the 3 vs. 9 case) but possible under the conditions of this experiment, and subjects reported that their performance was hard to characterize in terms of rules based on features (desirable if performance is to be associatively driven).

Subjects and Design

Subjects were 40 Cambridge undergraduates and graduates with an age range of 18 - 30. They were randomly assigned to two equal groups, one of which (Group Easy) was pre-trained on the easy problem for all four dimensions concurrently for a fixed number of trials (40 trials in total, five for each face), the other (Group Hard) was pre-trained on the hard version for all four dimensions for an equal number of trials. After the pre-training phase both groups were then trained on the hard problem for all four dimensions concurrently (again 40 trials in total, five for each individual face). This was followed by a final test phase in which performance on the hard problem for each dimension was assessed without giving the feedback used in pre-training and training. In this phase each face is also shown 5 times. The data of interest are the responses to the stimuli in this final test phase. If the discrimination between 5 and 7 is better learnt after pre-training on 3 and 9 then this would be evidence of TAC.

Procedure

In both the pre-training and training phases of the experiment, subjects were told that once they pressed the space bar, a constant stream of stimuli in the form of faces would appear on the screen, and that their task was to sort these stimuli into two categories. They were to do this by pressing one of two keys ('x' on the left or '.' on the right) and would receive immediate feedback as to the correctness of the response. If they did not respond within a few seconds (4.25 sec) they would be timed out. The subjects were told that the faces were randomly and equally allocated to either left or right key and that their task was to simply find out and remember which ones were 'right' and which ones were 'left'. Once the subject initiated the experiment, trials were continuous. Stimuli were presented singly, and each trial started with a '+' for 0.7 sec which was then replaced by a rectangular frame for 0.2 sec. Each face appeared and stayed in the screen for a maximum of 4.25 sec and disappeared once a response or time-out was made. Feedback was then given for 1 sec, either 'correct' displayed in the center of the screen or 'error' and a beep if the wrong key was chosen.

After they had completed the pre-training and training phases, subjects progressed to the test phase of the experiment. Subjects were told to categorize the stimuli into the two categories based on the judgments they had made in the training phase. So, if a face had been 'a left key stimulus' in the training phase, it was to be allocated again to the 'left key' category in the testing phase. This time no feedback was given. The procedure of stimulus presentation was as before with the exception that feedback was replaced in this phase by a 1 sec pause between the subject's response and the proceeding stimulus.

Results

The results of Experiment 1 are shown in Figure 4. One key, e.g. the left key, is designated the negative category (a press scores -0.5 for that stimulus) and the other right key the positive category (scores +0.5) during test.

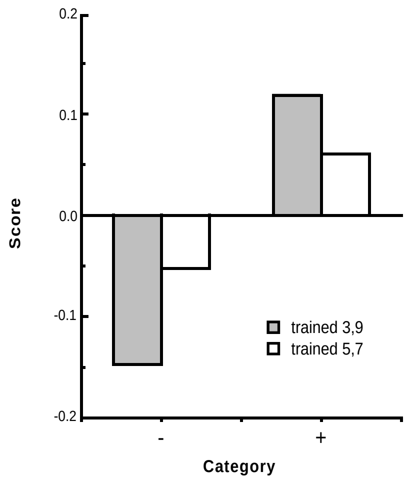

Figure 4: Results for Experiment 1.

Key assignments were counterbalanced across subjects so that the positive category has equal numbers of left and right key responses (at least by design). The test score indicates the average of the key presses across subjects, and would be zero if subjects were indifferent to which stimulus went with a given key, and ranges from +0.5 to -0.5. The group pre-trained on the easy problem (3,9) shows much better performance on the hard discrimination than the group pre-trained on that discrimination (5,7) itself. That is, the 3,9 group has more positive scores for its positive stimulus on test, and more negative scores for its negative stimulus.

These impressions are borne out by statistical analysis, in which all probabilities are two-tail unless otherwise specified. ANOVA on the results with a between subjects factor of type of pre-training (3,9, vs. 5,7) and a within subjects factor of type of stimulus (- vs. +) gave an $F(1,38) = 27.75$, $p<.001$ for the main effect of type of stimulus and $F(1,38) = 4.39$, $p<.05$ for the interaction between the two factors. The first effect refers to the fact that the positive stimulus is, overall, given a more positive score than its negative counterpart, the interaction reveals that the difference in score between positive and negative stimuli was significantly greater for the 3,9 group who were pre-trained on the easy problem. This demonstrates transfer along a continuum with these stimuli. Planned comparisons on the positive and negative stimuli for each group separately reveal that both groups are significantly better than chance on the test discrimination, $F(1,19) = 24.9$ and 5.52, both $p<.05$. Thus both groups can be said to have learned the discrimination.

Discussion

Experiment 1 provides a convincing demonstration of transfer along a continuum in human subjects using an artificial dimension constructed by morphing between similar faces. Performance on the hard problem after pre-training on the easy problem is much better than if pre-training had been on the hard problem used during training and test. Nevertheless, both groups were able to acquire the discrimination under the conditions of the experiment.

We are now in a position to ask if this TAC effect is simply due to exposure to the stimuli used in the easy problem, or if instead it requires that subjects be trained on the easy problem for the effect to occur. Experiment 2 seeks to answer this question by pre-exposing subjects to the stimuli of either the easy or hard problem instead of pre-training them.

Experiment 2

In this experiment the stimuli are the same as in Experiment 1, and two new groups of 20 subjects from the same population are assigned to two different pre-exposure conditions. These are equivalent to the pre-training conditions of Experiment 1 except that a) no response is required as it is pre-exposure and b) each stimulus is shown for a fixed duration of 2 sec. This duration was chosen to ensure that subjects in this experiment received the same or greater total time of exposure to the stimuli compared to all the subjects in Experiment 1. Thus subjects were pre-exposed to the stimuli that constituted either the easy or the hard problem, then trained on the hard problem exactly as in Experiment 1, then tested exactly as in Experiment 1. If the results of Experiment 1 were predominantly due to exposure to the stimuli of a given problem, then the results of this experiment should resemble those of the previous experiment. If, on the other hand, they were strongly dependent on the training element during pre-training then we might expect the results to differ in that evidence for any TAC effect should disappear.

Results

The results of Experiment 2 are shown in Figure 5.

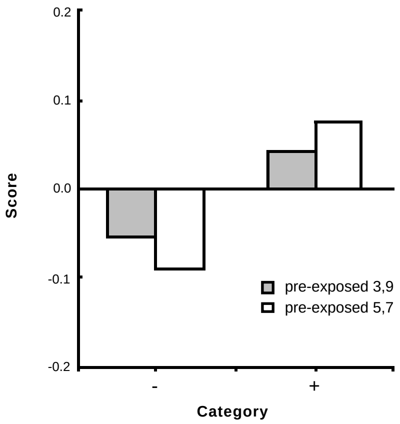

Figure 5: The results of Experiment 2.

Once again an ANOVA with one between factor of type of pre-exposure (3,9, vs. 5,7) and one within factor of stimulus type (5 vs. 7) was conducted which gave an overall main effect of stimulus type, $F(1,38) = 9.84$, $p<.005$, but no significant interaction between the two factors ($F<1$). Thus there is good evidence for acquisition of the discrimination, but no significant evidence that pre-exposure to either the easy or hard problem stimuli had any differential effect. Contrary to expectations on a differentiation account of TAC, the group pre-exposed to the hard problem stimuli was actually numerically better on test. Planned comparisons on the two groups revealed that the group pre-exposed to the hard problem was significantly better than chance on test, $F(1,19) = 7.26$, $p<.05$, whilst the other group was only marginal, $F(1,19) = 2.89$, $p(1\text{-tail}) = .052$.

As the two experiments are highly comparable in their stimuli, apparatus, procedures and subject populations we can compare them in a single analysis. When this is done there is a three way interaction ((pre-trained vs. pre-exposed) x (problem, 3,9 vs. 5,7) x (stimulus on test, 5 vs. 7) that indicates that the effect of pre-exposure in Experiment 2 is significantly different to the effect of pre-training in Experiment 1. Finally, pre-exposure to the easy problem in Experiment 2 was significantly less effective than pre-training on the easy problem in Experiment 1 $F(1,38) = 4.83$, $p<.05$. This is despite the fact that pre-exposure in Experiment 2 was at the maximum level observed in Experiment 1 (where the fluctuations were due to different speeds of response during pre-training).

Discussion

The results of Experiment 2, taken in conjunction with those of Experiment 1, do not support a differentiation account of transfer along a continuum. The effect of pre-exposure to the problem stimuli is seen to produce the converse pattern of results to pre-training, that is, with regard to learning the hard problem during the training phase, pre-exposure to the easy problem is less effective than pre-exposure to the hard problem, whereas pre-training on the easy problem is much more effective than pre-training on the hard problem. This pattern of results strongly suggests that the advantage that accrues as a result of pre-training on the easy problem is due to generalization of the associations acquired during that pre-training.

One loose end in this experiment concerns the extent to which pre-exposure can be said to have an effect at all, given that the two groups do not differ significantly. Some light can be cast on this issue by considering the data from a previous series of experiments (McLaren, 1997) which used the same procedures and stimuli, but merely trained the face discrimination (as here) but without any pre-training or pre-exposure. Under these circumstances the hard problem was not solved (mean difference between the positive and negative stimuli was only .037, $F<1$), and a comparison between these results and those of Experiment 2 reveals that the group pre-exposed to the hard problem is better than the group simply trained on the hard problem, $F(1,43) = 2.64$, $p(\text{one-tail}) = 0.055$). This means that pre-exposure to the hard problem has had a near significant beneficial effect (i.e. we have some evidence for perceptual learning), though this is not true of pre-exposure to the easy problem ($F<1$).

General Discussion

In this paper we have contrasted two classes of model of TAC, associative and non-associative versions. We should make it clear that while we have found no evidence that supports the non-associative account relative to the associative version, nor do our results falsify the non-associative position adopted by Saksida and others. What is needed to rescue this account is a parametrisation of the non-associative model that allows the generalization from pre-training to dominate any effects of representation formation and differentiation. In these circumstances the two types of models would in some sense be different instantiations of the same psychological theory.

Taking Saksida's account first: the perceptual learning effect seen in Experiment 2 would be due to representation formation and differentiation, as the initially overlapping patterns of activation for the two to-be-discriminated stimuli were pushed apart by competitive learning, and the TAC effect would be due to generalization associations to reward and non-reward formed during pre-training to training and testing. The model would be constrained so that the latter effect would be stronger than the former, which is not a typical feature of this class of model. The more natural account would attribute the difficulty of the hard problem to the need to establish well differentiated representations of the stimuli, a process that was aided by training on the easy problem. It may well be that the need both to model TAC and perceptual learning within this class of model may impose unsustainable constraints on its ability to function effectively, but this is a question for future research.

The associative account offered here follows McLaren *et al's* (1989) theory of association and representation. The explanation of TAC is the standard associative account given earlier in this paper, but the explanation of the perceptual learning effect seen in Experiment 2 may bear further exposition. On this theory, exposure to two similar stimuli that will be represented as overlapping patterns of activation results in a decrease in the salience (in this case this can be understood as the degree of activation) of the elements representing those stimuli. This occurs to the extent that they become predicted by associations from other elements. The reduction in salience will be greatest for the elements shared by the two stimuli (the overlap) because they are encountered, and hence engage in learning, twice as often as the elements unique to either stimulus. The effect is that the elements that make the discrimination difficult (because they are shared by the stimuli and lead to generalization between them) become relatively less salient than those that enable discrimination between the stimuli. This consequence of pre-exposure leads to the discrimination between the stimuli becoming easier, as the distinctive features (represented by the unique elements) of each stimulus are now able to preferentially engage in learning. The result is perceptual learning, in that the discrimination is learned faster after pre-exposure. The effect is predicted to be greater for more similar stimuli, which fits well with the greater pre-exposure benefit for the harder problem. This is because the more similar stimuli are taken to have a higher proportion of shared elements, and so the effect of reducing these elements' salience relative to the unique, distinctive elements of the stimuli is proportionately greater.

Conclusion

Associative theories of representation development and learning are adeqequate to model the transfer along a continuum effect reported here. Non-associative theories that appeal to competitive learning or some other mechanism for representation formation are probably able to instantiate the same psychological theory, but offer nothing new in modelling these data. The challenge is to find data that require this type of theory rather than an associative account.

Acknowledgments

This research was supported by an ESRC grant to IPL McLaren

References

Gibson, EJ (1969). *Principles of perceptual learning and development.* New York: Appleton-Century Crofts.

Lawrence, D.H. 1952. The transfer of a discrimination along a continuum. *Journal of Comparative and Physiological Psychology*, 45, 511-516.

McLaren, I.P.L. 1997. Peak-Shift on a Novel Stimulus Dimension. Paper given to the Conference on Associative Learning at Gregynog, Wales, UK.

McLaren, I.P.L., Kaye, H. and Mackintosh, N.J. 1989. An associative theory of the representation of stimuli: applications to perceptual learning and latent inhibition. In R.G.M. Morris (Ed.) *Parallel Distributed Processing - Implications for Psychology and Neurobiology*. Oxford. OUP.

Mackintosh, NJ (1983). *Conditioning and associative learning.* Oxford: Oxford University Press.

Saksida, LM. 1999. Effects of similarity and experience on discrimination learning: A nonassociative connectionist model of perceptual learning. *Journal of Experimental Psychology: Animal Behavior Processes,* 25, 308-23.

Thompson, RF (1965). The neural basis of stimulus generalization. In DI Mostofsky (Ed.). *Stimulus Generalization,* pp.154-178. Stanford: Stanford University Press.