Attention and Perceptual Learning

I.P.L. McLaren[1] , A.J. Wills[1] and S. Graham[2]
[1]University of Exeter and [2]National University of Singapore

Perceptual learning as a phenomenon.

We are all of us experts. Each one of us is able to effortlessly distinguish between a large number of people that we encounter in our everyday lives. When considered as stimuli in the abstract this is a difficult discrimination task as the stimuli are all rather similar to one another, yet we can identify an individual in a moment with an accuracy that would put a supercomputer to shame. We will argue that this particular expertise comes about because we are very experienced with the class of stimuli in question (in this case faces), and that it is simply a more commonplace example of a type of expertise that we can study in twitchers (experts in the identification of birds), field botanists (experts in plant identification in a particular habitat) and dog show judges (experts in fine discriminations betweens dogs drawn from the same class such as gundogs; as reported in Diamond and Carey, 1986). We can also manipulate experience with stimuli in the laboratory and show that an enhanced ability to discriminate between stimuli is a consequence of the right kind of experience with the right kind of stimuli (McLaren, Leevers and Mackintosh 1994; McLaren, 1997;Lavis and Mitchell, 2006; Mundy, Honey and Dwyer, 2007). This phenomenon goes by the name of perceptual learning, and many peoples first reaction to it is to make the claim that experience helps one learn what aspects of the stimulus to attend to. Perceptual learning, in its broadest sense, can be understood to be any enhancement of learning to a stimulus as a consequence of experience with that stimulus. Thus the proposal would be that perceptual learning is a consequence of the allocation of attention based on past experience with a particular stimulus type. This assertion is a natural (and plausible) response to the data, but is it the correct line to take? The question addressed in this chapter is whether perceptual learning phenomena can be explained by appealing to mechanisms that control attention (we will conclude that in some cases they can) or whether there is a need to appeal to other mechanisms as well (we hope to show that there is!). Starting with some examples of what might be considered perceptual learning phenomena taken from our early work, we go on to discuss them in the context of current research on perceptual learning, with an emphasis on analysing the possible role that attention might play in producing these effects. We will focus on perceptual learning in humans, but will allow ourselves to appeal to evidence from infra-human studies where appropriate.

Enhanced acquisition of a discrimination after stimulus exposure.

The general idea of the experiments reported here was to have participants learn discriminations in much the same circumstances as an animal (such as the rat) might. In the test or discrimination phase, a pair of stimuli would be presented side by side on a VDU with the arbitrarily designated 'positive' stimulus equally often on the left or right. Participants were allowed to choose one of the stimuli by means of a keypress, and received immediate feedback as to the correctness of the choice. One problem with perceptual learning experiments is the need to pre-expose participants to stimuli in such a way as to guarantee that they attend to the stimuli, but receive no information specific to the discrimination problem presented after the pre-exposure phase (otherwise the procedure would be one of pre-training rather than pre-exposure). The first two experiments we will consider exploited the fact that we were carrying out a series of experiments on the delayed matching of visual stimuli in which participants were exposed to 'pattern' stimuli and 'masks' over a substantial number of trials. The pattern stimuli were the experimental items in the delayed matching task, participants had to decide if a probe was the same as, or different from, a prime presented a few seconds before the probe. The masks intervened between prime and probe. Neither method of exposure to these stimuli relates to the discrimination learning task used in the experiments reported in this chapter, and so we exploited this fact by 'piggy-backing' perceptual learning experiments onto the delayed matching sessions.

The stimuli used in Experiment 1 were quasi-random dot patterns generated with certain constraints in operation to favour the formation of interesting, but difficult to describe, blobs or clumps of dots. Examples are shown in the figure below. This type of stimulus was used because it had little or no linguistic component, but was nevertheless complex and detailed. Each subject saw four different pairs of stimuli in this experiment, two of the pairs had been extensively pre-exposed (exposure to each stimulus was more than 96 sec. and less than 152 sec.), whilst the other two (control) pairs were taken from the stimuli another subject was pre-exposed to and so were novel for this subject.

Figure 1: Examples of the dot pattern stimuli used in Experiment 1.

As previously indicated, in the test or discrimination phase, a pair of stimuli would be presented side by side on a VDU with the arbitrarily designated 'positive' stimulus equally often on the left or right. Participants were allowed to choose one of the stimuli by means of a keypress, and received immediate feedback as to the correctness of the choice. The feedback given was either "correct" displayed in the centre of the screen or "error" and a beep if the wrong stimulus was chosen. It was emphasised to the participants that the designation of one member of a stimulus pair as "correct" was arbitrary, that this designation was consistent within the experiment, and that members of the pair would always be presented together but with either member equally often to the left or right of the other. Summary feedback on the number of errors and the mean correct reaction time was given at the end of each block. The results for the accuracy measure (our main dependent variable) for the pre-exposed and nonpre-exposed discriminations are shown in Figure 2 below. Statistical analysis confirmed that the effect of pre-exposure was significant, $F(1,18) = 8.26$, $p < .01$, with performance on the pre-exposed stimuli superior to controls. Latencies were in line with the error data, with time to respond significantly longer for the novel stimuli.

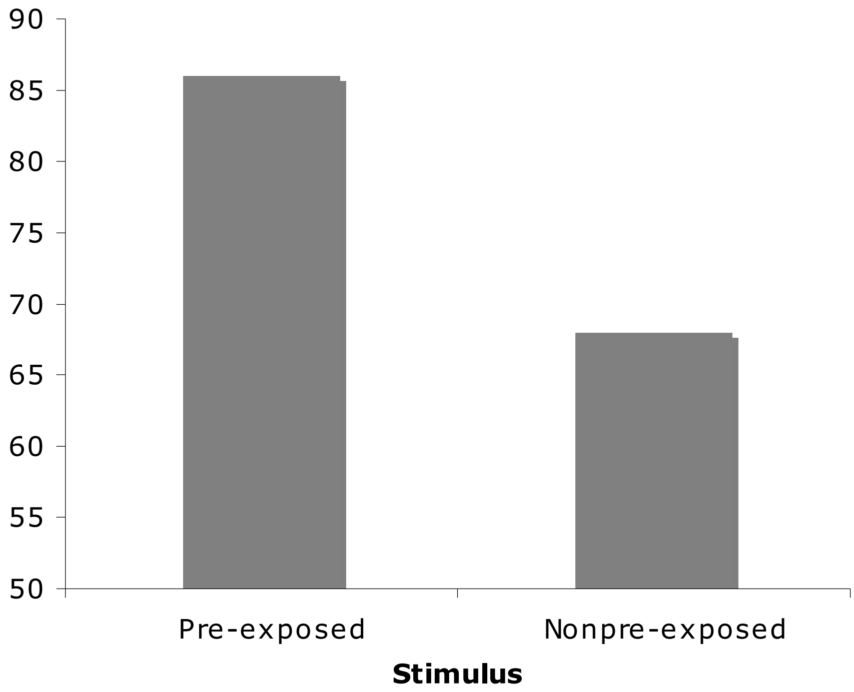

Figure 2: Results of Experiment 1. Chance performance = 50%.

What have we learned from this simple experiment? It is possible to conclude that experience with the stimuli used here facilitated the subsequent acquisition of a discrimination involving those stimuli. We should acknowledge at once that the type of pre-exposure to the stimuli given here is not that most commonly employed in typical perceptual learning experiments nowadays. The use of a paradigm that would have encouraged participants participating in the experiment to try and differentiate between the stimuli (a prerequisite to successfully make the same / different judgement that was required of them on each trial) would be considered suspect in that it contains an element of pre-training on the discrimination that follows. Nevertheless, the learning enhancement effect demonstrated in this experiment was, to our knowledge, the first of its kind at the time that it was conducted. The more typical result in previous research had been no effect of pre-exposure with adults and older children unless special procedures were used. Lubow, Caspy and Schnur (1982) reported no effect on discrimination learning with simple pre-exposure, and Lubow, Alek and Arzy (1975) found no effect of pre-exposure on simple learning. Pre-exposure in children resulted in a retardation in learning, a phenomenon termed the stimulus familiarisation effect by Cantor (1969) who reviewed a number of studies on this point (see also Kaniel and Lubow, 1986). The same result (i.e. a retardation in acquisition of the discrimination) can be obtained in adults or older children if pre-exposure takes place using an 'incidental learning' procedure (Lubow *et al*, 1982; Ginton, Urca and Lubow, 1975), which involved pre-exposing the stimuli as an irrelevant component of some other task. The only experiment to have produced an enhancement in discrimination learning after stimulus pre-exposure was one by Lubow, Rifkin and Alek (1976) with five year olds, where enhanced discrimination occurred after a change of context. Even in this study, if discrimination learning took place in the same context as pre-exposure (as it does here) then acquisition was significantly retarded compared to controls.

The results taken from the literature on stimulus pre-exposure at the time, then, suggested that first of all, the enhancement in discrimination learning observed here was unusual, and secondly, that if stimuli had instead been pre-exposed using an incidental learning procedure then a retardation in discrimination learning might have been expected. We shall defer any explanation of these results until we have presented the next experiment in which a set of eight stimuli unique to each subject were used as masks in the delayed matching task. This experiment naturally follows on from Experiment 1 and will enable us to make contact with the literature on incidental pre-exposure. The stimuli used in this experiment were somewhat different from the dot patterns used in Experiment 1, as they had to be easily discriminable from them to avoid participants erroneously responding to them, and played no role in the experiment other than that of masking the dot pattern stimuli.

Retarded acquisition of a discrimination after stimulus exposure.

The design of this experiment was exactly as for Experiment 1, as were the procedures used in it. Twenty participants participated. The only difference was in the stimuli used and the method of pre-exposure. Examples of the stimuli used in Experiment 2 are shown in Figure 3 below.

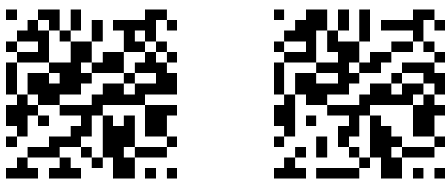

Figure 3: Examples of the chequerboard stimuli used in Experiment 2.

These 16 x 16 (where each square is 4 x 4 pixels) chequerboard type stimuli were generated as follows. Stimuli were generated at random for each subject by making the probability that a square would be black .5, then two more were generated by changing a randomly designated square (i.e. one of the sixteen) in each of the sixteen rows. This gave pairs of very similar stimuli (as shown in the example) for each subject. Control stimuli were taken from those used as experimental stimuli by another subject. Pre-exposure was accomplished by using these stimuli as masks in the delayed matching reaction time task described earlier. The masks were quite easily discriminated from the dot patterns, having a coarser grain to their features, and there was every reason for the subject to fixate on the masks because the test pattern on that delayed matching trial would soon appear in the same place. Total exposure for each mask was 72 sec over 72 separate presentations, half of which occurred in the same session as this experiment and just prior to it, with the other half taking place in an earlier session on the previous day.

The results of this study were quite different from those of Experiment 1 and are shown in Figure 4. Statistical analysis confirmed that the effect of pre-exposure was significant with the pre-exposed stimuli discriminated worse than the novel pair, $F(1,19) = 7.66$, $p < .025$. Reaction times showed a similar pattern. Thus pre-exposure now resulted in slower learning. Why is there this difference in kind between the results of our two experiments? Clearly the type of pre-exposure given is a key factor. We have already discussed how the pre-exposure regime in Experiment 1 provided an incentive for participants to learn to tell the stimuli in that experiment apart. The pre-exposure technique in Experiment 2 does no such thing, instead its incidental nature would, if anything, provide some incentive for participants to ignore these stimuli as they focussed on the ones that mattered (the dot patterns) during that phase of the experiment. Thus one general class of explanation of these results can be identified as being attentional in nature. It proposes that participants learn to attend, or not to attend, to the stimuli according to the type of pre-exposure given. Effectively, this raises or lowers stimulus salience, thus enhancing or retarding learning to these stimuli. Detailed models that fall into this class would be the conditioned attention theory of Lubow, Weiner and Schnur (1981), later refined by Lubow (1989), and the models dealing with negative priming introduced by Tipper (1985) and Tipper and Cranston (1985). Although the processes involved in negative priming were once thought to operate on a much shorter time scale, we now know that this does not have to be the case (Treismann and DeSchepper, 1996; Grison,Tipper, and Hewitt, 2005). If we adopt a Lubowian terminology for the moment, the general idea here is that stimuli can have attention conditioned to them – such that stimuli to which one is deliberately paying attention gradually become able to elicit an attentional response automatically (and hence can enter into learning more effectively), whereas stimuli that the subject is trying to ignore can instead come to suffer from conditioned inattention such that they automatically elicit the response of being ignored resulting in slower learning to these stimuli.

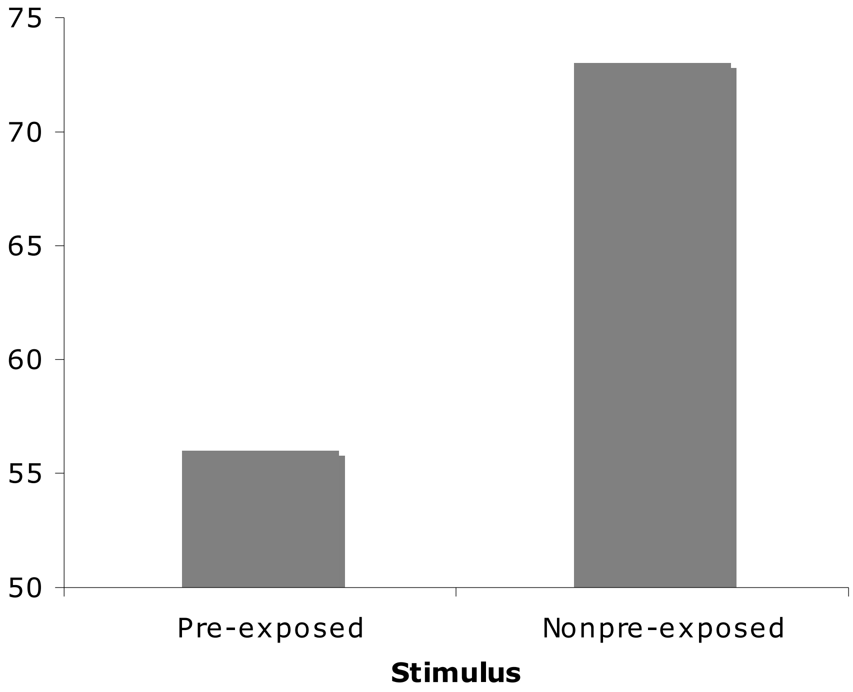

Figure 4: Results of Experiment 2. Chance performance = 50%.

Thus, in Experiment 1 participants needed to pay attention to the dot patterns, and over time this became an automatic response for the dot patterns they were trained with (but not to just any stimulus that was a dot pattern, at least, not to the same extent, as the results with the control stimuli show). The outcome would be that these patterns then elicited attention when encountered in a subsequent discrimination, and this would lead to faster acquisition of the discrimination. In Experiment 2, however, the stimuli were masks, and the response conditioned to them was to ignore them. Thus, when these stimuli were used in the final discrimination task the attentional response to them was to lower their salience, and this led to slower acquisition. One key aspect of these results is that they are stimulus specific: That is, only the stimuli experienced by the participants in the experiment demonstrate the full effect. It is not possible to explain these findings by reference to changes in some general level of attention to all stimuli encountered subsequently as this would predict no difference in acquisition to the discrimination between experimental stimuli and controls. Hence, the notion of some general gating of stimulus input that varies in the course of experience is not sufficient to explain the results that we have considered so far.

This analysis is further supported by a series of experiments performed by Graham and McLaren (1998) in which participants had to attend to one set of chequerboard patterns in order to classify them as members of one of two possible groups, whilst ignoring other chequerboard patterns onscreen (that they nevertheless had to look at). After this phase (Phase 1), they were then required to learn discriminations of exactly the type used here (Phase 2) between pairs of patterns they had either been attending to or ignoring. The result was that they showed faster acquisition than to control pairs for the attended to patterns, but slower acquisition (relative to controls) to the patterns they had been ignoring during pre-exposure. Importantly, Graham and McLaren were able to go one step further, and show that a similar result was obtained even if the to-be-discriminated patterns used in Phase 2 were small distortions of the patterns pre-exposed in Phase 1 (this technique is further elaborated when considering Experiment 3 shortly). The result would not be expected on the basis of theories which explain perceptual learning effects in terms of latent inhibition, as these theories would predict that the features of incidentally exposed stimuli would accrue latent inhibition, and that small changes to these stimuli would then stand out which in turn would aid discrimination based on these distortions. This latter result, which is the opposite of the effect observed by Graham and McLaren, has been obtained in experiments using simple pre-exposure in animals other than humans (e.g. Gibson, Walk, Pick, and Tighe, 1958; Forgus, 1958a;1958b), in particular in a very similar experiment using chequerboard stimuli with pigeons (Aitken, Bennet, McLaren and Mackintosh, 1996). It is at this point, then, that we part company with Lubow – who proposes that his conditioned attention mechanism can also serve as a model of latent inhibition in infra-humans. Instead, we feel that this result makes it all the more likely that what we have here is not some effect analogous to latent inhibition in animals, but rather a phenomenon that derives from attentional processes that make the transition from intentional control to automatic allocation in the course of experience with these stimuli. The argument would then be that the distortions of the pre-exposed stimuli were small enough to allow generalisation of the conditioned attention (or inattention) to them – leading to the same results as before (more on this later as well).

This analysis prompts us to ask "are all the effects of previous experience with stimuli due to conditioned attention or inattention to those stimuli?". Our answer will be "no" - there are perceptual learning phenomena that do not admit of this explanation and also effects whereby experience with certain stimuli retard learning to a stimulus in a way that can not be explained by appealing to attention. To prove this assertion, however, we will need to consider more complex pre-exposure paradigms. The basic approach to be used is illustrated in Experiment 3.

<u>Perceptual learning and categorisation.</u>

Experiment 3 examined the effect of category learning on the ability to discriminate both <u>between</u> two categories and <u>within</u> a given category. The experiment consisted of two phases. In the first the participants learnt to categorise stimuli into two categories. In the second they learnt discriminations as in Experiments 1 and 2, using category exemplars and category prototypes as the stimuli.

Twenty four participants participated in this experiment. The stimuli used were of the same type as the masks used in the delayed matching task in Experiment 2, but were generated separately for each subject and were not used as masks. Two base stimuli were generated randomly as before. These 'prototypes' were used to define two categories by generating a set of exemplars for each category by adding noise to the base patterns. The technique for adding noise involved changing whole horizontal lines of blocks by replacing that line with another randomly generated one. This was done stochastically, with the probability of a line being replaced being .39 so that on average 6.25 lines were changed. This resulted in considerable distortion of the base stimulus. Examples of the base stimuli and exemplars are shown in the figure below.

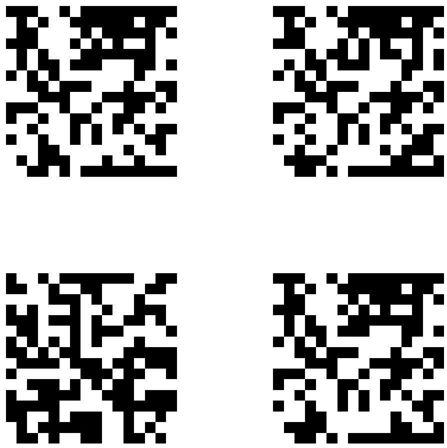

Figure 5: Example stimuli from Experiment 3.

The lower pair of stimuli are example prototypes, and the upper pair are minimal distortions of the lower right hand prototype or base pattern. This kind of exemplar is used in the discrimination task that follows on from the category learning phase of this experiment, but the exemplars used in categorisation training were much more distorted than the exemplars shown here. As many exemplars as were necessary for a given subject were generated online during the course of the experiment.

Participants were told that once they pressed the space bar a constant stream of stimuli would appear on screen, and that their task was to sort these stimuli into two categories. They were to do this by pressing one of two keys 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 participants were warned that the task would be quite difficult initially, as the stimuli would vary considerably and some would be quite ambiguous. Nevertheless the stimuli belonging to a given category would tend to share features in common, though the variability would be such that no particular feature would be a reliable index of category membership. For this reason they were encouraged to scan the stimulus before making a decision; speed of response was relatively unimportant, the need for accuracy was emphasised. This was further sharpened by telling them that after a certain minimum number of trials had been completed this phase of the experiment would end " when they were doing well enough". In fact at least 50 trials were required and the criterion required to finish was six correct responses in a row, which must include the last response made. Once the subject initiated the experiment, trials were continuous. Stimuli were presented singly, with 1 sec allowed for feedback and a 1 sec pause before the next stimulus came on screen, during which a fixation stimulus was displayed. No summary feedback was presented when the session terminated.

Five minutes after they had completed the categorisation phase, participants progressed to the discrimination phase of the experiment. The procedure for the discrimination task was exactly as in Experiments 1 and 2, with the following changes in stimuli and design. For each subject the 'pre-exposed' stimuli were the two base patterns from that subject's categorisation phase, and two new exemplars (from one category) that were generated so as to be very similar both to one another and to the prototype for their category. This was achieved by making the probability of altering a line only 1/16, so that only one line is expected to change. Checks were run to ensure that the resulting exemplars were not in fact the same. The effect of all this was to produce one pair of stimuli that were moderately similar (the two base stimuli) and one pair that were very similar but discriminable. Examples of two base stimuli and two exemplars used in this task were shown in the figure earlier. Control, nonpre-exposed stimuli were taken from another subject so that all stimuli were used in the same way in both the pre-exposed and nonpre-exposed conditions, but for different participants.

One subject was unable to complete the categorisation phase of the experiment and did not go on to the discrimination task; hence the results for 23 participants are reported here. The overall mean for the number of trials to acquisition was 68.9 with a standard error of 4.3. Our measure of how well participants can categorise the stimuli is given by the number of correct categorisations divided by the total number of trials attempted by a given participant expressed as a percentage. The last six correct trials were excluded from this analysis, so that a null hypothesis of 50% correct could be assumed to apply. Overall the mean percentage correct was 62.4 with a standard error of 2.3. The accuracy data for each of the four conditions during the discrimination phase of the experiment is shown in Figure 6.

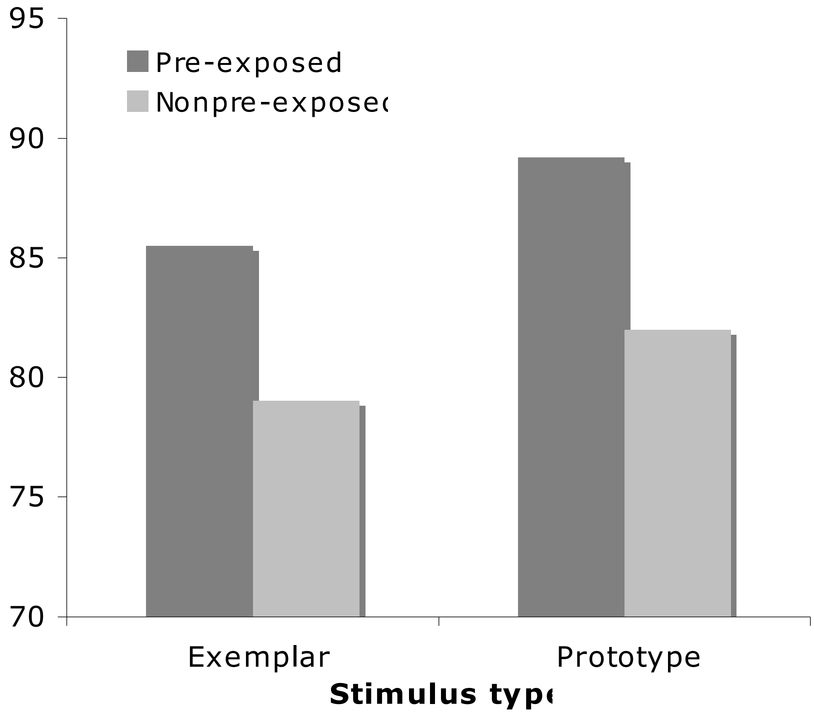

Figure 6: Results of the test phase in Experiment 3.

The graph indicates that pre-exposure resulted in faster acquisition of the discrimination to a roughly equal extent for both stimulus types, and that the prototypes were somewhat easier to discriminate than the exemplars. Analysis of variance revealed a significant effect of pre-exposure, $F(1,22) = 3.13$, $p(\text{1-tailed})$[1] $< .05$. Pre-exposure facilitated acquisition of the discrimination roughly equally for both stimulus types and there was a similar pattern in the latency to respond.

The results presented here demonstrate that learning to categorise stimuli that are appropriately characterised as distortions of a prototype enhances the discriminability between the prototypes for the two categories involved and between two novel exemplars drawn from one of the categories. As neither the prototypes nor the exemplars used in the test discrimination were shown during training, this extends the results of Experiment 1 considerably. Putting this difference to one side for the moment, the pre-exposure regime for the prototypes is perhaps closest to that used in Experiment 1, as in both cases the stimuli (dot patterns in Experiment 1, category exemplars in Experiment 3) have to be distinguished from one another either to make a same / different decision (Experiment 1) or to allocate the exemplar to the appropriate category (Experiment 3). If this is necessary for attention to accrue to the stimuli then we might expect attention to become conditioned to the two categories of stimuli used for each subject and then

[1] A one-tail test because the effect was predicted on the basis of theory (see later) and other experimental work not considered in this chapter.

this would be elicited by the prototypes during discrimination training (as in some sense they are the "best" exemplars of each category) and lead to enhanced acquisition. This account obviously requires some mechanism for generalising conditioned attention between category exemplars to be successful. What that mechanism might be is further constrained by the result obtained with the exemplar discrimination. This was that a discrimination between two very similar exemplars drawn from the same category was acquired more rapidly if that category was familiar. First of all note that in this case, there is now no incentive for participants to have previously tried to distinguish between the exemplars that belonged to one category during training, so the implicit requirements in play during pre-exposure were somewhat different for these stimuli and if we are to appeal to the same attentional mechanisms to explain the effect then we must take this into account. But perhaps the strongest inference that can be made on the grounds of this result would be that generalisation of conditioned attention must be driven by a whole stimulus comparison rather than operating on a feature by feature basis. The argument is that if attention is allocated to features of stimuli, then those features that best discriminated between the two categories will be the ones that have attention conditioned to them. These will be the prototypical features (which would help explain enhanced discrimination of the prototypes). But the exemplars created for the test discrimination have not been seen before, and they differ only on features on which they also differ from the prototype. Put another way, they are very similar, and this similarity is because they both have a very large number of prototypical features. If these features elicit an attentional response, this will make discrimination more difficult, as the subject will be preferentially attending to the features that are the same for the two stimuli rather than those on which they differ. Not a recipe for successful discrimination.

In fact this argument can also be deployed when considering the results of the final experiment in Graham and McLaren (1998). In that experiment, distortions similar to those used to make the exemplars here were performed on a base pattern that had previously been incidentally exposed to create two stimuli from that subsequently had to be discriminated. If conditioned inattention had been to the features of the base pattern, then the distortions, that is the changed features in the two variants that had to be discriminated, would have been relatively salient and this would have helped discrimination. The fact that all the features common to the two stimuli would have been ignored would also have been helpful in acquiring the discrimination, and this would not easily have produced the retardation in acquisition relative to controls that was the result of that experiment.

Our conclusion, then, has to be that generalisation takes place on a "whole stimulus' basis if we wish to continue with our conditioned attention / inattention account of the phenomena considered to date. With this proviso, the results of Experiment 3 can perhaps be accommodated and we have yet to find circumstances in which the conditioned attention approach to changes in learning as a consequence of experience with stimuli has difficulties in explaining the results. This, however, will change when we consider the next variation on this design.

Perceptual learning, categorisation, and inversion.

McLaren (1997) was able to demonstrate that one of the perceptual learning effects commonly found in face recognition, that performance on upright faces is much better than on inverted faces and that this difference is significantly bigger for faces than for other stimuli such as houses (Yin, 1969), can also be induced in the laboratory with novel stimuli. The stimuli used were chequerboards of the type considered in Experiment 3. Two categories were generated by creating prototypes and then adding noise to them to create exemplars of each category. Participants were trained on these categories until they were able to classify the stimuli successfully as in Experiment 3, then transferred to a test discrimination which this time involved pairs of stimuli that were similar exemplars drawn from one category. Four pairs were used concurrently as before, two experimental and two control pairs taken from another participants experimental stimuli. The two experimental pairs were in one case two upright exemplars of one of the categories that had not yet been seen by the subject, and in the other case two inverted (rotated through 180 degrees) exemplars from the same category (also unseen to that point). The question asked in this experiment was how inversion would interact with the perceptual learning effect that was expected to follow from the experience with the categories prior to learning these discriminations. One very appealing feature of this experiment was that an inverted chequerboard is still a chequerboard, thus minimising any cues the participants might have access to with regard to the experimental manipulation that was in play. In the case of inverted faces you know immediately that you are dealing with an upside down face and this may alter the way you go about dealing with it. In the case of the chequerboards this was less likely to be an issue.

The results were clear and strongly supported the notion that the face inversion effect is, at least in part, a perceptual learning phenomenon. Performance on the upright familiar exemplars was significantly better than on the inverted familiar exemplars (an inversion effect), and was also significantly better than on controls, replicating the effect of better discrimination within a category contingent on pre-exposure to that category. In addition, the inversion effect for the familiar exemplars was significantly greater than any such effect in the controls (in fact there wasn't even a hint of one there). This result is an analogue of the face inversion effect and can also be explained by our attentional mechanism for perceptual learning. We simply have to assume that attention is conditioned to the upright exemplars that participants are trained on and generalises to other novel exemplars during test, but that inversion renders the resultant chequerboards sufficiently unfamiliar to prevent this generalisation occurring. The analogy to the face inversion effect could not be considered complete, however, until it was possible to show that other types of stimuli did not generate this effect even though participants were familiarised with them. Recall that part of the face inversion effect result is that the difference in performance in a recognition task between upright and inverted faces is much bigger than for other types of stimuli such as houses. The problem here was to come up with a chequerboard stimulus that could play the role of the image of a house the same way that the exemplars drawn from a prototype-defined category had played the role of faces in this experiment. The solution adopted was to construct exemplars from base patterns by shuffling or permuting rows of squares as units rather than adding noise at random. Thus, after generating a base pattern, the first exemplar simply shuffled, at random, the sixteen rows that made up the base pattern. Additional exemplars were generated in the same way, with a random permutation of all the rows in the chequerboard stimulus to create the exemplar. This produces a category with a quite different structure to the prototype-defined ones previously used. There is now no "special" exemplar (this would be the prototype itself in the prototype-defined category), all exemplars have equal status, even the base pattern, and averaging all the exemplars generated would not result in the base pattern as was the case when adding random noise to generate exemplars. This technique also promised to generate categories that could still be learned and this did indeed turn out to be the case. Using the same procedures as before 24 participants were able to learn to categorise these stimuli (two categories used) in a mean of 76 trials with a mean percentage correct of 64.2%. These were slightly better, but entirely comparable results to those obtained with prototype-defined stimuli in this experiment (and very similar to those in Experiment 3), and there were no significant differences in categorisation between the two groups trained on the different stimulus types. The key results, however, are those from the discrimination test after categorisation training. The results are shown in Figure 7 below.

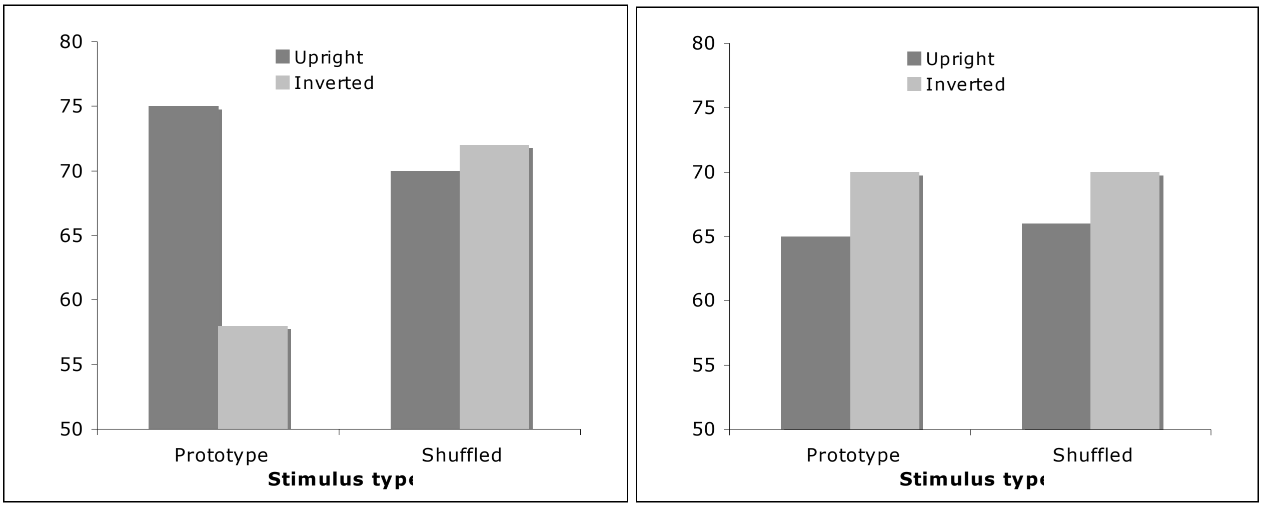

Figure 7: The effect of stimulus type and orientation for the pre-exposed stimuli (left panel) and the control stimuli (right panel).

There was no inversion effect in this case (right panel of figure), no perceptual learning effect either and there was an interaction between these results and the results with the prototype-defined stimuli (left panel of figure) which forces the conclusion that the inversion effect was contingent not just on familiarity with the stimuli that made up the category concerned, but also on the category being prototype-defined, i.e. possessing a certain stimulus structure. It is this role for stimulus structure in generating the perceptual learning effect that conditioned attention theories of perceptual learning cannot explain. If attention is conditioned during the categorisation phase, then this should apply equally well whether the category is prototype-defined or created by row permutation, given that both categories are equally easy to learn. Taking this last point, the learning data show no significant differences and a slight numerical advantage in this case for the shuffled stimuli. Post categorisation tests showed that upright shuffled exemplars were classified at 70% accuracy and inverted ones at 59% accuracy with upright prototype-defined exemplars at 66% and inverted ones at 59%. All these results are above chance, and whilst performance on the upright stimuli is significantly better than for the inverted stimuli (as might be expected, note that the inverted shuffled stimuli will be rather different to the upright shuffled stimuli because of the left/right reversal contingent on being rotated by 180 degrees), there is no effect of category type, and no interaction with this factor. Given this, it is hard to explain why a conditioned attention mechanism would not predict exactly the same pattern of effects for the shuffled stimuli as for the prototype-defined stimuli – but this is very much not what is observed. We are forced to search for some alternative mechanism to explain the perceptual learning effects found here, and to re-evaluate the role of attention in perceptual learning.

Perceptual learning as a consequence of latent inhibition

In fact, the experiments just reported were motivated by a quite different theory of perceptual learning, one based on a model of latent inhibition first proposed by McLaren, Kaye and Mackintosh (1989) and subsequently developed in McLaren and Mackintosh (2000, 2002). In this model latent inhibition is considered to be caused by a reduction in stimulus salience due to the lower activation of units representing the features of stimuli, which itself is a consequence of the associations that have formed between these units in the course of experience. The more reliable and more frequent the co-occurrence of certain features in a stimulus, the more the units representing those features become associated. This then lowers their activation, making them slower to enter into new association formation.

This model can explain how familiarisation with a prototype-defined category can lead to better discrimination of exemplars taken from that category. Exposure to exemplars of the category leads to strong associations forming between units representing prototypical features of that category as they will co-occur reliably. This will lead to a reduction in the activation of the units representing the prototypical features, which are exactly those features that tend to be shared by exemplars drawn from that category. It will be the non-prototypical features of an exemplar (those introduced by the addition of noise) that will be relatively more salient as they will not suffer from latent inhibition in the same way. As these features can be used to differentiate between exemplars, the prediction of enhanced discrimination between exemplars drawn from a familiar prototype-defined category follows.

But the real strength of this account is that it can explain why the same result would not be expected in the case of a category that is not defined by a prototype but instead has the structure of our shuffled stimuli. In this case there are no prototypical features shared by most exemplars that are represented by units that then lose salience as a result of association formation. Most of the features of an exemplar drawn from a category with this structure will tend to be relatively unpredicted and hence salient. As a consequence this mechanism cannot operate and confer any benefit contingent on familiarity with the category that will reduce the perceived similarity between two exemplars. Thus, perceptual learning will be weak or non-existent, and this will undermine any inversion effect that relies on it. The reader may object that if there are no shared prototypical features, then how is it that our participants are able to discriminate successfully between categories? Surely there must be some set of features held in common by the exemplars of a particular category that makes classification possible? There may well be; one example that comes readily to mind is the overall luminance of the stimuli. If one category happens to have more black squares than the other – or even more black squares on the left than the other category does – then this holistic attribute of the stimuli will enable reliable categorisation and units representing it may come to suffer latent inhibition. But the more local features of any exemplar due to a particular arrangement of squares will not accrue much by way of latent inhibition, and it is these features that will be needed to discriminate between exemplars. Many of these features will be shared (by chance) by any given pair of exemplars and these common features will be as salient as the unique features which will make acquisition of the discrimination between them difficult.

The perceptual learning that accrues as a result of experience with a category defined by a prototype may not be explicable by means of conditioned attention then. The attentional account cannot be rescued by shifting to a modified version whereby attention is conditioned to individual features rather than the stimulus as a whole, because, as we have already seen, this would lead to the prediction of retarded rather than speeded acquisition in many cases. It would be the units representing prototypical features that would elicit attention the most strongly and this would tend to make a discrimination between category members more difficult rather than easier. The same argument applies to Graham and McLaren's (1998) finding (though in this case to give the converse result) as here conditioned inattention to units representing the features of the pre-exposed base patterns would, when distortions are introduced, mean that units still representing the remaining original features would be less salient, the units representing the distortions more salient, and so discrimination on the basis of the distortions would be enhanced. We must acknowledge, however, that our latent inhibition-based mechanism for perceptual learning has no way of explaining the basic conditioned inattention result in Experiment 2 or in Graham and McLaren (1998). Some attempt can be made to appeal to salience reduction and to postulate that this is severe in the case where incidental pre-exposure is used, but the inevitable consequence of deploying a theory that operates at an elemental or feature-based level is that, given the initial requirement to explain the retarded learning consequent on incidental pre-exposure, when noise or distortions of some kind are introduced the prediction is that these distortions are relatively salient and that discrimination in the sort of scenario employed by Graham and McLaren will be facilitated. As this is clearly not the case, we are left with the need to acknowledge that a conditioned inattention explanation of these results is required, and if this is allowed then it would seem odd to deny that conditioned attention could play at least a part in the results of Experiments

1 and 3 (in the latter experiment particularly for the discrimination between the prototypes of the two categories, as the requirement during familiarisation is to distinguish between the categories defined by these prototypes). The requirement would seem to be for two classes of theory here, one operating at an elemental level (latent inhibition) and one at a whole stimulus or configural level (attention).

Latent inhibition in humans

Before we can feel entirely secure in this conclusion, however, it would be as well to give further consideration to how these two classes of mechanism might complement one another. In particular, given that one of them is a theory of latent inhibition, we must discuss latent inhibition in humans and the somewhat vexed question of why it is so hard to find simple, straightforward evidence of latent inhibition in humans. Note here that we are not directly addressing the question of whether latent inhibition can be demonstrated in humans. This is an interesting question in its own right, and there have been reports of such an effect in a variety of conditioning preparations (e.g. Schnur and Ksir, 1969; Surwit and Poser, 1974; Siddle, Remington and Churchill, 1985; Klosterhalfen, Kellermann, Stockhorst, Wolf, Kirschbaum, Hall, & Enck, 2005), but it is not, perhaps, the most interesting question to ask. The problem for the elemental theory of latent inhibition put forward by McLaren, Kaye and Mackintosh (1989) and others like it was always that when applied to humans it seemed to suggest that if they were simply pre-exposed to a stimulus, then had to learn about that stimulus, the consequence should be that learning would be slow relative to control stimuli. This has been known to be the case for infra-humans since the seminal work of Lubow and Moore (1959) and has proven a straightforward and reliable phenomenon to demonstrate, but has by comparison been an elusive target for human studies. One proposal (Lubow, 1980) had been that incidental pre-exposure is required to reveal latent inhibition in humans, which would fit well with Experiment 2 and some of Graham and McLaren (1998) but cannot survive the finding that distortions of incidentally pre-exposed stimuli are still learned about more slowly than controls, which is quite unlike the result typically found with infra-humans (see Hall, 1980 for a review of early evidence on this point, and it is reinforced by the results of Aitken *et al*, 1996). The question is, then, why latent inhibition in humans should prove so difficult to demonstrate if it is the case that it plays a key role in producing perceptual learning phenomena that are, by comparison, readily observable? The answer we shall offer to this conundrum is that it is a consequence of the fact that both classes of mechanism, those based on latent inhibition and attention respectively, are typically operating in studies in which stimuli are pre-exposed, and that the effects of one mechanism (based on conditioned attention) often conceals at least some of the impact of the other (based on latent inhibition).

Consider the most basic type of pre-exposure to a stimulus. A human subject is repeatedly shown a stimulus on a screen, let's say it's a chequerboard pattern to be both specific and topical. After some time they enter a new phase in which they have to learn that, when this pattern is shown on screen, if they press a key they will receive a monetary reward. This is as close to a typical latent inhibition experiment in infra-humans as one can come without employing more specialist conditioning preparations (eyeblink, electrodermal, induced nausea would be examples of these). If a thirsty rat is pre-exposed to a tone or light in an operant chamber then trained that the tone or light signals water delivery it will learn this relationship more slowly than nonpre-exposed controls (e.g Mclaren, Bennet, Plaisted, Aitken and Mackintosh, 1994), so why should we not expect a similar effect here with our human participants? Our explanation is that during pre-exposure attention was conditioned to the pre-exposed stimulus at the same time latent inhibition of the elements making up that stimulus also took place. In combination these two effects could produce almost any result, but in terms of simple learning to the stimulus they clearly act in opposition to one another making little detectable impact on learning a not unlikely result. If instead we now consider the case where two rather similar chequerboard patterns are pre-exposed before training our subject to discriminate between them then the analysis is the same but the expected outcome is rather different. The effects of latent inhibition will be to reduce the salience of the common elements shared by the two stimuli rather more than the elements unique to a particular stimulus. The effects of conditioned attention will be to raise the salience of all the elements in each stimulus, which will leave the differential effects of latent inhibition on the shared and unique features of the stimuli unaltered. The prediction is now quite clear, discrimination between the two stimuli should be enhanced as their representations have been fine-tuned to make them more discriminable without incurring much by way of a "latent inhibition penalty". The two mechanisms that are invoked by stimulus pre-exposure have here combined to produce perceptual learning in a sophisticated fashion that improves on what each mechanism on its own would be capable of.

A similar analysis follows for pre-exposure leading to conditioned inattention. Incidental exposure to our stimulus will produce a combination of latent inhibition and conditioned inattention that clearly predicts that simple learning will be slower to this stimulus. The case of discrimination between two similar and incidentally pre-exposed stimuli is a more ambiguous scenario, with the differential effects of latent inhibition to some extent opposing the effects of overall latent inhibition and conditioned inattention, but the expectation will be for the overall loss in salience to quite often outweigh the differential effects on latent inhibition.

Reverse perceptual learning.

Armed with this hypothesis about the effects of pre-exposure being mediated by two different classes of mechanism, one elemental and resulting in latent inhibition, one holistic and operating by means of conditioned attention, we can now ask if it is possible to find circumstances where, without using special conditioning procedures (which have a somewhat controversial status with regard to what actually is learned when they are used) or incidental pre-exposure (which is likely to invoke conditioned inattention and so confound our results), it will be possible to demonstrate retarded acquisition as a consequence of pre-exposure in humans? In fact, it is possible to find a set of circumstances where this type of effect should manifest in human experiments using simple pre-exposure, but they are quite restrictive and unlikely to occur in the normal run of things. This, in itself, we take to be a good thing, in that it indicates that the combination of mechanisms for attention and latent inhibition is very robust in preventing what might be seen as the undesireable consequences of latent inhibition affecting the general rate of learning, whilst maintaining its utility as a mechanism for producing perceptual learning.

One requirement for obtaining poorer learning as a consequence of simple pre-exposure is to use stimuli that do not benefit from the differential latent inhibition of common elements. We can achieve this by using the shuffled stimuli from McLaren (1997). But recall that in that experiment there was no evidence of retarded learning to the shuffled stimuli, it was more that they did not benefit from the perceptual learning effect that accrued to the prototype-based stimuli. We will return to this point shortly, but the conclusion we reached was that we also needed to find a task that involves stimuli that vary from trial to trial. The particular paradigm that meets these requirements is given in Wills, Suret and McLaren (2004), (see also Wills and McLaren, 1998 and Wills and Mclaren, 2002). In this paradigm pre-exposure to the stimuli is given by means of a running recognition task. Stimuli are shown one at a time to a subject who has to decide if it is the first (new) or second (old) time in the block that that particular stimulus has been seen. The pre-exposed stimuli were drawn from two categories, with the exemplars shown generated by shuffling or permuting the rows of a base pattern as before. After this pre-exposure phase, the test phase was not the type of discrimination experiment used in the other studies discussed here. Instead, this phase involved participants learning to categorise stimuli from the two categories that had been pre-exposed, with the categorisation task being performed in exactly the way used to provide pre-exposure in our earlier experiments. A separate control group was also run on this test phase after experience of running recognition using quite different stimuli. The dependent measure in this experiment was mean number of trials to criterion, with criterion being a run of six consecutive correct responses. The results of this experiment are shown in Figure 8. They indicate that pre-exposure via a running recognition task, which can hardly be classified as incidental pre-exposure, increases the number of trials taken to learn to successfully categorise the shuffled stimuli to a criterion of six correct responses in a row. Statistical analysis showed that this difference was significant, $F(1,20) = 4.67$, $p<.05$. In order to check that this result was genuinely a result of the stimulus structure of these categories and not an inevitable consequence of our procedures another condition using exemplars from prototype-defined categories was also run (i.e. running recognition followed by categorisation training). This gave the opposite pattern of results, with pre-exposure now leading to faster acquisition (i.e. fewer trials to criterion) as expected, and there was a significant interaction between the effects of pre-exposure and stimulus type used in the experiment, $F(1,40) = 8.66$, $p<.01$. Thus we have strong evidence that simple exposure to the shuffled stimuli can retard learning involving them, in this case learning to classify them as members of one category or another.

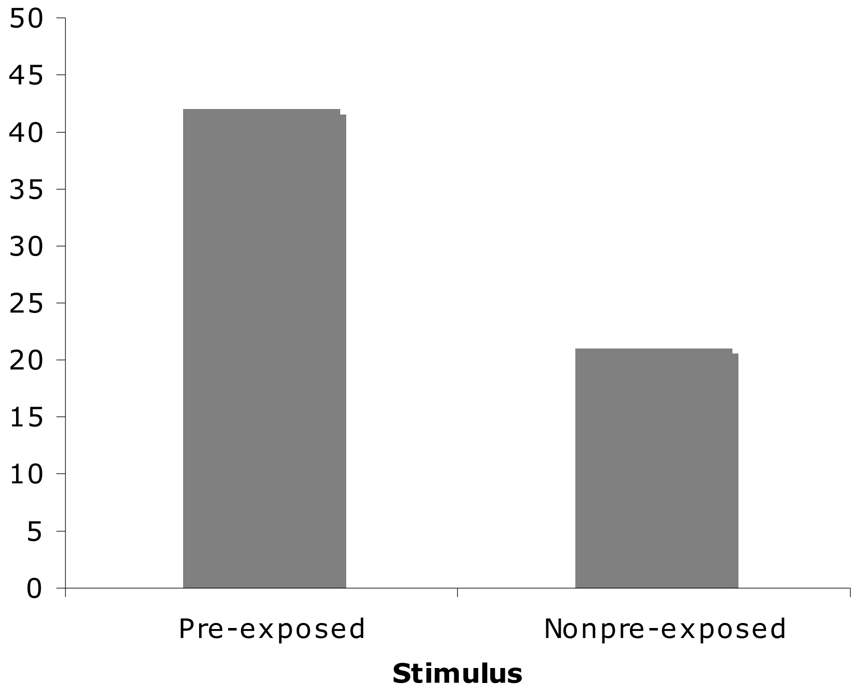

Figure 8: Results of the categorisation test phase for shuffled stimuli after pre-exposure in a running recognition task. Higher scores indicate worse performance.

Having obtained this result we can now ask why it follows from the analysis presented in this chapter. The key to understanding this lies in the fact that the latent inhibition mechanism is operating on the shuffled stimuli, and that as a consequence it will modulate the relative salience of some features of these stimuli considerably. The full analysis of what happens in this experiment is complex, but a simple thought experiment will make the general idea clear. Imagine for a moment that the two base patterns that are used to generate the category exemplars differ in the number of black and white squares that occur in a given column. Perhaps the third column from the left has many black squares in Base Pattern 1, but rather few in Base Pattern 2. This feature of these stimuli will be invariant under the row shuffling transformation, and as a consequence of experience with these stimuli we assume that participants learn that the chance of a white square in this column is rather low for the stimulus set defined by Base Pattern 1, but rather high for the set defined by Base Pattern 2. By this we do not mean that the participants have learned to categorise the stimuli at this stage during pre-exposure, but that other features present in the stimuli predict that the squares in the third column will be black in the case of stimuli from Category 1 and white in the case of stimuli from Category 2. Now, when the subject moves onto the categorisation phase they will be attempting to learn which key to press to any given stimulus by trial and error. Any learning that takes place will tend to involve the most salient features of the stimuli on screen, and we can ask what these will be for the third column of our hypothetical stimuli. For members of Category 1 where a black square will be expected in any position in this column it will be the few white squares that have been shuffled into position that will be most salient, the majority black squares less so. The converse will apply to Category 2. Hence a white square in the third column will tend to be associated to Category 1 membership, and a black square to Category 2 membership. But, in fact, the opposite contingency is in force, and a black square is predictive of this exemplar being a member of Category 1. Recall that the position of the minority colour squares will change from exemplar to exemplar, so it is only the general contingency between colour of square and category that will prove to be reliable over time in this task, and this will eventually lead to acquisition, but the counteracting effect of stimulus salience will slow this process and retard acquisition of the discrimination between these categories compared to nonpre-exposed control categories. Hence, the retardation observed can be said to be a consequence of the latent inhibition mechanism operating under conditions of simple exposure, but it relies on differential latent inhibition in this case emphasising the wrong features to aid discrimination rather than the right ones as in the case of prototype-defined categories. As such, it may perhaps be better to consider it as more a case of reverse perceptual learning rather than latent inhibition per se. Note also how this explanation relies on the stimuli changing from trial to trial in this particular task, if instead we had used our conventional discrimination design taking two exemplars of one of the familiar categories, then the salient features of each (the unexpected squares in a column) could have been exploited to speed acquisition somewhat.

Perceptual learning and attention: A model and some conclusions.

There is one more topic to discuss before we can summarise with some concluding remarks about attention and perceptual learning, and that is to address the issue of what a model that combined latent inhibition and conditioned attention might look like. Figure 9 gives one possible instantiation of the ideas contained in this chapter.

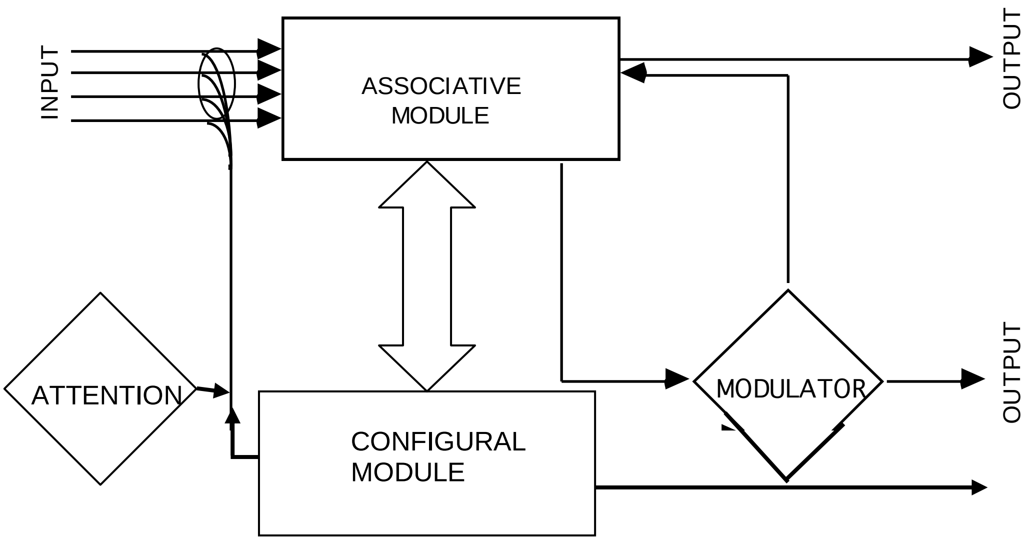

Figure 9: A combined elemental latent inhibition and configural conditioned attention model of stimulus representation and development.

In essence this is the McLaren, Kaye and Mackintosh (1989) model with a configural module added that can detect when certain combinations of stimulus elements or features occur and so recognise that a particular stimulus is present. This configural representation then has connections that can gate input to the associative module that represents stimulus features. If attention is being paid to the stimulus then these connections are strengthened enhancing input to the associative module, if the stimulus is being ignored then these connections can become negative and suppress input to the associative module. This process of raising or lowering attention is, in itself a controlled and active one which can be flexibly determined by the individual, but the connections themselves persist over time, such that a stimulus that has been repeatedly attended to will elicit attention when presented again because there will be strong positive weights from its configural representation (which will be activated by presentation of the stimulus) gating stimulus input. Equally, a stimulus that has been ignored in the past will have developed strong negative weights from its configural representation that will suppress stimulus input when it is presented again (other things being equal). It is important to note that the configural module, which serves the purpose of recognising specific stimuli with some history of a given level of attention being paid to them, must be set-up so that once a configural unit is activated its effect in gating stimulus input is general and not just confined to the features of the stimulus that activated it. If this were not the case then the effect of conditioned attention to stimuli that were small distortions of previously experienced stimuli would be incorrect – and would lead to elemental-type results. Instead a stimulus similar to one that has some past history of conditioned attention or inattention will partially activate the appropriate configural unit which must then gate stimulus input for that entire stimulus to some extent. This will ensure that conditioned attention will generalise in the correct fashion to complement elemental latent inhibition.

I must, of course, acknowledge that the proposed model leaves many questions unanswered, and, indeed, that there are many issues to do with the relationship between attention and perceptual learning that have not been resolved in the course of this discussion. We believe that we have made a powerful case that there is a role for attention to play in perceptual learning in humans. The data seem to require that there be two separate mechanisms active during pre-exposure, one elemental and modulating unit salience on the basis of how predicted that unit is by others concurrently active, whilst the other is configural in nature and captures the history of controlled and flexibly allocated attention to a given stimulus, but we have not been able to provide unequivocal proof for the positive effects of conditioned attention. Instead, we have had to rely on the somewhat stronger evidence for conditioned inattention and argue that it would be rather strange if inattention could be conditioned and attention could not. The reader might argue that simple priming phenomena (e.g. repetition priming) provide the necessary evidence for this conjecture, but it would be better if there were evidence from the perceptual learning domain that supported the notion of conditioned attention unconfounded by other possible explanations such as those provided by the differential latent inhibition of common elements mechanism. We think it likely that the results of Experiment 1 are due, in part at least, to conditioned attention, but fall short of being able to prove that this is the case.

We have not considered other models of perceptual learning (e.g. Goldstone, 1998) and other theories of latent inhibition (e.g. Pearce and Hall, 1980) in this treatment. Our reason for this is that we wished to explore how attention and latent inhibition might be integrated into a single model without entering into arguments about which model of perceptual learning or latent inhibition was the best. Obviously the type of model of perceptual learning or latent inhibition one starts with will constrain any attempt to incorporate attentional factors into the model and probably change the nature of the final hybrid system, but we are happy to leave this exercise to the authors of the various models currently under development with the proviso that they will have to rise to the challenge provided by the data presented in this chapter. We have also not considered the strand of research on human (and infra-human) associative learning that makes the case for variation in the attention paid to stimuli as a consequence of past associative history (e.g. Mackintosh, 1975; Lochman & Wills, 2003; Le Pelley and McLaren, 2003; Le Pelley, Oakeshott, Wills, and McLaren, 2005). Again, our reason is that this would have diluted our focus on the issue at hand (which is more the effects of pre-exposure than pre-training), but we acknowledge that this research points towards the role that attention may play in parameterising associative learning. We certainly believe that the role of attention in representation development as a consequence of stimulus pre-exposure is a topic that is worthy of a great deal more study than has so far been the case, and look forward to a coming together in the next few years of conventional research on attention and the parallel strand of research on perceptual learning in humans and infra-humans that increasingly seems to contain hints of convergence with that on attentional processes. The models that emerge from that collision may be the first mature models of human representation development that are both rooted in research on infra-humans yet also have the capacity to model human abilities that go beyond the merely associative and take us into the realm where association and cognition interact to produce truly complex behaviours.

References

Aitken, M.R.F., Bennett, C.H., McLaren, I.P.L., & Mackintosh, N.J. (1996). Perceptual differentiation during categorization learning by pigeons. Journal of Experimental Psychology: Animal Behavior Processes, 22, 43-50.

Cantor, G.N. 1969. Stimulus familiarisation effect and the change effect in children's motor task behaviour. Psychological Bulletin, 71, 2, 144-160

Diamond, R. and Carey, S. 1986. Why faces are and are not special: An effect of expertise. Journal of Experimental Psychology: General, 115, 107-117

Forgus, R.H. 1958a. The effect of different kinds of form preexposure on form discrimination learning. Journal of Comparative and Physiological Psychology. 51, 75-78.

Forgus, R.H. 1958b. The interaction between form preexposure and test requirements in determining form discrimination. Journal of Comparative and Physiological Psychology, 51, 588-91.

Gibson, E.J., Walk, R.D., Pick, H.L., and Tighe, T.J. 1958. The effect of prolonged exposure to visual patterns on learning to discriminate similar and different patterns.

Ginton, A., Urca, G., & Lubow, R.E. (1975). T he effects of preexposure to a non-attended stimulus on subsequent learning: Latent inhibition in adults. Bulletin of the Psychonomic Society, 5, 5-8.

Goldstone, R. L. (1998). Perceptual Learning. Annual Review of Psychology, 49, 585-612.

Graham, S. and McLaren, I.P.L. **(**1998). Retardation in human discrimination learning as a consequence of preexposure: Latent inhibition or negative priming? Quarterly Journal of Experimental Psychology, 51B, 155-172.

Grison, S., Tipper, S.P., and Hewitt, O. (2005). Long-term Negative Priming: Support for retrieval of prior attentional Processes. Quarterly Journal of Experimental Psychology 58A.

Hall, G. (1980). Exposure learning in animals. *Psychological Bulletin, 88,* 535–50.

Kaniel, S. and Lubow, R.E. 1986. Latent inhibition: A developmental study. British Journal of Developmental Psychology, 4, 367-375

Klosterhalfen, S., Kellermann, S., Stockhorst, U., Wolf, J., Kirschbaum, C., Hall, G., & Enck, P. (2005). Latent inhibition of rotation-chair induced nausea in healthy male and female volunteers. Psychosomatic Medicine, 67, 335-340.

Lavis Y., Mitchell C. (2006). Effects of preexposure on stimulus discrimination: An investigation of the mechanisms responsible for human perceptual learning. Quarterly Journal of Experimental Psychology, 59, 2083-2101.

Lubow, R.E. , (1989). Latent inhibition and conditioned attention theory. Cambridge, UK:

Cambridge University Press.

Lubow, R.E., Alek, M. and Arzy, J. 1975. Behavioural decrement following stimulus preexposure: Effects of number of preexposures, presence of a second stimulus, and interstimulus interval in children and adults. Journal of Experimental Psychology: Animal Behaviour Processes, 104, 2, 178-188

Lubow, R.E., Caspy, T. and Schnur, P. 1982. Latent inhibition and learned helplessness in children: Similarities and differences. Journal of Experimental Child Psychology, 34, 231-256

Lubow, R.E., Rifkin, B. and Alek, M. 1976. The context effect: The relationship between stimulus preexposure and environmental preexposure determines subsequent learning. Journal of Experimental Psychology: Animal Behaviour Processes, 2, 1, 38-47

Lubow, R.E., & Moore, A.U. (1959). Latent inhibition: T he effect of non-reinforced preexposure to the conditioned stimulus. Journal of Comparative and Physiological Psychology, 52, 415-419.

Lubow, R.E., Weiner, M. and Schnur, P. 1981. Conditioned attention theory. In G.H. Bower (Ed.), *The psychology of learning and motivation.* Vol. 15. N.Y. AP

McLaren, I.P.L. (1997).Categorisation and perceptual learning: an analogue of the face inversion effect. Quarterly Journal of Experimental Psychology, 50A, 257-73.

McLaren, I.P.L., Bennett, C., Plaisted, K., Aitken, M. and Mackintosh, N.J. (1994). Latent inhibition, context specificity, and context familiarity. Quarterly Journal of Experimental Psychology, 47B, 387-400.

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.

McLaren, I.P.L., Leevers, H.L. & Mackintosh, N.J. (1994). Recognition, categorisation and perceptual learning. In C. Umilta & M. Moscovitch (Eds.) Attention & Performance XV.

McLaren, I.P.L. and Mackintosh, N.J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. Animal Learning and Behavior, *38(3),* 211-246.

McLaren, I. P. L. & Mackintosh, N. J. (2002).Associative learning and elemental representation: II. Generalization and discrimination. Animal Learning & Behavior, 30, 3, 177-200.

Mundy, M. E., Honey, R.C., & Dwyer, D.M. (2007). Simultaneous presentation of similar stimuli produces perceptual learning in human picture processing. Journal of Experimental Psychology-Animal Behavior Processes. 33, 124-138.

Pearce, J.M. and Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychological Review, 87,532-52.

Schnur, P., & Ksir, C. J. (1969). Latent inhibition in human eyelid conditioning. Journal of

Experimental Psychology, 80, 388-389.

Siddle, D. A.T., Remington, B., & Churchill , M. (1985). Effects of conditioned stimulus pre-exposure on human electrodermal conditioning. Biological Psychology, 20, 113-127.

Surwit, R.S. , & Poser, E. G. (1974). Latent inhibition in the conditioned electrodermal response. Journal of Comparativ e and Physiological Psychology, 86, 543-548.

Tipper, S.P. (1985). The negative priming effect: Inhibitory priming by ignored objects. Quarterly Journal of Experimental Psychology, 37A, 571-590.

Tipper, S. P., & Driver, J. (1988). Negative priming between pictures and words: Evidence for semantic analysis of ignored stimuli. Memory and Cognition, 16, 64-70.

Tipper, S. P., Weaver, B. , & Milliken, B. (1995). Spatial negative priming without mismatching: Comment on Park and Kanwisher (1994). Journal of Experimental Psychology: Human Perception and Performance, 21, (5), 1220-1229.

Treisman, A., & DeSchepper, B. (1996). Object tokens, attention, and visual memory. In T. Inui & J.L. McClelland (Eds.), Attention & performance XVI (pp. 15-46). Cambridge, MA: MIT Press.

Wills, A.J., and McLaren, I.P.L. (1998). Perceptual learning and free classification. Quarterly Journal of Experimental Psychology, 51B.

Wills, A. J., & McLaren, I. P. L. (2002). Frequency, predictability and salience.Unpublished manuscript, www.willslab.co.uk.

Wills, A.J., Suret, M.B. and McLaren, I.P.L. (2004). The role of category structure in determining the effects of stimulus pre-exposure on categorisation accuracy. Quarterly Journal of Experimental Psychology, 57B, 79-88.