

COMPARATIVE COGNITION & BEHAVIOR REVIEWS

2008

Volume 3, pp 115 -133

Use of multiple dimensions in learned discriminations

Stephen E. G. Lea and A. J. Wills
School of Psychology, University of Exeter, UK

Many naturally occurring categories vary across multiple stimulus dimensions (e.g. size, color, texture). When humans categorize multidimensional stimuli on the basis of a single dimension this has been taken to indicate use of a rule that could be verbalized. Sorting on the basis of all the stimulus dimensions ('overall similarity' or 'family resemblance') has been taken to indicate a more basic, implicit, automatic, perhaps associative process. However, a review of the literature on animal discrimination learning shows that animals often discriminate on the basis of one dominant dimension. In recent experiments, situations conducive to more complex cognitive processes have increased family resemblance sorting in humans. In an effort to resolve this apparent paradox, experiments were conducted in which humans and pigeons were exposed to multidimensional category discrimination tasks under closely similar conditions. Preliminary results show no evidence that even a non-verbal rule can be said to be involved in pigeons' choices in these conditions, despite the fact that under some conditions a single dimension may dominate their behavior.

Keywords: pigeons, humans, category learning, multidimensional stimuli, family resemblance

Background: Concept discrimination in animals and rule elaboration in humans

The natural world is full of objects that humans and other animals have to discriminate in order to carry on their daily lives. Such objects typically differ from one another on many different discriminable dimensions. However, it does not follow that either people or animals will use all the information available to them: they might use only a little of it, some of it, most of it, or all of it. Given that information

processing has a cost, it is not obvious which of these approaches we should expect to encounter.

Experiments in which people are asked to sort multidimensional stimuli into two categories have revealed two kinds of sorting that commonly take place. The classic result (e.g. Medin, Wattenmaker, and Hampson, 1987) is that most participants sort the stimuli on the basis of a single dimension only. Under other circumstances, however, people sort multidimensional stimuli on the basis of overall similarity or family resemblance, that is to say, basing their categorization on all or most of the available stimulus dimensions.

To introduce the concept of overall similarity clearly begs the question of how similarity is to be defined. Different definitions suit different purposes. The distinction between single-dimension versus overall similarity sorting is a question of how much of the available stimulus information is used, rather than a question about the precise nature of the information stored (e.g. cue validities or whole exemplars), or a question about the precise manner in which information from different dimensions is integrated (e.g. additive or multiplicative summation). These are interesting questions, but they are not the topic of the current article. Consistent with this approach, for the purposes of the current article,

Send correspondence to Stephen Lea, University of Exeter, School of Psychology, Washington Singer Laboratories, Exeter EX4 4QG, United Kingdom (s.e.g.lea@exeter.ac.uk). The experimental work described here, and the writing of this article, were supported by the New and Emerging Science and Technologies activity of the European Community Framework Programme 6, under grant no. 516542, "From Associations to Rules in the Development of Concepts" (FAR), and by grant RES-000-22-1779 from the UK Economic and Social Research Council. The unpublished experiments described here were carried out by Catherine Bryant, Britta Osthaus, Catriona Ryan, Mark Suret and Lou Millar. We also gratefully acknowledge ideas contributed during discussion of these experiments by other members of the FAR network, by Tom Zentall and Gerd Gigerenzer, and by delegates at the 2007 meeting of the Comparative Cognition Society.

we adopt the simple operational definition that the overall similarity between two stimuli is correlated with the proportion of the available stimulus dimensions that take the same value in both stimuli.

Many factors have been found to increase the likelihood of overall similarity sorting in humans. They include stimulus factors such as stimulus integrality (e.g., Garner, 1974; Handel & Imai, 1972; Lockhead, 1972), and presentational factors – family resemblance sorting is more likely with sequential rather than simultaneous presentation of exemplars (e.g. Regehr & Brooks, 1995), though unidimensional sorting can occur with sequential presentation (Ashby, Queller & Berretty, 1999). Procedural factors also have an effect: for example, increases in time pressure can increase the likelihood of family resemblance sorting (Smith & Kemler Nelson, 1984). Also, in general, children are more likely to sort according to family resemblance than adults (Smith & Kemler, 1977).

The distinction between unidimensional and overall similarity categorization has been taken to reflect two different kinds of cognitive processing in humans (e.g. Smith & Kemler Nelson, 1984). Unidimensional sorting has been assumed to reflect a more explicit, analytic, verbal or verbalizable, kind of process. Many authors have referred to this kind of process as sorting by rule. The hypothesis that unidimensional sorting is due to the use of rules is consistent with at least some of the evidence cited above: for example, it is more likely to occur in adults rather than children, and under relaxed testing conditions rather than under time pressure. Overall-similarity-based sorting, on the other hand, is taken to be characteristic of a more implicit, non-analytic, non-verbalizable, similarity-based process, for which the shorthand ‘associative process’ is sometimes used. This “rules vs. similarity” distinction has become very important in the analysis of human categorization (e.g. Ashby et al., 1998; Erickson & Kruschke, 1998; Nosofsky, Palmeri & McKimley, 1994; Pothos, 2005), and for this reason we have taken it as the focus of our recent experimental work. While no sensible person would deny that humans can verbalize rules and that these may govern their categorization, one of the purposes of the present paper is to challenge the value of unidimensional sorting as a diagnostic for recognizing when rule-based categorization is taking place.

While humans are presumably capable of either rule-based or associative categorization, most authors would assume that in other animals only the associative learning mechanism is available (for an exception to this assumption, see Beckers, Miller, De Houwer & Urushihara, 2006). We would therefore expect, on the basis of the human findings described above that, in animal cognition overall similarity sorting would be the usual finding. It is difficult to do sort-

ing experiments as such with animals, however, so there are no data from directly comparable procedures that would enable us to say whether animals do sort on the basis of overall similarity. However there is a well-established literature on a closely related question which is usually taken to imply that animals do indeed make discriminations on a family resemblance basis.

An early, influential, demonstration of this was the experiment by Herrnstein and Loveland (1964), in which pigeons were taught to discriminate between hundreds of different colored photographic slides, which were projected onto the side of an operant chamber. If the picture contained a human, the pigeon was rewarded for pecking a key; if the picture did not contain a human, pecks had no consequences. The humans in the picture were highly variable – some slides had only parts of people, some had single people, some had groups; and the people were in different postures, were in close-up or long-shot, were dressed in different ways, and so forth. Nonetheless the pigeons achieved high accuracy within “just a few” 80-trial sessions; Herrnstein and Loveland do not give more precise details of the training required, but in our laboratory such discriminations are typically learned in between 5 and 30 sessions, depending on the stimuli concerned.

There have been many similar demonstrations since. Most have involved birds as subjects (see, for example, the review by Huber, 2001), but a fair number have involved primates (e.g. D’Amato & Van Sant, 1987; Ferreira, Keller, Saint-Dizier, Perrin & Levy, 2004; Roberts & Mazmanian, 1988; Schrier & Brady, 1987; Vonk & MacDonald, 2002, 2004). These experiments involve what are usually called “natural concepts”, by which we mean concepts that are referred to in everyday human speech. The human experimenters choose stimuli that members of their language community agree are exemplars of the natural language concept under examination. In the latter half of the twentieth century, linguistic philosophers (e.g. Ryle, 1949; Wittgenstein, 1953/1968) observed that such natural language concepts frequently differ from the kind of concepts considered in classical logic, in that there are no discoverable necessary or sufficient conditions for membership in them. Rather they are defined by a “family resemblance” or fuzzy rules, with the result that there are cases whose membership of the concept is open to question. It is a matter of common experience that humans can nonetheless operate with such concepts fluently (Lea, 1979). Ryle termed such concepts “polymorphous” in recognition that exemplars of them could take many different forms.

Experimenters in animal cognition have sought to study polymorphous concepts by using artificial stimulus sets whose membership depends on the values taken by numer-

ous stimulus dimensions. A simple way of doing this is to specify that if a certain number of dimensions take the values characteristic of the concept, the stimulus is an exemplar of that concept. This approach was first taken in connection with human category learning (Shepard, Hovland & Jenkins, 1961, type IV stimuli; Dennis, Hampton & Lea, 1973) but was introduced into the animal discrimination literature by Lea and Harrison (1978) and has since featured in many other experiments on birds, such as Lea, Lohmann and Ryan (1993), Huber and Lenz (1993), Jitsumori (1993, 1996) and Lea, Wills and Ryan (2006). Similar stimuli have been used with other animals, e.g. rhesus monkeys (Jitsumori, 1994). Such experiments have generally found that, like humans, animals can successfully discriminate artificial polymorphous stimuli of this kind, though it can be very difficult both for humans and for birds (e.g. Dennis et al., 1973; Jitsumori, 1993).

At first glance, the fact that animals can be trained to make good discriminations between categories defined by family resemblance is what one might expect if overall similarity sorting indicates an associative process, since associative processes should be available to any bird or mammal (and probably to other kinds of animal as well). We would not however expect animals to be able to sort by using a rule.

Some inconvenient facts

So far, therefore, our look at the broad trends in the literature on human categorization and animal category discrimination has given us a tidy picture, of animals demonstrating discrimination by overall similarity while humans have the capacity for elaborating verbal rules and thus show unidimensional categorization in some situations. It is also possible to support this picture by pointing out that the human brain regions that seem to be implicated in overall similarity classification (e.g. the striatum) have clear analogues in most vertebrates, whilst the human brain regions implicated in single dimension classifications (e.g. prefrontal cortex) do not; see Ashby et al. (1998) for a review of these data. Nevertheless, the tidiness of the picture is illusory, as it ignores a number of facts about both animal and human behavior towards multidimensional stimuli.

We consider first some results from animal discrimination learning which, taken together, suggest that unidimensional categorization may come more easily to animals than making use of overall similarity. Some of these results do not directly concern categorization learning: what they have in common is the implication that when multiple discriminable dimensions are available, animals commonly fasten on only one, or a minority of them, and if that is not an adequate solution to the problem, they may be slow to find such a solution.

The first point is that although animals can learn to discriminate multidimensional stimuli, they do not always do so quickly or easily. The speed with which Herrnstein and Loveland's pigeons learned to discriminate people from non-people is usually taken as surprising, but discriminations based on a single stimulus dimension are learned a great deal faster. Direct comparisons in the literature, with training under the same conditions, are rare simply because the point is so obvious. However, in our laboratory, pigeons will typically take 5-10 80-trial sessions to reach a good level of discrimination between two fairly simple sets of pictorial stimuli (e.g. pictures of cats and dogs on a plain background, as in Ghosh, Lea and Noury, 2004, Table 1), and might take 40 or more such sessions to reach a criterion on a more complex task. Our usual criterion is sustained performance with the p statistic of Herrnstein, Loveland and Cable (1976) at or above 0.8. In contrast, simple discriminations between red and green keys under the same conditions, which we often use in pretraining, are learned to p values of almost 1.0 in one or two sessions. Whilst this may seem obvious, an entirely parallel associative process would predict at least as rapid learning to the multidimensional discrimination as to the single-dimension discrimination (one can even make the prediction that the multidimensional task would be learned faster, see Pearce & Redhead, 1993).

Second, when animals are being trained even in a simple discrimination, in which there is just one stimulus dimension that is perfectly correlated with reinforcement, and all others have no correlation with reinforcement at all, they may nonetheless come strongly under the control of the "wrong" dimension, to the exclusion of the "right" one. The most notorious case is the high salience of location for rats in food-reward tasks, which was a major problem for experimenters throughout the period (from the 1930s to the 1970s) when the usual apparatus for testing discrimination learning was the Lashley jumping stand. Indeed the problem of "position habits" led to Krechevsky's (1932) formulation of a theory according to which animals were only able to attend to one stimulus dimension at a time. Krechevsky's view clearly went too far: Eninger (1953), Mahut (1954) and Turner (1968, as reported in Sutherland & Mackintosh, 1971, p. 95) demonstrated that, even when a rat was in the grip of a position habit as far as its choice responses were concerned, other aspects of its behavior could come under the control of the experimenter-defined relevant dimension. But such control was generally weak and partial. More recently, operant chambers have replaced jumping stands as the apparatus of choice for investigating visual discrimination, and pigeons have replaced rats as the subject species of choice, and as a result there has been more use of successive discriminations. In a simple go/no-go successive discrimination procedure, position habits are not an issue. But experience in our

laboratory is that, even with pigeons in operant chambers, the moment different response locations are brought into an experiment (as in a conditional go-left/go-right discriminations) position biases are likely to become an issue with at least some subjects.

Thirdly, when we examine the dimensions that have gained control over behavior in natural or near-natural concept discriminations, they are frequently rather few, and they control behavior to very different extents. For example, in unpublished analyses of results from the experiments on discrimination of groups of letters reported by Lea and Ryan (1983), where a regression analysis on letter characteristics was carried out to determine what features were correlated with high response rates, we have found that two out of the twelve features that were identified as distinguishing between the two groups (total area of the letter and the height above baseline of its centre of gravity) had much greater impact on performance than the remaining ten. Similarly, Troje, Huber, Loidolt, Aust and Fieder (1999) showed that pigeons' discrimination between male and female faces could largely be explained by the control over behavior exerted by a single dimension, texture; and Jitsumori and Yoshihara (1997) found that, when pigeons discriminated between emotional expressions on human faces, although they used both the available stimulus dimensions (eyes/eyebrows and mouth) to do so, one or the other of the dimensions tended to dominate.

Finally, experiments with artificial polymorphous concepts allow us to control the feature content of stimuli, and therefore allow a more exact determination of the extent to which different stimulus dimensions control behavior. In a series of such experiments, we have found it impossible in many cases to achieve control over behavior by all the stimulus dimensions, and in all cases we have found that different dimensions gained control to very different extents or at very different rates – despite the facts that all the dimensions offered equally valid cues as to the availability of reward, and all dimensions had to be taken into account in order to achieve 100% correct performance. In the experiments of Lea and Harrison (1978) there were three dimensions: sample stimuli are shown in Figure 1. Control by all three was achieved but analysis of errors during training showed that they did not control behavior to the same extent. Similarly, Jitsumori (1993, Experiment 2; see Figure 4 of the paper) found uneven control by the different dimensions when pigeons had reached an overall criterion in a three-dimension task. With more than three dimensions, it has proved almost impossible to get control by all dimensions. Lea, Lohmann and Ryan (1993) used five simple geometrical dimensions (the sample stimuli are shown in Figure 2), and one of them never came to control behavior, despite the fact that when it was used on its own in a simple discrimination, it gained



Figure 1. Examples of the stimuli used by Lea and Harrison (1978). Upper stimuli were used in Experiment I, lower stimuli in Experiment II. The examples illustrate the two values of the three dimensions used: in Experiment I, background hue, superimposed shape, and brightness of superimposed shape; in Experiment 2, hue and shape of the three elements and the orientation in which they were arranged.

control over behavior without particular difficulty. Von Fersen and Lea (1990) used five dimensions that were themselves polymorphous (sample stimuli are shown in Figure 3), and did achieve control by all five, but for several of their subjects they had to use special remedial training with one or more individual dimensions. The experiments cited above all involved pigeons. Lea et al. (2006) carried out an experiment with chickens, using the five-dimensional stimuli illustrated in Figure 4. They found that the hue dimension rapidly acquired control over behavior, but control by the others remained marginal at best, and could only be demonstrated with difficulty. Yet in all these cases, we demonstrated that the birds were capable of discriminating all five of the dimensions if they were presented in isolation, or if they were the only cue correlated with reward.

A reasonable generalization from these results would be that birds rarely if ever use all the dimensions of stimulus variation that are available in a discrimination task, even when taking more dimensions into account would allow higher levels of accuracy; and furthermore, the subset of dimensions that they do use is not necessarily the best available, and may be quite far from the best. Our earlier generalization, that birds categorize multidimensional stimuli

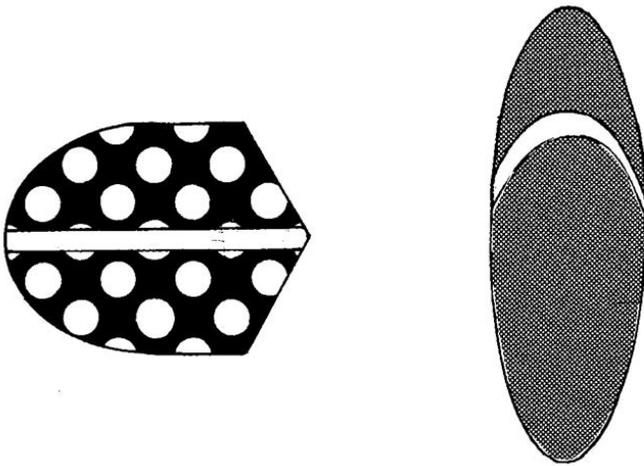


Figure 2. Examples of the stimuli used by Lea, Lohmann and Ryan (1993); the full set is shown in Figure 1 of that paper. Note the five dimensions of these “pseudo-seed” stimuli: orientation of major axis (horizontal/vertical), eccentricity of shape (low/high), size of spots (large/minute), orientation of “split” relative to major axis (parallel/perpendicular), and shape of end (angular/rounded). Less detailed illustrations of these stimuli were included in Figure 1, page 22, “Discrimination of five-dimensional stimuli by pigeons: Limitations of feature analysis”, by S. E. G. Lea, A. Lohmann and C. M. E. Ryan, *Quarterly Journal of Experimental Psychology*, 46B, 19-42. Copyright 1993 by the Experimental Psychology Society. Adapted with permission.



Figure 3. Examples of the stimuli used by Von Fersen & Lea (1990). These are original color versions of the examples sketched in Figure 2 of that paper. Note the five polymorphous dimensions used: The upper picture shows the site Crown and Sceptre, weather sunny, distance near, in orientation horizontal and with camera height aerial, while the lower picture shows the site Northcote House, weather cloudy, distance far, in orientation oblique and with camera height ground level. Photographs taken by Lorenzo Von Fersen. Tracings of these stimuli appeared as Figure 2, page 73, “Category discrimination by pigeons using five polymorphous features”, by L. von Fersen and S. E. G. Lea, *Journal of the Experimental Analysis of Behavior*, 54, 59-64. Copyright 1990 by the Society for the Experimental Analysis of Behavior, Inc. Adapted with permission. Original photographs by Lorenzo von Fersen, used with permission.

in terms of overall similarity, is clearly called into serious question. What about its converse, that humans tend to categorize by a single dimension rule, at least under conditions favoring rule elaboration, with family resemblance sorting occurring only under conditions favoring associative learning?

Some of the best-known evidence for the conventional view comes from what is probably the simplest possible sorting task, the triad task. In this task, three two-dimensional stimuli (call them A, B and C) are simultaneously presented. Stimuli A and B are identical on one dimension but very different on the other. Stimuli B and C are similar but not identical on both dimensions: Figure 5 presents the design schematically. Participants are asked to perform some task such as picking the two that go together. In this design, putting A and B together is a dimensional response, whilst putting B and C together is an overall similarity (family resemblance) response. As noted above, the tendency to group B and C has been shown to increase as a function of numerous variables. Not surprisingly these variables include whether the stimulus dimensions are separable or integral (e.g., Garner, 1974; Handel & Imai, 1972; Lockhead, 1972). According to Garner, stimulus dimensions are integral (as against separable) if it is difficult to attend to one of the dimensions whilst

ignoring the others. For example, color patches that vary in their hue and saturation are (relatively) integral, whilst L-shaped stimuli that differ in the length of their vertical and horizontal bars are (relatively) separable. But other variables also affect performance in the triad task, including procedural factors such as increasing time pressure or concurrent load (Smith & Kemler Nelson, 1984), and age, with younger participants producing more family resemblance sorts than

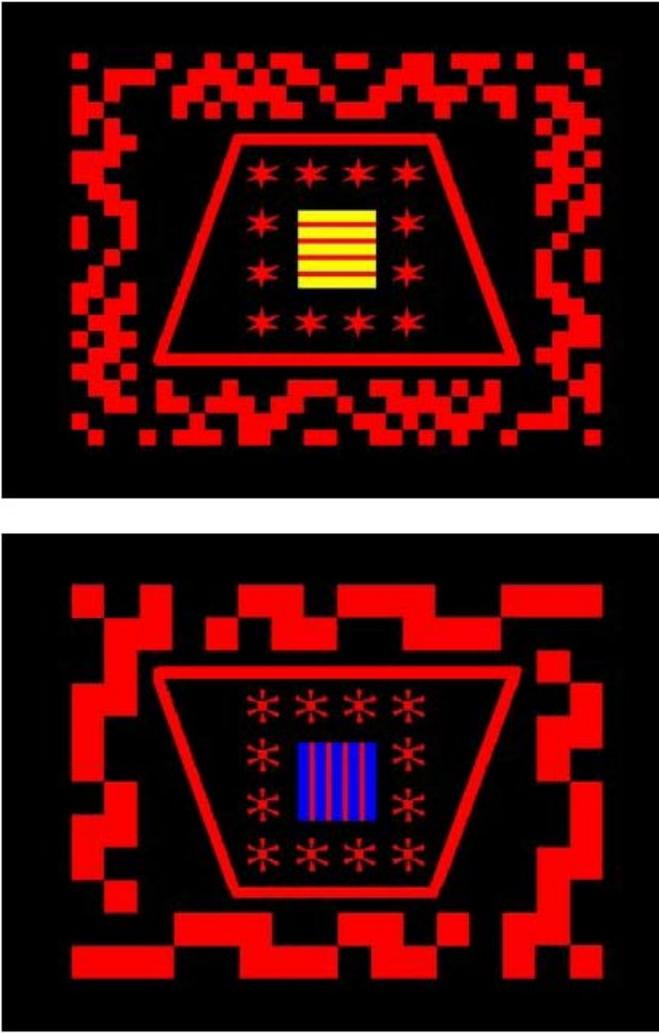


Figure 4. Examples of the stimuli used by Lea et al. (2006). These are full color versions of the examples shown in Figure 1 of that paper. Note the five dimensions used: Stripe orientation on the central square: (horizontal or vertical), Central square background color (yellow or blue), Shapes forming the field around the central square (stars or flower shaped), Trapezium (upright or inverted, giving it a long or short baseline), and Flanker (border) spatial frequency, high or low. Grey-scale versions of these stimuli appeared as Figure 1, page 255, “Why are artificial polymorphous concepts so hard for birds to learn?”, by S. E. G. Lea, A. J. Wills and C. M. E. Ryan, *Quarterly Journal of Experimental Psychology*, 59, 251-267. Adapted with permission.

adults (Smith and Kemler, 1977).

However, over the past few years, results from our laboratory have shown that things are more complex than these classic results might be taken to suggest. We have used the slightly more complex task of match-to-standards. In this procedure, participants are shown two stimuli and told they are exemplars of different categories, and then asked to sort

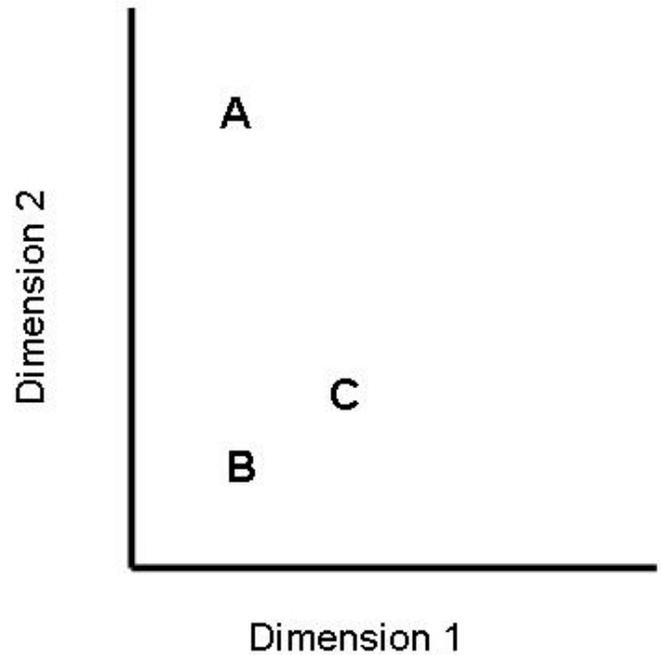


Figure 5. Abstract representation of stimuli in the triad task. The participant is presented with three stimuli, A, B and C. There are two stimulus dimensions (e.g. size and hue). A and B are identical on one dimension but very different on the other. B and C are similar but not identical on both dimensions. Participants are asked which two stimuli are most similar. An “A and B” response is described as a dimensional response; a “B and C” response is described as an overall similarity response. An “A and C” response is described as a haphazard response.

a number of further stimuli (in our case ten) into those categories. Figure 6 illustrates the procedure. Whilst we can replicate the triad-task results under triad task conditions, we also find that under our slightly different conditions (i.e. the match-to-standards task) we can put many of the trends into reverse. Thus increasing the spatial separation of the stimulus dimensions, as illustrated in Figure 6, can increase the tendency to overall similarity sorting instead of decreasing it (Milton & Wills, 2004), and increasing time pressure or adding a concurrent task can reduce the tendency to overall similarity sorting instead of increasing it (Milton et al., 2007). Furthermore, in experiments that are so far unpublished, we find that giving participants instructions to be meticulous and careful increases the tendency to overall similarity sorting, while participants with higher working memory capacity or lower impulsivity are more likely to show overall similarity sorting (Longmore, Milton & Wills, in preparation).

As an aside, it should be noted that although all the stimuli used by Milton & Wills (2004) are likely to be separable by Garner’s definition, separability-integrality is often considered to be a continuum rather than a dichotomy, and it seems

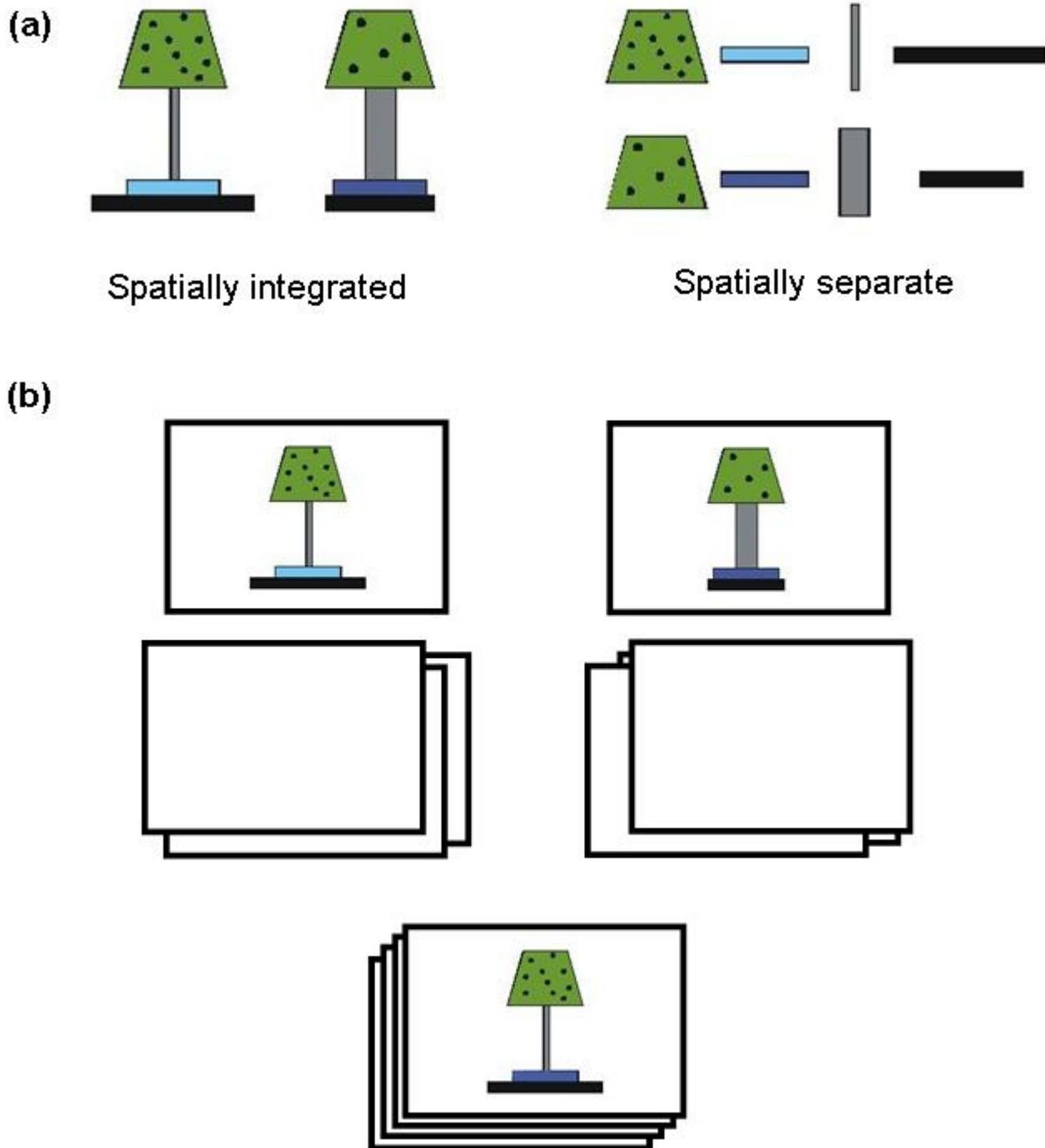


Figure 6. (a) Spatially integrated and spatially separate versions of the ‘lampshade’ stimuli used by Milton and Wills (2004, Experiment 5). (b) Illustration of the paper-based match-to-standards procedure employed by Milton and Wills (2004). The figure shows the subject’s view at the point of making a decision about the 7th of 10 cards. The top two cards are the ‘standards’ (category prototypes) and are placed by the experimenter at the beginning of the task. The bottom pile of cards are sorted sequentially by the participant. The participant does this by placing each card face down underneath the standard they feel it most resembles.

uncontentious to suggest that spatially integrated stimuli are more integral than spatially separate stimuli. One might therefore predict, on the basis of previous results, that the spatially integrated stimuli would evoke more family resemblance sorting. Actually, the opposite happened.

It seems, therefore, that we have pigeons and other birds showing a tendency towards unidimensional sorting, whereas human behavior in conditions that seem to encourage its uniquely human qualities can sometimes be characterized by overall similarity. According to the argument outlined

above that unidimensional sorting is likely to be rule-based and overall similarity sorting is likely to be associatively-based, we would be forced to conclude that pigeons typically elaborate rules, whilst humans, at least when applying their full cognitive resources, can sometimes show associatively-based performance. This is obviously absurd, and it is clear that the simple equation of unidimensional response control with the use of rules cannot be sustained. What can be done to replace it?

Some possible resolutions

There are some obvious steps that can be taken to resolve this apparent paradox. We consider first some ideas that come from the literature on animal learning, that offer alternative interpretations of unidimensional response control.

The first possibility is to focus on the question of salience, and argue that however we try to balance the different dimensions in set of complex stimuli, one or two of them will always be much more discriminable than the rest. It has always been recognized in theories of animal learning that stimuli, or stimulus dimensions, differ in the rate at which they acquire control over behavior, a property that is called salience (Pearce & Hall, 1980; Sutherland & Mackintosh, 1971, p. 127ff). So rats' position habits, or the dominance of luminosity in pigeons' letter set discrimination, or chickens' failure to discriminate anything but color cues, are explained by saying that the cues in question have high salience – a conclusion we have ourselves drawn in a previous paper (Lea et al., 2006).

An explanation of unidimensional discrimination of multidimensional stimuli in terms of salience has some weaknesses. First, unless we make an independent assessment of the salience of each dimension, it does no more than label the phenomenon. Although possible, such independent assessment is seldom done. Second, when the most salient cue has only partial validity and other cues are independently and equally valid (as in artificial polymorphous discriminations), many learning theories predict that even the least salient cues will eventually acquire some control over behavior (though they will do so more slowly than cues of equal validity but greater salience). This is because learning in many contemporary theories is driven by prediction error, and therefore any available cue that can improve performance (reduce prediction error) will gain some control over responding.

A second possible explanation is limited attentional capacity: the idea that animals can only process a certain number of stimulus dimensions at any one time. This is the essence of theories such as those of Krechevsky (1932) or Sutherland and Mackintosh (1971), and Lea et al. (2006) concluded from an experimental analysis that it was one of the strongest factors behind the difficulty birds have in learning

artificial polymorphous category discriminations.

A third possibility is to recognize that, while the experimenter may be aiming for 100% correct discrimination performance, the animals in the experiment have no concept of that. In this they differ from human experimental participants, who can be given explicit instructions that they are faced with a problem that has a perfect solution. A position habit in a jumping stand will bring a rat reward on 50% of trials; how is the animal to know that a higher rate of reward is available? Experimenters commonly recognize this difficulty by giving animals non-choice trials in spatial choice apparatuses like jumping stands. On such trials, only the non-preferred alternative is available, so as to expose the subject both to the possibility of making that response and to its reinforcement. However, it is not clear that this account fully explains the dominance of single dimensions in experiments such as those of Lea et al. (2006). In a three-out-of-five artificial polymorphous concept discriminations of the sort used by Lea et al., control by a single dimension will only lead to correct response to 22 of the 32 possible stimuli, but it is common practice to precede the experimental discrimination by a simple, "practice" discrimination in which subjects quickly achieve nearly 100% accuracy. This comes as close as one can to telling the animal that it could do better than rely on a single dimension.

Some further ways of resolving the problem are rooted in research on human cognition. The recent work from our laboratory (Milton & Wills, 2004; Milton, Longmore & Wills, 2007; Longmore, et al., in preparation) indicates that one should not simply equate single dimension responding in humans with rule use. Single dimension rules are probably the most easily accessible and easily verbalized of rules (e.g. "red vs. blue") and, hence, it is perhaps not surprising that they are a characteristic output of a rule-based process, even if not a defining one. Nevertheless, humans clearly have the capacity to formulate and verbalize more complex rules (e.g. that a prime number is a number divisible only by one and itself). Perhaps what our recent human results show is that there are conditions which promote (and/or inhibit) the development and application of these more complex rules. In a match-to-standards task, one rule that produces overall similarity responding is "Count the number of category A features in the stimulus, and call it #A. Now count the number of category B features in the stimulus, and call that #B. If #A > #B, place the stimulus into category A. Otherwise, place it into category B". In post-experimental discussion, participants in our match-to-standard task who produce family resemblance sorts can often report an explicit dimensional summation rule of this type.

The idea that multidimensional rules are possible, but typically have a low probability of being employed, is an ex-

PLICIT part of certain multi-process accounts of categorization in humans (e.g. Ashby, Alfonso-Reese, Turken & Waldron, 1998; Nosofsky et al., 1994). That said, some formal simulations of such models assume that the probability of multidimensional rules is actually zero (e.g. Ashby et al., 1998), and some approaches to the ‘rules vs. similarity’ question appear to assume that rules are, definitionally, things that operate on a relatively small subset of the available information (Pothos, 2005). Neither approach seems consistent with our recent work.

There still remains the question of how one can reconcile the varied results in the human literature. For example, why does time pressure sometimes increase, and sometimes decrease, the prevalence of family resemblance sorting? In a recent paper (Milton et al., 2007), we demonstrated, within a single experiment and procedure, that changes in time pressure can have either effect – the direction of the effect seems to depend on the absolute level of time pressure. This result is consistent with the idea that family resemblance sorting can arise from two quite distinct processes. At high levels of time pressure, perhaps participants are treating the stimuli essentially as ‘blobs’, and this leads to family resemblance sorting. As time pressure decreases somewhat, they are able to start to apply rules, but only simple ones – and hence unidimensional responding dominates. As time pressure decreases further then, if the other conditions (e.g. the experimental procedure) are conducive, participants start considering more complex rules, such as the dimensional summation rule described above.

An empirical approach

Both the data and the theoretical approaches we have summarized here suggest that it is premature to draw the generalization that, faced with the task of responding categorically to multiple, multi-dimensional stimuli, humans (sometimes) elaborate single-dimension rules while other animals use multi-dimensional family resemblances. At the same time the data – and common sense – do suggest that there are different cognitive processes at work in humans and other animals, and to gain a proper comparative understanding of this aspect of cognition, it is important to characterize the differences accurately.

At least in the simpler situations we are interested in, there are already parallel experiments in the human and animal categorization literatures. However, there are usually multiple procedural differences between experiments on humans and other species, and this makes it hard to interpret any differences in results; in other situations, we have found apparent cross-species differences sometimes dissolve when such procedural differences are removed (Goto, Wills & Lea, 2004). Accordingly, we have embarked on a series of

experiments in which we have exposed humans and pigeons to multi-dimensional stimuli in, as far as can reasonably be achieved, identical conditions. Because the animal laboratory is more constraining than human testing, our general strategy has been to develop experimental procedures using pigeons, and then once we have established a robust paradigm, devise the closest human analogue we can.

In most of our experiments, we use spatially separate stimulus dimensions, in which the different dimensions are located in different positions. Figure 7 shows an example of the stimulus elements we have used. In each stimulus, there are four elements, which can take either of two values on a particular stimulus dimension; we called the elements in the Figure 7 stimuli the “bar”, “doughnut”, “checks” and “lozenge”. The advantage of spatially separate stimulus dimensions is that the impact of each dimension can be tested in the complete absence of interference from, or interaction with, any other dimension – dimensions not of interest in any particular trial can simply be omitted. With some other types of stimuli, all that can be done is to give them a putatively neutral value, and such a value can only be determined with the help of questionable assumptions. The psychological midpoint between two stimuli does not necessarily lie at their physical midpoint (e.g. Emmerton & Renner, 2006; Matheson, Asher & Bateson, 2007); and a previously unused value on a continuum could conceivably elicit a startle response even though it lies within a continuum.

The stimuli shown in Figure 7 have another design feature. To maximize the chances of getting multi-dimensional stimulus control in the pigeons, we avoided two features which previous experiments indicate frequently dominate pigeons’ behavior in multidimensional tasks, namely color and overall black or white area (which may well be mediated by the total luminous flux from the stimulus). All eight elements (i.e. the two values on each of the four dimensions) consisted of white shapes on a black background, and the overall background on which the elements appeared was also black. The number of white pixels in each element was near-identical, the variation being less than 1%.

Although our approach to experimental details has been to design an effective pigeon experiment and then mimic it for humans, the general situation was chosen by reference to the human literature. Most of the relevant human experiments are framed as categorization tasks, in which there is a positive, correct response that can be made to each stimulus, and the two (or more) categories that are involved all have identical association with reward. Much of the animal literature, in contrast, has been based on go/no-go discrimination procedures, where for one category there is no correct response that can be made. This asymmetry between categories has the potential to mislead, particularly when novel test stimuli

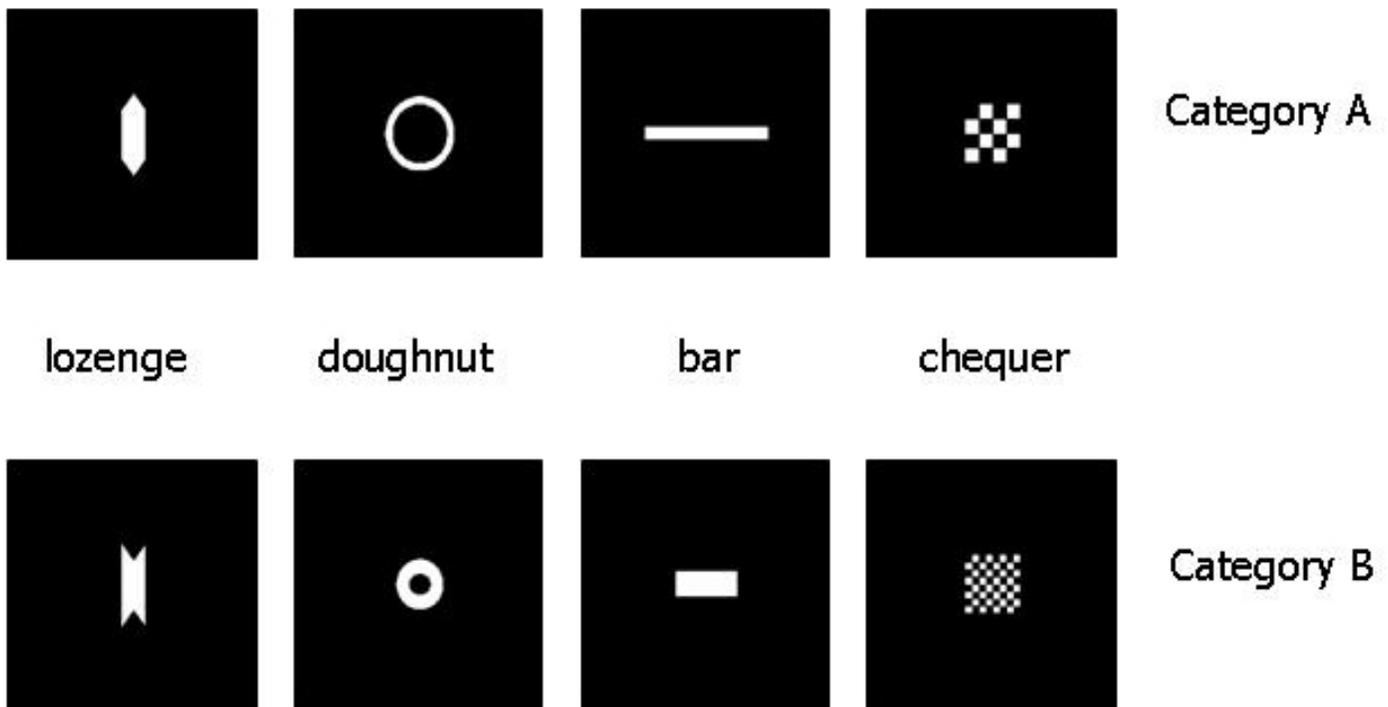


Figure 7. Elements of the stimuli used in the first phase of training of the first pair of experiments

are introduced: pigeons' response to an unknown stimulus may either be to treat it as positive (most go/no-go discrimination learning takes the form of reduction in the rate of response to negative stimulus), or to avoid it (a form of stimulus neophobia). Accordingly, in the pigeon experiments we used conditional (go-left/go-right) discriminations. In pilot work we found that acquisition of such discriminations could be very slow, so to speed learning we used a differential outcomes procedure of the sort introduced by Sheldon (1967) and Overmier, Bull and Trapold (1971).

In this review, we will give a brief account of two pairs of experiments we have conducted within this general approach. These experiments are currently unpublished, and full experimental details are not given here.

Simple categorization of multidimensional stimuli

The first pair of experiments (one involving pigeons and the other involving humans) involved a simple discrimination between stimuli based on those shown in Figure 7. Following our general procedure, we carried out the pigeon experiment first. One obvious way in which pigeon and human experiments are likely to differ is that we know in advance that all the dimensions we put into stimuli are going to be reasonably obvious to humans, because we as experimenters are also humans, and we have chosen them as obvious to us. They may or may not be obvious to subjects of another species. So we set out to ensure that the birds were discriminat-

ing all four dimensions of the stimuli in Figure 7, by giving initial training in which we presented only one element at a time, in a random sequence, as in the "Features in parallel" procedure of Lea et al. (2006). This procedure should eliminate extreme versions of a differential salience account of any unidimensional control over behavior in subsequent stages of the experiment (for example, the possibility that certain dimensions are psychophysically indistinguishable). It turned out that it was difficult to establish reliable performance on all dimensions, and only one out of the eight birds we used reached a criterion of 75% correct on each dimension within a successive session. We therefore relaxed our criterion, and trained the remaining birds until they reached a weaker criterion of 75% correct on at least 3 dimensions in a session; all birds did reach this level, though they took between 25 and 30 sessions to do so.

Our initial intention had been to then continue with a second phase of training using stimuli in which all four dimensions were present on each trial. Given that most birds only mastered three dimensions, we instead used stimuli containing three of the elements shown in Figure 7, arranged in a triangle; for each bird we used the three dimensions that it had mastered in phase 1. Examples of such stimuli are included in Figure 8. Once a criterion of accurate performance on these three-element stimuli had been reached, there were tests in which we used the six possible "one-away" stimuli – the stimuli formed by reversing the value of just one of the

Prototype stimuli

'One-aways'

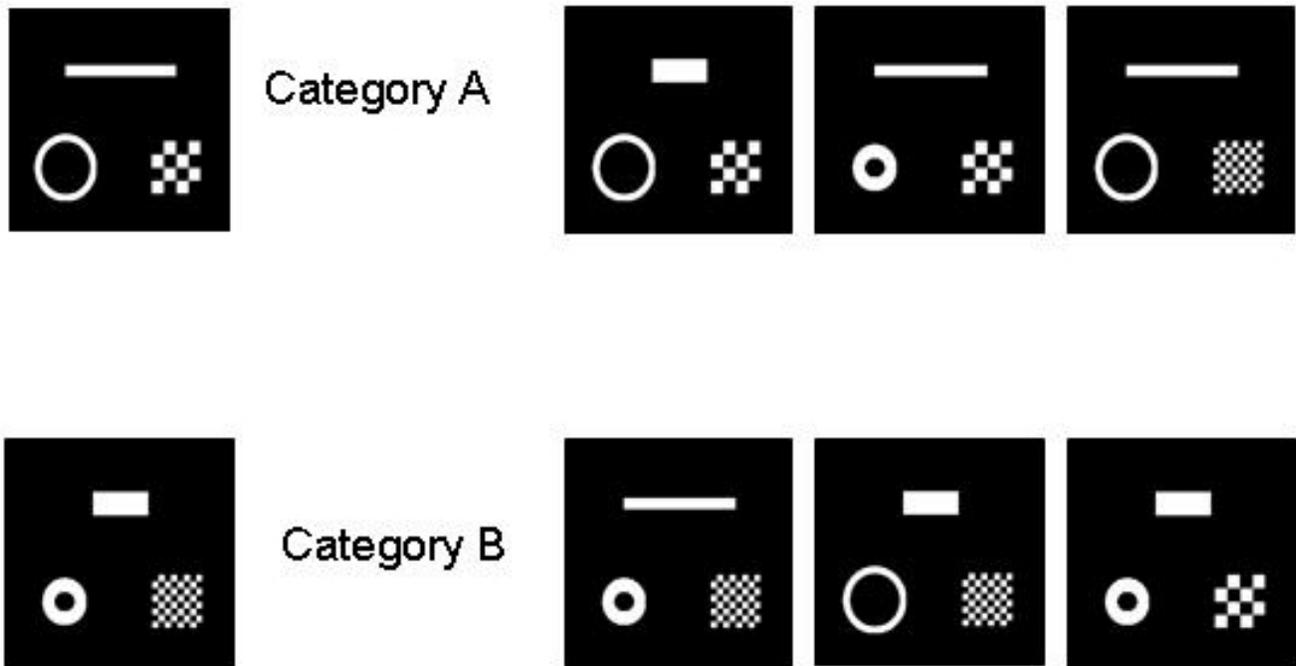


Figure 8. Stimuli used in the second training and test phases of the first pair of experiments.

three dimensions in the stimulus (see Figure 8). This test procedure is a direct descendant of that used in a classic experiment by Reynolds (1961), in which pigeons were trained to discriminate two stimuli that differed in both color and a superimposed shape, and were then given tests in which the color and shape cues were put into conflict. With more than two dimensions, categorization of such cue-conflict stimuli can be used as a sensitive test of whether unidimensional or family resemblance sorting is being used (e.g. Medin, Wattenmaker & Hampson, 1987; Regehr & Brooks, 1995; Milton & Wills, 2004). If unidimensional sorting is being used, then all and only the stimuli that have a particular value on one of the dimensions will be placed into one of the categories, for example all the stimuli with a short fat bar in Figure 7; if family resemblance sorting is in use, then the stimuli placed into a category will be those in which most of the dimensions take the same value as the sample or training stimulus, with no one dimension being privileged, for example all the stimuli in the upper row of Figure 7.

Both pigeons and humans saw the stimuli on a computer screen of similar dimensions. The pigeons categorized the stimuli by responding directly to one of two choice zones on the screen, while the humans responded by pressing keys on the keyboard with a roughly similar spatial relationship to the stimulus zone. In training, choice responses were either correct or incorrect, and correct responses were followed

by appropriate reward (grain for the pigeons, and presentation of a schematic smiley face associated with a promise of a small money payment or course credit for the humans), while incorrect responses resulted in the screen going blank and the next trial being offered after a suitable interval. In the test trials, reward was given regardless of the choice response made, to avoid biasing the subjects. Intervals between events, and response requirements, were, in so far as it was possible, the same for the humans and the pigeons, with two unavoidable exceptions. First, the inter-trial interval was much longer for the pigeons than for the humans: long inter-trial intervals help to maintain responding in animals, presumably through Pavlovian processes (cf. Balsam & Payne, 1979), while in our experience they are likely to lead to inaccurate responding and outright rebellion among human participants. Secondly, as noted above, the pigeons required numerous 80-trial sessions to reach a good level of performance in the first stage of the task, while the human participants required either two or three blocks of 40 trials. The criterion used, however, was essentially the same for both species – sustained performance at or above a level of 75% correct on at least three of the four available dimensions. Unlike the pigeons, however, the humans typically comfortably exceeded this target by the end of training, and discriminated all four dimensions correctly. However, to maintain comparability of procedures, we chose each participant's best three dimensions for use in the next phase, as we

had with the pigeons.

On the basis of their responses to the one-away stimuli, only three out of the twelve human participants could be classified unambiguously as unidimensional sorters, since all or nearly all their test trial responses could be predicted by the value of a single stimulus dimension. Eight of the remaining people clearly sorted by family resemblance, since nearly all their test trial responses could be predicted by counting the number of stimulus dimensions that indicated a left or right response. Under the conditions of the present experiment, therefore, human participants show the sort of result that has been obtained from matching to sample procedures in the past (e.g. Milton & Wills, 2004). The pigeons presented a more complicated picture, partly because their overall level of performance was further from 100% so the distinction between different sorting strategies is harder to make. We therefore used logistic regression to see how well the three stimulus dimensions predicted the direction of response on test trials. One bird showed a very strong position bias in the test trials, and while this is evidence of a unidimensional strategy of a kind, it did not allow us to examine the relation between the stimulus dimensions and responding, so this bird was dropped from the analysis. Of the remaining birds, one showed significant evidence that all three stimulus dimensions controlled behavior, and so could be classified as a true overall similarity sorter, and two others showed significant evidence that two of the three dimensions influenced behavior. For the remaining four birds, only one of the dimensions (not the same dimension for all birds) was significantly related to choice. It would be tempting to classify these birds as unidimensional sorters, but that might be misleading. Human unidimensional sorters tend to make 100% of their choices on test trials consistent with the single dimensions that governs their behavior. Our pigeons did not, and examination of their behavior on trials when their choice was not consistent with the dominant dimension showed that there were non-significant tendencies for the other dimensions to influence them. So if we had been able to give very large numbers of test trials (not a feasible strategy as the birds would have learned that reward on test trials was non-contingent), it may well be that even the apparently unidimensional sorters would have shown control by multiple dimensions.

From this pair of experiments, therefore, it is clear that our initial generalization does not stand up when humans and pigeons are tested under comparable conditions. Our pigeons did not seem any less likely to show unidimensional sorting than humans. And they are less likely than humans to show clear family resemblance sorting. Typically, their categorization of ambiguous stimuli tends to be strongly influenced by one or two dimensions at the expense of others available; and this is despite the fact that they were pretrained to dis-

criminate each dimension up to a criterion.

One reason decision-making?

The second pair of experiments used a situation where one might expect even pigeons to be pushed towards true unidimensional sorting. Gigerenzer and his colleagues (e.g. Gigerenzer and Goldstein, 1996) have argued strongly that categorizing objects on the basis of a weighted sum of multiple features is maladaptively slow and inefficient, and instead people should and do rely on a strategy of “one-reason decision-making”: identifying one dimension that does a good enough job of categorizing the objects, and ignoring everything else. This is often referred to as the “Take The Best” strategy, and has been extensively investigated (e.g. Broder & Schiffer, 2003; Martignon & Hoffrage, 2002; Newell & Shanks, 2003).

Once again we started by finding a procedure that was effective for pigeons, and then conducted an analogous experiment with humans. The basic procedure, of a conditional discrimination with differential outcomes, and no scheduled consequences for wrong responses on the choice keys, was the same as in the previous pair of experiments, and the stimuli we used were essentially the same, except that the features on the “lozenge” dimension were exaggerated (see Figure 9), since this dimension exerted least control over the behavior of the pigeons in the experiment above. The stimuli had the same design properties as before, avoiding color and overall flux as dimensions, and ensuring that all eight elements had the same pixel count. In this experiment, all four elements of the stimulus were presented together in the majority of sessions: we called these “Prototype” sessions. One session in three, however was run under a different condition, which we called “Wait to Reveal”. In this condition, each trial began with the exposure of a randomly-chosen single element of the stimulus, plus the choice keys. If the pigeon waited for 2.5s before pecking either choice key, another element appeared, and so on until all four elements were present. But once the pigeon had pecked either choice key, no further elements would appear, though the elements already exposed remained on view. Given what is known about pigeons’ inability to choose to delay gratification when immediate reward is available (e.g. Ainslie, 1974), we were concerned that the pigeons would always peck the choice keys at an early stage if this led to immediate reinforcement. Accordingly, the schedule of reinforcement on the correct choice key (in both Prototype and Wait to Reveal conditions) was a tandem FT10s-VI2s, so there was no advantage to the pigeon in making an early choice. However we expected that if the birds had learned to use a single dimension, they might make a choice response as soon as a feature of that dimension appeared.

The results of this procedure were curious. The birds did learn the discrimination within a reasonable time (all but one bird had reached around 80% correct within 40 sessions, the remaining bird requiring a further 12 sessions) and in the Wait to Reveal condition they did make their initial choice

before all the elements of the stimuli were present. Most strikingly, however, their performance under the Wait to Reveal condition tended to be worse than under Prototype conditions, particularly earlier in training, and as training progressed, they tended to wait for more elements to appear

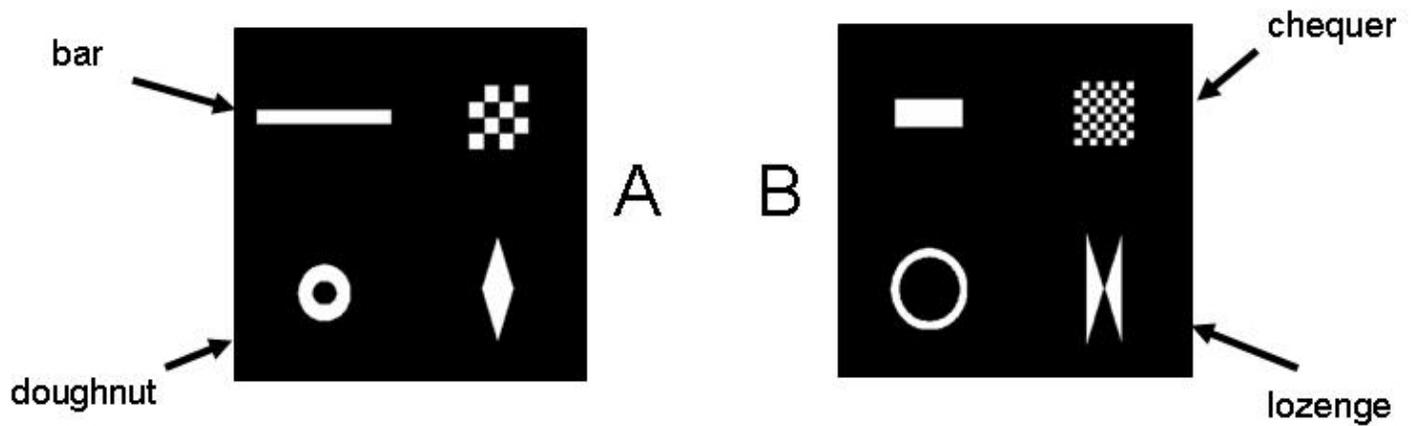


Figure 9. Stimuli used in the second pair of experiments (Prototype forms)

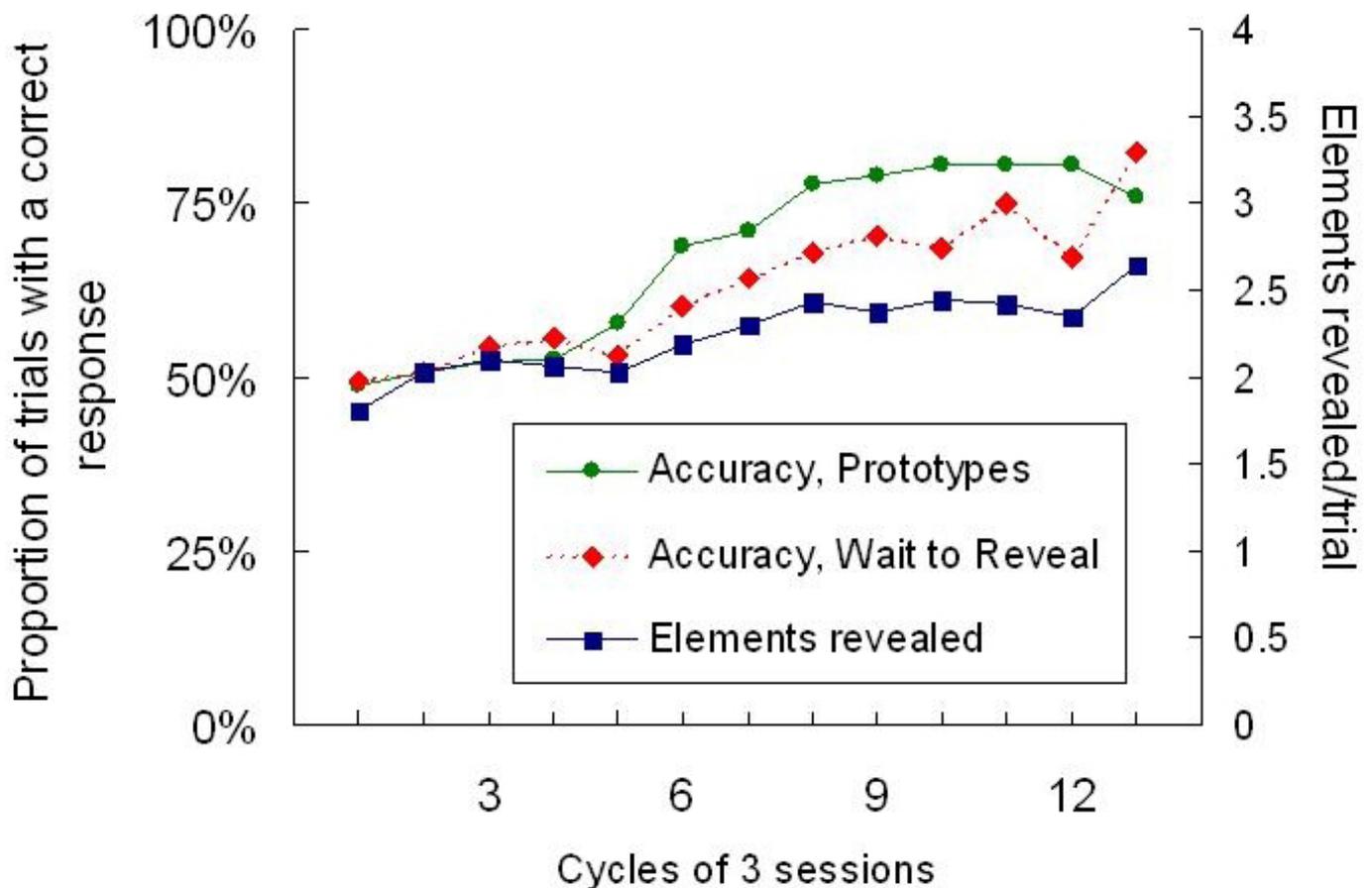


Figure 10. Course of acquisition of discrimination by the pigeons in the second pair of experiments. Sessions occurred in blocks of three including two in “Prototype” conditions and one in “Wait to Reveal” conditions. All data are means across the 6 birds in the experiment.

(see Figure 10).

The results might be taken to imply that the birds needed to use all dimensions of the stimuli in making an accurate choice, and what they had to learn in the Wait to Reveal condition was to wait until all elements had been revealed. That was not the case, however. Because the birds did not always wait for all four elements had been revealed, and because order in which the dimensions were revealed varied from trial to trial, it was possible to analyze how their performance in the Wait to Reveal sessions was affected by the presence or absence of each dimensions. Figure 11 summarizes such an analysis, taken over the final five cycles of three sessions (each cycle included one Wait to Reveal session). In this analysis, we used logistic regression to assess how well the presence or absence of each dimension at the time of choice predicted whether the choice response would be correct or not. The dependent variable was the bird's initial choice, to left or right; the independent variables were whether or not each dimension was present at the time of choice. The session cycle was included as an additional independent variable, to remove variance due to any continuing increase in accuracy of performance. Note that this analysis treats

the dimensions as additive, and so assumes that there are no compound rules such as “peck right if the bar is long and the doughnut is thin”. It can be seen that the dimensions varied greatly in their control over the choice response, and we subsequently confirmed this conclusion by running trials with the one-away stimuli, as in the previous experiments, once the criterion of learning had been reached. Despite these differences in control between dimensions, however, the birds did not learn to wait for the dimensions that were most helpful to them. Figure 12 shows what information the birds had at the time they made their initial choice response: it reports the probability that each dimension was present at the time of choice. In contrast to Figure 11, there is absolutely no variation between dimensions. What this means is that the pigeons' timing of their choice responding was essentially independent of how much of the stimulus had appeared. In other words, their waiting was completely unselective between stimulus dimensions, despite the fact that the dimensions exerted selective control over their behavior.

Clearly the Wait to Reveal condition is a rich one, and it opens up a number of possibilities for both experimental and theoretical analysis, which we are currently exploring in fur-

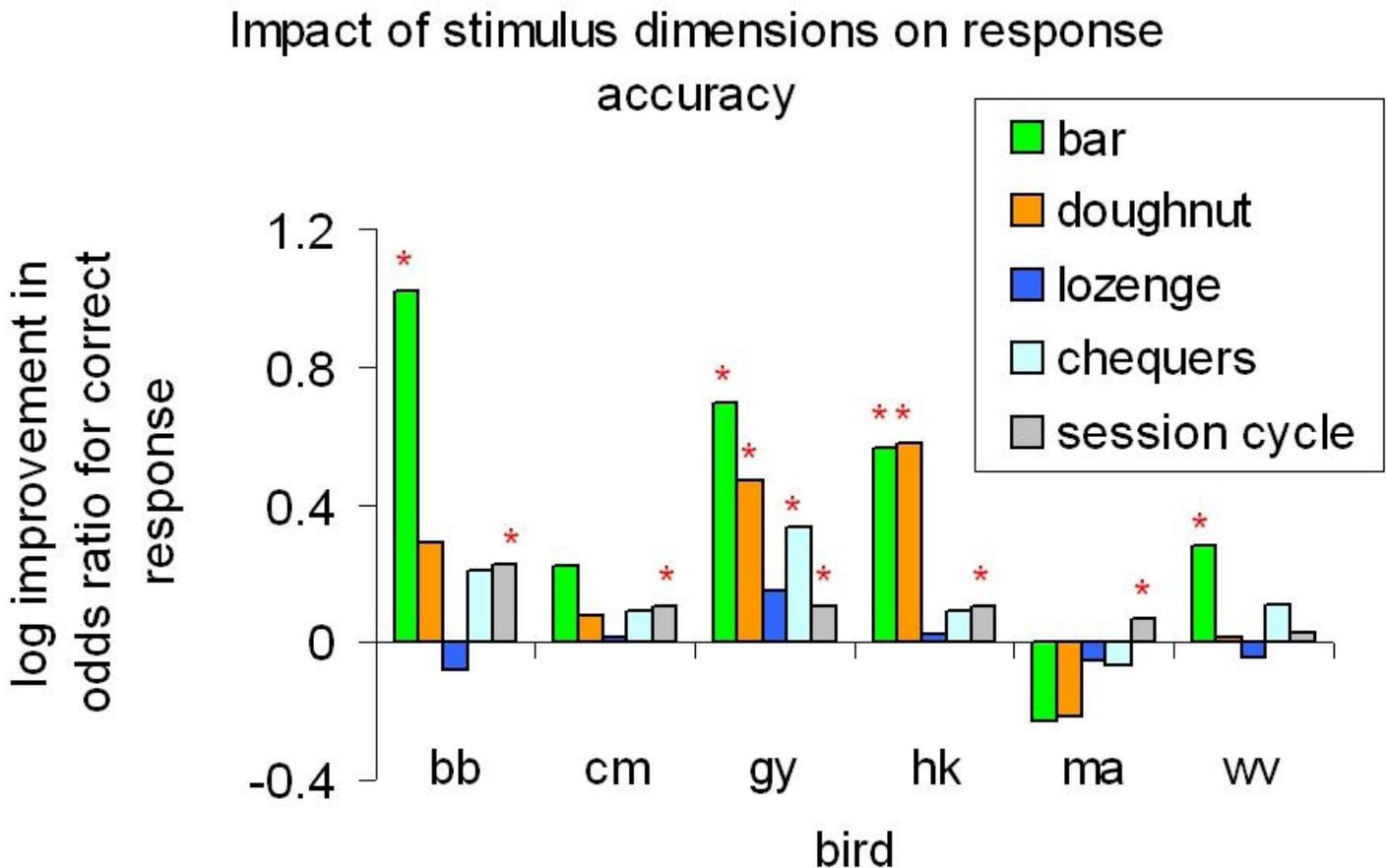


Figure 11. “Wait to Reveal” procedure with pigeons: Impact of the presence of the four stimulus dimensions on choice accuracy, assuming additive combination of cue information.

* $P < 0.05$; ** $P < 0.01$

Stimulus information available at time of first choice peck

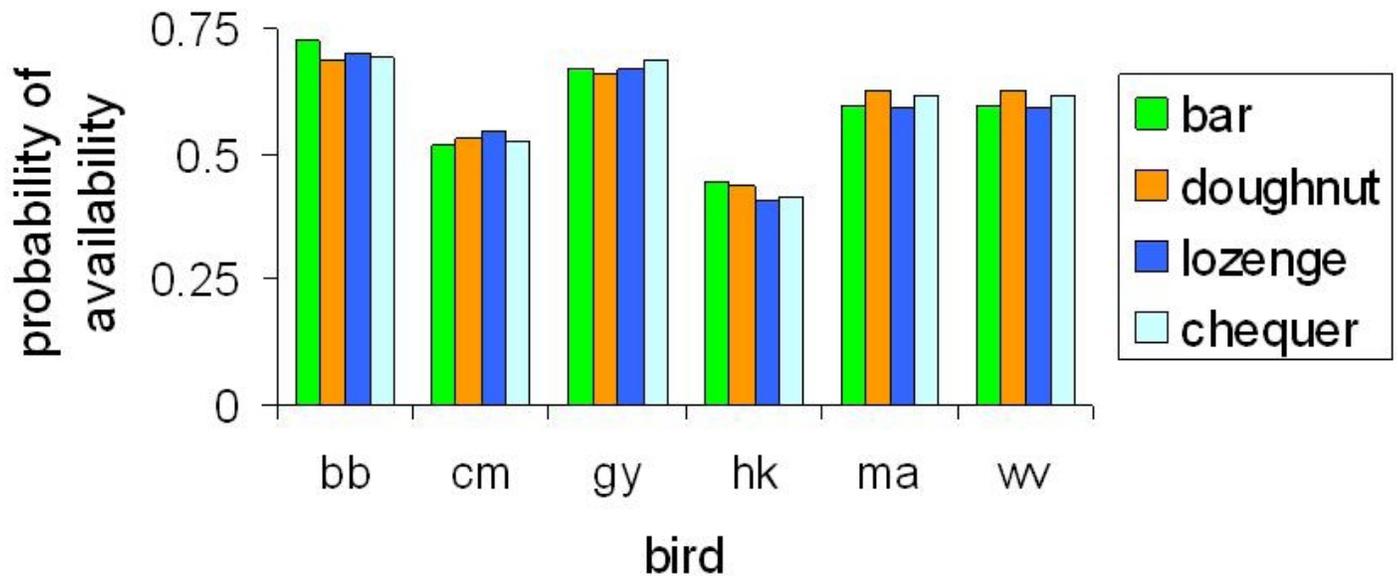


Figure 12. “Wait to Reveal” procedure with pigeons: Probability of each stimulus dimension being present at the point where the bird first pecks a choice key, thus preventing the appearance of further cues.

ther experiments. In particular its relation to impulse control (Ainslie, 1974) poses some challenging questions. Here we have considered only some relatively simple implications of the results, focusing on what they can tell us about the discrimination of multidimensional stimuli.

How do human participants get on under these conditions? The experiment was very similar except that, because humans learn this kind of task so much faster than pigeons, we carried out the entire experiment in a single session. To give the human participants the same kind of experience as the pigeons, of Prototype and Wait to Reveal procedures occurring throughout learning, in the trial sequence for humans two Prototype trials were followed by a single Wait to Reveal trial. In effect, therefore, each session of pigeon training was replaced by a single trial of human training. The timing of trials was essentially the same as for the pigeons, except for the usual reduction of the inter-trial interval. Once again the humans behaved quite differently from the pigeons. On the Wait to Reveal trials, although they occasionally made a choice before all the elements of the stimulus had arrived, they did so so rarely (and they learned so quickly) that it was impossible to analyze what information they were using on these trials. However, after criterion had been reached, we gave “one-away” test trials as in the previous pair of experiments. Out of eight people we have tested so far, five showed

unambiguous unidimensional sorting in these tests (using between then all four of the dimensions), two showed overall similarity sorting, and one used a mixed strategy. Looking back at the Wait to Reveal trial outcomes, therefore, we can see that the unidimensional sorters among the human participants showed a pattern that was almost exactly opposite to the pigeons. Whereas the pigeons really needed to wait selectively for the most useful information to arrive, but did not, the humans could have responded the moment their preferred dimension arrived, but instead waited. If we were to accept the classic use of unidimensional sorting as a marker for rule formation, we could say that the human subjects showed more evidence of rule formation than the pigeons, which is a sensible outcome; but even on that analysis, they did not seem to deploy their rules intelligently in order to make a prompt choice response. Very possibly, of course, they were not as motivated to get immediate reward as the pigeons, because although humans like other animals show excess preference for immediate reward, the phenomenon is much more acute in pigeons (Ainslie, 1975).

Associations, Rules and Verbalizations

We opened this paper with the idea that using a unidimensional strategy when sorting multidimensional stimuli might be a reliable sign of rule-based cognitive processing, which

only humans should be capable of. Sorting on the basis of overall similarity, on the other hand, would indicate associationistic processes, of the sort that humans are thought to share with other animals species. We challenged that view by showing that non-humans often come under the control of single dimensions in multidimensional tasks, while humans sometimes use multiple dimensions in conditions that seem to encourage complex cognitive processing. We sought to resolve this paradox by comparing the behavior of humans and pigeons in more closely comparable experimental conditions than have been used before. Given the results of those experiments, how do we now stand on the issues of the use of single or multiple dimensions by humans and non-humans, and their interpretation?

Before we seek to interpret our results, we should return to a caveat that we raised earlier. Although in both pairs of experiments we tested humans and pigeons under closely comparable conditions, we could not make the conditions identical. Indeed, it is never possible to know that we have made experimental conditions for different species identical, for all the reasons rehearsed long ago, for example by Bitterman (1960). Furthermore, there was one striking difference between the experiences of our pigeon subjects and human participants: the pigeons required several weeks and thousands of trials to reach the standard of performance we required for our tests, whereas the humans' training was complete in a single session. Amount of training certainly affects the representations that are formed of complex stimuli, in both pigeons and humans (e.g. Cook & Smith, 2006). Our view would be that the use of a common criterion of performance before moving on to testing should mean that at the point of test, learning has reached the same stage, as closely as we can make it, in both species. However, this is a matter that certainly needs more investigation in the future.

With that caveat, however, our results give no support to the classic position that we put forward at the beginning of this paper. In the first experiment, if anything fewer humans than pigeons sorted one-away stimuli in accordance with a one-dimensional rule. In the second pair of experiments, humans were somewhat more likely to show unidimensional stimuli, but the effect was far from overwhelming. Furthermore, our results do not support a strong overall-similarity account of pigeon categorization: in our first experiment, only one out of eight birds appeared to use all three dimensions, and in our second experiment, Figure 10 suggests that none of them did so (note, however, that this is a different test than we used in the first experiment). It is also true that training with all the elements of the stimuli separately, which happened explicitly in our first experiment, and in a less systematic way as a result of the Wait to Reveal in the second experiment, may have biased the participants towards using more information, and thus towards overall similarity cate-

gorization. However, as both birds and humans experienced the same single-element training, it seems unlikely that this possible bias directly underlies any species differences.

At least some of our pigeons, in at least some tests, seemed to be using single dimensions. How should we interpret this behavior? According to criteria sometimes suggested for use in human cognition by Pothos (2005) and others, it is evidence for the use of a rule rather than merely associative similarity-based processes. But if we are to describe such behavior in pigeons as rule-governed we have to give "rule" some meaning other than a verbal statement, or even a statement that the participant could verbalize. What could a non-verbal rule be like?

One way of giving meaning to the idea of a rule in a non-verbal animal is to say that a rule does not just govern a particular behavior in a particular context, but is available for other purposes in other contexts. This fits with the way we assess verbal rules: having observed that a human participant's behavior is consistent with a certain rule, we ask them questions about it, and find they can describe the rule they were using. So, as well as governing behavior within the experiment, the rule is available in the different context of the post-experimental interview, for the different purpose of satisfying the experimenter's curiosity. By this criterion, the pigeons in our second pair of experiments miserably failed to show any evidence of rule use. As Figure 10 shows, their behavior when choosing which side key to peck was strongly governed by one or two of the four available dimensions; but as Figure 11 shows, if this was the result of a rule, that rule was not available to govern even the closely related behavior of choosing whether or not to wait for relevant elements to be exposed. However, by the same criterion, our human subjects did not make very intelligent use of rules either. The one-away trials in the second human experiment suggested that most of them did use one-dimension rules to govern choice. But like the pigeons, they did not seem to apply these rules when it came to deciding when to make their choices.

The task of using behavioral criteria to decide when a subject is using a rule is a difficult one. With human participants, we can always investigate after an experiment whether someone could verbalize a rule, but it does not automatically follow that they were using the same rule within the experiment. If we are to continue to use the unidimensional sorting criterion to establish when humans are using rules, we will have to recognize that some of the rules concerned must be supposed to have very limited scope, and will not be applied as they logically could be to govern other behaviors within the experiment.

Might a different sort of account have any better success

at explaining the results we have discussed? One possibility would be to try to construct an associative model that could embrace both the human and bird results. Within such an account, it would be possible to parameterize the capacity for broad attention to multiple cues. We could then express the hypothesis that birds have difficulty in maintaining broad attention across stimulus dimensions, as concluded by Lea et al. (2006) following an experimental analysis of the origins of the difficulty of training artificial polymorphous category discriminations in chickens in terms of an interspecific difference in this parameter. For such a model to be adequate, it would have to be possible to find a parameter setting that would reproduce the results of our first pair of experiments, in which even though the pigeons were successfully trained to use all three cues that subsequently appeared in tests, they apparently could not maintain their attention to them all in tests.

The development of such a model lies in the future. Conceptually, however, it lies at the heart of what we have been investigating in this paper. There are undoubtedly differences between the way birds and humans learn about and classify multidimensional stimuli. But are those differences quantitative or qualitative? Specifically, are there processes that are available to humans that are not available to non-humans, or at any rate to birds? We do not claim to have provided a definitive answer in this paper, but we hope that we have shown two things: methodologically, that understanding the differences between human and avian categorization requires comparative experiments that put the two species more closely to the same test than has been the custom in the past; and substantively, that it is no longer sustainable to simply identify unidimensional sorting with a kind of verbal rule elaboration that is only available to humans.

References

- Ainslie, G. (1974). Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior*, *21*, 485-489.
- Ainslie, G. W. (1975). Specious reward: A behavioral theory of impulsiveness and impulse controls. *Psychological Bulletin*, *82*, 463-496.
- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A formal neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442-481.
- Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, *61*, 1178-1199.
- Balsam, P. D., & Payne, D. (1979). Inter-trial interval and unconditioned stimulus durations in autoshaping. *Animal Learning & Behavior*, *7*, 477-482.
- Beckers, T., Miller, R. R., De Houwer, J., & Urushihara, K. (2006). Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference. *Journal of Experimental Psychology: General*, *135*, 92-102.
- Bitterman, M. E. (1960). Toward a comparative psychology of learning. *American Psychologist*, *15*, 704-712.
- Broder, A., & Schiffer, S. (2003). Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, *132*, 277-293.
- Cook, R. G., & Smith, D. J. (2006). Stages of abstraction and exemplar memorization in pigeon category learning. *Psychological Science*, *17*, 1059-1067.
- D'Amato, M. R., & Van Sant, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 43-55.
- Dennis, I., Hampton, J., & Lea, S. (1973). New problem in concept formation. *Nature*, *243*, 101-102.
- Emmerton, J., & Renner, J. C. (2006). Scalar effects in the visual discrimination of numerosity by pigeons. *Learning & Behavior*, *34*, 176-192.
- Eninger, M. U. (1953). The role of generalised approach and avoidance tendencies in brightness discrimination. *Journal of Comparative and Physiological Psychology*, *45*, 604-608.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107-140.
- Ferreira, G., Keller, M., Saint-Dizier, H., Perrin, G., & Levy, F. (2004). Transfer between views of conspecific faces at different ages or in different orientations by sheep. *Behavioural Processes*, *67*, 491-499.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac MD: Erlbaum.
- Ghosh, N., Lea, S. E. G., & Noury, M. (2004). Transfer to intermediate forms following concept discrimination by pigeons: Chimeras and morphs. *Journal of the Experimental Analysis of Behavior*, *82*, 125-141.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650-669.
- Goto, K., Wills, A. J., & Lea, S. E. G. (2004). Global-feature classification can be acquired more rapidly than local-feature classification in both humans and pigeons. *Animal Cognition*, *7*, 109-113.
- Handel, S., & Imai, S. (1972). Free classification of analyzable and unanalyzable stimuli. *Perception & Psychophysics*, *12*, 108-&.
- Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science*, *146*, 549-551.
- Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*, 285-302.

- Huber, L. (2001). Visual categorization in pigeons. In R. G. Cook (Eds.), *Avian visual cognition*. Available online at www.pigeon.psy.tufts.edu/avc/huber/
- Huber, L., & Lenz, R. (1993). A test of the linear feature model of polymorphous concept discrimination with pigeons. *Quarterly Journal of Experimental Psychology: Comparative & Physiological Psychology*, *46B*, 1-18.
- Jitsumori, M. (1993). Category discrimination of artificial polymorphous stimuli based on feature learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *19*, 244-254.
- Jitsumori, M. (1994). Discrimination of artificial polymorphous categories by rhesus monkeys (*Macaca mulatta*). *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology*, *47B*, 371-386.
- Jitsumori, M. (1996). A prototype effect and categorization of artificial polymorphous stimuli in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*, 405-419.
- Jitsumori, M., & Yoshihara, M. (1997). Categorical discrimination of human facial expressions by pigeons: A test of the linear feature model. *Quarterly Journal of Experimental Psychology*, *50B*, 253-268.
- Krechevsky, I. (1932). Hypotheses in rats. *Psychological Review*, *39*, 516-532.
- Lea, S. E. G. (1979). The art of being an experimental psychologist. *New Universities Quarterly*, *33*, 167-176.
- Lea, S. E. G., & Harrison, S. N. (1978). Discrimination of polymorphous stimulus sets by pigeons. *Quarterly Journal of Experimental Psychology*, *30*, 521-537.
- Lea, S. E. G., Lohmann, A., & Ryan, C. M. E. (1993). Discrimination of 5-dimensional stimuli by pigeons: Limitations of feature analysis. *Quarterly Journal of Experimental Psychology*, *46B*, 19-42.
- Lea, S. E. G., and Ryan, C. M. E. (1983). Feature analysis of pigeons' acquisition of concept discrimination. In M. L. Commons, R. J. Herrnstein and A. R. Wagner (Eds.), *Quantitative analyses of behavior, Vol 4: Acquisition processes*, pp. 239-253. Cambridge, MA: Ballinger.
- Lea, S. E. G., Wills, A. J., & Ryan, C. M. E. (2006). Why are artificial polymorphous concepts so hard for birds to learn? *Quarterly Journal of Experimental Psychology*, *59*, 251-267.
- Lockhead, G. R. (1972). Processing dimensional stimuli. *Psychological Review*, *79*, 410-419.
- Mahut, H. (1954). The effect of stimulus position on visual discrimination by the rat. *Canadian Journal of Psychology*, *8*, 130-138.
- Martignon, L., & Hoffrage, U. (2002). Fast, frugal and fit: Simple heuristics for paired comparison. *Theory and Decision*, *52*, 29-71.
- Matheson, S., Asher, L., & Bateson, M. (2007). Larger, enriched cages are associated with 'optimistic' response biases in captive European starlings (*Sturnus vulgaris*). *Applied Animal Behaviour Science*, *109*, 374-383.
- Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology*, *19*, 242-279.
- Milton, F., Longmore, C. A. and Wills, A. J. (2007). Processes of overall similarity sorting in free classification. *Journal of Experimental Psychology: Human Perception and Performance*, in press.
- Milton, F., & Wills, A. J. (2004). The influence of stimulus properties on category construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 407-415.
- Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision making. *Journal of Experimental Psychology: Learning Memory and Cognition*, *29*, 53-65.
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53-79.
- Overmier, J. B., Bull, J. A., & Trapold, M. A. (1971). Discriminative cue properties of different fears and their role in response selection in dogs. *Journal of Comparative & Physiological Psychology*, *76*, 478-482.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning - variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532-552.
- Pearce, J. M., & Redhead, E. S. (1995). Supernormal conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 155-165.
- Pothos, E. M. (2005). The rules versus similarity distinction. *Behavioral and Brain Sciences*, *28*, 1-49.
- Regehr, G., & Brooks, L. R. (1995). Category organization in free classification - the organizing effect of an array of stimuli. *Journal of Experimental Psychology: Learning Memory and Cognition*, *21*, 347-363.
- Reynolds, G. S. (1961). Attention in the pigeon. *Journal of the Experimental Analysis of Behavior*, *4*, 203-208.
- Roberts, W. A., & Mazmanian, D. S. (1988). Concept-learning at different levels of abstraction by pigeons, monkeys, and people. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 247-260.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
- Schrier, A. M., & Brady, P. M. (1987). Categorization of natural stimuli by monkeys (*Macaca mulatta*) - effects of stimulus set size and modification of exemplars. *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 136-143.
- Sheldon, M. H. (1967). Some effects of discrimination

- goal-box conditions on the learning of a successive discrimination.. *Quarterly Journal of Experimental Psychology*, 19, 319-326.
- Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, 75(517).
- Smith, L. B., & Kemler, D. G. (1977). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology*, 24, 279-298.
- Smith, J. D. & Kemler Nelson, D. G. (1984). Overall similarity in adults' classification: The child in all of us. *Journal of Experimental Psychology: General*, 113, 137-159.
- Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.
- Troje, N. F., Huber, L., Loidolt, M., Aust, U., & Fieder, M. (1999). Categorical learning in pigeons: The role of texture and shape in complex static stimuli. *Vision Research*, 39, 353-366.
- Von Fersen, L., & Lea, S. E. G. (1990). Category discrimination by pigeons using 5 polymorphous features. *Journal of the Experimental Analysis of Behavior*, 54, 69-84.
- Vonk, J., & MacDonald, S. E. (2002). Natural concepts in a juvenile gorilla (*Gorilla gorilla gorilla*) at three levels of abstraction. *Journal of the Experimental Analysis Of Behavior*, 78, 315-332.
- Vonk, J., & MacDonald, S. E. (2004). Levels of abstraction in orangutan (*Pongo abelii*) categorization. *Journal of Comparative Psychology*, 118, 3-13.
- Wittgenstein, L. (1968). *Philosophical investigations* (3rd edn; Trans G. E. M. Anscombe). Oxford: Blackwell. Originally published, 1953.