1

Labelling and Family Resemblance in the discrimination of polymorphous categories by pigeons

5

Elizabeth Nicholls, Catriona M. E. Ryan,

Catherine M. L. Bryant, Stephen E. G. Lea

8

9 University of Exeter
10 School of Psychology
11 Washington Singer Laboratories
12 Perry Road
13 Exeter
14 EX4 4QG
15
16 Email: s.e.g.lea@exeter.ac.uk
17 Tel: +44 1392 264612
18 Fax: +44 1392 264623
19

21 This draft last saved 1 July 2010 by SEGL

22

23

24 **Abstract**

25

26 Two experiments examined whether pigeons discriminate polymorphous categories

27 on the basis of a single highly predictive feature or overall similarity. In the first

28 experiment, pigeons were trained to discriminate between categories of photographs

29 of complex real objects. Within these pictures, single features had been manipulated

30 to produce a highly salient texture cue. Either the picture or the texture provided a

31 reliable cue for discrimination during training, but in probe tests, the picture and

32 texture cues were put into conflict. Some pigeons showed a significant tendency to

33 discriminate on the basis of the picture cue (overall similarity or family resemblance),

34 whereas others appeared to rely on the manipulated texture cue. The second

35 experiment used artificial polymorphous categories in which one dimension of the

36 stimulus provided a completely reliable cue to category membership, whereas three

37 other dimensions provided cues that were individually unreliable but collectively

38 provided a completely reliable basis for discrimination. Most pigeons came under

39 the control of the reliable cue rather than the unreliable cues. A minority, however,

40 came under the control of single dimensions from the unreliable set. We conclude

41 that cue salience can be more important than cue reliability in determining what

42 features will control behaviour when multiple cues are available.

43 **Key words:** Pigeon, Category discrimination, Feature learning, Family

44 resemblance, Labelling, Salience

45

General Introduction

The ability to categorise natural objects is expected to be widespread across the animal kingdom, since it permits a reduction in the amount of information an animal must acquire about its environment in order to respond adaptively (Makino and Jitsumori 2007). Since the pioneering experiments of Herrnstein and Loveland (1964) birds have been shown to be capable of discriminating a wide range of categories, with category members varying from simple artificially constructed stimuli (e.g. Lea and Harrison 1978) to photographs of complex natural-language categories such as trees (Herrnstein et al. 1976; for a review see Huber 2001).

A typical categorisation experiment involves pigeons learning to associate one stimulus set or category with a food reward, when there is no single feature that is necessary or sufficient for category membership. However it is not clear what pigeons learn to associate with the food reward. It could be that the pigeon has elaborated a concept corresponding to the category. Typically however it is assumed that successful category discrimination does not necessarily imply that pigeons have a concept corresponding to the experimenter-defined category (despite the terminology of early experiments e.g. Herrnstein et al. 1976), or in the case of natural photographs, the ability to generalise from images to the corresponding real objects (Bovet and Vauclair 2000). But if pigeons are not using concepts to categorize complex stimuli, how might they be doing it? There are two well studied possibilities. It could be that they learn how to respond to one or more specific exemplars, with other stimuli being categorised by a process of generalisation. Alternatively, it may be that they learn to respond to a collectivity of features that are between them good enough predictors of

71 category membership, even if none of them is individually necessary or sufficient.

72 These possibilities are usually referred to respectively as exemplar learning and

73 learning by family resemblance or overall similarity.

74

75 It is often useful to compare the cognitive abilities of non-human animals and human

76 infants, because in both cases subjects have to complete tasks without the elaborate

77 language competence that facilitates so many cognitive performances in adult

78 humans. The present experiment draws on two results that have been established in

79 the field of infant categorization, in order to pose questions about categorization in

80 pigeons.

81

82 Both phenomena involve the use of single, highly predictive features within sets of

83 multidimensional stimuli. However, they are to some extent contradictory. On the

84 one hand, under conditions where adults and older children typically categorise items

85 according to a single stimulus dimension, infants have been found to group items

86 according to overall similarity or family resemblance across numerous stimulus

87 dimensions (Smith and Kemler 1977; Smith 1981). On the other hand, the provision

88 of a verbal label which is invariantly associated with the members of one category has

89 been found to facilitate children's category acquisition (Waxman and Markow, 1995),

90 even among infants too young to have functional speech (e.g. Balaban and Waxman,

91 1997); so in contrast to the control by family resemblance seen in spontaneous

92 categorization, the presence of a single reliable feature is found to facilitate category

93 acquisition.

94

95   The first of these findings, the tendency of younger children to categorize on the basis

96   of overall similarity, has supported a general arguments that such categorization must

97   require simpler cognitive mechanisms than using a single stimulus dimension (e.g.

98   Ashby et al. 1998; Pothos 2005).   Unidimensional discrimination is assumed to

99   require the ability to verbalise a rule, something that a pre-verbal infant cannot do. In

100   accordance with this assumption, Couchman et al. (2010) found that, when they used

101   stimulus sets of a kind introduced by Kemler Nelson (1984), which can be

102   discriminated either on the basis of a single dimension or on a family resemblance

103   across three other dimensions, human participants categorised them unidimensionally

104   whereas two rhesus monkey subjects categorized them mainly by family resemblance.

105

106   However, it is not the case that non-human animals always categorize complex stimuli

107   by overall similarity rather than unidimensionally, or that unidimensional

108   categorization implies more complex cognitive processing (e.g. Lea and Wills 2008;

109   Wills et al. 2009). Although it is certainly possible for pigeons to use multiple

110   stimulus dimensions in making complex discriminations (e.g. Blough 1972;

111   Herbranson et al. 1999; Kirkpatrick-Steger and Wasserman 1996; Soto and

112   Wasserman 2010; von Fersen and Lea 1990), this is not necessarily what they do most

113   easily or spontaneously. Rather than categorizing in terms of overall similarity, which

114   would involve all available features, birds in categorization experiments with complex

115   stimuli often come under the control of just one or two features (e.g. Lea et al. 1993,

116   2006); and it would be absurd to suppose that pigeons can verbalise rules (cf. Lea and

117   Wills, 2008). Furthermore, the fact that infant categorization proceeds more easily in

118   the presence of a verbal label – which is really nothing but a single highly valid

119   stimulus dimension (Plunkett et al., 2008) – suggests that discriminating a single

120   dimension may be a cognitively simpler task than discriminating overall similarity.

121

122   The aim of the present study is to explore whether, under conditions where both

123   strategies would be equally effective, pigeons would solve categorization tasks by

124   using a single highly predictive feature, or by using the overall similarity across a

125   group of features of more modest individual predictiveness. However, previous results

126   have shown that when pigeons do make use of single features, they do not always

127   select the most predictive feature: a less valid but more salient dimension of stimulus

128   variation may acquire exclusive control over behaviour (Lea et al. 2009).

129   Accordingly, the experiments were designed to enable us to dissociate the effects of

130   cue salience from those of cue validity.

131

132   Von Fersen and Lea (1990) noted that categorization in pigeons can be investigated in

133   two different ways: by using natural photographs, videos and objects, which may be

134   referred to as an analytic approach, and by using artificially designed stimuli, the

135   synthetic approach. The advantage of using artificial stimuli is that structure and

136   feature content can be carefully controlled, but such control is usually at the expense

137   of the richness and detail associated with natural exemplars. The use of photographs

138   or videos enhances the verisimilitude of categorization experiments, because it

139   presumably mimics more closely the kind of discrimination that birds have to make in

140   the wild (though it cannot do so exactly, and it cannot necessarily be assumed that the

141   birds recognize the pictures as representations of the corresponding real objects). In

142   the present paper, we used both techniques, so as to gain the advantages of both, and

143   also to provide systematic replication of our main manipulation. Experiment 1 used

144 natural photographs (though with some artificial modification), while Experiment 2

145 used wholly artificial stimuli   In both experiments, the stimuli included a single

146 feature that was a perfectly reliable predictor of category membership; by analogy

147 with the work of Waxman and others cited above, we refer to this as the "label". In

148 addition the stimuli included several other features that were individually imperfect

149 predictors but in combination provided enough information to allow perfect

150 discrimination.   Once good performance had been achieved, probe trials were

151 introduced, in which the single, perfectly reliable cue and the remaining cues were put

152 in conflict with each other, allowing us to discern which cues were controlling

153 behaviour more strongly.

154

155 **Experiment 1**

156

157 The first experiment used photographs of everyday objects as stimuli. The two

158 categories were houses and cars, which differ along multiple stimulus dimensions. To

159 provide a single highly salient "label" cue, the roofs of the houses and the windows of

160 the cars were replaced with a standard uniform texture in a contrasting colour. If the

161 pigeons in this situation base their discrimination on the label alone, then when probe

162 stimuli are introduced in which the labels have been reversed between categories, it is

163 expected that probes containing the label previously associated with the positively

164 reinforced category (i.e. positive texture probes) will be treated as positive. If pigeons

165 respond positively instead to probes which contain the label originally associated with

166 the negatively reinforced category (positive picture probes), this would suggest they

167 are attending more to the overall similarity of exemplars. Although the stimuli were

168 based on natural photographs of objects, it was not required by the design that the

169   pigeons should recognize them as depictions of objects that they had seen;

170   photographs were used only to ensure that the stimulus sets had the kinds of variation

171   that are characteristic of functional categories in the everyday world.

172

173   Materials and Methods

174

175   *Subjects*

176   Twenty pigeons were used. They were acquired as discards from pigeon fanciers, and

177   so had visual experience of the world outside the laboratory. Some had previous

178   experience of the touchscreen arrangement used in the present experiment, but in

179   experiments with monochrome, geometric stimuli and a different training procedure

180   (Wills et al. 2009, Experiment 1a); the remainder were experimentally naive. The

181   pigeons were housed in an indoor aviary, measuring 2.2 m by 3.4 m by 2.4 m, and

182   given constant access to water and grit. Prior to testing, pigeons were held in

183   individual cages in which they had access to water and were released back into the

184   aviary when testing had finished for the day. All pigeons were maintained on a 12:12

185   hr light/dark cycle at 95% of their free-feeding weight.

186

187   *Apparatus*

188   Four identical operant chambers (internal size; 640 mm x 430 mm x 470 mm) were

189   used. Each consisted of a plywood box, with a 15-inch (39-cm) touch-monitor (Elo

190   Touchsystems Accutouch (resistive) or CarrollTouch (infra-red) model 1547L)

191   mounted in the front wall. The screen resolution was set to 1024 x 768 pixels. The

192   monitor was controlled by software written in Visual Basic using the Whisker control

193 system (Cardinal and Aitken 2001) running on a computer supplied by Quadvision

194 (Quadvision Ltd., Dorset, UK). Two food hoppers, positioned one on each side of the

195 main screen, could be used to deliver a 2:1 mixture of hemp seed and conditioner to

196 the pigeons, for 2.5 seconds. Each box had a webcam fitted into the side wall, 250

197 mm above the floor, allowing the pigeons' behaviour to be observed from outside the

198 test room using the imaging software ViewCommander (Internet Video and Imaging,

199 Ltd.). Each pigeon was assigned its own test chamber for all stages of the experiment:

200 six pigeons used resistive touchscreens while the remaining pigeons used infra-red

201 touchscreens.

202

203 *Stimulus Materials*

204 The images used were natural photographs comprising two stimulus sets; houses and

205 cars. There were 24 pictures of each type of object. Photographs were manipulated

206 using the GNU Image Manipulation Program (GIMP ver. 2.2). Each image was

207 isolated from the original setting, placed on a black background and then scaled to

208 ensure images were approximately the same size (160 x 120 pixels). To produce a

209 highly salient "label" feature, similar areas in the upper part of the objects were

210 selected and given a new artificial texture. For houses this was the roof, and for cars

211 the windows. Thus 12 of the houses had their roofs replaced with a leopard-skin

212 texture, and the other 12 had their roofs replaced by a blue "swimming pool" wave

213 texture. Similarly, 12 of the cars had all their windows replaced with the leopard-skin

214 texture and the other 12 had them replaced with the blue wave texture. Figure 1

215 shows examples of the stimuli. Fourteen of the pigeons were trained to discriminate

216 the houses with leopard-skin roofs from the cars with blue windows, and for these

217 pigeons probe stimuli consisted of houses with blue roofs and cars with leopard-skin

218   windows. For the remaining pigeons, the training stimuli were the houses with blue

219   roofs and the cars with leopard-skin windows, and the probe stimuli were the houses

220   with leopard-skin roofs and the cars with blue windows. Within each of these groups,

221   some of the pigeons were trained with houses as positive stimuli and the remainder

222   with cars as positive stimuli.

223

224   *General Procedure*

225   The pigeons were pre-trained using conventional methods to find food in the grain-

226   feeders, and to peck lighted discs on the touch-screen for food reinforcement. They

227   were then trained in a multi-stimulus discrimination procedure similar to that

228   described by Huber et al. (2005) and Wills et al (2009, Experiment 2b). Initially, they

229   were trained using this procedure to discriminate white filled hexagons (8mm across)

230   within a 25-mm diameter black circle, outlined in white, from a blank circle. Those

231   pigeons that mastered this discrimination proceeded to the task involving the house

232   and car stimuli.

233

234   In the multi-stimulus discrimination procedure, each session started when the house-

235   lights in the box came on, and a white disc of diameter 4cm (the observing key) was

236   displayed on the screen. When the pigeon pecked the observing key it disappeared, to

237   be replaced by an array of photographs. The array consisted of eight cells, arranged as

238   shown in Figure 2. In order to 'open' a cell, pigeons had to peck it twice in rapid

239   succession, causing the image to disappear. If the opened cell was positive, a side key

240   appeared to one side of the array, in the nearer of the two locations shown in Fig. 2

241   (arbitrarily, the left side key was used when the middle top cell was opened, and the

242   right side key when the middle bottom cell was). The pigeons were required to peck

10

243 the side key to activate the food hopper on that side for 2.5 seconds (Figure 2). If a

244 negative cell was pecked this was recorded as a miss; the image disappeared but no

245 side key appeared and the screen became unresponsive for 2.5 seconds. If a pigeon

246 pecked a blank cell this was recorded but there were no scheduled consequences.

247 When all the positive cells had been opened, the array disappeared, to be replaced

248 after an inter-trial interval that varied from 1 to 5 s by the observing key. Sessions

249 consisted of six or seven arrays depending on the stage of the experiment.

250

251 *Training*

252 Three pigeons failed to complete pre-training and were dropped from the experiment.

253 Of the remaining 17 pigeons, 11 were trained using the houses with leopard-skin roofs

254 and the cars with blue windows; of these, five were assigned houses as the positive

255 stimulus, and six were car-positive. The remaining six pigeons were trained using the

256 houses with blue roofs and the cars with leopard-skin windows, and of these three

257 were house-positive and three were car-positive. Arrays were made up of four cars

258 and four houses, pseudo-randomly arranged, with the constraints that stimuli from the

259 same set were never presented in the same place in more than three consecutive arrays

260 and that no more than three stimuli from the same set were placed next to each other

261 or reinforced on the same side. Each training session contained six arrays, so that each

262 of the 12 positive and negative stimuli was seen twice within a session.

263 Discrimination during the training phase was determined using the ρ statistic of

264 Herrnstein et al. (1976), which was used to measure the overlap between stimulus sets

265 in terms of the order in which the cells were opened. A ρ value of 0.5 indicates

266 random responding, whereas a value of 1 indicates perfect discrimination (i.e. that all

267 the positive stimuli have been opened before any of the negative stimuli). After each

268 pigeon had reached a criterion of a ρ value of 0.8 or more in all six arrays in a session,

269 it was exposed to probe sessions.

270

271 *Probe Sessions*

272 Probe sessions consisted of four training arrays, alternated with three probe arrays.

273 Each session began and ended with a training array. A probe array contained four

274 training stimuli, two from each stimulus set, and four probe stimuli in which the

275 artificial manipulations were reversed. Two of the probes in each array, one from each

276 stimulus set, were assigned positive reinforcement contingencies and the other two

277 were assigned negative contingencies. Probes that were designated positive the first

278 time they were displayed were not reinforced when next shown. Pigeons were

279 required to open the two positive training cells and all four probe cells in order to

280 complete a probe array. Each pigeon received two probe sessions, which were

281 alternated with training sessions, to ensure pigeons recovered their original training

282 performance. Following the procedure of Wills et al. (2009, Experiment 2b),

283 responses to probe stimuli were categorised according to the order in which the

284 pigeons opened them within the array. A probe that was one of the first four stimuli

285 opened in an array was considered to have been treated as positive, and a probe that

286 was one of the last four stimuli opened was considered to have been treated as

287 negative.

288

289 Results

290

291 The pigeons reached criterion in a median of 5 training sessions (range 2-8 sessions).

292 This corresponds to the presentation of a median of 30 arrays, i.e. 240 stimuli. There

293 were no significant differences in the rate of acquisition between the house-positive

294 and the car-positive groups, or the leopard-skin positive and blue-wave positive

295 groups, nor was the interaction between the positive picture and the positive texture

296 significant.

297

298 In probe sessions, responses to training stimuli continued at the same high level of

299 accuracy. Different pigeons behaved differently towards the probe stimuli. Probes

300 included either a positive picture (house or car, depending on which was positive

301 during training) and a negative texture (blue wave or leopard-skin), or a negative

302 picture and a positive texture. Figure 3 shows, for each pigeon, the proportion of

303 probe stimuli that were treated in accordance with the texture cue present (that is,

304 responded to as positive if they included the positive texture or as negative if they

305 included the negative texture). In all, 7 of the 17 pigeons (Ro, Ma, Fr, Sn, Su, Ba and

306 Ze) responded to more than half the probes in accordance with the texture cue, 9

307 pigeons (Io, Le, Ri, Ru, Rs, Ti, Ot, Ry, and Se) responded to more than half the

308 probes in accordance with the picture cue, and the remaining pigeon (At) treated both

309 kinds of probes equally. Chi-square tests were used to determine whether the

310 proportions of probe stimuli treated in accordance with the texture or picture cue

311 differed significantly from 0.5 for individual pigeons. Four pigeons responded to the

312 probes in accordance with the texture cue on significantly more than half the trials

313 (Ro, Fr, Sn and Su; χ^2 values of 15.04, 14.09, 8.52 and 4.34, df=1, p<0.0001, 0.0001,

314 0.01 and 0.05 respectively). Three pigeons responded to the probes in accordance

315 with the picture cue on significantly more than half the trials (Io, Ri, and Ry; χ^2 values

316 of 15.04, 9.38 and 7.04, df=1, p<0.0001, 0.01 and 0.01 respectively). The sum of the

317 individual χ^2 values across the pigeons was 83.05; by virtue of the additive property of

318 the chi-square distribution (Weatherburn, 1957, p. 177), this can be tested as a χ^2 value

319 against the sum of their degrees of freedom. The resulting significance level is <

320 0.0001, showing that despite the fact that not all results were in the same direction, we

321 can reject the null hypothesis that all pigeons had a 50% chance of treating any probe

322 in accordance with the picture cue.

323

324

325 Discussion

326

327 The rapid learning shown by all the pigeons confirms the conclusion of Huber et al.

328 (2005) and Wills et al (2009) that multi-stimulus training methods are an efficient

329 means of establishing complex pattern discriminations. It is a property of such

330 methods, as they have been implemented previously and in the present experiment

331 that, within an array, a correct response makes the remainder of the task more difficult

332 (because it reduces the ratio of positive to negative stimuli remaining) and an

333 incorrect response makes it easier. This makes the task of predicting the probability

334 of a correct response mathematically more difficult, but is not a cocnern when as here,

335 the intention is to compare the probabilities of choosing different probe stimuli that

336 are presented at the same frequency.

337

338 Responses to probe stimuli were reinforced non-differentially. This was because

339 substantial numbers of probe trials were required, and the alternative, non-

14

340 reinforcement, could have led to the pigeons learning that all probe stimuli were

341 essentially negative. Non-differential reinforcement carries the alternative risk that

342 subjects will come to behave at random towards probes. However in the present

343 experiment, sustained non-random behaviour towards probes was observed in most

344 pigeons, showing that non-differential reinforcement had no major effect.

345

346 The pigeons were almost equally split between showing greater control by the picture

347 and showing greater control by the texture (the "label" cue). This was not a result of

348 random behaviour, but of systematic behaviour that differed between individuals.

349 Seven of the 17 pigeons showed significant differences in their responses to probes

350 containing the positive picture rather than the positive texture, but in three of these it

351 was the picture cue that controlled probe responding, and in the other four it was the

352 texture cue. As regards the remaining pigeons, given that performance on training

353 stimuli remained highly accurate throughout the probe sessions, the most plausible

354 conclusion is that they were under the control of both the picture and the texture. The

355 distribution of control by the picture or texture was not even between the groups, and

356 in particular all four of the pigeons that showed a significant trend to texture control

357 had the leopard-skin texture as positive. If the leopard-skin was more salient than the

358 blue-wave texture for the pigeons, this could be accounted for as a feature-positive

359 effect (Jenkins and Sainsbury 1970), but in the absence of independent evidence about

360 relative salience, this can only be a speculation. Furthermore one of the pigeons that

361 showed a significant trend towards control by the picture cue (Io) also had the

362 leopard-skin texture as positive, weakening the argument that the overall pattern of

363 results could be due to high salience of the leopard-skin cue.

364

365 Because the pictures were natural photographs, we cannot easily specify which

366 features were supporting discrimination between them, but inevitably the features

367 available will have varied somewhat between instances. We therefore interpret

368 discrimination based on the picture cue as categorization by overall similarity or

369 family resemblance. The texture cue, by contrast, was (to the human eye at least)

370 highly salient, and it was consistent across the training and probe stimulus sets.

371 Discrimination of the texture cue can therefore be regarded as unidimensional

372 categorization. Nevertheless, it only exerted dominant control over behaviour in

373 probe trials for four pigeons.

374

375 It is possible that the pigeons learned to discriminate the categories by learning each

376 exemplar of a house or car separately. Previous results make this unlikely given the

377 number of exemplars used: an experiment by Bhatt, cited by Wasserman and Bhatt

378 1992, showed that pigeons switch from exemplar control to featural control in

379 category discriminations where the number of exemplars rises above about six. In

380 any case it would not matter to the present experiment, whose point was to investigate

381 whether the pigeons would come under the control of the single feature provided by

382 the texture cue or the multiple features provided by the picture cue: if the pigeons

383 discriminated houses from cars on the basis of exemplars, that would involve the

384 learning of even more different features than doing so by extracting a few general

385 features that were positively but not perfectly correlated with reinforcement.

386

387 The roughly equal distribution of subjects between control by the texture cue and

388 control by the picture cue is an unexpected result, given that pigeons have a tendency

389 to be dominated by single features (Lea et al. 1993, 2006). However, while it is true

390   that birds rarely use all the features available, they can certainly be trained to use

391   more than one (e.g. von Fersen and Lea 1990; Lea et al. 2006). Indeed, some of the

392   pigeons used in the present experiment had been trained in Experiment 1a of Wills et

393   al. (2009), in which they had to learn to discriminate several different features

394   presented on separate trials (though the stimuli and the procedure were very different

395   from those of the present experiment). In the generalisation tests of that experiment,

396   when the features were combined, the behaviour of one of the pigeons that was re-

397   used here (Io) was shown to be under the control of multiple features. However,

398   bringing behaviour under the control of multiple stimulus dimensions often requires

399   special training procedures (as in von Fersen and Lea 1990) or extended training,

400   whereas the pigeons in the present experiment learned very quickly.

401

402   Because the picture cues were based on natural photographs, it remains an assumption

403   that discrimination of the pictures was on the basis of overall similarity. It is possible

404   that within the pictures there were other highly predictive features apart from the one

405   manipulated, which were consistent within each category and which were salient for

406   the pigeons. All the car stimuli, for example, contain wheels and so pigeons might

407   have learnt the discrimination on the basis of the presence or absence of silver

408   ellipses. We did our best to ensure there was no such single predictive feature for

409   each category, for example the angle at which the photographs of the stimuli were

410   taken was varied, as was the colour of the cars (so that although all the cars had

411   wheels, those wheels themselves formed a polymorphous set), but there can be no

412   certainty that such attempts will succeed. Conversely, although the label features

413   were much more consistent, they did have some variability: for example, the shapes of

414   the house roofs and car windows to which the textures were applied varied between

415 instances. To avoid these problems while investigating the question of whether birds

416 are more likely to rely on single features or family resemblance, in Experiment 2 we

417 took the alternative approach of using artificial compound stimuli, and so bringing

418 feature content under control.

419

420 **Experiment 2**

421

422 Experiment 2 was formally very similar to Experiment 1, but used artificial

423 multidimensional stimuli. The stimuli in the present experiment were made up of four

424 spatially separated elements, each of which constituted a stimulus dimension; they

425 were based on those used by Wills et al. (2009, Experiment 1a) and Lea et al. (2009).

426 The elements are referred to as the Annulus, the Bar, the Checks and the Diamond,

427 and examples are shown in Figure 4. Each element could be used in four graded

428 versions, two of them positively and two negatively correlated with the occurrence of

429 reinforcement. The stimulus sets had exactly the same formal structure as those used

430 by Kemler Nelson (1984) with children and Couchman et al (2010) with rhesus

431 monkeys, but the stimulus dimensions used were different.

432

433 Corresponding to the reliable, salient, texture cue used as a label in Experiment 1, one

434 of the dimensions was designated as a label cue for each pigeon. This dimension was

435 used only in its extreme versions, and one of these always occurred in the presence of

436 reinforcement while the other one never did. The remaining three dimensions were

437 used to construct a two-out-of-three artificial polymorphous stimulus of the sort

438 introduced by Dennis, Hampton and Lea (1973) and used with pigeons by Lea and

Harrison (1978): that is, whenever at least two of the three dimensions took its positive value, the stimulus as a whole was designated as positive. For these dimensions, all four versions were used. These three dimensions collectively corresponded to the picture feature in Experiment 1: if a pigeon was to discriminate on the basis of these cues, it would have to come under the control of the overall similarity of the stimuli to an ideal positive form if it was to achieve 100% correct performance, since each individual dimension within this set was imperfectly correlated with reinforcement. These artificial stimulus sets had several additional advantages. First, the spatial separation of the stimulus elements ensured that the dimensions of the stimuli could be manipulated entirely independently. Second, the four dimensions could be used in a balanced way, with each dimension being used to provide the label (reliable) feature for some pigeons. Thirdly, within the set of features used to create the polymorphous concept, it was possible to assess empirically whether all three features were controlling behaviour, and therefore whether it is appropriate to describe the pigeons as coming under the control of overall similarity. Finally, on the basis of results with similar stimuli (Lea et al 2009, Wills et al 2009) we had reason to think that the saliences of the feature differences used on the four dimensions of the stimuli were comparable.

Materials and Methods

Subjects

Sixteen pigeons were used in this experiment. None of them had previous experience of this kind of discrimination task. They were maintained under the same conditions as the pigeons used in Experiment 1.

Apparatus

465 The same apparatus was used as in Experiment 1, except that only infra-red

466 touchscreens were used.

467

468 *Stimuli and Experimental design*

469 The stimuli were modified from those used in Wills et al. (2009, Experiment 1a) and

470 Lea et al. (2009). Examples are shown in Figure 4. All stimuli consisted of a square

471 array of four elements. Elements of the same type were always placed in the same

472 location. The element types were an annulus (A), a horizontal bar (B), a square

473 chequerboard (C), and a diamond shape made up of equal-width stripes (D). There

474 were four versions of each element, designated as the X, x, y, and Y forms. All

475 versions of all elements were placed on a black background, and contained the same

476 number of white pixels (within 2%). The specifications of the four forms of each

477 element are listed in Table 1, and they are illustrated in Figure 4. Note that because of

478 the constraint that all elements should have the same area, some elements varied on

479 two dimensions simultaneously: for example, when the Annulus was made smaller, it

480 was also made wider. With four versions of each of four elements, there were 256

481 possible stimuli, but not all pigeons experienced all stimuli. All versions of all

482 elements fitted within a 60 x 60 pixel square with some black border, so that the entire

483 stimulus including borders fitted within a 120 x 120 pixel square; at a pigeon's typical

484 pre-peck viewing distance, 120 pixels subtended about 25° of arc, and would thus fit

485 within the extent of the pigeon's frontal, binocular viewing field (Martin and Young,

486 1983).

487

For each pigeon, one of the four dimensions was designated as Reliable, and either the X or the Y value of it was designated as positive; the less extreme (x and y) values of the Reliable dimension were not used either in training or in probe stimuli. The remaining three dimensions were designated as Unreliable, and either all their X and x values, or all their Y and y values, were designated as positive. Positive training stimuli always included the extreme positive value of the Reliable dimension, and positive values (either extreme or less extreme) of at least two of the Unreliable dimensions. Negative training stimuli always included the extreme negative value of the Reliable dimension, and negative values (either extreme or less extreme) of at least two of the Unreliable dimensions. As an example, consider Pigeon Mo, for which the Reliable dimension was the Annulus and the X values of both the Reliable and Unreliable dimensions were designated as positive (see Table 2). For this pigeon, positive stimuli always contained the X value of the Annulus, and at least two of the other three dimensions (Bar, Chequerboard and Diamond) in either their X or their x forms. The negative stimuli always contained the Y form of the Annulus, and either the y or the Y form of at least two of the other three dimensions. There were 32 stimuli in each of the positive and negative categories. These categories could be discriminated perfectly in either of two ways (or by a mixture of them). The pigeon could use the Reliable dimension alone, and ignore the three Unreliable dimensions. Alternatively, it could ignore the Reliable dimension, and respond on the basis of the preponderance of values of the three Unreliable dimensions. Each Unreliable dimension took one of its negative values in a quarter of the positive training stimuli, and a positive value in a quarter of the negative training stimuli, so each Unreliable dimension considered on its own can be described as being 75% valid during training; their collectivity, however, was 100% valid. The Reliable cue thus served as a label,

513 while the Unreliable cues defined a 2-out-of-3 polymorphous stimulus set. Stimuli

514 that included the positive value of the Reliable dimension with negative values of at

515 least two of the Unreliable dimensions, or the negative value of the Reliable

516 dimensions with positive values of at least two of the Unreliable dimensions, were

517 ambiguous, and were not used in training. There were 64 such ambiguous stimuli for

518 each pigeon, and a selection of these were used as probes, in tests conducted after

519 training was complete.

520

521 The way the categories were used was varied between pigeons so as to balance the use

522 of the different dimensions, as shown in Table 2. Each dimension was assigned as

523 Reliable for four of the pigeons. For two of these, the X value of the Reliable

524 dimension was assigned as positive, and for the other two its Y value was assigned as

525 positive. For one of the pigeons for which each value of the Reliable dimension was

526 assigned as positive, the X and x values of the Unreliable dimensions were assigned

527 as positive, and for the other one the Y and y values of the Unreliable dimensions

528 were assigned as positive.

529

530 To reduce the risk that the pigeons would learn the contingencies associated with

531 probe stimuli, only 36 probe trials were given to each pigeon, so not all the 64

532 possible probe stimuli were used. The stimuli used as probes are summarised in Table

533 3. Four of these stimuli were used 6 times each, so as to provide a strong test of the

534 basic question of whether the pigeons' behaviour was governed by the Reliable or the

535 Unreliable dimensions. These stimuli involved:

536 The positive value of the Reliable dimension, combined with the extreme negative

537 values of all three Unreliable dimensions

538   The negative value of the Reliable dimension, combined with the extreme positive

539   values of all three Unreliable dimensions

540   The positive value of the Reliable dimension, combined with the less extreme

541   negative values of all three Unreliable dimensions

542   The negative value of the Reliable dimension, combined with the less extreme

543   positive values of all three Unreliable dimensions.

544

545   In addition six other probe stimuli, as indicated in Table 3, were used twice each.

546   These stimuli involved the positive value of the Reliable dimension and the extreme

547   positive value of one of the Unreliable dimensions, and the extreme negative values of

548   the other two Unreliable dimensions; or the negative value of the Reliable dimension

549   and the extreme negative value of one of the Unreliable dimensions, and the extreme

550   positive values of the other two Unreliable dimensions. Each of the Unreliable

551   dimensions was paired with the Reliable dimension in an equal number of probe

552   stimuli. These stimuli allowed a test of which of the Unreliable dimensions was

553   controlling behaviour. Each probe stimulus was associated an equal number of times

554   with the reinforcement contingencies appropriate to positive and negative stimuli.

555

556   *Procedure*

557   The pigeons were pretrained as in Experiment 1. They were then trained, using the

558   same procedure as in Experiment 1, to discriminate between two positive and negative

559   training categories. Training was continued for a maximum of 20 sessions, but was

560   stopped earlier if a pigeon reached a criterion of a ρ value of 0.8 or more in all six

561   arrays in a session. Three probe sessions where then given. As in Experiment 1,

562   probe sessions consisted of seven arrays: four training arrays, alternated with three

563   probe arrays. Each session began and ended with a training array. A probe array

564   contained four training stimuli and four probe stimuli, of which two were associated

565   with the reinforcement contingencies appropriate to positive stimuli and two with

566   those appropriate to negative stimuli. Responses to probe stimuli were classified as

567   positive if they occurred within the first four stimulus cells pecked in an array, and as

568   negative otherwise.

569

570   Results

571   Ten of the 16 pigeons reached criterion before their 20th training session; the number

572   of sessions required ranged from 9 to 18 (median, including the pigeons that did not

573   reach criterion, 17).   The other six pigeons were showing no obvious further

574   improvement in performance after 20 sessions, though all but one of them was

575   showing ρ values consistently above 0.5; the performance of the remaining pigeon

576   was erratic.

577

578   For comparison with Fig. 3, Fig. 5 shows the proportions of probe trials in which the

579   response was correctly predicted by the Reliable stimulus dimension rather than the

580   preponderance of the Unreliable dimensions.   For 10 of the 16 pigeons, this

581   proportion was greater than 0.5, and over the group as a whole the deviation from 0.5

582   was significant (1-sample, 2-tailed Wilcoxon test, $T = 23.5$, $N=16$, $p<0.05$). For each

583   of these pigeons individually, the proportion deviated from 0.5 significantly (2-tailed

584   binomial test).   For one of the six pigeons where the proportion was below 0.5 (Ba),

585   the deviation was significant (2-tailed binomial test, $p < 0.001$). The pigeons that did

586   not respond according to the Reliable stimulus on the probe trials had taken

587   significantly longer to reach criterion than those that did (2-tailed Mann-Whitney test,

588    $U = 10.5$, $N_1 = 6$, $N_2 = 10$, $p = 0.03$).

589

590   Figure 6 shows, for each pigeon, the proportions of probe trials in which the response

591   was correctly predicted by each dimension of the stimulus, whether it was reliable or

592   unreliable for that pigeon. In this figure, any proportion above 0.67 or below 0.33

593   would be significantly different from 0.5 on an individual test, though the number of

594   data points involved and the correlations between dimension values mean that simple

595   significance tests cannot be interpreted directly. It can be seen that for each of the

596   pigeons where the unreliable dimensions predicted the response to probe trials

597   markedly better than the reliable dimension (Mo, Bn and Jk), there was one of the

598   unreliable dimensions that predicted response particularly well (Checks for Mo and

599   Jk, and Annulus for Bn). The same trend can be seen more weakly in two of the

600   pigeons where the dominance of the unreliable dimensions was more marginal, Mr

601   and Sa, where Checks and Annulus respectively seemed to have majority control over

602   behaviour. The remaining pigeon, Cr, showed apparently random behaviour towards

603   probe stimuli. For the pigeons where behaviour towards the probe stimuli was

604   dominated but not 100% predicted by the reliable dimension, there was no evidence

605   that individual unreliable dimensions contributed disproportionately to controlling

606   behaviour.

607

608   Because the values of the unreliable dimensions used in the training stimuli varied, it

609   was possible to examine the relative control over behaviour of these dimensions under

610   training as well as probe conditions.   Table 4 shows the results of such an analysis,

611   carried out on the training stimuli that were used within the probe sessions so as to

612 maximise comparability with the probe stimulus results shown in Figures 5 and 6.

613 There were 96 such trials, so any proportion greater than 0.58 or less than 0.42 would

614 be significantly different from 0.5 in a single analysis, though the same cautions about

615 the number of tests and the non-independence of the dimensions must be applied as

616 with Figure 6. These data confirm the dominance of the Annulus dimension for

617 pigeons Sa and Ba, and weakly support the dominance of Checks for Mr, but do not

618 support the dominance of Checks for Mo or Jk. On the other hand they suggest that

619 Cr may have been somewhat under the control of the Bar dimension. It is notable

620 (and logically inevitable) that in those pigeons where one of the unreliable dimensions

621 exerted disproportionate control over behaviour, overall discrimination performance

622 was poorer than in the other pigeons.

623

624 Discussion

625 In Experiment 1, we found that the texture and picture cues (which we interpret

626 respectively as a single reliable dimension, and a collectivity of presumably unreliable

627 cues) were about equally likely to control behaviour, and in many cases individual

628 pigeons probably came under the control of both types of cue. However the two kinds

629 of cues differed in nature. In Experiment 2, where the same cues were used as

630 reliable and unreliable dimensions in a fully balanced way, we found a clear trend for

631 reliable dimensions rather than a collectivity of unreliable dimensions to control

632 behaviour, even though either of them could have enabled the pigeons to perform with

633 perfect accuracy. There were some individuals in which the unreliable dimensions

634 exerted dominant control over behaviour, but in at least some cases it appears that this

635 was not because the collectivity of those dimensions was enabling perfect

636 discrimination, but rather because one of those dimensions was controlling behaviour

637 to the exclusion of the others and of the reliable dimension, with the result that the

638 pigeon's discrimination accuracy was limited. The pigeons that did not come under

639 the control of the Reliable dimension were slower to reach criterion (indeed, 4 out of

640 these 6 pigeons did not reach it within 20 sessions, as against only 2 of the 10 pigeons

641 that did show Reliable dimension control), and if they were coming under the

642 exclusive control of a less valid dimension, this is a more or less inevitable result.

643 Our results thus differ from those obtained by Couchman et al (2010), using stimulus

644 sets that were structurally identical to those used here, though with different elements:

645 Compared with humans exposed to the same task, Couchman et al's rhesus monkeys

646 always showed a greater tendency to categorise novel stimuli in terms of overall

647 similarity rather than in terms of the label dimension, and the authors concluded that

648 this was because the monkeys could not use verbal rules to categorize complex

649 stimuli. Since we do not believe that pigeons use verbal rules, yet in our experiment

650 they showed a clear tendency to use the label dimension rather than family

651 resemblance, we conclude that categorization by family resemblance is not an

652 inevitable consequence of failing to use verbal rules. We therefore also conclude that

653 unidimensional categorization is not a reliable sign that a verbal rule is being used, in

654 accordance with the conclusions of Lea and Wills (2008) and Wills et al (2009). The

655 difference between our results and those of Couchman et al may be due to the subject

656 species, or it may be due to differences in the details of the stimuli involved, a matter

657 to which we return in the General Discussion.

658

659 Part of the reason why clear dominance of the reliable dimension was found in the

660 present experiment is that with the values on them used in the present experiment, the

661 salience of the four dimensions seems to have been, if not equal, at least comparable,

662   as we predicted on the basis of our previous results with similar stimuli (Wills et al

663   2009, Lea et al 2009). As Table 4 shows, when used as the reliable dimension, all

664   four dimensions achieved dominant control over behaviour in at least one of the four

665   pigeons for which they were Reliable; assuming that salience is reflected in the

666   number of pigeons for which this was so, the order for salience was Checks >

667   Annulus> Diamond > Bar. The data on dominance by Unreliable dimensions (also

668   shown in Table 4) confirm this pattern, with Checks and Annulus being the only

669   dimensions to achieve dominant control when unreliable. Acquisition data also show

670   that the pigeons for which these dimensions were Reliable tended to learn more

671   quickly than the others.

672

673   Learning in Experiment 2 was substantially slower than in Experiment 1. In

674   Experiment 1, pigeons took a median of 5 training sessions (range 2-8 sessions) to

675   reach criterion; in Experiment 2, the median number of sessions to the same criterion

676   was 12, with the number required ranging from 3 to over 20.   This difference is

677   consistent with the fact that in several studies where polymorphous concept

678   acquisition has been slow or incomplete artificial stimuli have been used (e.g. Lea et

679   al., 1993, 2006), whereas otherwise quite similar studies using natural photographic

680   stimuli have found more rapid acquisition (e.g. von Fersen and Lea, 1990). It may be

681   that there is something about abstract geometric stimuli that makes it hard for pigeons

682   to learn about them. The present results do however rule out one explanation for the

683   ease of learning discriminations based on natural photographs, which is the possibility

684   that natural stimuli contain artefacts that enable the subjects to make an easy single-

685   dimensional discrimination, whereas the more fully controlled artificial stimuli do not.

686 In Experiment 2, all the discriminations could have been solved by the use of a single

687 dimension of demonstrated salience, and yet they were not easy for the pigeons.

688

689 **General Discussion**

690 Both of the present experiments examined the effect on pigeons' category

691 discrimination of providing a single, salient dimension, in effect a category label,

692 alongside more variable information that was nonetheless sufficient when taken as a

693 whole to support perfect discrimination. In Experiment 1, the label was a distinctive

694 texture, applied to parts of natural photographs. In Experiment 2, it was one of four

695 artificial geometric elements, distinguished by the fact that it had 100% validity as a

696 cue to reward, whereas the other elements only had 75% validity individually, though

697 collectively they were fully valid. In Experiment 1, some pigeons clearly based their

698 discrimination on the label, but slightly more of the pigeons used the pictorial

699 information instead. In Experiment 2, on the other hand, almost all the pigeons used

700 the label (the Reliable feature); of the small number that did not, most came under the

701 control of just one of the Unreliable features, and were thereby prevented from

702 achieving accurate discrimination.

703

704 The results of Experiment 1 demonstrate a tendency that is evident in the literature on

705 complex visual discriminations in pigeons; that detailed photographic material can

706 serve as highly effective stimuli in such discriminations, competing in effectiveness

707 with simple unidimensional cues on highly salient dimensions such as colour. On the

708 other hand, taking the results of the two experiments together, they do not support the

709 idea that pigeons discriminate photographs readily because the birds have a strong

710 tendency to integrate the wealth of different and possibly unreliable cues that pictures

711   contain. Rather, the results support the conclusions of Lea and Wills (2008), Wills et

712   al. (2009) and Lea et al. (2009), that pigeons have a strong tendency to come under

713   the control of single salient cues when exposed to complex discrimination tasks.

714   What photographs offer is a wide range of strong cues, so there is a good chance that

715   any pigeon will find one that is salient for it – salience being presumably, in part, a

716   function of the individual's previous experience. The carefully balanced design used

717   in Experiment 2 enables us to conclude that, other things being equal, if there is a

718   fully reliable cue available it is likely that it will achieve dominant control over

719   behaviour. Nonetheless, it is not certain that this will happen; it is possible for more

720   salient but less reliable cues to dominate, resulting in imperfect discrimination. The

721   extreme case of this is the occurrence of position habits in experiments where animals

722   have to make spatial choices in a discrimination task; by the design of the experiment,

723   position is normally a cue that carries no information at all, but it can still achieve

724   dominant control over behaviour that continues despite evidence that an animal has

725   detected the truly predictive contingencies (e.g. Mahut 1954).

726

727   Despite the strong tendency for pigeons to come under the control of single

728   dimensions, it is clearly not the case that they cannot be controlled by more than one

729   dimension. Nor is it the case that control by multiple dimensions only occurs when it

730   is essential for perfect discrimination, as in a compound discrimination. These

731   generalizations were again confirmed in the present experiments: In Experiment 1, it

732   appeared that a majority of the pigeons came under the control of both the picture and

733   the texture, and in Experiment 2 at least a few of the pigeons showed evidence of

734   control by more than one of the unreliable dimensions.

735

736 While control by multiple dimensions is not impossible, it seems from the present

737 results, and previous data, that it is not the natural tendency for pigeons. It may be

738 that it is easier for primates, and this would be one account of the difference between

739 our results and those of Couchman et al (2010) with rhesus monkeys.   If

740 multidimensional control does not come naturally to pigeons, tasks that require it,

741 such as polymorphous discriminations, are likely to be difficult for pigeons or other

742 birds to learn, and previous data support this position (e.g. Lea and Harrison 1978;

743 Lea et al. 2006). Lea et al. concluded that the reason was a limitation of attention: any

744 task that requires the processing of more than one stimulus dimension is inherently

745 difficult for a bird. The present data are consistent with that position.

746

747 It is possible that the difficulty of attending to multiple dimensions is exacerbated

748 when the dimensions are spatially separated, as in the present experiments. It is

749 notable that in the stimuli used by Couchman et al (2010), who obtained

750 categorization by overall similarity from rhesus monkeys, the four elements were

751 spatially contiguous, and this could be an alternative account of the difference

752 between their results and ours. Spatial grouping does affect pigeons: for example,

753 Sainsbury (1971) found that the feature positive effect in pigeons, which depends on

754 the elements of a stimulus being processed separately, was attenuated when the

755 elements were grouped more closely together. It is also possible that some kinds of

756 stimulus dimension are more separable than others for cognitive rather than perceptual

757 reasons (Soto and Wasserman 2010). However it is not a foregone conclusion that

758 spatial separation or cognitive compatibility will lead to a greater tendency towards

759 categorization by overall similarity: they could instead act to mitigate confusion

760 between the dimensions. This is a matter that would repay experimental

761 investigation.

762

763 The task used in the present Experiment 2 had many points in common with a

764 standard *m*-out-of-*n* artificial polymorphous discrimination task. The essential

765 difference was that there was an additional, completely reliable stimulus dimension (a

766 label in the terms used by Waxman and Markow, 1995), whereas in a standard

767 polymorphous task, all dimensions are equally unreliable. It is therefore not

768 surprising that the task was learned relatively quickly, especially when the complex

769 and abstract nature of the stimuli is taken into account. The speed of learning may

770 have been partly due to the multiple simultaneous presentation procedure, which is

771 known to produce faster learning than a simple go/no-go task (Huber et al. 2005,

772 Wills et al. 2009), and it would be interesting to try a standard polymorphous

773 discrimination using this procedure. However the present design does raise an

774 intriguing possibility. We normally think of the different dimensions of a stimulus as

775 competing for a subject's limited capacity for attention, and this is the basis of most

776 attention-based theories of discrimination learning, e.g. Sutherland and Mackintosh

777 (1971). Indeed, earlier attention-based theories assumed that animals could only

778 attend to a single stimulus dimension at once (e.g. Krechevsky, 1932). Our results are

779 certainly consistent with the idea that animals' learning of complex stimuli is limited

780 by their attentional capacity. It is conceivable, however, that a highly reliable

781 dimension could act to inform an animal that a task can be learned, and this might

782 heighten attention to other dimensions rather than diminishing it, especially if the

783 reliable dimension was removed once learning had taken place; something of this sort

784 might underlie the "labeling" phenomenon as it occurs in young children (Waxman

and Markow, 1995; Balaban and Waxman, 1997; Plunkett et al., 2008). This possibility gives more potential empirical bite to limited-attention theories, which can seem to add little to the empirical generalization that multiple-dimension discriminations are difficult. It could perhaps be investigated by exploring the mechanisms of attention in more detail by comparing the amount learned about unreliable stimulus dimensions in the presence or absence of more reliable cues.

References

Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM (1998) A formal neuropsychological theory of multiple systems in category learning. Psychol Rev 105:442-481

Blough DS (1972) Recognition by the pigeon of stimuli varying in two dimensions. J Exp Anal Behav 18:345-367

810 Bovet D, Vauclair J (2000) Picture recognition in animals and humans. Behav Brain

811   Res 109:143-165

812 Cardinal, RN, Aitken, MRF (2001) Whisker (version 2). Cambridge University

813   Technical Services Ltd, Cambridge

814 Couchman JJ, Coutinho MVC, Smith, JD (2010) Rules and resemblance: Their

815   changing balance in the category learning of humans (*Homo sapiens*) and

816   monkeys (*Macaca mulatta*). J Exp Psychol: Anim Behav Proc 36: 172-183

817 Dennis I, Hampton JA, Lea S (1973) New problem in concept formation. Nature

818   243:101-102

819 Herbranson WT, Fremouw T, Shimp, CP (1999) The randomization procedure in the

820   study of categorization of multidimensional stimuli by pigeons. J Exp

821   Psychol: Anim Behav Proc 25: 113-134

822 Herrnstein RJ, Loveland DH (1964) Complex visual concept in the pigeon. Science

823   146: 549-551. doi:10.1126/science.146.3643.549

824 Herrnstein RJ, Loveland DH, Cable C (1976) Natural concepts in pigeons. J Exp

825   Psychol Anim Behav Proc 2: 285-302. doi:10.1037.0097-7403.2.4.285

826 Huber L (2001) Visual categorization in pigeons. In: Cook RG (ed) Avian visual

827   cognition. [On-line], Available www.pigeon.psy.tufts.edu/avc/huber/

828 Huber L, Apfalter W, Steurer M, Prossinger H (2005). A new learning paradigm

829   elicits fast visual discrimination in pigeons. J Exp Psychol Anim Behav Proc

830   31, 237-246. doi:10.1037/0097-7403.31.2.237

831 Kemler Nelson, DG (1984) The effect of intention on what concepts are acquired. J

832   Verb Lng Verb Behav 23:734-759

833   Kirkpatrick-Steger K, Wasserman EA (1996) The what and the where of the pigeon's

834   processing of complex visual stimuli. J Exp Psychol Anim Behav Proc 22:60-

835   67. doi:10.1037/0097-7403.22.1.60

836   Krechevsky I (1932) Hypotheses in rats. Psychol Rev 39:516-532

837   Jenkins HM, Sainsbury RS (1970) Discrimination learning with the distinctive feature

838   on positive or negative trials. In: Mostofsky DI (ed) Attention: Contemporary

839   theory and analysis. Appleton-Century-Crofts, NewYork, pp 239-273

840   Lea SEG, Harrison SN (1978) Discrimination of polymorphous stimulus sets by

841   pigeons. Q J Exp Psychol 30:521-537

842   Lea SEG, Lohmann A, Ryan CME (1993) Discrimination of five-dimensional stimuli

843   by pigeons: Limitations of feature analysis. Q J Exp Psychol 46B:19-42.

844   doi:10.1080/14640749308401093

845   Lea SEG, Wills AJ (2008) Use of multiple dimensions in learned discriminations.

846   Comp Cogn Behav Rev 3, 115-133. doi:10.3819/ccbr.2008.30007

847   Lea SEG, Wills AJ, Leaver LA, Ryan CME, Bryant CML, Millar L (2009) A

848   Comparative analysis of the categorization of multidimensional stimuli: II.

849   Strategic information search in humans (*Homo sapiens*) but not in pigeons

850   (*Columba livia*). J Comp Psychol 123:406-420 . doi:10.1037/a0016851

851   Lea SEG, Wills AJ, Ryan CME (2006) Why are artificial polymorphous concepts so

852   hard for birds to learn? Q J Exper Psychol 59:251-267 doi

853   10.1080/02724990544000031

854   Mahut H (1954) The effect of stimulus position on visual discrimination by the rat.

855   Canad J Psychol 8:130-138

856   Makino H, Jitsumori M (2007) Discrimination of artificial categories structured by

857   family resemblances: a comparative study in people (*Homo sapiens*) and

858   pigeons (*Columba livia*). J Comp Psychol 121, 22-33. doi:10.1037/0735-

859   7036.121.1.22

860   Martin GR, Young SR (1983) The retinal binocular field of the pigeon (*Columba*

861   *livia*: English racing homer). Vision Research 23:911-915

862   Pothos EM (2005) The rules versus similarity distinction. Behav Brain Sci 28:1-49

863   Sainsbury RS (1971) Effect of proximity of elements on the feature-positive effect. J

864   Exp Anal Behav 16:315-325

865   Smith LB (1981) Importance of the overall similarity of objects for adults' and

866   children's classifications. J Exp Psychol Human Percep Perf 7, 811-824.

867   doi:10.1037/0096-1523.7.4.811

868   Smith LB, Kemler DG (1977) Developmental trends in free classification; Evidence

869   for a new conceptualisation of perceptual development. J Exp Child Psychol

870   24, 279-298. doi:10.1016/022-0965(77)90007-8

871   Soto FA, Wasserman EA (2010) Integrality/separability of stimulus dimensions and

872   multidimensional generalization in pigeons. J Exp Psychol Anim Behav Proc

873   36:194-205 doi :10.1037/a0016560

874   Sutherland NS, Mackintosh NJ (1971) Mechanisms of animal discrimination learning.

875   Academic Press, New York

876   Von Fersen L, Lea SEG (1990) Category discrimination by pigeons using five

877   polymorphous features. J Exp Anal Behav 54, 69-84.

878   doi:10.1901/jeab.1990.54-69

879   Wasserman EA, Bhatt RS (1992) Conceptualization of natural and artificial stimuli by

880   pigeons. In: Honig WK, Fetterman JG (eds) Cognitive aspects of stimulus

881   control. Erlbaum, Hillsdale NJ, pp 203-223

882   Weatherburn CE (1957) A first course in mathematical statistics, 2nd Edn corrected.

883   Cambridge University Press, Cambridge

884   Wills AJ, Lea SEG, Leaver LA, Osthaus B, Ryan CME, Suret MB, Bryant CML,

885   Chapman SJA, Millar L (2009) A comparative analysis of the categorization

886   of multidimensional stimuli: I. Unidimensional classification does not

887   necessarily imply analytic processing; evidence from pigeons (*Columba livia*),

888   squirrels (*Sciurus carolinensis*) and humans (*Homo sapiens*). J Comp Psychol

889   123:391-405. doi:10.1037/a0016216

890

891   Table 1. Experiment 2: The four forms of each dimension of the stimuli

Element	dimension of variation	X value	x value	y value	Y value
Annulus	width:radius ratio	5:1	4:1	2:1	1.5:1
Bar	length:width ratio	7.5:1	6:1	3.3:1	2:1
Chequerboard	number of elements	2 x 2	4 x 4	6 x 6	8 x 8
Diamond	orientation of stripes	vertical	60º	30º	horizontal

892

893

894   Table 2

895   Experiment 2: Use of stimulus dimensions in training stimuli for each pigeon

Pigeon	Reliable dimension (label)	Positive value	Unreliable dimensions	Positive values
Mo	Annulus	X	Bar, Chequerboard, Diamond	X, x
Ct	Annulus	X	Bar, Chequerboard, Diamond	Y, y
Ch	Annulus	Y	Bar, Chequerboard, Diamond	X, x
Kc	Annulus	Y	Bar, Chequerboard, Diamond	Y, y
Bn	Bar	X	Annulus, Chequerboard, Diamond	X, x
Mr	Bar	X	Annulus, Chequerboard, Diamond	Y, y
Sa	Bar	Y	Annulus, Chequerboard, Diamond	X, x
Cu	Bar	Y	Annulus, Chequerboard, Diamond	Y, y
Rg	Chequerboard	X	Annulus, Bar, Diamond	X, x
Hy	Chequerboard	X	Annulus, Bar, Diamond	Y, y
Yw	Chequerboard	Y	Annulus, Bar, Diamond	X, x
Bu	Chequerboard	Y	Annulus, Bar, Diamond	Y, y
Fl	Diamond	X	Annulus, Bar, Chequerboard	X, x
Cr	Diamond	X	Annulus, Bar, Chequerboard	Y, y
Dd	Diamond	Y	Annulus, Bar, Chequerboard	X, x
Jk	Diamond	Y	Annulus, Bar, Chequerboard	Y, y

896

897

898 Table 3. Example showing the Probe stimuli used in Experiment 2. In this example,

899 A was the Reliable dimension, with its X value as positive, and B, C and D were the

900 Unreliable dimensions, with their X and x values as positive. The dimensions used as

901 reliable and unreliable, and the values used as positive and negative, varied between

902 birds in a balanced fashion.

Stimulus dimension				
Annulus (Reliable)	Bar	Chequerboard (Unreliable)	Diamond	Number of times used
X	Y	Y	Y	6
Y	X	X	X	6
X	y	y	y	6
Y	x	x	x	6
X	X	Y	Y	2
Y	Y	X	X	2
X	Y	X	Y	2
Y	X	Y	X	2
X	Y	Y	X	2
Y	X	X	Y	2

903

904

Table 4. Experiment 2: Proportions of training trials within test sessions in which the

response was correctly predicted by each dimension of the stimulus

Reliable dimension	Pigeon	Positive stimulus (Reliable dimension and majority of unreliable dimensions)	Proportions of trials correctly predicted by value of			
			Individual unreliable dimensions			
			Annulus	Bar	Checks	Diamond
Annulus	Mo	0.64		0.63	0.58	0.58
Annulus	Ct	0.87		0.76	0.67	0.70
Annulus	Ch	0.98		0.77	0.77	0.78
Annulus	Kc	0.98		0.72	0.79	0.79
Bar	Bn	0.70	0.85		0.55	0.50
Bar	Mr	0.64	0.53		0.60	0.58
Bar	Sa	0.66	0.69		0.58	0.53
Bar	Cu	0.95	0.80		0.68	0.73
Checks	Rg	0.88	0.77	0.61		0.73
Checks	Hy	0.98	0.79	0.73		0.77
Checks	Yw	0.89	0.80	0.70		0.66
Checks	Bu	1.00	0.81	0.76		0.77
Diamond	Fl	0.86	0.72	0.68	0.68	
Diamond	Cr	0.46	0.41	0.61	0.38	
Diamond	Dd	0.86	0.74	0.70	0.68	
Diamond	Jk	0.71	0.77	0.64	0.51	

Figure Legends

909

910 Fig. 1. Experiment 1: Examples of stimuli showing each of the two picture types

911 (house and car) associated with each of the two artificially introduced textures

912 (leopard-skin and blue wave). Each stimulus measured 160 x 120 pixels. (From

913 original photographs by Casper Addyman, used with permission)

914

915 Fig. 2. Diagram of touch screen display, showing size and position of array and two

916 side keys. Cells were numbered 1-8, clockwise from top left. Cells 3-6 were

917 reinforced on the right of the screen: cells 7, 8, 1 and 2 reinforced on the left.

918

919 Fig. 3. Proportions of probe stimuli responded to in accordance with the texture cue.

920 Data are shown separately for each pigeon, separated by the stimulus type that was

921 positive in training (*=Difference from 0.5 significant at 0.05 level).

922

923 Fig. 4. Examples of the stimulus sets used in Experiment 2. The upper two panels

924 show the A and a versions of each dimension, the lower two panels the b and B

925 versions.

926

927 Fig. 5. Experiment 2: Proportions of probe trials in which the stimulus was responded

928 to in accordance with the Reliable stimulus dimension. Data are shown separately for

929 each pigeon, grouped by the dimension that was designated as Reliable (*=Difference

930 from 0.5 significant at 0.05 level).

931

932   Fig. 6. Experiment 2: Proportions of probe stimulus trials on which each of the

933   Unreliable dimensions correctly predicted response. Data are shown separately for

934   each pigeon, grouped by the dimension that was designated as Reliable. Gaps within

935   the cluster of bars for each pigeon correspond to the Reliable dimension

936

Figure 1.

Fig. 1 Experiment 1: Examples of stimuli showing each of the two picture types (house and car) associated with each of the two artificially introduced textures (leopard-skin and blue wave). Each stimulus measured 160 x 120 pixels. (From original photographs by Casper Addyman, used with permission)

Figure 2

Fig. 2 Diagram of touch screen display, showing size and position of array and two side keys. Cells were numbered 1-8, clockwise from top left. Cells 3-6 were reinforced on the right of the screen: cells 7, 8, 1 and 2 reinforced on the left.
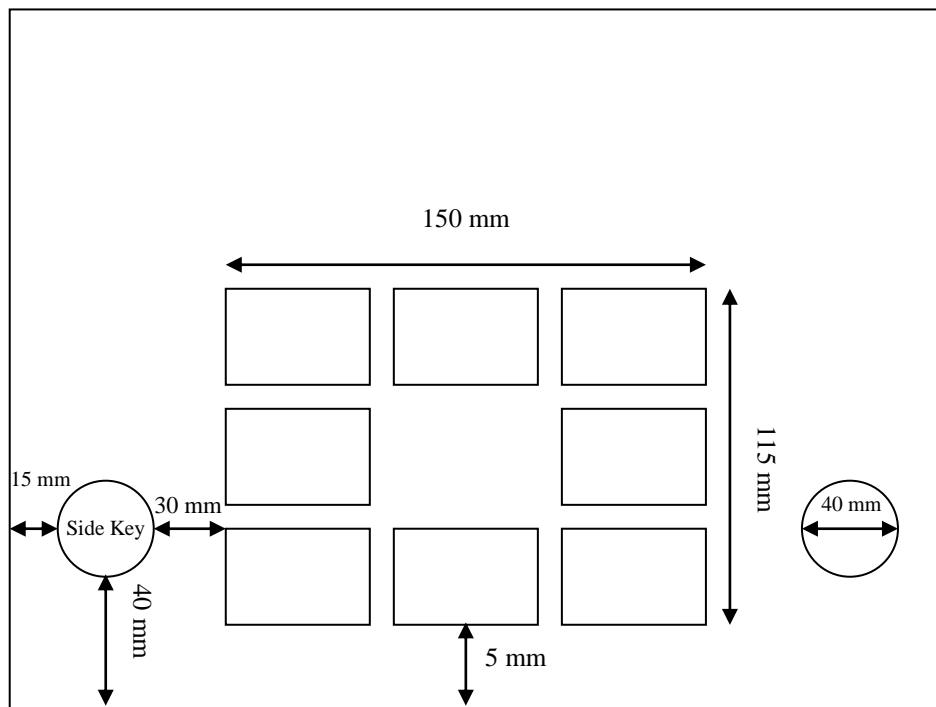
Figure 3

Fig. 3 Experiment 1: Proportions of probe stimuli responded to in accordance with the texture cue by each pigeon, separated by the stimulus type that was positive in training (*=Difference from 0.5 significant at 0.05 level).
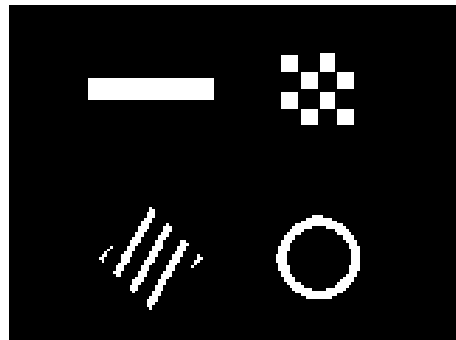
[see separate file]

Fig. 4. Examples of the stimulus sets used in Experiment 2. The upper two panels show the X and x versions of each dimension, the lower two panels the y and Y versions.
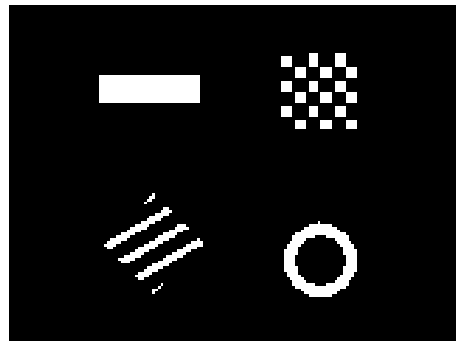
X form of all dimensions
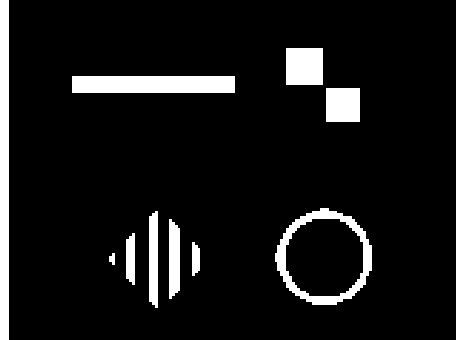
x form of all dimensions

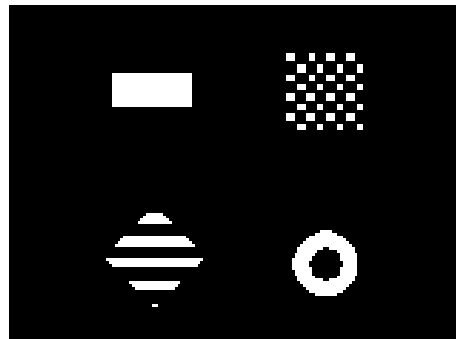y form of all dimensions

Y form of all dimensions

Fig. 5 Experiment 2: Proportions of probe trials in which the stimulus was responded to in accordance with the Reliable stimulus dimension. Data are shown separately for each pigeon, grouped by the dimension that was designated as Reliable (*=Difference from 0.5 significant at 0.05 level). Gaps within the cluster of bars for each pigeon correspond to the Reliable dimension

[see separate file]

Fig. 6. Experiment 2: Proportions of probe stimulus trials on which each of the Unreliable dimensions correctly predicted response. Data are shown separately for each pigeon, grouped by the dimension that was designated as Reliable.