Matching based on Biological Categories in Orangutans (*Pongo abelii*) and a Gorilla *(Gorilla gorilla gorilla)*

[p] Following a series of experiments in which six orangutans and one gorilla discriminatedphotographs of different animal species in a two-choice touch screen procedure,Vonk and MacDonald, (2002, 2004) concluded that orangutans, but not thegorilla, seemed to learn intermediate level category discriminations, such asprimates versus non-primates, more rapidly than they learned concrete leveldiscriminations, such as orangutans versus humans. In the current experiments,four of the same orangutans and the gorilla were presented with delayedmatching-to-sample tasks in which they were rewarded for matching photos ofdifferent members of the same primate species; golden lion tamarins, Japanesemacaques, and proboscis monkeys, or family; gibbons, lemurs (Experiment 1), andsubsequently for matching photos of different species within the followingclasses: birds, reptiles, insects, mammals, and fish (Experiment 2). Members ofboth Great Ape species were rapidly able to match the photos at levels abovechance. Orangutans matched images from both category levels spontaneouslywhereas the gorilla showed effects of learning to match intermediate levelcategories. The results show that biological knowledge is not necessary to formnatural categories at both concrete and intermediate levels.

2   Jennifer Vonk

3   Oakland University

4   **Please address correspondence to:**

5   **Jennifer Vonk**
6   **Dept. of Psychology**
7   **Oakland University**
8   **2200 N Squirrel Rd.**
9   **Rochester MI, 48309**
10   **vonk@oakland.edu**

11   **Introduction**

12   Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976) suggested a chronological and

13   hierarchical structure for concepts, such that humans first learn basic level concepts (e.g. dog)

14   and only later are able to learn subordinate (e.g. poodle) and superordinate (e.g. mammal or

15   animal) concepts (see also Mervis & Rosch, 1981). This chronology of concept learning seems

16   tied to language in that the basic level category words are the first learned and the most

17   commonly used. It has been suggested that the formation of superordinate categories relies less

18   upon perceptual feature analysis and more on an understanding of how the category coheres,

19   across significant perceptual variance (Spalding & Ross, 2000). Superordinate categories are thus

20   thought of as being more conceptually based or "abstract" relative to basic level categories, and it

21   has been speculated that language is necessary for the formation of these later categories (Benelli,

22   1988; Keil, 1988; Nelson, 1988; Premack, 1983). Organisms may be inherently preprogrammed

23   to distinguish between items at the level of the basic concept perceptually (Cerella, 1979; Eimas

24   & Quinn, 1994). However, Mandler (2000) has articulated an opposing view, supported by

25   several studies, in which the conceptualization of superordinate categories by human infants may

26   actually precede that of basic level categories. Mandler's description highlights the distinction

27   between perceptual and conceptual categorization, a distinction not traditionally acknowledged

28   by other researchers. In her view, exemplars within basic level categories may be associated

29   together *on a perceptual basis* before exemplars from superordinate categories are associated

30   perceptually. In contrast, when forming *conceptual* categories, which are based on shared,

31   underlying properties as opposed to perceptual similarity, more global, abstract categories such as

32   animals, foods etc. may emerge first. This may be the case because children learn about broad

33   categories, such as animate/inanimate distinctions prior to learning specific distinctions such as

34   between reptiles and mammals.

35   Recent work (Coley, 2007) has indicated that children may make global distinctions such

36   as between animals and non-animals, but are less likely, compared to adults, to consider humans

37   to be similar to other primates or non-primate animals. This recent work, highlighting differences

38   between children and adults with regard to between categorical judgments of similarity,

39   reinforces the notion that such categorization is not made on a purely perceptual basis. In

40   addition, Gelman and Davidson (2013) have shown that both children and adults use category

41   membership more than similarity to make basic level category inferences, although their studies

42   involved experimenter-created rather than actual natural categories. Thus there is some

43   disagreement as to what sorts of categorization depend upon physical similarity, depend upon

44   language and emerge first in the human infant. By studying concept formation in non-human

45   primates one can determine parallels in the emergence of non-language based categorization. In

46   the current study, two species of great ape (orangutans and a gorilla) were required to match

47   images based on biological classifications at the level of species, family or class.

48   Although taxonomic class groupings may be considered basic level concepts (Roberts &

49   Mazmanian, 1988), it has been suggested that the ability to make taxonomic classifications of

50   natural stimuli depends upon biological or scientific knowledge (Eimas & Quinn, 1994;

51   Hampton, 1998; Inagaki, 1989), as well as language (Anggoro, Medin, & Waxman, 2010;

52   Benelli, 1988; Gelman, 1989; Nelson, 1988). Quinn and Tanaka (2007) found that expertise

53   within the same basic level category aided infants in the ability to discriminate other concrete

54   level categories, emphasizing the role of expertise in category discrimination. Coley (2007)

55   showed that children of eight years or older were more likely to categorize humans as being

56   similar to primates and primates as being more similar to mammals than non-mammals, but

57   younger children were not sensitive to these taxonomic groupings, further supporting the notion

58   that intermediate level biological categories are learned rather than perceived. At least in humans,

59   it appears that discriminations are not made solely on the basis of similarity of perceptual

60   features. Categorizing stimuli according to biological taxonomies may then be presumed to be a

61   uniquely human tendency.

62   Against this supposition, exciting recent work from neuroscience has demonstrated a

63   common code for inferior temporal (IT) object representations in monkeys and humans (Kiani et

64   al., 2007; Kriegeskorte et al., 2008). This work has demonstrated global representations along a

65   continuum of inanimate to animate objects, but also within category distinctions between

66   different taxonomic groups. Connolly et al. (2012) have suggested a continuum of activation

67   representing categorical structure from insects to primates that mirrors the continuum from

68   inanimate to animate objects. Connolly and colleagues suggest a categorical structure within the

69   domain of animate objects that reflects the biological relations among species, suggesting that

70   such categorization may be innate within primates. Furthermore, Murai and colleagues (Murai et

71   al., 2004, Murai et al., 2005), using a familiarization-novelty preference task, have suggested that

72   infant monkeys, chimpanzees and humans may spontaneously form categories at least at the

73   global level, using categories of mammals, furniture and vehicles. The current experiments

74   investigated the ability to make explicit classifications of more finite natural class distinctions in

75   two other species of Great Ape; orangutans and gorillas. In the case of non-humans, categories

76   would be based on shared observable features rather than on underlying knowledge of taxonomic

77   class structures, particularly given that the only information provided involves visual features in

78   two dimensional photographs. However, it was of interest to determine whether exemplars of

79   more closely related groupings are more readily categorized together compared to more distantly

80   related members of the same class. We predicted that orangutans may readily categorize stimuli

81   from both concrete and intermediate level categories, whereas the gorilla might categorize stimuli

82   more readily at the concrete level.

83   Subordinate categories, for example, '*poodles*', are described as being concrete (or least

84   abstract) along a hierarchy of abstraction because exemplars within such categories share many

85   readily perceived features and can be easily distinguished from exemplars from other categories

86   on a purely perceptual basis. Basic level categories, such as 'dogs', are described as being

87   intermediate in terms of abstractness. The variability within an intermediate category is greater

88   than the variability within a concrete level category. For instance, the more general category of

89   dogs includes not only poodles but hounds, terriers, and many other types of dogs, which may

90   vary in terms of size, color and distinctive features such as the long, short body of the dachshund.

91   However, category distinctiveness is also increased such that members within an intermediate

92   level category share fewer features with members of other intermediate level categories, whereas

93   concrete exemplars may share many features with exemplars from other concrete level categories

94   subsumed within the same intermediate category. For example, toy poodles may look a lot like

95   the bichon frise but dogs may not look much like other mammals such as whales or marsupials.

96   At the most abstract level, such as 'animal', superordinate category exemplars share even fewer

97   perceptual features within a category but also share even fewer features with members of other

98   abstract categories. The category 'mammal' may be considered superordinate to the basic level

99   'dog', but, in keeping with previously published studies (Roberts & Mazmanian, 1988; Vonk &

100   MacDonald, 2002, 2004; Vonk, Jett & Mosteller, 2012; Vonk, Jett, Mosteller & Galvan, 2013),

101   'mammals' will be considered intermediate with 'animal' being considered the most abstract level

102   category.

103   Of course, level of expertise determines category level to some degree. For an expert

104   herpetologist, insects may be easily categorized at much finer levels than would be the case for a

105   third grade science student, or a novice adult, for that matter. Thus, while the most generic level

106   of categorization e.g. insects may be a basic or intermediate level of category for the novice, a

107   more highly specified category such as hymenoptera may serve as a basic level category for the

108   herpetologist, with categories of wasps, bees, and ants – or even particular species of each,

109   serving as subordinate or concrete levels. For the purpose of describing the current study, it is

110 assumed that apes have no special expertise regarding unfamiliar members of the primate order,

111 or broader class of mammals, birds, fish, insects, so intermediate level will be used to refer to

112 class while concrete level will be used to refer to species discriminations.

113   Often abstract level categories can be discriminated only with some additional conceptual

114 knowledge, rather than by relying on perceived shared attributes. For instance one would not

115 know that insects, birds, amphibians, and mammals all belonged to the category 'animal' if one

116 was unaware of their unobservable properties such as the ability to eat, breathe, reproduce etc.

117 The ease with which category membership can be determined on strictly a perceptual basis thus

118 declines as one moves from concrete to abstract level categories. The terms concrete and

119 subordinate, intermediate and basic, and abstract and superordinate have been used

120 interchangeably. The former terms will be used exclusively throughout the remainder of this

121 manuscript in keeping with previous research (Vonk & MacDonald, 2002, 2004; Vonk et al.,

122 2012; Vonk et al., 2013).

123   There is much evidence for concrete level discrimination learning in primates, (Fujita,

124 1987; Fujita & Matsuzawa, 1986; Fujita, Watanabe, Widarto & Suryobroto, 1997; Marsh &

125 MacDonald, 2008; Yoshikubo, 1985) as well as in pigeons (Herrnstein, 1979; Herrnstein,

126 Loveland, & Cable, 1976), but there is much less evidence for intermediate or abstract level

127 natural concept discrimination learning (although see Brooks et al., 2013 for evidence in rats).

128 The difficulty with Brooks et al. (2013) is that all of the discriminations were also between

129 animate and inanimate categories (e.g. chairs versus flowers and cars versus humans) – allowing

130 the rats to make more global level discriminations. Recently, investigators (Autier-Dérian,

131 Deputte, Chalvet-Monfray, Coulon, & Mounier, 2013) have shown that domestic dogs are

132 capable of categorizing many diverse dog species together into a single basic level "dog"

133 category. The dogs were also able to discriminate between the species, demonstrating concrete

134 level categorization. Caution is appropriate when making cross species comparisons as the reader

135   must take into account possible differences in perceptual systems in different organisms.

136   However, the species tested in relevant paradigms share sophisticated visual acuity and color

137   vision making them suitable candidates for categorical discrimination research. In a more

138   relevant study, Tanaka (2001) demonstrated that chimpanzees were able to discriminate

139   exemplars both *within* intermediate level categories (at the concrete level) and *between*

140   intermediate categories. The chimpanzees were trained to match exemplars that belonged to the

141   same concrete level category and subsequently matched those exemplars to other members of the

142   same intermediate, but different concrete level category, when a concrete level match was no

143   longer an option. Because Tanaka's chimpanzees were trained to make concrete level matches

144   and then tested on intermediate level matches to the same stimuli it is difficult to compare the

145   chimpanzees' performance on the two levels of discrimination.

146   In one of only five published attempts to compare concept learning at different levels of

147   abstraction in non-humans, Roberts and Mazmanian (1988) studied concept discrimination by

148   pigeon, squirrel monkey and human subjects across three levels of abstraction, where abstraction

149   was defined as the breadth of the category to be learned. At the concrete level, the subjects were

150   asked to discriminate between photographs of one bird species (kingfishers) and other birds. At

151   the intermediate level they were asked to discriminate between birds and other animals. Finally,

152   at the most abstract level they were asked to discriminate between animals and non-animals. The

153   authors found that humans easily discriminated concepts at all levels of abstraction, whereas both

154   monkeys and pigeons had difficulty with the intermediate level discrimination. This finding was

155   somewhat surprising because this intermediate level corresponds to the basic level that Rosch

156   presumed was easiest for humans to learn (Keil, 1988; Mervis & Rosch, 1981; Rosch et al.,

157   1976). Ross, Medin, Coley, and Atran (2003) suggest that children acquire such concepts by the

158   age of six years. However, the results may confirm Mandler's predictions (2000), particularly if

159   the apes are performing conceptual rather than perceptual categorizations.

160   Vonk and MacDonald (2002, 2004) trained orangutans and a gorilla to discriminate

161   between photographs of members of their own species and humans, and between members of

162   their own species and other primates (concrete level), between primates and other species

163   (intermediate level) and between animals and non-animals (abstract level). The orangutans

164   quickly learned the concrete and intermediate level discriminations and also learned the most

165   abstract discrimination, but with slightly more difficulty. The gorilla subject appeared to have the

166   most difficulty with the intermediate level discrimination, although she also learned all

167   discriminations and showed significant positive transfer to novel stimuli. Thus both species of

168   Great Ape demonstrated the ability to learn concepts at each level of abstraction, but the

169   orangutans appeared to learn the basic or intermediate level discrimination the most readily while

170   the gorilla learned it with the most difficulty. More recent follow-ups have shown chimpanzees to

171   have the most difficulty with the most abstract discriminations (Vonk et al., 2013) in that they

172   required more sessions to reach criterion as discriminations became more abstract, and were less

173   likely to show significant transfer. In contrast, black bears showed significant transfer at all levels

174   of concept discrimination, even when trained on the most abstract problems first (Vonk et al.,

175   2012).

176   Some of the orangutans in the previous study (Vonk & MacDonald, 2004) learned to

177   select photos of primates very rapidly (e.g. within three 10-trial sessions), despite the fact that the

178   primates presented to them were of unfamiliar species. It was possible that they simply selected

179   photos that they preferred.. A spontaneous preference indicates that subjects may not have learned

180   to abstract particular concepts or categories as being "correct" within the context of the

181   experiment but implicitly preferred photos that happened to belong in the same taxonomic group.

182   However, the very existence of a preference suggested that they did perceive primates as distinct

183   from other species regardless of whether they explicitly recognized primates as belonging to a

184   coherent category. Another possibility was that they selected photos of primates because these

185 photos were more similar to photos that had previously been presented, although selecting these

186 photos had not been reinforced. The latter explanation is unlikely, because of latent inhibition, the

187 finding that it becomes more difficult to learn to respond to previously unreinforced stimuli due

188 to retroactive interference. The fact that this was not the case for the orangutans suggests that

189 they were attending to the category of the photo and not the specific exemplar itself. In addition,

190 different and diverse primate species were presented in each subsequent transfer photo set, and

191 transfer performance remained high. However, the question remained as to whether or not these

192 apes would be equally likely to learn other intermediate level discriminations, including those

193 more analogous to the bird/other animal discrimination tested by Roberts and Mazmanian (1988).

194 Both chimpanzees and black bears were tested on an intermediate problem in which the two

195 categories were equally inclusive and novel (carnivores for chimpanzees, or primates for bears,

196 versus ungulates). The black bears performed better than the chimpanzees in terms of acquisition

197 and transfer, suggesting the possibility of different mechanisms for forming the discriminations.

198 Here, it was of interest to directly contrast the acquisition of matching intermediate level concepts

199 to their acquisition of matching concrete level concepts in an identical procedure for the

200 orangutans and gorilla tested previously.

201 The subjects were presented with a delayed matching-to sample (DMTS) task in order to

202 examine their understanding of several categories simultaneously. In contrast to the previous

203 studies, where the subjects learned to discriminate a single category at a time, use of a DMTS

204 procedure made it possible to have them categorize stimuli from several different species or

205 classes within a single session. Each trial involved presentation of a sample that was replaced by

206 two comparison photos once the subject attended to the sample. The reinforced comparison was a

207 different photo of the same or different member of the same species or family as the sample

208 (Experiment 1), or a photo of a different species from the same class as the sample (Experiment

209 2). The non-reinforced comparison was a photo of a member of a different primate species

210   (Experiment 1), or a member of a different class (Experiment 2). Experiment 1 tested for concrete

211   level discriminations in that the categorizations could be made by matching perceptual features of

212   the stimuli. In Experiment 2 an effort was made to select photographs of species belonging to the

213   same taxonomic class that were nonetheless perceptually quite distinct from one another. For

214   instance, a photo of a stingray shared few features with a photo of the head of a blenny fish, but

215   both belonged to the fish category. Backgrounds varied both within and between categories.

216   Because the comparison stimuli did not share many perceptual features with the sample stimulus,

217   the subjects' ability to correctly match the photos might indicate the capacity for forming

218   concepts at the intermediate level of abstraction.

219   Experiment 1

220   In Experiment 1, subjects were presented with photos of five different primate species

221   (Japanese macaques, golden lion tamarins, proboscis monkeys) or families (gibbons, lemurs) and

222   were rewarded for matching based on the species or family. This was a concrete level

223   discrimination in which the exemplars within a category shared several physical features, such as

224   color, and body shape. However, in two of the categories, (lemurs and gibbons) members of the

225   same genus but different closely related species were included, and it was predicted that the

226   subjects might have more difficulty when the match was a close relative compared to when the

227   match was of the same species.

228   *Materials and Methods*

229   *Subjects.* One female western lowland gorilla, Zuri, age 4, three male Sumatran

230   orangutans, Dinding, (44 years), Dinar, (13 years), Molek, (22, years) and one female Sumatran

231   orangutan, Abby, (42 years), participated in these experiments. All subjects were housed at the

232   Toronto Zoo, Toronto, Ontario, Canada. The orangutan subjects were group-housed in an indoor

233   exhibit and Zuri was housed separately at the time of testing, although she had auditory and

234   visual access to the other gorillas at the zoo. The gorillas had access to both indoor and outdoor

235 exhibits. Occasionally Zuri was integrated with the other gorillas for brief periods. Orangutans

236 could view gibbons from their exhibit, as well as various bird, reptile, fish, and insect species.

237 Dinding and Molek had been housed at the Yerkes primate research center many years prior and

238 could see chimpanzees and bonobos from their enclosures there. Zuri could not view other

239 primates from her enclosure but also had exposure to birds, reptiles, and insects. Four of the

240 subjects had participated in one prior touch screen experiment (Vonk & MacDonald, 2002, 2004),

241 whereas Abby had participated in two prior touch-screen experiments (Vonk, 2003; Vonk &

242 MacDonald, 2004). Only Abby had previously participated in a DMTS procedure (Vonk, 2003).

243 Testing was approved by the Animal care review board of York University, Canada under the

244 direction of Suzanne MacDonald.

245   *Materials.* The photo set included 20 color photographs of 5 different primate species.

246 There were 4 gibbon photos; two white-handed gibbons (*Hylobates lar*), one Mueller's or Gray

247 gibbon (*Hylobates muelleri*) and one dark-handed gibbon (*Hylobates agilis*). There were four

248 lemur photos; three ring-tailed lemurs (*Lemur catta*), and one collared brown lemur (*Eulemur*

249 *fulvus collaris*). Examples of these images appear in Figure 1. There were also four photos each

250 of Japanese macaques (*Macaca fuscata fuscata*), golden lion tamarins (*Leontopithecus rosalia*),

251 and proboscis monkeys (*Nasalis larvatus*). Each species was depicted in various positions and

252 orientations and in both close-up and far away shots. Some photos showed only the face of the

253 subject whereas other photos showed the full body. Most photos depicted an individual whereas

254 some included several individuals of the same species. The backgrounds of the photos varied

255 both within and between categories. By varying such dimensions the extent to which irrelevant

256 features might control responding in the task was minimized. A list and brief description of the

257 photos used appears in Appendix A. All photos were novel for all subjects.

258   *Procedure.* The experiment was programmed in Filemaker Pro 3 software for Macintosh.

259 The photos were presented on a 13" Apple touch screen monitor and were approximately 3" by

260   4" on the screen, separated by approximately 1.5" and horizontally aligned. The monitor was

261   placed against the bars of the subjects' housing and they were required to either reach through the

262   mesh holes to touch the screen (orangutans) or to reach underneath and touch the monitor

263   (gorilla). In these experiments the experimenter sat behind the laptop, which was covered by a

264   protective covering, connected to the touchscreen which was pushed right up against the subjects'

265   enclosure. The images on the touchscreen were mirror-reversed from the images on the laptop

266   and the experimenter always gazed directly at the midpoint of the screen. The experimenter could

267   not see the subject's face or fingers, or the front of the touchscreen, when they made a response

268   so could not react to the correctness or incorrectness of the choice until after the choice was

269   made.

270   Subjects were tested individually at the same time each day. The orangutans received one

271   to four sessions per day, two or three times a week, whereas the gorilla was given between five

272   and ten sessions a day four days a week. Number of sessions was dependent on the availability

273   of the subjects and their keepers. Each session consisted of ten trials. During a trial, a sample

274   photograph was presented in the center of the monitor and stayed on screen until the subject

275   attended to the photo and touched it, activating the touch screen. The two comparison photos

276   subsequently appeared on the screen after a short delay (approximately 3 seconds). The subject

277   was then required to select, by touching, only one of the two comparison photos. If he or she

278   selected the photo that matched the species of the sample he or she was given a small highly

279   preferred food reward by hand (M&Ms or dried fruits and nuts for the orangutans, and dried

280   fruits or nuts for the gorilla) during presentation of a blank screen. If the subject made an

281   incorrect response the screen advanced immediately to the blank screen and then to the next

282   sample photo with no reward and no time-out. Thus intertrial intervals varied but were always

283   less than one minute. Sessions continued until all ten trials were completed. Intersession intervals

284   also varied but were always at least two minutes in length.

285   During each session, half of the photos were presented once and half of the photos were

286   presented twice. If a photo had appeared as a correct comparison, it appeared as either a sample

287   or as an incorrect comparison the next time it appeared within that session. This method

288   discouraged a strategy of attending to individual exemplars and encouraged attending to the

289   relationship between the sample and the comparison stimuli. Order of presentation and pairing of

290   the photographs was randomized for each session. Thus each photo appeared as a sample, as a

291   correct comparison or as an incorrect comparison during the course of the experiment. Each

292   photo also appeared in both left and right positions on the screen across sessions. Within a

293   session, half of the correct comparisons appeared on the left side of the screen, and half appeared

294   on the right. Photos that appeared twice during some sessions appeared only once during other

295   sessions.

296   Each species in the photos was represented in the sample twice and thus as a correct and

297   incorrect match twice as well within each session. Thus gibbons, proboscis monkeys, tamarins,

298   lemurs and Japanese macaques each appeared as samples on two of the ten trials within a session

299   in random order. The photo used as a correct match was always different from the sample photo

300   so that there were no identity matching trials. The same photos were used on each session.

301   For the first two sessions of the task, the subjects (except for Abby, due to previous

302   training on a DMTS task) were given a small reward for simply touching the sample photo, as

303   well as for making the correct choice. This was done so that the animals would learn to attend to,

304   and select, the sample photo. After the second session, touching the sample photo was no longer

305   reinforced. The experiment was considered complete when the subject was performing

306   consistently, after a minimum of four blocks of five sessions (20 sessions or 200 trials). Subjects

307   received four to six blocks of five sessions, depending on their level of performance. Abby

308   completed the two experiments simultaneously as a control for the order of testing. She first

309   completed five sessions of Experiment 2 and then five sessions of Experiment 1. From that point

on, she completed one or two sessions of each experiment on each test day. Order of presentation

of the two tasks was not counterbalanced for the other subjects because it was felt that

presentation of the more visibly similar samples and exemplars in the concrete discrimination

task would facilitate acquisition of the DMTS task that they had no prior experience with. Recall

that these four subjects had not received any prior training on MTS procedures and none of the

five had ever received identity-matching trials.

Results and Discussion

As shown in Figure 2 and confirmed by binomial tests, each subject performed at a level

significantly greater than chance (50% correct) overall, all p's < .001. Separate binomial tests for

each subject also indicated how many sessions were required to reach levels of responding that

were significantly above chance. Molek's performance was significantly above chance by the

second session, N = 20, p = .04. Dinding's performance was significantly above chance by the

fourth session, N = 40, p = .02. Abby and Zuri were performing significantly above chance by the

sixth session, N = 60, p = .007 and .05 respectively. For Dinar, performance did not exceed

chance levels until the 18th session, N = 180, p = .04.

The fact that four of the five subjects reached above chance levels within the first six

sessions of testing is impressive, particularly because the correct category differed on every trial.

In addition, no training on the DMTS procedure occurred prior to the first session of testing. By

contrast, subjects in similar experiments typically undergo extensive training with identity

matching before being tested in conceptual matching procedures.

Recall that stimuli were randomly paired within each trial so that sometimes stimuli that

had been paired on previous trials in prior sessions were re-paired and sometimes pairings were

novel. It is possible that the subjects did not map the photographs on to concepts for each unique

species, but, rather, that they were rapidly able to learn associations between particular exemplars

based on reward contingencies. In order to argue against this latter interpretation, paired t-tests

were conducted to show that performance on novel stimulus pairings did not differ from

performance on previous pairings, for the first six sessions, for any of the individual subjects,

highest $t_2 = 1.67$, $p = .24$. Only the first six sessions were considered because, after that, the

likelihood of novel pairings decreased substantially. These data appear in Table 1.

In addition, one sample t-tests compared performance on both novel and prior pairings to

chance (50%) for each subject for the first six sessions. Only Molek was above chance on both

novel ($M = 84\%$, $SD = 15\%$, $t_4 = 5.09$, $p = .007$), and prior pairings ($M = 78\%$, $SD = 9\%$, $t_4 = 6.78$, $p = .002$). Abby was above chance on only novel pairings ($M = 78\%$, $SD = 18\%$, $t_4 = 3.07$,

$p = .05$). The other three subjects were not above chance on either novel or prior pairings within

the first six sessions. These results lend no support to the idea that subjects relied upon

associations that they formed during the course of the experiment between particular stimuli

pairings and reward in performing this task.

In order to determine if subjects were differentially accurate at matching photos

depending on the species or family of the animal depicted, separate univariate ANOVAs of the

subjects' scores on each session, with sample species (species of the animal in the sample

stimulus) and incorrect species (species of the animal in the non-reinforced comparison stimulus)

as independent variables were conducted for each subject. ANOVA assumes normality of data

and homogeneity of variance – conditions that were satisfied by the current data. Table 2 displays

the average scores for each subject for each type of discrimination, according to which specie

was depicted in the sample. Dinding was influenced by both the species of the sample stimulus,

$F_{4, 25} = 5.66$, $p = .002$, and of the non-reinforced comparison, $F_{4, 25} = 4.23$, $p = .01$. He scored

above 66% for all species except when the sample was a Japanese Macaque; on those trials he

was below chance. He also tended to score near chance levels when the non-reinforced

comparison was a proboscis monkey or a golden lion tamarin, indicating that he preferred to

select those images.

360   Dinar's choices were influenced by the sample stimulus, $F_{4, 25} = 8.7$, $p = .001$. He

361   matched accurately when the sample photo was of a golden lion tamarin but matched

362   inaccurately with every other type of sample. He was apparently distracted by the non-reinforced

363   comparisons, $F_{4, 25} = 8.78$, $p = .001$. He made many errors when the lion tamarin was the

364   incorrect choice, indicating a preference for photos of tamarins regardless of whether they were

365   the correct matches or not. He also tended to select photos of gibbons and lemurs when they were

366   not correct choices. It is not unexpected that he might prefer to choose gibbons over other

367   primates because gibbons were visible from the orangutan exhibit and were the only other

368   primate species, other than humans or orangutans that the orangutans were familiar with. Dinar's

369   results suggested that he did not use a generalized concept to perform the task, but are not

370   inconsistent with an ability to discriminate the different species of primates.

371   Molek's choices were affected by the non-reinforced comparison, $F_{4, 20} = 5.63$, $p = .003$.

372   He performed accurately on all trials except when the non-reinforced comparison depicted a

373   Japanese macaque, indicating that he selected the Japanese macaque photos often when they were

374   not correct.

375   Abby's choices were not significantly affected by the sample species or by the non-

376   reinforced species, although the effect of non-reinforced species approached significance, $F_{4, 20} =$

377   2.64, $p = .08$.

378   Zuri's choices were influenced by the sample species, $F_{4, 15} = 13.06$, $p = .001$. She scored

379   above 70% on all trials except for when the sample photos were of proboscis monkeys. She was

380   also affected by the species depicted by the non-reinforced comparison photos, $F_{4, 15} = 3.83$, $p = .$

381   02. She chose accurately on trials when proboscis monkeys were depicted in the incorrect

382   comparisons, consistent with the idea that she simply avoided photos of proboscis monkeys

383   regardless of whether they were correct matches or not. She also did well when tamarins were the

384   non-reinforced comparisons but not as well when gibbons, Japanese macaques or lemurs were the

385   non-reinforced comparisons.

386   None of the subjects, with the exception of Dinar, appeared to have significantly greater

387   difficulty with the gibbon or lemur species, despite the fact that these photo sets included

388   different species within the same family. This finding implies that the greater perceptual

389   difference between the sample stimulus and the comparison stimulus did not influence

390   responding for most of the subjects.

391   The results indicated, for the most part, that the subjects matched above chance at this

392   concrete level. They sometimes displayed idiosyncratic preferences for photos of particular

393   species but, overall, they discriminated between members of various primate species, consistent

394   with earlier research in which they selected photos of orangutans or gorillas and avoided

395   selecting photos of other primates (Vonk & MacDonald, 2002, 2004). The present experiment

396   extends this finding to photographs of unfamiliar species, and to a new procedure. Use of a

397   DMTS task has an advantage over the two-choice procedure in that no single exemplar is

398   associated with reward or non-reward. Instead it is the relationship between, or the pairing of,

399   stimuli that determines the reward contingency. The fact that the subjects were no less accurate

400   with novel than prior pairings is thus an important finding and argues against simple stimulus

401   reward associations. Other experiments with the same subjects revealed that they do not learn all

402   DMTS tasks equally rapidly (Vonk, 2002; 2003), implying that there is something special about

403   the nature of these discriminations making the categories readily perceivable by these species.

404   The subjects differed from each other in terms of which species they preferred to select

405   and which species they had the most difficulty matching. In general, however, they seemed to

406   have difficulty with the proboscis monkey photos, which was unexpected because those photos

407   were mostly, although not exclusively, of the same animal from various positions but with the

408   identical background, whereas photos of the other species varied along more dimensions and

409 sometimes even depicted different species. This result therefore argues against a reliance on

410 perceptual similarities to make the discriminations or attention to irrelevant background details.

411 The difficulty may have stemmed from the fact that the proboscis monkey subjects tended to fill

412 a smaller percentage of the entire photo.

413   Above chance levels of performance were obtained within the first six sessions by all but

414 one of the subjects, indicating that accurate responding did not require extended training or

415 learning of associations between the exemplars. Performance was no more accurate on trials

416 where specific exemplars had been paired before than on those trials on which stimulus pairings

417 were novel. However, Dinar's pattern of responding suggested that he often simply selected

418 photos that he preferred. Concrete level discriminations can be made on a perceptual basis and

419 are not necessarily demonstrative of abstract concepts representing the species depicted. Thus

420 performance may be expected to decline in Experiment 2 where photos had to be matched

421 according to intermediate level categories. Exemplars within these categories shared fewer

422 physical features.

423   Experiment 2

424   The second experiment was of critical interest. Few researchers have examined whether

425 non-human primates spontaneously classify other species according to biological or taxonomic

426 categories (see Brown & Boysen, 2000 for one example). Could these subjects match photos of

427 members of various taxonomic classes (such as birds, reptiles, insects, mammals, and fish)

428 despite the fact that exemplars within a category would share only some features, and may also

429 share features with exemplars from another category? This experiment tested their abilities to

430 form concepts at an intermediate level of abstraction, and might corroborate previous findings

431 (Vonk & MacDonald, 2002, 2004).

432 *Method*

433   *Subjects.* The subjects were the same five animals that participated in Experiment 1.

434 *Materials.* The photo set included thirty novel color photos, six from each of the

435 following taxonomic class categories: birds, insects, mammals, fish and reptiles. Photos included

436 single or several individuals, pictured close-up or at a distance, faces or entire bodies. The

437 subjects in the photos were also pictured in a variety of orientations and postures. Figure 3

438 depicts sample images from the categories of birds and reptiles. Each photo appeared once during

439 each session. Whether or not the photo appeared as a sample, correct, or incorrect choice was

440 randomly determined on each session. Photos were also randomly paired. Within each session

441 one different exemplar from each taxonomic group appeared as a sample twice and as an

442 incorrect choice twice as well. Thus, out of ten trials within a session, two of the sample photos

443 were fish; two were birds, insects, mammals and reptiles. The same thirty photos were used in

444 each session. A list and description of the photographs appears in Appendix B.

445 *Procedure.* The procedure was the same as in Experiment 1, except for the different

446 materials noted above. In addition, the subjects did not have to be rewarded for touching the

447 sample photo on the first two sessions, because they had already mastered the DMTS procedure.

448 Each subject received four or five blocks of five sessions depending on the number of sessions

449 required to reach a stable level of responding.

450 *Results and Discussion*

451 As shown in Figure 4, and confirmed with binomial tests, each subject performed

452 significantly above chance (50%) overall, N = 200 or 250, all *p*'s < .001. Individual binomial

453 tests were conducted to determine how many sessions were required for each subject to achieve

454 above chance levels of performance. Molek and Dinar required only two sessions to perform

455 above chance, N = 20, both *p*'s = .04. Abby performed above chance after only four sessions, N =

456 40, *p* = .04, and Dinding performed above chance after only six sessions, N = 60, *p* = .05. Zuri,

457 on the other hand, required 14 sessions to reach above chance levels of performance, N = 140, *p*

458 = .04.

459   In order to argue against an interpretation favoring rapid learning of associations between

460   particular stimuli, paired t-tests were conducted to show that performance on novel stimulus

461   pairings did not differ from that on previous pairings, for the first six sessions, for any of the

462   subjects. These data appear in Table 3. Performance did differ between novel and prior pairings

463   but only the difference for Dinar reached significance, and his performance was actually better

464   for novel (M = 74%, SD = 5%) than for prior pairings (M = 52%, SD = 11%), $t_4 = 3.31$, $p = .05$).

465   Again, there was no evidence that subjects were learning to choose correctly based on

466   remembering associations between the stimuli and patterns of reward.

467   In addition, one sample t-tests compared performance on both novel and prior pairings to

468   chance (50%) for each subject for the first six sessions. Performance was above chance on only

469   novel pairings for Molek, ($t_5 = 9.62$, $p < .001$) and Dinar, ($t_5 = 11.44$, $p < .001$). Performance on

470   prior pairings alone was significantly above chance for Abby ($t_4 = 2.73$, $p = .05$) and Dinding (t_4

471    $= 2.92$, $p = .04$). Zuri's performance was not above chance for either novel or prior pairings

472   within the first six sessions, highest $t_5 = 0.44$, $p = .68$.

473   Individual univariate ANOVAs of the subjects' scores with class (of the sample photo) as

474   the independent variable, for each subject, revealed an effect of class that was significant for

475   Dinar alone, $F_{4, 20} = 6.82$, $p < .001$. Dinar scored above 69% correct on all discriminations except

476   for birds, on which he performed close to chance. Dinding also had difficulty with bird trials as

477   well as with mammals, $F_{4, 20} = 2.40$, $p = .08$. The class of the animals depicted in the sample did

478   not significantly affect the performance of Molek, ($F_{4, 14} = 1.54$), Abby, ($F_{4, 14} = 2.06$), or Zuri,

479   ($F_{4, 14} = 1.01$) all p's > .05. Average percent correct on trials with each type of discrimination are

480   displayed in Table 4.

481   Only one of the subjects, Dinar, showed significant differences in accuracy based on

482   which class the sample photo belonged to. Dinar may have been distracted by preferences for

483   particular photos in both experiments. However his accuracy in this experiment was higher and

484   his performance was more consistent. The other subjects were not significantly distracted by

485   preferences for photos of animals belonging to particular taxonomic categories.

486   The orangutans were able to rapidly discriminate amongst species from different

487   taxonomic classes and did not have to learn to make associations between the exemplars.

488   However, the gorilla did not reach significant levels of performance until the third block of

489   sessions suggesting that she was not as predisposed to making these discriminations initially. In

490   addition, her accuracy tended to be better on Experiment 1, whereas the reverse was true for at

491   least two of the orangutan subjects, Molek and Dinar, (although these tendencies were

492   statistically significant only for Dinar when performance was compared across experiments,

493   (repeated measures ANOVA, $F_{1,24} = 8.18$, $p < .009$). The increase in performance for these two

494   orangutans might be attributed to prior experience with the DMTS procedure in Experiment 1;

495   However the same increase was not found for other similarly trained subjects (Zuri and Dinding).

496   In addition, Abby, who had been previously trained on DMTS tasks in a different study (Vonk,

497   2003), and who was tested on both experiments simultaneously, performed equally well across

498   experiments. This finding is not consistent with the claim that performance on the more concrete

499   level task might be superior if task ordering was not a factor.

Conclusions

501   It appears that orangutans learn intermediate level discriminations at least as readily as

502   they learn concrete level discriminations (see also Vonk & MacDonald, 2004). This is consistent

503   with the prediction made for humans (Keil, 1988; Rosch et al., 1976), and possibly with findings

504   from chimpanzees (Tanaka, 2001, although see Vonk et al., 2013), but not from squirrel monkeys,

505   pigeons (Roberts & Mazmanian, 1988) or gorillas (Vonk & MacDonald, 2002). The gorilla

506   subject in the present experiments seemed to learn the concrete level discrimination more rapidly

507   despite the fact that she was simultaneously learning the DMTS task for the first time. However,

508   it is difficult to make cross-species comparisons when fewer than five subjects of each species is

509 tested, particularly given the extreme individual subject differences in these experiments and

510 others. Research with young humans suggests that the intermediate or basic level concepts are

511 learned before subordinate or concrete level discriminations (Rosch et al., 1976). A general

512 assumption is that categories acquired earlier in ontogeny are also more accessible to more

513 phylogenetically removed species. Although the current study was not designed to address

514 developmental changes in concept acquisition, the fact that at least one gorilla showed greater

515 facility with concrete level categories is interesting and is worthy of further exploration.

516   Intermediate level categories are those that maximize within-category similarity and

517 distinctiveness relative to between-category similarity (Medin & Smith, 1984; Mervis & Rosch,

518 1981). It is unclear whether these categorizations are readily made because of an inherent

519 tendency to detect perceptual similarities within classes and dissimilarities between classes, or

520 whether such categorizations must be learned through experience with the exemplars or natural

521 instances (Medin & Smith, 1984). It is also unclear whether experiments such as these direct the

522 creation of a concept, or merely provide evidence for pre-existing concepts in the subjects

523 (Huber, 1999; Roitblat & Von Fersen, 1992). However, the fact that subjects achieved above

524 chance levels of responding with both of these discriminations more rapidly than they did with

525 different kinds of MTS tasks (Vonk, 2002, 2003), despite learning the DMTS procedure for the

526 first time here, indicates that the categories may have been spontaneously perceived, or at the

527 very least, relatively easy for these subjects to acquire. The current experiments provide the most

528 direct comparison between acquisition of various concrete and intermediate level category

529 discriminations made at a perceptual level in the absence of additional information, such as labels

530 or biological facts. They therefore provide evidence that categories at various levels of

531 abstraction can be acquired in the absence of linguistic labels or biological knowledge, in at least

532 two of our closest primate relatives.

533   Previous studies have shown that items may be correctly classified spontaneously and

534   without regard to experimental training but have not directly contrasted classification at various

535   levels of abstraction (Murai et al., 2004; 2005). Cerella (1979) investigated the ability of pigeons

536   to discriminate oak leaves from leaves of other species. Results from his series of experiments

537   lent support to the idea that these taxonomic classifications were made spontaneously and did not

538   involve induction. Significant transfer was made to multiple unique positive exemplars after

539   training with a single exemplar, and transfer did not depend upon contrasting the original instance

540   to negative instances. Typically, learning of experimentally defined categories requires

541   presentation of both positive and negative stimuli (Sutton & Roberts, 2002) and is improved by

542   the presentation of multiple exemplars (Katz, Wright & Bachevalier, 2002; Sutton & Roberts,

543   2002; Wright & Katz, 2007). That Cerella's pigeons did not require this experience might suggest

544   a pre-existing concept for oak leaves. The number of exemplars presented was also limited in the

545   current experiments.

546   Furthermore, Cerella's pigeons had difficulty discriminating between instances of oak

547   leaves that varied on specific dimensions, despite the fact that they could discriminate between

548   leaves of different species that sometimes shared features. These findings also suggest that

549   pigeons may be hardwired to perceive oak leaves as belonging to the same species, unique from

550   other species of leaves, as opposed to learning the discriminations by attending to distinct

551   physical features in the stimuli. In the current study, at least three orangutans learned rapidly to

552   judge animal members of different taxonomic classes as belonging to the same category, despite

553   lacking experience with many of the species depicted in the test stimuli. The rapid learning could

554   not be attributed to learning associations between specific exemplars because both the orangutans

555   and the gorilla were at least as accurate with newly paired as with previously paired exemplars.

556   Thus it is possible that concepts for intermediate level categories are readily extracted from

557   shared perceptual information between the stimuli. The gorilla subject appeared to learn to

558 classify the stimuli similarly after several sessions, suggesting that she was capable of perceiving

559 distinctions between the categories tested but did not do so immediately.

560   In an interesting test of spontaneous classification, Brown and Boysen (2000) presented

561 chimpanzees with pairs of photos depicting different species of animals and required them to

562 identify the pairs as being either the same or different. The chimpanzees showed above chance

563 categorization despite not being differentially reinforced for their responses. Interestingly,

564 whereas they were slightly more likely to classify tigers and housecats as the "same", they were

565 not more likely to classify gorillas and chimpanzees as the "same", relative to the other

566 comparisons tested. Therefore, whereas there was some evidence for more intermediate level

567 discriminations, in general, the spontaneous discriminations made by the chimpanzees were more

568 analogous to concrete level discriminations. Findings from this study are limited based on the fact

569 that chimpanzees were presented with specific pairs and might indicate them to be different but

570 could not also indicate that they were *more* similar to each other than to other potential pictures.

571 For example, they may have indicated that chimpanzees were different from gorillas but still

572 might have had the capacity to categorize chimpanzees and gorillas as more alike than

573 chimpanzees and lions. This capacity was not tested, however. In one of only two other studies to

574 investigate various levels of abstraction in chimpanzee concept formation (Tanaka, 2001) the

575 procedure does not allow comparison of the relative ease with which the two levels of concept

576 discrimination were achieved. In Vonk et al. (2013), only two chimpanzees were tested in

577 different orders as an attempt to control for order effects in learning the discriminations at various

578 levels. One chimpanzee did not show evidence of concept acquisition at any level, and the other

579 found discriminations more difficult as they became more abstract. The current study was able to

580 test for multiple possible category matches in the same session without setting one category as

581 correct or incorrect as in Vonk et al. (2013). In the current study, we were also able to assess

582   whether matches were more difficult for particular species or class comparisons and whether

583   errors revealed associative learning, perceptual confusions or untrained preferences.

584   In Brown and Boysen's study (2000), chimpanzees were less likely to judge two different

585   chimpanzees as being the same, relative to their judgments for cats, tigers, fish and gorillas. This

586   result is not surprising, given that animals may be more inclined to detect individual differences

587   between members of their own species (Martin-Malivel & Okada, 2007). Face processing in

588   particular may be specific to one's own species in other primates (Dufour, Pascalis & Petit,

589   2006). In the current experiments, orangutans were slightly less likely to correctly match

590   members of the same primate species, as compared to their performance for members of broader

591   taxonomic groups. Perhaps this finding is due to a greater tendency to perceive individual

592   differences amongst other primate members. It is possible that basic or intermediate level

593   discriminations are most critical to an animal's survival and that, at the more concrete level, even

594   finer perceptual discriminations are made.

595   The results of the current experiments cannot rule out a perceptual basis for making

596   taxonomic classifications, because a certain degree of perceptual overlap is necessarily evident

597   among members of a class. However sufficiently diverse stimuli were used here, to rule out the

598   use of single or few features. The subjects were required to make judgments based on relatively

599   few exemplars and reached high levels of performance after few sessions. This result may be

600   considered more consistent with the idea of an innate mechanism for distinguishing among

601   members of a class on a perceptual basis (Cerella, 1979), as opposed to inducing rules about the

602   associations between stimuli (Gelman, 1989; Mandler, 2000). However, subjects who did not

603   initially perform at high levels did learn to match at very high levels. This finding suggests that

604   even when categorizations are not made spontaneously, other non-human Great Ape species do

605   possess the capability for applying rules based on the presence or absence of multiple relevant

606   features.

607   Hampton (1998) demonstrated the influence of biological knowledge on humans'

608   classifications of similar natural stimuli; birds, fish, insects and animals. This knowledge

609   influenced classification more than it influenced typicality judgements. The present results are

610   not consistent with the idea that language (Benelli, 1988; Nelson, 1988) or scientific knowledge

611   (Inagaki, 1989) is necessary for making natural taxonomic classifications – at least not at the

612   perceptual level of categorization. Instead it would appear that our ability to categorize organisms

613   based on biological similarities is shared with other members of the animal kingdom, at least with

614   other non-human primates. These distinctions can clearly be made on the basis of presentation of

615   two-dimensional stimuli depicting species with which the subjects have had no experience.

616   In accord with the hypothesis that language is not needed to support the categorization of

617   natural stimuli, human infants have been shown to categorize in a manner that corresponds to

618   those categorizations made by adults on the basis of biological knowledge (Eimas & Quinn,

619   1994). Clearly infants do not yet hold conceptual representations of biological categories and yet,

620   the fact that these perceptually based discriminations correspond to mature adult conceptual

621   discriminations is important. As concepts are learned they may be influenced by cultural contexts,

622   particularly the labels provided by natural languages (Davidoff, 2001; Whorf, 1956); However,

623   Rhodes and Gelman (2009) have shown that cultural context is important for the categorization

624   of artifacts, but less so for natural objects. Following a meta-analysis of concept studies across

625   cultures, Malt (1995) concludes that there is significant structure provided by natural categories

626   to evoke similar categorization in different cultures that vary with regard to familiarity with the

627   objects (e.g. bird species). Across cultures, there was great convergence between folk concepts

628   and scientific concepts. However, she concludes that the issue is more complex in that culturally

629   specific beliefs may influence categorization at more abstract levels. Data from such cross

630   cultural studies and comparative studies (such as this one) converge to suggest that the ability to

631   designate linguistic labels for categories is not necessary for creating at least *perceptual*

632   representations of biological categories (see also Wasserman & DeVolder, 1993). Perhaps both

633   infants and non-human primates have acquired the ability to *perceive* categories, an ability that

634   may underlie the capacity to develop *abstract concepts* whose full emergence may yet depend

635   upon the development of language. This latter conjecture has yet to be proven. The current

636   experiments are a start in that direction.

637   In defining "concepts" it is important to distinguish between perceptual versus conceptual

638   processes (Eimas &Quinn, 1994; Huber, 1999; Mandler, 2000; Premack, 1983). Because the

639   subjects in the current study were able to correctly match members of various taxonomic groups

640   after a brief delay between sample and test stimuli, it is possible that they maintained a

641   representation of the category itself and not only of the particular sample exemplar. This

642   possibility is made more likely by the important finding that the subjects did not rely upon

643   remembering previous pairings of stimuli. In addition the subjects here were required to make an

644   instrumental response to the stimuli. These results might thus constitute evidence for a conceptual

645   versus a purely perceptual representation. Use of a concept is implicated when an individual is

646   able to form a coherent category from exemplars displaying some shared and yet many

647   distinctive features (Spalding & Ross, 2000). Analysis of single features in isolation is not

648   sufficient for the formation of abstract concepts. Instead, the individual must combine and

649   compare various features and make some determination as to which features are deemed critical

650   for category membership. In the current experiment, orangutans and one gorilla were able to

651   analyze various perceptual aspects of two-dimensional photographs and to use this information to

652   represent distinct categories. Further work is needed to clearly discern the extent to which such

653   categories are perceptual versus conceptual in nature.

654   **Acknowledgements**

655   The cooperation and support of the staff at the Toronto Zoo was greatly appreciated.

656   Special thanks to Dianne Devison, Bev Carter, Vanessa Phelan, Gerri Mintha, Charles Guthrie,

657   Tim McCaskie, Connie Wiederman, Heidi Minicke, Michelle Smith, Karyn Tunwell, Bridget

658   Burke-Johnson, Andrea Beatson, Rick Vos, Mark Bongelli, David Partington, Des Macguire,

659   John Armstrong, and Jackie Craig, without whose assistance this study would not have been

660   possible. I would like to thank Suzanne MacDonald and Robert Sorge for comments on an

661   earlier version of this paper. I also thank Jon Toth for providing protective equipment.

662   **References**

663   Anggoro, F. K., Medin, D. L., & Waxman, S. R. (2010). Language and experience influence

664   children's biological induction. *Journal of Cognition and Culture, 10*, 171-187.

665   doi:http://dx.doi.org/10.1163/156853710X497220

666   Autier-Dérian, D., Deputte, B. L., Chalvet-Monfray, K., Coulon, M., & Mounier, L. (2013).

667   Visual discrimination of species in dogs (*canis familiaris*). *Animal Cognition*,

668   doi:http://dx.doi.org/10.1007/s10071-013-0600-8

669   Benelli, B. (1988). On the linguistic origin of superordinate categories. *Human Development, 31,*

670   20-27.

671   Brooks, D. I., Ng, K. H., Buss, E. W., Marshall, A. T., Freeman, J. H., & Wasserman, E. A.

672   (2013). Categorization of photographic images by rats using shape-based image

673   dimensions. *Journal of Experimental Psychology: Animal Behavior Processes, 39*(1), 85-

674   92. doi:http://dx.doi.org/10.1037/a0030404

675   Brown, D.A. & Boysen, S.T. (2000). Spontaneous discrimination of natural stimuli by

676   Chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology, 114,* 392-400.

677   Cerella, J. (1979). Visual classes and natural categories in the pigeon. *Journal of Experimental*

678   *Psychology: Human Perception and Performance, 5,* 68-77.

679   Coley, J. D. (2007). The human animal: developmental changes in judgments of taxonomic and

680   psychological similarity among humans and other animals. *Cognitie, Creier,*

681   *Comportament / Cognition, Brain, Behavior, 11*, 733-756.

682   Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y., . . . Haxby, J.

683   V. (2012). The representation of biological classes in the human brain. *The Journal of*

684   *Neuroscience, 32*, 2608-2618. doi:http://dx.doi.org/10.1523/JNEUROSCI.5547-11.2012

685   Davidoff, J. (2001). Language and perceptual categorisation. *Trends in Cognitive Sciences, 5*,

686   382-387. doi:http://dx.doi.org/10.1016/S1364-6613(00)01726-5

687   Dufour, V., Pascalis, O., & Petit, O. (2006). Face processing limitation to own species in

688   primates: A comparative study in brown capuchins, tonkean macaques and humans.

689   *Behavioural Processes, 73*, 107-113. doi:http://dx.doi.org/10.1016/j.beproc.2006.04.006

690   Eimas, P.E. & Quinn, P.C. (1994). Studies on the formation of perceptually based basic-level

691   categories in young infants. *Child Development, 65,* 903-917.

692   Fujita, K. (1987). Species recognition by five macaque monkeys. *Primates, 28,* 353 – 366.

693   Fujita, K. & Matsuzawa, T. (1986). A new procedure to study the perceptual world of animals

694   with sensory reinforcement: Recognition of humans by a Chimpanzee. *Primates, 27,* 283-

695   291.

696   Fujita, K., Watanabe, K., Widarto, T.H. & Suryobroto, B. (1997) Discrimination of macaques by

697   macaques: the Case of Sulawesi species. *Primates, 38*, 233-245.

698   Gelman, S.A. (1989). Children's use of categories to guide biological inferences. *Human*

699   *Development, 32,* 65-71.

700   Gelman, S. A., & Davidson, N. S. (2013). Conceptual influences on category-based induction.

701   *Cognitive Psychology, 66*, 327-353.

702   doi:http://dx.doi.org/10.1016/j.cogpsych.2013.02.001

703   Hampton, J.A. (1998). Similarity-based categorization and fuzziness of natural categories.

704   *Cognition, 65,* 137-165.

705   Herrnstein, R.J. (1979). Acquisition, generalization, and discrimination reversal of a natural

706   concept. *Journal of Experimental Psychology: Animal Behaviour Processes, 5,* 116-129.

707   Herrnstein, R.J., Loveland, D.H. & Cable, C. (1976). Natural concepts in pigeons. *Journal of*

708   *Experimental Psychology: Animal Behaviour Processes, 2,* 285-302.

709   Huber, L. (1999). Generic perception: Open-ended categorization of natural classes. *Current*

710   *Psychology of Cognition, 18,* 845-887.

711   Inagaki, K. (1989). Developmental shift in biological inference processes: From similarity-based

712   to category-based attribution. *Human Development, 32,* 79-87.

713 Johnson, K. E. (2001). Impact of varying levels of expertise on decisions of category typicality.

714   *Memory & Cognition, 29*, 1036-1050. doi:http://dx.doi.org/10.3758/BF03195765

715 Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level

716   of categorization. *Journal of Experimental Psychology: General, 126*, 248-277.

717   doi:http://dx.doi.org/10.1037/0096-3445.126.3.248

718 Katz, J.S., Wright, A.A., Bachevalier, J. (2002). Mechanisms of *Same/Different* Abstract-

719   Concept learning by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental*

720   *Psychology: Animal Behaviour Processes, 28,* 358-368.

        Keil, F.C. (1988). Commentary: Conceptual heterogeneity versus developmental homogeneity

721   (on chairs and bears and other such pairs). *Human Development, 31,* 35-43.

722 Kiani, R., Esteky, H., Mirpour, K., & Tanaka, K. (2007). Object category structure in response

723   patterns of neuronal population in monkey inferior temporal cortex. *Journal of*

724   *Neurophysiology, 97*, 4296-4309. doi:http://dx.doi.org/10.1152/jn.00024.2007

725 Martin-Malivel, J., & Okada, K. (2007). Human and chimpanzee face recognition in

726   chimpanzees (*pan troglodytes*): Role of exposure and impact on categorical perception.

727   *Behavioral Neuroscience, 121*, 1145-1155. doi:http://dx.doi.org/10.1037/0735-

728   7044.121.6.1145

729 Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology, 29*,

730   85-148.

        Mandler, J.M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition &*

731   *Development, 1*, 3-36.

732 Marsh, H. L., & MacDonald, S. E. (2008). The use of perceptual features in categorization by

733   orangutans (*pongo abelli*). *Animal Cognition, 11*, 569-585.

734   doi:http://dx.doi.org/10.1007/s10071-008-0148-1

735 Medin, D.L. & Smith, E.E. (1984). Concepts and concept formation. *Annual Review of*

736   *Psychology, 35,* 113-138.

737 Mervis, C.B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of*

738   *Psychology, 32,* 89-115.

739 Murai, C., Tomonaga, M., Kamegai, K., Terazawa, N., & Yamaguchi, M. K. (2004). Do infant

740   Japanese macaques (*macaca fuscata*) categorize objects without specific training?

741   *Primates, 45*(1), 1-6. doi:http://dx.doi.org/10.1007/s10329-003-0053-9

742 Murai, C., Kosugi, D., Tomonaga, M., Tanaka, M., Matsuzawa, T., & Itakura, S. (2005). Can

743   chimpanzee infants (*pan troglodytes*) form categorical representations in the same

744   manner as human infants (*homo sapiens*)? *Developmental Science, 8*(3), 240-254.

745   doi:http://dx.doi.org/10.1111/j.1467-7687.2005.00413.x

746 Nelson, K. (1988). Where do taxonomic categories come from? *Human Development, 31,* 3-10.

747 Premack, D. (1983). The Codes of man and beasts. *The Behavioral and Brain Sciences, 6,* 125-

748   167.

749 Quinn, P. C., & Tanaka, J. W. (2007). Early development of perceptual expertise: Within-basic-

750   level categorization experience facilitates the formation of subordinate-level category

751   representations in 6- to 7-month-old infants. *Memory & Cognition , 35*, 1422-31.

752 Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure

753   of animal, artifact, and human social categories across two cultural contexts. *Cognitive*

754   *Psychology, 59*(3), 244-274. doi:http://dx.doi.org/10.1016/j.cogpsych.2009.05.001

755 Roberts, W.A. & Mazmanian, D.S. (1988). Concept learning at different levels of abstraction by

756   pigeons, monkeys, and people. *Journal of Experimental Psychology: Animal Behaviour*

757   *Processes, 14,* 247-260.

758 Roitblat, H.L. & Von Fersen, L. (1992). Comparative cognition: Representations and processes

759   in learning and memory. *Annual Review of Psychology, 43,* 671-710.

760   Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects

761   in natural categories. *Cognitive Psychology, 8,* 382-439.

762   Ross, N., Medin, D. L., Coley, J. D., & Atran, S. (2003). Cultural and experiential

763   differences in the development of biological induction. *Cognitive Development, 18,*

764   25-47.

765   Spalding, T.L. & Ross, B.H. (2000). Concept learning and feature interpretation. *Memory and*

766   *Cognition, 28,* 439-451.

767   Sutton, J.E., & Roberts, W.A. (2002) Failure to find evidence of stimulus generalization within

768   pictorial categories in pigeons. *Journal of the Experimental Analysis of Behavior, 78,*

769   333-343.

      Tanaka, M. (2001). Discrimination and categorization of photographs of natural objects by

770   chimpanzees (*Pan troglodytes*). *Animal Cognition, 4*, 201-211.

771   Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in

772   the eye of the beholder? *Cognitive Psychology, 23*, 457-482.

773   Vonk, J. (2003). Gorilla (*Gorilla gorilla gorilla*) and Orangutan (*Pongo abelii*) understanding

774   of first and second order relations. *Animal Cognition, 6,* 77-86.

775   Vonk, J. (2002). Can Orangutans (*Pongo abelii*) and Gorillas (*Gorilla gorilla gorilla*) acquire

776   relationships for social concepts? *International Journal of Comparative Psychology, 15,*

777   257-277.

778   Vonk, J., Jett, S. E., Mosteller, K.W. & Galvan, M. (2013). Natural category discrimination in

779   chimpanzees (*Pan troglodytes*) at three levels of abstraction. *Learning and Behavior*.

780   10.3758/s13420-013-0103-0

781   Vonk, J., Jett, S.E., & Mosteller, K.W. (2012). Concept formation in American black bears

782   (*Ursus americanus*). *Animal Behaviour, 84,* 953-964.

783   Vonk, J. & MacDonald, S.E. (2002). Natural concept formation in a juvenile gorilla (*Gorilla*

784   *gorilla gorilla*) at three levels of abstraction. *Journal of the Experimental Analysis of*

785   *Behavior, 78,* 315-332.

786   Vonk, J. & MacDonald, S.E. (2004). Levels of abstraction in Orangutan (*Pongo abelii*)

787   categorization. *Journal of Comparative Psychology, 118,* 3-13.

788   Wasserman, E.A. & Devolder, C.L. (1993). Similarity and nonsimilarity-based

789   conceptualization in children and pigeons. *The Psychological Record, 43,* 779-793.

790   Whorf, B.L. (1956) The relation of habitual thought and behavior to language. In *Language,*

791   *Thought and Reality: Essays by B.L. Whorf* (Carroll, J.B., ed.), pp. 35–270, MIT Press.

792   Wright, A. A., & Katz, J. S. (2007). Generalization hypothesis of abstract-concept learning:

793   Learning strategies and related issues in *macaca mulatta, cebus apella,* and *columba*

794   *livia. Journal of Comparative Psychology, 121*, 387-397.

795   doi:http://dx.doi.org/10.1037/0735-7036.121.4.387

796   Yoshikubo, S. (1985). Species discrimination and concept formation by Rhesus monkeys

797   (*Macaca mulatta*). *Primates, 26*, 285-299.

Performance in Exp. 1

Percentage of Correct Responses in Experiment 1 on trials where exemplars comprised

novel or prior pairings (Standard deviations in parentheses) rimina ÿ Ṣ H

Table 1.

Subject	Novel Pairings	Prior Pairings
Abby	77.5 (17.9)	61.7 (21.9)
	N =26	N = 34
Dinar	58.2 (39.0)	54.5 (13.8)
	N = 26	N = 34
Dinding	66.2 (20.6)	65.2 (16.9)
	N = 28	N = 32
Molek	84.3 (15.0)	78.3 (9.0)
	N = 23	N = 37
Zuri	69.4 (30.3)	64.9 (17.6)
	N = 28	N = 32

Table 2_(on next page)

Performance by category

Average percent correct for each subject across each type of discrimination (according to the image depicted in the sample) for Experiment 1. Standard deviations in parentheses.

Table 2.

	Zuri	Molek	Dinar	Dinding	Abby
Tamarin	79 (03)	73 (13)	84 (10)	78 (12)	75 (13)
Gibbon	78 (10)	84 (09)	52 (12)	69 (15)	63 (10)
Jap. Mac.	81 (08)	76 (21)	58 (13)	42 (13)	75 (10)
Lemur	71 (14)	74 (09)	46 (10)	70 (13)	70 (20)
Proboscis	36 (13)	60 (07)	61 (14)	66 (18)	73 (13)

Performance in Exp. 2

Percentage of Correct Responses in Experiment 2 on trials where exemplars comprised novel or prior pairings (Standard deviations in parentheses)

Table 3.

Subject	Novel Pairings	Prior Pairings
Abby	49.2 (20.9) N = 39	81.9 (26.1) N = 21
Dinar	74.5 (5.2) N = 43	51.8 (11.2) N = 17
Dinding	58.8 (16.3) N = 41	78.4 (21.7) N = 19
Molek	77.9 (7.4) N = 41	71.6 (29.9) N = 19
Zuri	51.3 (7.4) N = 44	51.0 (47.5) N = 16

Performance by category.

Average percent correct for each subject across each type of discrimination (according to the image depicted in the sample) for Experiment 2. Standard deviations in parentheses.

Table 4.

	Zuri	Molek	Dinar	Dinding	Abby
bird	65 (13)	75 (17)	53 (06)	57 (06)	80 (08)
fish	63 (05)	85 (06)	73 (08)	69 (16)	68 (10)
insect	60 (14)	65 (06)	71 (05)	68 (04)	70 (14)
mammal	80 (27)	85 (10)	72 (05)	58 (08)	73 (10)
reptile	60 (16)	73 (22)	70 (11)	72 (11)	60 (08)

Figure 1

Example images used in Exp. 1

Example images from two categories in Exp. 2: gibbons (A-D) and lemurs (E-H).

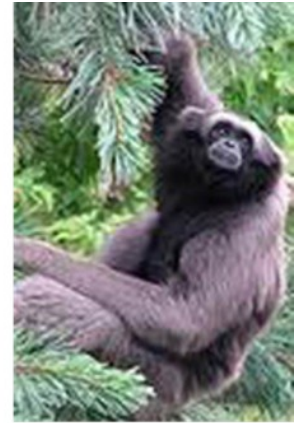
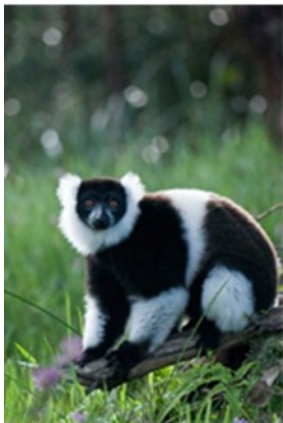
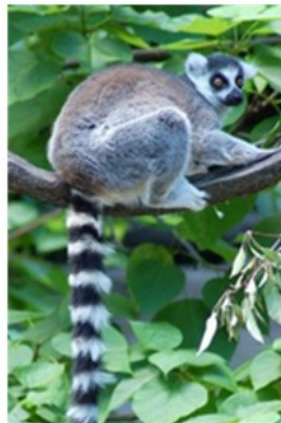
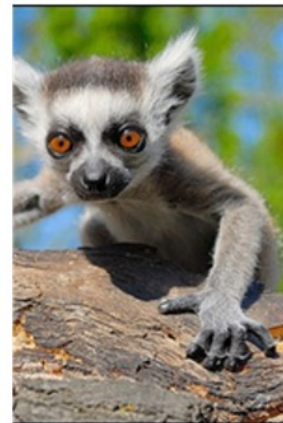
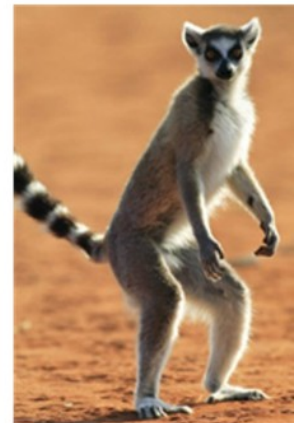



Figure 2

Results from Exp. 1

Average percent correct across blocks of 5 sessions (50 trials) for each subject in Experiment 1.

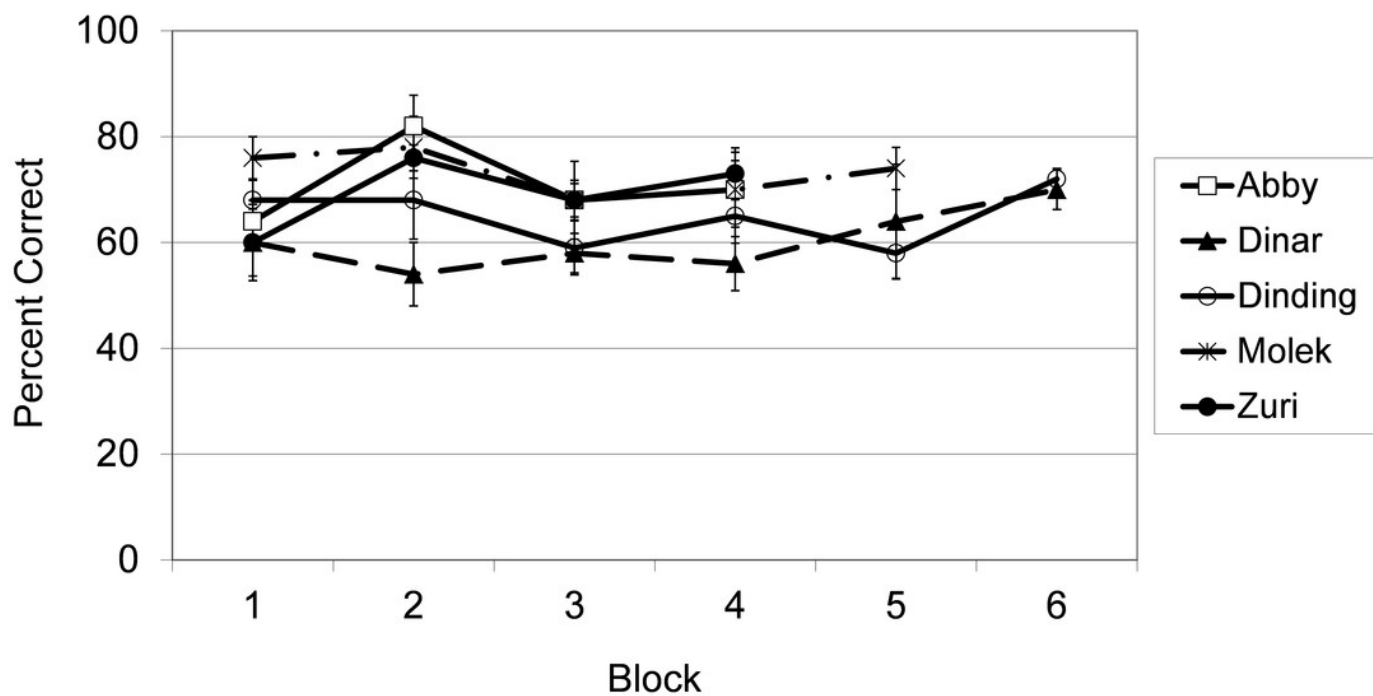
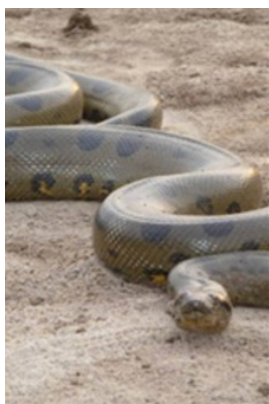



Figure 3

Example Images from Exp. 2

Examples of images from two categories used in Exp. 2: Reptiles (A-F) and Birds (G-L)

A B C

D E F

G H I

J K L

Figure 4

Results from Exp. 2

Average percent correct across blocks of 5 sessions (50 trials) for each subject in Experiment 2.