Running Head: ANALOGY BY ABSTRACTION

Flemming, T., Thompson, R.K., & Fagot, J. (2013). Baboons, like humans, solve an analogical matching task by categorical abstraction of relations. Animal Cognition, 16, 519-524

Baboons, like humans, solve analogy by categorical abstraction of relations

Timothy M. Flemming
CNRS, Aix – Marseille University, Pôle 3C, Laboratory of Cognitive Psychology

Roger K. R. Thompson
Franklin & Marshall College, Department of Psychology

Joël Fagot
CNRS, Aix – Marseille University, Pôle 3C, Laboratory of Cognitive Psychology

Address correspondence to:

Timothy Flemming
timothy.flemming@gmail.com

404.323.3813

CNRS, Aix – Marseille University, Pôle 3C
Laboratory of Cognitive Psychology
3 Place Victor Hugo, Bât. 9, Case D
13331 Marseille cedex 1, FRANCE

Abstract

Reasoning by analogy is one of the most complex and highly adaptive cognitive processes in abstract thinking. For humans, analogical reasoning entails the judgment and conceptual mapping of relations-between-relations, and is facilitated by language (Gentner 1983; Premack 1986). Recent evidence, however, shows that monkeys like 'language-trained' apes exhibit similar capacity to match relations-between-relations (Fagot & Thompson 2011; Flemming, Thompson, Beran & Washburn 2011). Whether this behavior is driven by the abstraction of categorical relations, or alternatively by direct perception of variability (entropy) is crucial to the debate as to whether nonhuman animals are capable of analogical reasoning. In the current study, we presented baboons (*Papio papio*) and humans (*Homo sapiens*) with a computerized *same-different* relational matching task that in principle could be solved by either strategy. Both baboons and humans produced statistically identical patterns of responding. Both species responded *different* when the perceptual variability of a stimulus array fell exactly between or even closer to that of a *same* display. Overall, these results demonstrate that categorical abstraction trumps perceptual properties and, like humans, old-world monkeys can solve the analogical matching task by judging the categorical abstract equivalence of *same/different* relations between relations.

Keywords: Analogy, Analogical reasoning, RMTS, Monkey, Same/Different, Abstract Thinking, Representation

Baboons, like humans, solve analogy by categorical abstraction of relations

As a measure of abstract thinking, reasoning by analogy proves to be one of the most complex cognitive processes. To solve an analogical problem, one must determine the relational concept instantiated between or among one set (stimuli, words, problem-space, etc.) and subsequently search for that same structural or functional relation among sets of novel and unfamiliar exemplars. As a foundation for critical thinking (Gentner 1999), success in analogical reasoning entails judging and mapping conceptually abstract relations-between-relations rather than mere surface perceptual features (Hofstadter 2001; Gentner 1983). In humans this ability occurs at an early age and, importantly, is facilitated by the emerging capacity to represent abstract relations in concrete symbolic terms via linguistic labeling, as is the case also for language-trained or symbol-using chimpanzees (e.g., Premack 1986; Rattmermann & Gentner 1998; Thompson, Oden & Boysen 1997; see also Clark & Thornton 1997; Flemming, Beran, Thompson, Kleider & Washburn 2008). Crucial to analogical reasoning, whether by human or nonhuman, thus are [1] examination of underlying structural similarities rather than perceptual features and [2] re-encoding or reinterpretation of the abstract relations employed.

Recent theoretical perspectives on analogical reasoning assert that it is not only a hallmark of higher-order reasoning, but also a uniquely human cognitive trait (Penn, Holyoak & Povinelli 2008). In this view, the symbolic reinterpretation of analogical relations by other species, including all nonhuman primates, is impossible due to their lack of a generalized abstract representational capacity (c.f., Thompson & Oden 2000) and therefore their abstract conceptual abilities are grounded solely on directly perceptible surface features. Hence, relational-matching behavior by nonhuman animals would reflect their expert sensitivity to perceptual characteristics, rather than analogical relations *per se*.

Supporting this claim is evidence from relational-matching tasks that the discrimination of *same* from *different* in multiple-icon arrays by both pigeons and monkeys is profoundly affected by the magnitude of perceptual variability (i.e., entropy) in a visual array (Young & Wasserman 1997; for review see Wasserman & Young 2010; Zentall, Wasserman, Lazareva, Thompson & Rattermann 2008). If relational matching performance is based on entropy/perceptual variability alone, then judgments of the concepts *same* and *different* would thus be derived from a perceptually grounded *continuous* scale (from *mostly sameness* to *mostly difference*) rather than *categorically* (*same* or *different*) as presumably they are by humans (Smith, Redford, Hass, Coutinho & Couchman 2008). Hence, if animals match relations solely in terms of similar or dissimilar levels of perceptual variability, they cannot be said to be representing abstract concepts categorically.

Recent evidence for reliable success by old-world monkeys in generalized 2x2 item relational matching-to-sample (RMTS) tasks (e.g., match AA with BB, not CD and match EF with CD, but not BB) by monkeys (e.g., Fagot & Thompson 2011; Flemming et al. 2011), suggests that non-symbol/language-trained primates are indeed capable of rudimentary analogical reasoning. However, as yet there is no unequivocal evidence of commonalities or differences between the cognitive processes underlying the RMTS performances of these old-world monkeys and humans.

Do non-linguistic monkeys accomplish the RMTS task by categorically representing, or ‘reinterpreting’ relations as is the case for humans or alternatively, is their performance mediated via rudimentary perceptual judgments? The present study aimed to settle the above debate. In our study, we presented 8 baboons (*Papio papio*) and 8 humans (*Homo sapiens*) with a computerized relational matching task involving different combinations of 4-item geometrical shapes (e.g., of the type AAAA, ABBA, ABCD) as stimuli, with no item sharing between sample and comparison stimuli (see Figure 1). In principle, this task could be solved either by considering the abstract (same/different) relation instantiated by these stimuli, or by matching the stimuli on the basis of their perceptual entropy. Our results provide unique evidence for the categorical representation of abstract relations on a RMTS task by both human and nonhuman primates, implicating continuity of underlying cognitive processes for analogical reasoning.

Method

Participants and Apparatus

Eight Guinea baboons (*Papio papio*) age 3 to 7 years participated on computerized tasks. Eight human (*Homo sapiens*) participants (6 female) age 23 to 34 years were also recruited from the CNRS field site in Rousset, France. All participants were naïve to testing hypotheses and were not familiar with any related topics being investigated.

Baboon participants belong to a large social group of 29 animals living in 670 m^2 enclosure at the CNRS primatology station (Rousset-sur-ARC, France). All baboons had a radio frequency identification (RFID) microchip implanted in each form-arm that served for auto-identification within the test boxes. The baboons were tested with a new test apparatus named “Automatic Learning Device for Monkeys" [ALDM, see their detailed description in Fagot and Paleressompoulle (2009) and Fagot & Bonté (2010)]. Each of the 10 identical ALDM test systems used in the research consisted of an experimental chamber (0.7 x 0.7 x 0.8 m) freely accessible from the living enclosure. Upon entering the test chamber, a baboon faces an opaque panel within which a (7 x 7 cm) viewing port and two (8 x 5 cm) hand ports are positioned to see and touch a 19-inch LCD touch-screen monitor (model ET1739L from Elotouch; Berwyn, IL) at eye level 25 cm behind the view port. Each ADLM system had a network-controlled PC computer (Dell precision 67, 3.2-GHz**)** and comprised a ‘homemade’ dispenser delivering dry-wheat reinforcers (www.ebly.co.uk) inside the test chamber. The test program, written in E-prime (V2 professional, Psychology Software Tools, Pittsburgh), automatically identified an animal via the microchip when it placed its forearm within an arm port, and determined its last stopping point in trial presentations, in order to assign the independent variables to be experienced during the trial. The stimulus choices were recorded in each trial. With this system, the participants could experience identical test programs at their own pace, independently of the test system they chose. Participants from all groups were neither food nor water deprived.

All baboons in the present experiment had recently participated in a 2x2 relational-matching task (Fagot & Thompson 2011) that utilized the same geometric shapes as stimuli, and reliably performed this task above chance.

Procedure

During initial training, all participants completed relational matching trials with uniform same/different arrays each composed of four geometrical shapes. When converted to letters for expository purposes, this task can be formalized as “match AAAA with BBBB, not CDEF.” The baboons completed trials until they reached 80% or greater accuracy in two consecutive 100-trial blocks. Correct choices resulted in the delivery of a wheat-grain reward; incorrect choices resulted in no reward and the presentation of a green screen for 1 sec before the initiation of the next trial. Baboons completed this basic training after *M* = 400 trials, *SD* = 129.1.

Human participants followed a parallel procedure with 2 slight modifications. Humans completed trials until they reached 80% or greater accuracy in two consecutive 12-trial blocks. Feedback was given following each trial with a display of “OUI” (yes) or “NON” (no). The 80% criterion was reached in training after *M* = 34.5 trials, *SD* = 16.3.

In test sessions, mixed-array probe trials of three kinds were inserted randomly within a trial block and rewarded nondifferentially: AAAB (entropy = 0.81), AABB (entropy = 1.0) and AABC (entropy = 1.5), with the location of each item randomized across trials. Choice arrays on both baseline and probe trials were comprised of *all same* (DDDD: entropy = 0) or *all different* (EFGH: entropy = 2.0) stimuli. See Figure 1 for a depiction of trial types. Our rationale was that an entropy-based strategy should be demonstrated by a greater frequency of *same* responses than *different responses* to the 0.81 entropy arrays, because the degree of variability within those arrays is much closer to *all same* (entropy = 0) than to *all different* (entropy = 2) arrays. Similarly, no response bias should be observed in response to the arrays with an entropy of 1.0, if the subject used an entropy-based strategy, because that value of entropy falls exactly in between that of the *all same* (entropy = 0) and *all different* (entropy = 2) arrays. Probe trials were neither determinately “correct” nor “incorrect” as their purpose was to discover how the participants spontaneously matched mixed arrays that could be classified as either *same* or *different*, dependent upon cognitive mechanism utilized (perceptual/conceptual), and how relations were defined (continuous/categorical).

Each trial began with the presentation of a sample 4-item array (uniform baseline or mixed probe). This sample remained centered in the upper third of the computer screen until touched (see Figure 1, top left). After the participant touched the array, it was removed from the screen and two 4-item choice arrays (all baseline-type, e.g. AAAA vs. BCDE) were presented in the lower third of the screen at the left and right extremes (see Figure 1, top right). Arrays were composed of simple white geometric shapes from a set of 10 items. By using a small stimulus set, the opportunity for confounding matching strategies was high, providing more support for a relational match if observed. Baseline trials were differentially rewarded in a manner identical to training and probe trials were non-differentially rewarded as described above.

Eighty percent of randomly-selected probe trials were rewarded regardless of response (correct or incorrect) to reflect levels of reward during baseline trials. Baboons completed 10 blocks of 124 trials (24 probe trials each block) resulting in approximately 240 total probe trials for each subject. Human participants completed 2 blocks of 100 trials (60 probe trials each block) resulting in 120 total probe trials of completely trial-unique configuration for each subject, making rote memorization unlikely.

Results

As illustrated in Figure 2, the selection of choice arrays provides evidence of categorical representation of relations by baboons, wherein only baseline identity samples (AAAA) were reliably matched to *same* ($z = 46.89$, $\alpha = 0.001$). Regardless of entropy level (i.e., perceptual variability), all other trial types were matched to *different*. As a group, the baboons responded “different” significantly more often than chance on all probe-trial types [entropy = 0.81 (AAAB), 1.0 (AABB), 1.5 (AABC) and 2.0 (*baseline* ABCD)], z-values = 4.85, 8.99, 15.62 and 16.40 respectively, all $\alpha = 0.01$.

Human participants matched according to an identical pattern of responding, and the correlation between the humans’ and baboons’ data was very close to 1, $r^2 = .997$, $p < 0.0001$. Humans reliably matched baseline identity samples to *same* ($z = 14.39$, $\alpha = 0.001$). As was the case for baboons, all other display types (including 2.0 baseline and mixed probe arrays from entropy 0.81-1.5) were reliably matched to *different*. Human participants responded “different” significantly more often than chance on all probe-trial types [entropy = 0.81 (AAAB), 1.0 (AABB), 1.5 (AABC) and 2.0 (*baseline* ABCD)], z-values = 3.43, 7.96, 11.58 and 13.91 respectively, all $\alpha = 0.01$, see Figure 2.

Whereas an effect of entropy within *different* classifications was observed across all trial types for baboons, the effect was driven largely by significant differences between the *all same* (entropy = 0) and all other types. Using logit transformed proportion of *different* responses, an ANOVA revealed a significant effect of entropy value, $F(4,35) = 92.54$, $p< 0.01$. Tukey HSD post-hoc tests (α level 0.01) revealed significant differences in responding between all entropy levels, except between 0.81 and 1.0, and between entropy level 1.5 and 2.0. The exact same pattern of results was obtained for human participants, as a main effect of entropy level was also observed, $F(4,35) = 37.22$, $p< 0.01$. Tukey HSD post-hoc tests (α level 0.01) revealed significant differences between all entropy levels, except between entropy levels within the different category (i.e., 0.81 to 2.0). Further, no differences were observed between the entropy levels within the different category. Critically, no reliable differences between percentage of *different* selection for baboons and humans at any entropy level were observed: entropy = 0.0 (*baseline* AAAA), 0.81 (AAAB), 1.0 (AABB), 1.5 (AABC) and 2.0 (*baseline* ABCD), z-values = 1.74, 0.14, 1.68, 0.69 and 1.94 respectively, $\alpha = 0.05$.

At entropy level 0.81 (e.g. AAAB, ABAA), wherein matches based on perceptual variability would predict classification as *same*, we observe categorical *different* selection for both humans and baboons. However, whereas at the group level all probe trials of this type were matched to *different*, responding to configuration types with repeated identity (AAAB and BAAA) was at levels of chance for both humans and baboons, z-values 1.47 and 1.22, respectively, α = 0.05, see Figure 3a. This pattern of responding suggests for the two species a "fuzzy" boundary for *different* when presented with higher levels of repeated identity.

Because AABB trial types fall at the midpoint between *same* and *different*, responding at this trial type is predicted to be random if driven by mechanisms of perceptual variability comparison as is observed in pigeons (see Young & Wasserman 1997, Figure 4). Contrary to this hypothesis, and in accordance with the above-reported findings at the 0.81 entropy level, we find responding at the 1.0 entropy level to be categorically defined as *different*. Regardless of configuration (AABB, ABAB and ABBA) all responding to different was significantly above chance for *Papio*, *z*-values = 3.66, 5.69, 6.23 respectively, and *Homo*, *z*-values = 3.96, 5.10 and 4.46 respectively , α = 0.01. See Figure 3b.

Whereas we observed a small increase in responding *same* to mixed arrays of entropy 0.81 (e.g., AAAB), matches to *different* (e.g., CDEF) were still significantly above chance at this and all samples of greater variability. Even when a display contains mostly identical (e.g., AAAB) or an equal number of identical and nonidentical (e.g., AABB) items, both humans and baboons reliably classify relations as *different*, implying a categorical abstraction and representation of relations. Analyses of the different spatial arrangements of the shapes within each level of entropy further stressed identical patterns of responding in humans and baboons.

Discussion

That both human and nonhuman primates follow a remarkably identical categorical pattern of responding to *same* and *different* when matching relations provides unique support for relational abstraction by nonhuman primates. Importantly, these results provide good evidence of continuity in the evolution of the cognitive requisites for analogical reasoning in primates. This would not have been the case had the monkeys' response decisions been determined predominantly by the perceptual strategy of matching the degree of stimulus variability instantiated by the different levels of entropy.

Admittedly, both baboons and humans perceived and were secondarily sensitive to changes in entropy. Interestingly, predominant effects of entropy as a basis for relational matching have been observed in up to 20% of human participants in a similar task (Young & Wasserman 2001), and have repeatedly been found in both monkeys (Fagot, Wasserman & Young 2001; Flemming 2011; Flemming, Beran & Washburn 2007) and pigeons (for review, see Wasserman & Young 2010). Importantly however, both monkey and human in the present task were primarily predisposed to reinterpret relations in an abstract categorical manner.

The present observed effect of entropy within the *different* category for both humans and baboons, suggests that *same* and *different* are “quantifiable” in that some mixed displays are relationally *more same* or *more different* than others, as Smith et al. (2008) suggest. Nevertheless, even the quantifiable nature of these relations does not preclude the predominance of abstract categorical representation, as observed in the present experiment for both human and baboon.

Because the explicit employment of concepts (rather than percepts) is essential to analogy-making by even a strict definition (i.e., Penn et al. 2008), we provide here the first conclusive demonstration that analogical relational matching is carried out by such means in nonhuman animals other than apes (c.f., Gillian, Premack & Woodruff 1981; Oden, Thompson & Premack 2001). More broadly, the present results implicate an abstract rule-based approach to relational matching by monkeys suggestive of their capacity to re-encode and abstractly represent relations categorically. This cognitive capacity further brings into question the extent and limits of abstract thinking in the absence of language or symbolic codes and the nature of the cognitive scaffolding effects they facilitate.

This adaptive capacity for abstract reasoning, which likely emerged prior to language, may well have its evolutionary origins in the social domain wherein animals have to deal with complex networks of relations between relationships (Dasser 1988). Given this evidence of continuity between human and nonhuman animal in the evolution of the requisite cognitive foundation for analogical reasoning, we can better entertain and investigate other contexts under which nonhuman animals may or may not similarly represent their world abstractly

Author Note

Timothy M. Flemming supported by Fyssen Foundation. Joël Fagot supported by project ANR-2010-BLANC-1908-01, and European Project EupriùNet-II. Roger Thompson supported by the Franklin & Marshall College Faculty Research & Development Fund. Address email correspondence to first author at: timothy.flemming@gmail.com

References

Clark A, Thornton C (1997) Trading spaces: Computation, representation, and the limits of uniformed learning. Behav Brain Sci 20: 57-90.

Dasser V (1988) Mapping social concepts in monkeys. In Byrne RW, Whiten A, editors. (Eds.), Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans. New York: Oxford University Press. pp 85-93.

Fagot J, Paleressompoulle D (2009) Automated testing of cognitive performance in baboons maintained in social groups. Behav Res Meth 41: 396-404.

Fagot J, Bonté E (2010) Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free-ranging baboons (*Papio papio*). Behav Res Meth 42: 507-516.

Fagot J, Thompson RKR (2011) Generalized relational matching by Guinea baboons (*Papio papio*) in two by two-item analogy problems. Psychol Sci 22: 1304-1309.

Fagot J, Wasserman EA, Young ME (2001) Discriminating the relation between relations: The role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*). J Exp Psychol: Anim Behav Process 27: 316-328.

Flemming TM (2011) Conceptual thresholds for *same* and *different* in old- (*Macaca mulatta*) and new-world (*Cebus apella*) monkeys. Behav Process 86: 316-322.

Flemming TM, Beran MJ, Thompson RKR, Kleider HM, Washburn DA (2008) What meaning means for same and different: Analogical reasoning in humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), and rhesus monkeys (*Macaca mulatta*). J Comp Psychol 122: 176-185.

Flemming TM, Beran MJ, Washburn DA (2007) Disconnect in concept learning by rhesus monkeys: Judgment of relations and relations-between-relations. J Exp Psychol: Anim Behav Process 33: 55-63.

Flemming TM, Thompson RKR, Beran MJ, Washburn DA (2011). Analogical reasoning and the differential outcome effect: Transitory bridging of the conceptual gap for rhesus monkeys (*Macaca mulatta*). J Exp Psychol: Anim Behav Process 37: 353-360.

Gentner D (1999) Analogy. In Wilson RA, Keil F, editors. The MIT encyclopedia of the cognitive sciences. Cambridge, MA: MIT Press. pp. 17-20.

Gentner D (1983) Structure-mapping: A theoretical framework for analogy. Cogn Sci 7: 155-170.

Gillian DJ, Premack D, Woodruff G (1981) Reasoning in the chimpanzee: I. Analogical reasoning. J Exp Psychol: Anim Behav Process 7: 1-17.

Hofstadter DR (2001) Analogy as the core of cognition. In Gentner D, Holyoak KJ, Kokinov BN, editors. The analogical mind: Perspectives from cognitive science. Cambridge, MA: The MIT Press. pp. 499-538.

Oden DL, Thompson RKR, Premack D (2001) Can an ape reason analogically? Comprehension and production of analogical problems by Sarah, a Chimpanzee (*Pan troglodytes*). In Gentner D, Holyoak KJ, Kokinov BN, editors. The analogical mind: Perspectives from cognitive science. Cambridge, MA: The MIT Press. pp. 471-497.

Penn DC, Holyoak KJ, Povinelli DJ (2008) Darwin's mistake: Explaining the discontinuity between human and nonhuman animal minds. Behav Brain Sci 31: 109-130.

Premack D (1986) Minds with and without language. In Weiskrantz L, editor. Thought without language. New York: Oxford University Press. pp. 46-65.

Rattermann MJ, Gentner D (1998) More evidence for a relational shift in the development of analogy: Children’s performance on a causal mapping task. Cogn Dev 13: 453-478.

Smith JD, Redford JS, Haas SM, Coutinho MVC, Couchman JJ (2008) The comparative psychology of same-different judgments by humans (*Homo sapiens*) and monkeys (*Macaca mulatta*). J Exp Psychol: Anim Behav Process 34: 361-374.

Thompson RKR, Oden DL (2000) Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. Cogn Sci 24: 363-396.

Thompson RKR, Oden DL, Boysen ST (1997) Language-naïve chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. J Exp Psychol: Anim Behav Process 23: 31-43.

Wasserman EA, Young ME (2010) Same-different discrimination: The keel and backbone of thought and reasoning. J Exp Psychol: Anim Behav Process 36: 3-22.

Young ME, Wasserman EA (1997) Entropy detection by pigeons: Response to mixed visual displays after same-different discrimination training. J Exp Psychol: Anim Behav Process 23: 157-170.

Young ME, Wasserman EA (2001) Entropy and variability discrimination. J Exp Psychol: Learn Mem Cogn 27: 278-293.

Zentall TR, Wasserman EA, Lazareva OF, Thompson RKR, Rattermann MJ (2008) Concept learning in animals. Comp Cogn Behav Rev 3: 13-45.

Figures

Figure 1. Depiction of baseline and probe trial types in 4-item relational matching-to-sample task completed by *Homo sapiens* and *Papio papio*. Each sample configuration was presented first on a black background centered on the computer screen (top left) and disappeared upon touch allowing for the subsequent presentation of two choice arrays (top right). Exemplary sample arrays at each level of entropy level (0 to 2.0) are depicted along bottom.

Figure 2. Proportion of *different* responses for humans and baboons for all trials types (baseline and probe) combined. Error bars indicate standard error (SE); dashed horizontal line represents chance performance. All matches to *different* at entropy levels 0.81 to 2.0 are significantly above chance levels of responding.

Figure 3. Percent *different* matches for various configuration types at entropy levels 0.81 and 1.0 for baboons and humans. At entropy = 0.81 (a), *lead/end* refers to to configuration types wherein the unique stimulus either lead or ended the array (BAAA and AAAB). *Imbedded* configuration types were those wherein the unique object was in a middle position (ABAA or AABA). Dashed horizontal line represents chance selection.