Fagot, J. & Thompson, R.K.R (2011). Generalized Relational Matching by Guinea Baboons (Papio papio) in two by two-item analogy problems. Psychological Science, 22, 1304–1309

Generalized Relational Matching by Guinea Baboons (*Papio papio*)

in two by two-item analogy problems

Joël Fagot
CNRS - Université de Provence
Marseille France

&

Roger K. R. Thompson
Franklin & Marshall College
Lancaster PA
USA

Corresponding author : Joël Fagot, CNRS, LPC, Université de Provence, 3 Place Victor Hugo, Bât. 9, Case D 13331 Marseille cedex 1, France. Email: joel.fagot@univ-provence.fr

Abstract

Analogical reasoning is considered the “hallmark” of human reasoning, but can a monkey reason analogically as it was demonstrated in language or symbol trained chimpanzees? Past evidence suggests not. In our extensive study, 6 baboons of a group of 29 tested acquired a 2 x 2 item relational matching to sample (RMTS) task. Five of them further transferred their performances to novel exemplars of the identity and nonidentity relations, and continued to do so when the incorrect comparison choice contained a shape present also in the sample. When tested again 12 months later, the baboons RMTS performances revealed significant savings. Whereas language or the provision of symbol training boosts conceptual thinking, they appear not to be prerequisites for analogical reasoning as demonstrated in our symbol-naïve baboons.

Analogical reasoning is often considered the “hallmark” of human reasoning, but can a nonhuman primate likewise reason analogically? This question has been subject to vigorous debate since the seminal research by David Premack and his colleagues showing that Sarah, a language-trained chimpanzee, could construct analogical relationships as instantiated by geometric figures and objects (e.g., Gillan, Premack, & Woodruff, 1981). Much of the research driving the debate since Premack has focused on comparative analyses of birds and monkeys performances on two-item Relational Matching to Sample (RMTS) tasks (e.g., Fagot, Wasserman & Young, 2001; Thompson, Oden & Boysen, 1997; Young & Wasserman, 1997). In a typical RMTS task, a single relational exemplar consisting of two items that are either identical (e.g., AA) or nonidentical (e.g., AB) is presented as the sample. Once the animal has made an observing response, it is faced with two additional relational exemplars of the identity and nonidentity relations, both composed of novel items (e.g., CC & DE). The animal is rewarded for choosing the comparison stimulus pair that is relationally the SAME as that exemplified by the sample. Success on is task implies that the animal is judging ‘relations-between-relations, and this ability arguably is a cognitive foundation for analogical reasoning.

Results from this task and others have led to the conclusion of a “Profound Disparity’ (*pace* Premack, 1983; Thompson & Oden, 1996) between apes and humans on the one hand, and other animals on the other. Despite the high performance levels of both monkeys (e.g., Wasserman, Fagot & Young, 2001) and birds (e.g., Wasserman, Hugart & Kirkpatrick-Steger, 1995) on Same/Different discrimination problems, learning to associate condition cues to these relations does not facilitate acquisition of 2-item RMTS in these species (cf., Blaisdell & Cook, 2005; Flemming, Beran & Washburn, 2007). In contrast, experience with conditional cues via ‘token training’ facilitates acquisition of RMTS by chimpanzees (Thompson et al., 1997), possibly by providing the representational scaffolding for the symbolic encoding of the relations (Premack, 1983).

Language/token naïve baboons and pigeons acquire generalized RMTS task when the “Same’ & ‘Different’ relational displays contained up to 16 icons (Fagot, Wasserman, & Young, 2001; Cook & Wasserman, 2007), but their performance in Same/Different discrimination problems or RMTS tasks systematically decreased with smaller numbers of icons (Fagot et al. 2001; Young & Wasserman, 1997), suggesting perceptual constraints on the processing of relations between relations. Recently, Fagot & Parron (2010) showed that baboons can solve the 2-item RMTS task, but this performance could only be obtained when the elements comprising the relations were spatially very close.

All these recent studies suggest that the assumptions of the ‘profound disparity’ as well as the contribution of language/token training to relational thinking both warrant further assessment, which we carried out with baboons in the study reported here. Here we provide compelling evidence that a nonhuman primate other than an ape, language and token naïve, can acquire, generalize and retain over a 12-month period a two-item RMTS task. These results are indicative of a more broadly construed abstract conceptual ability in primates that one might suspect if it is a necessary cognitive foundation for analogical reasoning in humans.

Our study is unique in several ways. First, the N of 29 baboons makes it the most comprehensive examination of relational abilities in an animal species. Second, all our animals received an exceptionally large number of trials (from thousands to tens of thousands), providing information on the effect of sustained training on performance. Third, we were able to retest baboons that served as subjects in earlier studies. Finally, performance was retested one year after the initial training period, providing information on the long-term retention of abstract relations.

Experiment 1: Acquisition and Transfer of Relational Matching to Sample

Experiment 1 assessed individual differences in RMTS abilities and to test their performances on transfer trials involving stimuli different from the training set.

Methods

Twenty-nine male and female guinea baboons (*Papio papio*) varying in age from 4 social groups at two sites participated in this study (Table 1). Only the baboons at Marseille (N=7) had previously been tested on RMTS tasks. Two of them (B03 and B08) had participated in both the Fagot et al. (2001) and Fagot & Parron (2010) studies, and the other 5 (B05, B06, B07, B11 and B15) had participated only in the latter study. All the other baboons were experimentally naïve with respect to the RMTS procedure, but had prior experience in two-alternative forced choice (2AFC) and physical matching-to-sample (MTS) tasks (e.g., Fagot & Bonté, 2010).

Insert Table 1 about here

The task was presented with a series of automatic learning devices for monkeys (ALDM test system, see supporting information and Fagot & Paleressompoulle, 2010) which were freely accessible from the living enclosures. Each test system comprised a touch screen for stimulus display and responding, and a food dispenser for reward delivery. The RMTS procedure was used in both training and testing. In each trial, the monkeys firstly saw a sample pair composed of two identical (identity trial) or two different (nonidentity trial) shapes randomly selected from the stimulus set. These were displayed adjacent to one another on a black background in the middle of the screen. The sample disappeared when the animal touched it and two comparison pairs of shapes appeared in the center of on the left and right hemi-screen. Each comparison pair was comprised of shapes different from those of the alternate comparison pair or the sample. One comparison pair instantiated the identity relation (two identical shapes); the other instantiated the nonidentity relation (two different shapes). The baboons were rewarded when touching the comparison pair showing the same (identity or non identity) relation as the sample pair. Randomized sessions of 100 trials (50% of SAME and 50 of DIFFERENT trials), using an initial set of 10 images, were continuously repeated until a performance of 80% correct or more was maintained in three consecutive sessions. After this criterion was reached, the baboons received a total of 9 transfer tests each using a set of 10 new images (90 new images altogether). The transfer procedure is described in Table S1 in the supporting information available on-line.

Results

Table 1 shows the number of training sessions performed by each baboon, and its average final performance for the final 3 sessions. Six baboons (DRE, CAU, CLO, BAR, BOB and B15) learned the task to criterion. The relatively smooth profile of their learning curves, as shown in Figure 1, suggests that no baboon spontaneously comprehended the RMTS rule of matching relations (same & different) with relations (identity and nonidentity). Data were analyzed with an ANCOVA using the final performance as the dependant variable, and Gender, Test site, Age and number of Training sessions as independent factors. The main effects of Gender, Age and Training sessions were significant (Gender: F (1,17)= 20.9, p<. 001; Age: F (1,17)=12,41, p<. 01; Training sessions: F (1,17)=10,06, p<. 01) indicating greater performance levels by the successful males and with increasing age and number of training sessions. The scores correlated positively with the Number of Training sessions (partial correlation, partial r=. 26) and negatively with Age (partial r=-.22). There was in addition a negative correlation between Age and the number of sessions performed (partial r=-.33). Importantly, the effect of Test site (Marseille vs. Rousset) was not significant, F

(1,17)=2.22, p>.15), indicating that the prior experience or expertise on RMTS by the animals in Marseille (Fagot & Parron, 2010) had not facilitated their performances on the present task. Also, none of the possible two- or three ways interactions was significant (all ps >.5).

Insert Figure 1 about here

Regarding the transfer tests, only BOB showed an above chance performance in the first transfer test with 10 new images, but much more positive transfers were obtained in subsequent transfer tests (see Table S1). All transfer tests combined, all successful learner baboons except B15 positively transferred their RMTS performance to 90 novel shapes.

Experiment 2: Pitting object similarity against relational similarity

The apparent difficulty of the baboons to acquire the RMTS tasks as opposed to physical identity matching suggests that consistent with the assumptions of the 'Profound Disparity' the physical identity of objects is initially more salient for baboons than the abstract relations they instantiate (Thompson & Oden, 2000). Recent results on the transitory facilitative effect of differential outcome effects on RMTS by macaque monkeys (Flemming, Thompson, Beran & Washburn, in press) further indicate that attention to the relational content of stimuli in the RMTS task by monkeys is cognitively demanding. Hence, in Experiment 2, we tested if the baboons' attention to relational similarity, as reflected in their RMTS performances, might be adversely affected by a competing object similarity of one of the elements from sample and choice pairs.

Method

Prior to testing, the successful learners of Experiment 1 received a training regimen to teach them to solve the RMTS task with an increased separation (up to 60 pixels) between the two shapes of each pair (see supporting information and Table S2 for a description of the training procedure). They were then tested in 6 160-trial “Probe Distracter-Foil” sessions consisting of a balanced number of (SAME & DIFFERENT) baseline, identical to the training trials of Experiment 1. The structure of the probe trials was of the type AA (sample), BB (positive comparison) and AC (foil) for identity trials. It was AB (sample), CD (positive) and AA (foil) in non-identity trials. Note the sharing of the A between the sample and foil pairs. Hence, the animal should choose the incorrect foil in probe trials if shape similarity trumps relational similarity, or select the correct relational comparison stimulus if relational similarity prevails, as already found in children (Christie & Gentner, 2007**)** and chimpanzees (Matsuzawa & Premack 1985, unpublished data as cited in Premack 2010).

Results and discussion

All five baboons performed above chance on both baseline and probe trials (two tailed binomial test, all ps <. 05, see Figure 2), although the average performance was reliably higher on average in former (77.11%) than in latter trials (72.3%, one way ANOVA, F (1,4)=47.14, p<. 01). The above chance performances in probe trials suggest that the baboons preferentially, if not exclusively, attended to relational rather than perceptual similarity.

Insert Figure 2 about here

Experiment 3: Long term retention of relational processing

Flemming et al. (in press) showed that macaques could acquire the RMTS task under conditions of both differential reinforcement and differential punishment, but the RMTS performances dropped to chance levels when the monkeys experiences either differential reward or punishment alone, or neither differential contingency. This ‘now you see it, now you don’t,” aspect of the monkeys’ RMTS abilities suggests that the ability the retention of the relational matching rule may well be only transitory. Such a transitory memory trace may limit potential transfers of knowledge to other domains of relations. In Experiment 3, we investigated the long-term retention of our animals’ ability to perform the RMTS task one year after their initial training, in the absence of any RMTS practice during that time.

Methods

One year after Experiments 1-2, we replicated the initial training procedure with the 5 baboons who had acquired the RMTS task in Experiment 1. In that context, any reduction of the number of sessions to criterion relative to Experiment 1 would provide compelling evidence of long-term (one year) retention of the relational abilities in a monkey.

Results and discussion

DRE, CAU, CLO, BAR and BOB respectively required 39, 17, 77, 89 and 21 100-trial training sessions to reach the learning criterion in Experiment 3. Comparison of these frequencies to those of Experiment 1 (i.e., 179, 154, 321, 169 and 179, respectively) shows that the baboons required fewer training sessions to reach criterion in Experiment 3 than in Experiment 1 (one-tailed Wilcoxon test, p<. 05). The findings therefore show a massive saving of RMTS performance, suggesting that the baboons can retain the cognitive (relational) strategies learned to solve the RMTS task over a one year period.

General Discussion

The present results provide convincing evidence that prolonged training over hundreds or thousands of trials – what Premack (1983) aptly labeled ‘Dogged Training” - can result in the acquisition and long-term retention over 12 months of a generalized RMTS task by monkeys. Six of 29 baboons acquired a RMTS task and five of these transferred their RMTS performances to novel exemplars of the Identity and nonidentity relations. That so few baboons acquired the present RMTS task, in addition after tenths of thousand trials, leaves open the possibility that past failures to demonstrate 2-item RMTS in monkeys may be a consequence of inappropriate selection of the subjects and/or the premature termination of training after hundreds rather than the thousands of acquisition trials, as was possible with our automatic test-systems.

Interestingly, none of the animals that acquired the two-item color RMTS task in the Fagot & Parron (2010) study reached criterion in the present two-item geometric RMTS task. This finding suggests that the baboons' comprehension of the abstract relational matching rule is task (context) specific and narrowly construed, unlike patterns of analogical reasoning by humans that normally extends beyond a particular domain of knowledge (e.g., Gentner, 2003).

The baboons above chance RMTS performances were maintained when the incorrect comparison choice on both Identity and Nonidentity RMTS trials contained a 'distracter foil' comprised of a geometric stimulus element present also in the relational sample. This result implies that the baboons, much like 4.5 to 8.5 year old children (Christie & Gentler, 2007), and chimpanzees (Matsuzawa & Premack, 1985 as cited in Premack 2010), selectively attended to the analogical relational information in the displays and not to the perceptual information within the paired stimulus elements.

Remarkably, the baboons showed significant savings when after 12 months during which time they had no experience with RMTS tasks. This remarkable achievement shows that the lack of positive RMTS transfer across tasks (and presumably domains) is unlikely explained in baboons by an inability to retain relational information in the long-term.

Overall, the present results provide strong evidence that a strict ‘Profound disparity’ between ‘analogical’ hominoids (humans & apes) and ‘paleological’ perceptually bound old-world monkeys can no longer stand as originally postulated (Premack, 1983; Oden & Thompson 2000). It would be premature to speculate too broadly as to the precise, possibly associative mechanisms, reflected in the initially prolonged acquisition of the RMTS task. However, what cannot be denied is that old-world monkeys (at least some in our study) possess the requisite abstract cognitive foundation for analogical reasoning. Language training or the provision of symbol systems may facilitate acquisition of RMTS, but they are not necessary.

References

Blaisdell, A. P. & Cook, R. G. (2005). Two-item same-different concept learning in pigeons. *Learning & Behavior, 33,* 67-77.

Cook, R.G. & Wasserman, E.A (2007). Learning and transfer of relational matching-to-sample by pigeons. *Psychonomic Bulletin & Review, 14*, 1107-1114.

Christie, S. & Gentner, D. (2007). Relational Similarity in Identity Relation: The Role of Language. In Vosniadou, S. & Kayser, D. (Eds.). *Proceedings of the Second European Cognitive Science Conference*.

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. *Behavioral Research Methods,* 42, 507-512.

Fagot, J. & Parron, C. (2010). Relational matching in baboons (*Papio papio*) with reduced grouping requirements. *Journal of Experimental Psychology: Animal Behavior Processes, 36,* 184-193

Fagot, J., Wasserman, E. A., & Young, M. E. (2001). Discriminating the relation between relations: The role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes, 27,* 316-328.

Flemming, T. M., Thompson, R. K.R., Beran, M. J. & Washburn, D. A. (in press). Analogical reasoning and the differential outcome effect: Transitory bridging of the conceptual gap for rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*.

Flemming, T.M, Beran, M.J. & Washburn, D.A. (2007). Disconnect in concept learning by rhesus monkeys (*Macaca mulatta*): Judgment of relations and relations-between-relations. *Journal of Experimental Psychology: Animal Behavior Processes*, 33, 55–63.

Gentner, D. (2003). Why we’re so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195-235). Cambridge, MA: MIT Press.

Gillian, D. J., Premack, D., & Woodruff, G. (1981). Reasoning in the chimpanzee: I. Analogical reasoning. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 1-17.

Premack, D. (1983). Animal cognition. *Annual Review of Psychology, 34,* 351-362.

Premack, D. (2010). Why human are unique: Three theories. *Perspectives on Psychological Science, 5,* 22-32.

Thompson, R. K. R., & Oden, D. L. (1996). A profound disparity revisited: Perception and judgment of abstract identity relations by chimpanzees, human infants, and monkeys. *Behavioural Processes, 35,* 149 –161.

Thompson, R. K. R., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science, 24,* 363-396.

Thompson, R. K. R., Oden, D. L., & Boysen, S. T. (1997). Language-naïve chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 31-43.

Wasserman, E. A., Fagot, J. & Young, M. (2001). Same-different conceptualization by baboons (Papio papio): The role of entropy. *Journal of Comparative Psychology*, 115, 42-52.

Wasserman, E.A, Hugart, J.A. Kirkpatrick-Steger, K. (1995). Pigeons show same-different conceptualization after training with complex visual stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, 21, 248-252

Young, M.E. & Wasserman, E.A. (1997). Entropy detection by pigeons: Response to mixed visual displays after same-different discrimination training. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, *157-170.*

Acknowledgments

JF was supported by the ANR grant ANAFONEX. The authors acknowledge the technical support of Jean Christophe Marin and Georges Di Grandi. Data originally presented at the 16th annual International Conference on Comparative Cognition, Melbourne, FL, March 18-21, 2010.

Baboon	**Site**	**Sex**	**Age**	**N-Sessions**	**% Correct**
DRE	**Rousset**	**F**	**1.6**	**179**	**81.3**
CAU	**Rousset**	**M**	**1.9**	**154**	**81.7**
CLO	**Rousset**	**M**	**2.1**	**321**	**81**
BAR	**Rousset**	**M**	**2.7**	**169**	**80**
ARI	Rousset	F	3.4	406	65.6
ART	Rousset	M	3.4	222	76.6
BOB	**Rousset**	**M**	**3.4**	**179**	**84.7**
ANG	Rousset	F	3.8	302	61.7
VIO	Rousset	F	4.2	376	61.7
VIV	Rousset	M	4.4	241	72
VAN	Rousset	F	4.7	239	53.3
URA	Rousset	F	5.2	181	51
TAR	Rousset	F	6.7	337	49.3
ROM	Rousset	F	8.9	93	49.7
PIP	Rousset	M	10.4	32	49.7
ATM	Rousset	F	11	165	49.3
MON	Rousset	F	12	55	51.3
LEA	Rousset	F	13.1	75	50
MIC	Rousset	F	13.1	105	51
KAL	Rousset	F	13.5	48	50.3
BRI	Rousset	F	23.8	130	49.7
B03	Marseille	M	24	150	52.7
B05	Marseille	M	24	203	56
B06	Marseille	F	24	129	57.7
B07	Marseille	M	24	157	69.3
B08	Marseille	F	24	209	51.7

B11	Marseille	M	24	198	60.3
B15	**Marseille**	**M**	**24**	**270**	**81.3**
MIL	Rousset	M	32.2	126	50

Table 1. Demographic data on the participants and training information. Columns indicate (from left to right) the participants' name, living\test site, sex, age, number of 100-trial training sessions voluntarily performed and the mean final performance on the last 3 100-trial training sessions.

Figure captions

Figure 1. Learning curves of the 6 successful learners.

Figure 2. Performance obtained in the baseline and probe trials of Experiment 3.

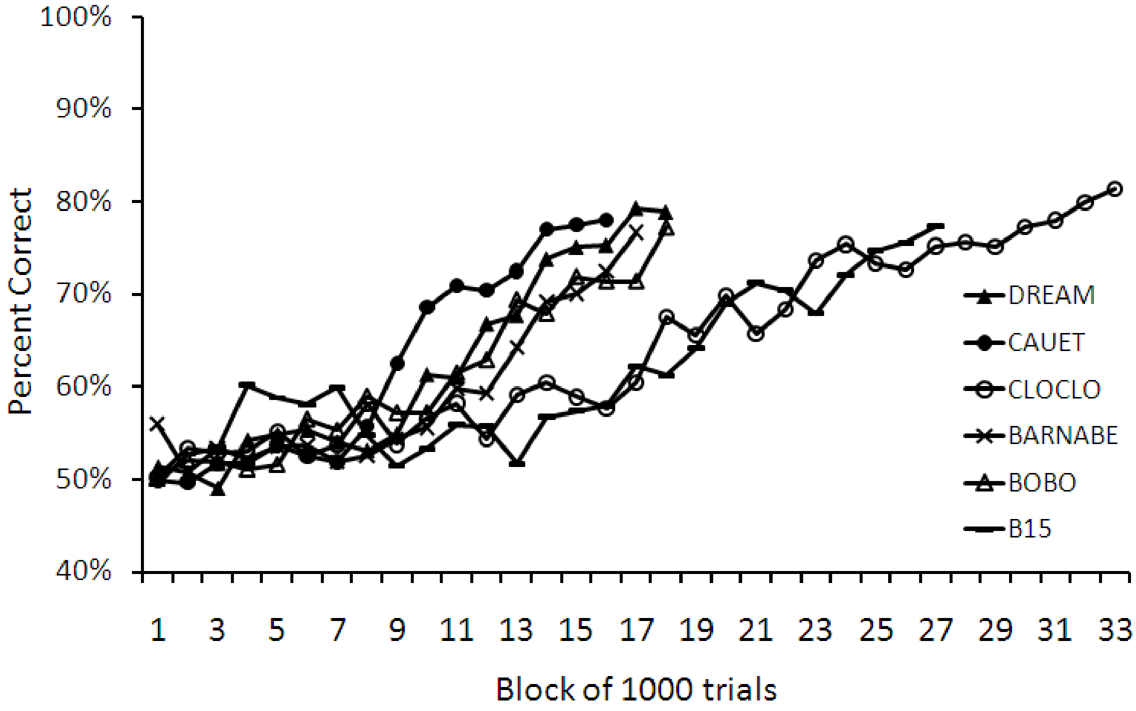

100%
90%
80%
70%
60%
50%
40%
Percent Correct
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
Block of 1000 trials
DREAM
CAUET
CLOCLO
BARNABE
BOBO
B15

Figure 1

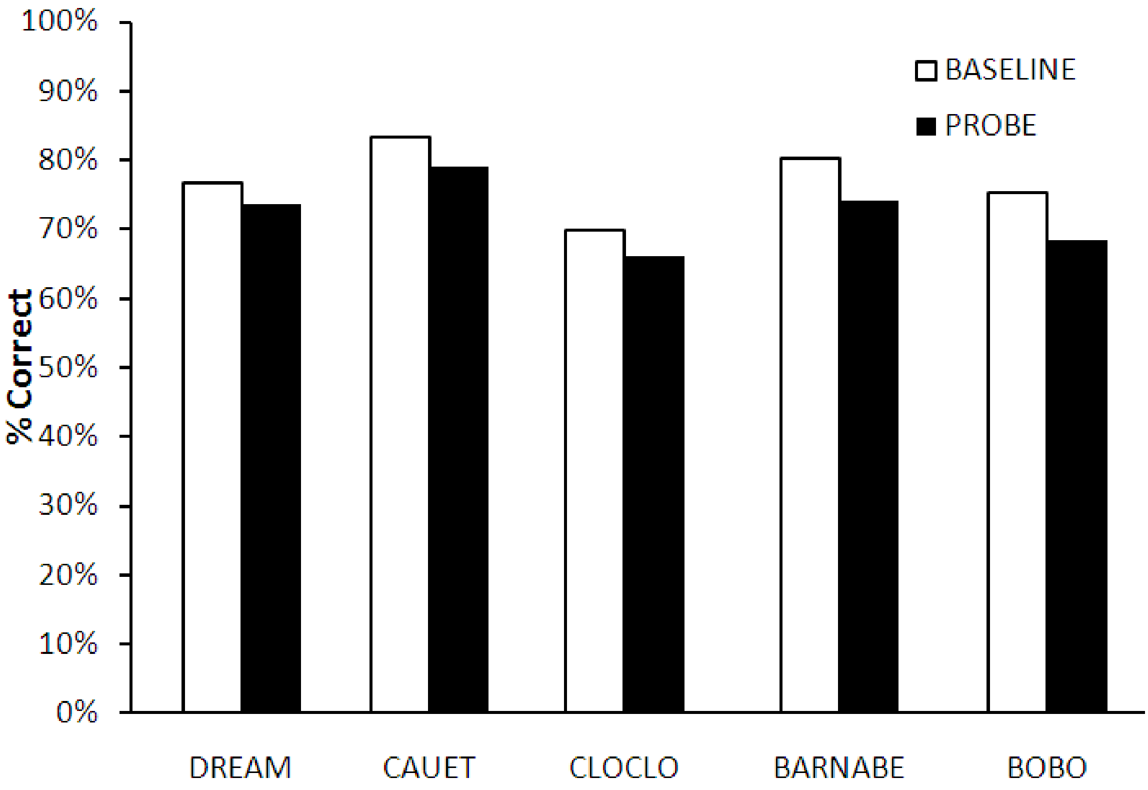

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
% Correct
☐ BASELINE
■ PROBE
DREAM
CAUET
CLOCLO
BARNABE
BOBO

Figure 2

SUPPORTING INFORMATION

GENERAL METHOD

The participants (N=29) belong to four different social groups from two test sites: Marseille and the Primate Center of Rousset-sur-Arc, France. The largest group from Rousset lives in a 670-m^2 enclosure connected. Each of the three smaller groups, based at Marseille, lived in one of three indoor areas connected to one of three outdoor. All baboons had a radio frequency identification (RFID) microchip implanted in each form-arm that served for auto-identification within the test boxes.

Baboons from the two test sites were tested with a new test equipment referred to as the Automatic Learning Device for Monkeys [ALDM, see their detailed description in Fagot and Paleressompoulle (2009) and Fagot & Bonté (2010)]. Each of the 13 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, through an opened sliding door on its rear side. 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 never food deprived.

EXPERIMENT 1: TRAINING AND TRANSFER PROCEDURE

All the stimuli used in training and test trials measured 100 X 100 pixels, and were painted white on a black background. Two pixels separated the inner side of two items from the same pair. Use of such a small separation between the pair items was aimed to promote relational processing, as suggested by Fagot & Parron (2010). The touch-screen turned black during an inter-trial interval (ITI) of 3-s, after a RMTS trials was accomplished. Correct responses were food rewarded. Incorrect responses gave rise to a 3-s time out during which the touch-screen was uniformly green before turning grey, which cued the animal that it could initiate a new trial via self-identification as described above. A trial was aborted if the animal failed to respond to either the sample or a comparison stimulus during their initial 5-sec display period. A grey screen cued the baboon that it could move on to the next trial via self-identification. Training continued until the baboons performed above 80% correct responses over three consecutive 100-trial sessions (50 identity and 50 non identity trials in randomized order), but ended after approximately one month of training if an animal failed to reach that criterion. The number of test sessions performed by each animal necessarily varied because participation as not coerced. Despite these variations, all participants completed a uniquely large number of training trials which (range 3,200 to 40,600 trials, mean = 18,790; SD=9,552, grand total = 545,100 training trials for the group).

The procedure used in the different transfer sessions is explained in Table S1. The transfer sessions consisted of a mixture of ‘baseline’ trials, initially drawn from the first training set of 10 geometric shapes, and ‘probe’ trials using a new set of 10 shapes in each transfer. Once a transfer was accomplished, each new set of stimuli was added to the old training set serving for baseline trials. With that procedure, transfer 1 used the initial training set of 10 stimuli for baseline trials, and 10 new stimuli for probe trials. The final transfer test 9 used a total of 90 familiar shapes for baseline trials in addition to 10 new shapes for probe trials. To avoid the confound effect of stimulus repetition, our analysis of the transfer trials of the successful learners included only the overall performance scores for the initial 20 probe trials performed during each transfer test. That performance is reported on the left hand side of Table S1.

A total of six baboons learned the task and five of them showed a positive transfer to the new shapes. Although not reported here, we continued to train some unsuccessful learners with the same procedure as described here, but with a larger number of stimuli (up to 100), after completion of this experiment. Two baboons (ART and VIV, see their initial performance in Table 1) acquired the RMTS task at the criterion of >80% correct RMTS after, respectively, a total of 339 and 379 100-trial training sessions. It is unclear if their success was a result of the increased number of trials, of the use of a larger stimulus set, or a combination of the two.

EXPERIMENT 2: METHOD AND PROCEDURE

Prior to completion of Experiment 2,_the 5 successful learners of Experiment 1 received training sessions to teach them to solve the task when the distance separating the two shapes of each pair was increased from 2 to 60 pixels. That training procedure consisted of 2 phases, in which the separation between the individual shapes composing the sample pair (phase 1) or comparison pairs (phase 2) was increased in a step-wise manner once an animal's performance was 70% correct or greater in the previous session. The training procedure and the number of training sessions performance by each participant during that training is reported in Table S2. Altogether, the completion of both Phases 1 & 2 by all 5 animals occurred over 29 to 55 100-trials training sessions (SD=16.9). The test session of Experiment 2 employed the training set of 10 shapes of Experiment 1. That experiment was run immediately after completion of Experiment 1.

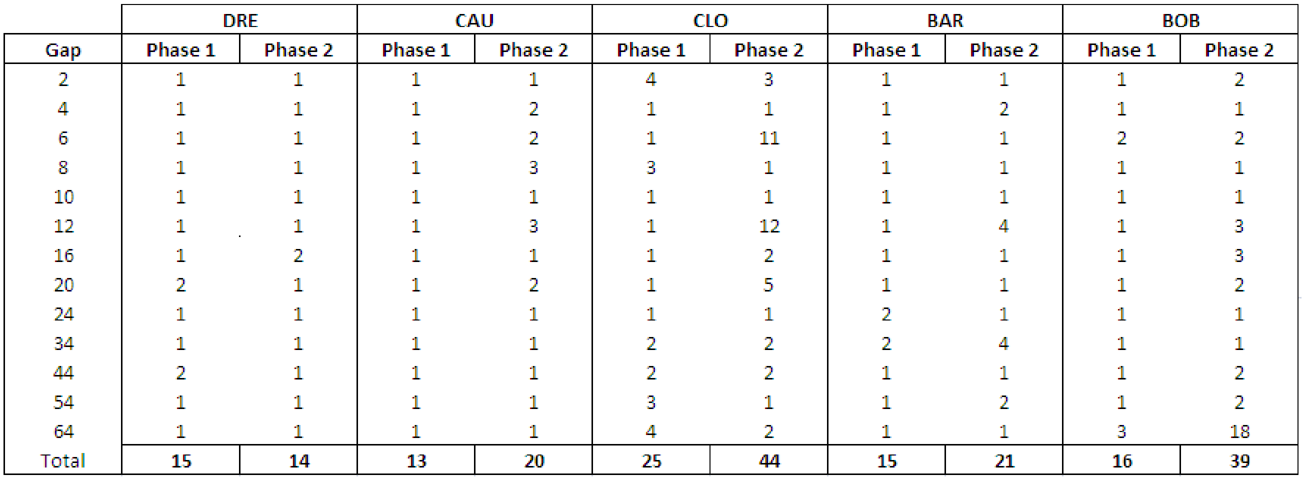
EXPERIMENT 3: METHOD AND PROCEDURE

The unique difference between the training procedure of Experiment 1 and 3 was the inter-item distance of the stimuli, which was now set to 32 instead of 2 pixels in all pairs. A larger gap sizes increased the difficulty of the RMTS task (Fagot & Parron, 2010). This experiment was run approximately 1 year after completion of Experiments 1 and 2.

Transfer	***New shape In***	***N-old Shape***	***N-Trials Baseline/Probe***	***N Sessions***	***DRE***	***CAU***	***CLO***	***BAR***	***BOB***	***B15***
1	Comparison	10	80/20	5	10	10	10	10	**15**	10
2	Sample	20	80/20	5	10	11	**17**	10	13	12
3	Comparison	30	96/4	10	9	12	12	6	12	11
4	Sample	40	80/20	5	**15**	**17**	8	13	12	11
5	Sample	50	80/20	5	**14**	**15**	12	**16**	10	7
6	Sample	60	80/20	5	11	**14**	**16**	**14**	**15**	**16**
7	Sample	70	80/20	5	**14**	11	**14**	11	**15**	12

8	Sample	80	80/20	5	**15**	11	**16**	**14**	**14**	13
9	Sample	90	80/20	5	**16**	**18**	**17**	**14**	**17**	10
				TOTAL	114	119	122	108	123	102
				% Correct	63.3%	66.1%	67.8%	60%	68.3%	56.7%
				Prob.	P<.05	P<.05	P<.05	P<.05	P<.05	P<.10

Table S1. Description of the nine transfer tests along with the performance of the 6 baboons that learned the RMTS task. From left to right, the columns indicate (1) number of the transfer test, (2) whether the new shapes were presented on the sample or comparison pairs during probe trials, the number of old (previously presented) shapes used for baseline trials, the frequency of baseline and probe trials within a transfer session, the number of transfer sessions, and the frequency of correct responses observed during the 20 first probe trials of each transfer test. Bold numbers indicate above chance performance within the 20 first probe trials of each transfer test (binomial two tailed test, $p<.05$). The sub-lines show the "grand total" and percentage correct obtained when considering the 20 first probe trials of 9 transfer tests. Prob shows the results of two-tailed binomial test ($p<.05$) comparing the grand total to chance performance (i.e., 90).



	DRE		CAU		CLO		BAR		BOB	
Gap	**Phase 1**	**Phase 2**	**Phase 1**	**Phase 2**	**Phase 1**	**Phase 2**	**Phase 1**	**Phase 2**	**Phase 1**	**Phase 2**
2	1	1	1	1	4	3	1	1	1	2
4	1	1	1	2	1	1	1	2	1	1
6	1	1	1	2	1	11	1	1	2	2
8	1	1	1	3	3	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1
12	1	1	1	3	1	12	1	4	1	3
16	1	2	1	1	1	2	1	1	1	3
20	2	1	1	2	1	5	1	1	1	2
24	1	1	1	1	1	1	2	1	1	1
34	1	1	1	1	2	2	2	4	1	1
44	2	1	1	1	2	2	1	1	1	2
54	1	1	1	1	3	1	1	2	1	2
64	1	1	1	1	4	2	1	1	3	18
Total	**15**	**14**	**13**	**20**	**25**	**44**	**15**	**21**	**16**	**39**

Table S2. Number of 100-trial sessions to criterion required by each baboon in the pre-training phase of Experiment 2. The pre-training procedure progressively increased the separation between the shapes composing the sample (Phase 1) and comparisons (phase 2), from 2 to 64 pixels in each pair, using a titration procedure. At each training step, the gap size was increased once the baboon achieved 70% correct in one training session. Each session comprised an even number of SAME and DIFFERENT trials presented in a random order. The other aspects of the training procedure were the as as for the training of Experiment 1.