Behavioural Processes 123 (2016) 74–83

Contents lists available at ScienceDirect

Behavioural Processes

journal homepage: www.elsevier.com/locate/behavproc

Can old-world and new-world monkeys judge spatial above/below relations to be the same or different? Some of them, but not all of them

Roger K.R. Thompson [a,*], Timothy M. Flemming [b], Carl Erick Hagmann [c]

[a] Franklin & Marshall College, United States
[b] Georgia State University, Spelman College, United States
[c] Franklin & Marshall College, Syracuse University, United States

ARTICLE INFO

Article history:
Received 14 July 2015
Received in revised form 2 November 2015
Accepted 6 November 2015
Available online 12 November 2015

Keywords:
Same-different
Above/below spatial relations
Analogical reasoning
Primates
New- and old-world monkeys

ABSTRACT

Chimpanzees (*Pan troglodytes*) with the aid of token training can achieve analogical reasoning, or the ability to understand relations-between-relations (e.g., Premack, 1976; Thompson, Oden, & Boysen, 1997). However, extraordinarily few numbers of old- and new-world monkeys have demonstrated this ability in variants of relational matching to sample tasks. Moreover, the rarity of replications leaves open the question of whether the results are normative for other captive colonies of the same species. In experiment one we attempted to replicate whether old world rhesus monkeys (*Macaca mulatta*) might demonstrate the same level of proficiency on a spatial above/below relational matching task as reported for old world baboons (*Papio papio*). None of the rhesus monkeys attained above chance performances over 10,000 training trials. In experiment two we attempted to replicate results demonstrating that new-world capuchin monkeys (*Cebus apella*) match above/below relations. The capuchin monkeys performed above chance only in the absence of 'Clever Hans' controls for cuing of the correct choice by the experimenters. These failures to replicate previously reported results demonstrate that some, but definitely not all monkeys can judge the equivalence of abstract 'relations between relations' and warrant further investigations into the behavioral and cognitive characteristics that underlie these similarities and differences within population and between individuals of different primate species.

1. Introduction

Comparative studies of concept learning have revealed good evidence that animals are capable of forming perceptual, associative, functional and relational concepts (e.g., Thompson, 1995; Zentall et al., 2008). Until very recently, however, studies with nonhuman primates pointed to an apparent 'Profound Disparity' between humans and chimpanzees on the one hand, and monkeys on the other, in their ability to explicitly judge the equivalence of 'relations-between-relations' (Premack, 1976a,b, 1978, 1983, 2007; Thompson and Oden, 1996, 2000).

The basis for this strong claim rested in large part on results from two-item relational matching-to-sample studies with chimpanzees (*Pan troglodytes*) and old world rhesus monkeys (*Macaca mulatta*) (e.g., Flemming et al., 2007; Thompson et al., 1997). In the standard two-item relational matching task an animal is presented with a sample comprised of either a pair of identical items (e.g., AA) or non-identical items (e.g., BC). The comparison choice stimuli are comprised of a pair of identical items (e.g., DD) and a pair of non-identical items (e.g., EF). Importantly, none of the individual elements comprising the sample and the comparison pairs are identical. In the present example: given the AA sample the correct choice is DD because the relationship of within-pair identity, exemplified by AA, is relationally the same as the within pair identity relation exemplified by DD and different from the within-pair nonidentity relation exemplified by EF. Conversely, given the BC sample the correct choice is EF because the relation of within-pair nonidentity exemplified by BC is relationally the same as the within-pair non-identity relation of EF and different from the identity relation of DD.

Hence, as described above, success on a 2×2 relational matching-to-sample task hinges on an animal's ability to discriminate 'relations (i.e., same or different) between relations' (i.e., identity or nonidentity). Transfer of this ability to trial-unique pairs of samples and comparison stimuli is taken as evidence of a 'higher-order relational concept' that is arguably the requisite cognitive foundation upon which the formal analogical reasoning demonstrated by humans and some chimpanzees rests (Gentner, 1989; Gillan et al., 1981; Premack, 1983; Oden et al., 1998, 2001).

* Corresponding author.
E-mail address: roger.thompson@fandm.edu (R.K.R. Thompson).

http://dx.doi.org/10.1016/j.beproc.2015.11.010
0376-6357/

Successful acquisition of relational matching-to-sample and more formal analogical reasoning tasks by chimpanzees is facilitated by their having previously learned to associate tokens or icons with the 'same' and 'different' relations (Premack, 1983). It may well be that the tokens function as symbolic representations of abstract relations thereby making perceptually 'concrete' and meaningful that which is abstract (Premack, 1983; Thompson and Oden, 2000; Thompson et al., 2001). However, Flemming et al. (2008) found that, in contrast to humans and 'language trained' chimpanzees (e.g., Brakke and Savage-Rumbaugh, 1996), token training—regardless of prior symbolic meaning (Washburn and Rumbaugh, 1991)—had no such scaffolding effect on acquisition of a RMTS task by old world rhesus monkeys.

Demonstrations that old world Guinea baboons (*Papio papio*) could discriminate and match same/different relations presented in arrays of same or different icons varying in number (Wasserman et al., 2001; Fagot et al., 2001) prompted Flemming et al. (2007) to present old world rhesus monkeys (*M. mulatta*) an array of horizontally aligned identical (same) pictorial icons together with an equal number of horizontally aligned different icons in a two-choice conditional discrimination task. The correct choice of same or different icons was indicated by the color of the background on which they were presented simultaneously. All of the monkeys acquired the task with novel 8×8 same/different icons per trial and then continued to choose correctly when the number of icons was systematic reduced from 8×8 to 2×2 item same/different displays via a performance driven titration procedure. Nevertheless, all of the animals subsequently failed the 2×2 item relational matching task perhaps because no 'recoding' of relations-between-relations is required in this two-choice discrimination.

Flemming (2011) similarly trained both rhesus (*M. mulatta*) and capuchin (*Cebus apella*) monkeys on a conditional same/different discrimination task in which 2, 4 or 6 S/D stimulus arrays were displayed on a common background the color of which identified the correct relational choice. The number of pellet rewards associated with each level was inversely related to the number of icons; 4 pellet rewards for correct choices on two-item S/D displays, two pellets for correct 4-item choices and one pellet for 6 item choices. Presentation of 2×2, 4×4 or 6×6 S/D icons displays was titrated up or down contingent upon performance levels across trial blocks. Both the rhesus and capuchin monkeys acquired the task. Interestingly, however, the level of perceptual stimulus variability mattered less for the capuchins than for the rhesus monkeys. Two of the 6 tested capuchins, for example, spontaneously judged two-item pairs as 'same' or 'different' within 200 trials whereas rhesus monkeys succeeded within an average of 150 trials.

Given Fleming's (2011) above results might not rhesus and capuchin monkeys**,** be capable of two-item RMTS given the appropriate procedural scaffolding? Flemming et al. (2011) provided the first clear evidence of significant, albeit transitory, RMTS performances by rhesus monkeys using a differential outcomes procedure in which both reward magnitude (number of pellet ratios) and punishment duration (Time-Out ratios) were differentially applied to trial types. Only those animals experiencing both the reward and punishment conditions attained performance levels significantly above chance. However, under non-differential outcome conditions the performance of all the animals were at chance levels.

Fagot and Thompson (2011) found that training old-world Guinea baboons over thousands of trials with multiple stimulus sets of computer generated icons was instrumental in their acquiring a two-item RMTS task. However, of the 29 animals in the training cohort only 6 met the statistically significant above chance performance criterion, and of these, 5 subsequently generalized their RMTS ability to novel exemplars of the same/different relations including those in which the incorrect comparison choice contained an icon present also in the sample pair. Significant savings were obtained in the animals' performances when they were tested 12 months later. Importantly, Flemming et al. (2013) subsequently demonstrated these baboons, like human participants, solved the RMTS task by categorical abstraction of relations and not by direct perception of stimulus variability as measured in units of entropy (cf, Fagot et al., 2001).

As noted above, two of the capuchin monkeys in the Flemming (2011) study judged two-item pairs as 'same' or 'different' within 200 trials raising the question of whether or not this new world species might be more predisposed than old-world monkeys to acquire a two by two item RMTS task. Results reported by Truppa et al. (2011) suggest not. Truppa et al. (2011) trained 5 tufted capuchins on a series of simultaneous RMTS sets in which novel relational pairs were introduced when the animals met the correct performance criterion—typically over many thousands of trials—with prior sets.

In a second experiment all 5 animals were tested on the RMTS task with large novel stimulus sets including sample and choice stimuli comprised of 4 and 2 icons. Only one animal's RMTS performance was significantly above chance, again after thousands of training trials with the two-icon sets, and then only after she had first reached criterion, again after many thousands of trials, with 4-icon sets; a level which none of the other 4 animals achieved.

Kennedy and Fragaszy (2008) also found significant individual differences in the acquisition of a Relational Relative Size matching search task by 4 capuchin monkeys that had years of prior experience with tasks involving spatial relations and tool use (e.g., Cummins-Sebree and Fragaszy, 2005; Fragaszy and Cummins-Sebree, 2005). After having seen the experimenter retrieve a reward from a set of differently sized cups, the animal's task was to retrieve a hidden food reward from the analogously same sized cup relative to the others in its set. All 4 animals reached greater-than-chance performance criterion over a range of 309–1113 trials on a preliminary physical matching task in which they were rewarded for choosing the cup from their set that was the exact same size as the baited cup in in the experimenter's set. Three of the animals, however, subsequently failed to reach criterion of the relational matching phase after 600 trials and were not tested further. However, the fourth capuchin reached criterion on this RMTS task not after thousands of trials, as was the case for Truppa et al. (2011) single successful animal, but after a training phase of only 143 trials. He continued to perform above chance on all subsequent test phases—over a total of 281 trials—including the final one in which his stimulus set differed in color, shape, and size from that the experimenter.

Apart from the performance differences revealed in the above cited studies the suggestion that not all relational judgments are necessarily the same is supported also by the results of Dépy et al. (1999) and Spinozzi et al. (2004). These studies showed that baboons and capuchin monkeys, respectively, acquired a spatial relational matching task in which the relations 'above' and 'below' were instantiated in the sample and comparison choice stimuli by a dot stimulus presented either above or below a horizontal line stimulus.

In the study by Dépy et al. (1999), 5 baboons met the performance criterion of 80% correct spatial matches with the initial matching task with two line-dot distances (0.5 and 2.3 cm) after, on average, 3024 trials (range 1095–5696). In a second experiment the same baboons' performances remained above chance over 4 consecutive 144-trial sessions (576 trials) regardless of systematic differences in the distance between the dot and line stimuli instantiating the sample and the comparison stimuli. In a third experiment two of the baboons were tested on their ability to make same/different 'above/below' relational judgments in a 'go/no-go' task with stimuli comprised of variations of the fonts used to produce the letter B and the number 3. The two animals reached