



Published in final edited form as:

*Cognition*. 2013 December ; 129(3): . doi:10.1016/j.cognition.2013.08.001.

## Fading Perceptual Resemblance: A Path for Rhesus Macaques (*Macaca mulatta*) to Conceptual Matching?

**J. David Smith,**

Department of Psychology, University at Buffalo, The State University of New York

**Timothy M. Fleming,**

Department of Psychology, Reed College

**Joseph Boomer,**

Department of Psychology, University at Buffalo, The State University of New York

**Michael J. Beran,** and

Language Research Center, Georgia State University

**Barbara A. Church**

Department of Psychology, University at Buffalo, The State University of New York

### Abstract

Cognitive, comparative, and developmental psychologists have long been intrigued by humans' and animals' capacity to respond to abstract relations like sameness and difference, because this capacity may underlie crucial aspects of cognition like analogical reasoning. Recently, this capacity has been explored in higher-order, relational matching-to-sample (RMTS) tasks in which humans and animals try to complete analogies of sameness and difference between disparate groups of items. The authors introduced a new paradigm to this area, by yoking the relational-matching cue to a perceptual-matching cue. Then, using established algorithms for shape distortion, the perceptual cue was weakened and eliminated. Humans' RMTS performance easily transcended the elimination of perceptual support. In contrast, RMTS performance by six macaques faltered as they were weaned from perceptual support. No macaque showed evidence of mature RMTS performance, even given more than 260,000 training trials during which we tried to coax a relational-matching performance from them. It is an important species difference that macaques show so hesitant a response to conceptual relations when humans respond to them so effortlessly. It raises theoretical questions about the emergence of this crucial capacity during humans' cognitive evolution and during humans' cognitive development.

### Keywords

same-different; concept learning; relational matching; primate cognition; comparative cognition

---

Cognitive psychologists have long been interested in humans' and animals' capacity to respond to abstract relations like sameness and difference (Wasserman & Young, 2010).

---

© 2013 Elsevier B.V. All rights reserved.

Correspondence concerning this article should be addressed to J. David Smith, 346 Park Hall, University at Buffalo, The State University of New York, Buffalo, NY 14260, psysmith@buffalo.edu.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

James (1890/1950) referred to relational concepts as the keel and backbone of thinking. Relational concepts are one basis for humans' capacity for analogical reasoning and problem solving (e.g., Hummel & Holyoak, 2003). They figure prominently in humans' cognitive-developmental change (Gentner, 2003). They ground theoretical discussions of language's role in supporting symbolic representation (Gentner & Ratterman, 1991; Goldin-Meadow, 2003). They could represent a pivotal discontinuity between humans' and animals' cognitive system (e.g., Locke, 1690; Penn, Holyoak, & Povinelli, 2008). For these reasons, a full understanding of the phylogenetic emergence of relational concepts is one of comparative psychology's premier goals. That understanding could help explain the emergence of this crucial capacity during humans' cognitive evolution, the timing of that emergence, and perhaps the distinctive character of the human cognitive system.

The issue of relational learning was crucial even in psychology's early behaviorism-cognitivism debates (Kohler, 1918/1958; Spence, 1937). It was realized that relational (e.g., Same-Different) tasks can require a cognitive abstraction beyond the primary stimulus qualities, requiring animals to transcend behaviorist response tendencies. From this derives a principal idea in the area that relational concepts like same and different are cognitively sophisticated and phylogenetically restricted (Herrnstein, 1990). In fact, animals find these kinds of relational judgments difficult and their same-different performances are sometimes fragile (Carter & Werner, 1978; Cumming & Berryman, 1961; Farthing & Opuda, 1974; Fujita, 1982; Holmes, 1979; Premack, 1978; Rumbaugh & Pate, 1984; Shields, Smith, & Washburn, 1997; Washburn & Rumbaugh, 1991; Wright, Shyan, & Jitsumori, 1990).

For example, pigeons sometimes have difficulty transferring relational rules from one stimulus domain to another (Zentall & Hogan, 1974, but see Cook, 2002). Premack (1978) concluded that it would be difficult to "talk to a pigeon" because they respond so weakly to relational cues and so strongly to absolute stimulus cues. On a cross-species shopping trip to buy a wall-hanging, for example, the human would be thinking "this will go great with my couch" (a relational judgment). The pigeon would be thinking "red" (a perceptual judgment). The tradeoff between the perceptual and relational levels of processing in matching tasks is crucial to the theoretical and empirical context of the present article (see also Gibson & Wasserman, 2003, 2004). This tradeoff also illuminates by contrast humans' status as the premier abstract and analogical cognitive system, demonstrating that other cognitive organizations are possible and extant in phylogeny.

Old World primates have had more success on relational tasks (Katz, Wright, & Bachevalier, 2002; Shields et al., 1997; Wasserman, Young, & Fagot, 2001; Wright, Cook, & Kendrick, 1989; Wright, Rivera, Katz, & Bachevalier, 2003; Wright, Santiago, & Sands, 1984; Wright, Shyan, & Jitsumori, 1990). However, their Same-Different (SD) concepts are also limited. In Shields et al. (1997), macaques responded persistently to absolute stimulus cues in an SD task before achieving successful performance. In D'Amato and Colombo (1989), monkeys failed to transfer an SD concept from static to dynamic stimuli. Monkeys have also failed to show robust, transferable SD performance when small training-item sets led them to favor response strategies based on specific-item associations (D'Amato, Salmon, & Colombo, 1985; Katz et al., 2002).

These limitations set the stage for the discovery of species differences across the primates in higher-order tasks of relational judgment, especially the relational matching-to-sample (RMTS) task. In this task, given a pair of same objects (AA) as a sample, subjects should choose a second pair of same objects (BB) instead of a pair of different objects (CD). Or, given a pair of different objects (AB), they should choose another pair of different objects (CD) and not a pair of identical objects (EE). Thompson et al. (1997) found evidence using this paradigm of relational matching in four of five chimpanzees (*Pan troglodytes*). The

chimpanzee Sarah also performed in relational-cognition tasks that required her to complete and even create functional analogies (Gillian, Premack, & Woodruff, 1981; Premack, 1976, 1986).

Thus arose the influential idea that apes are *analogical* (Thompson & Oden, 2000) in the sense of dealing fluently with analogies and abstract relations. However, it is important that Thompson et al.'s (1997) chimpanzees had a history of learning conditional discriminations involving tokens matched to object pairs, a kind of pair naming or symbol training that might have facilitated their RMTS performance. The chimpanzee Sarah had received extensive training with plastic tokens symbolizing “same” and “different.” Premack (1976; 1986), Gillian et al. (1981), and Thompson and Oden (1996) all acknowledged the critical role of her symbols for *same* and *different* in fostering her relational cognition. Therefore, these apes might have been analogical as a result of symbolic pre-exposure or as a result of some aspect of their biology such as large brain size. (The story of the emergence of humans' analogical capacity could be very different depending on which of these explanations turned out to be correct.)

For one or both of these reasons, the case for non-apes is different. Fagot, Wasserman, and Young (2001) presented RMTS tasks to baboons. Given 16-icon arrays as sample and choices, baboons demonstrated successful relational matching. However, additional studies showed that they were not responding purely relationally, but in a manner correlated with the visual-entropy or information-theoretic complexity of the stimulus arrays. When the strength of the visual-entropy cue was weakened, by reducing the sizes of the stimulus arrays, baboons' performance fell sharply to reveal a constraint on their abstract conceptual ability. In another experiment, however, these baboons successfully matched 2-item choice arrays to 16-item sample arrays. Perhaps the detection of visual entropy in the sample arrays primed conceptual processing as a correct relational choice was made. In a very recent study, Flemming, Thompson, & Fagot (2013) also showed that one could not explain baboons' matching performances based solely on the cues of perceptual variability or visual entropy. To the contrary, it appears that some conceptual processes add in along with perceptual processes in comparisons of four-item arrays. In a sense, Flemming et al. (2013) may be thought of as foreclosing the possibility that primates' array-matching performances only reflect the use of a visual-entropy cue.

Findings like these ground the idea in the literature that Old World monkeys are on the verge of true relational-matching performances if only the appropriate facilitating conditions were found. This article joins others in trying to establish these facilitating conditions.

For example, Fagot and Parron (2010) used a spatial fostering paradigm. Their RMTS task used adjacent color-patch stimulus elements that initially were spatially conjoined. Thus, both RMTS samples and the choices could initially be considered as single to-be-matched stimuli. Later, gaps between the color blocks were introduced to foster relational comparisons—now between spatially separated pairs of stimuli. Baboons' performance collapsed to chance with a spatial gap of just 30 pixels. Fagot and Parron (2010) suggested that the contiguous stimuli gave baboons a way to successfully complete the task using a local mode of processing. By contrast, to group the elements separated by gaps into a higher-order structure proved difficult or impossible for monkeys (Fagot & Deruelle, 1997; Spinozzi, De Lillo & Truppa, 2003).

Flemming et al. (2011) used a hedonic fostering paradigm. Rhesus monkeys completed an RMTS task with differential outcomes corresponding to same and different trials (e.g., correct same and different responses might earn either double or single food rewards; incorrect same and different responses might garner either 45 s or 10 s timeout periods). The

idea was to recruit sharp attention to the character of the two trial types. However, macaques' performance, although above chance during differential-outcome training (in only the coupled reward and timeout differential condition), collapsed to chance during nondifferential transfer sessions. Flemming et al. (2011) concluded that the differential reward and punishment contingencies in combination provided an effective but transitory relational scaffolding for monkeys. Still, this methodology did not establish relational matching as an independent or predominant strategy that outlasted the differential hedonic-valence manipulation.

Fagot and Thompson (2012) fostered baboons' RMTS performance using the approach that Premack (1983) called "dogged training." Six of 29 baboons met a criterion of 80% correct within 15,000–30,000 trials on an RMTS task constructed using 10 repeating geometric shapes. Five of these baboons also showed above-chance RMTS performance 12 months later with novel stimuli. Though limited by the potential dependence on specific-item or exemplar-based associative strategies in these tasks, the results suggest that baboons have some cognitive foundation for detecting and matching relational sameness and difference.

The overall pattern of findings for non-apes indicates the fragility of relational matching in the absence of auxiliary supports to performance (visual entropy in Fagot et al., 2001, contiguous stimuli supporting local processing in Fagot & Parron, 2010, hedonic overtones in Flemming et al., 2011). However, it may not indicate the complete absence of the capacity for relational matching. Yet if monkeys have the relational capacity, why does it require these supports to manifest? And under what conditions would it manifest robustly—absent those supports?

Our approach toward finding these conditions can be grounded in the literatures on concepts and conceptual development. For human children and adults, surface similarities seem to be integral in determining whether participants recognize relational similarity in order to solve a problem (Catrambone, 2002; Dunbar, 2001; Gentner, Rattermann, & Forbes, 1993; Ross, 1987). Moreover, Goldstone and Barsalou (1998) pointed out that there are commonalities between perceptual and conceptual processes and that abstract conceptualization has a deep-rooted dependence on perception. Therefore, one can see that perceptual similarity might be a good carrier-wave signal with which to broadcast relational information to nonhuman primates.

Therefore, we tested whether perceptually-mediated matching might give way naturally to relational matching. We introduced the RMTS task with same and different stimulus pairs that had to be matched to perceptually similar pairs with the same relation. Then, following established and quantifiable methods of statistical shape distortion, we decreased the perceptual similarity between to-be-matched pairs, fading the perceptual cue, weaning participants from this signal, and leaving behind only the relational cue. We hypothesized that this technique might foster successful RMTS performance in our macaques. Thus, our RMTS paradigm instantiated Quine's career-of-similarity hypothesis (Quine, 1969, p. 138) by transitioning animals from brute, perceptual similarity toward perceiving more sophisticated and abstract forms of similarity. The task also controlled for array size and visual entropy by always presenting only two-item arrays. It controlled for item separation so that global processing was required of animals. Yet it still allowed us to gradually wean animals toward a full relational performance—flexibly at different rates, reversibly if needed, depending on the individual's tolerance for having the perceptual-similarity support weakened. Given our comparative theoretical perspective—to understand relational matching and analogical reasoning broadly across species and to trace its phylogenetic roots—we included both human and macaque participants in our research.

## Experiments 1A, 1B, and 1C: Humans

Experiments 1A-1C explored relational matching-to-sample performance by humans. In Experiment 1A, we studied humans' response as perceptual support was gradually removed from within an RMTS task. In Experiment 1B, we studied humans' transition to mature RMTS responding as perceptual support faded. And, in Experiment 1C, we studied humans' mature RMTS responding absent any perceptual-support training. These experiments gave us a grounding human perspective for considering macaques' capacity for relational matching.

### Method

**Participants**—Participants were 88 undergraduates from the University at Buffalo, the State University of New York (UB) who participated in a session lasting about an hour to fulfill a course requirement. Ten and nineteen participants, respectively, completed Experiments 1A and 1B. Thirty and twenty-nine participants, respectively, completed the two conditions of Experiment 1C. Participants were in their late teens or early twenties with normal or corrected-to-normal visual acuity.

**Dot-distortion stimuli**—Our task required a stimulus domain within which the similarity relationships among stimuli were well-understood, quantifiable, and manipulable. Accordingly, the stimulus materials for the RMTS task were created with a well-established method that generates families of polygons from originating prototypes (Posner, Goldsmith, & Welton, 1967), extended in human research by Smith and Minda (2001, 2002), and extended in cross-species research on non-human primates by Smith, Redford, and Haas (2008a,b). Using this method, we could control the overall perceptual similarity between two pairs of same stimuli and between two pairs of different stimuli. The dot-distortion stimulus domain lent additional strengths to our RMTS task. First, these stimuli let us create single-use trials indefinitely, sampling stimulus pairs from an essentially infinite population without replacement or repetition. Thus, we could study performance uncontaminated by any specific-item memorization (see also Brooks & Wasserman, 2008). Second, these stimuli let us study performance within a psychological similarity space of high dimensionality and complexity, perhaps capturing something of the complex, probabilistic similarity relationships among members of natural-kind categories.

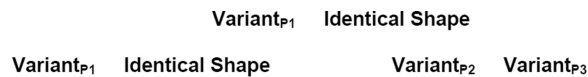
In our method, shapes were defined by randomly selecting nine points from within a 30 by 30 grid. These nine points became the vertices of 9-pointed polygons. After selecting the 18 coordinates for a polygon (9 X coordinates and 9 Y coordinates), the shapes were centered within the 30 by 30 grid using a normalization algorithm. With the dot positions chosen for a polygon, the shape was magnified to be more visible. Each pixel position in the 30 × 30 grid was mapped to a 3 by 3 pixel square on the screen, and the dot was placed in the center of the appropriate 9-pixel cell on the screen. In this way the stimulus patterns were magnified threefold from being created within a virtual 30 by 30 coordinate space to being shown on an actual 90 by 90 pixel space on the screen. Finally, the DrawPoly procedure within Turbo Pascal 7.0 connected successive dots by lines and filled the resulting polygon shape in yellow. This followed the common practice of presenting the dot distortions as random polygon shapes (Homa, Rhoads, & Chambliss, 1979; Homa, Sterling, & Trepel, 1981).

**RMTS trials**—Each trial consisted of two polygon shapes presented side by side in the center-top of an 11.5-inch computer screen against a black background. On some trials, these two polygons were identical. On different trials, the two polygons were very different, having been created based on two different random selections of points from within the 30 × 30 stimulus grid. Two more shape pairs—a same pair and a different pair—were presented

at the screen's bottom-left and bottom-right. The placement of the same or different pair to left or right was decided randomly on each trial. A cursor controlled by the participant's keypresses was illuminated between the response pairs. As the top pair of shapes was same or different, respectively, the participant was to move the cursor to touch the pair on the bottom that was same or different.

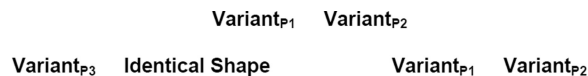
For correct responses, participants heard a computer-generated whooping sound and earned one point. For incorrect responses, participants heard a computer-generated buzzing sound, lost one point, and experienced a 2-s timeout period.

**The method of fading similarity**—The crucial aspect of our method was to use the prototype-variant technique that has been so influential in the literature on human and animal categorization (e.g., Smith & Minda, 2002; Smith, Redford, & Haas, 2008a). In our RMTS task, Same Trials had this general structure:



The rationale behind this was as follows. Variants of Prototype 1 (P1) will generally share a family-resemblance—that is, will look somewhat alike. Therefore, we could make the correct answer to a same trial actually look perceptually like the same pair at the top of the screen that the participant was responding to. This provided a perceptual-similarity cue supporting the participant in making the correct response. However, the prototype-variant method also gave us indefinite scope to make the correct answer to a Same trial look less perceptually like the same pair that the participant was responding to. Thus, we could progressively fade the cue of perceptual resemblance, and in the end ask whether participants can learn to go without it.

By a similar logic, Different Trials had this general structure:



The rationale behind this was as follows. Variants of Prototype 1 (P1) and Prototype 2 (P2) will generally share a family-resemblance to other variants of P1 and P2, respectively—that is, they will look somewhat alike. Therefore, we could also make the correct answer to a different pair look perceptually like the different pair at the top of the screen that the participant was responding to. This provided a supporting perceptual pointer toward correct responding. But then, by increasing the range of the variants, we could reduce this perceptual similarity systematically, fade the perceptual-similarity cue, and in the end ask whether participants can go without it.

**The fading-similarity algorithm**—The degree of perceptual similarity available in the trials depended on variables Level and Move. At Level 91, all 18 X-Y coordinates in a 9-dot prototype were moved by 0 pixel positions (Move=0), so the all variants were identical to the prototype and to each other. The perceptual-similarity was maximally strong. From Levels 90 down to 82, 1 dot, 2 dots, and so forth, up to 9 dots, were displaced in both their X and Y coordinates one pixel position (Move=1) from their place in the prototype. Now variants could differ but only slightly. The process continued. At Levels 73, 64, 55, 46, 37, 28, 19, 10, 1, respectively, the value of Move for all 18 coordinates was 2, 3, 4, 5, 6, 7, 8, 9, and 10 pixels. Move was applied to coordinates of a shape as follows. For each vertex in the polygon, one of its coordinates was displaced the full range of Move (e.g., Move=10, -10 or +10, nothing in between). The other coordinate was displaced somewhere in the range of

Move (e.g., Move=10, -10, -9, ..., +9, +10). Which coordinate received which treatment was decided randomly for each dot for each shape. As the value of move increased, the overall similarity between variants of the same prototype was decreased. Occasionally, the shape distortion algorithm may have created items that could have been the original pair of items viewed from another perspective. This was another source of initial perceptual support in the task, and of course this support also faded as perceptual support was withdrawn.

For large values of Move, coordinates could be displaced far across the original  $30 \times 30$  space within which stimuli were created. These large variations in some cases amounted to just creating the next random polygon. Thus, this method did progressively and profoundly degrade the availability and strength of the perceptual-similarity cue.

**Mature or zero-similarity trials**—In addition, we added a final phase to the task given to the 19 participants in Experiment 1B and to participants in one of the two conditions in Experiment 1C. In that final phase, some trials had this general structure.

	Variant <sub>P1</sub>	Identical Shape		
Variant <sub>P2</sub>	Identical Shape		Variant <sub>P3</sub>	Variant <sub>P4</sub>

Notice that in these zero-similarity trials, it simply was not possible to rely on any kind of perceptual cue at all. The same alternative to below-left is created from a qualitatively different prototype, and so there could be no helpful resemblance guiding the participants' decisions. We call these zero-similarity trials mature RMTS trials, because these trials stand at the end of the perceptual-weaning process, when perceptual support had been removed completely and only the conceptual relation was available for matching.

By a similar logic, the mature different trials in the final phase of Experiment 1B and in 1C had this general structure:

	Variant <sub>P1</sub>	Variant <sub>P2</sub>		
Variant <sub>P3</sub>	Identical Shape		Variant <sub>P4</sub>	Variant <sub>P5</sub>

In this case, too, it was impossible to rely on any perceptual cue at all. The different alternative to below-right is created from qualitatively different prototypes, and so there could be no helpful resemblance guiding participants to choose that different pair.

Figure 1 (left column) shows same trials constructed at similarity levels 90, 78, 66, 54, 42, 30, 18, and 0 (a mature same trial). The right column shows different trials constructed at the same levels. Depending on the algorithm's random choices in producing variants, and the configural shape of the prototypes involved, the perceptual cue could be slightly stronger or weaker from pair to pair at the same Level. Nonetheless, the degree of perceptual similarity between the sample and the relational match, and the strength of the perceptual cue available to the participant, did weaken systematically as the task progressed, Level decreased, and Move increased.

**Instructions**—Participants were given minimal instructions for the RMTS task, leaving them—like the macaques—to mainly construe the task themselves. Participants were told: “On each trial, you will see a pair of shapes at the top of the screen. To respond, use the left arrow key to pick the left pair of shapes. Use the right arrow key to pick the right pair of shapes. If you are correct, you will hear a WHOOP and GAIN 1 point. If you are INCORRECT, you will hear a BUZZ and LOSE 1 point.” We point out that these instructions give humans some performance advantage over macaques, who can be given none. Doubtless humans also self-instruct in tasks, possibly providing them additional

performance advantages, including the availability of explicit verbalizable rules that stabilize and optimize performance.

**Task progression**—In Experiment 1A, the decrease of Level, and the weakening of the supporting perceptual cue, proceeded as follows. Every 11 trials, the controlling program assessed ongoing performance. If the participant was  $\geq 75\%$  correct on the previous 20 trials, Level decreased by 1 step, weakening slightly the perceptual-similarity cue. If performance was below 75%, Level increased by one step, strengthening the perceptual-similarity cue slightly.

In Experiment 1B, the decrease of Level, and the weakening of the supporting perceptual cue, proceeded as follows. Every 4 trials, the controlling program assessed ongoing performance. If the participant was  $\geq 75\%$  correct on the previous 20 trials, Level decreased by 1 step. If performance was below 75%, Level increased by one step. Thus, the ramp of decreasing perceptual similarity was steeper and went faster. The purpose of this speed was to leave time in Experiment 1B for the final phase of mature trials. When similarity level fell below 26, only mature same and different trials were presented. This phase evaluated whether humans' RMTS performance would transcend the final removal of all perceptual support.

In Experiment 1C, 30 participants received the task progression just described for Experiment 1B. Relative to those participants, we were interested in evaluating humans' ability to solve the RMTS when never given any perceptual support from the beginning of training. Therefore, twenty-nine participants in this experiment received only mature, zero-similarity trials from the task's beginning and throughout.

### Results: Experiment 1A

The 10 humans completed 4,907 RMTS trials—about 490 per participant. There were 2,394 same trials and 2,513 different trials. Humans were .991 correct overall—a strong and accurate RMTS performance with only rare motoric perseverations or transposed response mappings that produced occasional random errors.

Figure 2 (two upper left panels) shows the progression of humans' RMTS performance in 10-trial blocks. Humans were constantly meeting criterion, drawing similarity level downward so that the strength of support offered by the perceptual-similarity cue waned. Humans' proportion correct was never dented by the removal of this perceptual support. Judging by our experience in this task, the perceptual cue becomes too weak to use reliably (by us!) as the similarity level falls through the 60's into the 50s. Over this range, cognitive control is spontaneously handed off to the conceptual or relational cue that will thereafter control performance. Humans manage this cognitive transition easily. Humans did make a large proportion of their errors in this range of the task, though the absolute number of errors was very small even here. These errors may have reflected the conceptual transition in the task, but the reflection is faint because humans are so seamless.

### Results: Experiment 1B

The 19 humans completed 8,685 RMTS trials—about 457 per participant. There were 4,354 same trials and 4,331 different trials. Humans were .974 correct overall, a highly accurate RMTS performance again.

Figure 2 (bottom-left panels) shows the progression of humans' RMTS performance in 10-trial blocks. Again humans steadily met criterion, drawing Similarity Level downward and weakening the support offered by the perceptual-similarity cue. Again humans' proportion



correct was never threatened by the removal of this support. At Trial Block 24, humans (with some slight misalignments based on how quickly they met criterion through the training trials) transitioned to the last phase of the task in which they received only trials that provided no perceptual-similarity support. These trial blocks are graphed in red symbols. There was no performance reduction during the task's conceptual transition.

### Results: Experiment 1C

Thirty humans in the transition condition completed 15,696 RMTS trials—about 523 per participant. There were 7,927 same trials and 7,769 different trials. Humans were .945 correct overall, a highly accurate RMTS performance again.

Figure 2 (top-right panels) shows the progression of these humans' RMTS performance in 10-trial blocks. Similarity Level steadily declined as humans kept meeting performance criteria. Near Block 25, the onset of the red symbols shows that 24 participants seamlessly made the conceptual transition in the task, maintaining nearly perfect performance. The continuing black symbols show that 6 participants did not immediately make that transition. Several were struggling to progress with the task with the similarity level of perceptual support set in the 50s. After they transitioned, several more participants still struggled to progress with the similarity level set in the 60s or 70s. In the end, all participants did complete the task's conceptual transition.

The 29 humans in the zero-similarity condition of this experiment completed 15,454 RMTS trials—about 533 per participant. There were 7,724 same trials and 7,730 different trials. Humans were .929 correct overall, still an accurate RMTS performance.

Figure 2 (bottom-right panels) shows humans' RMTS performance in 10-trial blocks. Humans, facing the zero-similarity version of the task at its outset, did show an initial period of lower performance as they came to construe the task appropriately. This lasted some tens of trials depending on the participant. From then on RMTS performance was highly accurate.

We also plotted a backwards learning curve for the zero-similarity condition. We aligned the trial blocks at which participants reached and sustained .8 correct in the RMTS task. We then examined performance back from that point to show the path by which humans solved the task. Figure 3 shows participants' proportion correct relative to Block 0 which represents the aligned criterion block. In that aligned criterion block, humans were .915 correct. In the immediately preceding block, they were .543 correct. This striking change in performance over just 10 trials strongly suggests that humans gained control of the mature, conceptual version of the RMTS task through an explicit realization of the task's relational rule. This could provide an important clue in understanding species differences in relational matching (see *Discussion*). Probably humans in Experiments 1A and 1B underwent a similar sudden realization as the perceptual cue faded, but this realization cannot be seen clearly because task control was handed off between two strategies that produced very high performance.

## Experiment 2: Macaques

### Method

**Participants**—Male rhesus monkeys (*Macaca mulatta*) Hank (27 years old), Obi (7 years old), Gale (27 years old), Murph (17 years old), Han (8 years old), and Chewie (11 years old) were tested. They had been trained using procedures described elsewhere (Rumbaugh, Richardson, Washburn, Savage-Rumbaugh, & Hopkins, 1989; Washburn & Rumbaugh, 1992) to respond to computer-graphic stimuli by manipulating a joystick. They had had experience in a large variety of cognitive tests prior to this experiment, including tests of

quantity judgment (e.g., Beran, 2008), what-where-when memory (e.g., Hoffman, Beran, & Washburn, 2009), implicit and explicit category learning (e.g., Smith, Beran, Crossley, Boomer, & Ashby, 2010), and metacognition (e.g., Beran & Smith, 2011; Smith, Redford, Beran, & Washburn, 2010). As a group, these monkeys consistently show high levels of performance in these kinds of tests. Overall, these monkeys had failed to show robust RMTS performance in other tasks (e.g., Flemming et al., 2007, 2011), while showing, as discussed in the introduction, some hints of a relational-matching capacity that could be fostered. All in all, they seemed to be excellent candidates for testing the perceptual-support hypothesis.

Macaques were tested in their home cages at the Language Research Center of Georgia State University, with *ad lib* access to the test apparatus, working or resting as they chose during long sessions. The animals were neither food deprived nor weight reduced for the purposes of testing, and they had continuous access to water.

**Apparatus**—The monkeys were tested using the Language Research Center’s Computerized Test System (Washburn & Rumbaugh, 1992), comprising a computer, a digital joystick, a color monitor, and a pellet dispenser. Monkeys manipulated the joystick through the mesh of their home cages, producing isomorphic movements of a computer-graphic cursor on the screen. Contacting appropriate computer-generated stimuli with the cursor brought them a 94-mg fruit-flavored chow pellet (Bio-Serve, Frenchtown, NJ) using a Gerbrands 5120 dispenser interfaced to the computer through a relay box and output board (PIO-12 and ERA-01; Keithley Instruments, Cleveland, OH). Correct responses were accompanied by a computer-generated whooping sound that bridged the monkeys to their reward. On incorrect responses, the screen froze with the wrong response visible, and there was a computer-generated buzzing sound and a trial-less timeout period. The monkeys were all highly experienced with the general methods and response modality of computer-based, joystick-controlled tasks.

**RMTS trials**—Each RMTS trial consisted of two polygons presented side by side in the center-top of an 11.5-inch computer screen against a black background. On some trials, the two polygons were identical. On different trials, the two polygons were very different, having been created based on two different random selections of points from within the 30 × 30 stimulus grid. A cursor controlled by the animals’ joysticks was illuminated below the sample pair of shapes. In all phases of testing, the monkeys moved the cursor up to the pair of sample shapes, confirming their readiness to perform the trial and hopefully garnering some visual contact with the sample pair of shapes. Subsequently, two more shape pairs—a same pair and a different pair—were presented at the screen’s bottom-left and bottom-right. The placement of the same or different pair to left or right was decided randomly on each trial. A cursor controlled by the animals’ joysticks was illuminated between the two response pairs. As the top pair of shapes was same or different, respectively, the animal would move the cursor to touch the pair on the bottom that was also same or different. The program tracked constantly the monkeys’ proportion correct across the most recent 20 trials. This allowed aspects of the program to be changed dynamically as described now.

At the beginning of testing, the prevailing similarity level was set high, so that the animals had a strong perceptual cue supporting their correct choice of the same or different pair. For Hank, Obi, and Murph, that level was set at 78 (it was set at 80 for humans in Experiments 1A and 1B). For the rest of the subjects, that level was set at 97. Periodically, the program software evaluated whether recent performance was high enough so that the similarity level could be reduced, weakening the perceptual support, or low enough so that the similarity level should *be* increased, strengthening the perceptual support. In this way, animals found their own performance space at some level of the perceptual-support continuum. From session to session, in all phases of testing, the starting similarity level for a session was

always set 2 steps higher than the ending level for the previous session, to give animals a brief warm-up in the task with slightly more perceptually supportive trials.

The macaques experienced eight variants of the RMTS program as we describe now. However, as indicated by the results that follow, none of these variants differentially affected the animals' performances.

We explored a variety of criteria for when to decrease the similarity support on good performance and for when to increase the similarity support on bad performance. In version 1 of the program, every 11 trials, the program software evaluated whether recent performance was  $\geq .75$  over the last 20 trials, and if so it reduced the similarity level, weakening the perceptual support. Also, every 11 trials, it evaluated whether recent performance was  $< .75$ , and if so it increased the similarity level, strengthening the perceptual support. One may call this set of criteria {11, .75, 11, .75}. In version 2 of the program, we used the set of criteria {10, .75, 30, .65}. This program pressed slightly harder for animals to move down the progression of similarity levels, because it evaluated more seldom (every 30 trials) whether it should ease the task, and the animal had to perform worse ( $< .65$ ) for this easing to occur. In version 3 of the program, we used the criteria {10, .70, 30, .60}. In versions 4–5 of the program, we used the criteria {5, .60, 30, .60}. In version 6–8 of the program, we used the criteria {5, .60, 20, .60}.

Beginning with version 3 of the program, we infused zero-similarity trials in among the trials presented at the prevailing level of similarity support. This also occurred dynamically based on monkeys' performance. In version 3 and versions 6–8, at similarity levels  $> 50$ , all trials were at the prevailing similarity level. At similarity level 50, 49, 48, 47, and so forth, 90%, 88%, 86%, and 84% of trials were at the prevailing similarity level but the rest of the trials were at similarity level 0. The more that the monkeys progressed by bringing similarity levels lower, the more zero-similarity trials they received. In version 4 of the program, we gave 100% zero-similarity trials to Hank and Obi to force them to have extensive experience with that kind of trial. In version 5, at similarity levels  $> 50$ , all trials were at the prevailing similarity level. At similarity levels 50, 49, 48, 47, and so forth, 90%, 87%, 84%, and 81% of trials were at the prevailing similarity level but the rest of the trials were at similarity level=0.

In versions 5–8 of the program, we tried to spur macaques on to careful effort and attention on the zero-similarity trials that gave them no perceptual support at all.

Therefore, they received double rewards (auditory whoops and food pellets) for correct responses, and doubly long timeouts for incorrect responses.

In versions 7 and 8 of the program only, the three pairs of shapes on the screen were enclosed in boxes. The goal of this was to help the monkeys treat the stimulus pairs more configurally, and as unitary objects to be matched, as suggested by the spatial proximity results in Fagot and Parron (2010).

These variations in procedure were applied as we realized that monkeys were not able to smoothly progress downward toward low levels of perceptual support. Accordingly, we tried to find the optimal downward pressure on the perceptual-support level that might foster their progress, and the optimal mixture of perceptually supported trials and mature, zero-similarity trials that might help their transition to true relational matching.

**Pilot study: direct onset of mature RMTS trials**—The present study considers a potentially important new methodology for fostering relational cognition in macaques. It would have been inappropriate to first run the six macaques with mature RMTS trials before

the fostering manipulation, because this could easily have undermined the fostering capacity of perceptual support. Therefore, to ask whether our RMTS task was especially learnable, or whether the training history of LRC's macaques made them somehow RMTS adept, we gave one of our strongest performing macaques in psychophysical tasks (Lou, 17 years old) the RMTS task with zero-similarity trials from the outset, with no perceptual support ever provided. Lou had not been available at the outset of the main experiment because he was involved in other projects, and so his availability later allowed us to pose to him the conceptual RMTS task at similarity level = 0. Over more than 20,000 trials (21,305 trials), Lou was unable to master the RMTS task. Over his last 15 sessions, (about 8,000 trials), his proportion correct was .497. This finding replicates, and is supported by, the findings from Flemming et al. (2007, Experiment 3) in which Murph, Lou, Hank and Gale all failed to master a different RMTS task over more than 10,000 trials. This suggests that our macaques and our RMTS task were generally suitable for exploring the new fostering manipulation here. It also speaks against the possibility that something in our perceptual-support manipulation somehow blocked the RMTS learning that might have occurred otherwise. The data from the human studies also speak against that possibility.

## Results

**Hank**—Hank completed 115,250 trials in the RMTS task. He was .599 correct overall on the trials that were presented at his prevailing similarity level of perceptual support. He was .497 correct overall on the mature trials that offered him no perceptual support.

Figure 4 shows his data through the whole experiment by 250-trial blocks. Hank participated in all eight versions of the RMTS task. He transitioned to versions 2, 3, 4, 5, 6, 7, and 8 at trial blocks 62, 122, 270, 375, 386, 414, and 434. Here and in all figures showing macaque data task transitions are shown by an alternation between filled and open black symbols. In addition, Hank's participation in version 4 of the RMTS task is shown by the absence of any black symbols (i.e., the absence of any trials at his prevailing similarity level while he received only mature trials). Hank's first two transitions are sharply marked in his similarity levels, which immediately fell when we adopted criteria that let the similarity level fall more freely.

We can provide a statistical yardstick to Figure 4 for those preferring this approach to the data. At Hank's levels of performance of .5 to .7 through most of the experiment, if he completed 4, 8, 16, or 32 250-trial blocks, or 1000, 2000, 4000, or 8000 trials, the .01 confidence interval around the observed proportion correct, to include the real population value of his performance 99% of the time, would be  $\pm 0.04$ ,  $\pm 0.03$ ,  $\pm 0.02$ , and  $\pm 0.01$ , respectively. This confidence-interval yardstick can be equivalently applied to the performance of all six monkeys.

So as not to minimize what Hank accomplished in this experiment, we point out that Hank sustained correct performance at .600 for many thousands of trials at similarity levels near 40. By the statistical yardstick, this performance was above chance.

However, Hank's performance on mature trials leaves open the possibility that he was receiving some perceptual support from these trials near similarity level 40. That humans transitioned in Experiment 1 to relational matching at similarity levels near 60 does not mean that no residual perceptual support remains below that level. Macaques, less facile with relational cues that are less salient to them, might be far more dependent on any available perceptual support and might hang in with the perceptual strategy to essentially the bitter end, whereas humans might comfortably transition earlier given a salient relational cue that starts to trump the weakening perceptual cue. Hank would have only needed to find some perceptual support on 2 in 10 trials to achieve the level of .600 correct he actually did.

The addition of zero-similarity trials is marked by the red symbols beginning at trial block 122 in the figure. These trials were only occasional, and therefore the estimates of performance in each trial block are somewhat variable being based on fewer trials. Those estimates sharpened during Hank's performance on version 4 of the RMTS task that presented only mature trials. Near the end of the experiment, we gave Hank several additional experiences with similarity-level trials and mature trials intermixed, in the hope that he would be able to extend some of his limited success with similarity-supported trials to trials that provided no perceptual support. He did not do so. He was at chance on mature trials throughout the experiment.

**Obi**—Figure 5 summarizes 39,000 trials—in 250-trial blocks—by macaque Obi in the RMTS task. Obi was .610 correct overall on the trials that were presented at his prevailing similarity level of perceptual support. He was .498 correct overall on mature trials that offered him no perceptual support.

Obi participated in all eight versions of the RMTS task. He transitioned to versions 2, 3, 4, 5, 6, 7, and 8 at trial blocks 32, 55, 97, 116, 118, 133, and 144. The first two transitions are sharply marked in his similarity levels, which immediately fell from 70 to 50 and from 50 to 40 when we adopted criteria that let the similarity level fall more freely. Note that Obi also sustained correct performance at .600 for thousands of trials at similarity levels between 40 and 50. He was able to perform somewhat successfully with only minimal perceptual support provided by those low similarity levels

However, Obi, too, may still have been relying on some perceptual support. The addition of mature trials is marked by the red symbols beginning at trial block 55. These trials were only occasional, and therefore the estimates of performance in each trial block are somewhat variable being based on fewer trials. Those estimates sharpened during Obi's performance on version 4 that presented only mature trials and they centered on the .5 or chance level. Near the end of the experiment, we gave Obi several additional experiences with similarity-level trials and mature trials intermixed, to see if he would be able to extend some of his limited success with similarity-supported trials to trials that offered no perceptual support. Like Hank, he did not do so.

**Gale**—Figure 6 summarizes Gale's 39,000 trials. Gale was .580 correct overall on the trials that were presented at his prevailing similarity level of perceptual support. He was .467 correct overall on mature trials that offered him no perceptual support. Gale participated in versions 1, 2, 3, 6, 7, and 8 of the RMTS task. He transitioned to versions 2, 3, 6, 7, and 8 at trial blocks 21, 40, 132, 141, and 152. Gale also performed above chance for many thousands of trials at similarity levels near 50 that offered only weak perceptual support. Nonetheless, Gale's chance performance on mature trials—beginning at trial block 55—indicates that he was never able to successfully transfer that performance to trials that offered no perceptual support.

**Murph**—Figure 7 summarizes Murph's 19,500 trials. He was .621 correct overall on the trials that were presented at his prevailing similarity level of perceptual support. He was .480 correct overall on mature trials that offered him no perceptual support. Murph participated in versions 1, 2, and 3 of the RMTS task. He transitioned to versions 2 and 3 trial blocks 19 and 34. Like his peers, he was able to perform at about 60% correct on trials with similarity levels in the range of 40 to 50, but he was never able to translate that performance into any success on the mature trials that offered no perceptual support.

**Han**—Figure 8 summarizes Han's 23,000 trials. Han was .580 correct overall on the trials presented at his prevailing similarity level. He was .462 correct overall on mature trials that

offered no perceptual support. Han participated in versions 1, 2, and 3 of the RMTS task. He transitioned to versions 2 and 3 at trial blocks 3 and 16. As Murph did, Han performed above chance for thousands of trials near similarity level 50 (black symbols). Yet he also performed throughout at chance on mature trials (red symbols).

**Chewie**—Figure 9 summarizes Chewie’s 28,000 trials. Chewie—our weakest performer—was .574 correct overall on the trials that were presented at his prevailing similarity level (black symbols). He was .444 correct on mature trials (red symbols). Chewie participated in versions 1, 2, and 3 of the RMTS task. He transitioned to versions 2 and 3 at trial blocks 12 and 28. He struggled to be above chance even on trials in the range of similarity level 60. He also had no success on the mature trials.

## General Discussion

In Experiments 1A-1C, humans predictably showed a strong RMTS performance. Somewhere in the 50–70 range of similarity levels, humans relinquished their use of the perceptual matching cue and began to match relationally. They did so seamlessly, with no drop in nearly perfect performance. Their performance remained nearly perfect (Experiments 1B, 1C) for zero-similarity RMTS trials that offered no perceptual support. In Experiment 1C, humans were slower to learn the RMTS task given zero-similarity trials from the outset. Contrary to a perceptual-blocking hypothesis, the perceptual support did not confuse or mislead them: it evidently facilitated their entry into the RMTS task. Finally, Experiment 1C indicated that humans struck the RMTS solution instantaneously, signaling that they experienced an explicit realization of the task’s solution. The sudden transition, and its explicit-cognition implications, may be important clues in understanding species differences in relational cognition.

In sharp contrast, six macaques showed no evidence of using a relational-matching capacity that could be applied to zero-similarity trials, even over 260,000 trials during which we systematically tried to foster that capacity. They were persistently at chance on those trials. They did show some ability to perform as the similarity dropped lower and the perceptual support weakened. However, they may still have been picking up on some perceptual resonance at similarity levels 40–50, and their performance on the zero-similarity trials supports this interpretation. It was a strong feature of our paradigm that it gradually led macaques as far as they could go toward the weakening of the perceptual cue and toward the use of true analogical processing. They showed clear limits in making this progression.

The macaques’ perceptual dependence at lower similarity levels is not contradictory with our human findings. Humans may transition at higher similarity levels than 40–50 because the perceptual cues become difficult to use and because humans naturally use relational hypotheses. The tradeoff point for them from perceptual to conceptual comes earlier for them because the relational solution is relatively more salient. Macaques probably struggle onward to lower similarity levels, still using whatever shred of perceptual support remains, because really they almost have no relational-matching capacity to swap in to replace the perceptual strategy. This framework accounts well for all the human and animal data, for humans’ higher point of transition, and for humans’ sudden performance transformation in Experiment 1C.

Clearly, the perceptual-support methodology that granted humans a seamless entry into the RMTS task did not do so for macaques. This does not invalidate the claims in Goldstone and Barsalou (1998) about the connections between perceptual and conceptual systems in human cognition. To the contrary, the results may suggest that those connections are viable and visible given robust perceptual and conceptual systems, the latter possessed by humans but less so by macaques.

Therefore, our results join other findings (Fagot & Parron, 2010; Fagot et al., 2001; Flemming et al., 2007; Flemming, 2011) confirming that there is a pronounced species difference between humans and monkeys in their capacity to bring conceptual relations to bear on relational-matching tasks. As here, macaques and baboons have generally proved largely dependent on ancillary associative cues and learning mechanisms involving visual entropy, spatial grouping affording local processing, differential outcomes, and so forth, instead of showing full-fledged analogical processing in RMTS tasks. Even in Fagot and Thompson (2011), only a small proportion of their large subject group had success in the relational task, even after dogged training that spanned tens of thousands of trials in a task that might also have allowed some associative learning mechanisms to prime analogical performance because the original training set of stimuli was small and restricted. Overall, the empirical literature reveals a strong human-monkey disparity in the character of RMTS performance. However, additional research in this area is welcome and warranted, for the exact nature and sharpness of species differences in this capacity are still being debated, mapped, measured and defined.

In particular, there is ongoing debate on whether species disparities in relational cognition should be considered qualitative and discontinuous. Thompson and Oden (2000) distinguished the paleological monkey (lacking analogical cognition) from the analogical ape. Penn et al. (2008) placed the qualitative discontinuity higher, between all nonhuman primates and humans. In contrast, other scholars have seen continuity in the RMTS performances of humans and animals (e.g., Castro & Wasserman, 2013; Fagot et al., 2001; Fagot & Maugard, 2013; Flemming et al., 2013; Wasserman et al., 2001).

Our interpretation is mixed. Monkeys can perform an RMTS task based on arrays of items, perhaps basing their relational cognition on the informational variability and visual entropy present in the arrays. The use of variability to answer RMTS trials is not a confound. Variability and change are essential to the definition and psychological appreciation of difference, and therefore the use of variability could be foundational to building a capacity for abstract conceptual matching. Macaques may possess that foundation and thus share some continuities with humans' relational cognition. Moreover, Flemming et al. (2013) showed that the array-matching performances of monkeys do go beyond the use of the cues of entropy and perceptual variability.

However, humans may also bring a different set of processes and cognitive systems to bear on RMTS problems that macaques can barely recruit because their corresponding processes and systems are nascent. In particular, Experiment 1C provided evidence of humans' use of an explicit, rule-discovery process in solving the RMTS task (Figure 3). Smith Redford, Haas, Coutinho, and Couchman (2008) also found that humans' Same-Different concept was qualitative, categorical, and rule-based.

There could be diverse explanations for this qualitative aspect of the monkey-human relational-cognition species difference. These were elegantly surveyed in Penn et al. (2008) and the accompanying commentaries. Humans have more strongly developed prefrontal cortical systems that could afford the hypothesis generation and testing that would underlie RMTS rule discovery (Semendeferi, Lu, & Schenker, 2002). Categorization research has shown that macaques are impoverished relative to humans in their learning of category rules (Smith, Minda, & Washburn, 2004). A related hypothesis is that humans' working memory could be more dynamic and agile, affording the flexible evaluation of conceptual relations (Halford, Phillips, & Wilson, 2008). Humans have language and even a language of relations that could be another affordance for relational cognition (Bickerton, 2008; Gentner & Christie, 2008). This affordance could explain why, in the main, it has been the language- and symbol-trained chimpanzees that have most clearly demonstrated RMTS competence

(Gilliam et al., 1981, Premack, 1976, 1986; see also Thompson and Oden, 1996). This could also help explain why humans hit on the RMTS solution in these tasks so qualitatively suddenly (Figure 3 above). It is consistent with the sudden realization that a hypothesis is confirmed and will afford nearly perfect performance thereafter. Further research will be required to establish the extent to which monkeys can show these kinds of rule-based relational performances. A recent, promising first step was taken by Fagot and Maugard (2013). There, using an interesting multidimensional Same-Different task, they showed that macaques could in a sense report after the fact the sameness or differentness of a pair of multi-dimensional stimuli *along some particular cued dimension*.

Humans' relational aptitude could also be partly engendered from external causes—the character of humans' cognitive environment (Barrett, 2008), the character of human symbol systems (Lupyan, 2008), and so forth. For example, relational comparisons are an important feature of many aspects of humans' educational system, from the first felt boards of preschool programs through the Miller Analogies Test. Therefore, we must remember that humans are relationally focused not only by the pure strength of their cognitive systems, but also from their education and symbolic language training. This could also explain their RMTS adeptness. Conversely, macaques could be perceptually focused not only because of weakness in their conceptual system, but also because of ecological consideration that favored—for fitness's sake—the use of perceptual processes, not analogical processes, in cognition. That is, monkeys in their ecology deal constantly with primary perceptual similarity regarding their foraging and predator-avoidance categories: that eagle is like that other eagle, Careful!; that termite is like those other termites, Eat!. However, they almost never encounter natural RMTS problems: the two eagles are the same, now find the pair of terrestrial mammals that share the relation of sameness.

It is difficult to know whether human educational systems favor relational training because humans appreciate relations so well, or whether humans appreciate relations so well because our educational systems favor relational training. The perceptual-support paradigm might help resolve this issue. Any human population, with or without formal schooling, would be able to enter our task and perform it with a robust perceptual strategy. But then one could ask about the transition points to the relational strategy of different human populations that had received different amounts of formal Western schooling. In this way one might illuminate the relational character of human minds less biased by formal education.

The present RMTS paradigm also has promise to let researchers study the cognitive development of the relational-matching capacity. Our paradigm is instantiable without any instructions. The initial perceptual-strategy would be readily apparent to children at all ages as it is to macaques. Therefore, one could engage children with the task using the perceptual support, and then study their spontaneous capacity at different ages to make the transition away from the use of perceptual similarity toward the use of true conceptual relations. Children who had a stronger sense of conceptual relations would make the jump to relational matching earlier. Children who had a weaker relational capacity would depend on the waning perceptual cue for longer, and might even, as macaques do, fail to make the transition. Our RMTS task is constructive because it sets the table for a transition to conceptual relations, without demanding when that transition occurs. Children would make the determination of their transition point spontaneously for themselves. And, in that way, they would report the relative salience for them of the perceptual and relational cues in the task.

Similarly, the perceptual-support paradigm could help explain why some chimpanzees have succeeded with RMTS paradigms. Language-naïve and symbol-naïve chimpanzees would find the perceptual-support RMTS task perfectly natural to engage. One could then ask



about their ability to make the transition to mature relational matching. Thus one might address the issue of whether chimpanzees are inherently analogical compared to macaques, or whether it is more the case that symbol- and language-trained chimpanzees are analogical.

Finally, the perceptual-support paradigm might let researchers study the cognitive organization of humans' RMTS performances. In its perceptual-support phase, the task could be performed even given a concurrent cognitive load that used up humans' working memory resources or occupied their executive attention. However, the concurrent load might delay or disrupt humans' transition to the mature relational strategy, if that transition is somehow dependent on the executive level of cognition. This disruption would be a confirmation that the relational strategy in the task is adjudicated at a higher, more executive and perhaps more explicit cognitive level. And of course this would have important implications for theoretical development in the comparative study of relational matching, because it might feed back to help explain in what processes and at what cognitive level macaques are challenged as they enter RMTS tasks.

## Acknowledgments

The preparation of this article was supported by Grant HD-060563.

## References

- Barrett L. Out of their heads: Turning relational reinterpretation inside out. *Behavioral and Brain Sciences*. 2008; 31:130–131.
- Beran MJ. Monkeys (*Macaca mulatta* and *Cebus apella*) track, enumerate, and compare multiple sets of moving items. *Journal of Experimental Psychology: Animal Behavior Processes*. 2008; 34:63–74. [PubMed: 18248115]
- Beran MJ, Smith JD. Information seeking by rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *Cognition*. 2011; 120:90–105. [PubMed: 21459372]
- Bickerton D. Darwin's last word: How words changed cognition. *Behavioral and Brain Sciences*. 2008; 31:132.
- Brooks DI, Wasserman EA. Same/different discrimination learning with trial-unique stimuli. *Psychonomic Bulletin & Review*. 2008; 15:644–650. [PubMed: 18567268]
- Burdyn LE, Thomas RK. Conditional discrimination with conceptual simultaneous and successive cues in the squirrel monkey (*Saimiri sciureus*). *Journal of Comparative Psychology*. 1984; 98:405–413. [PubMed: 6210175]
- Carter DE, Werner TJ. Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior*. 1978; 29:565–601. [PubMed: 16812079]
- Castro L, Wasserman EA. Humans deploy diverse strategies in learning same-different discrimination tasks. *Behavioural Processes*. 2013; 93:125–139. [PubMed: 23073499]
- Catrambone R. The effects of surface and structural feature matches on the access of story analogs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2002; 28:318–334.
- Cook RG. The structure of pigeon multiple-class same-different learning. *Journal of the Experimental Analysis of Behavior*. 2002; 78:345–364. [PubMed: 12507008]
- Cook RG, Smith JD. Stages of abstraction and exemplar memorization in pigeons' category learning. *Psychological Science*. 2006; 17:1059–1067. [PubMed: 17201788]
- Cook RG, Wasserman EA. Learning and transfer of relational matching-to-sample by pigeons. *Psychonomic Bulletin & Review*. 2007; 14:1107–1114. [PubMed: 18229483]
- Cumming WW, Berryman R. Some data on matching behavior in the pigeon. *Journal of the Experimental Analysis of Behavior*. 1961; 4:281–284. [PubMed: 13718829]
- D'Amato MR, Colombo M. On the limits of the matching concept in monkeys (*Cebus apella*). *Journal of the Experimental Analysis of Behavior*. 1989; 52:225–236. [PubMed: 2584914]

- D'Amato MR, Salmon DP, Columbo M. Extent and limits of the matching concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*. 1985; 11:35–51. [PubMed: 3989475]
- Dunbar, K. The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In: Gentner, D.; Holyoak, KJ.; Kokinov, BN., editors. *The analogical mind: Perspectives from cognitive science*. Cambridge, Massachusetts: The MIT Press; 2001. p. 313-334.
- Fagot J, Deruelle C. Processing of global and local visual information and hemispheric specialization in humans (*Homo sapiens*) and baboons (*Papio papio*). *Journal of Experimental Psychology: Human Perception and Performance*. 1997; 23:429–442. [PubMed: 9104003]
- Fagot J, Maugard A. Analogical reasoning in baboons (*Papio papio*). Flexible reencoding of the source relation depending on the target relation. *Learning & Behavior*. 2013
- Fagot J, Parron C. Relational matching in baboons (*Papio papio*) with reduced grouping requirements. *Journal of Experimental Psychology: Animal Behavior Processes*. 2010; 36:184–193. [PubMed: 20384399]
- Fagot J, Thompson RKR. Generalized relational matching by guinea baboons (*Papio papio*) in two-by-two-item analogy problems. *Psychological Science*. 2012; 22:1304–1309. [PubMed: 21934135]
- Fagot J, Wasserman EA, Young ME. 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*. 2001; 27:316–328. [PubMed: 11676083]
- Farthing GW, Opuda MJ. Transfer of matching-to-sample in pigeons. *Journal of the Experimental Analysis of Behavior*. 1974; 21:199–213. [PubMed: 16811738]
- Flemming TM, Beran MJ, Washburn DA. Disconnect in concept learning by rhesus monkeys: Judgment of relations and relations-between-relations. *Journal of Experimental Psychology: Animal Behavior Processes*. 2007; 33:55–63. [PubMed: 17227195]
- Flemming TM. Conceptual thresholds for *same* and *different* in *old* (*Macaca mulatta*) and new-world (*Cebus apella*) monkeys. *Behavioural Processes*. 2011; 86:316–322. [PubMed: 21238555]
- Flemming TM, Thompson RKR, Beran MJ, Washburn DA. 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*. 2011; 37:353–360. [PubMed: 21574737]
- Flemming TM, Thompson RKR, Fagot J. Baboons, like humans, solve analogy by categorical abstraction of relations. *Animal Cognition*. 2013; 16:519–524. [PubMed: 23334572]
- Fujita K. An analysis of stimulus control in two-color matching-to-sample behaviors of Japanese monkeys (*Macaca fuscata*). *Japanese Psychological Research*. 1982; 24:124–135.
- Gentner D, Gentner D, Goldin-Meadow S. Why we're so smart. *Language in mind: Advances in the study of language and thought*. 2003:195–235.
- Gentner D, Christie S. Relational language supports relational cognition in humans and apes. *Behavioral and Brain Sciences*. 2008; 31:136–137.
- Gentner, D.; Rattermann, MJ. Language and the career of similarity. In: Gelman, SA.; Byrnes, JP., editors. *Perspectives on language and thought: Interrelations in development*. Cambridge University Press; 1991. p. 225-77.
- Gentner D, Rattermann MJ, Forbes KD. The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*. 1993; 25:431–467.
- Gibson BM, Wasserman EA. Pigeons learn stimulus identity and stimulus relations when both serve as redundant, relevant cues during same-different discrimination training. *Journal of Experimental Psychology: Animal Behavior Processes*. 2003; 29:84–91. [PubMed: 12561136]
- Gibson BM, Wasserman EA. Time-course of control by specific stimulus features and relational cues during same-different discrimination training. *Learning and Behavior*. 2004; 32:183–189. [PubMed: 15281390]
- Gillan DJ, Premack D, Woodruff G. Reasoning in the chimpanzee: I. Analogical reasoning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1981; 7:1–17.

- Goldstone RL, Barsalou LW. Reuniting perception and conception. *Cognition*. 1998; 65:231–262. [PubMed: 9557384]
- Halford GS, Phillips S, Wilson WH. The missing link: Dynamic, modifiable representations in working memory. *Behavioral and Brain Sciences*. 2008; 31:137–138.
- Herrnstein RJ. Levels of stimulus control: A functional approach. *Cognition*. 1990; 37:133–166. [PubMed: 2269005]
- Hoffman ML, Beran MJ, Washburn DA. Memory for 'what,' 'where,' and 'when' information in rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*. 2009; 35:143–152. [PubMed: 19364223]
- Holmes PW. Transfer of matching performance in pigeons. *Journal of the Experimental Analysis of Behavior*. 1979; 31:103–114. [PubMed: 429955]
- Homa D, Rhoads D, Chambliss D. Evolution of conceptual structure. *Journal of Experimental Psychology: Human Learning and Memory*. 1979; 5:11–23.
- Homa D, Sterling S, Trepel L. Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology : Human Learning and Memory*. 1981; 7:418–439.
- Hummel JE, Holyoak KJ. Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*. 1997; 104:427–66.
- Katz JS, Wright AA, Bachevalier J. Mechanisms of *same/different* abstract-concept learning by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*. 2002; 28:358–368. [PubMed: 12395493]
- Kohler, W. Simple structural functions in the chimpanzee and in the chicken. In: Ellis, WD., editor. *A source book of Gestalt psychology*. London: Routledge & Kegan Paul; 1918/1938. p. 217-227.
- Locke, J. *An essay concerning human understanding*. Philadelphia: Troutman & Hayes; 1690.
- Lupyan G. Taking symbols for granted? Is the discontinuity between human and nonhuman minds the product of external symbol systems? *Behavioral and Brain Sciences*. 2008; 31:140–141.
- Penn DC, Holyoak KJ, Povinelli DJ. Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*. 2008; 31:109–178. [PubMed: 18479531]
- Posner MI, Goldsmith R, Welton KE. Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology*. 1967; 73:28–38. [PubMed: 6047706]
- Premack, D. *Intelligence in ape and man*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1976.
- Premack, D. On the abstractness of human concepts: Why it would be difficult to talk to a pigeon. In: Hulse, SH.; Fowler, H.; Honig, WK., editors. *Cognitive processes in animal behavior*. Hillsdale, NJ: Erlbaum; 1978. p. 423-451.
- Premack, D. Minds with and without language. In: Weiskrantz, L., editor. *Thought without language*. New York: Oxford University Press; 1986. p. 46-65.
- Quine, WV. *Ontological relativity and other essays*. New York: Columbia University Press; 1969.
- Ross BH. This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1987; 13:629–639.
- Rumbaugh, DM.; Pate, JL. The evolution of cognition in primates: A comparative perspective. In: Roitblat, HL.; Bever, TG.; Terrace, HS., editors. *Animal cognition*. Hillsdale, NJ: Erlbaum; 1984. p. 569-585.
- Rumbaugh DM, Richardson WK, Washburn DA, Savage-Rumbaugh ES, Hopkins WD. Rhesus monkeys (*Macaca mulatta*), video tasks, and implications for stimulus-response spatial contiguity. *Journal of Comparative Psychology*. 1989; 103:32–38. [PubMed: 2924530]
- Semendeferi K, Lu A, Schenker N. Humans and great apes share a large frontal cortex. *Nature Neuroscience*. 2002; 5:272–276.
- Shields WE, Smith JD, Washburn DA. Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *Journal of Experimental Psychology: General*. 1997; 126:147–164. [PubMed: 9163934]
- Smith JD, Beran MJ, Crossley M, Boomer J, Ashby FG. Implicit and explicit category learning by macaques (*Macaca mulatta*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes*. 2010; 36:54–65. [PubMed: 20141317]

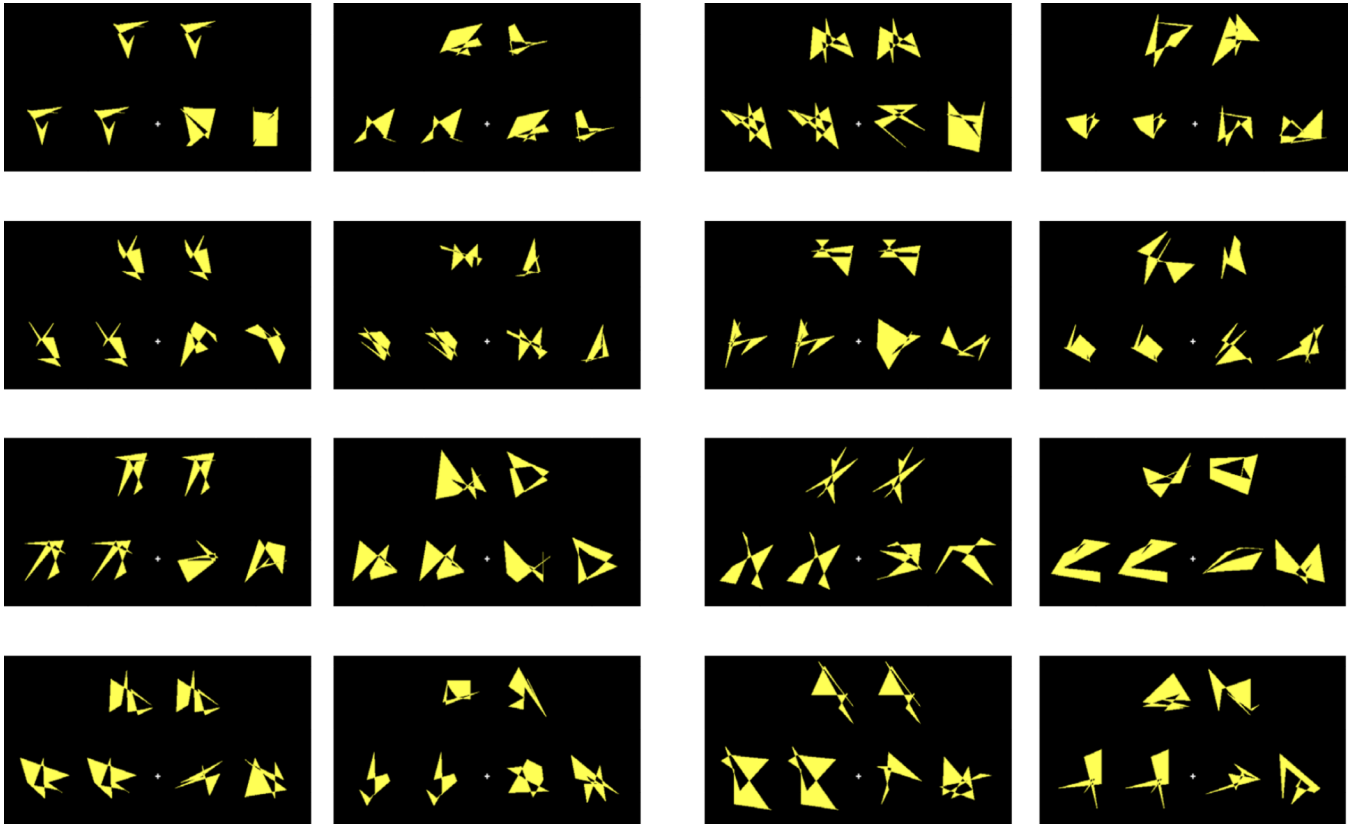
- Smith JD, Minda JP. Distinguishing prototype-based and exemplar-based processes in category learning. *Journal of Experimental Psychology: Learning Memory, and Cognition*. 2002; 28:800–811.
- Smith JD, Minda JP. Journey to the center of the category: The dissociation in amnesia between categorization and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2001; 27:984–1002.
- Smith JD, Minda JP, Washburn DA. Category learning in rhesus monkeys: A study of the Shepard, Hovland, and Jenkins tasks. *Journal of Experimental Psychology: General*. 2004; 133:398–414. [PubMed: 15355146]
- Smith JD, Redford JS, Beran MJ, Washburn DA. Rhesus monkeys (*Macaca mulatta*) adaptively monitor uncertainty while multi-tasking. *Animal Cognition*. 2010; 13:93–101. [PubMed: 19526256]
- Smith JD, Redford JS, Haas SM. Prototype abstraction by monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: General*. 2008; 137:390–401. [PubMed: 18473665]
- Smith JD, Redford JS, Haas SM. The comparative psychology of same-different judgments by humans (*Homo sapiens*) and monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*. 2008; 34:361–374. [PubMed: 18665719]
- Smith JD, Redford JS, Haas SM, Coutinho MVC, Couchman JJ. The comparative psychology of same-different judgments by humans (*Homo sapiens*) and monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*. 2008; 34:361–374. [PubMed: 18665719]
- Spence KW. The differential response in animals to stimuli varying within a single dimension. *Psychological Review*. 1937; 44:430–444.
- Spinozzi G, De Lillo C, Truppa V. Global and local processing of hierarchical visual stimuli in tufted capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*. 2003; 117:15–23. [PubMed: 12735359]
- Thompson RKR, Oden DL. A profound disparity revisited: Perception and judgment of abstract identity relations by chimpanzees, human infants, and monkeys. *Behavioural Processes*. 1996; 35:149–161.
- Thompson RKR, Oden DL. Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*. 2000; 24:363–396.
- Thompson RKR, Oden DL, Boysen ST. Language naïve chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes*. 1997; 23:31–43. [PubMed: 9008861]
- Washburn DA, Rumbaugh DM. Rhesus monkey (*Macaca mulatto*) complex learning skills reassessed. *International Journal of Primatology*. 1991; 12:377–388. [PubMed: 11539108]
- Washburn DA, Rumbaugh DM. Testing primates with joystick-based automated apparatus: Lessons from the Language Research Center's Computerized Test System. *Behavioral Research Methods and Instruments*. 1992; 24:157–164.
- Wasserman EA, Fagot J, Young M. Same-different conceptualization by baboons (*Papio papio*): The role of entropy. *Journal of Comparative Psychology*. 2001; 115:42–52. [PubMed: 11334218]
- Wasserman EA, Young M. Same-different discrimination: the keel and backbone of thought and reasoning. *Journal of Experimental Psychology: Animal Behavior Processes*. 2010; 36:3–22. [PubMed: 20141313]
- Wasserman EA, Young ME, Fagot J. Effects of number of items on the baboon's discrimination of same from different visual displays. *Animal Cognition*. 2001; 4:163–170.
- Wright AA, Cook RG, Kendrick D. Relational and absolute stimulus learning by monkeys in a memory task. *Journal of the Experimental Analysis of Behavior*. 1989; 52:237–248. [PubMed: 2584915]
- Wright AA, Rivera J, Katz JS, Bachevalier J. Abstract concept learning and list-memory processing by capuchin and rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*. 2003; 29:184–198. [PubMed: 12884678]
- Wright AA, Santiago H, Sands S. Monkey memory: Same/ different concept learning, serial probe acquisition, and probe delay effects. *Journal of Experimental Psychology: Animal Behavior Processes*. 1984; 10:513–529. [PubMed: 6491610]

- Wright AA, Shyan MR, Jitsumori M. Auditory same/different concept learning by monkeys. *Animal Learning and Behavior*. 1990; 18:287–294.
- Zentall TR, Hogan DE. Abstract concept learning in the pigeon. *Journal of Experimental Psychology*. 1974; 102:393–398.

### Highlights

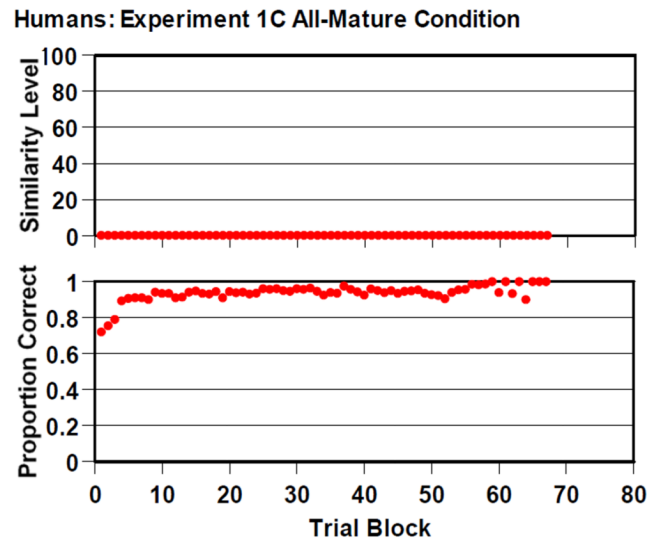
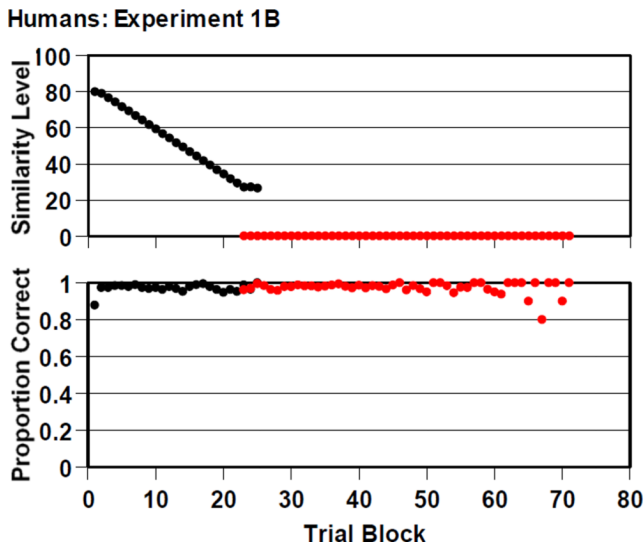
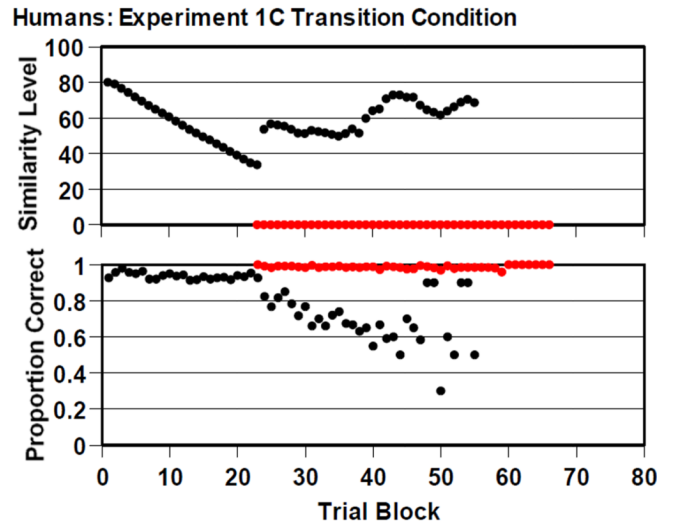
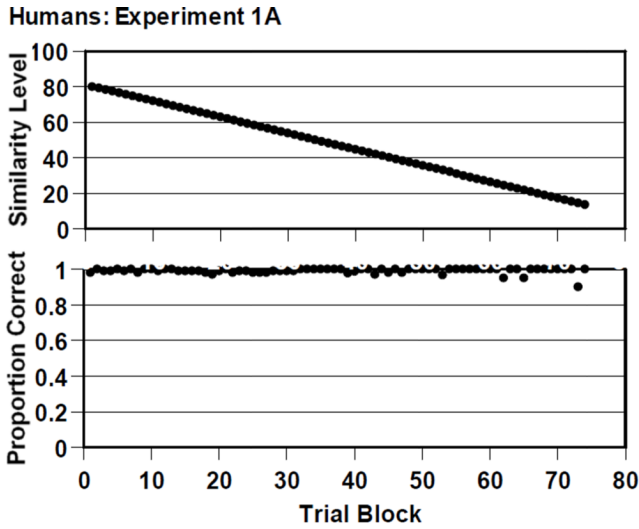
#### **Fading Perceptual Resemblance: A Path for Macaques to Conceptual Matching?**

We introduce a distinctively new approach toward fostering relational matching. Humans flawlessly transition from perceptual to relational matching strategies. We report one of the largest explorations of macaques' relational matching. Six macaques, given more than 260,000 trials, could not make this transition. This striking species difference illuminates the evolution of conceptual abilities.



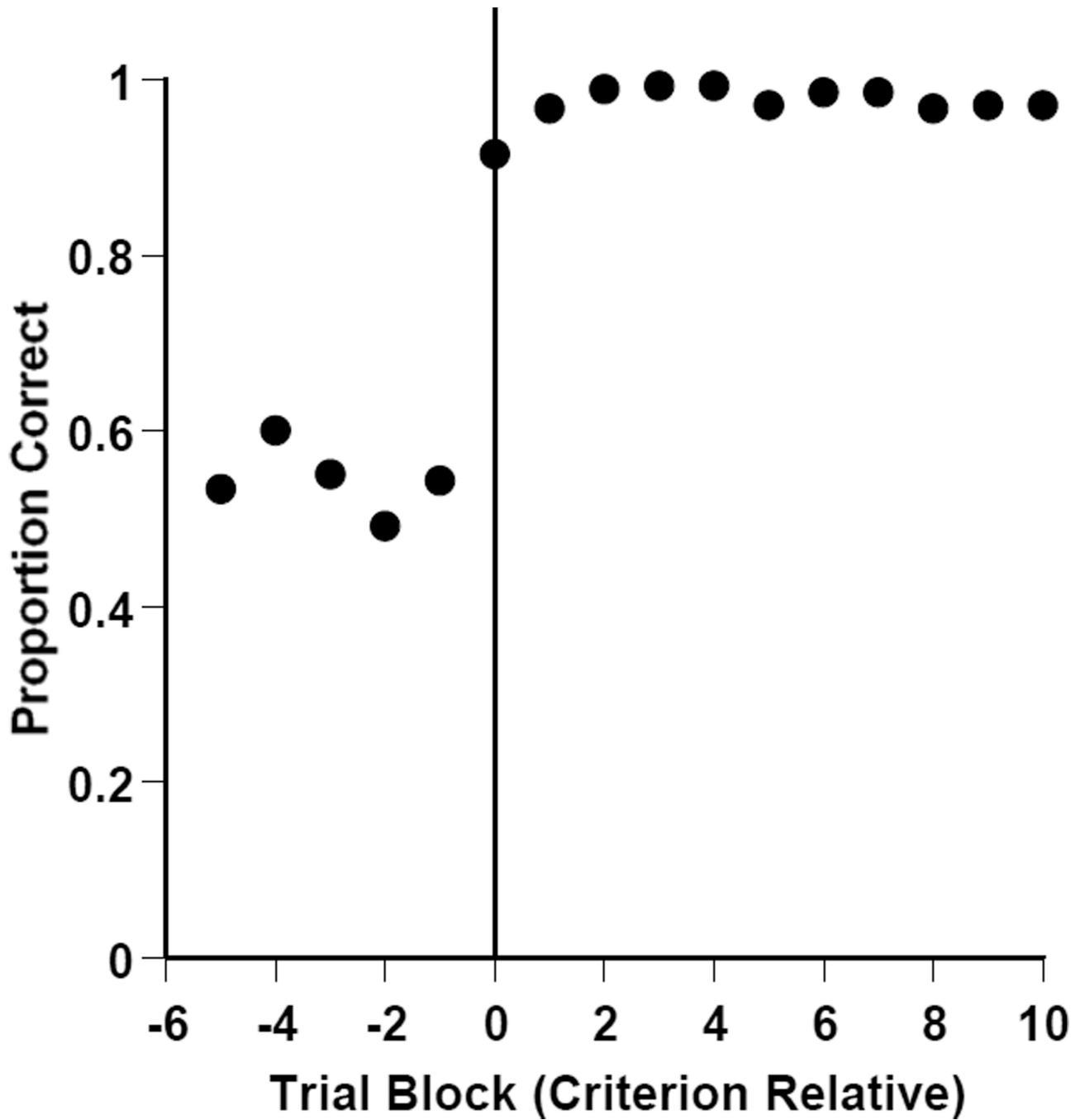
**Figure 1.**

Examples of trials from the relational match-to-sample (RMTS) task. Column 1 (farthest left): same trials, with the level of perceptual support set at 90, 78, 66, and 54. The definition of these levels of perceptual support is given in the text. Column 2: different trials, at the same four levels of perceptual support. Column 3: same trials, with the levels of perceptual support set at 42, 30, 18, and 0. Column 4 (farthest right): different trials, at the same four levels of perceptual support. For the figure's clarity, the same and different choice options, respectively, are always shown to the left and right on the bottom of the screen. These positional assignments actually varied randomly for each trial in the RMTS task.

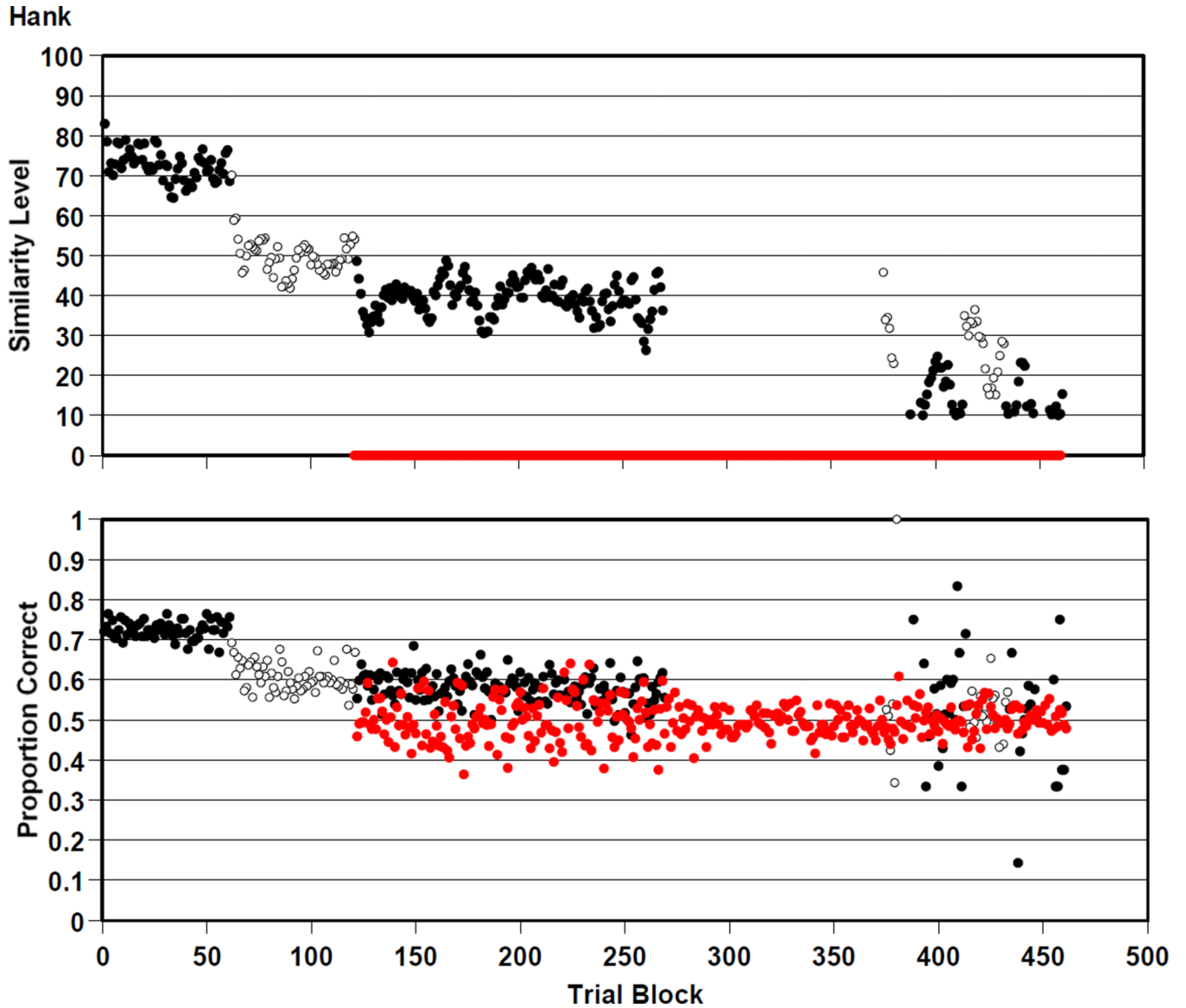


**Figure 2.** Humans' performance by 10-trial block in the RMTS task of Experiments 1A (top-left panels), 1B (bottom-left panels), and 1C (top-right and bottom-right panels). In The top panel gives the average level of perceptual support they experienced at each trial block. The definition of these levels of perceptual support is given in the text. The bottom panel gives humans' proportion correct for trials in each trial block. The red symbols summarized performance for the zero-similarity RMTS trials that offered no perceptual support.



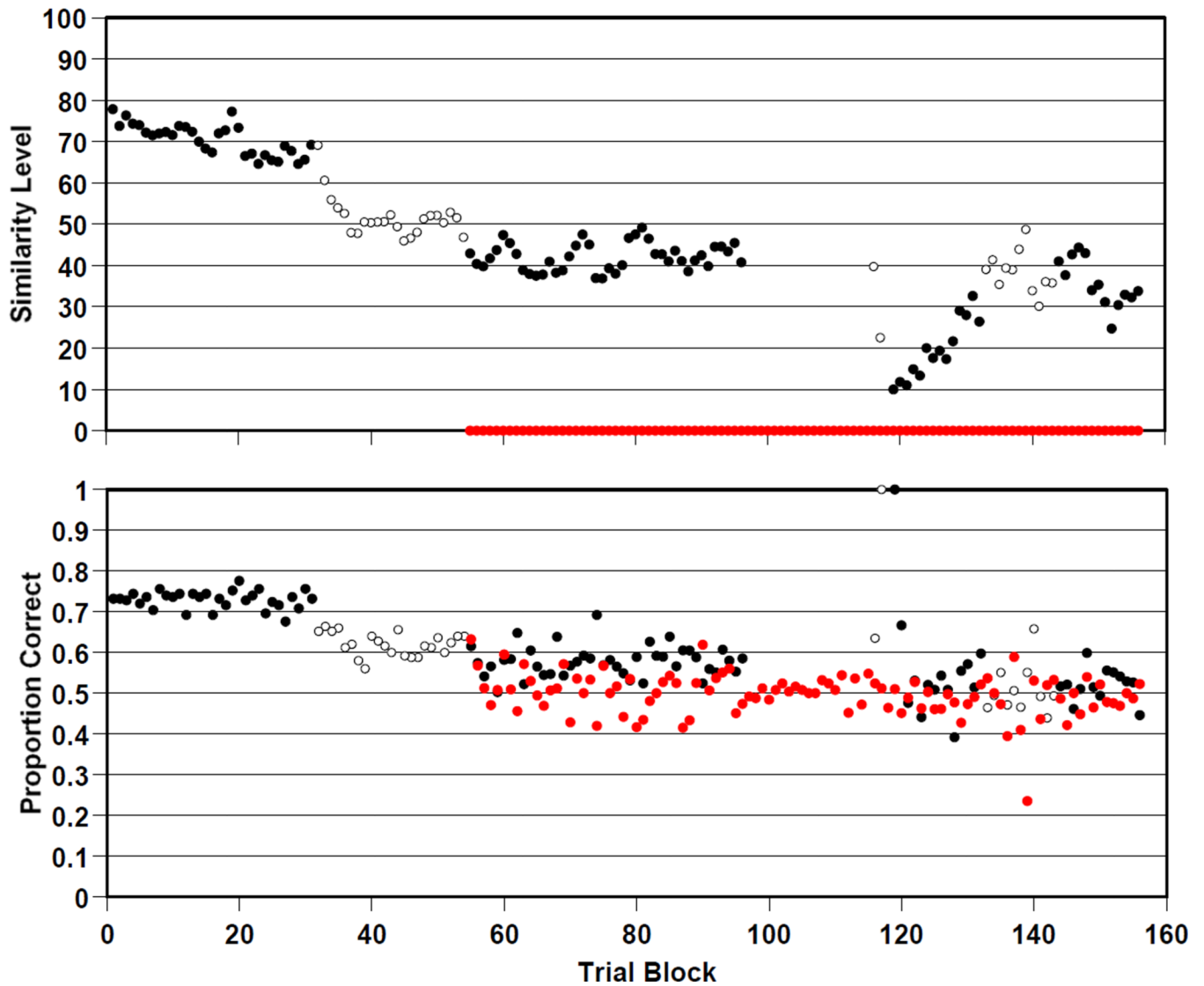


**Figure 3.** Humans' performance by 10-trial block in the RMTS task of Experiment 1C. Performance is shown for participants never given any perceptual support before and after their moment of task solution. Block 0 was defined to be the first block at which each participant achieved .80 correct and sustained that performance thereafter. Then, trial blocks were counted forward and backward from that criterion point.

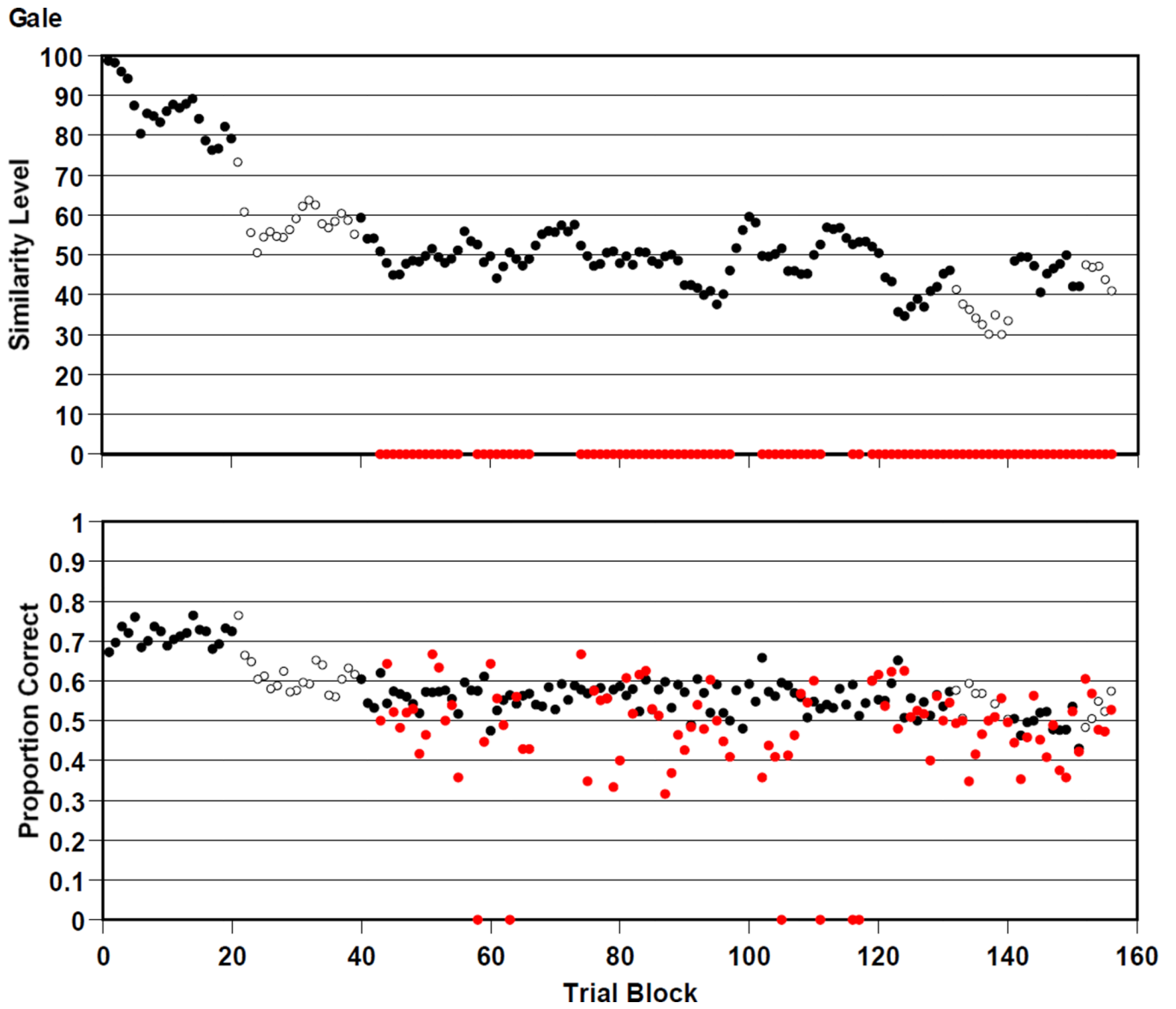


**Figure 4.** Hank's performance by 250-trial block in the RMTS task of Experiment 2. The top panel gives the average level of perceptual support Hank experienced at each trial block. The definition of these levels of perceptual support is given in the text. The succession of different tasks that Hank received is indicated by an alternation of filled and open symbols, or by a lapsing of the black symbols during a task method that only presented zero-similarity trials. These specific tasks are described in the text. The concurrent or sole presence of zero-similarity trials—that offered no perceptual support—is indicated by the red symbols at perceptual level 0. The bottom panel gives Hank's proportion correct for all trials in each trial block, with task alternation indicated as just described. Performance is shown on trials that potentially offered some level of perceptual support (black symbols) and on zero-similarity trials that offered no perceptual support.

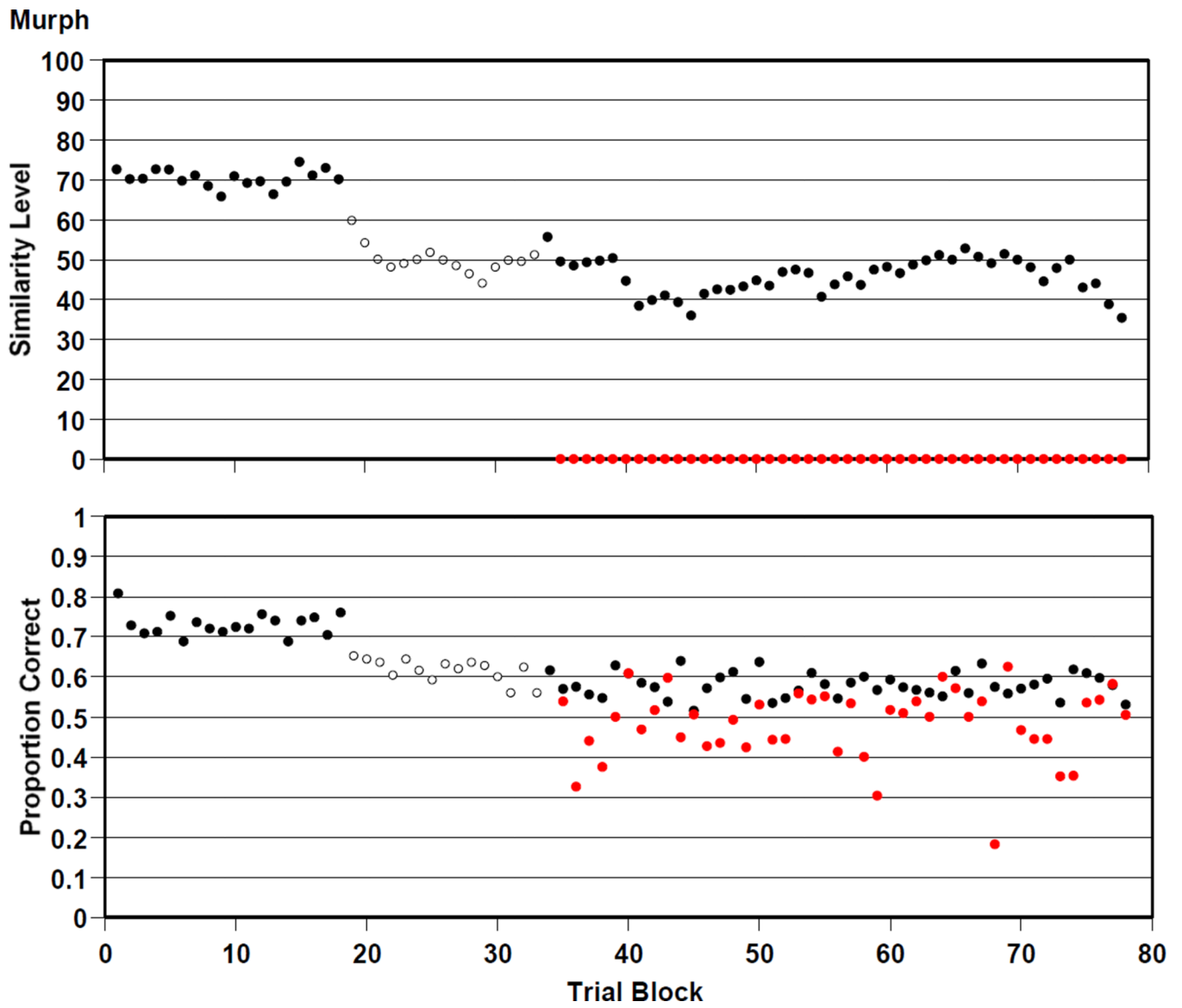
## Obi



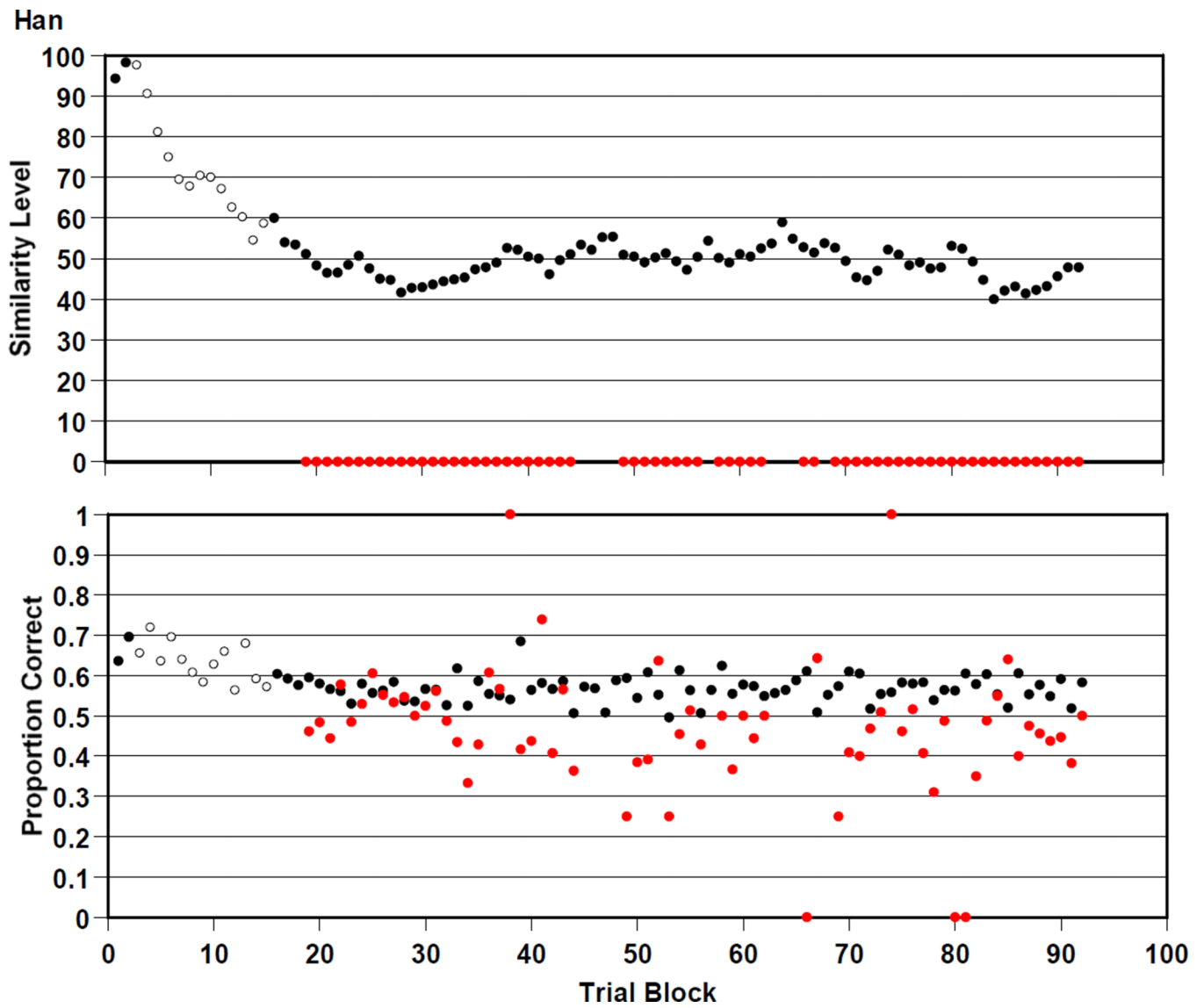
**Figure 5.** Obi's performance by 250-trial block in the RMTS task of Experiment 2, graphed as described in the caption to Figure 4.



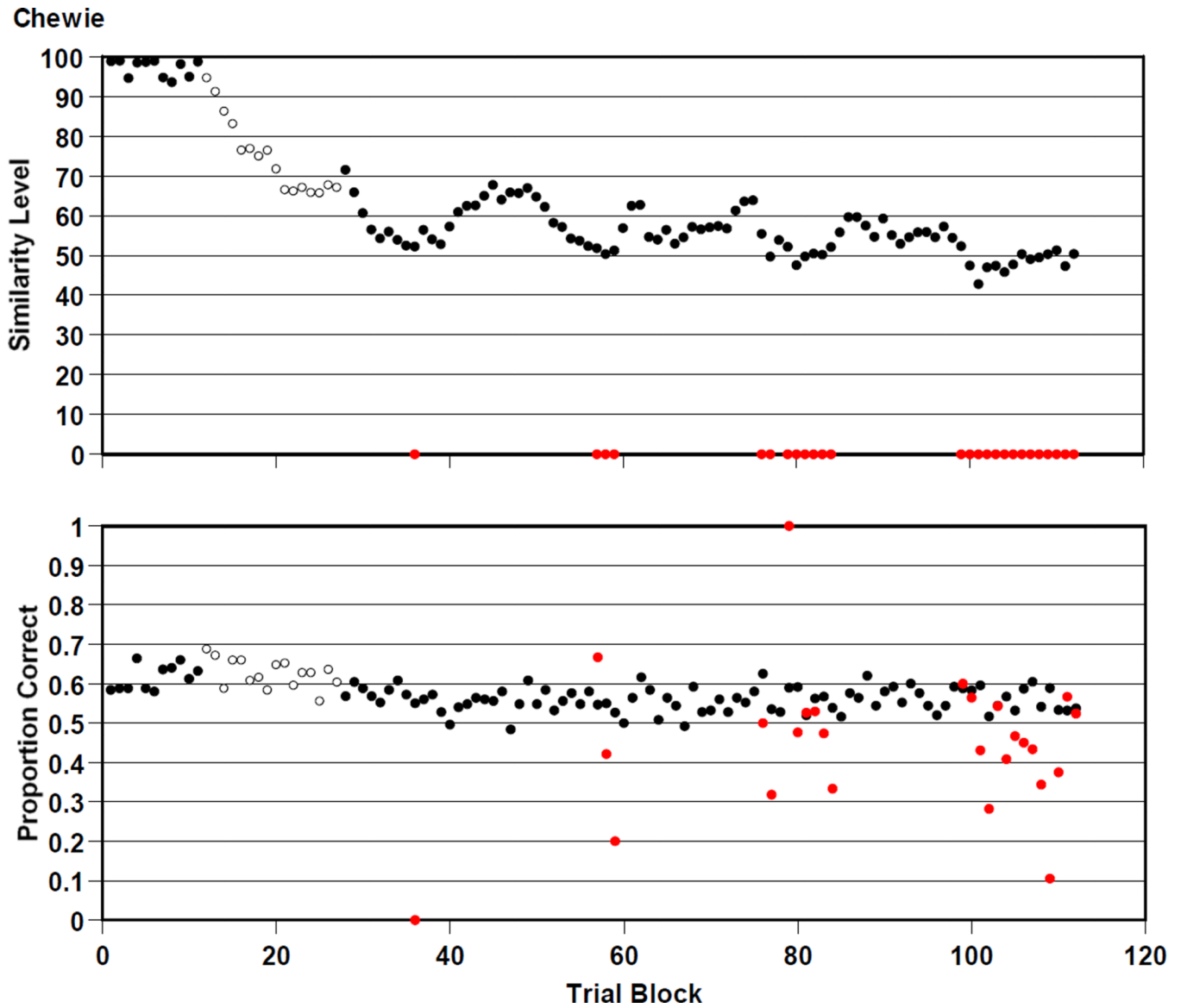
**Figure 6.** Gale's performance by 250-trial block in the RMTS task of Experiment 2, graphed as described in the caption to Figure 4.



**Figure 7.** Murph’s performance by 250-trial block in the RMTS task of Experiment 2, graphed as described in the caption to Figure 4.



**Figure 8.** Han's performance by 250-trial block in the RMTS task of Experiment 2, graphed as described in the caption to Figure 4.



**Figure 9.** Chewie’s performance by 250-trial block in the RMTS task of Experiment 2, graphed as described in the caption to Figure 4. On page 4 in the first sentence, why not take out "weakly" and give the reviewer the point he asked for, which is acknowledgment that pigeons do learn relationally sometimes. I think this is a smart move, too, as this person wanted some of those other papers cited, which we did not do, and I agree that in some other kinds of tasks there is better evidence of relational learning than in the RMTS tasks.