



Breaking the perceptual-conceptual barrier: Relational matching and working memory

J. David Smith¹ · Brooke N. Jackson¹ · Barbara A. Church²

Published online: 10 January 2019
© The Psychonomic Society, Inc. 2019

Abstract

Cognitive, comparative, and developmental psychologists have long been interested in humans' and animals' ability to respond to abstract relations, as this ability may underlie important capacities like analogical reasoning. Cross-species research has used relational matching-to-sample (RMTS) tasks in which participants try to find stimulus pairs that “match” because they both express the same abstract relation (same or different). Researchers seek to understand the cognitive processes that underlie successful matching performance. In the present RMTS paradigm, the abstract-relational cue was made redundant with a first-order perceptual cue. Then the perceptual cue faded, requiring participants to transition from a perceptual to a conceptual approach by realizing the task's abstract-relational affordance. We studied participants' ability to make this transition with and without a working-memory load. The concurrent load caused participants to fail to break the perceptual-conceptual barrier unless the load was abandoned. We conclude that finding the conceptual solution depends on reconstruing the task using cognitive processes that are especially reliant on working memory. Our data provide the closest existing look at this cognitive reorganization. They raise important theoretical issues for cross-species comparisons of relational cognition, especially regarding animals' limitations in this domain.

Keywords Same–different · Relational judgments · Analogies · Explicit cognition · Comparative cognition

Cognitive scientists have long focused on humans' ability to respond to abstract relations like sameness and difference (Wasserman & Young, 2010). James (1890/1950) described relational concepts as the backbone of thinking. Relational concepts underlie humans' analogical reasoning (Hummel & Holyoak, 2003). They mark cognitive-developmental change (Christie & Gentner, 2014; Ferry, Hespos, & Gentner, 2015; Gentner, 2003; Hochmann, Mody, & Carey, 2016; Hochmann et al., 2017). They may be a core executive function (Grafman & Litvan, 1999) that grounds higher cognition (Halford, Wilson, & Phillips, 2010). They may reveal an important discontinuity between human and animal cognition (Locke, 1690; Penn, Holyoak, & Povinelli, 2008). Thus, understanding the

cognitive organization of relational concepts is a significant empirical and theoretical goal.

In comparative psychology, the consensus is that relational concepts, like *same* and *different*, are sophisticated and phylogenetically restricted (Herrnstein, 1990). Relational judgments require an abstraction beyond the task's perceptual level. Many species find this abstraction difficult. Their same–different (SD) performances are fragile and difficult to train (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). Premack (1978) theorized a changing balance across evolution between concrete-perceptual and abstract-conceptual levels in cognition. In some species (pigeons, he thought), perceptual processing dominates; in others (apes and humans, he thought), relational processing dominates. By this narrative, humans should have a highly abstract and relational cognitive system—as in fact they do.

By this narrative, monkeys are middling. They succeed on some relational tasks (Katz, Wright, & Bachevalier, 2002; Shields et al., 1997; Wasserman, Fagot, & Young, 2001; Wright, Cook, & Kendrick, 1989; Wright, Rivera,

✉ J. David Smith
jsmith395@gsu.edu

¹ Department of Psychology, Georgia State University, 738 Urban Life Building, 140 Decatur Street, Atlanta, GA 30303, USA

² Language Research Center, Georgia State University, Decatur, GA, USA

Katz, & Bachevalier, 2003; Wright, Santiago, & Sands, 1984; Wright et al., 1990). In other cases, they have great difficulty or fail (D'Amato & Columbo, 1989; D'Amato, Salmon, & Columbo, 1985; Katz et al., 2002; Shields et al., 1997).

Monkeys' limitations extend to higher-order relational tasks, especially the relational matching-to-sample (RMTS) task that is our focus here. In the RMTS task, if the participant sees a *same* pair of objects (AA) as a sample, they should then respond by choosing a second *same* pair of objects (BB)—not a *different* pair of objects (CD). Or, given a *different* pair of objects (AB), they should respond by choosing a second *different* pair (CD)—not a *same* pair (EE). This task has been a staple of recent comparative research on relational cognition.

For example, Fagot and Parron (2010) used adjacent color patches as object pairs. At first, the color pairs were grouped so closely as to seem to be single stimuli. Baboons matched successfully; but when the color patches were separated spatially, so that the relation of object pairs had to be matched, baboons' performance collapsed. They did not group spatially separated objects into a relational pair whose same or different color relation could then be matched.

Flemming, Thompson, Beran, and Washburn (2011) tried to foster RMTS performance by using differential outcomes for same and different trials (e.g., big and small rewards for correct same and different responses). They hoped that differential rewards would psychologically demarcate the trial types. Monkeys still failed to achieve stably successful RMTS performance.

Fagot and Thompson (2011) used dogged training with extensive trial repetition. Six baboons (of 29) met an 80% criterion after 15,000–30,000 trials on their RMTS task constructed using 10 repeating geometric shapes. This study suggests that baboons, given extensive training, have some cognitive foundation for detecting and matching relational sameness and difference.

Fagot, Wasserman, and Young (2001) tried to foster baboons' RMTS performance using same or different multi-item arrays (e.g., up to 16 identical clip arts) instead of stimulus pairs (e.g., two identical clip arts). However, baboons matched arrays by relying on a low-level visual-entropy cue. That is, 16-item same and different arrays could be successfully differentiated because they are visually calm or visually jazzy, respectively. When the entropy cue was weakened by using object pairs instead of multi-item arrays, performance collapsed. Young, Wasserman, and Garner (1997) showed the same effect in pigeons. The present task uses only two-item pairs, so the entropy cue is minimized or eliminated (see also Castro & Wasserman, 2013).

Smith, Flemming, Boomer, Beran, and Church (2013) provided the closest model for the present study, and its

strongest motivation. They tried to foster monkeys' RMTS performance using perceptual cues. Rhesus monkeys were given a bistable RMTS task as follows. Each trial made available to the monkeys both a first-order perceptual level and a second-order relational level that could be used for successful task performance. So the monkeys had two routes to performance, and, in particular, they had a natural perceptual route into the task. Then, using a method described in detail below, researchers weaned subjects off the perceptual cue by progressively weakening its strength and usefulness. This weaning process should have left the monkeys, we thought, only to discover the still-available relational solution. Thus, we hoped to place monkeys in their strongest position to transition from the perceptual level of the task to the relational level, finally showing a robust, stand-alone, relational performance.

However, Smith et al.'s (2013) attempt failed (see Fig. 1). Macaques' RMTS performance collapsed as they were weaned from perceptual support. None showed successful RMTS performance, even after 260,000 trials during which we tried with special training techniques to coax a relational performance from them. In fact, Fig. 1 shows where monkeys' perceptual-relational barrier lay. Perceptual support could weaken from Similarity Level 90 (salient perceptual support) down to Level 0 (no perceptual support). Monkey Hank (and other macaques) performed well, given strong perceptual support, but they could not perform beyond chance for weak perceptual support. They failed to break through the perceptual-conceptual barrier. As the perceptual cue was faded, their capacity to match successfully was eliminated.

In contrast, Fig. 2 shows humans in the same task. They were 99% correct overall. They constantly met the task's performance criterion—that is, the criterion necessary for reducing similarity level and weakening the perceptual cue—so that perceptual support steadily waned. Humans' proportion correct was never reduced by this weakening. Judging by our experience in this task, the perceptual cue becomes too weak to use (by us!) as similarity level falls through the 60s into the 50s (Methods). Over this range, cognitive control was somehow spontaneously transferred over to the conceptual-relational cue that thereafter controlled performance. Humans managed this transition easily. They did make more errors through this range of similarity levels, though the absolute number of errors was small. These errors may reflect the conceptual transition in the task, but the reflection is faint because humans' transition so seamlessly.

The general failure of animals in RMTS tasks is a topic of sharp interest in the comparative literature. Monkeys' and humans' contrasting performance presents an unsolved information-processing mystery. What

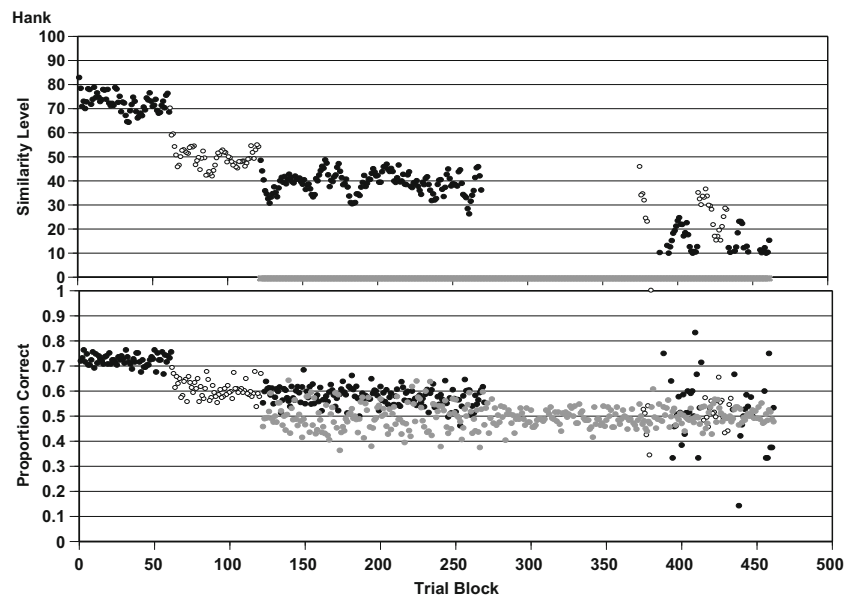


Fig. 1 The monkey Hank’s performance by 250-trial block in Smith et al.’s (2013) relational match-to-sample (RMTS) task. Top: Level of perceptual support given to Hank. These levels are defined in the text. Black and white symbols denote different fostering conditions with which Smith et al. tried to help Hank’s RMTS performance. Gray symbols indicate the trial blocks during which trials were interspersed that offered Hank no perceptual support and demanded a conceptual or

relational strategy from him. Bottom: Hank’s proportion correct for all trials in each trial block, depicted as already described. Adapted from “Fading perceptual resemblance: A path for rhesus macaques (*Macaca mulatta*) to conceptual matching?” by J. D. Smith, T. M. Flemming, J. Boomer, M. J. Beran, & B. Church, 2013, *Cognition*, 129, 598–614. Copyright Elsevier Ltd. 2013. Reprinted with permission

cognitive processes do monkeys lack that humans have? What are humans doing cognitively to discover the task’s conceptual organization? They might state verbal rules (e.g., “Match sameness to sameness”), which monkeys would not do. They might label stimulus pairs with abstract labels/words (same, different), which wordless monkeys cannot do. This article explores these questions. We also seek to build a dialog between human and

animal researchers concerning the best way to understand species’ differences in these tasks.

Accordingly, we placed humans into the RMTS task that is so influential in comparative psychology. This task brought the present study several advantages. First, it let us study humans in the task that causes animals such generalized failure. Second, it let us study the simplest form of humans’ relational cognition—the RMTS

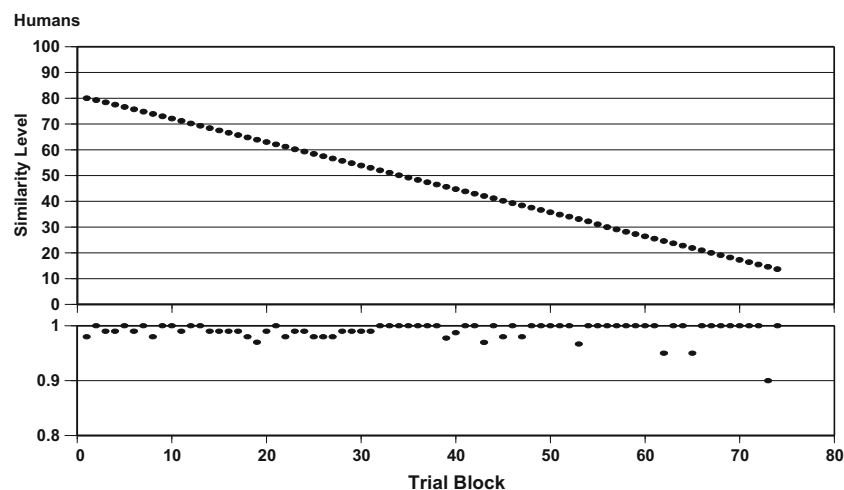


Fig. 2 Humans’ performance by 10-trial block in the relational match-to-sample (RMTS) task of Smith et al. (2013). Top: Average level of perceptual support they experienced at each trial block. Definition of these levels of perceptual support is given in the text. Bottom: Humans’ proportion correct for trials in each trial block. Adapted from “Fading

perceptual resemblance: A path for rhesus macaques (*Macaca mulatta*) to conceptual matching?” by J. D. Smith, T. M. Flemming, J. Boomer, M. J. Beran, & B. Church, 2013, *Cognition*, 129, 598–614. Copyright Elsevier Ltd. 2013. Reprinted with permission

task only involves pairwise matching, and its performance rules are perfectly transparent. Third, the comparative paradigm let us study relational cognition removed from human concepts, human narratives, human roles. Our task was different in these respects from the elegant tasks within the analogical-reasoning literature. Fourth, the comparative RMTS task let us study a language-free form of relational cognition by humans—that is, relational cognition understandable and performable without language.

Another advantage came from using Smith et al.'s (2013) bistable RMTS task. This task grants participants parallel solutions that are first-order perceptual and second-order relational. It ensures robust (perceptual) performance early on. Then, the perceptual can be faded later on. The participant is finally forced to reconstrue the task, to find the relational task approach, and we can study the information-processing character of that reconstrual. This let us focus on the cognitive reconstrual by which a perceptual task transitions into a relational task.

By understanding those reconstrual processes, one might illuminate not only how humans make conceptual discoveries but also why monkeys fail to do so. Our working hypothesis was that conceptual discoveries, or task reconstruals, involve the monitoring of failing performance as the perceptual cue weakens, the generation of alternative task hypotheses, the testing of those hypotheses in ongoing performance, and the acceptance of a new relational hypothesis given its success. Research in the cognitive neuroscience of categorization suggested that these processes would engage humans' explicit-declarative cognitive system, including prefrontal cortical circuits and the working-memory system (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Ell, 2001; Maddox & Ashby, 2004; Seger, & Miller, 2010; Smith et al., 2012; Smith et al., 2014; Smith, Zakrzewski, Johnson, Valteau, & Church, 2016).

Therefore, we tested a crucial working-memory manipulation, to see whether this might compromise participants' efforts toward relational-rule discovery. If it did, this would target working-memory resources as an important part of the cognitive processes by which humans' transition in the RMTS task from the perceptual to conceptual levels. In turn, this could help us interpret the sharp species differences in RMTS performance.

Experiment 1

Method

Participants Georgia State University undergraduates ($N = 120$), with normal or corrected vision, participated. They

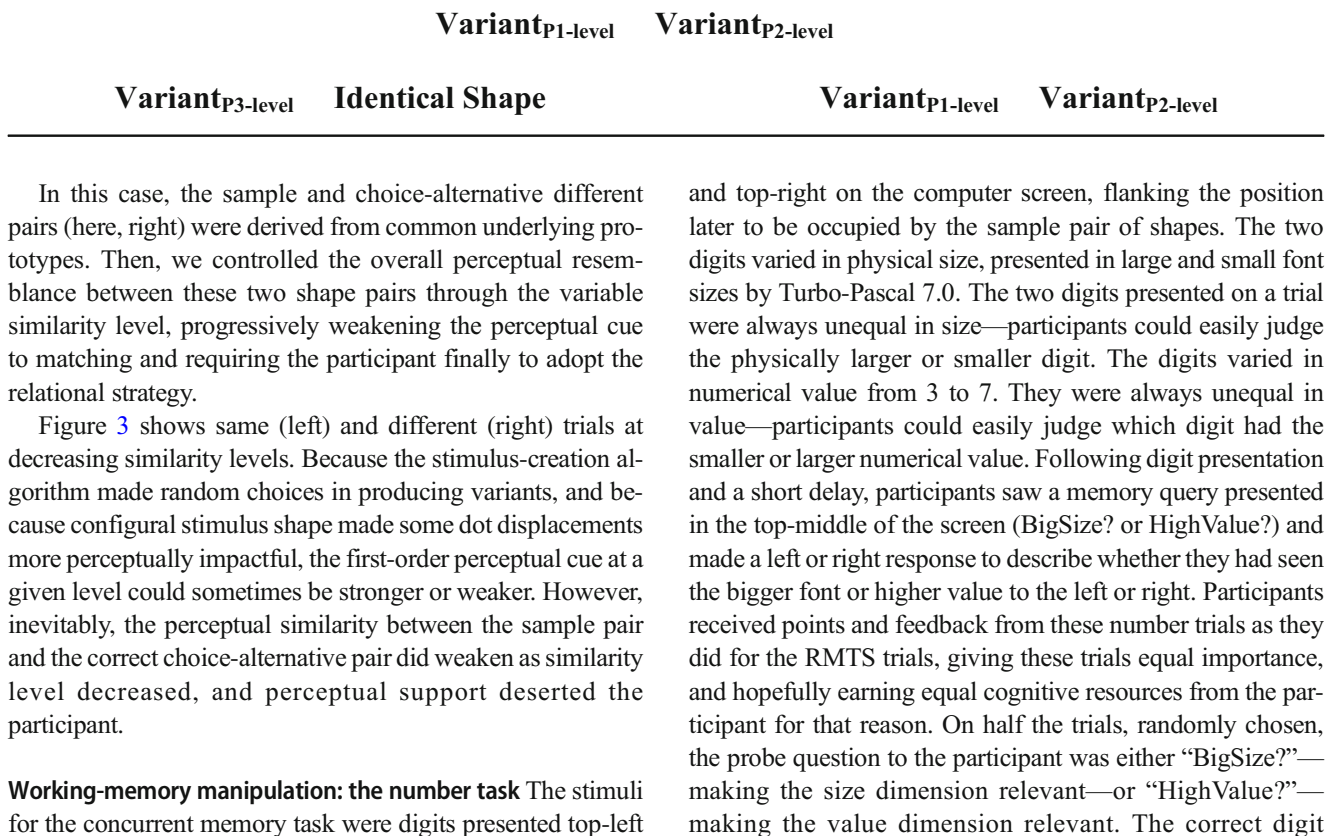
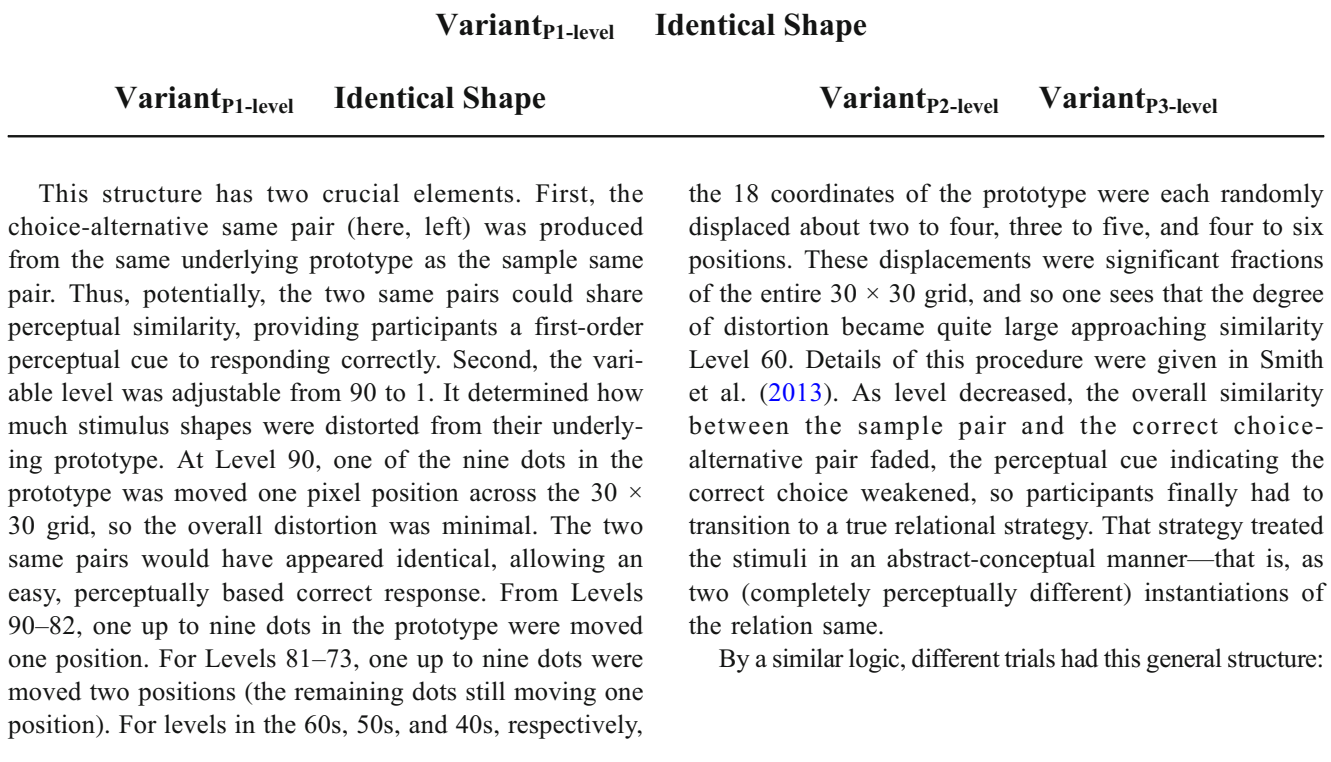
provided informed consent to be included and were compensated for participating by receiving partial course credit in a psychology course. We excluded 14 participants from analysis because they completed less than 300 trials (five and nine participants, respectively, in the control and concurrent conditions). We excluded one control participant for always making “left” responses. We excluded one control participant who solved the task and then went on to make 100% wrong responses to observe the effect. The data from 53 and 51 participants, respectively, were analyzed in the control and concurrent conditions.

Dot-distortion stimuli Stimuli were created using an influential method that generates polygon variants from originating prototypes (Posner, Goldsmith, & Welton, 1967), which has been used in human and monkey studies (e.g., Smith & Minda, 2001, 2002; Smith, Redford, & Haas, 2008; Smith, Redford, Haas, Coutinho, & Couchman, 2008). The method lets us control the perceptual similarity between stimulus pairs, the crucial element in our RMTS paradigm. It gives us endless possible trials, so we can present trials indefinitely without repetition (see also Brooks & Wasserman, 2008). It lets us present complex stimuli within a high-dimensional similarity space, perhaps mirroring the complex similarity relation among members of natural kinds.

Stimulus shapes are created as follows. Nine points are randomly selected from within a 30×30 grid as the vertices of a nine-pointed polygon designated a prototype. Variants of the prototype can be generated at different distortion levels discussed below. After selecting the 18 coordinates for a stimulus shape (i.e., 9 x, y coordinate pairs), the shape was centered within the 30×30 grid and was magnified to appear on a 90×90 pixel section of the screen. Finally, the DrawPoly procedure within Turbo Pascal 7.0 connected successive vertices of the shape by lines and filled the resulting complex polygon shape in yellow.

RMTS trials Each trial presented as a sample two polygon shapes centered at the top of a 16-inch computer screen on a black background. On same trials, they were the identical variant from the same prototype. On different trials, the two polygons were variants of different prototypes and thus extremely different perceptually. Two choice-alternative shape pairs—a same pair and a different pair—were presented at the screen's bottom left and bottom right (with left-right placement decided randomly on each trial). As the top pair of shapes was same or different, respectively, the participant was to make a left or right key press to choose the same or different pair on the bottom. For correct or incorrect responses, respectively, participants saw the message “+1” or “-1” on the screen. They also received an update on their total points in the session to that date.

The method of fading similarity Same trials in the RMTS task had this structure:



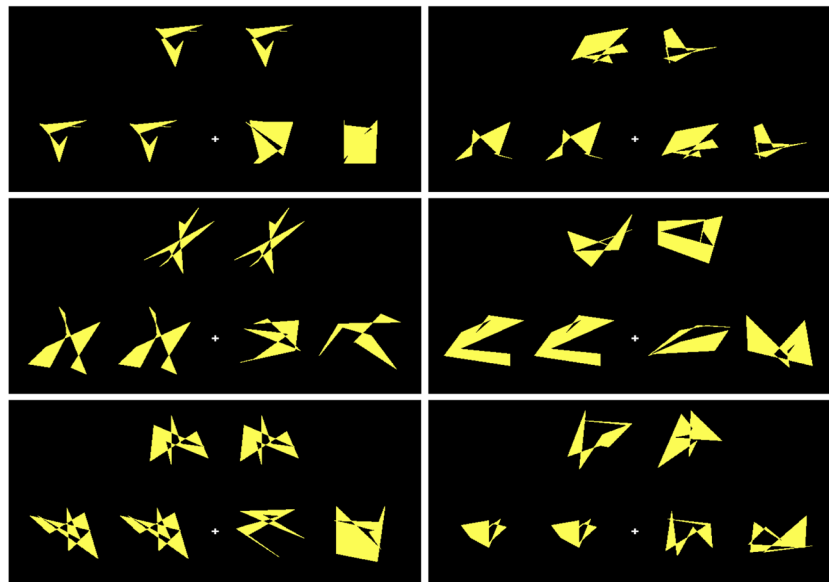


Fig. 3 Examples of trials from the relational match-to-sample (RMTS) task. Left column shows same trials, with the level of perceptual support set at 90, 42, and 18. These similarity levels are defined in the text. Right column shows different trials set at the same levels of perceptual support.

For clarity, the same and different choice options, respectively, are always shown to the left and right on the bottom of each black screen. These positional assignments were varied randomly for each trial type in the actual RMTS task

appeared to the left or right randomly across trials. We arranged it so that the irrelevant dimension's appearance was discrepant from the relevant dimension's appearance on 60% of trials. That is, if the relevant larger font was to the left, then that digit had the smaller numerical value on 60% of trials. We believe that this slight miscorrelation makes the concurrent task more difficult so that it involves a greater memory load. (This miscorrelation cannot be carried too far, however, or participants seize the miscorrelation to reduce their memory load. Humans, like monkeys, are not above using any shortcut that a task offers.) The digits were always presented for 0.6 s, in white on a black screen. Then they were masked over their entire area by a square white mask for 0.2 s.

Both conditions of the experiment included these number trials coordinated with the RMTS trials in contrasting manners as described now. In the control condition, the number and RMTS trials perfectly alternated, with each trial kind taken independently through to response and feedback before the other trial kind initiated. Thus, participants saw two digits for 0.6 s, these were masked for 0.2 s, and then participants received the number-task query already described. They responded and received right–wrong feedback. Then the next RMTS trial began. As a result of full alternation, the participants did not have to hold active number information in working memory as they processed RMTS trials.

In the concurrent condition, the number trial initiated, showing the digits for 0.6 s, to be encoded by the participant for a future query, with masking following (overcovering white rectangles) for 0.2 s. Then, the number trial was suspended, and the RMTS initiated exactly at that point. It was taken all the way through to

response and feedback. Then, the number trial was unsuspended, the query regarding the number trial was given, a response was offered by the participant, and right–wrong feedback was delivered. Given this coordination between trial types, the participants trying to complete RMTS trials, and possibly trying to reconstrue the RMTS task when the perceptual cue grew weak, needed to do so while holding enough information in working memory from the previous digit presentation to answer the query correctly. This memory information is itself quite interesting—to us, it appears to take the form of an extremely rapidly contrived verbal description of the left–right digit array.

Readers should be clear on exactly what aspect of this trial coordination was concurrent. The number information was never on the screen during the RMTS trial. The number query was never on the screen during the RMTS trial. The concurrent aspect of the task was invisible, but crucial: That is, the participant had his or her working memory occupied with previous digit information during the whole execution of the RMTS trial and during response and the receipt and interpretation of feedback.

Training and instructions Participants in both conditions received 40 number trials alone as the experiment began to familiarize them with the number-memory task. They were told that two numbers would appear on each trial and then be hidden by white rectangles. They were told to remember their value and their size, so that they could answer the questions: HighValue?, BigSize?. They were told they were to press the key labeled L (left) or R (right) to indicate which

of the two digits had HIGHER value or BIGGER size. And they were told they would gain or lose one point for correct or incorrect responses.

Combined trials: instructions Entering the phase of the experiment with number and RMTS trials combined, participants were told that they should still look at the numbers, remember their sizes/values, and respond to the number questions correctly. They were told that they would also see a pair of shapes at the top of the screen and a pair of shapes left and right at the bottom. They were told that they should press L (left) or R (right) to choose the correct pair of bottom shapes and that they would learn through practice how to be correct. They were also told that they would gain or lose points for correct or incorrect responses in both tasks. One sees that the instructions for the shapes task were minimal, silent on the task's dual perceptual and conceptual bases, leaving participants to construe the task and choose a task strategy for themselves.

Trials continued until participants completed 320 trials or until they reached a 50-minute time limit. The prevailing similarity level was initially high (level = 90) so that participants had a strong perceptual cue supporting their correct choice of the same or different pair. The crucial progression in the task was to gradually decrease level, weakening the low-level perceptual cue and eventually persuading participants toward the task's relational basis—if they could find it. This progression was based on the participant's performance. We monitored the participant's performance on the most recent 10 trials. Every two trials, the controlling software asked if it should reduce level. It reduced level by one, but not below one, if recent performance on shape trials and number trials had been above 0.80 and 0.70, respectively. It increased level by one, but not above 90, if recent performance on shape trials and number trials had been below 0.80 or 0.70, respectively. Level kept its current value if these conditions were not in effect. So, strongly performing participants paved their own path to a weakening perceptual cue. Weakly performing participants could be rescued as the program restored their perceptual support. A strong aspect of this task progression was that we could observe the levels of perceptual support at which participants stalled out in their progression in the two conditions, and from which they needed rescue.

Results

Overall performance The 53 and 51 participants in the control and concurrent conditions, respectively, were .913 and .865 correct in the matching task. They were .938 and .833 correct in the number task accompanying the RMTS trials—in alternation in the control condition and with number memory challenging matching in the concurrent condition.

Figure 4 shows the progression of performance for both conditions. Squares and triangles, respectively, show the

average proportion correct by five-trial blocks for matching and number memory. Diamonds show the task's similarity level on average during each block. To include similarity-level and accuracy measures on the same graph, we divided level (ranging from 90 down to 1) by 100 so that it ranged from 0.9 to 0.01. Level depicts the progression down the slope of fading perceptual support. Remember that these decreases were responsive to—and titrated by—humans' matching and memory performances. Both groups initially met the performance criteria steadily, so that similarity level steadily decreased, weakening the perceptual cue. Finally, the participant was forced to transition—if they could—to a relational strategy not based on perceptual similarity.

Figure 4 gives one view of the principal result. Control participants' RMTS performance (top)—indexed by the downward progression of similarity level—was never reduced by weakening perceptual support. They broke through the task's perceptual-conceptual barrier easily, with only a faint slowing of the similarity progression for trial blocks in the midteens.

In contrast, concurrent participants stalled (bottom). The progression of decreasing similarity levels found an asymptote at about 49 (0.40). This asymptote resembles that of macaques in Smith et al. (2013; see Fig. 1). The performance of these participants was compromised by the need to actively maintain memory material during matching, especially as perceptual support waned.

For statistical analysis, we averaged across every four trial blocks, creating 16 levels of consolidated trial blocks. We entered the progressions of similarity level into a two-way generalized linear model (GLM), with trial block as a within-participant factor and condition as a between-participants factor. There was a significant main effect for trial block, indicating that similarity level decreased across blocks, $F(15, 1530) = 234.558, p < .001, \eta_p^2 = .698$, and a significant main effect of condition, reflecting a higher similarity level in the concurrent condition, $F(1, 102) = 16.411, p < .001, \eta_p^2 = .139$. Most important, there was a significant interaction between trial block and condition, indicating that similarity level fell more strongly in the control condition than in the concurrent condition, $F(15, 1530) = 11.640, p < .001, \eta_p^2 = .102$. Figure 4 already made the character of this interaction plain.

Participant subgroups: control condition The graphs in Fig. 4 do not exhaust these data. Indeed, they likely average away important performance differences among participants. Illustrating this point, Fig. 5 (top) shows the data profile produced by a large subgroup of 43 control participants (80% of that group). This group's similarity level fell precipitously, as rapidly as allowable by the controlling software (one step each two trials).

In contrast, 20% of control participants produced the data profile in Fig. 5 (bottom). Similarity level fell at first, but only

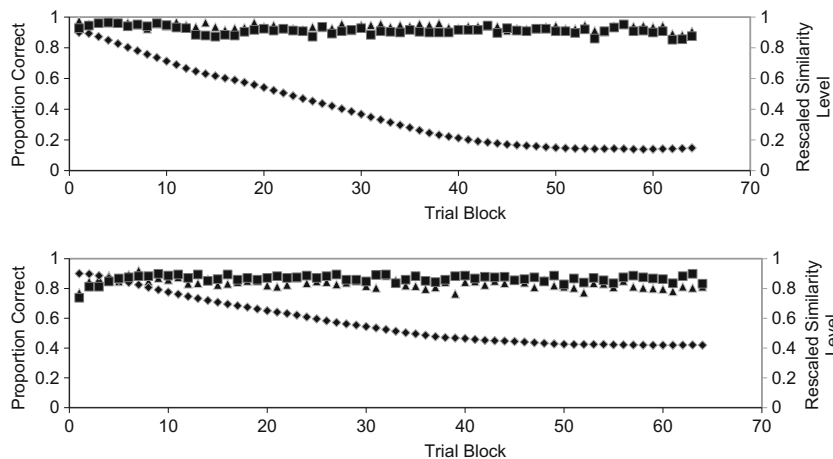


Fig. 4 Top: Performance of control participants in the relational-matching task by five-trial block. Diamonds show the progression of decreasing first-order perceptual similarity between the sample and the correct choice alternative. This was measured, as described in the text, as the extent to which the vertices of two complex polygons were different in coordinate space. Squares show participants’ proportion correct in the

matching task. Triangles show their proportion correct on the number-memory task that alternated with the matching task. Bottom: Performance of concurrent participants in the relational-matching task, depicted in the same way. In this case, the matching and memory tasks were interleaved in a way that caused the matching task cognitive interference

into the 60s. Their ongoing matching and memory performance fell apart to the point that similarity level no longer received its criterial permission to advance. Similarity level hit a wall in the 60s and never progressed beyond that point.

We defined these subgroups to be contrastive and non-overlapping in their lowest similarity level reached, and to include as many participants as possible. We chose this approach instead of predefining percentile groups because we did not know the sizes of the contrastive subgroups we would find. If instead one contrasted the 10% best progressing participants and the 10% worst progressing participants (measured by the final similarity level reached), then the results would be stronger than those we found. If one contrasted the best and worst

quartiles, then the results would be weaker (from including too many participants in the Fig. 5 bottom subgroup).

These struggling participants illustrate that—even when a concurrent task is not on the scene—breaking through the perceptual-conceptual barrier is not something that always happens, or happens automatically, or happens procedurally through associative learning and reinforcement. It is optional. It is probabilistic. It waits on cognitive processing achieving the appropriate realization. We believe that crossing the perceptual-conceptual barrier requires a qualitative reconstrual of the task, and that this construal is derived through active and explicit cognitive processes that may load the utility of working memory. On this hypothesis, we should

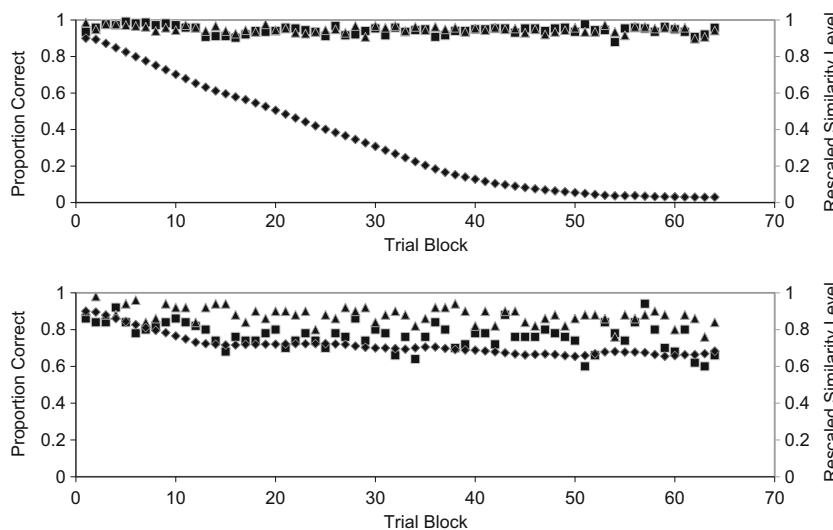


Fig. 5 Top: Performance of a subgroup of control participants in the relational-matching task, depicted as described in the caption to Fig. 4. Bottom: Performance of another subgroup of control participants, depicted in the same way

see participants' RMTS performance falter in the face of a concurrent cognitive load. We will evaluate this possibility shortly.

Finally, Fig. 6 suggests that even the strongest control participants had their own struggle in making the perceptual-conceptual transition. For these 14 participants, similarity level slowed in its downward course in the low 60s. This was accompanied by a slight drop in matching performance—this is what caused the controlling software to deny permission for similarity level to decrease during these blocks. Thus, all participants, in their own way, appear to have had a struggle in crossing the perceptual-conceptual threshold. This is the point when the task must be reconceptualized, reconstructed. Our data provide the closest existing look at this transition point.

This result also shows that it is this time of reorganization, the time of reconstruction, that causes humans the focal difficulty in the task. After that time, matching and memory performance remain strong. This suggests that the working-memory resources are needed to reconceive the task, developing the conceptual approach when the perceptual cue grows too dim.

Participant subgroups: concurrent condition The same subgroups appeared in the concurrent condition, though their relative sizes changed. Twenty-five participants (49% of the group, not 80% as in the control condition) produced the data profile shown in Fig. 7 (top). Their similarity level fell precipitously, to its floor. This performance was just like that in Fig. 5 (top). But now this subgroup was just half of the participant group. This reduction is one effect of the interference produced when working-memory material must be maintained during the performance of the matching trials.

In contrast, twenty-six participants (51%, compared with 20% in the control condition) produced the data profile shown in Fig. 7 (bottom). Their similarity level fell much less. This is another effect produced by the concurrent load. Similarity level was not permitted to advance. In some instances, it even regressed to strengthen the perceptual support provided to participants and make the task easier. Figure 7 (bottom) also shows that these participants had substantial difficulty managing the number-memory task as similarity level dropped.

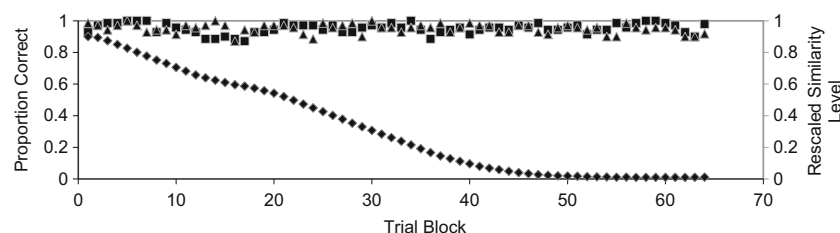


Fig. 6 Performance of 14 strong control participants in the relational-matching task, depicted as described in the caption to Fig. 4. Even they experience a distinctive period of cognitive reorganization, during the time in which they abandon the perceptual cue and adopt a relational

Probably, they were trying to spare resources, withdrawing them from the memory task, devoting them to the matching task as the perceptual cue weakened.

The concurrent task had different consequences for different participants (see Fig. 8). Sometimes, similarity level fell smartly until the 60s, but then stalled out at the task's perceptual-conceptual barrier. Sometimes, similarity level hit the barrier, backed off to regroup, then made another approach. Sometimes, similarity level hit the barrier, and then eased off for the duration. As psychologists who study macaques and humans comparatively, we appreciate the equalizing application of a concurrent load to turn human participants into "macaques" (compare Fig. 1, above).

Experiment 2

In Experiment 1, we weaned participants delicately, weakening the perceptual cue only given ongoing successful matching and memory performance. But there could be parallel interest in a complementary, harsher approach by which we would persist in weakening the perceptual cue, no matter the faltering performance. Then one might see participants' struggle facing the sharp, rapidly developing demand for a relational transition. Experiment 2 provides this complementary view.

Method

Participants One-hundred and twelve undergraduates from the same testing population participated for the same compensation. In the control condition, we excluded two participants for failing to complete 320 trials, one participant for failing to achieve 85% number matching in one of the first two trial blocks, and 10 participants for failing to achieve 85% correct in the RMTS task in one of the first two blocks. In the concurrent condition, we excluded one participant for a strong left-response bias, eight participants for failing to achieve 85% number matching in one of the first two trial blocks, and 10 participants for failing to achieve 85% correct in the RMTS task in one of the first two blocks. Based on effect sizes

strategy instead. The drop in matching performance (squares) and the shoulder in decreasing similarity levels (diamonds) reflect this reorganization

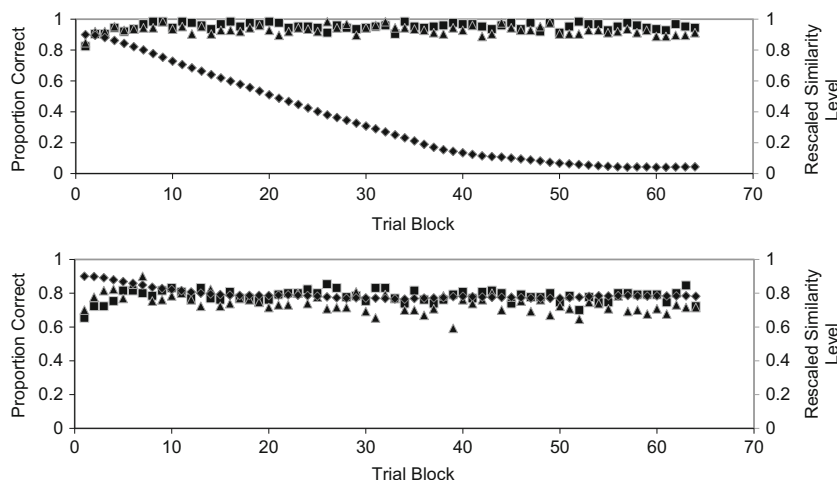


Fig. 7 Top: Performance of a subgroup of concurrent participants in the relational-matching task, depicted as described in the caption to Fig. 4. Bottom: Performance of another subgroup of concurrent participants, depicted in the same way

estimates, we used a stopping rule of 40 analyzable participants in each condition.

The RMTS task The RMTS task was just as that described in Experiment 1, except for different rules governing the weakening of the perceptual cue through the session. In Experiment 2, similarity level was decreased by one step every three trials, no matter the participant’s matching or memory performance.

The working-memory task The number-memory task was just as that described in Experiment 1. The RMTS and number-memory tasks were separated into different trial phases in the control condition as before, causing no working-memory interference on matching. The tasks were interleaved in the concurrent condition as already described, producing working-memory interference on matching.

Training and instructions All aspects of the training and the instructions were the same as before.

Results

Overall performance The 40 participants each in the control and concurrent conditions, respectively, were .903 and .889 correct in the RMTS task—strong group matching

performances. They were .939 and .868 correct in the number-memory task that accompanied the RMTS trials—in strict alternation in the control condition and with RMTS performance competing with number memory in the concurrent condition.

Figure 9 shows for both conditions the progression of performance. Squares and triangles, respectively, show the average proportion correct by twenty 16-trial blocks for matching and number memory. Similarity level is not shown now, because its value was perfectly correlated with trial/block number in synchrony across participants. Still, though, participants were finally faced with the need to transition to the relational strategy—if they could.

We entered the data in Fig. 9 into a three-way GLM, with trial block (1–20) and task (matching, number memory) as within-participant factors, and condition (control, concurrent) as a between-participants factor. There was a significant main effect for condition, indicating worse performance in the concurrent condition, $F(1, 78) = 6.670, p = .012, \eta_p^2 = .079$, and a significant main effect of block, indicating worse performance in later blocks, $F(1, 1482) = 10.645, p < .001, \eta_p^2 = .120$. There was also a significant interaction between task and condition, indicating that the concurrent condition had a larger gap between matching and memory performance, $F(1, 1482) = 9.618, p = .003, \eta_p^2 = .110$. Finally, there was a significant

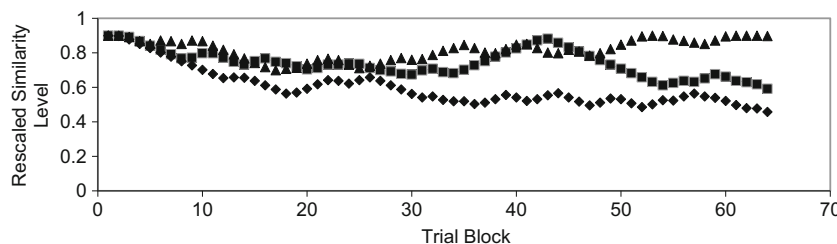


Fig. 8 Three selected profiles of similarity-level changing by trial block through the experimental session. Participants, by turns, hit their perceptual-conceptual barrier and stalled there (diamond symbols), or

hit their barrier and retreated from there (triangle symbols), or tried to break through their perceptual-conceptual barrier twice (square symbols)

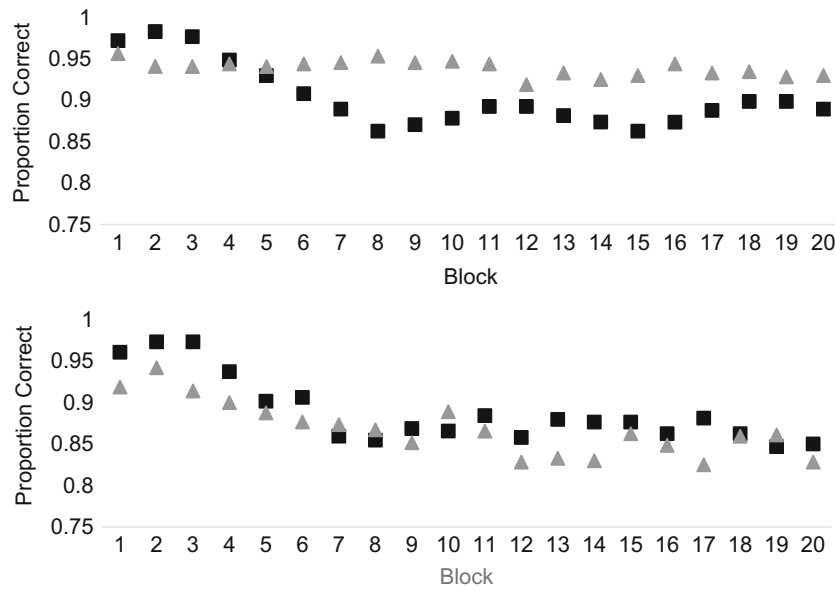


Fig. 9 Top: Performance of control participants in the relational-matching task (black squares) and the number-memory task (gray triangles) across the 20 blocks as the perceptual cue rapidly faded. Bottom: Performance of concurrent participants depicted in the same way

interaction between task and block, indicating that matching performance fell off more sharply over blocks than memory performance did, $F(1, 1482) = 3.842, p < .001, \eta_p^2 = .047$, but no significant three-way interaction ($F < 2$).

We will summarize these results intuitively. First, matching performance fell off through blocks to about the same degree in the control and concurrent conditions. This is intuitive because participants are having to derive and then execute the new, relational strategy. Second, memory performance fell off much more sharply in the concurrent condition. Working memory was a casualty of the cognitive trouble that participants faced. They sacrificed the working-memory task to have resources available to reconstrue their matching task when it began to go badly.

Figure 10 gives an alternative perspective on the results. Here, we tried to estimate for each participant their point of maximum cognitive difficulty, and we tried to align those blocks, concentrating observed effects. Block 0 is the aligned point of worst matching performance. Blocks are numbered forward and backward from that focal point. This is a version of a backwards learning curve. The matching performance levels at Block 0 are artifactually low, because that block was defined to be the lowest out of all of each participant's blocks. The number-memory performance levels are not artificial, because these levels did not enter the criterion definition of lowest performance.

These graphs are highly expressive of Experiment 2's main results. In both conditions, matching performance (squares)

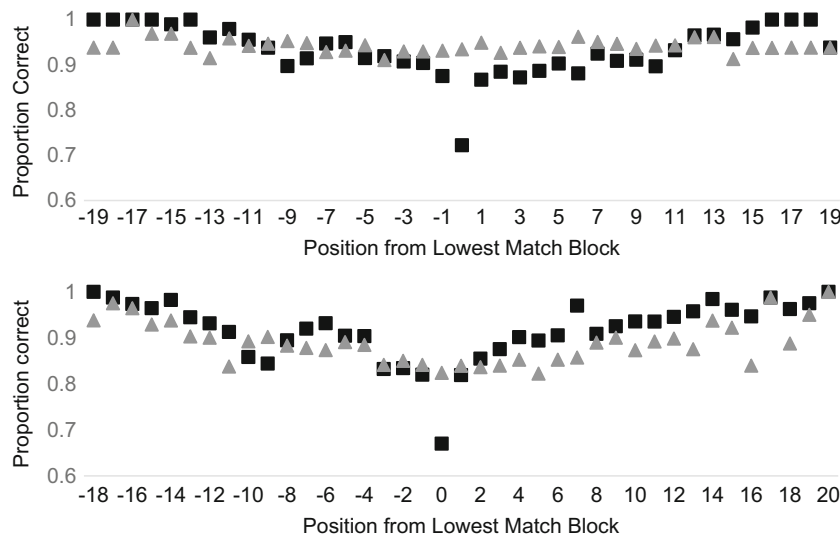


Fig. 10 Top: Performance of control participants in the relational-matching task (black squares) and the number-memory task (gray triangles), aligned by participants' worst block of matching performance (Position 0). Bottom: Performance of concurrent participants depicted in the same way

fell off near the low point Block 0. For many participants, this is likely the point of cognitive struggle with a weakening perceptual cue and a growing need to reconstrue the task. In this view of the data also, the working-memory task in the concurrent condition was a clear casualty during this phase of the task. Moreover, it was slow to recover. However, in the control condition, number memory performance stayed constant.

We compared the memory-performance levels for Blocks -1, 0, and 1 across the control and concurrent conditions. The latter performance was significantly lower (0.840 correct) than the former performance (0.938 correct), $t(78) = 4.089$, $p < .001$, Cohen's $d = 0.914$. We also showed that memory performance declined from the first three blocks of the experiment to the blocks labeled -1, 0, and 1 in the figures by 0.086 in the concurrent condition, but only by 0.008 in the control condition, $t(78) = 3.327$, $p = .001$, Cohen's $d = 0.744$. These analyses converge in their intuitive implications with the conclusions of the GLM analysis above. In Experiment 2, memory performance was sacrificed to free the resources needed to reconstrue the matching task in the concurrent condition.

General discussion

We adopted the RMTS task—in which participants express a basic form of relational matching—to understand better its cognitive character. Our bistable task gave participants an easy, perceptual-similarity entry into the task. There was no need for relational cognition—at first. Our task gives a look at humans' trying to break the perceptual-conceptual barrier, with and without a memory load. The load sharply reduced their ability to do so. In one experiment, half as many participants made the perceptual-conceptual transition. In a second experiment, participants had to drop their memory load to solve the RMTS task. This result suggests that the perceptual-conceptual transition depends on the task's relational reconstrual. In turn, the reconstrual evidently depends on working-memory processes. The epoch of the reconstrual was marked even in the performance of our strongest participants (see Fig. 6).

This result complements other findings on relational cognition. In verbal analogy and analogical reasoning tasks also, working-memory interference impairs performance (Morrison, Holyoak, & Truong, 2001; Waltz, Lau, Grewal, & Holyoak, 2000). For example, confirming analogies like noise:silence :: light:dark is impaired by manipulations that compete for working-memory resources (Morrison et al., 2001; Waltz, et al., 2000). These semantic-conceptual-language tasks are very different from our purely perceptual task drawn from the comparative literature, but perhaps they make a related point about human cognition. However, it is important to see that reconstruing a matching task onto the

relational level may make demands on cognitive resources even in a perceptual task when language is not at issue.

In one respect, our focus was different from that in the broader relational literature. We focused on the cognitive transition by which participants discover the task's relational solution. We did not focus on the working-memory resources expended as participants perform individual RMTS trials relationally. However, our results clearly show that the procedures of relational matching, once discovered, occur fluently even if working-memory resources are occupied. A full theoretical perspective needs to accommodate both the difficulty of initial discovery and the ease of subsequent execution.

Our findings fit perfectly within a current theoretical framework within neuroscience (e.g., Ashby & Ell, 2001; Smith & Church, 2018). Related research has also shown that brain structures involved in working memory (Glahn et al., 2002; Goldman-Rakic, 1987) support analogy tasks (Waltz et al., 1999; Wharton et al., 2000). Within this framework, a reasonable hypothesis about our results is that conceptual reconstrual is linked to humans' capacity for explicit-declarative cognition. This explicit system comprises executive attention (Posner & Petersen, 1990) and working memory (Fuster, 1989; Goldman-Rakic, 1987), capacities that would support rule formation and hypothesis testing (Brown & Marsden, 1988; Cools et al., 1984; Elliott & Dolan, 1998; Kolb & Whishaw, 1990; Rao et al., 1997; Robinson, Heaton, Lehman, & Stilson, 1980). This system learns by testing hypotheses. It learns rules that participants can describe verbally. These aspects of explicit cognition would serve humans well as they try to reconstrue a matching task.

Beyond humans, comparative psychologists try to understand the sharp species difference in relational cognition and to interpret those differences theoretically. In our view, it is important in this effort to enlist the cross-talk and synergy to be derived from related human research and theory. It is remarkable how this cross-talk has fallen silent. Accordingly, we consider our human result in the context of the human-animal species difference.

Fagot et al. (2001) illustrated the relevant species difference. They studied relational matching in humans and baboons using same and different stimulus arrays—not stimulus pairs—of clip-art icons. Through elegant stimulus manipulations, they were able to vary the degree of visual entropy (visual jazziness) presented by the arrays, so that entropy varied along a continuum. Humans adopted a highly restrictive, categorical *same* criterion (see Fig. 11, top). Only arrays with essentially zero visual entropy motivated *same* responses. By this qualitative criterion, humans labeled as *different* arrays with any *discernible entropy*. Monkeys (see Fig. 11, bottom) adopted a generous *same* criterion. They slowly relinquished *same* responding only as entropy increased to higher levels. The question is how we should characterize the different

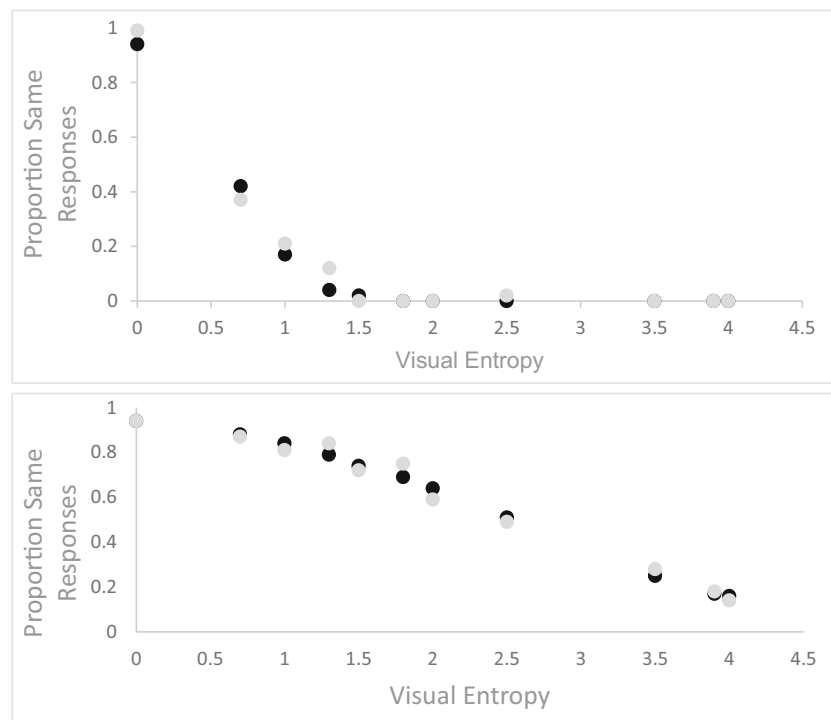


Fig. 11 Top: Observed performance (black symbols) of Human H01 in the relational-matching task of Fagot et al. (2001). The proportion of same responses is plotted against the level of visual entropy in the stimulus sample for the trial (a measure of the extent to which the 16 clip-art icons in the 16-item array were visually variable). The best-fitting predictions of Fagot et al.’s formal model are also shown (gray symbols). Bottom: Observed performance of Baboon B03 in the equivalent relational-

matching task, depicted in the same way. Adapted from “Discriminating the relation between relations: The role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*)” by J. Fagot, E. A. Wasserman., & M. E. Young, 2001, *Journal of Experimental Psychology: Animal Behavior Processes*, 27, 316–328. Reprinted with permission

cognitive systems that lie behind these two different performance patterns?

Smith, Redford, Haas, et al. (2008) found a strikingly converging finding. Humans and monkeys had to judge whether two complex polygon shapes were the same or different. They systematically varied the perceptual disparity between stimuli—to find both species’ disparity threshold for reporting that shapes were different. Figure 12 shows again that humans adopted a categorical, rule-based criterion for labeling stimulus pairs same. They distinguished zero-disparity pairs (same) from pairs *with any discernible disparity* (different). Monkeys adopted a generous, inclusive criterion for the stimulus pairs they would label same. They distinguished higher-similarity pairs (same) from lower-similarity pairs (different).

The dominant interpretation of these results is low level, parametric, and based in associative-learning theory. The idea is that monkeys and humans respond in relational tasks to the same, continuously varying perceptual cues of visual entropy or physical disparity, that there is a continuity across species of the basic processes involved, that the underlying cognitive capacity exhibited by both species is about the same, and that the species difference emerges because humans’ setting for the same–different criterion parameter is lower, tighter, more exclusive—yielding a highly restrictive same concept. This

idea has a powerful draw for several reasons. It emphasizes continuities between monkeys’ and humans’ minds, potentially providing insights about primate evolution and human emergence. It lets the minds of monkeys and humans be understood parsimoniously as the same discriminating, criterion-setting system that associative-learning theorists understand so well. It makes the minds of both species unitary, a preferred simplifying assumption in cognitive science and particularly in the area of animal learning. That is, one need not invoke the layering on of a qualitatively higher level of relational cognition with which comparative psychologists are not so comfortable. This unitary explanation is encouraged by the fitting of formal models to the performance of humans and monkeys. The result of that modeling is also shown in the two figures (gray symbols). The same formal model can be used to fit the performance of both species, with the difference being that the criterion parameter that separates same and different response regions is placed far differently.

The continuity idea was also expressed in an important theoretical article (Goldstone & Barsalou, 1998). In Goldstone and Barsalou’s (1998) theoretical reuniting of perceptual and conceptual processing, they found many continuities between these levels of information processing (e.g., even abstract properties of things can be represented quasiperceptually in analog

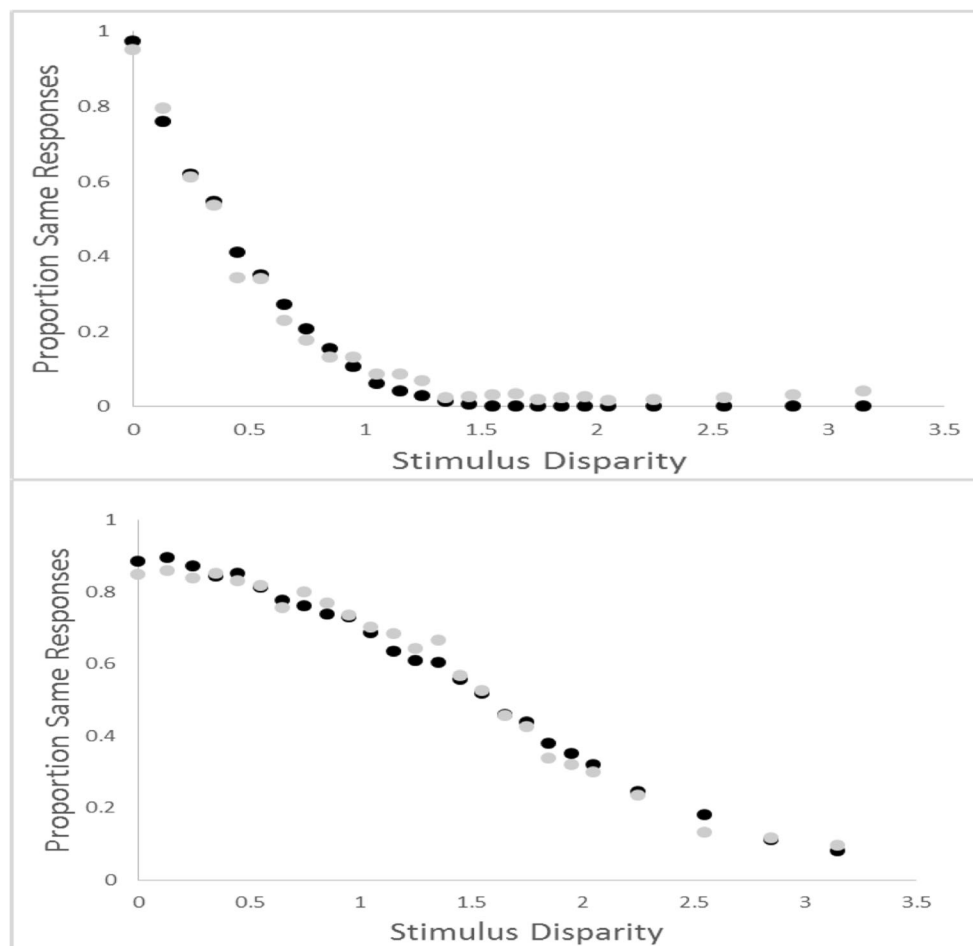


Fig. 12 Top: Observed performance (black symbols) of humans in the same-different task of Experiment 1b in Smith, Redford, Hass, et al. (2008). The proportion of “same” responses is plotted against stimulus disparity (a measure of the extent to which the vertices of two complex polygons are different in coordinate space). The best-fitting predictions of a standard signal-detection formal model are also shown (gray symbols). Bottom: Observed performance of Monkey Lou in the equivalent same-

different task, depicted in the same way. From “The comparative psychology of same-different judgments by humans (*Homo sapiens*) and monkeys (*Macaca mulatta*),” by J. D. Smith, J. S. Redford, S. M. Haas, M. V. C. Coutinho, & J. J. Couchman, 2008, *Journal of Experimental Psychology: Animal Behavior Processes*, 34, 365–370. Reprinted with permission

fashion; even abstract concepts sometimes have perceptual origins). They pointed to important commonalities and shared mechanisms between perception and conception.

However, the present results from humans suggest an alternative theoretical explanation. They suggest that reframing a task relationally is mediated by qualitatively separate cognitive processes akin to those generally labeled as explicit and declarative. Humans with working-memory loads may struggle to break the perceptual-conceptual barrier because the mediation of these explicit processes is then prevented. Likewise, monkeys may generally fail on RMTS tasks because they do not apply the hypothesis-reframing processes that would let them reconstrue the matching task. There could be different reasons for this failure. This theoretical approach would grant monkeys and concurrent humans interesting processing similarities (e.g., both would be stuck processing the task using first-order perceptual cues). This explanation has crucial theoretical differences from current theoretical descriptions in the area.

Goldstone and Barsalou (1998) expressed clear comfort with this alternative view. They likened (p. 243) the perception/conception distinction to the associative/rule distinction. They judged that the perceptual (associative) processes would change slower, be relatively automatic, use parallel processing and diffuse attention, and require practice and repetition. They judged that the conceptual (rule) processes would be labile, voluntarily controlled, serial in nature, sharply attending, and available with ad hoc immediate flexibility. These information-processing distinctions fit well with our theoretical understanding of the cognitive differences between perceptual matching processes in our RMTS task and the reframing processes that let humans discover the alternative, relational solution.

Ultimately, given all these information-processing distinctions, it becomes theoretically important to understand perceptual matching and humans’ relational discovery processes quite separately, probably underlain by different neural systems and possibly existing on different levels of cognitive

awareness. One can even envision the cognitive-neuroscience studies that could pursue this theoretical separation (e.g., Davis, Goldwater, & Giron, 2017). Thus, taking a perspective from the present results, and from theoretical perspectives in neuroscience, the species difference could turn out to be qualitative. Then, the species difference would no longer represent just two different parameter settings of the same discriminatory system. One could not simply dial M for monkey, or H for human, retuning the same processing system. Rather, the human and monkey modes of performance would be fundamentally different—perceptual and relational, associative and cognitive, procedural and declarative, implicit and explicit. Notice that this difference in views even goes to our understanding of the evolutionary emergence for human relational cognition. In one view, the same processing system would sharpen or tighten its criterial values, yielding humans' highly restrictive concepts of *same*. In the other view, the emergence of relational cognition would represent the layering on of a higher, explicit system that handles the task of relational reconstrual.

What might this explicit system be like representationally? On the one hand, working memory could be sufficient to serve these reconstrual processes, without language's intervention. That is, there might be prelinguistic relational processes (Ferry et al., 2015). Working memory would let multiple stimuli be represented simultaneously for comparison and relational judgment. It would let hypothesized task solutions be maintained while evaluated. On the other hand, humans' reconstrual processes that need working memory might also be facilitated by language. That is, language might be an important part of humans' relational toolkit (e.g., Gentner, 2016). We do not decide this issue here—indeed, our results do not resolve it. But these issues are crucial to theory in comparative research and to understanding humans' evolutionary emergence. We also caution against linking higher-level cognitive functions too closely to language. This linkage leaves unstudied what language allows and what is allowed in its absence. It also may be incorrectly exclusionary to other species.

We stress that we do not believe that animals, especially Old-World primates, are qualitatively lacking in their capacity to process relations or even to match relations. They continue to reveal intriguing glimmers of relational cognition (Fagot & Maugard, 2013; Flemming, Thompson, & Fagot, 2013; Obozova, Smirnova, Zorina, & Wasserman, 2015; Martinho & Kacelnik, 2016; Maugard, Marzouki, & Fagot, 2013; Pepperberg, 2013; Smirnova, Zorina, Obozova, & Wasserman, 2015; Vonk, 2003). These findings make the problem of relational cognition's evolutionary emergence far more interesting. Instead, this is what we are saying. As animals show these glimmers, it is not because a task or manipulation moves some cognitive rheostat continuously, as a volume control for relational cognition. Rather, it is possibly because one is successfully engaging in matching tasks' different neural structures,

different brain circuits, and different levels of cognition that especially require working memory.

In fact, the present human research provides empirical hints to comparative researchers trying to foster primates' relational-matching capacity. They should focus their efforts on supporting and easing animals' task reconstruals by which they discover the relational solution. For example, if one could provide monkeys with well-trained abstract symbols that connote same and different, this might strengthen their relational coding of the stimulus pairs, making the task's relational solution more accessible and salient.

We acknowledge that our one study cannot resolve these long-debated comparative issues, but we think that the difference in views is an intriguing one for future research to resolve. Smith and Church (2018) explained in more detail why further theoretical consideration of these issues is timely and essential to the next stage in the theoretical development of comparative psychology. Our hope here is that the present article shows why comparative *and* cognitive psychologists have equal standing to make a strong contribution to this area and why the theoretical discourse becomes stronger and richer through systematic cross-talk and interaction.

Author note The preparation of this article was supported by Grants HD093690 and HD060563 from NICHD. We thank the research assistants in the Complex Cognition Lab at Georgia State University for help with data collection.

Compliance with ethical standards

Conflict of interest The authors declare no conflicts of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442–481. <https://doi.org/10.1037/0033-295X.105.3.442>
- Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences*, *5*, 204–210. [https://doi.org/10.1016/S1364-6613\(00\)01624-7](https://doi.org/10.1016/S1364-6613(00)01624-7)
- Brooks, D. I., & Wasserman, E. A. (2008). Same/different discrimination learning with trial-unique stimuli. *Psychonomic Bulletin & Review*, *15*, 644–650. <https://doi.org/10.3758/PBR.15.3.644>
- Brown, R. G., & Marsden, C. D. (1988). Internal versus external cues and the control of attention in Parkinson's disease. *Brain*, *111*, 323–345. <https://doi.org/10.1093/brain/111.2.323>
- Carter, D. E., & Werner, T. J. (1978). Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior*, *29*, 565–601. <https://doi.org/10.3758/LB.37.2.204>
- Castro, L., & Wasserman, E. A. (2013). Humans deploy diverse strategies in learning same-different discrimination tasks. *Behavioural Processes*, *93*, 125–139. <https://doi.org/10.1016/j.beproc.2012.09.015>

- Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science*, 38, 383–397. <https://doi.org/10.1111/cogs.12099>
- Cools, A. R., van den Bercken, J. H., Horstink, M. W., van Spaendonck, K. P., & Berger, H. J. (1984). Cognitive and motor shifting aptitude disorder in Parkinson's disease. *Journal of Neurological and Neurosurgical Psychology*, 47, 443–453. Retrieved from <http://jnnp.bmj.com/content/jnnp/47/5/443.full.pdf>
- Cumming, W. W., & Berryman, R. (1961). Some data on matching behavior in the pigeon. *Journal of the Experimental Analysis of Behavior*, 4, 281–284. <https://doi.org/10.1901/jeab.1961.4.281>
- D'Amato, M. R., & Columbo, M. (1989). On the limits of the matching concept in monkeys (*Cebus apella*). *Journal of the Experimental Analysis of Behavior*, 52, 225–236. <https://doi.org/10.1901/jeab.1989.52.225>
- D'Amato, M. R., Salmon, D. P., & Columbo, M. (1985). Extent and limits of the matching concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 35–51. <https://doi.org/10.1037/0097-7403.11.1.35>
- Davis, T., Goldwater, M., & Giron, J. (2017). From concrete examples to abstract relations: The rostrolateral prefrontal cortex integrates novel examples into relational categories. *Cerebral Cortex*, 27, 2652–2670. <https://doi.org/10.1093/cercor/bhw099>
- Elliott, R., & Dolan, R. J. (1998). Activation of different anterior cingulate foci in association with hypothesis testing and response selection. *NeuroImage*, 8, 17–29. <https://doi.org/10.1006/nimg.1998.0344>
- Fagot, J., & Maugard, A. (2013). Analogical reasoning in baboons (*Papio papio*): Flexible reencoding of the source relation depending on the target relation. *Learning & Behavior*, 41, 229–237. <https://doi.org/10.3758/s13420-012-0101-7>
- Fagot, J., & Parron, C. (2010). Relational matching in baboons (*Papio papio*) with reduced grouping requirements. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 184–193. <https://doi.org/10.1037/a0017169>
- Fagot, J., & Thompson, R. K. R. (2011). Generalized relational matching by guinea baboons (*Papio papio*) in two-by-two-item analogy problems. *Psychological Science*, 22, 1304–1309. <https://doi.org/10.1177/0956797611422916>
- Fagot, J., Wasserman, E. A., & Young, M. E. (2001). Discriminating the relation between relations: The role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes*, 27, 316–328. <https://doi.org/10.1037/0097-7403.27.4.316>
- Farthing, G. W., & Opuda, M. J. (1974). Transfer of matching-to-sample in pigeons. *Journal of the Experimental Analysis of Behavior*, 21, 199–213. <https://doi.org/10.1901/jeab.1974.21.199>
- Ferry, A. L., Hespos, S. J., & Gentner, D. (2015). Prelinguistic relational concepts: Investigating analogical processing in infants. *Child Development*, 86, 1386–1405. <https://doi.org/10.1111/cdev.12381>
- Flemming, T. M., Thompson, R. K. R., Beran, M. J., & Washburn, D. A. (2011). 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*, 37, 353–360. <https://doi.org/10.1037/a0022142>
- Flemming, T. M., Thompson, R. K. R., & Fagot, J. (2013). Baboons, like humans, solve analogy by categorical abstraction of relations. *Animal Cognition*, 16, 519–524. <https://doi.org/10.1007/s10071-013-0596-0>
- Fujita, K. (1982). An analysis of stimulus control in two-color matching-to-sample behaviors of Japanese monkeys (*Macaca fuscata*). *Japanese Psychological Research*, 24, 124–135. <https://doi.org/10.1007/s10071-009-0217-0>
- Fuster, J. M. (1989). *The prefrontal cortex* (2nd ed.). Philadelphia, PA: Lippincott-Raven.
- Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–235). Cambridge, MA: MIT Press.
- Gentner, D. (2016). Language as cognitive tool kit: How language supports relational thought. *American Psychologist*, 71(8), 650–657. <https://doi.org/10.1037/amp0000082>
- Glahn, D. C., Kim, J., Cohen, M. S., Poutanen, V. P., Therman, S., Bava, S., ... Cannon, T. D. (2002). Maintenance and manipulation in spatial working memory: Dissociations in the prefrontal cortex. *NeuroImage*, 17, 201–213. <https://doi.org/10.1006/nimg.2002.1161>
- Goldman-Rakic, P. S. (1987). Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In F. Plum & V. Mountcastle (Eds.), *Handbook of physiology* (pp. 373–417). Bethesda, MD: American Physiological Society.
- Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, 65, 231–262. [https://doi.org/10.1016/S0010-0277\(97\)00047-4](https://doi.org/10.1016/S0010-0277(97)00047-4)
- Grafman, J., & Litvan, I. (1999). Importance of deficits in executive functions. *Lancet*, 354, 1921–1923. [https://doi.org/10.1016/S0140-6736\(99\)90438-5](https://doi.org/10.1016/S0140-6736(99)90438-5)
- Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences*, 14, 497–505. <https://doi.org/10.1016/j.tics.2010.08.005>
- Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition*, 37, 133–166. [https://doi.org/10.1016/0010-0277\(90\)90021-b](https://doi.org/10.1016/0010-0277(90)90021-b)
- Hochmann, J. R., Mody, S., & Carey, S. (2016). Infants' representations of same and different in match- and non-match-to-sample tasks. *Cognitive Psychology*, 86, 87–111. <https://doi.org/10.1016/j.cogpsych.2016.01.005>
- Hochmann, J. R., Tuerk, A. S., Sanborn, S., Zhu, R., Long, R., & Carey, S. (2017). Children's representation of abstract relations in relational/array match-to-sample tasks. *Cognitive Psychology*, 99, 17–43. <https://doi.org/10.1016/j.cogpsych.2017.11.001>
- Holmes, P. W. (1979). Transfer of matching performance in pigeons. *Journal of the Experimental Analysis of Behavior*, 31, 103–114. <https://doi.org/10.1901/jeab.1979.31.103>
- Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, 110, 220–264. <https://doi.org/10.1037/0033-295X.110.2.220>
- James, W. (1890/1950). *The principles of psychology*. New York, NY: Dover. (Original work published 1890)
- Katz, J. S., Wright, A. A., & Bachevalier, J. (2002). Mechanisms of same/different abstract-concept learning by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, 28, 358–368. <https://doi.org/10.1037/0097-7403.28.4.358>
- Kolb, B., & Wishaw, I. Q. (1990). *Fundamentals of human neuropsychology* (3rd ed.). New York, NY: Freeman.
- Locke, J. (1690). *An essay concerning human understanding*. Philadelphia, PA: Troutman & Hayes.
- Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes*, 66, 309–332. <https://doi.org/10.1016/j.beproc.2004.03.011>
- Martinho, A. I., II., & Kacelnik, A. (2016). Ducklings imprint on the relational concept of “same or different?”. *Science*, 353 (6296), 286–288. <https://doi.org/10.1126/science.aaf4247>
- Maugard, A., Marzouki, Y., & Fagot, J. (2013). Contribution of working memory processes to relational matching-to-sample performance in baboons (*Papio papio*). *Journal of Comparative Psychology*, 127, 370–379. <https://doi.org/10.1037/a0032336>
- Morrison, R. G., Holyoak, K. J., & Truong, B. (2001). Working memory modularity in analogical reasoning. *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, 23, 663–668. Retrieved from <https://escholarship.org/uc/item/90r5n01r>

- Obozova, T., Smirnova, A., Zorina, Z., & Wasserman, E. (2015). Analogical reasoning in amazons. *Animal Cognition*, *18*, 1363–1371. <https://doi.org/10.1007/s10071-015-0882-0>
- Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*, 109–178. <https://doi.org/10.1017/S0140525X08003543>
- Pepperberg, I. M. (2013). Abstract concepts: Data from a grey parrot. *Behavioural Processes*, *93*, 82–90. <https://doi.org/10.1016/j.beproc.2012.09.016>
- Posner, M. I., Goldsmith, R., & Welton, K. E. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology*, *73*, 28–38. <https://doi.org/10.1037/h0024135>
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. <https://doi.org/10.1146/annurev.ne.13.030190.000325>
- Premack, D. (1978). On the abstractness of human concepts: Why it would be difficult to talk to a pigeon. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 423–451). Hillsdale, NJ: Erlbaum.
- Rao, S. M., Bobholz, J. A., Hammeke, T. A., Rosen, A. C., Woodley, S. J., Cunningham, J. M., ... Binder, J. R. (1997). Functional MRI evidence for subcortical participation in conceptual reasoning skills. *Neuroreport*, *27*, 1987–1993. <https://doi.org/10.1097/00001756-199705260-00038>
- Robinson, A. L., Heaton, R. K., Lehman, R. A. W., & Stilson, D. W. (1980). The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *Journal of Consulting and Clinical Psychology*, *48*, 605–614. <https://doi.org/10.1037/0022-006X.48.5.605>
- Rumbaugh, D. M., & Pate, J. L. (1984). The evolution of cognition in primates: A comparative perspective. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 569–585). Hillsdale, NJ: Erlbaum.
- Seger, C. A., & Miller, E. K. (2010). Category learning in the brain. *Annual Review of Neuroscience*, *33*, 203–219. <https://doi.org/10.1146/annurev.neuro.051508.135546>
- Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *Journal of Experimental Psychology: General*, *126*, 147–164. [https://doi.org/10.1016/S0010-0277\(96\)00726-3](https://doi.org/10.1016/S0010-0277(96)00726-3)
- Smirnova, A., Zorina, Z., Obozova, T., & Wasserman, E. (2015). Crows spontaneously exhibit analogical reasoning. *Current Biology*, *25*, 256–260. <https://doi.org/10.1016/j.cub.2014.11.063>
- Smith, J. D., Berg, M. E., Cook, R. G., Murphy, M. S., Crossley, M. J., Boomer, J., ... Grace, R. C. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience and Biobehavioral Reviews*, *36*, 2355–2369. <https://doi.org/10.1016/j.neubiorev.2012.09.003>
- Smith, J. D., Boomer, J., Zakrzewski, A. C., Roeder, J. L., Church, B. A., & Ashby, F. G. (2014). Deferred feedback sharply dissociates implicit and explicit category learning. *Psychological Science*, *25*, 447–457. <https://doi.org/10.1177/0956797613509112>
- Smith, J. D., & Church, B. A. (2018). Dissociable learning processes in comparative psychology. *Psychonomic Bulletin & Review*, *25*, 1565–1584. <https://doi.org/10.3758/s13423-017-1353-1>
- Smith, J. D., Flemming, T. M., Boomer, J., Beran, M. J., & Church, B. (2013). Fading perceptual resemblance: A path for rhesus macaques (*Macaca mulatta*) to conceptual matching? *Cognition*, *129*, 598–614. <https://doi.org/10.1016/j.cognition.2013.08.001>
- Smith, J. D., & Minda, J. P. (2001). Journey to the center of the category: The dissociation in amnesia between categorization and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 984–1002. <https://doi.org/10.1037/0278-7393.27.4.984>
- Smith, J. D., & Minda, J. P. (2002). Distinguishing prototype-based and exemplar-based processes in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 800–811. <https://doi.org/10.1037/0278-7393.28.4.800>
- Smith, J. D., Redford, J. S., & Haas, S. M. (2008). Prototype abstraction by monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: General*, *137*, 390–401. <https://doi.org/10.1037/0096-3445.137.2.390>
- Smith, J. D., Redford, J. S., Haas, S. M., Coutinho, M. V. C., & Couchman, J. J. (2008). The comparative psychology of same-different judgments by humans (*Homo sapiens*) and monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, *34*, 361–374. <https://doi.org/10.1037/0097-7403.34.3.361>
- Smith, J. D., Zakrzewski, A. C., Johnson, J. M., Valteau, J. C., Church, B. A. (2016). Categorization: The view from animal cognition. *Behavioral Science*, *6*, 12. <https://doi.org/10.3390/bs6020012>
- Vonk, J. (2003). Gorilla (*Gorilla gorilla gorilla*) and Orangutan (*Pongo abelii*) understanding of first and second order relations. *Animal Cognition*, *6*, 77–86. <https://doi.org/10.1007/s10071-003-0159-x>
- Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., ... Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science*, *10*, 119–125. <https://doi.org/10.1111/1467-9280.00118>
- Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition*, *28*, 1205–1212. <https://doi.org/10.3758/BF03211821>
- Washburn, D. A., & Rumbaugh, D. M. (1991). Rhesus monkey (*Macaca mulatta*) complex learning skills reassessed. *International Journal of Primatology*, *12*, 377–388. <https://doi.org/10.1007/BF02547618>
- Wasserman, E. A., Fagot, J., & Young, M. (2001). Same-different conceptualization by baboons (*Papio papio*): The role of entropy. *Journal of Comparative Psychology*, *115*, 42–52. <https://doi.org/10.1037/0735-7036.115.1.42>
- Wasserman, E. A., & Young, M. (2010). Same-different discrimination: The keel and backbone of thought and reasoning. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*, 3–22. <https://doi.org/10.1037/a0016327>
- Wharton, C. M., Grafman, J., Flitman, S. S., Hansen, E. K., Brauner, J., Marks, A., & Honda, M. (2000). Toward neuroanatomical models of analogy: A positron emission tomography study of analogical mapping. *Cognitive Psychology*, *40*, 173–197. <https://doi.org/10.1006/cogp.1999.0726>
- Wright, A. A., Cook, R. G., & Kendrick, D. (1989). Relational and absolute stimulus learning by monkeys in a memory task. *Journal of the Experimental Analysis of Behavior*, *52*, 237–248. Retrieved from http://nba.uth.tmc.edu/homepage/wright/Assets/pdf/relat_absolute_1989.pdf
- Wright, A. A., Rivera, J., Katz, J. S., & Bachevalier, J. (2003). Abstract concept learning and list-memory processing by capuchin and rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, *29*, 184–198. <https://doi.org/10.1037/0097-7403.29.3.184>
- Wright, A. A., Santiago, H., & Sands, S. (1984). Monkey memory: Same/different concept learning, serial probe acquisition, and probe delay effects. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 513–529. Retrieved from http://nba.uth.tmc.edu/homepage/wright/Assets/pdf/monk_mem_SD_1984.pdf
- Wright, A. A., Shyan, M. R., & Jitsumori, M. (1990). Auditory same/different concept learning by monkeys. *Animal Learning and Behavior*, *18*, 287–294. Retrieved from http://nba.uth.tmc.edu/homepage/wright/Assets/pdf/aud_sd_monkeys_1990.pdf
- Young, M. E., Wasserman, E. A., & Garner, K. L. (1997). Effects of number of items on the pigeon's discrimination of same from different visual displays. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, 491–501. <https://doi.org/10.1037/0097-7403.23.4.491>