Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Current Opinion in
Behavioral Sciences

Population differences in performance on Relational Match to Sample (RMTS) sometimes reflect differences in inductive biases alone[☆]

Ivan Kroupin and Susan Carey

A much replicated finding is that only humans above the age of five succeed spontaneously on Premack's Relational Match to Sample task, which has been widely used in the comparative and developmental literatures to probe relational reasoning capacities. We review four different types of explanations for the failures of young children and non-human animals, two that posit capacity limitations and two that posit differences in learning histories alone. We review training studies that rule out capacity limitations, at least for crows, parrots, four-year-old children, and a variety of primate species. Finally, we review recent studies demonstrating that population differences sometimes reflect differences in inductive biases alone and discuss the crucial importance of inductive biases in relational reasoning.

Address
Harvard University, United States

Corresponding author: Kroupin, Ivan (ikroupin@g.harvard.edu)

Current Opinion in Behavioral Sciences 2021, **37**:75–83

This review comes from a themed issue on **Same-different conceptualization**

Edited by **Jean-Remy Hochmann, Ed Wasserman,** and **Susan Carey**

For a complete overview see the Issue and the Editorial

Available online 4th December 2020

https://doi.org/10.1016/j.cobeha.2020.11.006

2352-1546/© 2020 Elsevier Ltd. All rights reserved.

Introduction

One important goal of comparative psychology is to document capacity differences among species, accounting for them by characterizing the selection pressures that lead one species to have a given capacity that another lacks. Furthermore, centuries of debates in philosophy, biology, and psychology have concerned what, if any, fundamental cognitive capacity differences distinguish humans from other animals, allowing the cultural accumulation of knowledge and technology [1–7]. Parallel questions arise concerning ontogenesis: Does cognitive development involve the maturation of new capacities, making possible thoughts and reasoning processes that were not available to younger children?

Empirical investigations of the evolutionary and ontogenetic origins of relational reasoning have been an important case study in these debates. Relational reasoning – the ability to compare different sets of items in terms of the relations which hold between the objects in each set – underpins human achievements like scientific analogies and artistic metaphors and plays an important role in everyday language. All of these, in turn, are crucial to enabling our uniquely human cognitive niche [8–10].

Clearly, young children and non-human animals do not speak in metaphors or engage in mathematical modelling. The key question is what separates these groups from adult humans whose repertoire includes such feats of relational reasoning, feats that enable the human cognitive niche. Premack [11•] devised the Relational Match to Sample task (RMTS) to explore population differences in arguably the most basic case of analogical mapping — seeing the analogy between X X and Y Y on the basis of each pair instantiating the relations same, and seeing the analogy between A B and C D on the basis of each pair instantiating the relation different. Humans over the age of about five in industrialized, schooled, populations succeed without training on RMTS [11•,12•]. In contrast, younger children and non-human animals often persistently fail, even with error feedback, the latter sometimes even after tens of thousands of training trials ([12•], see below and Ref. [13•] for a review). The outstanding questions, then, are: *Why* do certain populations fail on this basic assessment of relational reasoning, and *What* do these failures reveal about the emergence of relational reasoning in general and concepts of sameness and difference specifically across phylogeny and ontogeny?

Plan of the present paper

First, we review four different proposals that have been offered to explain the differences between populations in performance on RMTS, two hypothesizing capacity limitations of animals and young children, and two hypothesizing differential learning histories. Next, we review research from training studies that shows that population differences do not *always* derive from capacity differences. Finally, we show that that in some cases population

---

[☆] The authors would thank the McDonnell Foundation network grant 'The Ontogenetic Origins of Abstract Combinatorial Thought' for support.

differences derive from differences in inductive biases *alone*. We endorse all four accounts as explaining *some specific* population differences; the challenge is specifying which apply where, and filling in the details of each.

Explaining population differences: capacity limitations

Many researchers, in the tradition of philosophers such as Descartes, Davidson, Bermudez [1–3], propose evolutionary and ontogenetic discontinuities in fundamental cognitive *capacities* — such that individuals of only some species beyond certain points in development are *capable* of relational reasoning [e.g. Refs. 6,14]. Two types of capacities have been implicated: capacities for representations of relations with specified properties, and/or computational capacities necessary to deploy available representations in the service of relational reasoning.

Account 1: representational capacities

As one example of a capacity limitation account, consider Penn *et al.*'s, [6] explanation of the failure on RMTS in terms of limitations in the nature of representations which are even potentially available to those individuals who fail. Their proposal is that animal populations failing RMTS entirely *lack the capacity to generate* representations (mental symbols) for the relations same and different 'which are (1) independent of any particular source of stimulus control, and (2) available to serve in a variety of further higher-order inferences in a systematic fashion' (p. 110; see Ref. [15•], for detailed proposals for how mental representations of sameness and difference might vary across populations).

Account 2: computational capacities

A different type of capacity limitation account describes failures in relational reasoning as a result of insufficient computational resources. These include sufficient working memory to contain the to-be-compared relations and at the same time [e.g. Ref. 16] as well as sufficient inhibitory control to prevent attention to non-relational properties of the comparands [e.g. Ref. 17]. Consistent with this account, populations clearly differ in such executive function capacities, both across human development [18], and between species [19].

Explaining population differences: differences in learning histories

A second *kind* of account aligns with thinkers like Darwin and Fodor [4,5] in characterizing population differences on tasks like RMTS as not necessarily reflective of an absolute limit in the *capacity* for relational reasoning. Instead, population differences may sometimes reflect differences in individuals' learning histories alone.

Account 3: lacking specific representations, in spite of the capacity to generate them

Many relational reasoning tasks require creating new specific representations, ones that the individual has never represented before. With respect to RMTS, individuals need not only the *capacity* to create abstract representations of same and different with the right properties to support success, but also to have *gone through the process of actually generating these representations*. There are many accounts of the process through which new specific representations are created, for example, bootstrapping mechanisms [20], abstraction through progressive alignment ([21•]; see below), conceptual combination of existing representations [22]. Animals and young children may have the representational capacities to generate abstract representations of sameness and difference, but nonetheless lack the specific learning experiences to have done so.

Account 4: differences in inductive biases

All tasks involve inductive inference; the participant must decide from the infinity of features of the stimuli before them *which* are relevant to the decision at hand. Inductive inference must be heavily constrained by inductive biases such that, in any given decision, a participant considers only a tiny subset of those features. Thus, if an individual has not developed inductive biases which would lead them to infer sameness and difference as relevant in the context of RMTS, then they will fail the task *even if* they are *able* to engage in relational reasoning *and already have* the necessary abstract representations.

Evidence from training studies: a representative sample

Despite repeated evidence of failure on RMTS by nonhuman animals and young children, various training regimes have produced success on RMTS in many of these populations. So long as success is not a result of learning to choose correct matches without using relational reasoning (e.g. matching by entropy [13•]), it is in apparent contradiction of accounts (1 and 2) which posit fundamental limitations in capacity.

'Dogged training'

The most extensively used RMTS training paradigm has been to provide individuals with correct/incorrect feedback on RMTS trials (over weeks or months, up to 60 000 trials). This has produced generalizable success on RMTS by some ape and monkey species, albeit generally in a minority of the individual primates so trained (e.g. Refs. [23–26]).

The sheer number of trials requires that stimulus triads must be repeated, leaving open the possibility that individuals were learning the choices that were correct on *specific stimulus triads* in training and may *not* have succeeded via relational matching [13•]. In line with this possibility, the majority of individual primates who were above chance on the training stimuli failed to generalize success to new stimulus triads. However, the fact that a minority of subjects succeeded in generalizing to new stimuli (albeit with a drop in performance) suggests that dogged training can in fact produce RMTS success in certain primates.

One possibility is that extensive training led to the generation of new abstract representations in the animals who generalized (Account 3) — perhaps in addition to learning the correct responses for specific trained triads. Another possibility is that those animals that succeeded *already had* the necessary representations and extensive training led to a gradual adjustment of primates' initial hypotheses regarding the correct bases of matching to favor these representations as the most likely bases of matching (Account 4).

Symbol training in chimpanzees

Premack [6] showed chimpanzees without special training fail RMTS. In contrast, Sarah — a chimp who had extensive training using tokens as communicative symbols, including tokens for same and different — succeeded (see also Refs. [27,28•]). Learning a new arbitrary (i.e. non-iconic) summary symbol for the concept *same* may constitute having generated a new abstract representation with new computational affordances [9,27] (Account 3). Alternatively such training may simply make *existing* relational representations of same and different more salient in the context of other tasks, such as RMTS (Account 4).

In summary, successful training with rhesus macaques, baboons, chimpanzees, parrots and crows (see below) rule out accounts (1 and 2) that posit absolute capacity limitations that preclude relational reasoning in these species. However, these results leave open both Accounts 3 (the absence of representations of specific relations) and 4 (different inductive biases) as explanations for the initial failures of individuals from these species on RMTS.

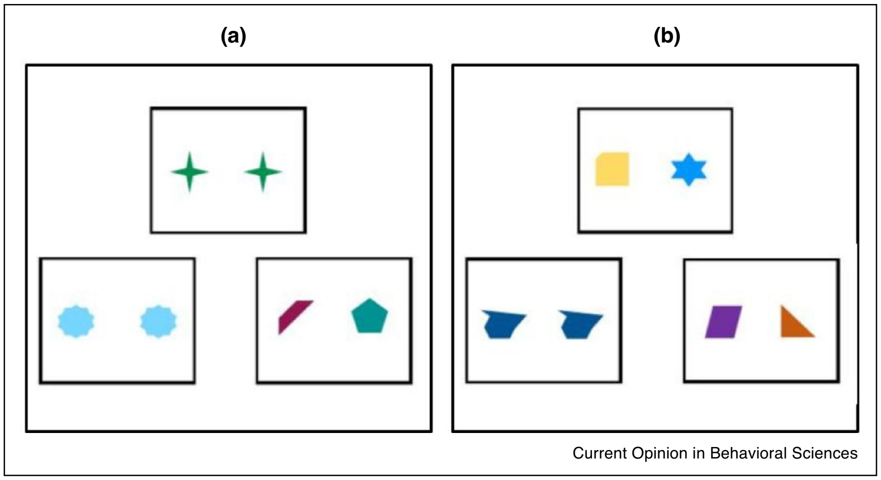
Progressive alignment

Kotovsky and Gentner [21•] pioneered a paradigm which has produced success on RMTS (albeit with relations other than same/different) by initially training children to match *both* on object features *and* relations. For RMTS, progressive alignment trials have the structure — 'Does AA or BC go with AA?'. When the object feature matches are removed, the task becomes a standard RMTS task (Does DE or FF go with GG? Figure 1; see Ref. [29•] for the role of progressive alignment in infants' success in tasks requiring representations of same or different, albeit not relational mapping). As in successful training studies with animals, success on the basis of progressive alignment rules out capacity limitation accounts (1 and 2) of children's failures on RMTS. Gentner *et al.* ([21•], also see Ref. [30]) have suggested that progressive alignment training works by producing a *new abstract relational representation* in the trainees (Account 3). However, this paradigm may also work by drawing individuals' attention to *pre-existing* representations of sameness and difference as the correct basis of matching (Account 4) as relational matches constantly co-occur with reinforced object matches.

Smirnova *et al.*; Obozova *et al.*; [31••,32••]

Two recent, successful training studies, with experimentally naive crows and parrots, establish that these species also have the representational and computational

Figure 1



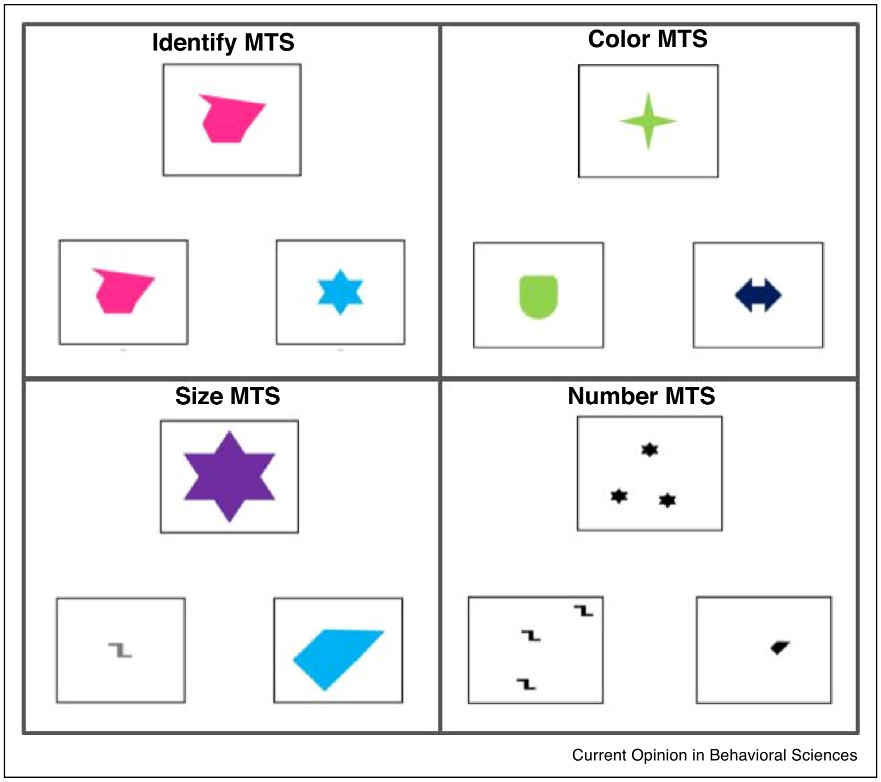
Examples of two RMTS trials. Top card in the triad is the sample, bottom two cards are choice cards. The correct choice card on each trial is the one which displays the same relation as the sample card: In trial A the sample relation is same and the correct choice card is on the left, in trial B the sample relation is different and the correct choice card is on the right.

capacities to succeed on RMTS. Furthermore, these studies are outliers in the relatively small amount of training on relational matching (see below) and the wide generalization of success: After training, the birds succeeded spontaneously (no error feedback) on three different RMTS tasks: size RMTS, color RMTS, and shape RMTS.

The Smirnova *et al.* [31••] and Obozova *et al.* [32••] studies are also particularly important as they provide the first potential evidence discriminating between Accounts 3 and 4 — that is, showing that successful training may sometimes be due to changes in inductive biases alone. These studies involved extensive training on several standard MTS tasks (see Figure 2). Notice that standard MTS tasks involve matching on the basis of individual object attributes (color, identity, shape, size) or array attributes (number), but do not require matching on the basis of relations, including same or different, that hold between individuals on a card. This is unlike the training regimes summarized above, *all* of which involve reinforced training on matching relations. Non-relational and relational matching are critically different in the representations involved: While MTS tasks involve finding the same value of an attribute across cards, this kind of matching does not require mental *representations of* the relations same and different, only a procedure for comparing attributes. This is in contrast to RMTS where matching same to same and different to different requires *representations of these relations*, in addition to a procedure for comparing them (see Refs. [6,11•,15•], and [16] for discussion). Thus, if MTS training trials in Refs. [31••] and [32••] were *sufficient* to lead to the observed success on the subsequent RMTS tasks, then it is unlikely to be due to changes in the nature of the underlying representations of sameness and difference (contra Account 3). Such representations are not involved in MTS.

However, we cannot know whether this is the case in the actual crow and parrot studies as the MTS training tasks were not the only training component in this study: The birds succeeded on the nondifferentially reinforced RMTS trials (Does AA match BB or CD?) from the first session, and as mentioned above, this spontaneous success is unique in the animal literature. However, relevant to us here, before every one RMTS trial, birds completed three differentially reinforced progressive alignment training trials (Does AA match AA or BC?) That is, the complete paradigm *also* included direct training in which responding on the basis of the relations same and different (as well as object features) was differentially reinforced, as in the many successful progressive alignment

Figure 2



Current Opinion in Behavioral Sciences

Types of MTS tasks used in Refs. [31••,32••,33••,34••] (stimuli taken from Refs. [33••,34••], not identical to those used in Refs. [31••,32••]). By Identity MTS we mean an MTS task in which sample and choice objects match on all dimensions.

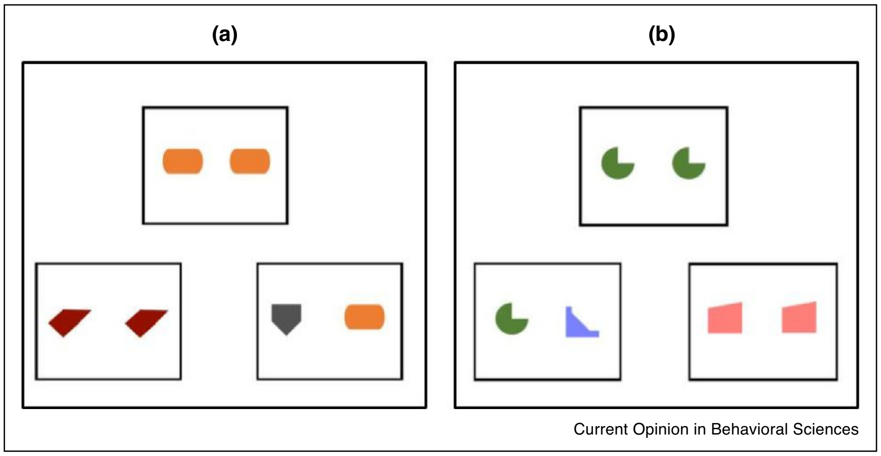
training experiments by Gentner *et al.* However, the birds' average rate of success on nondifferentially reinforced RMTS trials was numerically equal to their rate of success on differentially reinforced progressive alignment trials from the very first test session (78% and 76%, respectively). Even if progressive alignment trials were necessary, learning was extremely fast. There is reason to suspect, therefore, that – while likely important to birds' performance – progressive alignment trials do not fully account for birds' success on RMTS tasks and, as a consequence, that *non-relational* MTS training tasks played a significant role.

Ultimately, the relative importance of MTS training and progressive alignment in the case of crows and parrots remains to be determined since animals never received just one of these forms of training. That being said, as we noted above, *any* role the MTS training played *could not be* because it induced abstract representations of same and different for the first time (Account 3). *If* training on the MTS tasks were *sufficient* for success, there is no explanation other than Account 4; that MTS training altered birds' inductive biases in such a way as to make relational matches *more salient relative to alternative bases of matching*. Thus, a high priority would be to establish whether the MTS training tasks alone *are* sufficient to induce matching on the non-differentially reinforced test trials without the progressive alignment aspect of the training, and/or whether they are necessary, that is, whether the progressive alignment trials alone would be sufficient training leading crows and parrots to succeed on RMTS. We now turn to direct evidence for the sufficiency of MTS training, *on its own*, for increased relational responding in humans — both in adults who succeed on RMTS without training and in children who do not.

Training on MTS can lead to increased relational responding

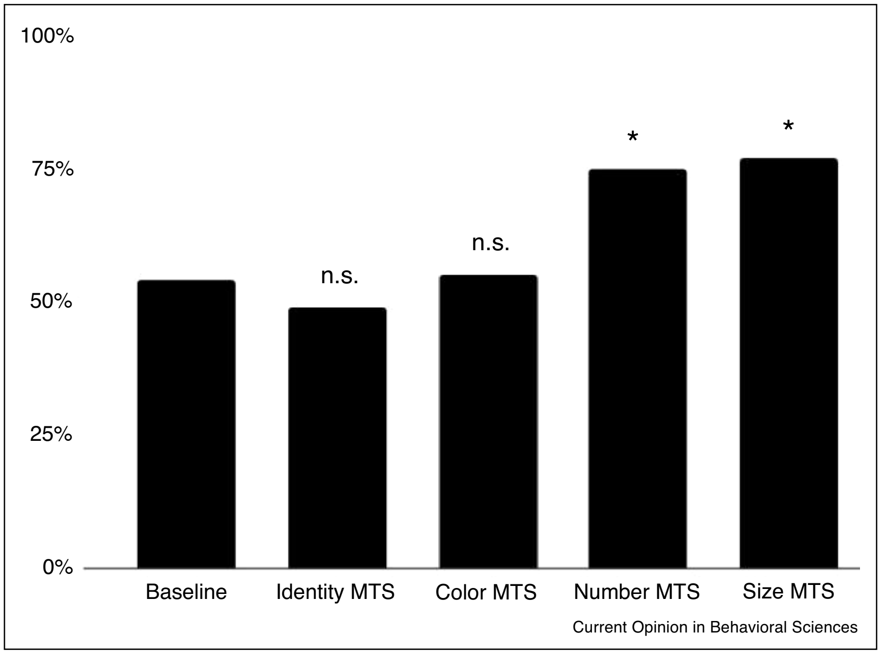
To reiterate, if training on MTS tasks (such as those in Figure 2) increases relational responding, then it *cannot* be because they produce new representations of relations (Account 3) because these tasks *do not involve matching on relations*. To test this possibility and further *ensure* that the MTS training affects *only* inductive biases (Account 4), we tested a population which *already has* the specific representations required for success on RMTS – human adults. Since adults spontaneously make relational matches on RMTS without training, we devised a task (OMTSvRMTS) which afforded two bases of matching – the relation same and an object match (Figure 3; see also Refs. [35••,36•]). Adults' initial inductive biases weighted both of these hypotheses about equally (i.e. they made matches on each basis about half of the time). Each adult received a mere *eight trials* of training on *just one* of the MTS training tasks from Smirnova *et al.* Number MTS *or* Size MTS (as depicted Figure 2) significantly increased relational responding on OMTSvRMTS, relative to baseline. In contrast, the other MTS training tasks from the crow and parrot studies (Identity MTS, Color MTS) did not do so (see Figure 4, [33••]). Given adults are clearly *already capable* of relational matching, and *already have* fully abstract representations of the relations same and different, the *only* way Number and Size MTS training could have increased relational responding was by changing inductive biases alone.

Figure 3



Two OMTSvRMTS trials. Top card in the triad is the sample, bottom two cards are choice cards. The sample card always displays the relation same, one choice card displays the relation same, the other an incomplete object match. In trial A the relational match is the card on the left and the incomplete object match is the card on the right. In trial B the incomplete object match is the card on the left and the relational match is the card on the right.

Figure 4



Percentage of relational matches made by adults in Ref. [34••], with MTS training task labelled on the Y axis. Baseline involves no training. Statistical significance of comparison to baseline indicated above the bars — n.s. (not significant), * ($p < .0001$).

Inductive biases *alone* can be the difference between success and failure on RMTS

Results with adults show that Size and Number MTS training can increase relational responding by changing inductive biases alone. The critical question, however, is whether changing inductive biases alone would be enough to produce spontaneous success on RMTS in a population which otherwise fails standard RMTS, such as four-year-old children [12•,34••]. We replicated the adult training study with four-year-olds, giving them the *very same* eight trials of either Identity MTS, Number MTS, or Size MTS as a training task, providing the correct basis of matching if they responded incorrectly. Unlike in adults, however, the test task was RMTS (Figure 1), with no feedback of any kind. The proportion of correct relational matches on the RMTS test trials after each MTS training task is depicted on Figure 5. Baseline performance with no training is not above chance responding [11•,12•]. The results were exactly parallel to those with adults: Eight trials of Number MTS training or of Size MTS training led to increased spontaneous relational matching on RMTS at above-chance levels, whereas Identity MTS training did not [34••]. Eight trials of MTS training *cannot* produce new abstract representations of the relations same and different for the first time (Account 3), since the training contains no relational matches. Therefore, this study provides the first unambiguous evidence for Account 4 — that *in some cases* population differences on RMTS (here, differences between adults who succeed spontaneously, and four-year-olds who persistently fail), are not differences in capacities or limitations in their representations of the relations same and different, but in inductive biases *alone*.

Investigating inductive biases

Ruling out, in some cases, all accounts other than Account 4 raises the question of the learning processes through which inductive biases are changed. Consider a mechanism through which Size MTS might lead to increased relational matching on a subsequent RMTS task. One must begin by characterizing an individual's initial inductive biases in a given context. In the above studies, we did so by discovering the level of spontaneous success (no error feedback) on (R)MTS tasks. To illustrate, we found, confirming much previous literature (e.g. Ref. [37]), that young children (in a middle-class US sample) have pre-existing inductive biases that make matching geometric figures by identity (i.e. when figures are the same on shape, color, and size) most likely, followed by matching by number, and least likely by matching by size alone and the relations same and different [34••]. For US adults, the ordering of hypotheses is identity matches, closely followed the relations same and different, next by color and number, and finally by size (which, like children, they are very unlikely to match on) [33••].

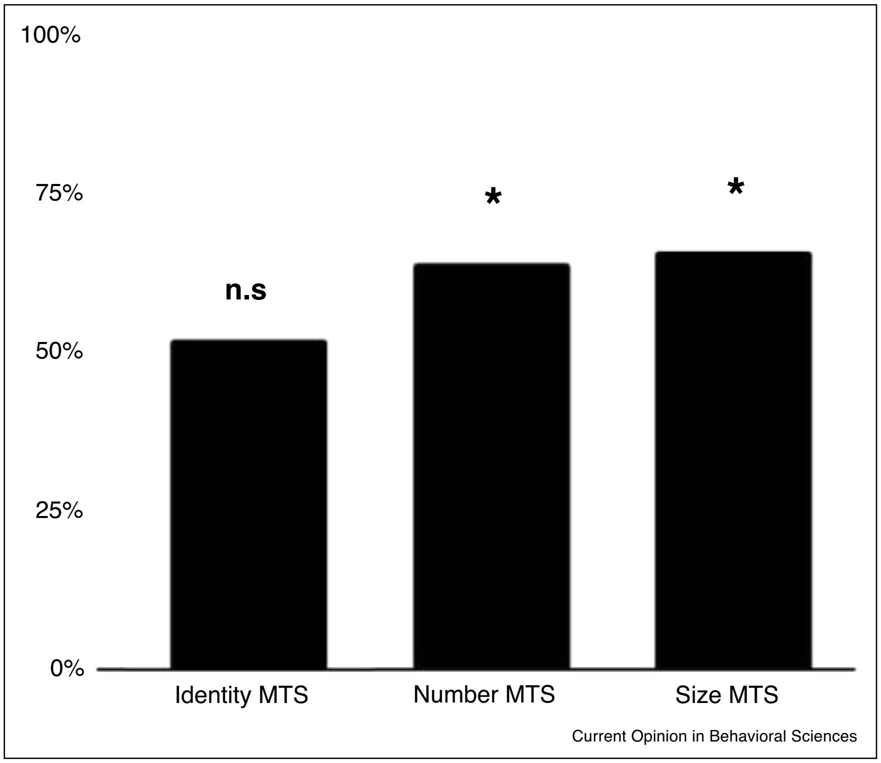
A training study might lead to increased relational responding on RMTS (or OMTSvRMTS) if it *either* changed inductive biases so as to make the specific relations same and different more likely to be inferred as the correct basis of matching, *or* if it decreased inductive biases on object identity — *specifically* object shape or color (the most likely to be inferred as correct across age), as being the correct basis of matching. Since size is not spontaneously inferred to be a relevant basis of matching, either by children or adults, when confronted by a Size MTS trial (Figure 2), participants may seek *partial* matches on dimensions they find more salient, that is, partial shape matches (e.g. two objects that are pointy) or partial color matches (e.g. two objects that are greenish). If they attempt such matches in our Size MTS training task, then they will almost certainly make an incorrect choice at least once and be told that the correct basis of matching is size. This correction tacitly informs them that shape/color are *not* correct bases of matching, potentially making them less likely to use these bases – and thus more likely to match on other bases, including relations – in a subsequent (RMTS) task.

This hypothesis generates the highly counterintuitive prediction that a task in which partial shape/color matches are *possible* but *incorrect* will lead to increased relational matching — even if the task has *no systematically correct basis of matching* (Random MTS). This prediction has been confirmed with adults [33••], in a paper which also details other specific, testable, hypotheses concerning how the MTS tasks from the crow/parrot studies might lead to increased relational responding on subsequent RMTS and OMTSvRMTS tasks.

Population differences, redux

As we mentioned at the outset, we endorse all four accounts in *some cases* of RMTS failure. The question, then, is: How can we identify the source of population differences in relational reasoning in a given case? Our work with young children provides one model of testing for differences in performance as a result of inductive biases alone: First one identifies which bases of responding individuals' inductive biases lead them to use spontaneously for the types of stimuli used in the study, next one attempts to change these biases so as to make relational responding more likely *without training on relations*. If this proves possible, Account 4 is confirmed and Accounts 1–3 are ruled out for this population. If it does not prove possible, then a further step is to train the population on a paradigm which stands to produce new relational representations (e.g. progressive alignment, or explicit external summary symbol training). In case this latter paradigm produces success, the evidence favors Account 3 and rules out Accounts 1 and 2. If no training is sufficient to produce success, then this supports Accounts 1 and 2. Notice that while such a procedure

Figure 5



Overall percentage correct on RMTS trials by children in Ref. [37] with MTS training tasks labelled on the Y axis. Statistical significance of comparison to chance (50%) indicated above the bars: n.s. (not significant), *($p < .01$).

of elimination can address the *descriptive* question of what kind of population difference may hold in any given case, it does not address the *explanatory* question of *why* such a difference exists. The former is merely a first step toward detailing the latter, which requires appeals to the specifics of an individual's experience and learning (Accounts 3 and 4) or evolutionary and maturational explanations (Accounts 1 and 2).

To conclude, our goal here has been to lay out these options and illustrate that studying the inductive biases that determine the *use* of relational reasoning capacities is an important and empirically tractable complement to existing literatures regarding the nature of these capacities/representations. After all, our human-unique repertoire of relational reasoning depends not only on the existence of our *capacities* and *representations*, but on our coming to know *how and when to use them*.

Conflicts of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as

- • of special interest
- •• of outstanding interest

1. Descartes R: **Discourse on the method**. In *Descartes: Selected Philosophical Writings*. Edited by Cottingham J *et al.*: Cambridge University Press; 1637/1985:20-56.
2. Davidson D: **Rational animals**. *Dialectica* 1982, **36**:317-327.
3. Bermúdez JL: *Thinking without Words*. Oxford: The Oxford University Press; 2003.
4. Darwin C: *The Descent of Man, and Selection in Relation to Sex*. John Murray; 1871.
5. Fodor JA: *The Language of Thought*. Harvard University Press; 1975.
6. Penn DC, Holyoak KJ, Povinelli DJ: **Darwin's mistake: explaining the discontinuity between human and nonhuman minds**. *Behav Brain Sci* 2008, **31**:109-130.
7. Tomasello M: *Becoming Human: A Theory of Ontogeny*. CITY: Belknap Press; 2019.
8. Holyoak KJ, Thagard P, Sutherland S: **Mental leaps: analogy in creative thought**. *Nature* 1995, **373**:572.
9. Halford GS, Wilson WH, Phillips S: **Relational knowledge: the foundation of higher cognition**. *Trends Cognit Sci* 2010, **14**:497-505.
10. Gentner D: **Why we're so smart**. *Language in Mind: Advances in the Study of Language and Thought*. 2003. 195235.
11. • Premack D: **The codes of man and beasts**. *Behav Brain Sci* 1983, **6**:125-136.

In this classic paper, the author develops the RMTS task in its current form and gives evidence that only children over the age of five or six and a single chimpanzee– Sarah, who had been trained in using a large vocabulary of symbols – succeed on the task.

12. • Hochmann J-R, Tuerk AS, Sanborn S, Zhu R, Long R, Dempster M, Carey S: **Children's representation of abstract relations in relational/array match-to-sample tasks**. *Cognit Psychol* 2017, **99**:17-43.

Besides confirming that three-year-olds to five-year-olds overwhelmingly fail classic 2-item RMTS, this paper establishes that three-year-olds and four-year-olds rely on entropy in 16-item Array Match to Sample, and that the few who solve the latter task on the basis of relations explicitly justify their responses using the words 'same' and/or 'different.' This is the first demonstration that the performance of preschool children resembles, in many details, that of non-human animals, and converges with other evidence that learning symbols for the relations same and different (i.e. the words 'same' and 'different'), may play a causal role in coming to solve RMTS tasks on the basis of relations.

13. • Wasserman E, Castro L, Fagot J: **Relational thinking in animals and humans: from percepts to concepts**. In *APA Handbook of Comparative Psychology: Perception, Learning, and Cognition.* Edited by Call J, Burghardt GM, Pepperberg IM, Snowdon CT, Zentall T. 2017:359-384.

A comprehensive review of relational reasoning work in human and non-human animals, including (but not limited to) work on RMTS.

14. Thompson RKR, Oden DL: **Categorical perception and conceptual judgments by nonhuman primates: the paleological monkey and the analogical ape**. *Cognit Sci* 2000, **24**:363-396.
15. • Hochmann J-R, Mody S, Carey S: **Infants' representations of same and different in match- and non-match-to sample**. *Cognit Psychol* 2016, **86**:87-111.

Fourteen-month-old infants learn to solve Match to Sample (MTS; e.g. match square to square rather than to triangle) and Non Match to Sample (NMTS; e.g. match square to triangle rather than to square) in less than 24 trials, which initially suggests they have access to representations of both same and different, learning the rule 'choose same' for MTS and 'choose different' for NMTS. However subsequent studies showed that they relied on representations of same for both tasks, learning the rules 'choose same' and 'avoid same.' The authors offer detailed proposals how representations of same (and/or) different might *differ* across populations.

16. Halford GS: *Children's Understanding: The Development of Mental Models*. Hillsdale, NJ: Erlbaum; 1993.
17. Richland LE, Morrison RG, Holyoak KJ: **Children's development of analogical reasoning: insights from scene analogy problems**. *J Exp Child Psychol* 2006, **94**:249-271.
18. Diamond A: **Executive functions**. *Annu Rev Psychol* 2013, **64**:135-168.
19. MacLean EL, Hare B, Nunn CL, Addessi E, Amici F, Anderson RC, Aureli F, Baker JM, Bania AE, Zhao Y *et al.*: **The evolution of self-control**. *Proc Natl Acad Sci U S A* 2014, **111**:E2140-E2148.
20. Carey S: **On learning new primitives in the language of thought: reply to Rey**. *Mind Lang* 2014, **29**:133-166.
21. • Kotovsky L, Gentner D: **Comparison and categorization in the development of relational similarity**. *Child Dev* 1996, **67**:2797-2822.

The authors established that preschool aged children fail an analogical mapping (relational reasoning) task involving relative positions of different sized identical objects (e.g. xXx goes with zZz rather than yyY). When trained under a progressive alignment regime, that is, (aAa goes with aAa rather than bbB, children of the same age subsequently succeeded at the original task on which there were no object matches in addition to the relational match. This was the first study showing the efficacy of progressive alignment training on relational reasoning.

22. Fodor JA: **On the impossibility of acquiring "more powerful" structures: fixation of belief and concept**. In *Language and Learning.* Edited by PiatelliPalmerini M. Harvard University Press; 1980:142-162.
23. Flemming TM, Beran MJ, Thompson RKR, Kleider HM, Washburn DA: **What meaning means for same and different: analogical reasoning in humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), and rhesus monkeys (*Macaca mulatta*)**. *J Comp Psychol* 2008, **122**:176-185.
24. Truppa V, Piano Mortari E, Garofoli D, Privitera S, Visalberghi E: **Same/different concept learning by capuchin monkeys in matching-to-sample tasks**. *PLoS One* 2011, **6**:e23809.
25. Fagot J, Maugard A: **Analogical reasoning in baboons (*Papio papio*): flexible re-encoding of the source relation depending on the target relation**. *Learn Behav* 2013, **41**:229-237.
26. Fagot J, Thompson RKR: **Generalized relational matching by guinea baboons (*Papio papio*) in two-by-two-item analogy problems**. *Psychol Sci* 2011, **22**:1304-1309.

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

28. Christie S, Gentner D: **Language helps children succeed on a**
• **classic analogy task**. *Cognit Sci* 2014, **38**:383-397.

In partial a RMTS task with only 'same' as the sample trials (i.e. match AA to BB or CD), which may lead to success on the task without relational reasoning). Priming with the words 'same' and 'different' *or* labelling of the relation 'same' with a novel word ('truffet') lead to early success.

29. Ferry AL, Hespos SJ, Gentner D: **Prelinguistic relational**
• **concepts: investigating analogical processing in infants**. *Child Dev* 2015, **86**:1386-1405.

A demonstration of the effect of progressive alignment on young infants' representations of same and different in a habituation paradigm. While this task does not require relational reasoning, it does require some abstract (generalizable to new stimuli) representations of the relations same and different, and young infants only succeed after a training regime that involved progressive alignment.

30. Gentner D, Hoyos C: **Analogy and abstraction**. *Top Cognit Sci* 2017, **9**:672-693.

31. Smirnova A, Zorina Z, Obozova T, Wasserman E: **Crows**
•• **spontaneously exhibit analogical reasoning**. *Curr Biol* 2015, **25**:256-260.

Two experimentally naive crows each showed spontaneous (i.e. no error feedback) on three separate RMTS tasks (shape RMTS, abstracting across color; color RMTS, abstracting across shape, and Size RMTS, abstracting across both shape and color). The RMTS test tasks contained three trials of progressive alignment (match AA to AA or BC) followed by one probe trial of standard RMTS (match DD to EE or FG), with no error feedback. Crows succeeded at a rate of above 70% on probe trials from the first block of testing, meaning that it took a *maximum* of 36 progressive alignment trials to reach such robust success. This level of performance held throughout almost all subsequent testing blocks for all three RMTS tasks. This experiment establishes that crows have the representational and computational capacities for success on RMTS, and given the spontaneous success on three separate RMTS tasks.

32. Obozova T, Smirnova A, Zorina Z, Wasserman E: **Analogical**
•• **reasoning in amazons**. *Anim Cognit* 2015, **18**:1363-1371.

Replication of the Smirnova *et al.* [34••] paradigm with two parrots. See above for annotation.

33. Kroupin I, Carey S: *The Importance of Inference in (Human-Unique)*
•• *Relational Reasoning: Relational Match to Sample as a Case Study*. . Manuscript under revision 2020.

A study testing whether MTS training tasks from Ref. [34••] increase the likelihood that human adults make relational matches in a modified RMTS task. Identity and Color MTS were *not* found to increase adults' relational responding, while Number and Size MTS did do so. An account of these effects at the level of specific inductive biases was developed.

34. Kroupin I, Carey S: *Failures in Relational Matching as Differences in*
•• *Specific Inductive Biases: Exploring Effects of Non-Relational Training Tasks*. . Manuscript under revision 2020.

A study training four-year-olds on MTS tasks from Ref. [34••], identical to the training used in Ref. [36••] and testing on standard RMTS. The pattern of results mirrored that in training with adults [36••]: Four-year-olds did not perform above chance on RMTS without training or after training on Identity MTS. After *eight trials* of training on Number or Size MTS, alone, however, children performed above chance on RMTS. This is the first unequivocal demonstration that a population difference on Premack's classic relational reasoning task can be a result of differences in inductive biases *alone*.

35. Christie S, Gentner D: **Relational similarity in identity relation:**
•• **the role of language**. In In *Proceedings of the Second European Cognitive Science Conferenc.* Edited by Vosniadou S, Kayser D. *Proceedings of the Second European Cognitive Science Conferenc* 2007. {is this referenced in the paper? Isn't it just the conference proceedings version of 33••?.

36. Vendetti MS, Wu A, Holyoak KJ: **Far-out thinking: generating**
• **solutions to distant analogies promotes relational thinking**. *Psychol Sci* 2014, **25**:1-6.

Adults, after completing distant analogy problems (e.g. fireplace:wood:: stomach:?) were more likely to make relational comparisons between two scenes– that is, identifying two objects as the same by the role they play and not by appearance. This is one of first explicit discussions of the role of inductive biases in relational reasoning.

37. Chan JYC, Mazzocco MM: **Competing features influence children's attention to number**. *J Exp Child Psychol* 2017, **156**:62-81.