eScholarship
International Journal of Comparative Psychology

Title
Relational and Analogical Reasoning in Comparative Cognition

Permalink
https://escholarship.org/uc/item/2225s37c

Journal
International Journal of Comparative Psychology, 29(1)

ISSN
0889-3675

Authors
Dymond, Simon
Stewart, Ian

Publication Date
2016

DOI
10.46867/ijcp.2016.29.00.24

Copyright Information

Relational and Analogical Reasoning in Comparative Cognition

Simon Dymond[1,2] and Ian Stewart[3]

[1] Swansea University, United Kingdom
[2] Reykjavík University, Iceland
[3] National University of Ireland, Galway, Ireland

Several nonhuman animal species have been claimed to successfully pass tests indicative of relational matching and to therefore engage in analogical reasoning. Here, we address these claims by focusing on one recent case study. We illustrate several potential methodological limitations that make it uncertain as to whether the subjects in this particular study were indeed showing relational matching. To the extent that similar or analogous limitations apply in other studies, this undermines the claim of relational matching. Apart from this, however, even if relational matching was to be conclusively demonstrated in non-humans, this behavior alone is profoundly different from analogical reasoning as performed by humans. Substantial converging evidence now suggests a critically important difference between humans and nonhumans at the level of behavioral process that explains why nonhumans do not engage in complex language and therefore do not engage in processes that require complex language, including analogy. In accordance with both these arguments, we suggest that caution is needed in the comparative cognition literature when extrapolating from nonhuman to human cognitive capacity.

There has been much to crow about in the comparative cognition literature of late, not least the demonstrations of seemingly uniquely human abilities in corvids. For instance, Smirnova, Zorina, Obozova, and Wasserman (2015) reported that two hooded crows (*Corvus corone*) can successfully pass relational matching to sample (RMTS) tasks and that this shows "the most convincing evidence yet of analogical reasoning in a nonprimate species" (p. 256). Here, we evaluate such claims and urge caution when extrapolating from the findings of such research to the cognitive capacities of humans. We first emphasize that methodological limitations within this particular study should preclude a definitive, parsimonious claim from being made about the capacity of crows to show relational matching to sample (Vonk, 2015). Second, we argue that, even if the latter is shown conclusively in this or another species, this by itself would still not be a display comparable to analogical reasoning as seen in humans, because while relational matching bears certain similarities to analogical reasoning in that sense, it is missing a core element, namely *arbitrarily applicable relational responding* (e.g., Stewart & Roche, 2013).

In what follows, we first describe the main features of the matching to sample tasks employed by Smirnova et al. (2015) as a representative example of relational matching paradigms and outline some potential limitations. Then, we describe the concept of *arbitrarily applicable relational responding,* which evidence suggests only humans show to any significant degree and which seems to be the process that facilitates the symbolism and generativity of human language and cognitive performance including analogical reasoning.

Please send correspondence to Simon Dymond, Experimental Psychopathology Lab, Department of Psychology, Swansea University, Singleton Campus, Swansea, SA2 8PP, United Kingdom. (Email: s.o.dymond@swansea.ac.uk)

Outline of Smirnova et al.'s (2015) Task

In their study, Smirnova et al. (2015) did not offer a formal definition of analogical reasoning, arguing instead that:

> …relational matching-to-sample (RMTS) effectively captures the essence of analogy, in which the relevant logical arguments are presented visually (Thompson & Oden, 2000). In RMTS, choice of test pair BB would be correct if the sample pair were AA, whereas choice of test pair EF would be correct if the sample pair were CD. Critically, no items in the correct test pair physically match items in the sample pair, thus demanding that only relational sameness or differentness is available to support accurate choice responding. (p. 256)

In other words, the authors see same/different relational responding based entirely on formal, nonarbitrary properties of the various choice arrays, as involving the same core processes as analogical reasoning. Accordingly, their experimental paradigm incorporated trials based on physical similarity or identity matching to sample (IMTS) and trials based on relational properties or RMTS. Crucially, the RMTS trials followed an extensive history of IMTS pretraining and testing and trials of both types were subsequently presented under different outcome contingencies during three defined assessment phases.

During the IMTS pretraining and pretesting, the birds were trained on sample and comparison arrays consisting of color, Arabic numerals, and number of items before being tested with stimuli based on size. The stimulus dimensions were printed on small, laminated cards placed over three food cups (one each for the sample and two comparisons), which were arranged on a tray slid into the bird's cage. The experimenter did not have visual access to the bird and, therefore, she:

> . . .could only judge the outcome of each trial by the sounds made by the bird. If she heard the sound of a card falling from the cup and the later sound of the bird pecking the cup containing the worms, then the choice was correct. If no pecking sound occurred, then it meant that the choice was incorrect and the experimenter quickly removed the tray to prevent the bird from uncovering the second cup (Supplementary material).

Importantly, differential reinforcement contingencies were in effect for IMTS trials throughout the study; that is, food, consisting of two live mealworms (*Tenebrio molitor*) was available following correct responses only.

Identity matching to sample training was conducted across four stimulus sets (black/white, Arabic numerals, l/2 elements, and intermixed pairs), and the results showed that accuracy increased across stimulus sets for both crows. During IMTS (within) testing, three-fourths of test trials were from the training set, with one-fourth of trials from outside the training set. The final IMTS test block consisted of testing outside training categories. Here, on three-fourths of the trials, the correct comparison stimulus was an exact match to the sample. On another one-fourth of the trials, the correct comparison stimulus matched the sample in size, but depicted a different shape.

Next, the crucial RMTS trials were presented in assessment phases consisting of 36 IMTS trials and 12 RMTS trials presented across eight sessions, with sessions conducted 5 days a week. Non-differential reinforcement contingencies were adopted on RMTS trials only. That is, food was given after all choices, both correct and incorrect, with comparison cups baited with live mealworms. This was done because the

authors did not wish to "teach the crows the very behavior that we were assessing; accurate choice responding here could only be based on relational matches between the sample pair and the correct test pair" (Smirnova et al., 2015, p. 256).

Assessment phases were arranged in terms of the stimulus dimension under study; the first phase was based on size, the second shape and the third color. In the first assessment phase, on both IMTS and RMTS trials, the large shape always appeared on the left and the small shape always appeared on the right. In all assessment phases, the IMTS and RMTS trials always appeared at the same point in the eight trial-blocks. In the second assessment phase, "item pairs were arranged along the positive diagonal in order for the spatial arrangement to differ from the horizontal arrangement in the prior assessment phase" (Smirnova et al., 2015, p. 256). In the third assessment phase, half of the trials involved sample and comparison arrays identical in color (e.g., a sample consisting of two blue crosses, with comparisons consisting of two blue squares or two yellow squares) and the other half in non-identical color. The same shape stimuli used in the previous assessment phase were employed here but with the identical/non-identical color combinations described above. It is noteworthy that trial-unique combinations of samples and comparisons were employed in the second and third assessment phases, a procedure known to facilitate acquisition of generalized concept learning in pigeons (Wright, Cook, Rivera, Sands, & Delius, 1988).

Key Methodological Limitations

Regardless of the species under study, a nonautomated procedure like this raises potential methodological limitations (Dymond, Rehfeldt, & Schenk, 2005). First, the crows may have moved (quickly) from the incorrect cup to the correct cup on some trials. Second, the IMTS trial duration was unclear and the intertrial interval (ITI), which was estimated by the authors to range between 1 and 5 min, was rather long. Putting aside the extended trial and ITI durations, the authors reported that if no choice occurred after 2 min, a 5 min break followed, and the trial was repeated. The authors do not present any data on how many such trials were repeated or if this contingency was applied throughout both IMTS and RMTS. It could, however, be argued that this type of correction procedure, in which the trial configuration is repeated until a correct response is made (Kangas & Branch, 2008), may have biased correct over incorrect responding. Corrections procedures are effective at overcoming stimulus and positional biases in matching to sample and may therefore have contributed to the high accuracy observed in Smirnova et al. (2015).

Third, there was no observing response requirement. In procedures involving an observing response, the subject must respond to the sample stimulus (e.g., by pecking it) and only then do the comparison stimuli appear, typically after the sample has been removed. This procedure, which ensures attendance to the sample, is known to support reliable acquisition of stimulus control (Hogan, Zentall, & Pace, 1983). In the absence of such a procedure, behavior may not be coming under appropriate stimulus control, even while appearing to do so.

Fourth, while Smirnova et al. (2015) sought to address potential confounds for inadvertent visual cueing by the experimenter and auditory stimuli when using live bait as reinforcement (Supplementary Methods) by preparing trials in a fixed sequence, the possibility remains that the birds may have been able to hear any noises made by the two live mealworms. This seems particularly likely given that trials began when the tray containing the sample and comparison cups was partially slid into the cage for 2 to 3 s to allow the crows to view but not uncover the cups. Noises made by the mealworms that were heard in this way might have acted as a supplementary source of stimulus control. Apart from making noises, mealworms also emit

pheromones, and thus olfactory stimulation may also have influenced acquisition of IMTS. One way to control for auditory and olfactory stimuli such as this is to bait both the correct and incorrect cups but to only allow access to the correct well (Dymond, Gomez-Martin, & Barnes, 1996; Peña, Pitts, & Galizio, 2006; Wright & Delius, 1994).

Finally, it is recommended practice when conducting nonautomated research involving visual-based data collection methods to report observer-training methods and calculate interobserver agreement (IOA) scores (Johnston & Pennypacker, 1993; Poling, 1985; Saunders & Williams, 1998). Unfortunately, however, Smirnova et al. (2015) reported no such data observation protocol or IOA scores.

Taken together, these methodological limitations should at this moment in time preclude a definitive empirical statement from being made about the behavioral capacity or otherwise of crows to engage in RMTS. Further work controlling for these kinds of procedural and interpretative issues is essential.

Rules and the Matching and Relational Tasks

Recently, Vonk (2015) suggested that the birds in Smirnova et al.'s (2015) study may have been responding in accordance with a simple, non-relational rule in both the perceptual matching and analogical relational tasks. Moreover, the sample and comparison combinations employed may have lead the crows to view the sample and comparison arrays as representing single, unitary patterns consisting of one color, size or shape, and so on (i.e., a form of configural learning; see Katz, Bodily, & Wright, 2008). This seems especially likely given that on relational trials the birds' choices may have been aided by the perceptual rule and supported by the fact that on trials involving size the smaller item was always to the right of the larger comparison. Similarly, color trials could have relied on this *one or two-colors in the sample* perceptual account to determine comparison selection.

As Vonk (2015) points out, Smirnova et al.'s (2015) argument that the birds were responding to the sample and comparison arrays in terms of analogy rests on the assumption that they (a) discriminated samples from comparisons and did not process them configurally as one unitary pattern and (b) separately processed the elements contained in the samples as distinct objects. Vonk's (2015) perceptual processing account "supposes only that the bird recognizes the presence or absence of certain perceptual features in the entire stimulus." (R69). A suggested control condition to prevent the birds from responding in accordance with the perceptual rule might entail employing stimulus configurations where, for instance, a sample consisted of two different shapes of the same color presented with comparisons of differing shapes and colors. In fact, quantity based discrimination (i.e., larger/smaller) was not assessed by Smirnova et al. (2015). Other control conditions suggested by Vonk (2015) might manipulate the number of shared features between sample and comparisons.

Control procedures such as this should be incorporated into future work examining the relational matching abilities of this or other species. Of course, it is possible that such better controlled experiments might provide even stronger evidence that crows are capable of relational matching. Recent evidence (involving a very different and simpler paradigm and arguably providing a much better controlled demonstration) suggests that relational matching may be observed even in the context of filial imprinting (Martinho & Kacelnik, 2016); hence the idea that crows might be trained in such responding seems at least plausible. Nevertheless, as suggested previously, this by itself would not, in our opinion, constitute a

demonstration that they are capable of analogical reasoning because, as we are about to explain, such a performance would still be missing a key ingredient.

Analogical Reasoning in Humans Requires Arbitrarily Applicable Relational Responding

There is now compelling evidence that humans alone have the ability to show a clearly functionally defined pattern of behavior referred to as *arbitrarily applicable relational responding* (see Hayes, Barnes-Holmes, & Roche, 2001; Stewart, 2016). This is a type of relational responding in which the relations between the stimuli involved are not based on the physical properties of those stimuli (in contrast to the relational behavior shown by the crows in Smirnova et al.'s [2015] research) but on additional sources of environmental control. For example, imagine that I play a game with a typically developing child in which I tell her that "Mr. X is taller than Mr. Y" and then ask her who is smaller. She will very likely answer (correctly) "Mr. Y" despite the fact that the auditory stimuli "Mr. X" and "Mr. Y" bear no obvious comparative relationship, and therefore she cannot be basing her answer on physical relations. Instead, what she is doing here is relationally responding to these stimuli based on the contextual cues of "taller" and "smaller" that she has learned through previous exposure to her language community. This is referred to as *arbitrarily applicable* relational responding (AARR; see Hayes et al., 2001) because the relational responding depends not on physical properties but on contextual cues and thus can in principle be applied to any stimuli regardless of actual physical relations (Stewart & McElwee, 2009).

This example illustrates just one type of AARR; that is, responding based on a comparative (more/less) relation. However, accumulating evidence suggests that humans learn many other varieties of AARR also, including similarity (or coordination, e.g., "dog is the same as chien"), opposition (e.g., "day is opposite to night"), distinction (e.g., "mice are different from rats"), hierarchy (e.g., "an ant is a type of insect"), analogy (e.g., "X is to Y as W is to Z"), deixis (e.g., "If I was you I'd think that through"), and temporality (e.g., "pride comes before a fall") to mention just a few that have received attention. Substantial evidence now suggests that people can engage in a variety of different forms of AARR (Dymond & Roche, 2013).

All forms of AARR share a number of key properties. The first, *mutual entailment*, refers to the finding that a relation trained in one direction entails derivation of a second relation in the other (e.g., if A > B then B < A). The second, *combinatorial entailment*, refers to the fact that taught relations are combined to derive new relations (e.g., if A > B and B > C then A > C and C < A). Third, *transformation of functions* refers to the effect in which a stimulus in an arbitrarily applicable relation with another stimulus can change its functions simply through being in that relation (e.g., if A > B > C and if B is conditioned as aversive then A will become less aversive than B and C more aversive; see Dougher, Hamilton, Fink, & Harrington, 2007).

Based on these symbolic and generative properties, it has been argued that AARR can be considered as the key operant underlying human language and cognition (e.g., Hayes & Hayes, 1992). Furthermore, considerable work has supported this argument by (a) showing that AARR is not just correlated with language but emerges before development of an advanced linguistic repertoire (e.g., Luciano, Gomez & Rodriguez, 2007), (b) by demonstrating that training AARR with abstract stimuli can substantially boost verbal and intellectual ability (e.g., Cassidy, Roche, Colbert, Stewart, & Grey, 2016; Hayes & Stewart, 2016), (c) by reporting controlled empirical demonstrations of the origins and development of various forms of AARR (e.g., Berens & Hayes, 2007), (d) by convincingly showing that AARR meets criteria as an operant

behavior (e.g., Healy, Barnes-Holmes, & Smeets, 2000), and (e) by demonstrating AARR as a useful functional analytic unit for the analysis of human language and cognition (Dymond & Roche, 2013).

Searching for Arbitrarily Applicable Relational Responding in Nonhumans

Of particular importance in the current context is the related empirical finding that while humans can readily engage in AARR, nonhumans cannot. Research with nonhumans has failed to convincingly demonstrate arbitrarily applicable relations even in the case of the relation of sameness, which is arguably the simplest and most fundamental type (Dymond, 2014; Hayes, 1989; Lionello-DeNolf, 2009). For example, in a seminal study by Sidman et al. (1982), rhesus monkeys, baboons and 5-year-old children were all trained, using identical procedures, to choose particular colors in the presence of particular line orientations and were then tested to see if, having being trained in one direction (i.e., to put particular colors with particular line orientations; for example, match red to horizontal and match green to vertical), they could then derive (i.e., demonstrate without reinforcement) the correct mutually entailed relations in the opposite direction (i.e., to put particular line orientations with the correct colors based on the training; for example, match horizontal to red and match vertical to green), thus suggesting that they were treating the members of particular line-color pairings (e.g., red and horizontal) as equivalent to each other (i.e., treating them as functionally the same despite the fact that they were not physically the same). In repeated testing, only the children showed these derivations (referred to in the literature as derived symmetry responding), while the other species consistently failed to do so. Dugdale and Lowe (1990, 2000) have found similar failures to derive even in chimpanzees provided with specialized language-relevant training.

The search for robust demonstrations of coordinate AARR in nonhumans has as yet proven fruitless despite ingenious experimental paradigms, meticulous stimulus control, and heroic data collection efforts (Lionello-DeNolf, 2009). A variant of symmetry, called associative symmetry, which has recently generated research attention, appears restricted to successive discrimination learning tasks (Campos, Urcuioli, & Swisher, 2014; Urcuioli, 2008; Zentall, Wasserman, & Urcuioli, 2013). Claims to the demonstration of forms of coordinate AARR in other contexts, are rare and disputed and remain unreplicated (cf. Kastak, Kastak, & Schusterman, 2001; Schusterman & Kastak, 1993). The uncomfortable fact therefore remains: nonhumans do not derive and combine stimulus relations with the same ease with which humans do (Dymond, 2014; Hayes, 1989).

What conditions might provide the best test of AARR in nonhumans? Empirical data on the learning of AARR in humans suggests the crucial role of multiple exemplar training (MET) of the relational patterns involved. As such, MET should play a central role in any test with nonhumans. For example, in testing for the very basic derived relational response of symmetry, necessary to provide extensive experience in bidirectional responding across multiple stimulus sets before testing and re-testing with novel sets. It is notable that Schusterman and Kastak's (1993) study employed multiple exemplar methods like these and reported evidence of coordinate AARR (i.e., *stimulus equivalence*) in a single sea lion. To date, however, this potentially important study awaits replication. Future research should also employ stimulus features which maximize animals' evolutionary capacities wherever possible and be presented in fully automated analogs of real world contexts. Experimental paradigms capable of conducting vast numbers of training and testing sessions and which track individual performance across time already exist (e.g., Grainger, Dufau, Montant, Ziegler, & Fagot, 2012; Scarf et al., 2016) and may prove useful in the ongoing search for AARR in nonhumans.

Analogical Reasoning as Arbitrarily Applicable Relational Responding

Another area of research of key relevance to the current paper is work that has used AARR to model analogical reasoning (Stewart, 2016). From this point of view, analogy can be conceptualized as the arbitrarily applicable relating (in accordance with a relation of coordination) of two sets of arbitrarily applicable relations to one another. Barnes, Hegarty, and Smeets (1997) first demonstrated analogical reasoning conceptualized in this way. They administered a learning task in which a series of arbitrarily applicable sameness relations, referred to as equivalence relations, were derived between abstract stimuli. After participants showed the latter they also went on to derive relations between equivalence relations themselves. Hence, an empirically controlled pattern of AARR could be used to model analogy. This not only provided further evidence of AARR as a key process constituting language but also a means of studying analogy in an empirically controlled environment, with much less reliance on previous, unmeasured learning (Stewart, 2016).

This model has since then been used to explore features of analogical reasoning including its development in childhood, electrophysiology, complexity versus speed in the context of analogy, and the effects of training on demonstrations of analogy (e.g., Barnes-Holmes et al., 2005; Carpentier, Smeets, Barnes-Holmes, & Stewart, 2004; Lipkens & Hayes, 2009). The work on development of analogy in childhood constitutes perhaps the best example of the insights emerging from this new emphasis on AARR. This research involved a number of studies comparing the AARR model of analogy in various age groups including five-year-olds, nine-year-olds, and adults (Carpentier, Smeets & Barnes-Holmes, 2002, 2003; Carpentier et al., 2004). Findings were that whereas both the latter groups could show analogical AARR, the five-year olds were unable to without specialized remediation. This is a similar developmental division as found by other researchers who have investigated the origins of analogy (e.g., Sternberg & Rifkin, 1979). In addition, this stream of research has also provided evidence that when children younger than five appear to show analogy they may be engaged in a functionally simpler pattern of matching (Carpentier et al., 2004).

These findings are important for a number of reasons (Stewart, 2016). They further support the AARR-based approach by replicating the childhood developmental divide in analogical ability seen in traditional research. However, they also go beyond this work by providing a more precisely defined and better controlled demonstration of analogy than previously and by training this capacity in young children. The capacity to train analogy has obvious significance for the applied educational arena; such training might boost analogical ability as well as intellectual performance more generally in both developmentally delayed and typically developing children (see Stewart, Barnes-Holmes, & Weil, 2009).

Conclusion

To sum up, extensive evidence suggests that humans readily show AARR in multiple varieties while other species do not; this capacity appears to be a central process in human language and cognition; and AARR has been used to model analogical reasoning such that key cognitive developmental findings have been recapitulated and additional advances made as regards both technical understanding of this ability as well as practical intervention with respect to training it. Returning to our key point in the current paper, we reiterate that analogical reasoning thus conceptualized is qualitatively different than the performance shown by the crows in Smirnova et al. (2015). Whereas analogy as conceptualized and modelled within the current perspective incorporates AARR as a central component, in the Smirnova et al. study, the pattern shown was exclusively based on nonarbitrary relations in which the relations are based on physical similarity and/or

physical difference. Hence, even if the crows in that study could be claimed to have shown relational matching, this is still not a display of analogical reasoning as seen in humans.

One argument that might be raised in counterpoint is that Smirnova et al. (2015) are not claiming that relational matching is analogical reasoning *as seen in humans*, but that it is still analogical reasoning, albeit of a more basic, simple variety. Furthermore, this point might even be argued to be supported by the Relational Frame Theory (RFT) analysis just provided which sees analogical reasoning in the human realm as a combination of the relating of relations (i.e., relational matching) and AARR. On this basis, perhaps, it could be argued that relational matching by itself is indeed a simple, non-symbolically based form of analogical reasoning.

Essentially, this debate concerns the appropriate use of terms, which of course is critical in science. *Analogical reasoning* is a natural language term rather than a technical one so it is not surprising that there might be disagreement about how exactly it should be applied – for example, whether to include both simple relational matching as well as abstract (i.e., AARR-based or symbolic) relational matching or to confine the term to the latter alone. However, one aspect suggested by RFT studies such as those laid out above is that AARR leads to exponentially greater complexity and speed of acquisition of new learning (see, for example, O'Hora, Barnes-Holmes, Smeets, & Roche, 2004; Wulfert & Hayes, 1988). Relational matching combined with AARR would be expected to differ profoundly from simple relational matching by itself and indeed empirical evidence of this already exists (Stewart, Barnes-Holmes, Smeets, & Roche, 2004). Hence, whether we term the combination of AARR and relational matching, analogical reasoning or we develop a completely novel term for this pattern, there is at least a strong case that some terminological differentiation is required. We would argue that the most coherent course of action might be to continue to use the terms *relational matching* and *analogical reasoning* and to confine the latter to relational matching combined with AARR. From this point of view, then, relational matching is indeed a core element of analogical reasoning and hence is similar to it at one level, but relational matching without AARR is missing something utterly transformative and thus is a vastly simpler process. Without AARR, relational matching is a relatively simple, low level activity; with it, complex abstractive generative patterns of thought linked with the highest levels of human creativity and ingenuity become possible (Dymond, Roche, & Barnes-Holmes, 2003).

It is important to evaluate new evidence of the comparative abilities of nonhumans thoroughly and with careful attention paid to the procedural nuances of individual studies. Replication and refinement should be the default positions before claims with respect to the higher-order cognitive abilities of different species is developed and offered for further discussion, debate, and possible falsification (Dymond, Haselgrove, & McGregor, 2013). As the ongoing debate over whether higher-order cognition in nonhumans is best explained by either associative learning or propositional theories shows (De Houwer, Hughes, & Barnes-Holmes, 2016; Haselgrove, 2016), seemingly minor methodological factors can have enormous, conflicting theoretical implications.

As a separate issue from this methodological caution, as we have tried to lay out in this paper, substantial converging evidence now suggests one very important difference between humans and nonhumans at the level of behavioral process which in particular suggests caution before making simple extrapolations from nonhuman to human cognitive capacity, namely arbitrarily applicable relational responding (Dymond & Roche, 2013; Stewart, 2016).

According to Vonk (2015):

We must be cautious in granting abilities to animals that are interesting largely because they potentially break down the human erected divide between humans and other animals, and less so because they illuminate some aspect of animal cognition that explains that particular animal's umwelt (R71).

The evidence of the uniquely human capacity for arbitrarily applicable relational responding should strongly underline this suggestion.

References

Barnes, D., Hegarty, N., & Smeets, P. M. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *The Analysis of Verbal Behavior*, *14*, 57–83.

Barnes-Holmes, D., Regan, D., Barnes-Holmes, Y., Commins, S., Walsh, D., Stewart, I., Smeets, P. M., Whelan, R., & Dymond, S. (2005). Relating derived relations as a model of analogical reasoning: Reaction times and event related potentials. *Journal of the Experimental Analysis of Behavior*, *84*, 435–452.

Berens, N. M., & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: Experimental evidence for relational operants. *Journal of Applied Behavior Analysis*, *40*, 45–71.

Campos, H. C., Urcuioli, P. J., & Swisher, M. (2014). Concurrent identity training is not necessary for associative symmetry in successive matching. *Journal of the Experimental Analysis of Behavior*, *101*, 10–25.

Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2002). Matching functionally same relations: Implications for equivalence-equivalence as a model for analogical reasoning. *Psychological Record*, *52*, 351–370.

Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2003). Equivalence-equivalence as a model of analogy: Further analyses. *Psychological Record*, *53*, 349–371.

Carpentier, F., Smeets, P. M., Barnes-Holmes, D., & Stewart, I. (2004). Matching derived functionally-same relations: Equivalence-equivalence and classical analogies. *Psychological Record*, *54*, 255–273.

Cassidy, S., Roche, B., Colbert, D., Stewart, I., & Grey, I. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learning and Individual Differences*, *47*, 222–235. doi.10.1016/j.lindif.2016.03.001

De Houwer, J., Hughes, S., & Barnes-Holmes, D. (2016). Associative learning as higher order cognition: Learning in human and nonhuman animals from the perspective of propositional theories and Relational Frame Theory. *Journal of Comparative Psychology*, *130*, 215–225.

Dougher, M. J., Hamilton, D. A., Fink, B. C., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the Experimental Analysis of Behavior*, *88*, 179–197. doi: 10.1901/jeab.2007.45–05.

Dugdale, N. A., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.) *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 115–138). Brighton, UK: Erlbaum.

Dugdale, N. A., & Lowe, C. F. (2000). Testing for symmetry in the conditional discriminations of language–trained chimpanzees. *Journal of the Experimental Analysis of Behavior*, *73*, 5-22.

Dymond, S. (2014). Meaning is more than associations: Relational operants and the search for derived relations in nonhumans. *Journal of the Experimental Analysis of Behavior*, *101*, 152–155

Dymond, S., Gomez-Martin, S., & Barnes, D. (1996). Multi-modal conditional discrimination in rats: Some preliminary findings. *Irish Journal of Psychology*, *17*, 269–281.

Dymond, S., Haselgrove, M., & McGregor, A. (2013). Clever crows or unbalanced birds?. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(5), E336.

Dymond, S., Rehfeldt, R. A., & Schenk, J. J. (2005). Nonautomated procedures in derived stimulus relations research: A methodological note. *Psychological Record*, *55*, 461–482.

Dymond, S., & Roche, B. (2013). *Advances in relational frame theory: Research & application*. Oakland, CA: New Harbinger.

Dymond, S., Roche, B., & Barnes-Holmes, D. (2003). The continuity strategy, human behavior, and behavior analysis. *The Psychological Record, 53*, 333–347.

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., & Fagot, J. (2012). Orthographic processing in baboons (*Papio papio*). *Science*, *336*, 245–248.

Haselgrove, M. (2016). Overcoming associative learning. *Journal of Comparative Psychol*ogy, *130*, 226–240. doi:/10.1037/a0040180

Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *Journal of the Experimental Analysis of Behavior*, *51*, 385–392.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York, NY: Plenum Press.

Hayes, S. C., & Hayes, L. J. (1992). Verbal relations and the evolution of behavior analysis. *American Psychologist*, *47*, 1383–1395.

Hayes, J., & Stewart, I. (2016). Comparing the effects of derived relational training and computer coding on intellectual potential in school-age children. *British Journal of Educational Psychology*, 86, 397–411. doi: 10.1111/bjep.12114

Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of the Experimental Analysis of Behavior*, *74*, 207–227.

Hogan, D. E., Zentall, T. R., & Pace, G. (1983). Control of pigeons' matching-to-sample performance by differential sample response requirements. *American Journal of Psychology*, *96*, 37–49.

Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of human behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.

Kangas, B. D., & Branch, M. N. (2008). Empirical validation of a procedure to correct position and stimulus biases in matching to sample. *Journal of the Experimental Analysis of Behavior*, *90*, 103–112.

Kastak, C. R., Kastak, D., & Schusterman, R. J. (2001). Equivalence classification by California sea lions using class-specific reinforcers. *Journal of the Experimental Analysis of Behavior*, *76*, 131–158.

Katz, J. S., Bodily, K. D., & Wright, A. A. (2008). Learning strategies in matching to sample: If-then and configural learning by pigeons. *Behavioural Processes*, *77*, 223–230.

Lipkens, R., & Hayes, S. C. (2009). Producing and recognizing analogical relations. *Journal of the Experimental Analysis of Behavior*, *91*, 105–126.

Lionello-DeNolf, K. M. (2009). The search for symmetry: 25 years in review. *Learning & Behavior*, *37*, 188–203.

Luciano, C., Gómez-Becerra, I., & Rodríguez-Valverde, M. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of the Experimental Analysis of Behavior*, *87*, 349–365.

Martinho, A., & Kacelnik, A. (2016). Ducklings imprint on the relational concept of "same or different". *Science*, *353*, 286–288.

O'Hora, D., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2004). Derived relational networks and control by novel instructions: A possible model of generative verbal responding. *The Psychological Record*, *54*, 437–460.

Peña, T., Pitts, R. C., & Galizio, M. (2006). Identity matching-to-sample with olfactory stimuli in rats. *Journal of the Experimental Analysis of Behavior*, *85*, 203–221. doi: 10.1901/jeab.2006.111-04

Poling, A. (1985). Reporting interobserver agreement: Another difference in applied and basic behavioral psychology. *Experimental Analysis of Human Behavior Bulletin*, *1*, 5–6.

Saunders, K. J., & Williams, D. C. (1998). Stimulus control procedures. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behaviour* (pp. 193–228). New York, NY: Plenum.

Scarf, D., Boy, K., Reinert, A. U., Devine, J., Güntürkün, O., & Colombo, M. (2016). Orthographic processing in pigeons (*Columba livia)*. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 11272–1127.

Schusterman, R. J., & Kastak, D. (1993). A California sea lion (*Zalophus Californius*) is capable of forming equivalence relations. *The Psychological Record*, *43*, 823–839

Sidman, M., Rauzin R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons and children. *Journal of the Experimental Analysis of Behavior*, *37*, 23–44.

Smirnova, A., Zorina, Z., Obozova, T., & Wasserman, E. (2015). Crows spontaneously exhibit analogical reasoning. *Current Biology*, *25*, 256–260.

Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, *27*, 195–232.

Stewart, I. (2016). The fruits of a functional approach for psychological science. *International Journal of Psychology*, *51*, 15–27. doi:10.1002/ijop.12184

Stewart, I., Barnes-Holmes, D., Smeets, P. M., & Roche, B. (2004). A functional analytic model of analogy using the relational evaluation procedure. *The Psychological Record*, *54*, 531–552.

Stewart, I., Barnes-Holmes, D., & Weil, T. (2009). Training analogical reasoning as relational responding. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: applications for learners with autism and other developmental disabilities* (pp. 257–279). Oakland, CA: New Harbinger.

Stewart, I., & McElwee, J. (2009). Relational responding and conditional discrimination procedures: An apparent inconsistency and clarification. *Behavior Analyst*, *32*, 309–318.

Stewart, I., & Roche, B. (2013). Relational frame theory: An overview. In S. Dymond & B. Roche (Eds.), *Advances in relational frame theory: Research and application* (pp. 51–72). Oakland, CA: Context Press.

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

Urcuioli, P. J. (2008). Associative symmetry, antisymmetry, and a theory of pigeons' equivalence-class formation. *Journal of the Experimental Analysis of Behavior*, *90*, 257–282.

Vonk, J. (2015). Corvid cognition: Something to crow about?. *Current Biology*, *25*, R69–R71. doi:10.1016/j.cub.2014.12.001

Wright, A. A., Cook, R. G., Rivera, J. J., Sands, S. F., & Delius, J. D. (1988). Concept learning by pigeons: Matching-to-sample with trial-unique video picture stimuli. *Animal Learning & Behavior*, *14*, 436–444.

Wright, A. A., & Delius, J. D. (1994). Scratch and match: Pigeons learn matching and oddity with gravel stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*, 108–112.

Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior*, *50*, 125–144.

Zentall, T. R., Wasserman, E. A., & Urcuioli, P. J. (2013). Associative concept learning in animals. *Journal of the Experimental Analysis of Behavior*, *101*, 130–151.