



# The long and the short of it: Rule-based relative length discrimination in carrion crows, *Corvus corone*



Felix W. Moll, Andreas Nieder\*

Animal Physiology, Institute of Neurobiology, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

## ARTICLE INFO

### Article history:

Received 8 January 2014

Received in revised form 30 July 2014

Accepted 8 August 2014

Available online 22 August 2014

### Keywords:

Corvid songbird  
Continuous quantity  
Rule switching  
Executive function  
Intermediate relation  
Intermediate size

## ABSTRACT

Birds and other nonhuman animals can choose the larger of two discrete or continuous quantities. However, whether birds possess the conceptual grasp and cognitive control to flexibly switch between relative more-or-less-than judgments remains elusive. We therefore tested carrion crows in a rule-based line-length discrimination task to flexibly select lines presented on a touchscreen according to their relative length. In the first experiment, the crows needed to discriminate a shorter from a longer line, and vice versa. In the second experiment, the crows were required to choose a medium long line among three lines of different length (intermediate-size task). The crows switched effortlessly between “longer than/shorter than” rules, showing no signs of trial history affecting switching performance. They reliably chose the relatively longer and shorter line length, thus demonstrating a concept of greater than/less than with a continuous magnitude. However, both crows failed to discriminate a line of ‘medium’ length embedded in longer and shorter lines. These results indicate that relational discrimination exhibits different cognitive demands. While a greater than/less than concept requires only one relational comparison (with the respectively greater or smaller magnitude), the discrimination of a ‘medium’ magnitude demands to relate two or more comparisons, which might overburden crows and maybe animals in general.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

It is well established that nonhuman animals can choose between discrete numbers of items (numerosity) and make cardinal judgements. The earliest scientific evidence for quantitative cognitive capabilities in animals dates back to the 1930s and 1940s, when Koehler (1941, 1951) and his students pioneered this field of research. Interestingly, Koehler worked almost exclusively with birds to show that pigeons, corvids and parrots discriminate set size. Since then, different forms of numerical competence have been reported in pigeons (Xia et al., 2001; Emmerton and Renner, 2006, 2009; Scarf et al., 2011), chicks (Rugani et al., 2008), parrots (Pepperberg, 1994; Al Ain et al., 2009) and corvids (Zorina and Smirnova, 1996; Smirnova et al., 2000; Bogale et al., 2011). Like many other animals, birds spontaneously respond to quantity and use this information adaptively. In different avian species, quantity discriminations play a role in fighting brood parasitism (Lyon, 2003; White et al., 2009), caching food (Hunt et al., 2008), communicating about predators via alarm calls (Templeton et al., 2005), or recognizing individuals (Thompson, 1969).

Compared to representations of discrete quantities, much less attention has been given to the discrimination of continuous quantity, such as extent or length (Tudusciuc and Nieder, 2010). This is surprising, given that size information clearly is also of adaptive value and spontaneously exploited by animals in the wild in foraging situations (Al Ain et al., 2009; Beran, 2010) and mate choice. For instance, the size of body features, such as tail length (Moller et al., 1998; Moller, 1990; Andersson, 1992), is a prominent fitness signal in many bird species. Birds should thus be able to judge relative extent.

Simply being able to discriminate continuous quantity, however, is not by itself an advantage. In addition to quantity comparisons, items need to be chosen according to behavioral principles, or rules, that describe the conditional logic of a goal-directed task (Cantlon and Brannon, 2005; Bongard and Nieder, 2010). For instance, fish can be trained to prefer an absolute object size but seem to lack a concept of relative size (Mark and Maxwell, 1969), whereas American crows are able to choose objects based on relative size by following a single relational rule like ‘choose larger one’ (Coburn, 1914).

The ability to recognize the relationship between two (or more) stimuli irrespective of the precise appearance of those stimuli is classically investigated in transposition tasks (Lazareva, 2012). In its simplest layout, a single pair of stimuli is presented to a subject, and the subject is required to indicate the relation between the

\* Corresponding author. Tel.: +49 7071 29 75347.

E-mail addresses: [felix.moll@uni-tuebingen.de](mailto:felix.moll@uni-tuebingen.de) (F.W. Moll), [andreas.nieder@uni-tuebingen.de](mailto:andreas.nieder@uni-tuebingen.de) (A. Nieder).

two stimuli along a previously indicated stimulus dimension. For instance, subjects are trained to always respond to the brighter, longer, or more numerous stimulus in a pair of stimuli. In a more elaborate form, subjects are presented with three stimuli that vary along a predefined dimension to select the middle stimulus ('intermediate size problem'). If the subjects continue to make the right choice even when new stimuli in such pairs of stimuli are shown, relational learning might be evident. Spence (1937), however, argued based on theoretical accounts that preference for a relationally correct stimulus would not require an animal to actually respond to the relations between stimuli. Rather, his discrimination theory predicts that reinforcement history and the subsequent generalization behavior alone (establishing a 'postdiscrimination generalization gradient' in his model) would predict an animal's choice in transposition experiments. In other words, because the discriminative stimulus associated with reward (S+) is always the brighter, longer, or more numerous stimulus in a pair of stimuli, generalization behavior would suffice to explain an animal's choice in transposition tasks.

We therefore reasoned that relational learning could be demonstrated – and simple generalization behavior based on reinforcement history could be excluded – if the discriminative stimulus associated with reward (S+) would not be fixed, but change from trial to trial according to task instructions. This would prevent the formation of sign-fixed postdiscrimination generalization gradients as predicted by Spence (1937) and provide evidence for relational learning. In addition, to avoid an impact of reinforcement history, multiple pairs of training stimuli and a wide range of testing pairs that include nonadjacent testing stimuli are essential (Lazareva, 2012).

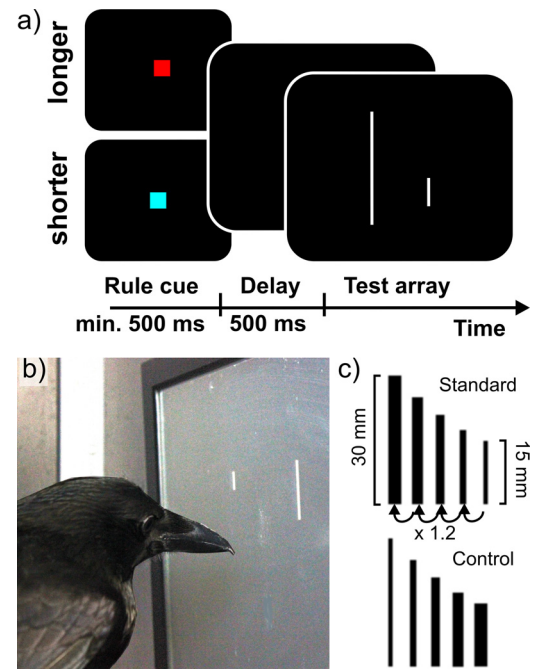
While transposition tasks with fixed discrimination rules are already demanding for animals, it is even more challenging for them to choose items according to switching rules (Stoet and Snyder, 2003; Vallentin et al., 2012; Eiselt and Nieder, 2013). Whether crows are able to switch flexibly between relational rules applied to magnitudes remains unknown. Rule switching is a hallmark of cognitive control functions and provides animals with superior behavioral flexibility. Corvids are well known for their impressive cognitive capabilities that rival even primates (Emery and Clayton, 2004). They manufacture and use tools (Hunt, 1996; Weir et al., 2002; Bird and Emery, 2009), represent hidden objects (Pollok et al., 2000; Zucca et al., 2007; Hoffmann et al., 2011), or take the status and prospective actions of conspecifics in social interactions into account (Emery and Clayton, 2001; Paz-Y-Miño et al., 2004). Moreover, corvids exhibit a high level of cognitive control and easily transfer relational rules to new sets of stimuli (Wilson et al., 1985).

In the present study, we therefore tested a corvid songbird, the carrion crow, in a rule-based line-length discrimination task. We investigated if and how crows can learn "longer than/shorter than" rules and apply them conceptually in the context of a longer and a shorter line. The crows' relational longer-versus-shorter performance was analyzed as a function of varying length ratios. An analysis of the trial-rule history was performed to detect potential performance decreases as a result of random trial-by-trial rule switches. In the final extension of the experiments, the discrimination of a 'middle' line length embedded in a longer and a shorter line was tested.

## 2. Methods

### 2.1. Subjects

Two tame male carrion crows (*Corvus corone*; bird T and bird M) weighting 540 g and 450 g were used. They were housed in a large indoor aviary (3.6 m × 2.4 m × 3.0 m) in a social group (Hoffmann



**Fig. 1.** Behavioral protocol and stimulus examples. (a) The crows initiated a trial by pecking on the colored 'longer/shorter than' rule cue on a touchscreen. After a 500 ms delay, two line stimuli appeared at randomly generated vertical positions. The crows had to peck the longer or shorter line, respectively, depending on the previous rule cue to gain a food reward. (b) Photograph of a working crow in front of the touchscreen during line stimulus presentation. (c) The standard (line area increases with length) and control line stimulus sets (line area decreases with length) from which the two displayed lines were randomly chosen per trial. Adjacent line stimuli differed in length by factor 1.2.

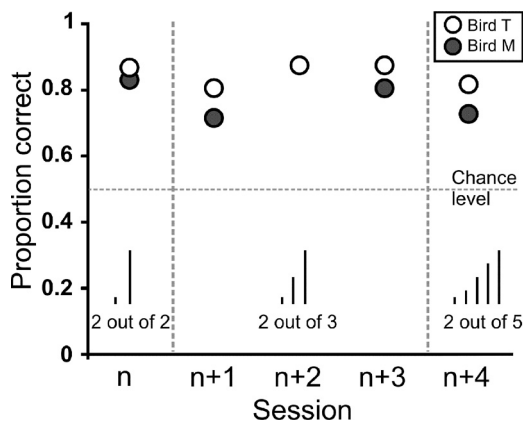
et al., 2011). The crows were taken from the Institute's breeding stock (Animal Physiology, University of Tübingen) at three weeks of age in June 2011 (bird T) and June 2012 (bird M) and raised by hand. The crows' age was 23 (bird T) and 11 (bird M) months at the time of testing. During behavioral sessions, they were maintained on a controlled feeding protocol and earned food during and after the daily tests. Both crows were equipped with leather jesses to secure them to their perch during experimental sessions in front of a touchscreen. All procedures were carried out according to the guidelines for animal experimentation and approved by the local authorities, the Regierungspräsidium Tübingen, Germany.

### 2.2. General procedure

The psychophysical set-up was composed of a fully controlled opaque operant conditioning chamber containing a 15 inch touchscreen (Fig. 1b; 3 M microtouch, 60 Hz refresh rate), a custom-made feeder, one night vision infrared video camera (iSImim321R Genius) and a wooden perch for the crows to stand on. Leather jesses secured the crows loosely to their perch. Apart from that, crows were able to move freely and could easily reach the touchscreen with their beak. Correct pecks on the screen were rewarded automatically with birdseed pellets and mealworms (*Tenebrio molitor* larvae). During reward phases, a small light integrated in the feeder lit up as additional positive feedback. For stimulus presentation and behavioral monitoring, personal computers running the CORTEX program (NIH) were used.

### 2.3. Behavioral protocol 1: longer/shorter than rule task

Our two crows were trained on a rule switching, line-length discrimination task. For this, the birds were placed in front of a



**Fig. 2.** Generalization performance of both birds from absolute (left) to relative (right) line length. Three successive test situations are shown (separated by dotted vertical lines): when two out of two lines were presented (left, training step 1), when two out of a set of three line lengths were shown (middle, training step 2), and when two out of a set of five line lengths were displayed (right, training step 3).

touchscreen monitor on which the rule cues and line stimuli were presented (Fig. 1a). A trial started when a colored rule cue ( $9 \times 9$  mm edge size; 500 ms in duration) appeared on the screen. A red square indicated that the longer-than rule was in effect, whereas a blue square signified the shorter-than rule. The two rule cues were presented pseudorandomly interleaved. The crows initiated a trial by pecking on the rule cue that disappeared thereafter. Next followed a delay (500 ms duration; black screen). Then, two visual target lines appeared side-by-side (distance 75 mm) at pseudorandomly generated vertical positions on the screen (with a jitter of  $\pm 33$  mm from center). Depending on the previously shown rule, the crows had to peck either the shorter (after the blue “shorter than”-rule) or the longer line (red “longer than”-rule) to gain a food reward. Wrong choices resulted in omission of food reward, a flashing gray feedback screen (500 ms duration) and a 1 s time-out during which the computer remained dormant. Every new trial started after a 2.5 s intertrial interval (black screen).

To ensure that the crows based their choice on the relative and not the absolute length of a given stimulus, the two line stimuli per trial were pseudo-randomly chosen from a set of five different line lengths (Fig. 1c; 15, 18, 21, 25 and 30 mm, adjacent stimuli differed by a factor of 1.19). However, the crows’ choice still could have been based on relative stimulus brightness rather than stimulus length. To control for this, we chose the stimuli from two different stimulus sets (Fig. 1c). In the standard set, the longest line was also the widest and therefore the brightest stimulus in the set, whereas in the control set the shortest line was the widest and thereby the brightest stimulus. Therefore, a choice based only on stimulus brightness (or width) would have resulted in chance performance. During the training phase (Fig. 2, training steps 1–3) we used a stimulus set, which was different from the stimulus sets depicted in Fig. 1c. The lines of these training stimuli were 7, 14, 28, 42 and 56 mm long and adjacent stimuli differed by the factors 1.33, 1.5 or 2.0. At all times, all relevant task parameters were balanced.

#### 2.4. Behavioral protocol 2: medium-length rule task

In the medium-length rule task, three line stimuli per trial were presented. In the first version of the medium-length rule task, three different rule cues (shorter: blue, medium: yellow and longer: red) were used. The line stimuli were pseudo-randomly chosen from the previously used longer/shorter rule task stimulus set (Fig. 1c, standard set) and were presented within a virtual triangle around the touch-screen’s center at random vertical positions. Following the rule, the shortest, medium or longest of the three presented

stimuli had to be chosen to earn a food reward. The three rule cues were presented pseudorandomly interleaved.

In the second, simpler, version of the medium-length rule task crows always had to choose the line stimulus of medium length. The longer/shorter than rule cues were not used any more. To make the task even easier, the three presented line stimuli per trial were pseudo-randomly chosen from a new stimulus set (Fig. 5; 3.5, 10, 20, 35 and 60 mm). Adjacent stimuli differed by the factor 1.7 at least and were all of the same width. To make length comparisons as easy as possible, lines were presented with their lower ends placed upon the same y-plane.

#### 2.5. Data analysis

For the behavioral protocol 1 (longer/shorter than rule task), we analyzed 20 consecutive sessions for each bird, each of which consisted of 160 trials. Behavioral protocol 2 (medium-length rule task) was analyzed over five consecutive sessions, each of which consisted of about 300 trials. We derived the median reaction times for single sessions and report the average of these medians over sessions throughout the paper. To analyze performance levels for different stimulus conditions within single sessions, the percent correct responses were derived and tested for individual crows using binomial tests. For all other comparisons (of data distributions), nonparametric tests (Wilcoxon and Friedman tests for paired data) were applied using SPSS 22 (IBM, Armonk, USA).

For our probability distributions (Fig. 6b–c) that described the correct response distribution for the medium-length rule task, line lengths were valued by their probability to appear as the correct (i.e. ‘medium’) stimulus. Line lengths ‘1’ and ‘5’ were never the correct line of medium length (0% probability), while lines ‘2’ and ‘4’ were correct in 30% of all trials and line ‘3’ was correct in 40% of the cases. Based on the latter five values (shown as ‘original’ distribution in Fig. 6a, filled circles), the correct response probability for the ten possible line-lengths-triplet displays was calculated (Fig. 6b). For example, the correct response probability for the line lengths triplet ‘1-2-3’ (Fig. 6b, first bar) was calculated as follows. The correctness probabilities 0% (line ‘1’), 30% (line ‘2’) and 40% (line ‘3’) were each divided by their sum, resulting in the response probability values 0% (‘1’), 42.9% (‘2’) and 57.1% (‘3’), i.e. now 100% in sum. Therefore, for the triplet ‘1-2-3’, our approach predicts a probability of 42.9% that the correct line stimulus of medium length (‘2’) is chosen (Fig. 6b, first bar). In addition, we fitted a Gaussian distribution ( $\sigma = 0.94$ ) to the original probability values (Fig. 6a, open circles). The same calculations as described above were then done with the resulting fitted correctness probabilities. This resulted in a more appropriate response distribution (Fig. 6c), which largely mirrored the crows’ real behavioral response distribution (Fig. 6d).

### 3. Results

#### 3.1. Generalization from absolute to relative line length

Two crows were trained to flexibly choose either the shorter or longer line from a display of two lines based on randomly switching longer than/shorter than rules. During the course of the initial training, we tested whether the crows could immediately generalize from absolute line lengths judgments to relative length discriminations. In the first training step (Fig. 2, left section), the crows had to choose between two test lines of constant length. During this phase, the crows still succeeded by responding to the absolute length of the lines. After reaching the performance criterion of 80.0% correct over one session, the crows were required to choose the longer or shorter, respectively, of two lines taken from a set of three line lengths. Depending on the stimulus combination, the stimulus of

middle length could now be the relatively longer or shorter of the two displayed lines, forcing the crows to make relative lengths judgments.

When switching from training step 1 (two out of two line lengths, Fig. 2, left section) to training step 2 (two out of three line lengths, Fig. 2, middle section), both birds performed significantly above chance for all three out of the three possible lengths combinations in the very first session (first-session data) of training step 2 (Fig. 2, middle section; pooled performance bird T: 80.5%; bird M: 70.1%; all stimulus combinations  $p < 0.001$ ; Binomial test; two-tailed; minimum number of trials per stimulus combination in either bird = 86 – maximum number of trials = 157).

To arrive at the final version of the task (training step 3; Fig. 2, right section), two test lines were pseudorandomly chosen from a set of five possible lengths. Both crows performed above chance in the very first session (first-session data pooled across all length combinations) (bird T: 81.4%; bird M: 72.2% correct performance;  $p < 0.001$ , binomial test, two-tailed). In bird T, performance for nine out of the ten possible lengths combinations was significant ( $p < 0.05$ , binomial test, two-tailed, minimum number of trials per combination = 31 – maximum number of trials = 45), with the exception of the line-length combination with the smallest length ratio (56 mm versus 42 mm;  $p > 0.05$ , Binomial test, two-tailed, number of trials = 49). In bird M, performance for 6 out of the ten possible lengths combinations was significant ( $p < 0.05$ , binomial test, two-tailed, minimum number of trials per combination = 19 – maximum number of trials = 27), with the exception of the four line-length combination with the smallest length ratio (56 mm vs. 42 mm; 42 mm vs. 28 mm, 28 mm vs. 14 mm, 42 mm vs. 14 mm;  $p > 0.05$ , Binomial test, two-tailed, minimum number of trials per combination = 26 – maximum number of trials = 36).

For all subsequent sessions ( $n = 20$  sessions), both crows significantly discriminated lines in all of the ten possible length combinations and reliably switched between the two rules (bird T data pooled across all length combinations: 84.2% correct performance; bird M: 84.7%; all stimulus combinations  $p < 0.001$ , one sample Wilcoxon test, test value = 0.5; two-tailed).

Performance was moderately better with the standard stimuli than with the control stimuli (Fig. 1c) for both crows (bird T: 86.2% versus 82.5%; bird M: 87.6% vs. 82.2%) (bird T: Wilcoxon test, two-tailed,  $U = -2.417$ ,  $p = 0.016$ ,  $n = 20$ ; bird M: Wilcoxon test, two-tailed,  $U = -3.659$ ,  $p < 0.001$ ,  $n = 20$ ). We suspect this mild bias to be a reflection of more intensive experience with the standard stimuli during initial training. Overall, these results confirm that the crows

reliably discriminated the length of the lines rather than low-level visual parameters.

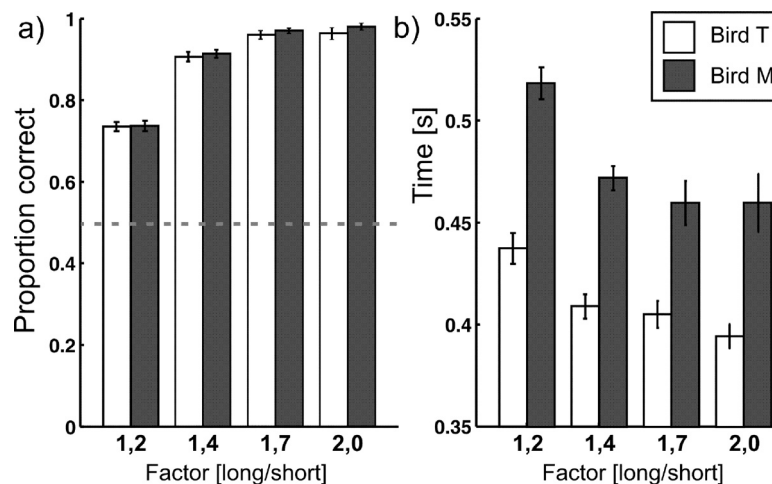
Reaction times (RTs) differed in both crows for correct versus error trials. In bird T, the average median RT (derived from the RT-medians of 20 single sessions) of 415 ms (standard deviation (SD) = 28 ms) in correct trials increased to 465 ms (SD = 37 ms) in error trials (Wilcoxon test, two-tailed,  $U = -3.920$ ,  $p < 0.001$ ,  $n = 20$ ). In bird M, the average median RT of 484 ms (SD = 26 ms) in correct trials increased to 601 ms (SD = 106 ms) in error trials (Wilcoxon test, two-tailed,  $U = -3.584$ ,  $p < 0.001$ ,  $n = 20$ ).

### 3.2. Performance varies with stimulus length ratio

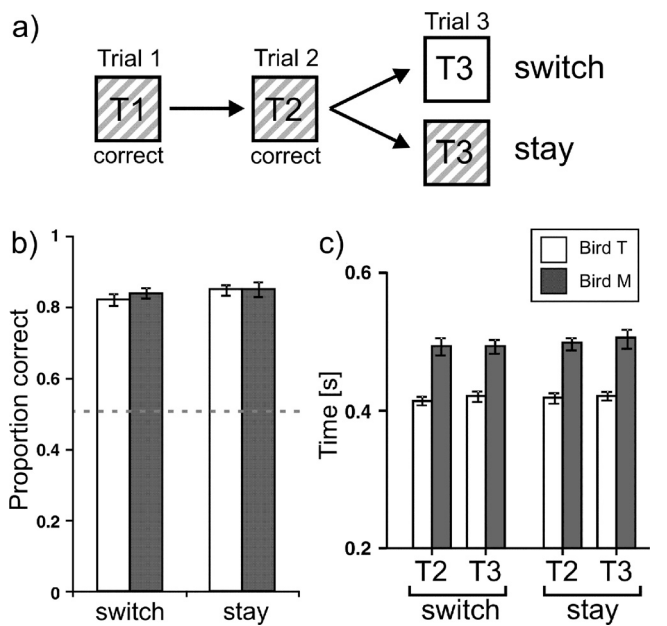
Next, we investigated whether the ratio between the lengths of the two test lines had an impact on the crows' performance. The crows' performance improved systematically with increasing relative length differences between the test lines (Fig. 3a). Because the two lines were chosen from a set of five different lengths, the relative length of line pairs differed by the factors (ratio) of 1.2, 1.4, 1.7 and 2.0 (Fig. 3a). The difference in performance between these four line ratios was significant (bird T: Friedman test,  $X^2(df 3) = 52.77$ ,  $p < 0.001$ ,  $n = 20$ ; bird M: Friedman test,  $X^2(df 3) = 58.43$ ,  $p < 0.001$ ,  $n = 20$ ) and reliably present in every session. Performance between length ratios 1.7 and 2.0 was almost perfect (96.0–98.0%, Fig. 3a) (ceiling effect). Additionally, the crows' reaction times in correct trials differed significantly between the four line ratios (Fig. 3b; bird T: Friedman test,  $X^2(df 3) = 22.47$ ,  $p < 0.001$ ,  $n = 20$ ; bird M: Friedman test,  $X^2(df 3) = 21.0$ ,  $p < 0.001$ ,  $n = 20$ ). Average median reaction time was at a clear maximum (bird T: 437 ms, SD = 34 ms; bird M: 518 ms, SD = 35 ms) for trials from the smallest line ratio 1.2 and at a minimum (bird T: 394 ms, SD = 26 ms; bird M: 460 ms, SD = 64 ms) for the highest ratio 2.0 (Fig. 3b).

### 3.3. Rule-switching performance is independent of trial history

Longer than/shorter than rules were presented in a pseudorandom order on a trial-by-trial basis. The crows switched between the longer than and the shorter than rule in a perfectly flexible manner. Both rules were mastered with equal proficiency (longer rule bird T: 85.7% and bird M: 84.8%; shorter rule bird T: 83.1% and bird M: 84.9%) (bird T: Wilcoxon test, two-tailed,  $U = -1.40$ ,  $p = 0.161$ ,  $n = 20$ ; bird M: Wilcoxon test, two-tailed,  $U = -0.261$ ,  $p = 0.794$ ,  $n = 20$ ). Therefore, there was no bias to preferentially pick the longer or shorter stimulus, just as there was no bias to



**Fig. 3.** Behavioral discrimination performance of both crows as a function of line lengths ratios. (a) The crows' correct performance level systematically increased with increasing line lengths ratio (i.e. factor). (b) Reaction times were highest for the smallest line lengths ratio, and lowest for the largest ratio, respectively. Error bars indicate the standard error of mean.



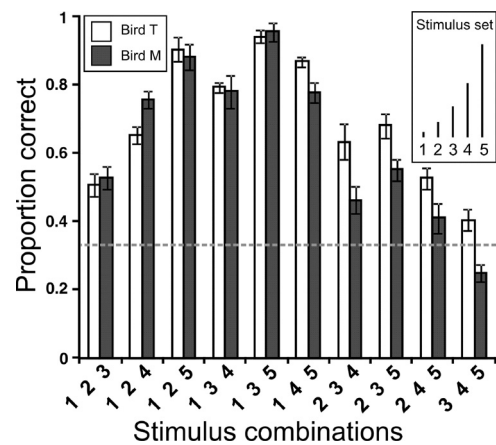
**Fig. 4.** Influence of rule switch on correct performance and reaction time. (a) Schematic layout of rule-type sequence. Every square (T1–T3) indicates one trial. The squares background pattern indicates the rule type (longer or shorter than). If the rule in the third trial (T3) was different from the second trial (T2), a “switch trial” was in effect, otherwise a “stay trial”. (b) Correct performance in switch and stay trials was unchanged. (c) Reaction time in switch and stay trials were unchanged for both crows.

preferentially pick the left (bird T: 84.0% and bird M: 84.7%) or right (bird T: 84.6% and bird M: 85.1%) stimulus on the screen (bird T: Wilcoxon test, two-tailed,  $U = -0.566$ ,  $p = 0.571$ ,  $n = 20$ ; bird M: Wilcoxon test, two-tailed,  $U = -0.373$ ,  $p = 0.709$ ,  $n = 20$ ).

To explore a putative effect of the trial history on performance, we compared correct performance and reaction times for trials of a given rule preceded by the same rule, (‘stay-trials’; e.g. a longer-than trial preceded by a longer-than trial) with trials in which the rule switched relative to the previous trial (‘switch-trials’; e.g. a longer-than trial after a shorter-than trial) (Fig. 4a). Correct performance in stay-trials (84.8% and 84.4%, bird T and bird M) was comparable to switch-trials (83.8% and 84.9%) (bird T: Wilcoxon test, two-tailed,  $U = -0.635$ ,  $p = 0.526$ ,  $n = 20$ ; bird M: Wilcoxon test, two-tailed,  $U = -0.299$ ,  $p = 0.765$ ,  $n = 20$ ) (Fig. 4b). The same was true for average median reaction times (Fig. 4c) in stay (bird T: 415 ms,  $SD = 27$  ms; bird M: 491 ms,  $SD = 63$  ms) and switch trials (bird T: 417 ms,  $SD = 31$  ms; bird M: 479 ms,  $SD = 48$  ms). Neither the rule type nor the trial history had a discriminative impact on the crows’ rule-switching behavior.

### 3.4. Failure to choose a “medium” line stimulus

After the crows had acquired the two longer than/shorter than rules, a third rule indicating the instruction to choose medium line length (second longest or shortest, respectively) was introduced. Instead of two test lines, three test lines were now shown. As long as three lines of constant lengths were shown (absolute length), the crows could discriminate the medium long line. However, when the three presented line stimuli per trial were evenly chosen from a set of five lines in succeeding generalization sessions, performance for the medium-length rule reached chance level of 33%. At the same time, only a slight drop in performance on the longer than and shorter than rule was found. In fact, the crows were perfectly able to choose the shortest or longest line out of three presented stimuli based on relative line length, whereas they were unable to discriminate the line of medium length.



**Fig. 5.** Performance in the ‘medium’-length rule task. Crows were unable to reliably choose the stimulus of medium length from a three-line stimulus test display. Error bars indicate the standard error of mean. The dashed line indicates chance level.

To help the crows to acquire a concept of “middle”, the task protocol was simplified by removing the rule cue and simply having the crows choose the medium long line out of a set of three lines. To reduce the task protocol even more, the lower ends of the three test lines were no longer presented at variable vertical positions, but always centered on the same horizontal plane. Moreover, the length ratios between the test lines were increased to differ by a factor of at least 1.7. Despite of all these simplifications, the crows had considerable difficulties to discriminate a stimulus of medium length. Fig. 5 presents correct performance for all ten possible line-length combinations averaged over five sessions of training on this simplified task. Performance averaged over these five sessions was only 64.0% and 55.0% correct (bird T and M). Correct performance depended strongly on line lengths combination (Fig. 5; bird T: Friedman test,  $X^2(df 9) = 39.34$ ,  $p < 0.001$ ,  $n = 5$ ; bird M: Friedman test,  $X^2(df 9) = 41.25$ ,  $p < 0.001$ ,  $n = 5$ ), with about chance performance at both ends of the line-length range, indicating that the crows’ choice was based on strong preferences for absolute rather than relative line length. This poor performance is likely not attributed to a lack of training because bird T had been trained for fourteen sessions and bird M had completed five training sessions prior to the collection of these data.

## 4. Discussion

The main findings of the current study are that (1) carrion crows grasp the concept “longer” and “shorter” applied to the lengths of lines, (2) switch effortlessly between “longer than/shorter than” rules, but (3) fail to learn to discriminate a line of “medium” length embedded in a longer and a shorter line. Critical aspects related to these findings will be discussed in the following section.

### 4.1. Scope and limit of relational length discrimination

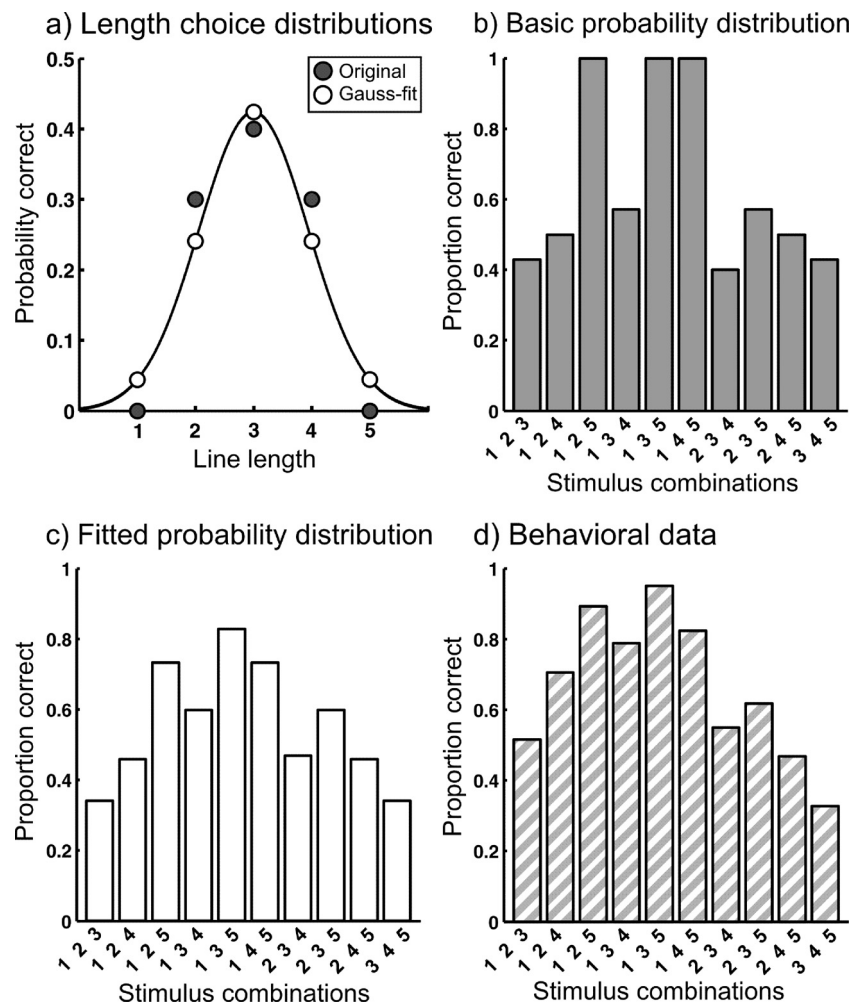
Our experiments show that crows master quantity discriminations based on absolute (fixed) length and grasp a concept of relative length. When faced with lines of variable length, they reliably chose the longer or shorter one, respectively, according to task contingencies. An analysis of accuracy as a function of whether the continuous variable of surface area was congruent versus incongruent with length showed that the crows’ behavior was controlled by length. As with classical finding for magnitude discriminations, the crows showed a length ratio dependent performance signature: Accuracy values were positively correlated with the line ratio, whereas reaction times were negatively correlated. Therefore, performance improved with greater differences between lengths, thus

reflecting a quantitative distance effect. Such a distance effect has also been reported in African gray parrots discriminating food amounts (Al Ain et al., 2009). We suspect this effect to be a reflection of a Weber Law signature classically present in a variety of magnitude discriminations, namely the finding that the discriminability of two magnitudes depends on their ratio, not absolute magnitude differences.

As mentioned in the introduction, Spence (1937) in his discrimination theory convincingly argued that successful behavior in transposition tasks may not be the result of relational discriminations. He showed that reinforcement history and the subsequent generalization behavior alone would predict an animal's choice in transposition experiments. In his model, excitatory (around the S+ stimulus associated with reinforcement) and inhibitory (around the S- stimulus associated with non-reinforcement) 'postdiscrimination generalization gradients' are becoming established with training and learning history. Importantly, the S+ and the S- are fixed relative to one another on the discriminative stimulus dimension. We think that – in addition to the wide range of standard and control testing pairs that included nonadjacent testing stimuli (Fig. 1c) – our approach of combining a length transposition task with a rule-switching task is well suited to avoid such a non-relational explanation. This is because in our experiment the S+ and S- were randomly and constantly exchanged according to the rule

that was active for any single trial. Thus, the formation of sign-fixed postdiscrimination generalization gradients as predicted by Spence (1937) would have left the crows at chance performance, because in half of the trials the reinforcement would not have been associated with the S+ but with the S-, and vice versa. We therefore argue that the discrimination behavior of our crows is best explained by relational discriminations. Our result of relational discriminations in crows is in agreement with more recent findings in pigeons that – albeit for different experimental reasons – also showed behavior based on relational factors (Lazareva et al., 2005, 2008; Lazareva, 2012).

The concept of relative length, however, seems to be restricted to the endpoints of the length continuum, because our crows failed to learn to discriminate a line of medium lengths amidst longer or shorter lines. Even after considerable simplifications of the stimuli and several days of training, performance was still around chance level and showed unexpected response distributions (Fig. 5). We therefore examined whether the probabilities of certain absolute line lengths to be correct 'medium' choices could explain the crows' behavior. With a set of five different line lengths as used in our study, the middle lengths of this set are more likely to be the correct choices (see 'original' distribution in Fig. 6a, filled circles) than the longer or shorter lines from this set. When we used this probability distribution to calculate the likelihood that the medium long



**Fig. 6.** Evaluation of 'medium' length discriminations based on the probability of certain absolute line lengths to be correct choices. (a) Original probabilities of the five used line stimuli to be the correct 'medium' choices (filled circles) and Gauss fitted correctness probabilities (open circles). (b) Basic response probability distribution resulting from the original correctness probabilities (a, filled circles). (c) Fitted response probability distribution resulting from the Gauss fitted correctness probabilities (a, open circles). (d) The crows' real behavioral response distribution (see also Fig. 5).

lines (lines 2–4) amidst the possible combinations with a longer and a shorter line were the correct choice, the resulting response distribution (Fig. 6b) showed characteristics of the crows' real performance (seen in Fig. 6d). Even more, when a Gauss function fitted to the original probability distribution was used (Fig. 6a, open circles), the calculated response distribution (Fig. 6c) largely mirrored the crows' real behavioral performance. Based on these clear predictions, we conclude that the crows were not discriminating the middle line to solve the medium-length rule task, but rather followed a preference to choose the absolute line length "3". Nevertheless, we can not exclude that more intense training for more than 18 sessions in bird T and 10 sessions in bird M might have enhanced the crows' intermediate size performance. Moreover, prior training to the longer/shorter than rule task could have deteriorated the crows' ability to learn an intermediate-size task. Still, as discussed below, poor performance of tested animals in intermediate-size tasks seems to be a general trend that does not seem to vanish after more intense training or a certain training history.

Solving intermediate size tasks not only seems difficult for our crows, but for animals in general. When Spence (1942) investigated the responses of chimpanzees on an "intermediate size problem" the animals seemed to respond on the basis of the absolute properties of the stimulus rather than the relative ones. Gonzalez et al. (1954) extended Spence's study and used a broader range of stimuli. A preference for the middle-sized stimuli suggesting preference for the relative properties was detected, but repeated test trials with the same two stimulus sets can not exclude that specific length configurations might have been learned by the chimpanzee. This protocol was later repeated with rhesus monkeys to find a similar above chance preference for the middle-sized stimuli (Gentry et al., 1959; Brown et al., 1959). Squirrel monkeys also seem to have problems to grasp the middle size relation. In the study by McGonigle and Jones (1978), monkeys failed to choose the object of medium size or brightness from a three-object display. 'Middleness' requires relational coordination, as it is required in transitive reasoning like 'when B is larger than A and C is larger than B, then C is largest' (McGonigle and Chalmers, 1977). It was suggested that monkeys do not logically coordinate the two available pieces of information in a transitive task. Instead, their choices could result from single binary decision making and are well predicted by a straightforward 'binary (statistical) decision model' (McGonigle and Chalmers, 1977).

Both in our and the above cited primate studies, animals were tested for inference of explicitly quantitative aspects. Interestingly, transitive inference of ordinal relations (rank position in a sequence) has been demonstrated multiple times in different animal species (Bond et al., 2003; Grosenick et al., 2007; Gazes et al., 2012). This may suggest that transitive inference of rank information is widespread in the animal kingdom, whereas intermediate-size transposition, i.e. transitive inference based on quantitative information (Merten and Nieder, 2009), is much more difficult for all animals including primates and crows.

#### 4.2. Rule-switching performance and trial history

In our study, crows managed to switch effortlessly from the rule "pick the longer line" to "pick the shorter line". Switching on demand from one cognitive task to another, thus rapidly selecting the appropriate course of action, is a fundamental manifestation of cognitive (executive) control. In mammals, this feat is primarily attributed to the prefrontal cortex (Miller and Cohen, 2001; Stoet and Snyder, 2009; Bongard and Nieder, 2010). Birds, however, lack a neocortex and thus also a prefrontal cortex altogether. Instead, they possess an integrative endbrain structure, called nidopallium caudolaterale (NCL) that subserves analog functions based on convergent evolution (Güntürkün, 2005). Neurons in this brain region show sustained activity during working memory tasks (Veit et al.,

2014) and signal abstract 'match' versus 'nonmatch' rules in a behaviourally relevant manner (Veit and Nieder, 2013), suggesting that also the "longer than/shorter than" decisions investigated in the current study are represented in the crows' NCL.

Behavioral flexibility in humans comes at a price and causes switch costs, a worsening in performance associated with changing versus repeating the task performed on the previous trial within mixed-task blocks (Jersild, 1927; Rogers and Monsell, 1995). Thus, humans performing a rule-switching task show robust switch costs in reaction time of about 35 ms (7.4% of mean reaction time) (Stoet and Snyder, 2007). Even extensive training (up to 30,000 trials) neither abolishes nor systematically reduces switch costs in humans (Stoet and Snyder, 2007).

In contrast to human data, our crows did not show any discriminative switch costs during the longer than/shorter than rule-switching trials. An absence of switch costs was also found in monkeys performing a rule-switching task fully comparable to human protocols (Stoet and Snyder, 2003, 2007). However, when Stoet and Snyder (2003) decreased the intertrial interval down to 170 ms, significant switch costs began to appear in monkey performance. Recently, Caselli and Chelazzi (2011) did report switch costs in monkeys that were comparable in size to those paid by humans. This indicates that switch costs are present in animals, but are very sensitive to the precise task contingencies applied. Whether corvids would still show switch costs under different task circumstances thus remains an unresolved question.

#### Author contributions

F.W.M. and A.N. designed the experiments and wrote the manuscript. F.W.M. trained the crows and analyzed the data. A.N. assisted in crow training and data analysis.

#### Competing financial interest

The authors declare no competing financial interests.

#### Acknowledgment

This work was supported by a DFG grant NI 618/3-1 to A.N.

#### References

- Al Ain, S., Giret, N., Grand, M., Kreutzer, M., Bovet, D., 2009. The discrimination of discrete and continuous amounts in African grey parrots (*Psittacus erithacus*). *Anim. Cogn.* 12, 145–154.
- Andersson, S., 1992. Female preference for long tails in lekking Jackson widowbirds – experimental evidence. *Anim. Behav.* 43, 379–388.
- Beran, M.J., 2010. Chimpanzees (*Pan troglodytes*) accurately compare poured liquid quantities. *Anim. Cogn.* 13, 641–649.
- Bird, C.D., Emery, N.J., 2009. Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10370–10375.
- Bogale, B.A., Kamata, N., Mioko, K., Sugita, S., 2011. Quantity discrimination in jungle crows *Corvus macrorhynchos*. *Anim. Behav.* 82, 635–641.
- Bond, A.B., Kamil, A.C., Balda, R.P., 2003. Social complexity and transitive inference in corvids. *Anim. Behav.* 65, 479–487.
- Bongard, S., Nieder, A., 2010. Basic mathematical rules are encoded by primate prefrontal cortex neurons. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2277–2282.
- Brown, W.L., Overall, J.E., Gentry, G.V., 1959. 'Absolute' versus 'relational' discrimination of intermediate size in the rhesus monkey. *Am. J. Psychol.* 72, 593–596.
- Cantlon, J.F., Brannon, E.M., 2005. Semantic congruity affects numerical judgments similarly in monkeys and humans. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16507–16511.
- Caselli, L., Chelazzi, L., 2011. Does the macaque monkey provide a good model for studying human executive control? A comparative behavioural study of task switching. *PLoS ONE* 6, e21489, <http://dx.doi.org/10.1371/journal.pone.0021489>.
- Coburn, C.A., 1914. The behaviour of the crow, *Corvus americanus*. *J. Anim. Behav. Boston* 4, 185–201.
- Eiselt, A.K., Nieder, A., 2013. Representation of abstract quantitative rules applied to spatial and numerical magnitudes in primate prefrontal cortex. *J. Neurosci.* 33, 7526–7534.

- Emery, N.J., Clayton, N.S., 2001. Effects of experience and social context on prospective caching strategies by scrub jays. *Nature* 414, 443–446.
- Emery, N.J., Clayton, N.S., 2004. The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306, 1903–1907.
- Emmerton, J., Renner, J.C., 2006. Scalar effects in the visual discrimination of numerosity by pigeons. *Learn. Behav.* 34, 176–192.
- Emmerton, J., Renner, J.C., 2009. Local rather than global processing of visual arrays in numerosity discrimination by pigeons (*Columba livia*). *Anim. Cogn.* 12, 511–526.
- Gazes, R.P., Chee, N.W., Hampton, R.R., 2012. Cognitive mechanisms for transitive inference performance in rhesus monkeys: measuring the influence of associative strength and inferred order. *J. Exp. Psychol. Anim. Behav. Process.* 38, 331.
- Gentry, G.V., Overall, J.E., Brown, W.L., 1959. Transpositional responses of rhesus monkeys to stimulus - objects of intermediate size. *Am. J. Psychol.* 72, 453–455.
- Gonzalez, R.C., Gentry, G.V., Bitterman, M.E., 1954. Relational discrimination of intermediate size in the chimpanzee. *J. Comp. Physiol. Psychol.* 47, 385–388.
- Grosenick, L., Clement, T.S., Fernald, R.D., 2007. Fish can infer social rank by observation alone. *Nature* 445, 429–432.
- Güntürkün, O., 2005. The avian 'prefrontal cortex' and cognition. *Curr. Opin. Neurobiol.* 15, 686–693.
- Hoffmann, A., Rüttler, V., Nieder, A., 2011. Ontogeny of object permanence and object tracking in the carrion crow (*Corvus corone*). *Anim. Behav.* 82, 359–367.
- Hunt, G.R., 1996. Manufacture and use of hook-tools by New Caledonian crows. *Nature* 379, 249–251.
- Hunt, S., Low, J., Burns, K.C., 2008. Adaptive numerical competency in a food-hoarding songbird. *Proc. R. Soc. B* 275, 2373–2379.
- Jersild, A.T., 1927. Mental set and shift. *Arch. Psychol.* (89), 5–82.
- Koehler, O., 1941. Vom Erlernen unbenannter Anzahlen bei Vögeln. *Naturwissenschaften* 29, 201–218.
- Koehler, O., 1951. The ability of birds to "count". *Bull. Anim. Behav.* 9, 41–45.
- Lazareva, O.F., Wasserman, E.A., Young, M.E., 2005. Transposition in pigeons: reassessing Spence (1937) with multiple discrimination training. *Learn. Behav.* 33, 22–46.
- Lazareva, O.F., Miner, M., Wasserman, E.A., Young, M.E., 2008. Multiple-pair training enhances transposition in pigeons. *Learn. Behav.* 36, 174–187.
- Lazareva, O.F., 2012. Relational learning in a context of transposition: a review. *J. Exp. Anal. Behav.* 97, 231–248.
- Lyon, B.E., 2003. Egg recognition and counting reduce costs of avian conspecific brood parasitism. *Nature* 422, 495–499.
- Mark, R.F., Maxwell, A., 1969. Circle size discrimination and transposition behaviour in cichlid fish. *Anim. Behav.* 17, 155–158.
- McGonigle, B.O., Chalmers, M., 1977. Are monkeys logical? *Nature* 267, 694–696.
- McGonigle, B.O., Jones, B.T., 1978. Levels of stimulus processing by the squirrel monkey: relative and absolute judgements compared. *Perception* 7, 635–659.
- Merten, K., Nieder, A., 2009. Compressed scaling of abstract numerosity representations in adult humans and monkeys. *J. Cogn. Neurosci.* 21, 333–346.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Moller, A.P., et al., 1998. Sexual selection and tail streamers in the barn swallow. *Proc. R. Soc. B* 265, 409–414.
- Moller, A.P., 1990. Male tail length and female mate choice in the monogamous swallow *Hirundo rustica*. *Anim. Behav.* 39, 458–465.
- Paz-Y-Miño, C.G., Bond, A.B., Kamil, A.C., Balda, R.P., 2004. Pinyon jays use transitive inference to predict social dominance. *Nature* 430, 778–781.
- Pepperberg, I., 1994. Numerical competence in an african gray parrot (*Psittacus erithacus*). *J. Comp. Psychol.* 108, 36–44.
- Pollok, B., Prior, H., Güntürkün, O., 2000. Development of object permanence in food-storing magpies (*Pica pica*). *J. Comp. Psychol.* 114, 148–157.
- Rogers, R., Monsell, S., 1995. Cost of a predictable switch between simple cognitive tasks. *J. Exp. Psych.: Hum. Percept. Perform.* 124, 207–231.
- Rugani, R., Regolin, L., Vallortigara, G., 2008. Discrimination of small numerosities in young chicks. *J. Exp. Psychol. Anim. Behav. Process.* 34, 388–399.
- Scarff, D., Hayne, H., Colombo, M., 2011. Pigeons on par with primates in numerical competence. *Science* 334, 1664.
- Smirnova, A.A., Lazareva, O.F., Zorina, Z.A., 2000. Use of number by crows: investigation by matching and oddity learning. *J. Exp. Anal. Behav.* 73, 163–176.
- Spence, K.W., 1937. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.* 44, 430–444.
- Spence, K.W., 1942. The basis of solution by chimpanzees of the intermediate size problem. *J. Exp. Psychol.* 31, 257–274.
- Stoet, G., Snyder, L.H., 2003. Executive control and task-switching in monkeys. *Neuropsychologia* 41, 1357–1364.
- Stoet, G., Snyder, L.H., 2007. Extensive practice does not eliminate human switch costs. *Cogn. Affect. Behav. Neurosci.* 7, 192–197.
- Stoet, G., Snyder, L.H., 2009. Neural correlates of executive control functions in the monkey. *Trends Cogn. Sci.* 13, 228–234.
- Templeton, C.N., Greene, E., Davis, K., 2005. Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science* 308, 1934–1937.
- Thompson, N.S., 1969. Individual identification and temporal patterning in the cawing of common crows. *Commun. Behav. Biol.* 4, 29–33.
- Tudusciuc, O., Nieder, A., 2010. Comparison of length judgments and the Müller-Lyer illusion in monkeys and humans. *Exp. Brain Res.* 207, 221–231.
- Vallentin, D., Bongard, S., Nieder, A., 2012. Numerical rule coding in the prefrontal, premotor, and posterior parietal cortices of macaques. *J. Neurosci.* 32, 6621–6630.
- Veit, L., Nieder, A., 2013. Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. *Nat. Commun.* 4, 2878.
- Veit, L., Hartmann, K., Nieder, A., 2014. Neuronal correlates of visual working memory in the corvid endbrain. *J. Neurosci.* 34, 7778–7786.
- Weir, A.A.S., Chappell, J., Kacelnik, A., 2002. Shaping of hooks in New Caledonian crows. *Science* 297, 981.
- White, D.J., Ho, L., Freed-Brown, G., 2009. Counting chicks before they hatch: female cowbirds can time readiness of a host nest for parasitism. *Psychol. Sci.* 20, 1140–1145.
- Wilson, B., Mackintosh, N.J., Boakes, R.A., 1985. Transfer of relational rules in matching and oddity learning by pigeons and corvids. *Q. J. Exp. Psychol.* 37, 313–332.
- Xia, L., Emmerton, J., Siemann, M., Delius, J.D., 2001. Pigeons (*Columba livia*) learn to link numerosities with symbols. *J. Comp. Psychol.* 115, 83–91.
- Zorina, Z.A., Smirnova, A.A., 1996. Quantitative evaluations in gray crows: generalization of the relative attribute 'larger set'. *Neurosci. Behav. Physiol.* 26, 357–364.
- Zucca, P., Milos, N., Vallortigara, G., 2007. Piagetian object permanence and its development in Eurasian jays (*Garrulus glandarius*). *Anim. Cogn.* 10, 243–258.