

Large quantity discrimination by North Island robins (*Petroica longipes*)

Alexis Garland · Jason Low · Kevin C. Burns

Received: 31 October 2011 / Revised: 25 June 2012 / Accepted: 9 July 2012
© Springer-Verlag 2012

Abstract While numerosity—representation and enumeration of different numbers of objects—and quantity discrimination in particular have been studied in a wide range of species, very little is known about the numerical abilities of animals in the wild. This study examined spontaneous relative quantity judgments (RQJs) by wild North Island robins (*Petroica longipes*) of New Zealand. In Experiment 1, robins were tested on a range of numerical values of up to 14 versus 16 items, which were sequentially presented and hidden. In Experiment 2, the same numerical contrasts were tested on a different group of subjects but quantities were presented as whole visible sets. Experiment 3 involved whole visible sets that comprised of exceedingly large quantities of up to 56 versus 64 items. While robins shared with other species a ratio-based representation system for representing very large values, they also appeared to have developed an object indexing system with an extended upper limit (well beyond 4) that may be an evolutionary response to ecological challenges faced by scatter-hoarding birds. These results suggest that cognitive mechanism influencing an understanding of physical quantity may be deployed more flexibly in some contexts than previously thought, and are discussed in light of findings across other mammalian and avian species.

Keywords New Zealand robin · Quantity discrimination · Large number representations · Relative quantity judgment · Avian cognition

Quantity discrimination is fundamental to numerical cognition and considered an ability that is at the roots of symbolic counting (Carey 2001; Gallistel and Gelman 1992). Not surprisingly then, studies examining relative quantity judgments (RQJs) have been conducted with human infants and adults (e.g., Barth et al. 2003; Feigenson et al. 2002) and primates (e.g., Beran and Beran 2004; Hanus and Call 2007; Hauser et al. 2000) as well as other species ranging from insects to elephants (e.g., Agrillo et al. 2008; Carazo et al. 2009; Irie-Sugimoto et al. 2009; Meck and Church 1983).

Non-linguistic number sense is understood to exist in humans and animals in two distinct systems: an indexing system, which processes or stores “object files” with a strict capacity limit of three to four items, and an analog magnitude system used in approximating larger numbers (Uller 2008; Trick and Pylyshyn 1994). Weber’s law appears to play a fundamental role in the underlying mechanism with which the perception of large numbers is approximated (Cantlon et al. 2009; Gallistel and Gelman 2005). According to this law, the size of the smallest perceptible difference in stimulus intensity varies as a function of a proportion of the original magnitude of that stimulus (Jordan and Brannon 2006). Within the framework of an analog magnitude system, ratio is the primary predictor of accuracy in discriminating between different quantities. There is still debate over whether there are two distinct core number systems. Some studies with monkeys and human infants support different cognitive signatures governing small versus large numbers (see Feigenson et al.

A. Garland (✉) · J. Low
School of Psychology, Victoria University of Wellington,
P.O. Box 600, Wellington 6012, New Zealand
e-mail: Alexis.Garland@vuw.ac.nz

K. C. Burns
School of Biological Sciences, Victoria University
of Wellington, Wellington, New Zealand

2004), while other studies have found that a singular mechanism based on Weber's law accounts for RQJs across a wide range of numerosities (e.g., Beran 2001; Cantlon and Brannon 2006).

Despite increasing comparative research interest in quantity discrimination, understanding the ecological factors that determine when and how mental representations of quantity are deployed remains a challenge that needs to be addressed (Dehaene et al. 2008; Feigenson et al. 2002; Hanus and Call 2007). Our goal was to investigate spontaneous quantitative cognition skills in wild New Zealand robins, specifically North Island robins (*Petroica longipes*), an endemic species of New Zealand (Hunt et al. 2008). The bulk of the data on spontaneous numerical discrimination is still primate and mammal centric (Carazo et al. 2009; Dadda et al. 2009; Uller et al. 2003). Considering the evolutionary distance of birds and mammals, examining RQJs in a scatter-hoarding songbird will contribute to a more phylogenetically diverse picture of numerical cognition (Beran 2008; Clayton and Krebs 1995). There is work on numerical abilities in birds, but they involve considerable amounts of training and/or relate primarily to discrimination of sets with small sizes (Farnsworth and Smolinski 2006; Honig and Stewart 1989; Koehler 1941; Pepperberg 2006; Rugani et al. 2008; Terrace 1987). However, convincingly, studies demonstrate animal learning and cognition after extensive training, interpreting what the findings mean in terms of animals' *natural* ability to represent number remains difficult (Davis and Memmott 1982; McComb et al. 1994).

There are some studies that hint at wild birds' adaptive number sense. For example, American coots experience nest parasitism from conspecific females, and egg number combined with recognition contributes to clutch size decisions. Lyon (2003) found that female coots that discriminated parasitic eggs from their own counted only their own eggs and thus avoided inadvertent clutch size reductions. Quantity discrimination has also been suggested to partly underpin Brown-headed cowbirds' (*Molothrus ater*, an obligate brood parasite) synchronization of nest parasitism with host incubation readiness. White et al. (2009) found that cowbirds spent longer periods investigating nests with the correct accumulated number of eggs in relation to elapsed days, suggesting that cowbirds appear to remember changes in egg number between visits. The few studies on number cognition in wild birds have, however, focused on only a narrow range of numerical values and with small set sizes (Farnsworth and Smolinski 2006; Low et al. 2009). Some of these sets are still relatively "large" in terms of being above object-file threshold of 3–4 items, but not substantially large enough to establish whether there is a sophisticated upper limit to birds' number discrimination or what representational system could underpin

such capacities. This is important to address when assessing numerosity in birds where certain food-storing species have been shown to display significant spatial attention and memory skills (Clayton et al. 2005; Hunt et al. 2008).

North Island robins, the species of the current study, are one of a very small number of food-hoarding birds in the Southern Hemisphere (Vander Wall 1990). Robins are monogamous and mated pairs reside on exclusive territories year-round (Burns and Steer 2006). This species is a medium-sized insectivorous passerine that is endemic to New Zealand and found both on the North and South Island. Like many other animals native to isolated oceanic islands, robins lack pronounced anti-predatory behaviors and are fearless of humans. They will consume and cache foods offered to them by hand and readily attend to and interact in experimental paradigms without extensive familiarization or training (Hunt et al. 2008). They show complex cognitive abilities in creating, protecting, and retrieving caches. Robins cache food in numerous different sites, regularly caching more than one item in each site. Further, when robins retrieve their own or pilfer their mates' caches, they fly to the exact location of these cache sites, even after leaving them for significant periods of time—suggesting accurate observational spatial memory (see Menzies and Burns 2008, for a review). Examining numerosity discrimination in a scatter-hoarding songbird may uncover what roles phylogeny, ontogeny, and ecology play in shaping universal and species-specific features in the cognitive mechanisms underlying quantity representation.

Initial RQJ testing by Hunt et al. (2008) indicated that robins chose the cache site with greater number more than half the time with comparisons involving 1 versus 2, 2 versus 3, 3 versus 4, 4 versus 5, and 4 versus 8, but chose at chance or below with 4 versus 6; and 6 versus 8 and 8 versus 10. Their analysis indicated that total set size rather than ratio predicted the accuracy of robins' numerical judgments. Precise numerical judgments could potentially form an integral part of robins' cache retrieval strategy. If robins are able to compute precisely how many pieces of prey items are in each of its (and its mates') cache sites, it would help prioritize efficient cache retrieval and pilfering (Burns and Steer 2006). Moreover, given that robins only store insect prey that are highly perishable, tracking and remembering the precise number of items stored in particular cache sites would help prioritize cache retrieval to minimize losses to spoilage. Hunt et al.'s findings are illuminating in their contrast with work on primates and human infants showing that the object-file system, while capturing performance with small sets, cannot process quantities greater than 3 or 4 items simultaneously. The results suggest that robins—in the context of food-storing and food-retrieval demands—may have developed an object indexing system with a higher upper limit.

In relation to the present study, though, it is still not clear whether robins have relatively sophisticated discrimination of numbers in an absolute sense, or whether a single fuzzy analog magnitude system for recognizing all quantities—large or small—underpins RQJs. First, in Hunt et al.'s (2008) study, all comparisons with ratios at 0.50, half at 0.67 and 0.75, and none at 0.80 were significant. This suggests that ratio, to some extent, may influence robins' numerosity discrimination, but it was undetected due to the limited range of comparisons being examined. Second, Hunt et al. tested RQJs by presenting the two sets of quantities item by item whereby the totality of items was never visually available at the time of cache selection, as prey was dropped into opaque containers and subsequently covered. Research indicates that apes (Hanus and Call 2007) and South American sea lions (Abramson et al. 2011) are poorer at spontaneously selecting the larger value when the two sets of quantities were presented item by item compared to when the two quantities are presented simultaneously as visually accessible whole sets. Importantly, ratio between quantities was found to account for apes and sea lions' spontaneous RQJ performance—analogue magnitude representations appear to be especially favored when items are visibly presented as a united set (see Shettleworth 2010, Ch 10).

There is compelling evidence in some studies (Beran 2001; Beran and Beran 2004) that chimpanzees are able to make highly accurate RQJs (70–95 % correct) even when larger numbers of food items are sequentially presented and hidden from view (5 vs. 8, 5 vs. 10, and 6 vs. 10). However, as Beran and Beran note (2004), it is not clear whether certain subjects' rich history of participation in numerical and other cognitive tasks may have scaffolded their ability to update and retain quantity information when sets are sequentially presented. Overall, while each of these methods (quantities presented simultaneously as whole sets or quantities presented item by item) has been frequently used to test animals' RQJs, there are few systematic comparisons of spontaneous performance between these two types of tasks (Abramson et al. 2011; Hanus and Call 2007). It is noteworthy that wild robins in Hunt et al.'s study could make RQJs successfully in a cognitively demanding task context where the item-by-item presentation meant that the birds had to deal with temporal discontinuity between prey items. Hunt et al. have suggested that an extended upper limit to robins' parallel individuation of food items may reflect these birds' sensory, motor, cognitive, and neuroanatomical specializations in behavior for scatter-hoarding, pilfering, and retrieving food. However, emphasizing a sophisticated object indexing system operating in robins instead of a generalized and evolutionarily primitive Weber-based magnitude representation system may be premature.

Comparing robins' RQJs across different sets of quantities (high and low numbers weighted across the ratio scale) and in different task contexts (item-by-item and simultaneous presentation of visually accessible whole sets) can help uncover the commonality and diversity in the nature of the cognitive representations governing robins' quantitative judgments. We aimed to test whether robins converged with other species in having both object-file representations and analog magnitude representations, yet may diverge in terms of when limits in precise discrimination are imposed and when fuzzy Weber's law-based discrimination of quantities is deployed.

We carried out three experiments to test robins' spontaneous RQJs and investigate what representational systems could characterize their abilities. In Experiment 1, we tested robins' RQJ using an item-by-item paradigm where 10 different number pair comparisons (from 1 vs. 8 to 14 vs. 16) were combined with 5 different ratios (from 0.125 to 0.875), so that the total number of worms remained hidden from view as each robin made its choice. In Experiment 2, we presented the same number comparisons (and ratio range) to a new group of robins, by pouring the two sets of worms simultaneously, so two quantities remained visually accessible at the time of choice. In Experiment 3, we tested another new group of robins on RQJs with the same visually accessible context in Experiment 2, using number comparisons 4–8 times larger than those presented in the first two experiments (see Table 1). Across all three experiments, we analyzed the effects of ratio between pairs of quantities and total set size across pairs of quantities to contrast the analog magnitude and object indexing systems.

Methods

Subjects

Twelve individual robins participated in each experiment (36 robins in total). No single subject participated in more than one of the three experiments described below. Each subject was identified with the unique combination of colored bands on the bird's legs (see Fig. 1). None of the robins used in this study had prior experimental history. This study was conducted in native New Zealand forest within the Karori Sanctuary. The sanctuary is located on the southern tip of the North Island of New Zealand (41°18'S, 174°44'E) in Karori, Wellington. As of 2008, it housed a population of approximately 150 color-banded birds, and a total population estimated to be roughly 600 birds (McGavin 2009). Robins used in trials were located auditorily and visually along a series of footpaths or transects traversing the Karori Sanctuary.

Table 1 Number comparisons and total number of worms displayed for each trial for the 5 ratios investigated across Experiments 1, 2, and 3

Ratio	Experiments 1 and 2		Experiment 3	
	Comparison	Total	Comparison	Total
0.125	1 versus 8	9	4 versus 32	36
	2 versus 16	18	8 versus 64	72
0.250	1 versus 4	5	8 versus 32	40
	2 versus 8	10	16 versus 64	80
0.500	4 versus 8	12	16 versus 32	48
	8 versus 16	24	32 versus 64	96
0.750	3 versus 4	7	24 versus 32	56
	6 versus 8	14	48 versus 64	112
0.875	7 versus 8	15	28 versus 32	60
	14 versus 16	30	56 versus 64	120

**Fig. 1** A robin makes his choice in a sequential number task

Materials

We examined RQJs by presenting subjects with mealworms (*Tenebrio molitor* larvae) in artificial caches comprising two plastic wells, sunk into a wooden encasement. The plastic lining of each well served to prevent the mealworms from being able to climb the rough wooden surface of the encasement, thus retaining even large quantities of mealworms within the encasement. Each well measured approximately 2 cm deep, and they were spaced approximately 35 cm apart, allowing for variations in the terrain of the individual robin's territory. In Experiment 1, the wells were each covered with a circular brown leather flap approximately 7 cm across. For Experiments 2 and 3, precounted containers of mealworms were poured into these wells from white plastic, 35-mm film canisters. No covers were utilized for the wells in Experiments 2 and 3.

Procedure

Mealworm prey was presented to wild robins in an experimental arena containing two artificial cache sites. Wild robins frequently turn over leaves in search of prey on the forest floor. As a result, all birds readily removed leather lids to access prey without training after a brief exposure period to the experimental materials. For all three experiments, subjects were not positively or negatively reinforced for a specific choice and allowed to retrieve mealworms for consumption or caching regardless of the choice made, but only from one of the two caches. We randomized the order of treatments in all three experiments in an attempt to control for observational learning. In each experiment, 12 individual robins participated, for a total of 36 subjects across the three experiments, all of whom were color-banded with unique identifying band combinations. Trials were conducted within the forest territory of each bird. Ten different number comparisons spanning 5 ratios were displayed to each robin (see Table 1), with the total combined number of worms displayed ranging between 5 and 120 across all three experiments. The number combinations in these three experiments (see Table 1) were specifically chosen for an even and broad distribution of ratios between 0 and 1 and to create essentially 4 tiers of varying set sizes for each ratio, using 5 different larger set quantities (4, 8, 16, 32, 64). There was no trial in which a choice was made and worms were not retrieved by the robin; every robin made a choice for the given set of 10 number comparisons. In all three experiments, presentation of the larger number on the right or left was counterbalanced and randomized, and the first number presented for Experiment 1 was also randomized for side (left, right) and quantity (more, less).

In item-by-item presentation, subjects are unable to see either the accumulation in each well as prey was added, or the final amounts (singly or simultaneously) as the opaque containers obscured prey, and were covered upon filling. Each individual mealworm is only visible as it is held in tweezers while being dropped into the well. It is important to note that the visible presentation procedure in Experiments 2 and 3 was specifically adapted for experimentation involving large quantities with a wild species, and prey was, therefore, not placed sequentially onto the platforms as in Experiment 1. In aiming to experiment with quantities of mealworms beyond a total number of 30, it was important to reduce the time it took to place the worms onto the platform, because in many cases, robins will not attend to a trial for longer than about 60s. To this end, pouring the worms simultaneously in trials, for example, using a total of 80 or 120 worms, was essential to the participation of the subjects. To be able to appropriately compare *both* of these changes in methodology (visibility

and manner of placement), Experiment 2 was also conducted using this pouring methodology and with the prey remaining visible, using the same smaller numbers as in Experiment 1.

Experiment 1

Hidden quantities of mealworms were presented item by item, numbering in combined total between 5 and 30. Ten specific number combinations consisting of five different ratios (see Table 1) were presented to all 12 subjects comprising 120 trials. Mealworms were placed in each site sequentially, at an approximate rate of 2 s per item, using metal tweezers, to prevent worms from being obscured as they were placed. Wells were separately filled with between 1 and 16 mealworms in each and then covered. Due to the depth of the well, and the position of the subjects sitting low to the ground, robins did not have visual access to the accumulating contents of each well either during or after filling. The robin was then allowed to make a choice between the two artificial caches by removing the cover and retrieving the worms for consumption or storage.

Experiment 2

Each of the same ten number combinations and five ratios presented in Experiment 1 (see Table 1) were simultaneously presented in whole sets to each of 12 different subjects. Subjects were presented with two caches, which were then simultaneously filled with between 1 and 16 mealworms each and remained visible. White plastic film canisters were preloaded with a specific number of mealworms. Mealworms were poured into the wells using the canisters.

Experiment 3

In the final experiment, we presented very large numbers (up to 64 in a single set) to explore the upper range of robins' spontaneous numerosity discrimination ability. Ten new number combinations (see Table 1) that spanned the same five ratios used in Experiments 1 and 2 were presented simultaneously as whole sets to each of the 12 new birds, totaling 120 separate trials (10 per individual). Testing very large number discrimination in the wild with item-by-item presentation is not practical: Ecological necessities faced by robins—vigilant monitoring of extant food caches and defending of territory from intruders—make it extremely unlikely that a given subject will attend to a sequential presentation of up to 120 worms. The experimental caches were simultaneously filled with between 4 and 64 mealworms in each well and remained visible. White plastic film canisters were preloaded with a

specific number of mealworms for each number combination. Mealworms were poured into the wells using identical white film canisters. The robin was then allowed to make a choice between the two artificial caches, and the outcome documented.

Results

In each of the three experiments, we examined the extent to which accuracy of RQJs varied with ratio and total set size and additionally performed a joint analysis of Experiments 2 and 3 across all visibly presented trials. Subjects' choice in each trial is defined as "successful" where the greater of the two numbers was chosen. The ratio of the mealworms displayed was calculated dividing the smaller number (in one well) by the larger (in the second well) (e.g., $1/8 = 0.125$). Total set size was the combined number of mealworms presented in both wells (e.g., 2 vs. 5 = 7). Robins did not appear to display a side-bias across the three experiments in making a choice (left, 51 %; right, 49 %).

A binary logistic regression analysis for both ratio and set size was conducted for each of the 3 Experiments, as well as a joint analysis of Experiments 2 and 3, with each analysis including robin (individual subject) as a categorical variable to account for the within-subjects design. There was considerable variation between individuals, ranging from 40 % by an individual in Experiment 3—100 % correct by two individuals in Experiments 1 and 2. Robin, as a variable, was never a significant factor in the resulting model across trials presented in Experiment 1 ($P = 0.863$, Wald = 6.149, $df = 11$), Experiment 2 ($P = 0.985$, Wald = 3.383, $df = 11$), or Experiment 3 ($P = 0.348$, Wald = 12.218, $df = 11$).

In a binary logistic regression, the odds ratio (Exp (B)) refers to the decrease in the dependent variable (percentage of correct RQJs), which occurs with every unit of increase in the independent variable (e.g., total number of worms). So, for example, when (Exp (B)) = 0.720, the model indicates that with each unit increase—in this example, with every worm added to the overall number shown—the odds of a subject choosing correctly are reduced to a projected average of 0.72 times the odds of the previous comparison. This effectively reduces the odds of success in this example by 28 % (or 0.28 times) with every added worm—decreasing the overall likelihood of success as set size gets larger.

Preliminary analyses indicated that there was no evidence of reinforced learning across trials in any of the three experiments (Exp. 1, $P = 0.451$; Exp. 2, $P = 0.161$; Exp. 3, $P = 0.657$), as shown in the results of a binary logistic regression. Individual robins were, therefore, not significantly more or less likely to be successful when responding

on their tenth trial than their first; subjects' responses were spontaneous and non-reinforced as intended. Across all three experiments, few number comparisons were performed at chance (Experiment 2, 8 vs. 16 and 14 vs. 16; Experiment 3, 28 vs. 32 and 56 vs. 64) or below (Experiment 3, 48 vs. 64).

Experiment 1

Robins' RQJs for sequentially displayed comparisons involving trials with a total of 5–30 prey items are shown in Fig. 2. Performances across the number combinations ranged from 58 % success (14 vs. 16) to 100 % success (1 vs. 4).

A binary logistic regression was conducted, including individual as a categorical variable, and prey ratio and total set size as independent covariates, with "success" as the dependent variable. This revealed total set size to be a significant predictor of RQJ performance ($P = 0.043$, Exp (B) = 0.936, Wald = 4.096, $df = 1$), whereas ratio was not ($P = 0.533$, Exp (B) = 0.577, Wald = 0.388, $df = 1$). The model indicated that with each unit increase in total set size for a number comparison, the odds were reduced to a projected average of 0.94 times the odds of the previous comparison, slightly decreasing (by 6 % with each additional worm) the overall likelihood of success as total set size gets larger. Further analysis of Experiment 1 is conducted and explained as part of the following joint analysis section.

Experiment 2

Robins' RQJs for simultaneously visible comparisons involving trials with a total of 5–30 prey items are shown in Fig. 3. Percent success overall ranged between 50 % (8 vs.

16 and 14 vs. 16) and 92 % (1 vs. 8), showing a slight but non-significant ($t(22) = 0.81$, $P = 0.42$) but slight overall decrease in accuracy ($n = 120$, $M = 72$ %, $SD = 0.453$, $SE = 0.041$) by comparison across the hidden trials of the same numbers in Experiment 1 ($n = 120$, $M = 76$ %, $SD = 0.430$, $SE = 0.039$).

As in Experiment 1, the logistic regression analyzing both prey ratio and total set size (with individual as a categorical variable) found total set size was significant ($P = 0.034$, Exp (B) = 0.936, Wald = 4.484, $df = 1$). The model indicated that with each unit increase (worm added) in total set size, the odds were reduced to a projected average of 0.94 times that of the previous comparison, slightly decreasing (by 6 % with each additional worm) the overall likelihood of success as total set size increases. Ratio was not a significant predictor of success ($P = 0.394$, Exp (B) = 0.494, Wald = 0.727, $df = 1$). While the model projected that the odds would be reduced to 0.49 times that of the previous comparison, drastically decreasing (by 51 % with each additional worm) the overall likelihood of success as ratio increases, it remained a non-significant predictor of success.

Experiment 3

Robins' RQJs for simultaneous visible comparisons involving trials with a total of 36–120 prey items are shown in Fig. 4. Percent success overall ranged from 33 % (48 vs. 64) to 83 % (4 vs. 32 and 8 vs. 64), on the whole lower ($n = 120$, $M = 63$ %) than Experiment 1 ($n = 120$, $M = 76$ %) or 2 ($n = 120$, $M = 72$ %). The overall range of percentage of "correct" decisions shows a slight decrease in accuracy in contrast to both Experiment 1 (a significant difference, $t(22) = 2.08$, $P = 0.049$) and Experiment 2

Fig. 2 Percentage of correct trials as a function of **a** ratio and **b** total number of worms (combined set size), for item-by-item presentation of comparisons containing up to 30 total prey items ($N = 12$)

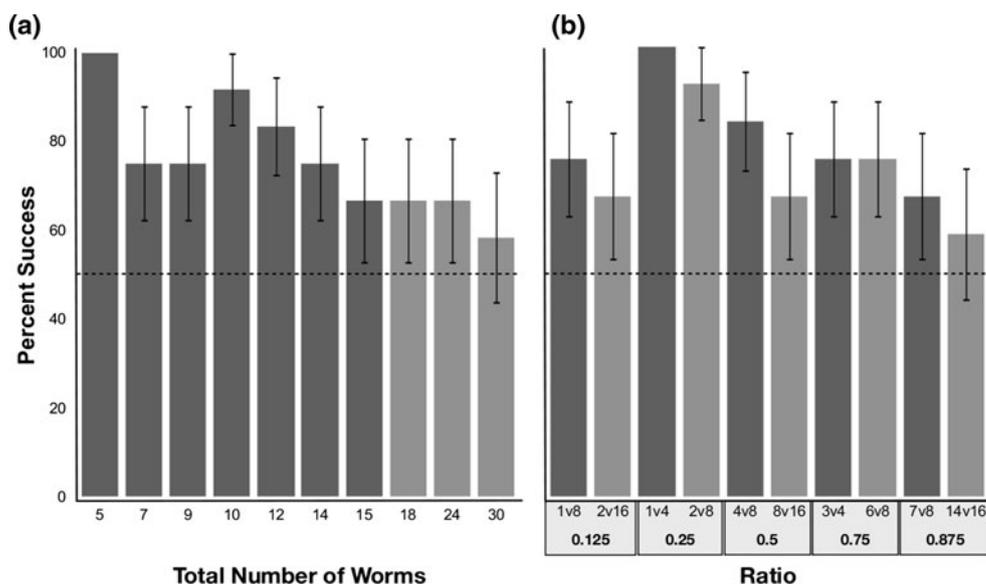
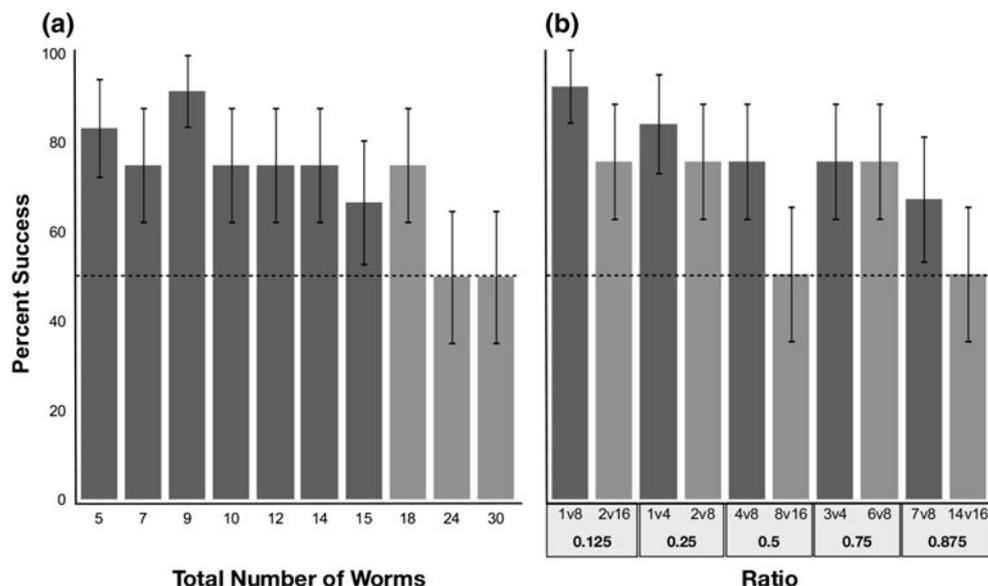


Fig. 3 Percentage of correct trials as a function of **a** ratio and **b** total number of worms (combined set size), for whole set presentation of comparisons containing up to 30 total prey items ($N = 12$)



(non-significant: $t(22) = 1.44$, $P = 0.17$), and response variance increased ($SD = 0.484$, $SE = 0.440$) as ratio between sets increased.

A logistic regression including prey ratio and total set size (with individual subject as a categorical variable and “success” as the dependent variable) revealed that ratio was a strong significant predictor of success ($P = 0.020$, $\text{Exp}(B) = 0.138$, $\text{Wald} = 5.412$, $df = 1$), while total set size was not ($P = 0.571$, $\text{Exp}(B) = 0.995$, $\text{Wald} = 0.320$, $df = 1$). According to this model, the odds of the likelihood of making a successful choice are reduced to a projected average of 0.14 times that of the previous comparison, sharply decreasing (by 86 % with each unit) the overall likelihood of success as prey ratio gets larger.

Joint analyses

In a joint analysis of Experiments 2 and 3, both of which used the same methodology of visibly presenting varying quantities of prey, ratio was a significant predictor of success ($P = 0.009$, $\text{Exp}(B) = 0.220$, $\text{Wald} = 6.817$, $df = 1$), while total set size was not ($P = 0.206$, $\text{Exp}(B) = 0.990$, $\text{Wald} = 1.603$, $df = 1$). This overall result is expected, when for both high and low numbers, the model showed an extremely strong potential reduction in success associated with ratio both in Experiments 2 (by 51 % per unit change) and 3 (by 86 % per unit change), although it was only a significant predictor in the latter. With comparison, total set size remains a weak effect in the models for both Experiments 2 (accuracy reduced by 6 % per unit change) and 3 (a less than 1 % drop in accuracy per unit change), although the set size effect was only significant in the former.

Most importantly, when percent success is looked at across trials in both experiments (see Table 2), a pattern

emerges indicating that in considering the larger of the two set sizes shown, success drops significantly ($t(18) = 3.780$, $P = 0.001$; see Fig. 5) from the percent success of smaller sets, with a larger set size of 4 or 8 ($n = 84$, $M = 77\%$, $SD = 0.421$), until the larger set size reaches 16 ($n = 36$, $M = 58\%$, $SD = 0.500$). Beyond this, a split occurs where number comparisons with smaller numerical distance between the two sets (ratios of 0.750 and 0.875) remain close to chance, and those with greater numerical distance (ratios of 0.125, 0.250, and 0.500) increase again in the percentage of successful responses. This serves as an important marker for defining and analyzing “small” and “large” quantities in reference to robins’ performance with comparison to other species and could be indicative of two systems at work.

In order to more accurately examine how ratio and total set size affect “small” and “large” numbers, we have analyzed Experiments 1 and 2 again, applying this new definition of “large” and “small” based on the findings of the joint analysis. To generate a data set of “small” set sizes, we removed all trials with a larger set size of 16 or more, effectively reducing the definition of “small” set sizes to those where the larger set is either 4 or 8 in these experiments, again using a binary logistic regression. To generate a data set of “large” set sizes, we used trials from Experiment 2 and Experiment 3 (both of which were presented in a simultaneous, visible manner) that had a larger set size of 16.

In Experiment 1, when trials with a larger set size of 16 or higher are removed from the model, neither ratio ($P = 0.344$, $\text{Exp}(B) = 0.247$, $\text{Wald} = 0.895$, $df = 1$) nor total set size ($P = 0.393$, $\text{Exp}(B) = 0.903$, $\text{Wald} = 0.730$, $df = 1$) is significant predictors of success. Likewise, in Experiment 2, when trials with a larger set size of 16 or

Fig. 4 Percentage of correct trials as a function of **a** ratio and **b** total number of worms (combined set size), for whole set presentation of comparisons containing up to 120 total prey items ($N = 12$)

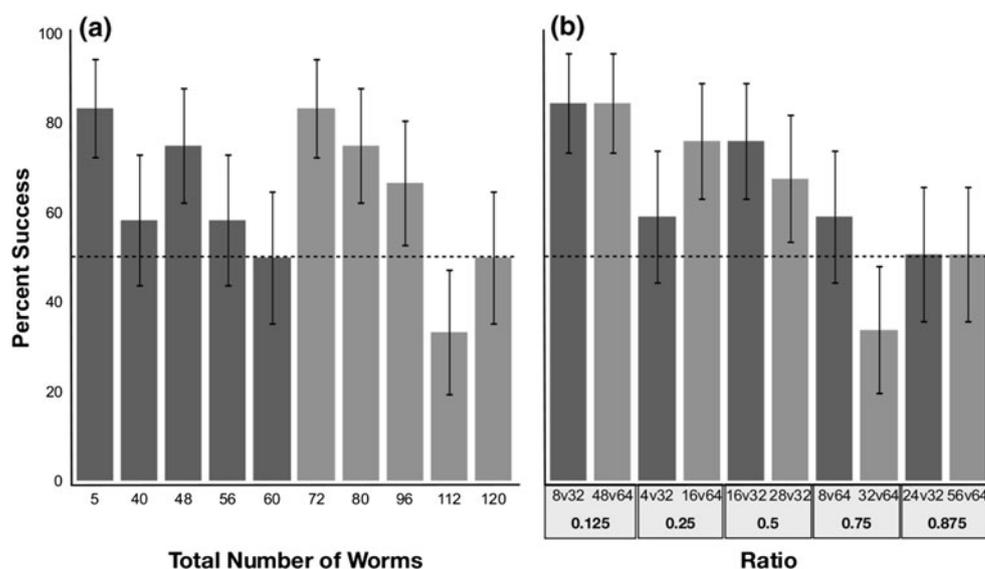


Table 2 Percentage of correct trials across Experiments 1, 2, and 3 in both “small” and “large” comparisons

Ratio	Experiment 1		Experiment 2		Experiment 3	
	Small	Large	Small	Large	Small	Large
0.125	1 versus 8	2 versus 16	1 versus 8	2 versus 16	4 versus 32	8 versus 64
	75 %	67 %	92 %	75 %	83 %	83 %
0.250	1 versus 4	2 versus 8	1 versus 4	2 versus 8	8 versus 32	16 versus 64
	100 %	92 %	83 %	75 %	58 %	75 %
0.500	4 versus 8	8 versus 16	4 versus 8	8 versus 16	16 versus 32	32 versus 64
	83 %	67 %	75 %	50 %	75 %	67 %
0.750	3 versus 4	6 versus 8	3 versus 4	6 versus 8	24 versus 32	48 versus 64
	75 %	75 %	75 %	75 %	58 %	33 %
0.875	7 versus 8	14 versus 16	7 versus 8	14 versus 16	28 versus 32	56 versus 64
	67 %	58 %	67 %	50 %	50 %	50 %

higher are removed from the model, neither ratio ($P = 0.344$, $\text{Exp}(B) = 0.247$, $\text{Wald} = 0.895$, $df = 1$) nor total set size ($P = 0.393$, $\text{Exp}(B) = 0.903$, $\text{Wald} = 0.730$, $df = 1$) is significant predictors of success. In Experiment 3, when trials from Experiment 2 with a larger set size of 16 are included in the model, again, ratio emerges as a strong significant predictor of success, and total set size is not significant. These results indicate that set sizes below 16 square with a more object-based paradigm, but do not fall in line with the strict 4-item limit imposed by a typical “object-file system,” while numbers 16 and above are subject to a ratio-based system.

Discussion

Robins were capable of selecting the larger of two quantities when they were presented item by item and when both were presented simultaneously as visually accessible

whole sets. These findings extend previous work by Hunt et al. (2008). In other ways, however, the findings are novel with respect to broader comparative RQJ research.

Consider the following themes in the literature. First, some studies have found that animals perform poorly when they have to process the sequential presentation of items so as to keep two sets of representations in working memory and compare quantities mentally (e.g., Abramson et al. in press, with sea lions; Hanus and Call 2007, with apes). Second, psychophysical representations—based on Weber’s law—are a clear signature of numerosity when animals are tested with quantities presented as a united set (e.g., Abramson et al., Barner et al. 2008). Third, while researchers have argued that the object-file tracking system is a better fit for capturing how animals represent small numbers irrespective of their ratio, the numerical representation involved here has a typical set size limit of 3–4 items (e.g., Feigenson et al. 2004; Trick and Pylyshyn 1994; Hauser et al. 2000). Robins’ RQJs were different on

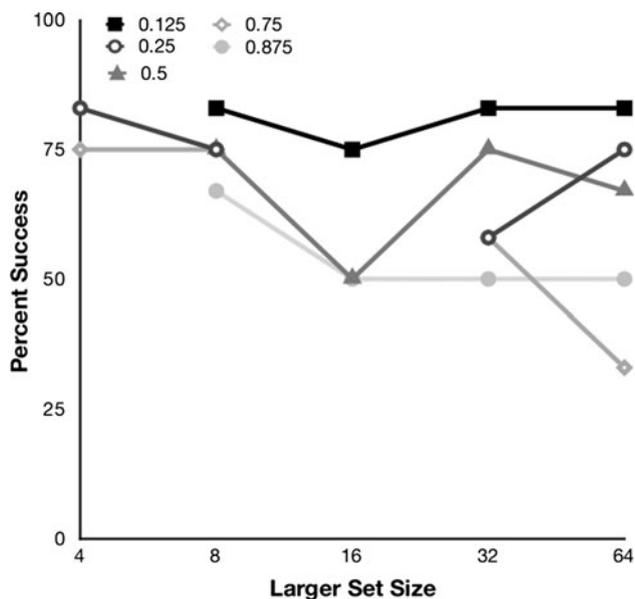


Fig. 5 Percentage of correct trials as a function of the larger number of worms of the two quantities shown in trials of Experiments 2 and 3

all three counts. Robins actually show slightly higher accuracy in selecting the larger number over a range of values—small (1–4) and large (8–16)—when quantities were presented item by item (Experiment 1) than when quantities were presented simultaneously (Experiment 2). Further, while there was a general decline in accuracy in numerical discrimination with increasing objective number in each method of RQJ assessment, accuracy in both contexts was predicted by the total set size of the quantities and not ratio. And finally, in Experiments 1 and 2, robins were able to make accurate RQJs for values far exceeding the typical object-file system limit of around 3–4 items; robins were accurate even for comparisons of 7 versus 8 (Experiments 1 and 2) and 14 versus 16 (Experiment 1). At the very least, robins’ object indexing system appears to operate over an extended number range of 1–16, and robins are able to individuate up to 8 or more items in parallel. While a detailed unpacking of robins’ sophisticated numerical competence will require further research, there may be merit to Hunt et al.’s (2008) explanations vis-à-vis robins’ food-caching ecology.

Initial RQJ testing by Hunt et al. (2008) indicated that robins chose the cache site with greater number with comparisons involving 1 versus 2, 2 versus 3, 3 versus 4, 4 versus 5, and 4 versus 8, but chose randomly with 4 versus 6; and 6 versus 8 and 8 versus 10. While the levels of accuracy in these results differ from the present study, this is likely explained by methodology adapted in the experiments presented here specifically for improved examination of extremely large numbers. For example, one important alteration is a faster drop rate of 2 s per worm in

the present study compared to 5 s in Hunt et al. (2008). This means that where a trial with 12 worms, for example, would be 60 s in duration in the Hunt et al. (2008) study, it would require only 24 s to display in the present study (2.5 times faster). The advantage therein lies in being able to overcome difficulties with loss of interest or attention by the subject, particularly when presentation of prey requires longer than 40–60 s, better maintaining full attention of the subject on the task at hand.

Cohesion has proven to be an essential feature in discrimination of quantities (van Marle et al. 2006; Cheries et al. 2008; van Marle and Wynn 2011) and may have played some role in influencing differing results between the hidden sequential presentation of worms in Experiment 1 and the visible poured presentation in Experiment 2 of the present study. Infants require a larger ratio difference to discriminate quantities of a non-solid substance, such as Cheerios (at least 1:4), compared to quantities of discrete objects, such as crackers (1:2). Strikingly, when Cheerios were then presented as discrete, individually placed objects rather than poured (so they no longer appeared as a non-countable “substance”) infants were successful at the 1:2 ratio they previously had failed (van Marle and Wynn 2011). For robins, this may well account for a slightly lower accuracy in the condition in which worms were poured in the present experiment, where robins fall to chance in the larger 0.50 ratio (8 vs. 16) in Experiment 2 (50 %), but not in the same comparison when it is sequentially placed in Experiment 1 (67 %) (see Table 2).

As a scatter-hoarding species, robins need to track each prey item to its endpoint—from being caught, stored, shifted, pilfered, and consumed. These physical inventories of stored food have high turnover rate due to perishability, are susceptible to pilfering by mates, fledglings, and neighbors, and are maintained on permanent year-round territories with seasonal pressures (summer provisioning, winter resource competition). Robins’ unusually high threshold for accuracy in quantity discrimination tasks complement evidence of neurological differences found not only between avian memory systems and mammals, but also between scatter-hoarding birds and other birds (e.g., Clayton 1995; Clayton and Krebs 1994). Of course, similar RQJ experiments on wild birds that do not hoard food have yet to be conducted; attributing the advanced numerical skills of robins to food hoarding *per se* remains speculative. That said, birds’ visual perception and spatial orientation are physiological systems that differ significantly from mammals and other animals (Cook 2000), and are likely to play a significant role in shaping how even such fundamental core knowledge systems as number, geometry, and object are evidenced in their natural behavior. Geospatial information attached to items tracked by scatter-hoarding species could very well augment behavioral evidence of

higher range object-based numerical system in such species the same way radio frequency identification (RFID) tags provide information enhancing a system's ability to track tagged items. For North Island robins, as scatter-hoarders, one ecological adaptation appears to be the use of a sophisticated number system that is able to precisely identify quantities larger than 3 or 4.

Importantly, our current findings also support more than one mechanism underlying the ability to enumerate quantities in the North Island robin. In Experiment 3, *ratio* was the only factor that significantly accounted for robins' RQJs involving very large values (16–64) presented simultaneously. Robins were able to discriminate between 32 and 64 prey items (0.50 ratio), but performance dropped to around chance level when 24 and 32 prey items were presented (0.75 ratio). With respect to large numbers that exceed the standard object-file range, mosquitofish have also been found to discriminate contrasts with ratios of 0.67 (8 vs. 12) but not ratios of 0.75 (9 vs. 12) (Agrillo et al. 2010). The developing precision of the analog magnitude representation system shown in human children also intersects with the ratio resolution in robins: Lipton and Spelke (2003) show that 6 month olds can discriminate a ratio of 0.50 (e.g., 8 vs. 16) but not 0.67 (e.g., 8 vs. 12), whereas 9 month olds can discriminate a ratio of 0.66 (e.g., 8 vs. 12) but not 0.80 (e.g., 8 vs. 10). As with Agrillo et al.'s mosquitofish, resolution of the ratio-based large number discrimination in robins parallels the resolution of the ratio-based large number discrimination shown between that of a 9-month-old human infant and a human preschooler. Given the correspondence of cognitive signatures between species as diverse as robins, fish, and human children's large number discrimination, our findings fit with the intriguing proposal that robins share with other species an analog magnitude system that is inherited from a common ancestor (Agrillo et al. 2008, 2010; Beran 2008; Cantlon and Brannon 2006).

Using New Zealand robins as a model system invites us to understand the representations underpinning numerical cognition in a new light. The analysis of the three experiments revealed that robins use two systems for coping with numbers. An object indexing system and a ratio-based analog magnitude system each produce an apparent and distinct "signal" across the number scale. The odds ratio for the object-file system hovered just below 1 across all three experiments (Exp (*B*)s ranged from 0.986 to 0.926) creating a signature of a gradual decline as objective number (total set size) increased, a signal that was insignificant for sets below 16. The signature which emerged for a ratio-based system was expressed by a very low odds ratio across all experiments (Exp (*B*)s ranged from 0.287 to 0.110), showing a drastic decline in probability of success as ratio increased.

The strength of the object indexing system is in providing a more accurate mechanism based upon enumerating distinct objects; it attempts to process objects as discrete items regardless of magnitude. The recent study with infants by van Marle and Wynn (2011) speculates on exactly this interplay between mechanisms in infants, reminding us that the analog magnitude system was initially based on data accounting for rats and pigeons perception of continuous stimuli such as time (duration) and number (Meck and Church 1983) and that analog magnitude representations themselves are also continuous, not discrete. They also suggest the possibility that estimation of substances is more variable than enumerating objects as a non-solid stimulus can change shape, dimension, or break apart upon movement—and that two mechanisms are likely to underlie the infant's quantification abilities. It is compelling that for both robins and human infants, the same non-cohesive material in each case (mealworms and Cheerios, respectively) can be perceived and possibly represented as either a continuous substance or as discrete items depending on context and presentation; and that in the case of robins especially, working memory and ecology appear to extend the item indexing limit typically associated with mammals tracking discrete items.

It is important to point out that robins' RQJs across all three experiments could be influenced by cues other than numerosness and are almost certainly likely to be affected by pilfering and hunting strategy specific to this species in ways that future experimentation will hopefully reveal in more detail. We did not control for cues such as differences in dropping duration (Experiment 1) or surface area (Experiments 2 and 3). Our findings do not definitively rule out that robins are making decisions based on *non-numerical* cues to solve numerical problems. Nonetheless, it is encouraging that the findings in Experiment 1 replicate work by Hunt et al. (2008) where dropping duration *was* controlled for. Hunt et al. also showed that other sensory confounds (odor and sound) do not affect robins' RQJs. Our attempt was primarily to simulate problems robins might encounter in the wild wherein the item-by-item presentation might approximate subjects observing a conspecific dropping insect prey into a cache one-by-one and the whole set presentation might approximate subjects turning over leaf litter to discover prey items. As such, even though our design cedes certain experimental controls for gains in ecological validity, our findings still afford important comparisons with other studies that have used similar protocols to measure spontaneous RQJs (e.g., Abramson et al. in press, with sea lions, and Hanus and Call 2007, with apes). Compared to those species where fewer subjects succeeded in the cognitively demanding item-by-item task context, robins were accurate in RQJs with sequential presentation as they were with presentation

of whole visible sets. As such, we have uncovered new findings about the representations that are at the root of robins' numerical abilities. Two systems appear to be at work when wild North Island robins perceive quantity: an approximate magnitude representation system that is effective for very large numbers and an object indexing system that remains effective at substantially higher thresholds than current theory posits.

Acknowledgments We thank the Karori Sanctuary staff and volunteers for making this research possible.

References

- Abramson JZ, Hernández-Lloreda V, Call J, Colmenares F (2011) Relative quantity judgments in South American sea lions (*Otaria flavescens*). *Anim Cogn* 14(5):695–706
- Agrillo C, Dadda M, Serena G, Bisazza A (2008) Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Anim Cogn* 11:495–503
- Agrillo C, Piffer L, Bisazza A (2010) Large number discrimination by fish. *PLoS One* 5:e15232
- Barner D, Wood J, Hauser M, Carey S (2008) Evidence for a nonlinguistic distinction between singular and plural sets in rhesus monkeys. *Cognition* 107:603–622
- Barth H, Kanwisher N, Spelke ES (2003) The construction of large number representations in adults. *Cognition* 86:201–221
- Beran MJ (2001) Summation and numerosness judgments of sequentially presented sets of items by chimpanzees (*Pan troglodytes*). *J Comp Psychol* 115:181–191
- Beran MJ (2008) The evolutionary and developmental foundations of mathematics. *PLoS Biol* 6:221–223
- Beran MJ, Beran MM (2004) Chimpanzees remember the results of one-by-one addition of food items to sets over extended time periods. *Psychol Sci* 15:94–99
- Burns KC, Steer J (2006) Dominance rank influences food hoarding in New Zealand Robins *Petroica australis*. *Ibis* 148:266–272
- Cantlon JF, Brannon EM (2006) Shared system for ordering small and large numbers in monkeys and humans. *Psychol Sci* 17:401–406
- Cantlon JF, Platt ML, Brannon EM (2009) Beyond the number domain. *Trends Cogn Sci* 13:83–91
- Carazo P, Font E, Forteza-Behrendt E, Desfilis E (2009) Quantity discrimination in *Tenebrio molitor*: evidence of numerosity discrimination in an invertebrate. *Anim Cogn* 12:463–470
- Carey S (2001) Cognitive foundations of arithmetic: evolution and ontogenesis. *Mind Lang* 16:37–55
- Cherries EW, Mitroff SR, Wynn K, Scholl BJ (2008) Cohesion as a constraint on object persistence in infancy. *Dev Sci* 11(3):427–432
- Clayton NS (1995) The neuroethological development of food-storing memory: a case of use it, or lose it. *Behav Brain Res* 70:95–102
- Clayton NS, Krebs JR (1994) Memory for spatial and object-specific cues in food-storing and non-storing birds. *J Comp Physiol* 174:371–379
- Clayton NS, Krebs JR (1995) Memory in food-storing birds: from behaviour to brain. *Curr Opin Neurobiol* 5:149–154
- Clayton NS, Dally JM, Gilbert J, Dickinson A (2005) Food caching by western scrub-jays (*Aphelocoma californica*) is sensitive to the conditions at recovery. *J Exp Psychol Anim Behav Process* 31:115–124
- Cook RG (2000) The comparative psychology of avian visual cognition. *Curr Dir Psychol Sci* 9:83–89
- Dadda M, Piffer L, Agrillo C, Bisazza A (2009) Spontaneous number representation in mosquitofish. *Cognition* 112:343–348
- Davis H, Memmott J (1982) Counting behaviour in animals: a critical evaluation. *Psychol Bull* 92:547–571
- Dehaene S, Izard V, Spelke ES, Pica P (2008) Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science* 320:1217
- Farnsworth GL, Smolinski JL (2006) Numerical discrimination by wild Northern mockingbirds. *Condor* 108:953
- Feigenson L, Carey S, Hauser MD (2002) The representations underlying infants' choice of more: object files versus analog magnitudes. *Psychol Sci* 13:150–156
- Feigenson L, Dehaene S, Spelke ES (2004) Core systems of number. *Trends Cogn Sci* 8:307–314
- Gallistel CR, Gelman R (1992) Preverbal and verbal counting and computation. *Cognition* 44:43–74
- Gallistel CR, Gelman R (2005) Mathematical cognition. In: Holyoak KJ, Morrison RG (eds) *The Cambridge handbook of thinking and reasoning*. Cambridge University Press, New York, pp 559–588
- Hanus D, Call J (2007) Discrete quantity judgments in the great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, *Pongo pygmaeus*): the effect of presenting whole sets versus item-by-item. *J Comp Psychol* 121:241–249
- Hauser MD, Carey S, Hauser LB (2000) Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proc R Soc Lond B Biol Sci* 267:829–833
- Honig WK, Stewart KE (1989) Discrimination of relative numerosity by pigeons. *Anim Learn Behav* 17:134–146
- Hunt S, Low J, Burns KC (2008) Adaptive numerical competency in a food-hoarding songbird. *Proc R Soc Lond B Biol Sci* 275:2373–2379
- Irie-Sugimoto N, Kobayashi T, Sato T, Hasegawa T (2009) Relative quantity judgment by Asian elephants (*Elephas maximus*). *Anim Cogn* 12:193–199
- Jordan KE, Brannon EM (2006) A common representational system governed by Weber's law: Nonverbal numerical similarity judgments in 6-year-olds and rhesus macaques. *J Exp Child Psychol* 95:215–229
- Koehler O (1941) Vom Erlernen unbenannter Anzahlen bei Vögeln/ The learning of unnamed amounts in birds. *Naturwissenschaften* 29:201–218
- Lipton JS, Spelke ES (2003) Origins of number sense: large-number discrimination in human infants. *Psychol Sci* 14:396–401
- Low J, Burns KC, Hauber ME (2009) Wild number sense in brood parasitic Brown-headed Cowbirds. *Ibis* 151:775–777
- Lyon BE (2003) Egg recognition and counting reduce costs of avian conspecific brood parasitism. *Nature* 422:495–499
- McComb K, Packer C, Pusey A (1994) Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Anim Behav* 47:379–387
- McGavin S (2009) Density and pair fidelity in a translocated population of North Island robin (*Petroica longipes*). *Notornis* 56:206–212
- Meck WH, Church RM (1983) A mode control model of counting and timing processes. *J Exp Psychol Anim Behav Process* 9:320–334
- Menzies IJ, Burns KC (2008) Food hoarding in the New Zealand robin: a review and synthesis. In: Weber EA, Krause LH (eds) *Animal behavior: new research*. Nova Science Publishers, New York, pp 163–183
- Pepperberg IM (2006) Grey parrot (*Psittacus erithacus*) numerical abilities: addition and further experiments on a zero-like concept. *J Comp Psychol* 120:1–11
- Rugani R, Regolin L, Vallortigara G (2008) Discrimination of small numerosities in young chicks. *J Exp Psychol Anim Behav Process* 34:388–399

- Shettleworth SJ (2010) Cognition, evolution, and behavior, 2nd edn. Oxford University Press, New York
- Terrace HS (1987) Chunking by a pigeon in a serial learning task. *Nature* 325:149–151
- Trick LM, Pylyshyn ZW (1994) Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol Rev* 10:80–102
- Uller C (2008) Developmental and evolutionary considerations on numerical cognition: a review. *J Evol Psychol* 6:237–253
- Uller C, Jaeger R, Guidry G, Martin C (2003) Salamanders (*Plethodon cinereus*) go for more: rudiments of number in an amphibian. *Anim Cogn* 6:105–112
- van Marle K, Wynn K (2011) Tracking and quantifying objects and non-cohesive substances. *Dev Sci* 14(3):502–515
- Vander Wall SB (1990) Food hoarding in animals. University of Chicago Press, Chicago
- vanMarle KL, Aw J, McCrink K, Santos LR (2006) How Capuchin monkeys (*Cebus apella*) quantify objects and substances. *J Comp Psychol* 120(4):416–426
- White DJ, 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