Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Three Chick (Gallus gallus) Breeds Differing in Domestication and Selection Status Show Proto-Arithmetic Abilities in Distinguishing Empty from Non-Empty Sets

Kimberly Brosche

kimberlyyvonne.brosche@studenti.unipd.it

University of Padua

Luna Dudine

University of Padua

Filippo Gambarota

University of Padua

Gianmarco Altoè

University of Padua

Rosa Rugani

University of Padua

Research Article

Keywords: Proto-arithmetic, Gallus gallus, domestication, zero, empty set

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11610664/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Three Chick (*Gallus gallus*) Breeds Differing in Domestication and Selection Status Show Proto-Arithmetic Abilities in Distinguishing Empty from Non-Empty Sets

Authors:

Kimberly Brosche*[1], Luna Dudine*[1,2], Filippo Gambarota[3], Gianmarco Altoè[3], Rosa Rugani[1]

*These authors contributed equally

[1] Comparative Cognition Lab, Department of General Psychology, University of Padua, Italy

[2] Department of Biology, University of Padua, Italy

[3] Department of Developmental Psychology and Socialisation, University of Padua, Italy

Corresponding authors:

Kimberly Brosche: kimberlyyvonne.brosche@studenti.unipd.it

Rosa Rugani: rosa.rugani@unipd.it

Orcids:

Kimberly Brosche: https://orcid.org/0000-0003-2809-8591

Luna Dudine: https://orcid.org/0009-0004-2585-1570

Filippo Gambarota: https://orcid.org/0000-0002-6666-1747

Gianmarco Altoè: https://orcid.org/0000-0003-1154-9528

Rosa Rugani: https://orcid.org/0000-0001-5294-6306

Abstract

Tracking the quantity of moving objects is advantageous in natural environments. Animals can monitor prey, predators, conspecifics, offspring, and other relevant items. For domesticated animals in captivity, the evolutionary pressure on this ability has likely changed. Here, we compare three breeds of day-old chicks differing in domestication status and selection for productivity (Ross308, Gallina Padovana, Red Junglefowl) in a task that required successive addition and subtraction of objects.

Chicks (n=48) were reared with five identical objects, which they treated as artificial social companions, naturally preferring the larger group. At test, chicks watched objects disappearing one-by-one behind two opaque screens. Subsequently, some objects were moved between screens, requiring chicks to add and subtract to locate the larger group. We assessed whether chicks can handle empty sets as the result of addition and subtraction. We found no evidence of breed differences, although the study was not powered to detect small differences. Ross308 and Gallina Padovana, but not Red Junglefowl (p=0.054), chicks chose the larger group in the 0*vs*.5 comparison (resulting from either 2-*2 vs*. 3+*2* or 3-*3 vs*. 2+*3*); but not in 2*vs*.3 (resulting from either 4-*1 vs*. 1+*1* or 4-*2 vs*. 1+*2*). In a follow-up experiment, chicks of the most productive breed (n=15) performed significantly above chance level also with the final comparison 0*vs*.3, showing that successful empty-set discrimination is not limited to large numerical contrasts. These results suggest that proto-arithmetic abilities involving empty sets are present in breeds that differ markedly in domestication and selection for productivity.

Keywords

Proto-arithmetic, *Gallus gallus*, domestication, zero, empty set

Introduction

The ability to combine and manipulate quantities, referred to as proto-arithmetic, is a fundamental component of numerical cognition, with deep developmental and evolutionary roots (Regolin et al. 2025; Dudine et al. under review). Proto-arithmetic ability has been shown in a variety of animal species, such as primates (vervet monkeys, *Chlorocebus pygerythrus* (Tsutsumi et al. 2011); chimpanzees, *Pan troglodytes* (Rumbaugh et al. 1987)), birds (domestic chicks, *Gallus gallus* (Rugani et al. 2009, 2017); pigeons, *Columba livia* (Brannon et al. 2001)), fish (zebrafish, *Danio rerio* (Potrich et al. 2024); Cichlids and stingrays (Schluessel et al. 2022)), insects (honeybees, *Apis mellifera* (Howard et al. 2019); ants, *Myrmica sabuleti* (Cammaerts and Cammaerts 2021)) as well as in pre-verbal human infants (e.g., Wynn 1992; McCrink and Wynn 2004; Gilmore 2015). Seminal evidence comes from two chimpanzees, which were trained to select the larger set of chocolates arranged on two trays (Rumbaugh et al. 1987). Each tray contained two wells that could hold 0–4 chocolates, thus the chimpanzees had to sum the chocolates contained in the two wells on each tray to estimate which tray contained more chocolates. Chimpanzees succeeded at a very high rate, also when one of the wells on the tray with the smaller total amount contained the largest individual set (e.g., 0 and 4 = 4 on one tray *vs*. 2 and 3 = 5 on the other tray). In a follow-up study, chimpanzees were also successful when food items were dropped into opaque containers, requiring them to perform addition (Beran 2004).

The ecological relevance of proto-arithmetic is particularly evident in foraging. Imagine an insectivorous animal observing insects as they disappear into the hollow of a tree. Then, the animal watches as some insects emerge again. Is it still worth to climb up the tree and take the remaining insects from the hollow? To solve this problem, the animal has to subtract the insects it saw leave from those that initially entered the hollow. In such scenarios, proto-arithmetic ability clearly increases foraging efficiency (Tsutsumi et al. 2011). Consistently, wild New Zealand robins looked longer at unexpected than expected outcomes of additions involving worms hidden in opaque containers (Garland and Low 2014). Nevertheless, with food items, that intrinsically function as rewards, it is difficult to control for continuous physical variables that co-vary with numerosity, such as overall volume, leaving open the possibility that animals mainly rely on quantitative cues (Davis and Pérusse 1988; Agrillo and Bisazza 2014; but see Stancher et al. 2015).

Social contexts provide another ecologically relevant domain in which proto-arithmetic abilities can emerge and guide adaptive behavior. Wild lionesses (*Panthera leo*) decided whether to approach a loudspeaker broadcasting the roars of one or three intruders based on the relative size and strength of their own group and the simulated opponent group (McComb et al. 1994). Dogs seem to apply a similar strategy in inter-group conflicts, with the decision to engage in fights with rival packs and the outcome of these conflicts varying with relative group size (Bonanni et al. 2011). Dogs (*Canis lupus familiaris*) responded most adaptively when both packs were small, but their performance declined when both groups included more than four members (for a review of similar studies see Benson-Amram et al. 2018). For birds, keeping track of changing quantities might also be relevant for reproductive success, to know their number of eggs and thus counter-act brood parasitism (Lyon 2003).

While proto-arithmetic abilities can provide adaptive advantages in the wild (Wynn and Bloom 1992; Dehaene 2011; Nieder 2020), it remains unclear how they have been affected by domestication and selection for productivity. Two competing hypotheses can be formulated. On the one hand, proto-arithmetic ability may become less relevant in captivity as animals are provided with food and shelter. Knowing that domesticated animals often reduce traits or abilities

they no longer need in favor of energy-costly traits, like production traits (Beilharz et al. 1993; Beilharz and Nitter 1998; Price 2002), we can hence hypothesize that domestication has *decreased* domesticated animals' proto-arithmetic abilities relative to their wild ancestors'. The effect of domestication might be exacerbated by intensive selection for productivity, further shifting animals' resource allocation (Beilharz et al. 1993; Beilharz and Nitter 1998; Price 2002). Thus, according to this hypothesis, we expect domesticated and highly selected animals to perform worst and wild animals to perform best in proto-arithmetic tasks.

On the other hand, domestication and selection for productivity could have led animals to optimize their information gain and memory. A study by Dudde et al. (2018) found that more productive hens outperformed less productive ones in a learning and memory task. Their results might extend to proto-arithmetic abilities, especially since proto-arithmetic is closely interconnected with memory, specifically working memory (DeStefano and LeFevre 2010; Cragg and Gilmore 2014). According to this second hypothesis, wild animals are expected to show lower performance in proto-arithmetic tasks than domesticated and highly selected animals.

In the present study, we investigate proto-arithmetic ability in young chicks (*Gallus gallus*), focusing on the effects of domestication and selection for productivity and on the processing of empty and non-empty outcomes. Specifically, we compared the proto-arithmetic performance of three breeds of chicks differing in domestication history and selection for productivity: i) the Red Junglefowl, the wild ancestor of domestic chickens (Al-Nasser et al. 2007), ii) the Gallinas Padovana, a domesticated breed selected for ornamental traits rather than production (Pozzato 2019), and iii) the Ross308, a highly productive commercial hybrid. We measured their accuracy in a proto-arithmetic task, a paradigm particularly suitable for this purpose because it requires chicks to represent numerosity while maintaining and updating information about hidden objects in working memory, thereby combining numerical processing– with working-memory demands.

As chicks are highly precocial and can undergo cognitive testing in the first days of life (Rugani et al. 2007, 2009, 2013a, b; Chiandetti and Vallortigara 2018; Rugani and Regolin 2021; Rosa-Salva et al. 2026), numerical abilities can be compared across breeds while minimizing differences in individual experience (Rugani 2018; Brosche et al. 2025; Rugani et al. 2025). Indeed, chicks have mastered previous proto-arithmetic tasks (Rugani et al. 2009, 2011, 2013a, 2017, 2020, 2022; Loconsole et al. 2023), providing a fruitful basis for the study of domestication effects. The seminal study on proto-arithmetic in day-old chicks (Rugani et al. 2009) exploited the chicks' tendency to approach the larger group of artificial objects they were reared with soon after hatching, toward which they developed social attachment (Rugani et al. 2010). To investigate proto-arithmetic abilities, the objects were hidden sequentially behind two opaque screens. Then, chicks were allowed to choose between the two screens. Hybro broiler chicks successfully selected the larger set in a 2 vs. 3 comparison, even when the two sets were equalized with respect to continuous variables and when the objects were moved after initial placement, thus requiring chicks to add and subtract (Rugani et al. 2009). This procedure has since been used to identify the upper limits of proto-arithmetic abilities (Rugani et al. 2011), to assess discrimination between small and large numerosities, revealing a limit at 3 versus 4 items (Rugani et al. 2013a). In Ross308 chicks, a faster-growing breed more strongly selected for meat production than Hybro, this limit can be overcome when the objects are grouped (Rugani et al. 2020) or when they bear individually distinctive face-like features, which may facilitate their individuation (Rugani et al. 2022). To date, the only study that has directly compared cognitive performance across chicken breeds differing in selection for productivity used an ordinal numerical task, which

required chicks to relocate the 4th item in a series of 10 identical items. Hybro chicks outperformed Ross308 chicks (Rugani et al. 2023).

Despite growing evidence for proto-arithmetic abilities in chicks, previous studies have focused on broiler breeds, and no study has directly compared breeds differing in domestication history and selection for productivity. In particular, proto-arithmetic has never been investigated in the wild Red Junglefowl, the ancestor of domestic chickens, or in Gallina Padovana, a domesticated breed not intensively selected for productivity. Moreover, previous studies tested different numerical comparisons in separate groups of chicks, preventing a direct assessment of how the same individual performs across different proto-arithmetic operations. By testing each chick across multiple numerical comparisons, the present study provides a more comprehensive assessment of individual proto-arithmetic abilities across breeds. To this end, we selected proto-arithmetic operations in which chicks initially observed objects being added one by one to two empty sets, requiring them to build a representation of each numerosity. These representations then had to be updated as objects were subsequently moved from one set to the other. Critically, the final outcome could consist of either two non-empty sets or an empty and a non-empty set. For example, in the 2 vs. 3 comparison, chicks initially saw 4 vs. 1 objects and subsequently experienced either 4−*1* vs. 1+*1* or 4−*2* vs. 1+*2*. In the 0 vs. 5 comparison, they initially saw 3 vs. 2 objects and experienced either 2−*2* vs. 3+*2* or 3−*3* vs. 2+*3*. Thus, both comparisons required chicks to update two initial quantities through subtraction and addition, while differing critically in whether one operation resulted in an empty set.

The capacity to represent empty sets is fundamental to proto-arithmetic, as an empty set provides the starting quantity to which subsequent items can be added to obtain the correct total. Evidence from adult birds suggests that representing the empty set is within the numerical repertoire of some species. Adult crows (*Corvus corone*) can discriminate empty sets from other small numerosities in delayed match-to-sample tasks. Their performance exhibits a typical distance effect, (i.e., accuracy increases as the difference between numerosities becomes larger Dehaene et al. 1998) with a clear discrimination between 0 *vs*. 4 but reduced discrimination between 0 *vs*. 1 (Kirschhock et al. 2021). Evidence in very young animals is more limited. In newly hatched chicks, Regolin et al. (2005) showed that social objects can be maintained in working memory even when temporarily occluded, indicating that chicks can represent presence, a prerequisite for representing absence (Nieder 2016). Extending this evidence, Szabó et al. (2022) adapted Wynn's violation-of-expectation paradigm (Wynn 1992) by using the unexpected appearance and disappearance of social objects and measuring chicks' looking times. Chicks' looking-time responses indicated sensitivity not only to the presence of hidden objects but also to their absence, suggesting that they can represent empty sets. A further conceptual step, however, involves integrating empty sets into proto-arithmetical tasks. Unlike simple numerical discrimination abilities, proto-arithmetic requires elementary addition and subtraction operations involving both the representation of numerosities and their manipulation in working memory (DeStefano and LeFevre 2010; Cragg and Gilmore 2014; Howard et al. 2019).

Within this domain, empirical evidence involving empty sets remains fragmented. A study on zebrafish implicitly suggests that zebrafish can treat the empty set as a possible outcome of numerical operations (Potrich et al. 2024). In this experiment, social stimuli were initially hidden behind panels and, before the fish was released, one stimulus from each group was made visible again. Under these conditions, zebrafish preferred the panel concealing the larger resulting numerosity in a 2 *vs*. 1 comparison, suggesting their ability to represent the empty set as a possible outcome of elementary proto-arithmetic operations (i.e., the larger group: 2 - 1 = 1; the smaller group: 1-1 = 0, (Potrich et al. 2024)). However, while it has been demonstrated that an

empty set can serve as the starting point of a proto-arithmetic operation (Rugani et al., 2009), it remains unknown whether young birds can integrate the empty set as the outcome of addition and subtraction in proto-arithmetical tasks. Addressing this gap is therefore essential for understanding whether empty sets can be integrated into early proto-arithmetical representations in precocial species such as domestic chicks.

Our study is innovative in a plethora of ways. First, to the best of our knowledge, it is the first to examine the effects of domestication on proto-arithmetic abilities in chickens by directly comparing multiple breeds. Indeed, no previous study has tested more than one breed in a proto-arithmetic task, and two of the breeds included here, the Red Junglefowl and the Gallina Padovana, have not previously been assessed for these abilities. Second, in previous proto-arithmetic studies with chicks (Rugani et al. 2009, 2011, 2013a, 2017, 2020, 2022), multiple numerical comparisons were tested across different individuals rather than within the same individual. For example, one chick was tested only on 2 *vs*. 3, whereas another was tested only on 3 *vs*. 4. Instead, here we are able to assess how individual chicks responded across different numerical comparisons.

Different outcomes can be predicted depending on the effects of domestication on cognitive abilities and on the functional relevance of numerical abilities. If domestication and selection for productivity have shifted resources away from proto-arithmetic ability towards production traits, wild and less selected breeds should outperform the highly selected one. If, however, memory has improved with domestication and selection for productivity (Dudde et al. 2018), the highly selected breed may perform better. If proto-arithmetic abilities are affected not only by domestication but also by selection for productivity, we would expect differences between the two domesticated breeds. Finally, if proto-arithmetic abilities represent a basic component of core numerical knowledge (Feigenson et al. 2004) that is robust to domestication and selection for productivity, we would expect comparable performance across breeds.

Methods – Main Experiment

Ethical approval

All procedures used in the experiments included in this study were examined and approved by the Ethical Committee of the University of Padova (Organismo Preposto al Benessere Animale – OPBA; approval number 27739, 89/2023 as well as amendment n. 87703).

Subjects and Housing

Subjects were 20 chicks of the domesticated breed Ross308, 20 chicks of the domesticated breed Gallina Padovana and 8 Red Junglefowl chicks. Among the Ross308 chicks were 10 males and 10 females, while for the other two breeds no phenotypical sexing was possible at their young age.

All three breeds were tested at the Department of General Psychology of the University of Padua Italy. The fertilized eggs of the domesticated breeds were obtained from a local commercial hatchery (Ross308: La Pellegrina, San Pietro in Gu, Padua, Italy; Gallina Padova: Istituto di Istruzione Superiore "Duca degli Abruzzi", Padova, Italia). For the Red Junglefowl, eggs were collected from natural nests in the game park “Oasi di Sant’Alessio”, Sant’Alessio con Vialone, Pavia, Italy (Società Pavese di Ornitologia). Eggs were incubated-at a temperature of 37-38 °C and a humidity of 60% in the light. After hatching, when the chicks were only a few hours old,

they were transferred to standard metal home cages (28 cm × 32 cm × 40 cm), whose floor was lined with absorbent paper towels. Fluorescent lamps (36 W) located 45 cm above the floor illuminated the cages. An automated lightning system regulated the light/dark cycle so that light was on from 7 a.m. to 7 p.m., and in the remaining time light/dark-cycles were alternated every 2 to 3 hours. In the rearing room, temperature and humidity were between 28–31 °C and 68%, respectively. Both food (chick starter crumbles) and water were available *ad libitum*. In addition, on the first day, each chick received three mealworms (*Tenebrio molitor*) to habituate them to the food reward to be used during testing.

Chicks were housed singly, in the presence of five rearing objects; i.e., two-dimensional red squares (4 cm × 4 cm) made of laminated paper. Taking advantage of the mechanisms of filial imprinting and social attachment (Lorenz 1970), numerous studies have shown that rearing with artificial objects leads chicks to perceive them as social companions. As a result, chicks are highly motivated to reunite with (the larger number of) these rearing objects when given the choice (Regolin et al. 2005; Lemaire et al. 2021; Rugani et al. 2022; Loconsole et al. 2023). The stimuli were hanging from transparent threads on the chicks' eye-level (the lowest part was approximately 2 to 6 cm above the ground) attached to a plastic frame above the cage. As the chicks used in this study were part of a project that assessed various cognitive and perceptual abilities intra-individually, they participated in other experiments before undergoing the proto-arithmetic task on days 4 and 5 of life. Specifically, chicks were tested in a human-hand approach test (day 1), a memory task (day 2) and a detour task (day 2) and a numerical discrimination task (day 3). Due to the requirements of the numerical discrimination task (for a similar procedure see Brosche et al. 2025), in addition to the rearing objects, during the first 2.5 days, chicks were exposed to panels showing numerical stimuli (arrays of dots) (Brosche et al., in preparation).

Experimental apparatus

The arena was circular with a diameter of 95 cm and a height of 30 cm. Inside the arena, adjacent to a wall, there was a start box (11 cm × 20 cm × 20 cm), whose front part was transparent, see Figure 1 (Rugani et al. 2009). Two opaque blue screens (16 cm x 8 cm each) were positioned in front of the starting box. Each screen had 3 cm-long side walls protruding to the back to prevent chicks from peeking behind the screen before circumnavigating it. During the familiarization (see *Training and Familiarization*) only one centrally positioned screen was presented in the arena (33.5 cm from the start box). During training and testing, both screens were presented and positioned as depicted in Figure 1.

In all training and test trials, the experimenter operated from behind the screens (i.e., from the side opposite the start box).

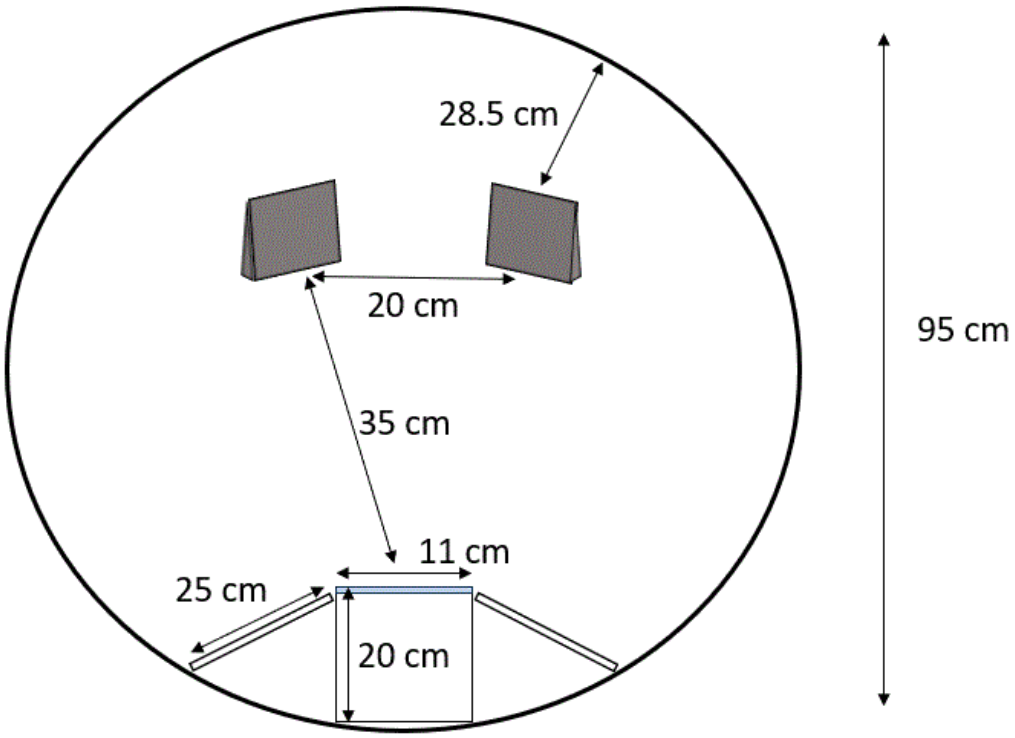


Figure 1: Schematic overview of the arena used in the proto-arithmetic test

Stimuli

In the proto-arithmetic task, the chick watched two sets of objects being hidden behind the screens. We exploited chicks' propensity to approach the larger number of objects they have been reared with when given the choice between two sets (Rugani et al. 2010, 2009, 2022; for red junglefowl chicks' preference to rejoin rearing objects see Kirkden et al. 2008). Testing stimuli resembled the objects chicks had been reared with. Specifically, they consisted of two-dimensional red laminated cardboard squares hanging from a transparent thread. However, as opposed to the rearing objects (4 cm × 4 cm), the test objects varied in size in some trials to control for the overall surface and perimeter of the two sets. Whenever a chick reached the rearing object(s) behind the screens in the training or the test, it was allowed to spend a few seconds in their proximity as a social reward.

In all cases in which the area of the objects was not or could not be controlled (all 0 *vs*. 5 trials and a third of the 2 *vs*. 3 trials), objects of the same size as the rearing objects (4 cm × 4 cm) were used. For the 2 *vs*. 3 area control condition (see *Test*), both sets had a total area of 36 cm^2, which we achieved by using three squares with an area of 12 cm^2 (3.46 cm × 3.46 cm) each for the set of three and an area of 18 cm^2 (4.24 cm × 4.24. cm) each for the set of two. For the 2 *vs*. 3 perimeter control condition, the overall perimeter of each set amounted to 36 cm. That is, the set of three squares contained squares with a perimeter of 12 cm per square (side length of 3 cm), while the set of two consisted of squares with a perimeter of 18 cm per square (side length of 4.5 cm).

Training and Familiarization

The familiarization and subsequent training started on the chicks' fourth day of life and could continue into the fifth if chicks did not reach the training criterion by the evening of the fourth

day. Compared with similar studies (e.g., Rugani et al. 2009), the familiarization in our study was relatively short (see below), as the chicks had already learned to circumnavigate the blue screens in a memory test they completed on their second day of life (see *Subjects and Housing*).

As a first step, chicks were allowed to freely explore the arena for 3 min with one of the objects to be used in the test (a standard 4 cm × 4 cm one) present in front of the single screen present during familiarization. After 3 min, the experimenter slowly moved the object (via a transparent thread and a hook) behind the screen. Once a chick promptly followed the object behind the screen it proceeded to the training.

In the training, the experimenter moved the object behind one of the two laterally positioned screens, alternatingly on the left or on the right. The chick watched the movements of the object from the start box and was released a few seconds after the end of the presentation. A few second after having made a choice or after 3 min, the chick was gently removed from the arena and placed back into the start box, before the next trial began. To start the test, the chicks needed to join the object behind the screen within 3 min in three consecutive trials (cf. Regolin et al. 2005; Rugani et al. 2020, 2022). If, after 10 trials, a chick did not pass the criterion, it was brought back to the home cage for approximately 1 hour. This could be repeated up to three times before the chick was excluded from the study.

Test

The test was conducted on the fourth or fifth day, depending on the length of the training sessions for each chick (Rugani et al. 2017). Each chick participated in 20 trials, divided into 4 blocks of 5 trials each. We decided to conduct relatively short blocks to maintain the chicks' attention, given the relatively long duration of the stimulus presentations before each trial (Rugani et al. 2017). The first block immediately followed the last phase of the training.

At the beginning of each trial, the chick saw the stimuli move one by one and hiding behind one of the two screens (initial placement). Then some objects moved from one screen to the other (operations phase) resulting in screens hiding different final sums (final comparison). As in the training, the chick watched the stimulus presentation from the start box and was released into the arena approximately 3 s after the end of the presentation.

A total of eight different trial types were performed. They differed in the final sum/comparison (2 *vs*. 3 or 0 *vs*. 5), the exact operation to be performed (how many objects were moved), and the controls, see Table 1. In 12 trials the final comparison was 2 *vs*. 3 and in eight trials the final comparison was 0 *vs*. 5.

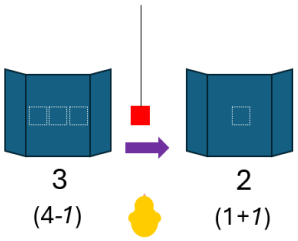
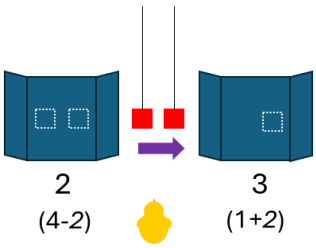
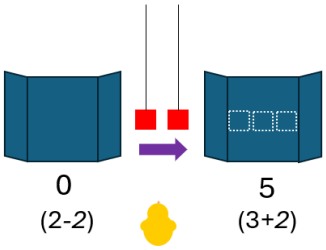
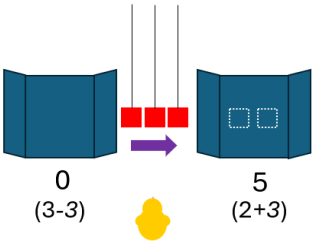
Within the 2 *vs*. 3 trials, we controlled for total area or total perimeter of the two final sums in some trials. Specifically, in four of the 12 2 *vs*. 3 trials, the two final sets were matched for total stimulus area, and in another four they were matched for total perimeter. No control for continuous variables was administered in the remaining four trials. For exact dimensions in each control, see section *Stimuli*.

In addition to the control for continuous variables, we included "larger initial sum" controls and "last movement" controls in the 2 *vs*. 3 trials. That is, half of all 2 *vs*. 3 trials (six trials, i.e., two of each of the three control types for continuous variables explained above) were "larger initial sum control" trials ("LISC"). This meant that, in order to choose the screen hiding the larger amount, i.e. the three objects, chicks needed to choose against the cue provided by the initial displacement of the objects (Rugani et al. 2009). In other words, the side hiding the larger sum of objects after the initial placement no longer hid the larger sum after all movements at the end

of the trial. The other half of the 2 *vs*. 3 trials (six trials, i.e., two of each of the three control types for continuous variables explained above) were controlled for a potential preference to follow the last movement of an object (directional cue), even if it leads to the incorrect screen. In these "last movement control" ("LMC") trials, the last object to be moved was moved towards the incorrect screen that hid the smaller sum in the end (4-*1 vs*. 1+*1*; numbers in italics denote objects being moved).

In the 0 *vs*. 5 trials, half of the trials (four out of eight) controlled for the initially larger sum. That is, in these trials, at the end of the initial placement, the larger set was located behind the screen which hid zero objects at the end of the trials, after all movements (one side 3-*3 vs*. other side 2+*3* numbers in italics indicate objects being moved). In the remaining four 0 *vs*. 5 trials, the larger sum stayed on the same side throughout the operations (2-*2 vs*. 3+*2*).

Table 1: Overview of proto-*arithmetic operations and resulting numerical comparisons employed in the* proto-*arithmetic task. Numbers in italics underneath the "operations" illustrations indicate the objects that were moved from one screen to another after the initial hiding. "Initial comparison" refers to the number of objects on each side before the displacements. Note that all objects were moved one after the other and that sides were counter-balanced across trials although just one version is depicted here for simplicity.*

Initial comparison	Operations	Final comparison
4 *vs*. 1		2 *vs*. 3
4 *vs*. 1		2 *vs*. 3
3 *vs*. 2		5 *vs*. 0
3 *vs*. 2		5 *vs*. 0

The side hiding the greater number of objects at the end of the presentation (correct side) was semi-randomized across trials in a way that the same side was correct in no more than two consecutive trials, following an established sequence (Fellows 1967; Rugani et al. 2015), i.e., L-R-L-L-R-R-L-R-L-R-L for half of the chicks and the reverse for the other half of the chicks. Also the fact which of the two sets, the initially smaller or the initially larger one, was hidden first was counter-balanced across trials in a similarly semi-random way. However, we always moved all objects of one set (behind the same screen) before moving any object behind the other screen, not to over-strain chicks' working memory capacity (cf. Moher and Feigenson 2013). During the

initial movement, before the proto-arithmetic operations, the objects were always moved behind the screens from the outside (i.e., the side of the screen facing away from the other screen; right for the right screen and left for the left screen). During the proto-arithmetic operations, the objects were always moved between screens, i.e., from the left side of the right screen to the right side of the left screen and vice versa. All movements were performed by the experimenter via a hook that could be attached to the transparent thread the objects were hanging from.

A trial ended when the chick circumnavigated either screen with at last ¾ of its body, which was the operationalization of a choice, or after 3 min (Rugani et al. 2009). If a chick did not make a choice within 3 min, the trial was repeated immediately after. Whenever a chick failed to make a choice in three consecutive trials, it was brought back to its home cage, and another attempt was started after a break of at least one hour. Chicks which scored three consecutive null trials three times during the test were excluded from the experiment.

All trials were video recorded for later analyses of chicks' choice (i.e., the first screen a chick circumnavigated in a trial).

Statistical Analysis

Inter-rater reliability

Approximately 20% of the videos, 174 trials, were coded by a second observer to assess inter-rater reliability. Cohen's kappa indicated substantial agreement (K = 0.80) (Landis and Koch 1977).

Excluded trials and subjects

Out of 960 theoretically planned trials (48 chicks × 20 trials), 899 could be analyzed. For the remaining 61 trials, no usable data existed due to technical failure (no video) for 25 trials across five chicks, experimenter error (the trial should have been repeated due to a null choice but was not repeated) for five trials across five chicks, and chicks' lack of motivation (leading to their exclusion after they had made three consecutive null choices for the third time) for 31 trials across three chicks.

Power analysis

Given that our sample size per breed was inherently limited due to practical limitations (availability of the Red Junglefowl chicks, time constraints and maximum number of trials chicks could complete), we did not calculate the desired sample size based on a threshold of statistical power we wanted to attain, but instead report and quantify the sensitivity of the fixed design.

For this purpose, we evaluated the expected statistical power of the planned analyses using Monte Carlo simulations under a set of plausible data-generating scenarios. Expected breed-level accuracies were fixed to empirically plausible values: 0.55 for Ross308 and 0.70 for Red Junglefowl (see for example Rugani et al. 2009, 2022). Gallina Padovana was set at the midpoint between these values on the logit scale. These values were intended to represent a monotonic pattern of increasing performance across breeds, from slightly above-chance performance in Ross308 to the highest expected performance in Red Junglefowl.

Simulated responses were generated from a binomial generalized linear mixed model with a logit link function. Subject-level heterogeneity was introduced through a random intercept on

the logit scale. We considered three levels of between-participant variability, with standard deviations of 0.30, 0.50, and 0.70, representing different plausible degrees of heterogeneity. No fixed effect of comparison and no breed × comparison interaction were included in the data-generating model. However, each simulated dataset was analyzed with the planned inferential model including the breed × comparison interaction, in order to reproduce the intended analysis strategy and provide a more conservative assessment of post-hoc contrasts across breed/comparison combinations.

For each simulated dataset, we extracted two families of planned contrasts: pairwise differences among breeds, averaged across comparisons, and one-sided tests of each breed/comparison combination against chance level. Within each family, we controlled the familywise error rate using a single-step multiplicity adjustment based on the multivariate distribution of the joint test statistics (Hothorn et al. 2008; Bretz et al. 2016). Thus, multiplicity correction was applied separately to the family of breed contrasts and to the family of tests against chance.

Here, we report a summary of the power estimates. The full simulation report is provided in in an online repository (see *Data Availability*). Under the scenario with moderate between-subject variability (σ_{id} = 0.50), power to detect breed differences was limited and depended on the contrasts. Power was moderate for Red Junglefowl *vs*. Ross308 (48.6%), the largest predicted contrast. It was lower for Gallina Padovana *vs*. Ross308 (22.4%) and Gallina Padovana *vs*. Red Junglefowl (12.3%). The average power across contrasts was 27.8%. Power for above-chance tests was higher on average (49.3%), albeit uneven across breeds: it was high for Gallina Padovana and Red Junglefowl (61.2%–76.8%),), but remained low for Ross308 (8.7%–13.3%). The design had moderate sensitivity for detecting above-chance performance in Gallina Padovana and Red Junglefowl (power = 61.2%–76.8%), but limited sensitivity for detecting breed differences (power ≤ 50%). This pattern mainly reflects the limited number of subjects and trials, together with the relatively narrow range of expected accuracies in the task.

The sensitivity analysis indicates that only differences of the size assumed there could be detected with reasonable power. Thus, non-significant results indicate lack of evidence of a difference, not ruling out the possibility of smaller, undetected effects, but not as evidence of absence.

Analysis

All analyses were performed in R version 4.6.1 (R Core Team 2024).

To analyze whether chicks' accuracy differed between breeds, we fitted a Generalized Linear Mixed Model (glmmTMB package version 1.1.14 (Brooks et al. 2017)) with a binomial family and a logit link function. The response variable was chicks' choice (1 = correct, 0 = incorrect) and the predictors were chicks' Breed (Ross308, Gallina Padovana, or Red Junglefowl), the Comparison (0 *vs*. 5 or 2 *vs*. 3) and the interaction between the two. Note that we did not interpret the statistical significance of the interaction. Subject was included as a random intercept with the random slope of Comparison, excluding the correlation between the two (i.e., a diagonal structure of the random effects matrix). The model included 899 observations across 48 subjects. We assessed model stability by excluding one subject at a time, see Supplementary Material 1.

Subsequently, we compared the three breeds and analyzed whether chicks performed significantly above chance level (0.5) for each combination of Breed (one-sided tests) and Comparison using emmeans version 2.0.3 (Lenth 2022) with the multivariate adjustment

(adjust = "mvt" argument of the emmeans function) for familywise error-rate control within each macro-hypothesis (i.e., breed differences and comparison to chance).

We did perform significance testing on chicks' performance in each trial type in the 2 *vs*. 3 final comparison. However, to descriptively illustrate performance across trial types, we fitted a GLMM and plotted the results. This model comprised all 536 trials in which the final comparison was 2 *vs*. 3. Chicks' response choice was the response variable, while Breed, Trial Type, and Trial number (z-transformed) were included as predictors. Subject was included as a random effect. For the random Intercept of Subject, the estimated standard deviation was 0, probably due to the limited number of trials.

Results – Main Experiment

Ross308 chose the larger final sum in an average of 57.87% (estimated probability 0.596, 95%-CIs = [0.543; 0.648]) of the trials, Gallina Padovana in 55.98% (estimated probability = 0.568, 95%-CIs = [0.517; 0.617]) of the trials, and Red Junglefowl chicks in 57.33% (estimated probability = 0.588, 95%-CIs = [0.505; 0.666]) of the trials. Breed did not have a significant effect on chicks' performance. Ross308 chicks' performance did not differ significantly from that of Gallina Padovana (z-ratio = 0.801, p = 0.699) or Red Junglefowl (z-ratio = 0.205, p = 0.977) chicks. Neither did Gallina Padovana and Red Junglefowl chicks' performance differ significantly (z-ratio = -0.404, p = 0.913).

Considering all breeds together, chicks chose the final larger sum in 51.87% of the 2 *vs*. 3 trials and in 64.46% of the 0 *vs*. 5 trials. Ross308 and Gallina Padovana chicks performed significantly above chance level in 0 *vs*. 5 trials (Ross308: z-ratio = 4.167, p < 0.001; Gallina Padovana: z-ratio = 2.635, p = 0.025), while Red Junglefowl chicks' performance was marginally significant (z-ratio = 2.360, p = 0.054). However, in the 2 *vs*. 3 trials, none of the breeds performed significantly above chance level (Ross308: z-ratio = 0.242, p = 0.958; Gallina Padovana: z-ratio = 0.817, p = 0.756; Red Junglefowl: z-ratio = 0.282 p = 0.951), see Figure 2. The different conditions in the 2 *vs*. 3 comparison are descriptively visualized in Figure 3. For detailed model outputs see Tables S1, S2, and S3 in the Supplementary Material 1.

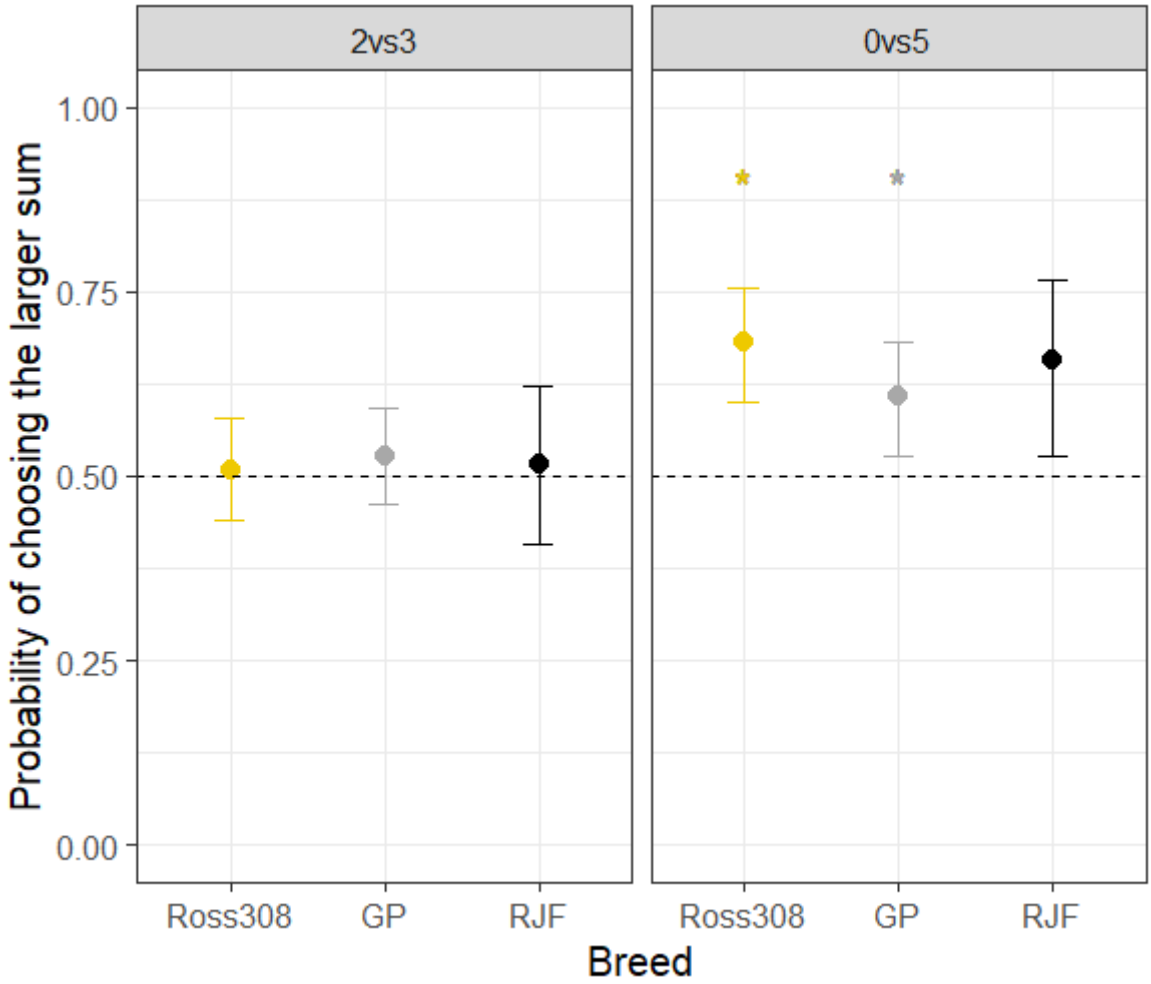


Figure 2: Predicted probabilities of choosing the screen hiding the larger sum across conditions for the three breeds: Ross308 (yellow), Gallina Padovana (gray), and Red Junglefowl (black). Points represent model-estimated probabilities of choosing the larger sum per breed, and error bars represent 95% confidence intervals. The gray dashed line indicates chance level. Asterisks denote significant deviation from chance level. Significant contrasts are corrected for multiple comparisons while confidence intervals are calculated from model summaries

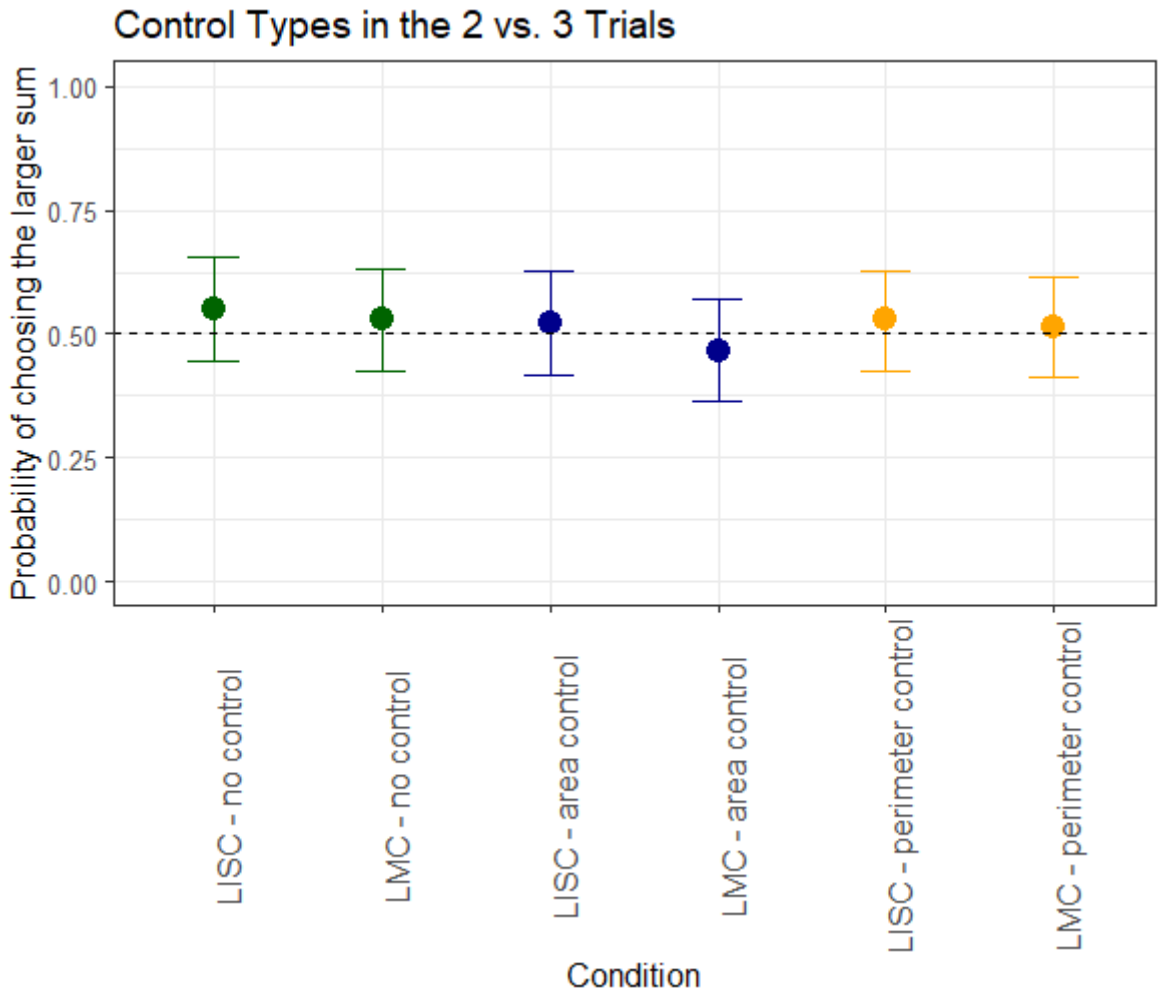


Figure 3: Probability of choosing the larger sum for all conditions in the 2 vs. 3 comparison. Green = no control for continuous variables; blue = control for overall area; orange = control for overall perimeter. LISC= larger initial sum control; LMC = last movement control. Trials in which continuous variables (area of perimeter) were not controlled are depicted in green, trials with control for overall area in blue and trials with control for overall perimeter in orange. The dashed line indicates chance level. Points represent model-estimated probabilities of choosing the larger sum per trial type, and error bars represent 95% confidence intervals.

Methods – Control Experiment

The results of our main experiment show that chicks of all breeds successfully selected the larger sum in 0 *vs*. 5 but not in the non-empty set comparison 2 *vs*. 3. To assess whether these results indeed reflect an ability to handle empty-set outcomes in a proto-arithmetic task, or can be attributed to low-level explanations (see below), we conducted a control experiment in which Ross308 chicks completed a proto-arithmetic task with the final comparison 0 *vs*. 3.

Specifically, we wanted to address the following alternative explanations for chicks' success in 0 *vs*. 5: (1) the numerical distance between 0 and 5 (distance 5) is greater than between 2 and 3 (distance 1); (2) the chicks could have peeked behind the screens from the center of the arena and more easily distinguishing 0 *vs*. 5 this way than 2 *vs*. 3 due to the absence of a middle barrier between the screens, and (3) the lower number of objects moved in 2 *vs*. 3 (one or two) than in 0 *vs*. 5 (two or three) paradoxically led chicks to perform at chance level in the former condition.

Subjects and housing

Subjects were 16 Ross308 chicks (8 males and 8 females). We conducted this control experiment only with Ross308 chicks as the two other breeds are difficult to obtain and, unlike for the Ross308, eggs are only available during breeding season in spring. Chicks were incubated and housed as described for the main study. However, given that they did not participate in other experiments, they were only exposed to the red squares in their cage but not to other stimuli.

Experimental apparatus and procedure

The experimental apparatus was identical to that used in the main experiment.

To improve the clarity of scoring an opaque white plastic barrier was placed between the screens. Viewed from above the barrier had a triangular shape (dimensions: the barrier protruded 37 cm into the arena (height of the triangle); distance between the apex of the triangle and each side of the arena, perpendicular to the height of the triangle = 46 cm, acute angle of the triangle = 30°, length of each side of the barrier = 38.3 cm; height of the barrier = 35 cm). Choice was defined by the chick passing the imaginary line that extends the front side of the panel to the middle barrier on the one side, and to the outer wall of the arena on the other side. See Figure S3 in the Supplementary Material 3. In this control experiment, the experimenter operated from behind the start box (i.e., from the side opposite the screens) during all training and test trials, thereby further minimizing the possibility of experimenter cueing.

The test was divided into two sessions with a break of at least 1 hour in between. In the first session, chicks saw the operation 1-*1 vs*. 2+*1*, i.e., one object was moved, and in the second session 2-*2 vs*. 1+*2,* i.e., two objects were moved (corresponding to a larger initial set control). In all trials, the standard 4 cm × 4 cm squares were used.

Statistical analysis

Inter-rater reliability

A second observer independently coded approximately 20% of the trials (53 trials). Cohen's kappa indicated almost perfect agreement (K = 0.90) (Landis and Koch 1977).

Excluded trials and subjects

Out of the 16 subjects for which valid test data exist, one was excluded due to having completed an insufficient number of trials (less than one complete session). For the remaining 15 subjects, 62 trials were not conducted or could not be analyzed due to chicks ceasing to participate (25 trials), experimenter error (13 trials) or equipment failure (24 trials). The remaining 226 trials were considered in the analysis.

Analysis

As for the main experiment, we calculated a GLMM with chicks' choice as a response variable and Session (equivalent to the operation being performed, move 1 or move 2) and trial (z-transformed) as fixed effects. Subject was included as a random intercept with the random slope of Session, excluding the correlation between the two. Based on this model, we calculated whether chicks performed significantly above chance level (one-sided tests) in each session (condition) and whether the performance differed significantly between sessions using emmeans with the multivariate adjustment (adjust = "mvt") for familywise error-rate control.

Results – Control Experiment

In the control experiment, chicks chose the final larger sum in 77.97% of the trials. Chicks performed significantly above chance level in both conditions of the control experiment (Condition 1-*1 vs*. 2+*1*: prob = 0.84, z-ratio = 5.273, p < 0.001; Condition 2-*2 vs*. 1+*2*: prob = 0.75, z-ratio = 3.525, p < 0.001). Their performance did not differ significantly between the two conditions (z-ratio = 1.486, p = 0.137). For details and model output tables see Tables S4, S5, and S6 in Supplementary Material 1.

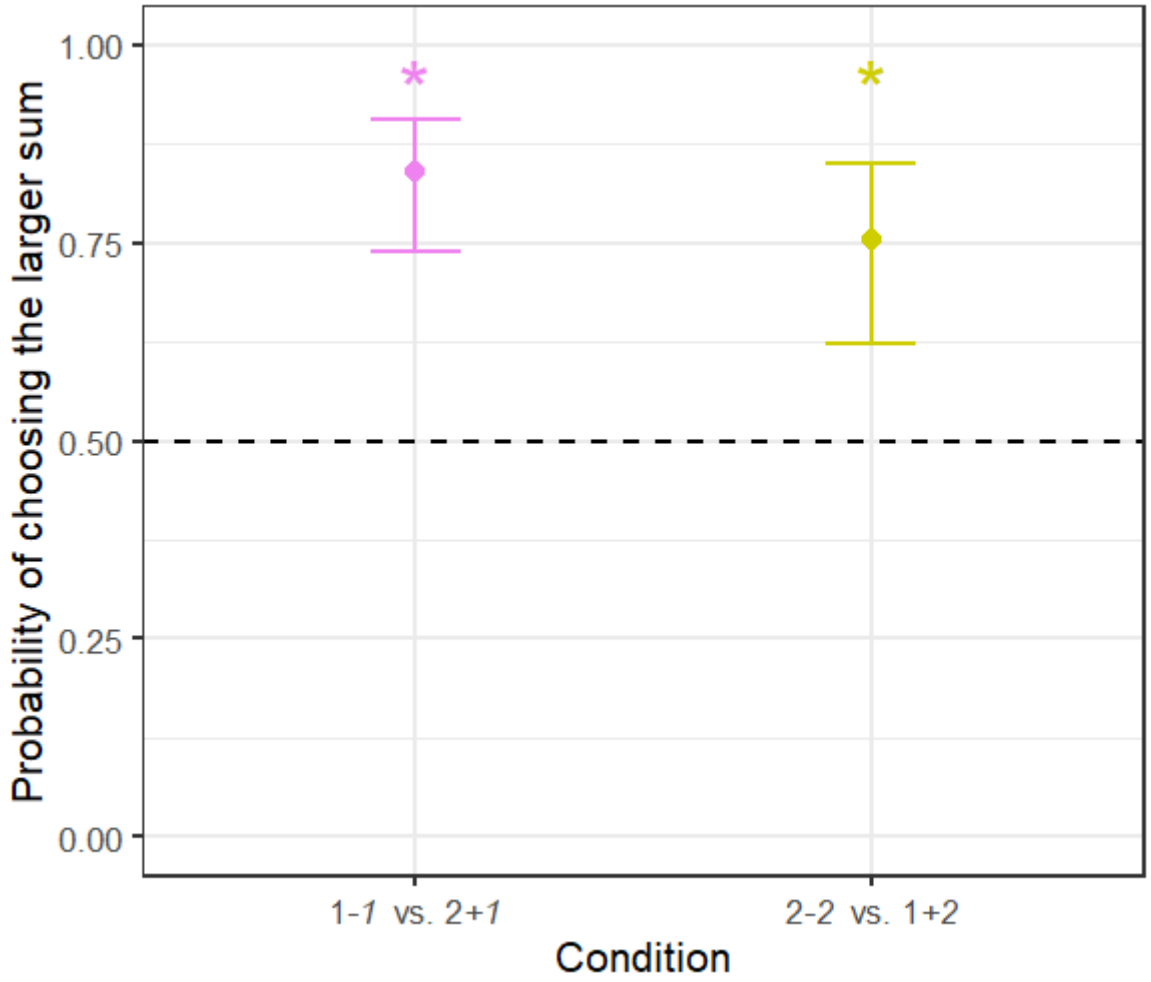


*Figure 4: Chicks' performance in the two conditions of the control experiment. The condition in which one object was moved (1-*1 vs. *2+*1*) is depicted in pink and the condition in which two objects were moved (2-*2 vs. *1+*2*) is depicted in yellow. Points represent model-estimated probabilities of choosing the larger sum per condition, and error bars represent 95% confidence intervals. The dashed line indicates chance level. Asterisks denote significant deviation from chance level.*

Discussion

In this study, we investigated whether domestication and selection for productivity have shaped chicks' proto-arithmetic abilities. To this end, we tested three breeds in a task requiring them to represent, maintain, and update hidden numerosity, including operations that resulted in either empty or non-empty sets.

Our main finding is that the Ross308, a highly productive domesticated breed, the Gallina Padovana, a domesticated breed not intensively selected for productivity, and the Red Junglefowl, domesticated chickens' wild ancestor, did not differ in their proto-arithmetic performance. However, our simulation-based sensitivity analysis showed that the design had limited power to detect breed differences (≤ 50% even for the largest expected contrast), so that relevant differences cannot be excluded. Making the sensitivity of a constrained design explicit, as done here, allows non-significant results to be interpreted for what they are: lack of evidence rather than evidence of absence. All breeds showed high performance when the final comparison was 0 *vs*. 5 but not in 2 *vs*. 3, although performance between the two conditions was not directly compared. Both final comparisons resulted from successive object displacements requiring chicks to update the representation of hidden sets in working memory.

Remarkably, the high success rate in the 0 *vs*. 5 trials suggests chicks' ability to solve proto-arithmetic problems involving empty-set outcomes. This is not trivial, as the representation of empty sets follows a special trajectory also in human ontogenetic development: before four years of age, children do not possess a representation of zero as a quantity and, although they are able to manipulate small numerosities, they fail when the outcome of an operation results in an empty set (Wynn 1998). Around four years of age, however, zero begins to be understood as a number, an acquisition that is strongly dependent on experience with numbers and numerical symbols (Bialystok and Codd 2000; Bialystok et al. 2000). Zero retains a special status even in adulthood (Nuerk et al. 2004) and that it requires longer processing times than other numbers in reading tasks (Brysbaert 1995).

The results of our control experiment confirm that chicks can solve comparisons involving an empty set also when the numerical distance is smaller than 5. To obtain even clearer results on chicks' choice in the control experiment, we modified methodological aspects beyond the numerical comparisons tested. These methodological modifications, however, mean that a direct comparison between chicks' success in 0 *vs*. 3 (control experiment) and 0 *vs*. 5 (main experiment) would be imprecise. Thus, future studies should specifically focus on comparing chicks' performance in various zero-sum comparisons to assess whether they exhibit a distance effect, performing better with larger numerical distances (Moyer and Landauer 1967; Kirschhock et al. 2021; Merritt et al. 2009).

In the comparisons involving an empty set, equalizing the overall area or perimeter between sets, as we did for the 2 *vs*. 3 comparison, was logically impossible. Hence, we cannot exclude that chicks might have relied on these variables when making their choice especially in 0 *vs*. 5 and 0 *vs*. 3. It is well known that subjects, such as infants (Feigenson et al. 2002), fish (Gómez-Laplaza et al. 2019) or birds (Bogale et al. 2011), may rely on continuous variables rather than numerosity, when solving numerical tasks. However, the above-cited tasks involved food items, for which choosing few large items can be as advantageous as choosing numerous smaller ones (Agrillo and Bisazza 2014; Agrillo 2015). Our study, in contrast, used social stimuli. Chicks generally prefer the larger set of rearing objects even when they are equalized for continuous variables like surface area or volume (Rugani et al. 2010). Thus, future studies could aim to run the same empty-set comparisons twice with different sizes of stimuli (e.g., 0 *vs*. 3 with small stimuli and 0 *vs*. 3 with large stimuli) to assess whether not just numerosity, but also area or perimeter, might create a distance effect in such comparisons.

Chicks' poor performance in 2 *vs*. 3 is at odds with previous results indicating that chicks succeed in 2 *vs*. 3 proto-arithmetic comparisons (Rugani et al. 2009). Several factors could explain this discrepancy of results. First, Rugani et al.'s chicks belonged to the Hybro breed, a

commercial hybrid similar to the Ross308, though characterized by a slightly lower production performance. Given that Hybro chicks performed differently from Ross308 chicks in an ordinal task (Rugani et al. 2023), it is conceivable that the divergent results could be attributed to breed differences. However, a breed-related explanation seems unlikely, as the three breeds tested here showed comparable performance despite differing more markedly in domestication history and selection for productivity than Hybro and Ross308, which are both broiler breeds. Second, the control of continuous variables could have increased task difficulty, as area and contour length were controlled in two thirds of the 2 vs. 3 trials. In previous studies, chicks successfully discriminated 2 vs. 3 even when objects were moved between sets after their initial disappearance and continuous variables, such as overall area and perimeter, were controlled (Rugani et al. 2009). By contrast, when processing larger numerosities, involving approximately 5–10 items, performance appears to benefit from redundant numerical and continuous information, such as overall area and perimeter (Rugani et al. 2011). Thus, reliance on continuous variables seems unlikely to explain the present results, particularly because chicks did not appear to perform better when these cues were available (see Figure 3 for the observed pattern). Third, unlike in previous studies, where different groups of chicks were tested in different numerical conditions, here each chick was tested across all conditions (Rugani et al. 2009, 2011, 2013a, 2017, 2020, 2022). While this allowed us two observe individuals' responses across conditions, it might have also confused chicks and decreased their performance relative to previous studies, making them succeed only in the presumably easier comparisons (0 *vs*. 5 as opposed to 2 *vs*. 3). The latter therefore appears to be the most plausible explanation, suggesting that repeated exposure to different proto-arithmetic operations may have increased task complexity and reduced performance.

Contrary to our hypothesis, the three breeds of chicks did not differ in their proto-arithmetic performance – or differences were too small to be detectable with the given statistical power. For other numerical abilities, such as quantity discrimination in dogs and wolves (Range et al. 2014) or ordinal ability compared between two domestic chicken breeds (Rugani et al. 2023), effects of domestication and selection for productivity have been detected, even if research in this area is scarce. Instead, our results suggest that basic proto-arithmetic ability may be similarly relevant to domesticated animals in captivity as it was to their wild ancestors. Besides the possibility of not detecting differences due to moderate statistical power, there are two possible speculative explanations for this lack of differences . On the one hand, it is conceivable that proto-arithmetic ability itself is not subject to evolution because it evolved as a by-product or exaptation of another trait (Gould and Lewontin 1979; Gould 1991). Thus, whether or not proto-arithmetic is still useful to domesticated animals in captivity does not affect animals' ability. On the other hand, it could be that the challenges proto-arithmetic ability is needed for still persist for domesticated animals captivity, maintaining the corresponding evolutionary pressure. The second explanation would be in line with Schluessel et al.'s (2022) study on cichlids and sting rays. The results suggest sophisticated proto-arithmetic abilities in the two species of fish that, presumably, do not strongly rely on proto-arithmetic in their ecological niche. Thus, the authors suggest that proto-arithmetic ability might be a cognitive byproduct conducive to social situations, aiding individual recognition or detecting (social) changes (Schluessel et al. 2022). If sociality is indeed a main driver of proto-arithmetic ability, this might explain the lack of significant domestication effects on this ability.

Social challenges continue to exert evolutionary pressure on domesticated animals, presumably more than those associated with foraging or predator avoidance. There are, however, plenty of reasons why changes in sociality should occur during domestication:

domesticated animals have to compete for food less (Schütz et al. 2001; Price 2002; Garnham and Løvlie 2018) and should exhibit higher social tolerance due to restricted space availability (Price 1984; Newberry 2017). Many domesticated animals are often immature for longer and thus have an extended socialization period (Frank 1980; Wilkins et al. 2014). These factors may contribute to the fact that domesticated forms often outperform wild ones in social cognition tasks (Hare et al. 2002; Rutkauskaite and Jensen 2022; reviewed in Ferreira et al. 2023). Indeed, differences in social behavior are found in various domesticated-wild species pairs, such as pigs (Collarini et al. 2022), mice (Venkatachalam et al. 2024), and guinea pigs/wild cavies (Künzl and Sachser 1999). Also chicken breeds differ in aspects like social learning (Rutkauskaite and Jensen 2022), the expression and frequency of social behaviors (Schütz and Jensen 2001; Schütz et al. 2001; Väisänen et al. 2005), the size of brain areas associated with social behavior (Agnvall et al. 2015), social tolerance (Ericsson et al. 2014; Hedlund et al. 2025), and reaction to social isolation (Gjøen et al. 2023). Nevertheless, social tasks that require simple proto-arithmetic might still pose similarly relevant challenges to domesticated and wild chickens. This might have allowed proto-arithmetic abilities to remain preserved throughout changes in selective pressures as drastic as the transition from a wild to a domesticated environment.

Conclusion

In summary, we found that three breeds of chicks differing in domestication status and selection for productivity showed no differences in proto-arithmetic performance when tested shortly after hatching, thus minimizing the influence of individual experience. This suggests that the effects of domestication and selection for productivity on proto-arithmetic ability are limited or too small to be detected with the given statistical power. Potentially, the selective pressures that necessitated the evolution of proto-arithmetic ability remain sufficiently strong even for domesticated animals in captivity. If so, one might tentatively speculate that proto-arithmetic ability is more, or at least equally, important in intraspecific social situations than in foraging and predator avoidance.

Remarkably, chicks' above-chance performance in 0 *vs*. 5 and 0 *vs*. 3 comparisons provides the first evidence of their ability to handle empty sets in proto-arithmetic operations. Future studies should build on this result by assessing chicks' performance in other empty-set comparisons, to detect potential distance effects, and by disentangling whether chicks rely on numerosity or continuous variables, like overall area and perimeter, in such tasks. This will generate crucial insights into the innate ability to process empty sets in proto-arithmetic operations in non-human animals.

Acknowledgements

We warmly thank the Oasi di Sant'Alessio and the Horti del Collegio Borromeo di Pavia for providing access to the Red Junglefowl chicks, as well as the Istituto Superiore Duca degli Abruzzi for providing access to the Gallina Padovana chicks. We further thank Casey Lee Kurmann, Oliver Condotta, Lorenzo Trevisan and Judit Martì for help with data collection and video coding. The authors gratefully acknowledge the support of the FIS 3 Advanced Grant (FIS-2024-06349) awarded to RR.

Author contributions

K.B., L.D. and R.R. conceptualized the study. K.B. and L.D. collected the data. F.G. conducted the power analysis. K.B. analyzed the data with assistance by L.D. and F.G. The research was supervised by R.R. and G.A. The original draft of the manuscript was written by K.B. and L.D. All authors were involved in reviewing and approved of the final version.

Statements and declarations

Competing interests

The authors have no competing interests to declare.

Data availability

The data and the R script used in the analyses are available here:

https://osf.io/6zakq/overview?view_only=c0bb24b292734a309aaf066475eb93bb

And will also be made available on the authors' institution's public repository after peer review.

References

Agnvall B, Katajamaa R, Altimiras J, Jensen P (2015) Is domestication driven by reduced fear of humans? Boldness, metabolism and serotonin levels in divergently selected red junglefowl (*Gallus gallus*). Biol Lett 11:20150509. https://doi.org/10.1098/rsbl.2015.0509

Agrillo C (2015) Numerical and arithmetic abilities in non-primate species. In: The Oxford handbook of numerical cognition. Oxford University Press, New York, NY, US, pp 214–236

Agrillo C, Bisazza A (2014) Spontaneous versus trained numerical abilities. A comparison between the two main tools to study numerical competence in non-human animals. J Neurosci Methods 234:82–91. https://doi.org/10.1016/j.jneumeth.2014.04.027

Al-Nasser A, Al-Khalaifah H, al Saffar O, et al (2007) Overview of chicken taxonomy and domestication. Worlds Poult Sci J 63:285–300. https://doi.org/10.1017/S004393390700147X

Beilharz RG, Luxford BG, Wilkinson JL (1993) Quantitative genetics and evolution: Is our understanding of genetics sufficient to explain evolution? J Anim Breed Genet 110:161–170. https://doi.org/10.1111/j.1439-0388.1993.tb00728.x

Beilharz RG, Nitter G (1998) The missing E: the role of the environment in evolution and animal breeding. J Anim Breed Genet 115:439–453. https://doi.org/10.1111/j.1439-0388.1998.tb00366.x

Benson-Amram S, Gilfillan G, McComb K (2018) Numerical assessment in the wild: insights from social carnivores. Philos Trans R Soc B Biol Sci 373:20160508. https://doi.org/10.1098/rstb.2016.0508

726 Beran MJ (2004) Chimpanzees (*Pan troglodytes*) Respond to Nonvisible Sets After One-by-One
727 Addition and Removal of Items. J Comp Psychol 118:25–36.
728 https://doi.org/10.1037/0735-7036.118.1.25

729 Bialystok E, Codd J (2000) Representing quantity beyond whole numbers: Some, none, and part.
730 Can J Exp Psychol Rev Can Psychol Expérimentale 54:117–128.
731 https://doi.org/10.1037/h0087334

732 Bialystok E, Shenfield T, Codd J (2000) Languages, scripts, and the environment: Factors in
733 developing concepts of print. Dev Psychol 36:66–76. https://doi.org/10.1037/0012-
734 1649.36.1.66

735 Bogale BA, Kamata N, Mioko K, Sugita S (2011) Quantity discrimination in jungle crows, Corvus
736 macrorhynchos. Anim Behav 82:635–641.
737 https://doi.org/10.1016/j.anbehav.2011.05.025

738 Bonanni R, Natoli E, Cafazzo S, Valsecchi P (2011) Free-ranging dogs assess the quantity of
739 opponents in intergroup conflicts. Anim Cogn 14:103–115.
740 https://doi.org/10.1007/s10071-010-0348-3

741 Brannon EM, Wusthoff CJ, Gallistel CR, Gibbon J (2001) Numerical subtraction in the pigeon:
742 Evidence for a linear subjective number scale. Psychol Sci 12:238–243.
743 https://doi.org/10.1111/1467-9280.00342

744 Bretz F, Hothorn T, Westfall P (2016) Multiple Comparisons Using R. Chapman and Hall/CRC,
745 New York

746 Brooks ME, Kristensen K, Benthem KJ van, et al (2017) glmmTMB Balances Speed and Flexibility
747 Among Packages for Zero-inflated Generalized Linear Mixed Modeling. R J 9:378–400

748 Brosche K, Regolin L, Zazio A, Rugani R (2025) Variability in cognitive strategies in a numerical
749 discrimination task: individual and comparative insights from day-old domestic chicks
750 (*Gallus gallus*). J Anim Sci 103:skaf351. https://doi.org/10.1093/jas/skaf351

751 Brysbaert M (1995) Arabic number reading: On the nature of the numerical scale and the origin
752 of phonological recoding. J Exp Psychol Gen 124:434–452. https://doi.org/10.1037/0096-
753 3445.124.4.434

754 Cammaerts M-C, Cammaerts R (2021) Ants Can Anticipate the Following Quantity in an
755 Arithmetic Sequence. Behav Sci 11:18. https://doi.org/10.3390/bs11020018

756 Chiandetti C, Vallortigara G (2018) Chicken – Cognition in the Poultry Yard. In: Amici F, Bueno-
757 Guerra N (eds) Field and Laboratory Methods in Animal Cognition: A Comparative Guide.
758 Cambridge University Press, Cambridge, pp 97–118

759 Collarini E, Gioia M, Cordoni G, Norscia I (2022) Does the Domestication Syndrome Apply to the
760 Domestic Pig? Not Completely. Animals 12:2458. https://doi.org/10.3390/ani12182458

761 Cragg L, Gilmore C (2014) Skills underlying mathematics: The role of executive function in the
762 development of mathematics proficiency. Trends Neurosci Educ 3:63–68.
763 https://doi.org/10.1016/j.tine.2013.12.001

764 Davis H, Pérusse R (1988) Numerical competence in animals: Definitional issues, current
765 evidence, and a new research agenda. Behav Brain Sci 11:561–579.
766 https://doi.org/10.1017/S0140525X00053437

767 Dehaene S (2011) The number sense: how the mind creates mathematics, Rev. and updated
768 ed.. Oxford University Press, New York :

769 Dehaene S, Dehaene-Lambertz G, Cohen L (1998) Abstract representations of numbers in the
770 animal and human brain. Trends Neurosci 21:355.
771 https://doi.org/https://doi.org/10.1016/s0166-2236(98)01263-6

772 DeStefano D, LeFevre J-A (2010) The role of working memory in mental arithmetic. Eur J Cogn
773 Psychol

774 Dudde A, Krause ET, Matthews LR, Schrader L (2018) More Than Eggs – Relationship Between
775 Productivity and Learning in Laying Hens. Front Psychol 9:.
776 https://doi.org/https://doi.org/10.3389/fpsyg.2018.02000

777 Dudine L, Felisatti A, Brosche K, Rugani R (under review) Perceiving and Acting on Numbers: A
778 Review on Numerical Cognition in Animals with a Focus on Domestic Chicks.
779 iPerception

780 Ericsson M, Fallahsharoudi A, Bergquist J, et al (2014) Domestication effects on behavioural and
781 hormonal responses to acute stress in chickens. Physiol Behav 133:161–169.
782 https://doi.org/10.1016/j.physbeh.2014.05.024

783 Feigenson L, Carey S, Spelke E (2002) Infants' discrimination of number vs. continuous extent.
784 Cognit Psychol 44:33–66. https://doi.org/10.1006/cogp.2001.0760

785 Feigenson L, Dehaene S, Spelke E (2004) Core systems of number. Trends Cogn Sci 8:307–314.
786 https://doi.org/10.1016/j.tics.2004.05.002

787 Fellows BJ (1967) Change stimulus sequences for discrimination tasks. Psychol Bull 67:87–92.
788 https://doi.org/10.1037/h0024098

789 Ferreira VHB, Lansade L, Calandreau L, et al (2023) Are domesticated animals dumber than
790 their wild relatives? A comprehensive review on the domestication effects on animal
791 cognitive performance. Neurosci Biobehav Rev 154:105407.
792 https://doi.org/10.1016/j.neubiorev.2023.105407

793 Frank H (1980) Evolution of Canine Information Processing under Conditions of Natural and
794 Artificial Selection. Z Für Tierpsychol 53:389–99. https://doi.org/10.1111/j.1439-
795 0310.1980.tb01059.x

796 Garland A, Low J (2014) Addition and subtraction in wild New Zealand robins. Behav Processes
797 109:103–110. https://doi.org/10.1016/j.beproc.2014.08.022

798 Garnham L, Løvlie H (2018) Sophisticated Fowl: The Complex Behaviour and Cognitive Skills of
799 Chickens and Red Junglefowl. Behav Sci 8:13. https://doi.org/10.3390/bs8010013

800 Gilmore C (2015) Approximate Arithmetic Abilities in Childhood. In: Cohen Kadosh R, Dowker A
801 (eds) The Oxford Handbook of Numerical Cognition. Oxford University Press, pp 310–330

802 Gjøen J, Jean-Joseph H, Kotrschal K, Jensen P (2023) Domestication and social environment
803 modulate fear responses in young chickens. Behav Processes 210:104906.
804 https://doi.org/10.1016/j.beproc.2023.104906

805 Gómez-Laplaza LM, Romero L, Gerlai R (2019) The role of item size on choosing contrasted food
806 quantities in angelfish (*Pterophyllum scalare*). Sci Rep 9:15305.
807 https://doi.org/10.1038/s41598-019-51753-1

808 Gould SJ (1991) Exaptation: A Crucial Tool for an Evolutionary Psychology. J Soc Issues 47:43–
809 65. https://doi.org/10.1111/j.1540-4560.1991.tb01822.x

810 Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: a
811 critique of the adaptationist programme. Proc R Soc Lond B Biol Sci 205:581–598.
812 https://doi.org/10.1098/rspb.1979.0086

813 Hare B, Brown M, Williamson C, Tomasello M (2002) The Domestication of Social Cognition in
814 Dogs. Science 298:1634–1636. https://doi.org/10.1126/science.1072702

815 Hedlund L, Wright D, Jensen P (2025) Do Laying Hens Form Stable Social Networks?—A Case
816 Study. Poultry 4:7. https://doi.org/10.3390/poultry4010007

817 Hothorn T, Bretz F, Westfall P (2008) Simultaneous Inference in General Parametric Models.
818 Biom J 50:346–363. https://doi.org/10.1002/bimj.200810425

819 Howard SR, Avarguès-Weber A, Garcia JE, et al (2019) Numerical cognition in honeybees
820 enables addition and subtraction. Sci Adv 5:eaav0961.
821 https://doi.org/10.1126/sciadv.aav0961

822 Kirkden RD, Lindqvist C, Jensen P (2008) Effects of domestication on filial motivation and
823 imprinting in chicks: comparison of red junglefowl and White Leghorns. Anim Behav
824 76:287–295. https://doi.org/10.1016/j.anbehav.2008.02.007

825 Kirschhock ME, Ditz HM, Nieder A (2021) Behavioral and Neuronal Representation of
826 Numerosity Zero in the Crow. J Neurosci Off J Soc Neurosci 41:4889–4896.
827 https://doi.org/10.1523/JNEUROSCI.0090-21.2021

828 Künzl C, Sachser N (1999) The Behavioral Endocrinology of Domestication: A Comparison
829 between the Domestic Guinea Pig (*Cavia aperea* f. *porcellus*) and Its Wild Ancestor, the
830 Cavy (*Cavia aperea*). Horm Behav 35:28–37. https://doi.org/10.1006/hbeh.1998.1493

831 Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data.
832 Biometrics 33:159–174

833 Lemaire BS, Rugani R, Regolin L, Vallortigara G (2021) Response of male and female domestic
834 chicks to change in the number (quantity) of imprinting objects. Learn Behav 49:54–66.
835 https://doi.org/10.3758/s13420-020-00446-1

836 Lenth RV (2022) emmeans: Estimated Marginal Means, aka Least-Squares Means

837 Loconsole M, Regolin L, Rugani R (2023) Passive Grouping Enhances Proto-Arithmetic
838 Calculation for Leftward Correct Responses. Symmetry 15:719.
839 https://doi.org/https://doi.org/10.3390/sym15030719

840 Lorenz K (1970) Studies in Animal and Human Behaviour. Volume I: Harvard University Press,
841 Cambridge, MA and London, England

842 Lyon BE (2003) Egg recognition and counting reduce costs of avian conspecific brood
843 parasitism. Nature 422:495–499. https://doi.org/10.1038/nature01505

844 McComb K, Packer C, Pusey A (1994) Roaring and numerical assessment in contests between
845 groups of female lions, *Panthera leo*. Anim Behav 47:379–387.
846 https://doi.org/10.1006/anbe.1994.1052

847 McCrink K, Wynn K (2004) Large-number addition and subtraction by 9-month-old infants.
848 Psychol Sci 15:776–781. https://doi.org/10.1111/j.0956-7976.2004.00755.x

849 Merritt DJ, Rugani R, Brannon EM (2009) Empty Sets as Part of the Numerical Continuum:
850 Conceptual Precursors to the Zero Concept in Rhesus Monkeys. J Exp Psychol Gen
851 138:258. https://doi.org/10.1037/a0015231

852 Moher M, Feigenson L (2013) Factors influencing infants' ability to update object
853 representations in memory. Cogn Dev 28:272–289.
854 https://doi.org/10.1016/j.cogdev.2013.04.002

855 Moyer RS, Landauer TK (1967) Time required for Judgements of Numerical Inequality. Nature
856 215:1519–1520. https://doi.org/10.1038/2151519a0

857 Newberry RC (2017) Behavioral, emotional, and cognitive effects of domestication. In: APA
858 handbook of comparative psychology: Basic concepts, methods, neural substrate, and
859 behavior, Vol. 1. American Psychological Association, Washington, DC, US, pp 315–329

860 Nieder A (2020) The Adaptive Value of Numerical Competence. Trends Ecol Evol 35:605–617.
861 https://doi.org/10.1016/j.tree.2020.02.009

862 Nieder A (2016) Representing Something Out of Nothing: The Dawning of Zero. Trends Cogn Sci
863 20:830–842. https://doi.org/10.1016/j.tics.2016.08.008

864 Nuerk H, Iversen W, Willmes K (2004) Notational Modulation of the SNARC and the MARC
865 (Linguistic Markedness of Response Codes) Effect. Q J Exp Psychol Sect A 57:835–863.
866 https://doi.org/10.1080/02724980343000512

867 Potrich D, Montel L, Stancher G, et al (2024) Proto-arithmetic abilities in zebrafish (*Danio rerio*).
868 Heliyon 10:e40585. https://doi.org/10.1016/j.heliyon.2024.e40585

869 Pozzato A (2019) Allevare la gallina Padovana, una razza dal gran ciuffo. Vita in Campagna:47–
870 49

871 Price EO (2002) Animal Domestication and Behavior. CABI Publishing, New York

872 Price EO (1984) Behavioral Aspects of Animal Domestication. Q Rev Biol 59:1–32.
873 https://doi.org/10.1086/413673

874 R Core Team (2024) R: A Language and Environment for Statistical Computing

875 Range F, Jenikejew J, Schröder I, Virányi Z (2014) Difference in quantity discrimination in dogs
876 and wolves. Front Psychol 5:. https://doi.org/10.3389/fpsyg.2014.01299

Regolin L, Loconsole M, Rosa-Salva O, et al (2025) Numerical cognition in birds. Nat Rev Psychol 1–15. https://doi.org/10.1038/s44159-025-00480-8

Regolin L, Rugani R, Pagni P, Vallortigara G (2005) Delayed search for social and nonsocial goals by young domestic chicks, *Gallus gallus domesticus*. Anim Behav 70:855–864. https://doi.org/10.1016/j.anbehav.2005.01.014

Rosa-Salva O, Rugani R, Felisatti A, Regolin L (2026) Evidence for abstract spatial concept learning in young animals. Cognition 275:106633

Rugani R (2018) Towards numerical cognition's origin: insights from day-old domestic chicks. Philos Trans R Soc B Biol Sci 373:20160509. https://doi.org/10.1098/rstb.2016.0509

Rugani R, Cavazzana A, Vallortigara G, Regolin L (2013a) One, two, three, four, or is there something more? Numerical discrimination in day-old domestic chicks. Anim Cogn 16:557–564. https://doi.org/10.1007/s10071-012-0593-8

Rugani R, Fontanari L, Simoni E, et al (2009) Arithmetic in newborn chicks. Proc R Soc B Biol Sci 276:2451–2460. https://doi.org/10.1098/rspb.2009.0044

Rugani R, Loconsole M, Koslowski M, Regolin L (2022) Processing Individually Distinctive Schematic-Faces Supports Proto-Arithmetical Counting in the Young Domestic Chicken. Animals 12:2322. https://doi.org/10.3390/ani12182322

Rugani R, Loconsole M, Regolin L (2017) A strategy to improve arithmetical performance in four day-old domestic chicks (*Gallus gallus*). Sci Rep 7:13900. https://doi.org/https://doi.org/10.1038/s41598-017-13677-6

Rugani R, Loconsole M, Simion F, Regolin L (2020) Individually distinctive features facilitate numerical discrimination of sets of objects in domestic chicks. Sci Rep 10:. https://doi.org/10.1038/s41598-020-73431-3

Rugani R, Macchinizzi M, Zhang Y, Regolin L (2025) Hatching with Numbers: Pre-natal Light Exposure Affects Number Sense and the Mental Number Line in young domestic chicks. eLife 14:. https://doi.org/10.7554/eLife.106356.2

Rugani R, Regolin L (2021) Numerical Abilities in Nonhumans: The Perspective of Comparative Studies. In: Danesi M (ed) Handbook of Cognitive Mathematics. Springer International Publishing, Cham, pp 1–33

Rugani R, Regolin L, Vallortigara G (2007) Rudimental numerical competence in 5-day-old domestic chicks (*Gallus gallus*): Identification of ordinal position. J Exp Psychol Anim Behav Process 33:21–31. https://doi.org/10.1037/0097-7403.33.1.21

Rugani R, Regolin L, Vallortigara G (2011) Summation of large numerousness by newborn chicks. Front Psychol 2:179. https://doi.org/https://doi.org/10.3389/fpsyg.2011.00179

Rugani R, Regolin L, Vallortigara G (2010) Imprinted numbers: Newborn chicks' sensitivity to number vs. continuous extent of objects they have been reared with. Dev Sci 13:790–7. https://doi.org/10.1111/j.1467-7687.2009.00936.x

Rugani R, Vallortigara G, Regolin L (2013b) Numerical Abstraction in Young Domestic Chicks (*Gallus gallus*). PLOS ONE 8:e65262. https://doi.org/10.1371/journal.pone.0065262

916 Rugani R, Vallortigara G, Regolin L (2015) The use of proportion by young domestic chicks
917 (*Gallus gallus*). Anim Cogn 18:605–616. https://doi.org/10.1007/s10071-014-0829-x

918 Rugani R, Zhang Y, Scarsi B, Regolin L (2023) Hybro chicks outperform Ross308 in a numerical-
919 ordinal task. Cognitive and behavioral comparisons between two broiler strains of
920 newborn domestic chicks (*Gallus gallus*). Poult Sci 103148.
921 https://doi.org/10.1016/j.psj.2023.103148

922 Rumbaugh DM, Savage-Rumbaugh S, Hegel MT (1987) Summation in the chimpanzee (*Pan*
923 *troglodytes*). J Exp Psychol Anim Behav Process 13:107–115.
924 https://doi.org/10.1037/0097-7403.13.2.107

925 Rutkauskaite A, Jensen P (2022) Domestication effects on social information transfer in
926 chickens. Anim Cogn. https://doi.org/https://doi.org/10.1007/s10071-022-01628-2

927 Schluessel V, Kreuter N, Gosemann IM, Schmidt E (2022) Cichlids and stingrays can add and
928 subtract 'one' in the number space from one to five. Sci Rep 12:3894.
929 https://doi.org/10.1038/s41598-022-07552-2

930 Schütz KE, Forkman B, Jensen P (2001) Domestication effects on foraging strategy, social
931 behaviour and different fear responses: a comparison between the red junglefowl
932 (*Gallus gallus*) and a modern layer strain. Appl Anim Behav Sci 74:1–14.
933 https://doi.org/10.1016/S0168-1591(01)00156-3

934 Schütz KE, Jensen P (2001) Effects of Resource Allocation on Behavioural Strategies: A
935 Comparison of Red Junglefowl (*Gallus gallus*) and Two Domesticated Breeds of Poultry.
936 Ethology 107:753–765. https://doi.org/10.1046/j.1439-0310.2001.00703.x

937 Stancher G, Rugani R, Regolin L, Vallortigara G (2015) Numerical discrimination by frogs
938 (*Bombina orientalis*). Anim Cogn 18:219–229. https://doi.org/10.1007/s10071-014-0791-
939 7

940 Szabó E, Chiandetti C, Téglás E, et al (2022) Young domestic chicks spontaneously represent
941 the absence of objects. eLife 11:e67208. https://doi.org/10.7554/eLife.67208

942 Tsutsumi S, Ushitani T, Fujita K (2011) Arithmetic-Like Reasoning in Wild Vervet Monkeys: A
943 Demonstration of Cost-Benefit Calculation in Foraging. Int J Zool 2011:806589.
944 https://doi.org/10.1155/2011/806589

945 Väisänen J, Håkansson J, Jensen P (2005) Social interactions in Red Junglefowl (*Gallus gallus*)
946 and White Leghorn layers in stable groups and after re-grouping. Br Poult Sci 46:156–
947 168. https://doi.org/10.1080/00071660500062638

948 Venkatachalam B, Biswa BB, Nagayama H, Koide T (2024) Association of tameness and
949 sociability but no sign of domestication syndrome in mice selectively bred for active
950 tameness. Genes Brain Behav 23:e12887. https://doi.org/10.1111/gbb.12887

951 Wilkins AS, Wrangham RW, Fitch WT (2014) The "Domestication Syndrome" in Mammals: A
952 Unified Explanation Based on Neural Crest Cell Behavior and Genetics. Genetics
953 197:795–808. https://doi.org/10.1534/genetics.114.165423

954 Wynn K (1992) Addition and subtraction by human infants. Nature 358:749–750.
955 https://doi.org/10.1038/358749a0

956 Wynn K (1998) An evolved capacity for number. In: The evolution of mind. Oxford University
957 Press, New York, NY, US, pp 107–126

958 Wynn K, Bloom P (1992) The origins of psychological axioms of arithmetic and geometry. Mind
959 Lang 7:409–416. https://doi.org/10.1111/j.1468-0017.1992.tb00313.x

960

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterialarithmetictaskdraft20261027.docx