UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Psicologia Generale

Scuola di dottorato di ricerca in: Scienze Psicologiche

Ciclo XXXI

New paradigms to study numerical cognition in animals

Direttore della scuola di dottorato: Ch.mo Prof Giovanni Galfano

Supervisore: Ch.mo Prof Angelo Bisazza

Co-Supervisore: Ch.mo Prof Marco Dadda

Dottorando: Elia Gatto

This thesis is dedicated to Maria, Gino, Savino, Luca, and, especially, Maria Pia.

New paradigms to study numerical cognition in animals.

Abstract

The mathematical abilities of animals have long fascinated scientists and the general public alike. However, it is only in the last two decades that researchers have systematically investigated this issue. Numerical abilities are widespread among mammals, birds and fish but little is known about their occurrence in other taxa. They play important functions in several ecological contexts, such as foraging, mate choice, and social interaction. The actual mechanisms underlying number sense in animals and the upper limits of these abilities are still a matter of debate. Conflicting results often emerge when comparing evidence from different species and studies of the same species. It is unclear to what extent this is due to differences in the complexity of the nervous system, to ecology, to phylogeny, or to other factors. Some have argued that intra- and inter-species differences may be due to differences in the used methodology. This is frequent in human literature, but the issue has received little attention in animals. To unravel these issues, we need to expand the number of studied species by including species that belong to a less studied taxa. In addition, there is a need for studies that directly compare different methodologies within the same laboratory.

In this thesis, I focus on the two latter issues. In two studies, I investigated the effect of the used methodology on the assessment of numerical abilities in a small teleost fish, *Poecilia reticulata*.

In the first study, I adopted an operant conditioning procedure to train fish to solve a numerical discrimination task using computer-generated stimuli. With this procedure, fish showed very low numerical discrimination performance, considerably lower than that reported in previous studies carried out in the same laboratory, in which fish had to solve a

similar numerical task in a more naturalistic setting. In a series of subsequent experiments, I attempted to identify the key factors that determine the different outcome of the two methods. Results only partly answer the question and it is possible that there are non-cognitive factors (i.e. side bias due to cerebral lateralization) that explain the observed difference.

Previous studies on guppies showed greater numerical acuity when tested with operant conditioning procedures than when tested with spontaneous preference methods, such as testing the preference for the larger of two social groups (shoal choice test). This difference could be due either to the different accuracy of the two methodologies or to within-species variation in acuity related to the examined context (foraging vs social). In the second study, I set up a modified version of the shoal choice test in order to eliminate some of the limitations of the method that are thought to determine a large imprecision in the measurement. With this new method, guppies were able to discriminate 4 vs. 5 companions, the same limit of numerical discrimination determined with operant conditioning procedures.

The last two studies focused on studying numerical cognition in two new species, an amphibian (*Hyla intermedia*) and an arthropod (*Acheta domesticus*) in a new context, shelter choice.

Both species showed a preference for the larger cluster of shelters and, in a series of subsequent experiments, I tried to determine the mechanism involved in such discrimination. *H. intermedia* appears to possess a genuine numerical system, being able to select the most numerous set even after I experimentally controlled the perceptual continuous variables, such as cumulative surface area and convex hull, that co-vary with number. Results suggest that *A. domesticus* have a less sophisticated system and make adaptive decisions, relying on continuous quantity as a proxy of number.

Table of contents

1. Introduction: current knowledge on numerical cognition in human and non-human animals

1.1 Numerical cognition

Quantity such as amount, space, time, and number, is one of the basic types of information that reflect a fundamental attribute of the environment (Dehaene & Brannon, 2011). Number, of course, is not a tangible attribute like other physical characteristic of stimuli, such as colour, shape, and size. The concept of number is an abstract and flexible property that allows defining an attribute of objects in many different contexts (Wiese, 2003). Wiese classified number assignments into three categories. The first category is ‘cardinal number’, which refer to the quantitative assignments for the elements of a set and concern the questions ‘How many?’ or ‘How much?’ Humans use cardinal numbers to identify the quantitative information of a set of elements. The category ‘ordinal number’ refers to the ranking position of a single element in a sequence and concerns the question ‘Which one?’ For example, humans use ordinal numbers when we identify the position of a football team during the season or the position of a runner during a marathon. The third category is the “nominal number”, which assign numbers as a proper name for objects within a set, such as “bus#3”, “salad#4”, “cr#7”. Nominal numbers are exclusively verbal and are found only in humans endowed with language (Wiese, 2003).

Historically, the capacity to perceive quantitative information, especially numbers, was considered a unique trait of human species. In particular, numerical cognition was thought to be developed as a consequence of language and verbal skills, as numerical competence in humans involves the use of syntax and symbols (Dehaene, 1992). Although certain numerical competence, such as mathematical calculus, are only performed by humans (but see Matsuzawa, 2009; Pepperberg, 2006), other numerical abilities do not require the use of symbolic representation. For example, it is possible to discriminate the relative numerosity in two distinct sets to understand that a group of 5 apples is larger than a group of 2 biscuits. The predisposition to easily and spontaneously use quantitative information without symbolic representation is defined as 'Number Sense' (Dehaene, 2001), and it has been reported in human infants and other animals (reviewed in Feigenson et al., 2004; Nieder, 2005).

Comparative psychology is concerned with the study of mechanisms possessed by animals to solve widespread ecological situations (Shettleworth, 2010). In particular, comparative psychologists have shown a growing interest in studying numerical cognition, a distinct discipline that aims to understand the origin and development of numerical cognition (Zorzi & Testolin, 2018).

1.1.1 Numerical cognition in vertebrates

The capacity to process quantitative information to make suitable decisions, or simply numerical cognition, may confer an advantage for the individual from an adaptive point of view (Gallistel, 1990). Contexts such as foraging, navigation, and avoidance of predators require individuals to make adaptive decisions by simultaneously processing different sources of quantitative information. Because human and non-human species have historically faced similar situations, evolutionary pressures have led to the development of comparable competencies to process quantitative information across different species (Gallistel, 1990; Haun et al., 2011).

During the past few decades, this hypothesis has been confirmed by a growing body of evidence that suggests a wide variety of non-human species share a non-symbolic system to process quantitative information (reviewed in Dehaene & Brannon, 2011; Gallistel, 1990; Mou, 2014). When we talk about numerical cognition in non-human species, we are referring to the non-symbolic system of processing quantitative information (e.g., ‘Number Sense’). The development of techniques traditionally used with infants have been widely used with non-human primates and birds in comparative investigations of numerical abilities (Boysen, 1997; Roberts et al., 2002; West & Young, 2002).

A first, reasonable question concerning the study of numerical competence in animals is why should number matter for non-human species? Numbers are salient, and, like other attributes of elements, they can be perceived, acquired, and processed to make adaptive decisions. For instance, animals maximize the quantity of food in relation to the cost of finding a food source by selecting the more profitable option (Hauser et al., 2000; Pantaleeva et al., 2013). Social species reduce the risk of predation by forming larger groups of conspecifics (Cresswell, 1994) or adapting different strategies in competitive behaviour in relation to the number of potential competitors (Bensom-Amram et al., 2011; Bonanni et al., 201; McComb et al., 1994). Opportunistic species use estimated fitness information (e.g., clutch size) of resident species as an indicator of quality of territory (Loukola et al., 2013), while parasitised species use numerical information (e.g., egg recognition) to reduce the cost of brood parasitism (Lyon, 2003). This selective pressure may explain why a large variety of species have developed numerical competence. However, it is important to note that in certain contexts numerical cognition is critical and has a direct impact on individual fitness (e.g., survival rate, offspring), while there are other situations in which it is unnecessary to discriminate between quantities. For example, a group of 4 apples provides a higher benefit than a group of 3 apples, whereas the discrimination between 13 and 12 apples does not afford the same advantages.

A second question is whether other animals perceive and process numbers in the same ways as humans. Considering performance patterns in numerical judgments, comparative studies have reported that human infants and non-human animals demonstrate similar performances when facing comparable numerical tasks (reviewed in Feigenson et al., 2004). In particular, different species showed different accuracy when performing quantitative discrimination between the range of small numbers (e.g., range 1–4 items) and larger numbers (e.g., > 4 items). These findings led several authors to suggest that vertebrates share two distinct non-symbolic quantificational systems (Beran, 2008; Feigenson, et al., 2004). The first system is the object-file system (OFS), which was developed from two theories on visual attention in adults: the FINST model (Pylyshyn & Storm, 1988) and the object file model (Kahneman et al., 1992). The OFS is based on a primitive attentive mechanism that permits identification of a limited number of elements as a unique symbol stored in the working memory (Trick & Pylyshyn, 1993). OFS accuracy is independent of numerical ratio, rather it depends on intrinsic memory limits and/or attentional constraints, which in human adults appears to be about four elements (Burr et al., 2010; Hyde, 2011; Pylyshyn, 2001). The second system is the approximate number system (ANS), which is involved in the estimation process of large numerosity. The ANS does not depend on an absolute limit of elements, but the accuracy of estimate numerosity is dependent on the numerical ratio between the elements. The imprecision of ANS is due to the distance effect (the discrimination 5 vs. 10 items is easier than 5 vs. 6 items) and the size effect (the discrimination 2 vs. 3 is easier than 10 vs. 12 items; Gallistel & Gelman, 1992). Recently studies have raised the issues on the existence of a single and unique system (ANS) that is activated in both small and large numerical estimation. According to this view, the ANS may be involved to estimate both small and large numerosity in a similar way because the ability to precisely estimate the difference between quantities depends on their ratio rather than their absolute difference (e.g., the difficulty of discriminating 3 vs. 4 objects is equal to 9 vs. 12 objects, because the ratio is 0.75 for both discriminations; Dehaene & Brannon, 2011; Gallistel, 1990; Halberda & Odic, 2015). However, several studies on mammals, birds, and fish have indirectly supported the existence of two system involved on quantity estimation (reviewed in Agrillo et al., 2015). These contrasting results may be explained by the possibility that specific condition may activate different mechanism of quantity estimation. It has recently been proposed that the ANS may be involved to estimate small numerosity when some specific properties of stimuli are modified (e.g., sensory modality, physical properties of stimuli) and the task requires the abstraction of the concepts of numbers (reviewed in VanMarle, 2015). The OTS fails to represent “amodal” stimuli because the mechanism involved visual attention based on the physical property of the stimuli. For this hypothesis, the activation of ANS instead of OTS in the range of small numbers depends on the experimental methods that are adopted (Chapter 1.2.3). Another possibility concerns different species have developed differently quantity estimation mechanisms in relation ecological pressures, such as other different cognitive abilities (Shettleworth, 2010). For example, in a dangerous situation, preyed species need to rapidly perform quantity discrimination when searching for the larger group; while predators do not necessarily need the same accuracy on quantity estimation. For this view, evolutionary pressures might play an important role in developing different forms of quantity discrimination across species (Chapter 1.3).

The third question addresses the origin of the mechanism to process quantitative information. Several authors have suggested that non-symbolic numerical systems may have a long evolutionary history (Beran, 2008; Feigenson, et al., 2004). It is not yet known whether the quantitative representation process originated in an ancient ancestor common to most distant phylogenetic species (homology) or whether a similar system has evolved independently in different species to afford similar ecological pressure (homoplasy). Despite differences in neural system architecture, humans, non-human primates, and birds actively display neural networks when performing quantitative discrimination tasks (reviewed in Nieder, 2018). Jones and colleagues (2014) attempted to solve this issue by performing a comparative study on human and non-human primates. The importance of this work is that the authors employed a common set of task parameters and stimuli to estimated numerical competence among species. The authors found an equivalent performance among all non-human primate species, which was qualitatively similar to that of humans. Neurobiological and behavioural findings from other comparative research (see Agrillo et al., 2012b; Hanus & Call, 2007; Jordan & Brannon, 2006) seem to validate the hypothesis of a homoplasy origin of numerical systems. The difference showed among species may be due to environmental factors and individual experience (Shettleworth, 2010). Despite the large amount of published data on numerical cognition in primates and birds, the origin of this system is still unclear. Several researchers have proposed investigating this capacity in less studied species to partially answer this question (e.g., Agrillo & Bisazza, 2018; Skorupski et al., 2018).

1.1.2 Numerical cognition in lower vertebrates

The growing interest in numerical competence in animals has focused only on mammals and birds. To gain a clear understanding of the origin and evolution of numerical cognition, it is essential to expand our knowledge of less studied species.

During the last decade, small fish such as the zebrafish (*Danio rerio*) and the guppy (*Poecilia reticulata*) have rapidly become popular animal models in many research fields, such as biomedicine, biomarker for assessing environmental quality, and comparative psychology. Factors that may have contributed to the recent boom in fish research include the reduced cost of experiments, the facilitate of maintaining fish in semi-wild conditions, and the velocity of reproduction compared to other vertebrates (Kalueff et al., 2016). Fish are interesting animal models because, despite placing at the root of the vertebrate phylogenetic tree, they share many physiological and cognitive traits similar to other vertebrates. Fish offer possibilities for investigating the evolution of cognitive functions shared among vertebrates (Brown et al., 2008; Bshary et al., 2002; Bshary & Brown, 2014).

The ecological pressure described for land vertebrates (Chapter 1.1) may lead fish to develop a similar ability to choose the larger/smaller quantity in different contexts. For example, fish show a spontaneous ability to select larger amounts when presenting a dichotomous choice between food patches (Lucon-Xiccato et al., 2015), a preference for larger groups of conspecifics when avoiding predators (Gómez-Laplaza & Gerlai, 2011), and improving mating opportunities (Agrillo et al., 2008a). Despite the smaller brains of fish, laboratory and fields studies have shown that the numerical cognition of fish is equal to that of many other vertebrates (reviewed in Agrillo & Bisazza, 2018).

In a spontaneous forced-choice discrimination task adapted from procedures developed for human infants and monkeys to assess numerical competence, red-backed salamanders (*Plethodon cinereus*) discriminated and chose the larger group of live prey up to 2 vs. 3 discrimination (Uller et al., 2003). Further experiments showed that *P. cinereus* performed similarly when the stimuli were live prey video or static images, suggesting that this species possess a stimulus-independent system for processing quantitative information (Krusche et al., 2010; Uller et al., 2003). A recent study by Stancher and colleagues (2015) investigated the spontaneous preference of oriental fire-bellied toads (*Bombina orientalis*) for different amounts of prey by presenting discrimination in the range of both small and large numerosity. Results suggested that *B. orientalis* possess two different systems for processing small and larger numerosity, as reported for other vertebrates. However, the evidence of quantity discrimination ability in amphibians is limited only in the context of prey selection. Few studies have suggested that amphibians may use quantitative information during social communication such as female mate choice (Arntzen, 1999) and male calling (Gerhardt et al., 2000; Rose, 2018), but no study has directly investigated their numerical competence in contexts other than prey selection.

1.1.3 Numerical cognition in invertebrate

In recent years, research on invertebrate cognition has raised growing interest among comparative psychologists. Historically, invertebrates were considered as automats that simply interact with their environment by innate reflexes, but recent behavioural and theoretical studies show that invertebrates, especially insects, are able to acquire, store, and process different source of environmental information (Chittka & Niven, 2009).

Authors	**Phylum**	**Class**	**Order**	**Species**	**Paradigm**
Wittinger, M., Wehner, R., & Wolf, H. (2006)	Arthropoda	Insecta	Hymenoptera	*Cataglyphis fortis*	Spontaneous preference
Wittinger, M., Wehner, R., & Wolf, H. (2007)	Arthropoda	Insecta	Hymenoptera	*Cataglyphis fortis*	Spontaneous preference
Reznikova, Z.I., & Ryabko, Y. (1993)	Arthropoda	Insecta	Hymenoptera	*Formica polyctena*	Spontaneous preference
Reznikova, Z.I., & Ryabko, Y. (1994)	Arthropoda	Insecta	Hymenoptera	*Formica polyctena*	Spontaneous preference
Reznikova, Z.I., & Ryabko, Y. (1996)	Arthropoda	Insecta	Hymenoptera	*Formica polyctena*	Spontaneous preference
Reznikova, Z.I., & Ryabko, Y. (2000)	Arthropoda	Insecta	Hymenoptera	*Formica polyctena*	Spontaneous preference
Tanner, C. (2006)	Arthropoda	Insecta	Hymenoptera	*Formica xerophila*	Spontaneous preference
Cronin, A. (2014)	Arthropoda	Insecta	Hymenoptera	*Myrmecina nipponica*	Spontaneous preference
Franks, N., Dornhaus, A., Metherell, B., Nelson, T., Lanfear, S.A.J, & Symes, W. (2006)	Arthropoda	Insecta	Hymenoptera	*Temnothorax albipenni*	Spontaneous preference
Leppik, E.E. (1953)	Arthropoda	Insecta	Hymenoptera	*Apis mellifera*	Spontaneous preference
Chittka, L., & Geiger, K. (1995a)	Arthropoda	Insecta	Hymenoptera	*Apis mellifera*	Spontaneous preference
Chittka, L., & Geiger, K. (1995b)	Arthropoda	Insecta	Hymenoptera	*Apis mellifera*	Spontaneous preference
Chittka, L., Geiger, K., & Kunze, J. (1995)	Arthropoda	Insecta	Hymenoptera	*Apis mellifera*	Spontaneous preference
Dacke, M., & Srinivasan, M. (2008)	Arthropoda	Insecta	Hymenoptera	*Apis mellifera*	Discrimination learning procedure
Gross H., Pahl, M., Si, A., Zhu, H., Tautz, J., & Zhang, S. (2009)	Arthropoda	Insecta	Hymenoptera	*Apis mellifera*	Discrimination learning procedure
Menzel, R., Fuchs, J., Nadles, L., Weiss, B., Kumbishinski, N., Adebiyi, D., .., & Greggers, U. (2010)	Arthropoda	Insecta	Hymenoptera	*Apis mellifera*	Spontaneous preference
Howard, S.R., Avarguès-Weber, A., Garcia, J.E., Greentree, A.D., & Dyer, A.G. (2018)	Arthropoda	Insecta	Hymenoptera	*Apis mellifera*	Discrimination learning procedure
Bar-shai, N., Keasar, T., & Shmida, A. (2011)	Arthropoda	Insecta	Hymenoptera	*Bumblebees*	Spontaneous preference
Mirwan, H., & Kevan, P. (2015)	Arthropoda	Insecta	Hymenoptera	*Bumblebees*	Spontaneous preference
Hemptinne, J.L., Dixon, A.F.G., & Coffin, J. (1992)	Arthropoda	Insecta	Coleoptera	*Adalia bipunctata*	Spontaneous preference
Carazo, P., Font, E., Forteza-Behrendt, E., & Desfilis, E. (2009)	Arthropoda	Insecta	Coleoptera	*Tenebrio molitor*	Spontaneous preference
Carazo, P., Fernandez-Perea, R., & Font, E. (2012)	Arthropoda	Insecta	Coleoptera	*Tenebrio molitor*	Spontaneous preference
Nelson, X., & Jackson, R. (2012)	Arthropoda	Arachnida	Araneae	*Portia Africana*	Spontaneous preference
Cross, F., & Jackson, R. (2017)	Arthropoda	Arachnida	Araneae	*Portia Africana*	Spontaneous preference
Rodriguez, R., Briceno, R.D., Briceno-Aguilar, E., & Hobel, G., (2015)	Arthropoda	Arachnida	Araneae	*Nephila clavipes*	Spontaneous preference
Yang, T., & Chiao, C. (2016)	Mollusca	Cephalopoda	Sepiida	*Sepia pharaonis*	Spontaneous preference

Tab 1. Current state of art of the studies where is investigated the numerical cognition in invertebrate

Invertebrates and vertebrates face similar ecological situations where numerical competence may provide a remarkable adaptive benefit to the individual by maximizing their fitness by discriminating and choosing the best available option among different in their environment. Numerical cognition research on invertebrates has mainly investigated the spontaneous preference for quantity (reviewed in Pahl et al., 2013; Tab. 1).

Bees and ants exhibit proto-counting abilities during navigation based on the number of overcome landmarkers and the covered distance from a food source (Chittka & Geiger 1995a; Wittlinger et al., 2006). Social and non-social species, such as ants, beetles, ladybirds, and spiders, assess the number of potential competitors to make proper decisions when searching for food sources, avoiding unequal conflicts, and choosing potential mates (Carazo et al., 2012; Hemptinne et al., 1992; Nelson & Jackson, 2012; Tanner, 2006). Recently, some studies have developed trained procedures to assay numerical competence in some social hymenopteran species. Bees are able to discriminate up to 3 vs. 4 items and understand the ‘concept of zero’ under control conditions (Gross et al., 2009; Howard et al., 2018).

The findings that invertebrates possess numerical competence comparable to vertebrates are surprising given the relatively small brain of invertebrates. However, theoretical and neuroactivity experiments have shown that high cognitive functions, such as attention, required a limited number of neurons. Invertebrates are able to solve complex tasks and perform as well as vertebrates without a high number of neurons (human: 85 billion; chimpanzee: 22 billion; grey parrot: 1.57 billion; mouse: 0.07 billion; zebrafish: 0.01 billion; bees: 0.001 billion; “List of animals by number of neurons”, n.d.). Indeed, a large brain seems not to be required for making quantitative discriminations, such as other cognitive functions (reviewed in Chittka & Niven, 2009). Neuroimaging studies have shown that non-symbolic numerical representation does not require a dedicated cortical module (Nieder, 2009, 2018), but may rather be an inherent process of visual perception that summarises sensory input and maintains a representation of the object in the working memory (Burr & Ross, 2008; Stoianov & Zorzi, 2012). This hypothesis is supported by theoretical models (Dehaene & Changeux, 1993; Stoianov & Zorzi 2012) and studies on infants (Cantlon et al., 2010; Hannula-Sormunen, 2015).

1.2 Methods adopted in studying numerical cognition in nonhuman animals

Two main procedures have been adopted for studying numerical cognition in animals, and they are based on ones traditionally developed for human infants: spontaneous forced-choice test and discrimination learning procedures. The former is generally conducted in nature or laboratory by presenting a choice in a biologically relevant situation; whereas discrimination learning procedures require sophisticated apparatus, prolonged training, and manipulation of stimuli, which are possible only under standard conditions in the laboratory.

1.2.1 Spontaneous forced-choice test

The spontaneous forced-choice test implies a spontaneous discrimination between biologically relevant stimuli (e.g., food). Animals are tested in their natural environment or in the laboratory under semi-natural conditions. This procedure allows to study the inherent numerical competence of a species in a situation that mimics its natural habitat. Observing how an animal behaves freely in critical conditions makes it possible to advance hypotheses on the adaptive importance of numerical cognition (Fig. 1).

Fig. 1. A New Zealand robin (*Petroica longipes*) receives a food reward (mealworm) in a spontaneous forced-choice discrimination. Photo is provided by Garland and colleagues, 2012.

Most of the studies that have employed this procedure focused on the animal's tendency to choose the set containing the larger number of elements, such as food patches or social companions. In consideration of basic foraging behaviour, discrimination and selection of the largest amount of food will maximize intake and subsequently increase survival rates (Stephens & Krebs, 1986).

Food is one of the most studied stimuli used to assess whether animal spontaneously show quantitative discrimination abilities. Animals spend most of their time in searching for food, so the ability to estimate food quantities plays a major role in foraging activity. In the classical paradigm, animals are presented with two amounts of food differing by the number or size of elements. Despite few cases (e.g., Pantaleeva et al., 2013), in general, animals maximize the quantity of food and the cost of searching by choosing the larger amount. Evidence from primates and other mammals, birds, reptiles, amphibians, fish, and insects has shown that the spontaneous ability to discriminate quantities is widespread among distant species (e.g., Baker et al., 2012; Call 2000; Garland et al., 2012; Lucon-Xiccato et al., 2015; Miletto Petrazzini & Wynne, 2016; Nelson & Jackson 2012; Uller et al., 2003).

A second common scenario exploits the social behaviour of many species, in particular fish. Social fish show a natural tendency to form larger groups under risk of predation and other stressful situations. The apparatus typically consists of an unfamiliar environment with no shelter in which fish try to join one of two available conspecific groups differed by the numbers of fish (review in Agrillo & Bisazza 2014). Indeed, in large groups, the individual risk of being caught is drastically diluted (Krause & Ruxton, 2002). By adopting this procedure, mosquitofish (*Gambusia affinis*), angelfish (*Pterophyllum scalare*), and the guppy (*Poecilia reticulata*) can discriminate between shoals up to 3 vs. 4 fish (Agrillo et al., 2008b, 2012; Gómez-Laplaza & Gerlai 2011).

1.2.2 Discrimination Learning Procedures

Discrimination learning procedures are based on the concept of operant conditioning already studied by Thorndike at the beginning of the twentieth century. During this procedure, subjects make a voluntary choice between discriminative stimuli to obtain a reward (reviewed in Hall, 2002).

To study numerical competence, animals are commonly faced with a dichotomous choice between two different quantities of items in which only one is reinforced. Subjects must associate the reinforced quantity with the reward. In this training phase, neutral or non-biologically relevant stimuli are commonly used, such as three-dimensional objects or two-dimensional figures displayed on a monitor. Subjects have to achieve a predeterminate criterion to pass at the test phase, generally 70% of correct choices in two consecutive trials or sessions (e.g., Bisazza et al., 2014b; Emmerton & Renner, 2009; Howard et al., 2018; Jordan et al., 2008; Vonk & Beran, 2012). In the test phase, subjects are faced with novel numerical discrimination. The experimenter evaluates the learning ability of animals by recording their behaviours in a pre-determined number of trials (Fig. 2).

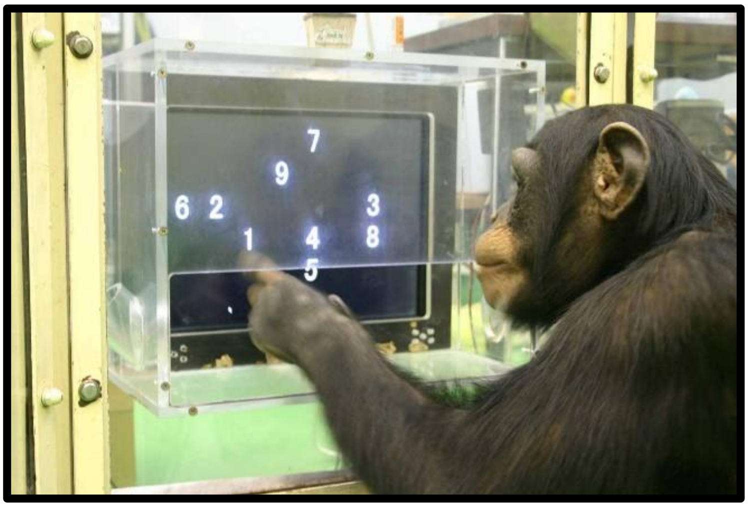


Fig. 2. A chimpanzee (*Pan troglodytes*) solves a numerical ordering task. Photo is provided by Prof Matsuzawa.

One of the main advantages of discrimination learning procedures is the administration of elevated number of trials for subjects. Studies on mammals and birds include thousands of reinforced trials with remarkable performance. Macaques exhibit an accuracy of 90% when discriminating quantities differing by a 0.90 numerical ratio (Beran, 2008); chimpanzees show 80% accuracy in performing quantity discrimination differed by a numerical ratio of 0.80 (Tomonaga, 2008); and birds' accuracy is 75% when the numerical ratio is 0.75 (Emmerton & Delius, 1993). The elevated number of trials and the standardized conditions permit comparing these performances to those exhibited by humans tested in comparable conditions (Cantlon & Brannon, 2006, 2007).

A second advantage of this procedure is the manipulation and control of presented stimuli. In fact, subjects may use non-numerical information or a pattern recognition mechanism to discriminate between quantities. Non-numerical information, such as convex hull, cumulative surface area, or amount of movement, is a perceptual aspect of objects that positively covaries with numerical information (Brannon et al., 2006). For example, a group of 10 apples has more elements and encompasses more cumulative area than a group of 3 same-size apples. Similarly, a larger group of flies produces more movement than a smaller group. Regardless of the exact perceptual cue involved, there are interspecific differences between which cues are spontaneously preferred in quantity discrimination. In tamarins, for example, the tendency to rely on non-numerical information prevails over numerical information when it is needed for mathematical calculation in the presences of food sources (Stevens et al., 2007), while mosquitofish primarily rely on numerical information rather than total area when discriminating between social groups by presenting one-by-one companions (Dadda et al., 2009). Hence, the use of artificial stimuli permits to manipulate their physical attributes (e.g., density, convex hull) in order to reduce the confound of non-numerical information. If animals solved quantitative discrimination tasks when stimuli were controlled for non-numerical information, one would be confident on the fact that subjects were make a choice relying on numerical information.

1.2.3 Pros and cons of the procedures adopted on studying numerical cognition in animals

Despite the benefits of both procedures described above, there are differences between the two regarding methodological aspects and ecological values of the cognitive function assessment (reviewed in Agrillo & Bisazza, 2014). The discrimination learning procedure affords testing subjects with thousands of trials and with a wide array of stimulus controls, while the spontaneous forced-choice test uses biologically relevant stimuli to investigate the natural tendency of individuals. First, there is a clear trade-off between stringent experimental control and ecological validity of numerical competence assessments. In spontaneous forced-choice tests, it is generally accepted that animals exhibit their natural decisioning behaviour when faced with biologically relevant stimuli. Results obtained with spontaneous forced-choice tests led us to understand in which contexts animals might have an advantage by possessing numerical abilities. On the other hand, discrimination learning procedures involve extensive training to provide direct evidence of numerical cognition under standardized conditions, which can be hardly compared to those in natural contexts. For example, numerical and non-numerical information are both presented in nature, and animals may use both to make quantity discrimination decisions.

Two important factors to consider are the motivations and individual preferences of subjects. Individual variability in numerical competence is marked in adult humans (reviewed in Piazza & Izard, 2009) but marginally considered in other species concerning numerical competence (e.g., Bisazza et al., 2014b; Jones et al., 2014). In discrimination learning procedures, animals are kept motivated by administering small quantities of food rewards when they choose the correct stimuli. The amount of food in each trial is not enough to fully satisfy subjects thus keeping animals motivated to work to obtain more rewards. Indeed, motivation is more difficult to control in spontaneous forced-choice test. For example, animals are highly motivated to choose the larger food quantity when the relative difference between numerosity is high (e.g., 2 vs. 10 items). On the contrary, when the relative numerical difference between sets (e.g., 2 vs. 3, or 8 vs. 9 items), animals may exhibit individual preferences or less motivation to select the larger quantity because two available options are equally favourable. The lack of discrimination exhibited by animals during training procedures is likely to reflect a true limit of their ability to process numerical information, whereas we cannot exclude the effects of internal factors (e.g., motivation, salience of stimuli, individual differences) when individuals spontaneously discriminate between quantities.

In discrimination learning procedures, few individuals may be enough to make proper conclusions about numerical competence of that species through the assumption that individuals of that species are equipped with the same neurocognitive systems able to potentially process quantitative information (Pepperberg & Brezinski, 1991). In spontaneous forced-choice tests, it is necessary to test a large number of subjects to study the broader effects across the population. It is commonly considered the first choice as an individual spontaneous preference for the stimulus. If the effect (e.g., preference for the larger stimulus) is large in the population, we will find a significant difference even with a reduced sample size. When the effect is very small, we will hardly find a significant preference by using a smaller sample size. As a first point, animals need time to look and choose between two objects (Chittka et al., 2009). The first choice could be random because individuals need time to explore each stimulus before making a decision. Second, subjects may have a motor asymmetry bias that conflicts with spontaneous preference for a stimulus. Cerebral and resulting motor lateralization has been well described in all species (reviewed in Andrew & Rogers, 2002). These factors may strongly influence the assessments of individual numerical competence when adopting a spontaneous forced-choice test; whereas extensive training may reduce the influence of individual and methodological bias. However, less studied is the possible effect of training on spontaneous preference. It is possible that individual performances observed during these kinds of procedures are an artefact of inherent abilities of species (Shettleworth, 2010). As Gallistel suggested (1990), none of the numerical cognition studies that adapted discrimination learning procedures describes the spontaneous behaviour of numerical ability because they all involved extensive training.

1.3 Two open issues in numerical cognition studies of non-human animals

Despite the general system for processing quantitative information appears similar among species, different results emerge when we compare performance on numerical tasks even within the same species. To explain the potential inconsistency in the literature, some researchers have recently suggested that the influence of procedure on the assessment of animal numerical competence is one of the main problems (reviewed in Agrillo & Bisazza, 2014). For example, some studies on human infants considered the time spent by subjects looking at matched trials (Izard et al., 2009; Jordan & Brannon, 2006; Starkey et al., 1990), but in other studies the time spent on looking at mismatched trials (Feigenson, 2011; Féron et al., 2006; Kobayashi et al., 2004; 2005). Intra- and inter-species differences emerge~~s~~ when comparing numerical accuracy among non-human animals. For example, the African grey parrot (*Psittacus erithacus*) spontaneously selects the larger amount of food up to 2 vs. 3 items (Al Aïn et al., 2009), but the same species shows higher numerical competence up to 5 vs. 6 items after extensive training (Pepperberg, 2006). Similar differences are reported in primates and other mammals, birds, and fish (reviewed in Agrillo & Bisazza 2014).

Differences in cognitive abilities are usually ascribed to ecological differences among species (Deaner et al., 2007; Lefebvre et al., 2004; Sherry, 2006). However, this variation could also be attributable to the procedure when it exploits the natural predisposition of a species to solve an ecological problem. Ecological pressure may favour species-specific predisposition to solve ecological problems. Therefore, adaptive specialization may involve morphological, physiological, behavioural, and cognitive traits that have evolved for ecological specialization (Shettleworth, 1972). It is well known that food-storing birds, such as parids and corvids, have a high capacity to memorize the location of thousands of hidden food sources, but, in other contexts, their mnemonic performance is similar to that of other birds (Shettleworth & Hampton, 1998). Food-storing birds possess a relatively larger hippocampus than other birds. This region is important for spatial memory in birds, as the homologous in mammals (Shettleworth & Hampton, 1998). Considering land and water species, individuals may evolve inherent predispositions for solving specific situations that rarely occur in other environments. Salwiczkek and colleagues (2012) reported that a cleaner fish (*Labroides dimidiatus*) outperformed primates in an experimental setting adopted to the fish's ecology but not the primates'. In a recent study that exploited its natural tendency of guppy on pecking and moving small objects on the bottom (Bisazza et al., 2014b; Fig. 3), the guppy achieved all discriminations up to 4 vs. 5 sets of items (e.g., yellow discs that mimic flowers or dead leaves), a much higher performance than previously reported in fish and higher than that other shown by some mammals and birds (reviewed in Agrillo & Bisazza, 2018).

Fig. 3. A female *Poecilia reticulata* attempts to solve a numerical task (e.g., 3 vs. 4 items). Under the yellow disc, a food reward is hidden. Photo is provided by Bisazza and colleagues, 2014.

On the contrary, the relaxation of ecological pressures may lead species to decrease the investment of some unnecessary cognitive functions. For example, Bisazza and colleagues (2014c) investigated numerical cognition in blind fish *Phreatichthys andruzzii*. This is a cave-dwelling species that separated from others teleost fish in the phreatic layer of the Somalia desert, a homogeneous environment with an absence of natural predators and a scarcity of food resources, approximately 2 million years ago. Bisazza and colleagues have trained fish to discriminate between two groups of three-dimensional sticks with different number of items to receive food rewards. *P. andruzzii* learned to discriminate between groups of objects up to 2 vs. 4, showing a numerical acuity lower than that reported in other teleost fish tested with visual stimuli.

A second important issue, already mentioned above (Chapter 1.1), is the complete lack of knowledge on numerical cognition in some taxa. Most numerical cognition studies have focused only on mammals, birds, and recently fish, and in only one or a few contexts for each species. Among vertebrates, few data are available for reptiles and amphibians. Indeed, there are fewer studies on numerical cognition in invertebrates despite that, for some taxa, such as arthropod and mollusc, the existence of sophisticated cognitive functions resembling those of vertebrates have been reported. Numerical cognition has been traditionally considered as a high cognitive function, limited to species that possess a relatively larger brain. Contrary to this view, studies on artificial neural networks have indicated that numerical competence does not require a dedicated cortical region or specialized neuron type for processing quantitative information (Dehaene & Changeux, 1993; Stoianov & Zorzi, 2012), rather this process requires a neural circuit activated by threshold-sensory mechanisms and the maintenance of object representation in the working memory (Feigenson et al., 2004). The preference for investigating numerical cognition in non-human primates and birds seems mainly related to the impossibility of finding a current procedure to study this capacity in other species. Since the beginning of this century, few studies have focused on numerical cognition on fish. After a decade, the number of published studied on this topic has dramatically increased (e.g., Agrillo & Bisazza, 2018), showing that a species (e.g., fish) that has historically been considered a lower-functioning vertebrate can solve complex tasks as well as other vertebrate species (Brown et al., 2008: Bshary et al., 2002; Bshary & Brown, 2014) More attention to experimental procedures and a better understanding of animal behaviour can help us clarify conflicting results (Thornton et al., 2012).

1.4 Aim of thesis

In the last section, I discussed two issues that afflict the comparative study of numerical cognition, and that have been recently highlighted by several authors in this discipline (Agrillo & Bisazza, 2018; Skorupski et al., 2018).

The first issue is that, in many cases, there is evidence, especially in the human literature, that the methods used for studying numerical abilities may influence the results. Moreover, the effect of the procedures may have a significant impact when comparing numerical abilities among species. Since different methods must often be used to study different species, especially if they are phylogenetically distant, when cognitive differences are found, it becomes difficult to understand whether these are due to real differences between species in cognitive abilities or to the fact that different methods have been applied.

In the first part of this thesis, I investigated how methodological changes influence the assessments of cognitive functions within the same species, using a teleost fish, the guppy (*Poecilia reticulata*), as a model.

In the first study, I used an operant conditioning chamber and computer-generated stimuli to train fish to a series of numerical discriminations. This is a replication of a recent study (Bisazza et al., 2014b) in which, training subjects in a naturalistic setting, guppies were able to discriminate 4 vs. 5 items. Since in my experiment, *P. reticulata* showed a much lower performance compared to previous study, I performed four additional experiments in which I considered different factors that might be responsible of the low performance shown with my setting. In particular, I manipulated the duration of the decision time, the type of stimuli (bi- vs. three-dimensional), the length of training, and the training schedule (introduction of a correction procedure).

In the second study, I set up and tested a modified version of a common paradigm used to study spontaneous quantity discrimination in fish, the shoal choice test. This paradigm exploits the spontaneous tendency of social fish to join the larger social group in a potentially dangerous context. In *P. reticulata*, numerical acuity is higher when assessed with an operant conditioning procedure than with spontaneous forced-choice tests, such as the shoal choice test (reviewed in Agrillo & Bisazza, 2018). It is not clear if this occurs because numerical acuity varies across contexts (e.g. foraging vs social behaviour) or if simply because spontaneous forced-choice tests are a less accurate method to study numerical cognition (Agrillo & Bisazza, 2014). I introduced three modifications of the apparatus and the procedure of the shoal choice test in the attempt to overcome some limitations that may determine significant imprecision in the measurement.

The second issue concerns the fact that our knowledge on numerical cognition is limited to a few taxa, namely mammals (chiefly primates) birds, and fish. Very few studies have investigated for example quantity discrimination in reptiles and amphibians although these taxa occupied a relevant position on the phylogenetical tree and such knowledge would be crucial for understanding the evolution of numerical cognition in vertebrates. We also have limited information about the presence of numerical abilities among invertebrates. To date this issue has been investigated only in a handful of species and there is evidence of a true numerical system only for one species, the honeybee. Studying numerical ability in arthropods and molluscan is particular interesting as numerical abilities very likely evolved independently in vertebrates and in these vertebrate taxa (Agrillo & Bisazza, 2018).

In the second part of this thesis, I investigated quantity discrimination in two less studied taxa, amphibians and an arthropod. In both studies, I investigated an ecological function, shelter choice, that had never been studied.

In the first study, I investigated whether the treefrog *Hyla intermedia* discriminate between two quantities when searching for potential refuges. In nature, this species showed an arboreal habit in order to escape from terrestrial predators. I replicated this situation in the laboratory by presenting a dichotomous choice between two sets comprising different numbers of vertical green bars (simulating grass clumps). In the second study, I exploited the scototactic reaction in the cricket *Acheta domesticus* to investigate the quantity discrimination ability of this species. The crickets were presented with a dichotomous choice between two sets of dark figures that simulated potential shelters. My hypothesis was that both species would prefer the larger quantity of stimuli, as they have a larger probability of finding suitable shelter.

The Species

Poecilia reticulata

Kingdom	Animalia
Phylum	Chordata
Class	Actinopterygii
Order	Cyprinodontiformes
Family	Poecillidae
Genus	*Poecilia*
Species	*P. reticulata*

Fig. 4. A male and female of an ornamental strain of *Poecilia reticulata*

Poecilia reticulata (Peters, 1859), also called the guppy from its discoverer (Robert J. Lechmere Guppy, 1836–1916), is a small, live-bearing fish native to Trinidad and Tobago and Venezuela. Currently, this species is widespread all over the world thanks to its high adaptability to many ecological conditions.

The guppy has small size, stretched shape, and elevated sexual dimorphism (Fig., 4). Males show a complex colour pattern polymorphism. The intensity of colour is related to environmental conditions (Houde, 1997), but the pattern is inherited from the father (Magurran, 2005). Females are larger than males but have a typical camouflage pattern to confound with the environment.

The popularity of this species is increasing as a new important animal model for studying sexual selection (Magurran, 2005), sociality (Brown & Laland, 2003), and, more recently, numerical cognition (Agrillo & Bisazza, 2018).

Hyla intermedia

Kingdom	Animalia
Phylum	Chordata
Class	Amphibia
Order	Anura
Famiglia	Hylidae
Genere	Hyla
Species	*H. imtermedia*

Fig. 5. An exemplar of *Hyla intermedia*

The Italian treefrog (*Hyla intermedia*; Boulenger, 1882) is species endemic to Italy. It is a colourful, small frog similar to the European treefrog (*Hyla arborea*). Its natural habitats are wetlands with easy access to water sources. During the breeding season, males gather near a pond and emit calls to attract female. Females release eggs on the vegetation near the pond and males externally fertilize them. A distinctive trait of this species is the adaptation for arboreal living. New metamorphosed juveniles leave the pond for vegetation (e.g., tree, brush) where they camouflage with their environment (Fig. 5).

Acheta domesticus

Kingdom	Animalia
Phylum	Arthropoda
Class	Insecta
Famiglia	Gryllidae
Genere	*Acheta*
Species	*A. domesticus*

Fig. 6. A male exemplar of *Acheta domesticus*

The house cricket (*Acheta domesticus*; Linnaeus, 1758) is native to Southwestern Asia although currently widespread all over the world. The natural habitat of this species is wet area, such as caves, burrows, and forests. Because of its attraction to heat, it is commonly found in urban areas. Females differ by presenting an ovipositor emerging from the rear. Dominant males establish territories near holes and attract female by producing a typical call (Fig. 6).

1.5 Ethical note

I adhered to the ASAB/ABS Guidelines for the Use of Animals in Research (see Buchanam et al., 2012). My experiments consisted of observations of animals' behaviour without manipulation. Subjects spontaneously participated in the experiments, otherwise they were substituted with new ones. The numbers of subjects participating in the experiments and the number of subjects that were substituted were reported for each experiment (see details below). At the end of the experiment, all subjects were released in tanks and terraria identical to the ones previously used for their maintenance. In discrimination learning procedures experiments (Chapter 3.1), fish were housed in a tank equipped with natural plants, bottom gravel, and social companions to minimize differences between maintenance tanks.

The experiments with vertebrate species (e.g., fish and treefrogs) complied with the laws of the country (Italy) in which they were performed (Decreto legislativo 4 marzo 2014, n. 26). The Ethical Committee of Università di Padova has approved the experimental procedures adopted for fish experiments (protocol n. 32/2015; Chapters 3 and 4) and for the treefrog experiments (protocol n. 388523; Chapter 5). For the invertebrate experiments (Chapters 6), the laws of my country (Italy) permit behavioural experiments on invertebrates without any approval by an ethical committee. However, all invertebrates were maintained and tested following the ASAB/ABS Guidelines for the Use of Animals in Research

.

2. Study 1: experimental setting affects the performance of guppies in a numerical discrimination task *

* The contents from this chapter have been published as Gatto, E., Lucon-Xiccato, T., Savaşçı, B. B., Dadda, M., Bisazza, A. (2017). Animal Cognition, 20: 187-198.

2.1 Abstract

A recent study found that guppies (*Poecilia reticulata*) can be trained to discriminate 4 vs. 5 objects, a numerical discrimination typically achieved only by some mammals and birds. In that study, guppies were required to discriminate between two patches of small objects on the bottom of the tank that they could remove to find a food reward. It is not clear whether this species possesses exceptional numerical accuracy compared with the other ectothermic vertebrates or whether its remarkable performance was due to a specific predisposition to discriminate between differences in the quality of patches while foraging. To disentangle these possibilities, we trained guppies to the same numerical discriminations with a more conventional two-choice discrimination task. Stimuli were sets of dots presented on a computer screen, and the subjects received a food reward upon approaching the set with the larger numerosity. Though the cognitive problem was identical in the two experiments, the change in the experimental setting led to a much poorer performance as most fish failed even the 2 vs. 3 discrimination. In four additional experiments, we varied the duration of the decision time, the type of stimuli, the length of training, and whether correction was allowed in order to identify the factors

responsible for the difference. None of these parameters succeeded in increasing the performance to the level of the previous study, although the group trained with three-dimensional stimuli learned the easiest numerical task. We suggest that the different results with the two experimental settings might be due to constraints on learning and that guppies might be prepared to accurately estimate patch quality during foraging but not to learn an abstract stimulus–reward association.

2.2 Introduction

Many vertebrates possess the ability to discriminate between sets of objects that differ in numerosity (reviewed in Agrillo & Bisazza, 2014). This ability may benefit several fitness-related activities, such as detecting the largest available food source (Normand et al., 2009), assessing the number of potential opponents in a contest (Benson-Amram et al., 2011), or choosing the largest social group to increase protection from predators (Hager & Helfman, 1991). Although the general mechanisms for numerical processing appear similar among vertebrates (reviewed in Feigenson et al., 2004), different species show different discrimination performance even within the same class. For example, apes can discriminate up to 9 vs. 10 items while dogs (*Canis lupus familiaris*) generally discriminate up to 2 vs. 3 items and domestic cats (*Felis silvestris catus*) can only discriminate up to 2 vs. 5 items (Hanus & Call, 2007; Bánszegi et al., 2016; Miletto Petrazzini & Wynne, 2016).

Inter-species differences in cognitive abilities are usually ascribed to differences in ecology or to the complexity of neural systems (Deaner et al., 2007; Lefebvre et al., 2004; Sherry, 2006). However, in many cases this variation could also be attributable to the different procedures that have been used in the different studies (Agrillo & Bisazza, 2014).

The effect of the procedure is also revealed by the fact that large between-studies differences in cognitive performance are sometimes observed even within the same species. For example, when presented with two sets of food items, the African grey parrot (*Psittacus erithacus*) selects the larger set in up to 2 vs. 3 items (Al Aïn et al., 2009), but with training this species is able to discriminate up to 5 vs. 6 items (Pepperberg, 2006).

A recent study on guppies (*Poecilia reticulata*) found that, when the length of the training is similar to that typically used in mammals and birds, this fish can discriminate up to 4 vs. 5 items (Bisazza et al., 2014b). Such numerical accuracy is higher than the maximum accuracy reported for many mammals and birds (reviewed in Agrillo & Bisazza, 2014). However, Bisazza and colleagues (2014b) used an experimental setting that deviated considerably from the setting of the classic two-choice discrimination task used with other species. The classic task generally consists in presenting two two-dimensional (2D) stimuli (e.g., sets of dots on a computer monitor) in front of the subject and releasing a food reward when it chooses the correct stimulus (Beran, 2008; Emmerton & Renner, 2006). Conversely, guppies were presented with two patches of yellow discs lying on the bottom of the test arena, each of which covered a small hole. The holes beneath the discs of the larger patch hid a food reward that the guppies could reach by pecking and dislodging the discs. This procedure was chosen because, in both the laboratory and the field, guppies spend much time pecking and moving small objects on the bottom (e.g., dead leaves and small stones) looking for hidden food (Dussault & Kramer, 1981; Rodd et al., 2002).

Exceptional cognitive abilities are usually associated with specific ecological adaptations such as homing in pigeons, food storing in parids and corvids, and vocal learning in parrots. For example, food-storing birds can memorize the location of thousands of food items, but in other contexts their memory performance is similar to that of other birds (Shettleworth & Hampton, 1998); among food-storing birds, performance in spatial tasks tends to positively correlate with dependence on cached food (Olson et al., 1995). Apart from these striking examples, biological predispositions to learn or to solve specific problems are extremely diffuse and can affect the comparison of cognitive abilities of the different species with laboratory tasks (Shettleworth, 1972; Hinde, 1973). For example, adopting a more ecologically relevant procedure dramatically affected the performance of two primate species in a complex learning task (Prétôt et al., 2016). Similarly, pigeons acquired matching- and oddity-to-sample tasks in less than 30 trials when the procedure exploited their natural foraging habit, but they required many more trials (from 1000 to 2500) to acquire the same task in a Skinner box (Wright & Delius, 1994). Inangas (*Galaxias maculatus*) were unable to associate predation risk with a specific place at a specific time of the day; however, they could promptly solve a similar time-place learning task to obtain a food reward (Reebs, 1999).

Because the study by Bisazza and colleagues (2014b) used a training method attuned to guppies' foraging behaviour, it is not clear whether guppies possess exceptional numerical abilities or if their remarkable performance was due to a specific predisposition to discriminate between differences in the quality of patches while foraging. In this study, we tried to disentangle these possibilities. In the first experiment, we trained guppies to select the larger quantity in the same numerical discriminations administered by Bisazza and colleagues (2014b), but we used a training setting more similar to that used in experiments with other vertebrates and which was previously used to study discrimination learning in zebrafish (Parker et al. 2012a, b; Proulx et al. 2014; Wang et al. 2015). The stimuli were sets of dots presented on a computer screen in two different chambers. We rewarded the guppies with food when they entered the correct chamber.

If guppies do possess numerical abilities equal or superior to many warm-blooded vertebrates, we expected the same numerical accuracy as Bisazza and colleagues (2014b), even with our training setting, similar to that used in experiments with other vertebrates. As we found a considerable difference between the performances in the two settings, we pursued four further experiments, varying the task parameters to identify the possible factors responsible for the difference.

2.3 Experiment 1

In experiment 1, we tested guppies on numerical discriminations of increasing difficulty starting from 2 vs. 3, as in the experiment of Bisazza and collaborators (2014b), but with a different training setting, a two-choice discrimination task devised for zebrafish (Parker et al., 2012a, b; Proulx et al., 2014; Wang et al., 2015). To keep the two experiments as similar as possible, we used subjects of same strain, sex and age, which were maintained under the same conditions until the experiment started.

2.3.1 Materials and methods

Subjects

The subjects were eight adult guppies of an ornamental strain (“snakeskin cobra green”) bred in our laboratory in the Department of General Psychology (University of Padova, Italy). Bisazza and colleagues (2014b) used only females in their experiment. In some contexts, guppies show sex differences in numerical discrimination abilities (Lucon-Xiccato et al. 2016); to avoid that these sex differences affected the comparison between our experiments and the previous study, here we used only females. The maintenance aquaria (150 L) had a gravel bottom, abundant natural and artificial plants, water filters, and 15-W fluorescent lamps (12h:12h light/dark photoperiod). We kept water temperature at 26 ± 1 °C and fed the fish with commercial food flakes (Fioccomix, Super Hi Group, Ovada, Italy) and *Artemia salina* nauplii three times per day.

Apparatus

The apparatus resembled the one used in previous work on fish (Parker et al., 2012a, b; Proulx et al., 2014; Wang et al., 2015; Fig. 7). In a 60 × 40 × 38 cm glass tank with gravel bottom and filled with 30 cm of water, we built a central compartment (60 × 16 cm) using green plastic. The two lateral compartments housed abundant natural vegetation and water filters to ensure that the water characteristics and odour were similar to the maintenance tanks. The central compartment also had two trapezoidal sectors made of transparent plastic, containing one immature guppy each as a social companion. We presented the stimuli at one end of the central compartment with an LCD monitor connected to a laptop. Here, we built two ‘choice chambers’ (8 × 9 cm) separated from the central compartment by means of a transparent guillotine door. Two pumps connected to the filters constantly drained water (4 litres/minute) from the choice chambers to the lateral compartments to remove the olfactory cues of the food reward. Two 15-W fluorescent lamps placed above the tank on the opposite side of the monitor illuminated the apparatus. Outside of the training sessions, we maintained the subjects in a 50 × 20 × 38 cm ‘home tank’ with immature guppies as social companions, vegetation, and water filters.

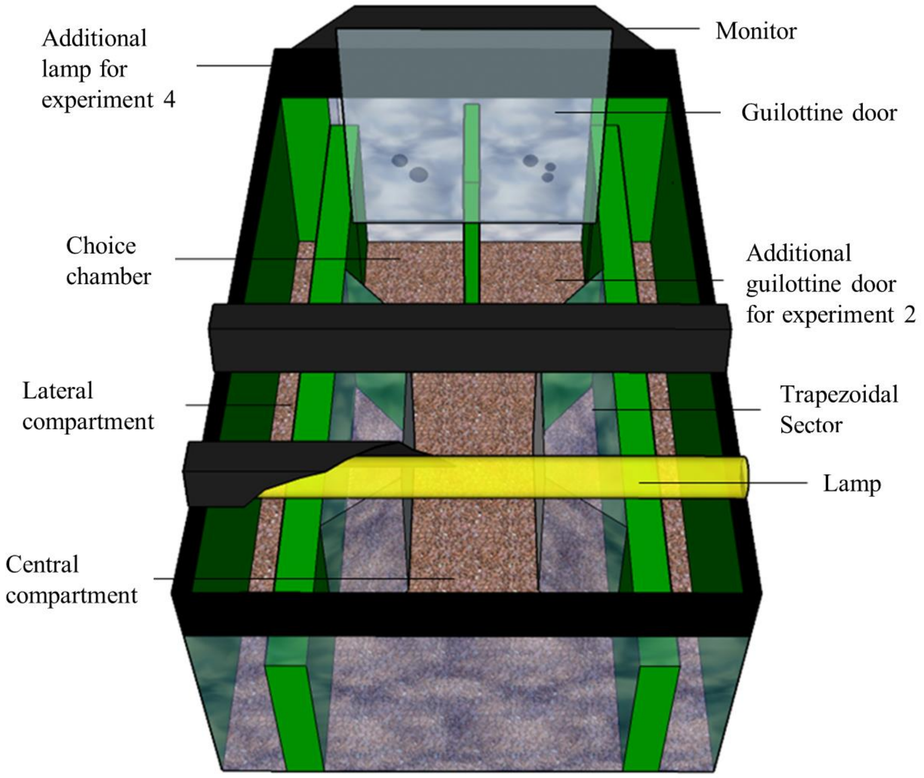


Fig. 7. Aerial view of the experimental apparatus. In experiment 2, an additional guillotine door (not drawn) was used to increase decision time; in experiment 4, an additional fluorescent lamp (not drawn) was placed above the choice chambers to light the 3D stimuli.

The stimuli were two sets of black dots on a white background made with Microsoft PowerPoint (Fig. 8a). During trials, we presented each set of dots in correspondence of one choice chamber. Two sets with a different number of objects may also differ by non-numerical attributes that covary with numerosity and fish might solve the discrimination exploiting these attributes. To prevent this confound, Bisazza and colleagues (2014b) controlled the stimuli for the three most important non-numerical attributes that covary with numerosity (cumulative surface area of the objects, overall space encompassed by the object arrays, and density of the objects) by using discs of different diameter and by displacing the discs on different arrays. We controlled our stimuli following the same scheme. We used dots with different diameter (range 0.3-0.5 cm) to control for the cumulative surface area: in one-third of trials, the ratio between the cumulative surface area of the smaller over the larger set was between 76 and 85%; in one-third, between 86 and 95%; and in one-third between 96 and 105%. Further, we varied the position of the dots in order to equate the overall space encompassed by the two arrays in half of the trials, and to equate the density of the dots in the other half of the trials.

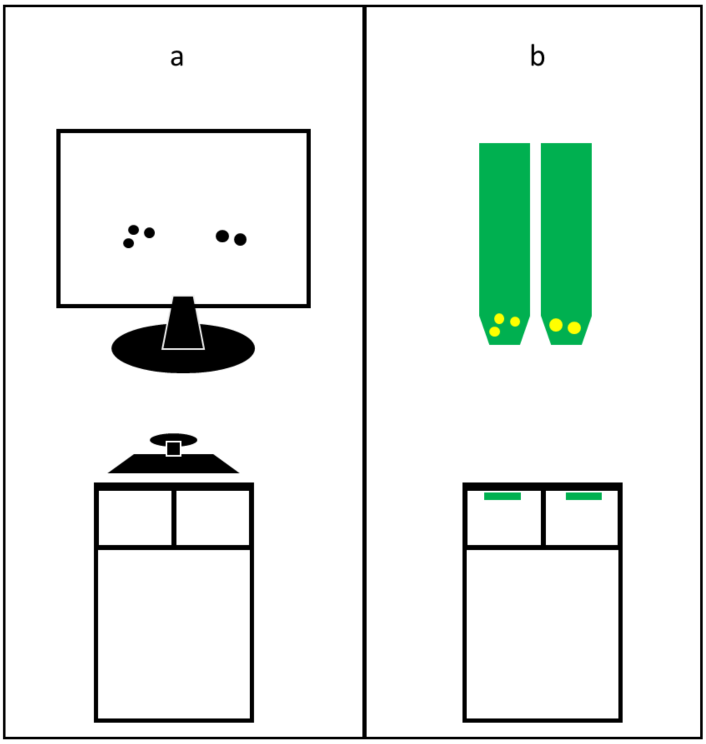


Fig. 8. Frontal view of the stimuli (above) and aerial view of their position during the trials (below). **a** In experiments 1, 2, 3, and 5, the stimuli were black dots on a white background projected on a LCD monitor; **b** in experiment 4, the stimuli were yellow discs fixed on green plastic panels

Familiarization with the apparatus and the procedure

The day before the experiment started, we randomly selected two female guppies from the maintenance tanks and moved them to the apparatus for habituation. In the strain of guppies that we tested, individual females show different fin coloration; we used this characteristic to recognize each individual subject. Only in this phase we provided three immature guppies in the central compartment to facilitate familiarization with the new environment (Lucon-Xiccato et al., 2015; Miletto Petrazzini et al., 2015). The computer monitor presented a white slide. When one subject swam inside a choice chamber, we supplied crumbled food flakes suspended in water with a Pasteur pipette. Usually, the remaining subject and the social companions rapidly joined to feed on the flakes. After both subjects finished the food and left the choice chamber, we closed the guillotine door and presented a black slide for 20 min. Then, the monitor presented again a white slide to attract the attention of the fish, and we repeated the procedure of above three times. As a consequence of this procedure, the subjects learned to seek food in the choice chambers when the white slide appeared. This resembled the situation of the training trials (see below) in which food could be obtained when the monitor presented the white slides with the stimuli. The subjects stayed all night in the apparatus, and the following morning we moved them to the home tank. The following two days, we moved one of the subjects to the apparatus and we administered 20 trials similar to those during the training phase. we then repeated the procedure for the second subject. In these trials, we presented 1 vs. 2 dots as stimuli on the monitor. If the subject entered the chamber corresponding to the set with the larger number of dots, we immediately rewarded it with food. In the initial 16 trials, if the subject chose the set with the smaller numerosity, we rewarded it when it switched to the other chamber. In the last four trials of this series, if the subject entered the chamber corresponding with the smaller set, we closed the guillotine door immediately after the fish exited from the first chosen chamber. Thus, the subject could not correct initial wrong choices, a situation similar to the one of the training phase. The following day, the training began.

Training

During the training period, we performed 12 daily trials, which were subdivided into two sessions of six trials each. The break between the two session was at least 2 h. In each session, we moved one subject from the home tank to the apparatus. The guillotine door was closed, and the monitor showed a black slide. Then, we presented the slide with the two sets of stimuli to be discriminated in the trial and raised the guillotine door to allow the subject to enter one of the choice chambers. If the subject chose the chamber associated with the larger numerosity, we administered a small quantity of food flakes as reward; if the subject chose the chamber with the smaller numerosity, we removed the stimuli and presented a black slide on the monitor. After the subject exited the choice chamber, we closed the guillotine door and started the following trial after 5 min. we varied the position (left or right) of the larger set of stimuli in each trial according to the pseudo-random scheme used by Bisazza and colleagues (2014b).

We initially presented the subjects with up to 120 trials (for a total of 10 days of training) with a 2 vs. 3 discrimination. Each day, we computed the number of correct choices in which the subject initially entered the chamber with the larger numerosity. Similarly to the previous study, we considered the subjects to achieve the discrimination if they reached the criterion of 75% correct responses in two consecutive days. We presented successful subjects with a more difficult discrimination (3 vs. 4). We also introduced a secondary, less stringent criterion: subjects that did not reach the primary learning criterion but reached a significant performance in the overall 10 days of training (chi-square test) were admitted to the more difficult discrimination task as well. we used the proportion of correct responses in the analysis (arcsine square root transformed to meet parametric assumptions; Sokal & Rohlf 1995). Statistical tests were two tailed; significance threshold was $\alpha = 0.05$, unless stated otherwise. In the text, the M ± SD percentage of correct choices are given.

2.3.2 Results

Only one out of eight subjects achieved the primary criterion in the 2 vs. 3 discrimination. This proportion of successful guppies was significantly smaller than the one observed in Bisazza and colleagues' (2014b) work, where seven out of eight subjects achieved the learning criterion in the 2 vs. 3 discrimination (Fisher's exact test: $P = 0.010$). Considering the secondary criterion, two additional subjects achieved the 2 vs. 3 discrimination (Tab. 2, column 1). In the 3 vs. 4 discrimination, none of these three subjects achieved the task through either the primary or the secondary criterion (62/120, 70/120, and 65/120 correct responses; all $P > 0.06$). An analysis with a chi-square test revealed that one subject expressed a significant preference bias for the left choice chamber.

Experiment 1 (*n*=8)	**Experiment 2 (*n*=6)**	**Experiment 3 (*n*=6)**	**Experiment 4 (*n*=6)**	**Experiment 5 (*n*=6)**
72/120, **P = 0.029**	64/120, P = 0.465	54/120, P = 0.273	64/120, P = 0.465	191/360, P = 0.246
65/120, P = 0.361	66/120, P = 0.273	53/120, P = 0.201	66/120, P = 0.273	190/360, P = 0.292
71/120, **P = 0.045**	67/120, P = 0.201	55/120, P = 0.361	**75/120,** **P = 0.006**	164/360, P = 0.091
60/120, P = 1.000	57/120, P = 0.584	69/120, P = 0.100	**63/96,** **P = 0.022 ***	181/360, P = 0.916
57/120, P = 0.584	52/120, P = 0.144	49/120, P = 0.045	68/120, P = 0.144	**199/360,** **P = 0.045**
29/48 **P = 0.149 ***	54/120, P = 0.273	61/120, P = 0.855	66/120, P = 0.273	175/360, P = 0.599
62/120 P = 0.715				
52/120 P = 0.144				

Tab. 2. Performance of individual guppies in the 2 vs. 3 discrimination. Number of correct choices and P-value calculated with chi-square tests (one degree of freedom). Each cell represents a different subject. Asterisks indicated subjects that achieved the primary criterion (75% correct responses in two consecutive days).

An overall analysis of the performance of the eight subjects failed to reveal a significant acquisition of the 2 vs. 3 discrimination (53.29 ± 6.30% correct responses; one-sample t test: $t_7 = 1.478$, $P = 0.183$). A linear mixed model (LMM) on the proportion of correct responses of each day of training (using the day as the fixed effect and the subject ID as a random factor) showed no significant improvement in the performance across the 10 days of the 2 vs. 3 discrimination ($F_{9,57} = 1.295$, $P = 0.260$).

2.3.3 Discussion

The guppies' performances in the numerical discrimination task were strikingly lower in our experiment compared to the ones observed in the previous study (Bisazza et al., 2014b). In our experiment, only one out of eight fish learned to discriminate up to 2 vs. 3 objects; in the previous study, seven out of eight fish did so. None of the subjects in our experiment achieved the 3 vs. 4 discrimination, while in the previous study, six out of eight fish were able to discriminate up to 3 vs. 4 objects and five out of eight fish were even able to discriminate up to 4 vs. 5 objects.

As in the two experiments we used subjects of the same sex, age and strain, we maintained them in identical conditions before the test, we followed the same training schedule and the cognitive problem was identical, the different outcome is most likely due to the differences in the experimental setting. In the previous study, the setting reproduced an ecologically relevant context in which guppies had to remove small objects on the bottom of a tank to find food concealed below (Rodd et al., 2002), while we used a training setting more similar to that used in experiments with other vertebrates. It is therefore possible that the improved performance of guppies in the former study was due to a predisposition to learn within the specific setting. Before considering this hypothesis, we need to evaluate whether other differences in task parameters might account for the different outcome of the two studies. We attempted to keep the conditions (characteristics of the subjects, numerical contrasts, control of non-numerical variables, number of trials) as similar as possible in the two studies, but there are at least three important differences.

The first difference is the decision time allowed to the subject. In the previous work, subjects were kept behind a transparent guillotine door to allow them to look at both stimuli for 10 s before making their choice. With our procedure, as soon as the stimuli appeared on the monitor, the subject swam to one of the two chambers and could no longer see the alternative stimulus. Consequently, the decision time was normally very short, around 1–2 s. In many animals, there is a trade-off between decision time and accuracy (Chittka et al., 2009), and it has been shown that a shorter decision time reduces the performance in quantity discrimination in fish (Krause et al., 1998). Thus, the reduced decision time for the guppies in experiment 1 might be insufficient for comparing the stimuli and making an appropriate choice. Moreover, fish often show high impulsivity (Danisman et al., 2010; Lucon-Xiccato & Bisazza, 2014; Miletto Petrazzini et al., 2015), which might cause guppies to choose the first stimulus they notice without comparing the two options.

The second important difference is the opportunity for fish to correct their first choice whenever it was wrong. Following the original procedure used with zebrafish (Parker et al. 2012a, b; Proulx et al., 2014; Wang et al., 2015), we did not allow correction, while Bisazza and colleagues (2014b) used a correction procedure. A non-correction procedure is believed to lead to greater accuracy (Parducci & Polt, 1958), but several studies have found that a correction method can facilitate learning, especially in the early phase (Hull & Spence, 1938).

Finally, the third major difference concerns the type of the stimuli used. In previous work, the stimuli were three-dimensional (3D) coloured objects, while here we used 2D images displayed on an LCD monitor. Virtual stimuli have often been used in fish research (for reviews see Fleishman & Endler, 2000, and Rosenthal, 2000), and it has been shown that fish respond properly to both simple, static stimuli (Agrillo et al., 2014) and complex, moving stimuli such as conspecifics (Nicoletto & Kodric-Brown, 1999). However, it is possible that 2D images are less salient than real objects leading guppies to experience difficulty with 2D abstract stimuli and show a reduced discrimination performance (Friedman et al., 2005; O'Hara et al., 2015; Truppa et al., 2015).

An additional potential problem is related to the number of training trials. Warm-blooded vertebrates can perform many training trials per day, and discrimination learning experiments with mammals and birds typically involve, overall, hundreds or even thousands of trials (Roberts & Mitchell, 1994; Jaakkola et al., 2005; Cantlon & Brannon, 2007). By contrast, many fish species, including guppies, can usually undergo a few reinforced trials per day (Agrillo et al., 2012a), and they become easily stressed in prolonged trainings or show reduced motivation to participate to the task (Sovrano et al., 2003; Lucon-Xiccato & Bisazza, 2016). For these reasons, discrimination learning experiments with fish normally have a pre-set limit in the number of training trials, typically one hundred or fewer (Bisazza et al., 2014b; Lucon-Xiccato & Bisazza, 2016). It is possible that with the setting used here, guppies require a longer training period to achieve the discrimination task. If two training settings lead to the same result but one requires a longer training, predetermining a small number of training trials can exaggerate the difference between the two methods.

In the following four experiments, we tested separately the influence of these four task parameters.

2.4 Experiment 2: Extension of decision time

Experiment 2 was similar to experiment 1, except that the fish were kept behind a transparent guillotine door and allowed to observe the stimuli for 20 s before they made their choice.

2.4.1 Materials and methods

Six naïve female guppies were used as subjects, which were different from the individuals used before. The only difference from the apparatus of experiment 1 is the presence of a second transparent guillotine door 25 cm away from the first door (Fig. 7). This second guillotine door formed a start box in the corridor in front of the choice chambers. At the beginning of each trial the experimenter waited until the subject spontaneously swam into the start box, and then closed the guillotine door. After 2 min, the stimuli were presented on the monitor. We released the fish from the start box after 20 s to allow it to observe the stimuli. We chose a decision time longer than the one adopted by Bisazza and colleagues (2014b) because the distance between the guillotine door and the stimuli was greater in our apparatus (34 cm) compared to the previous one (15 cm). Other details of the procedure were identical to experiment 1.

2.4.2 Results

None of the six subjects achieved the primary or the secondary criterion in the 2 vs. 3 discrimination (Tab. 1, column 2). This success rate was significantly lower than the one observed in the Bisazza and colleagues' (2014b) experiment (P = 0.005). One subject expressed a significant bias for the left choice chamber, and three subjects expressed a significant bias for the right choice chamber. An overall analysis of the performance of the six subjects failed to reveal a significant acquisition of the 2 vs. 3 discrimination (49.98 ± 5.40%; $t_5 = 0.008$, $P = 0.994$). There was no significant improvement in the performance across the 10 days of the 2 vs. 3 discrimination ($F_{9,45} = 0.759$, $P = 0.654$).

2.4.3 Discussion

In this experiment, we tested the possibility that the poor accuracy of guppies in experiment 1 was due to the reduced time allowed for observing the stimuli before making a decision. Indeed, in experiment 1, the fish observed the two stimuli for a few seconds before making the choice, and this may have contributed to the low performance observed (Chittka et al., 2009; Krause et al., 1998). Our attempt to extend the time before the decision yielded no evident improvement in learning compared to experiment 1. This result suggests that the greater accuracy shown in the previous study was not due to having forced the subjects to observe the task before the choice.

2.5 Experiment 3: correction procedure

This experiment was identical to experiment 1 except that we adopted a correction procedure.

2.5.1 Materials and methods

We observed six different female guppies in this experiment. The apparatus and the procedure were the same as in experiment 1. However, if the subject chose the wrong chamber, we allowed it to switch to the correct chamber to receive the food reward. As in previous experiments, we use the first choice to measure the performance.

2.5.2 Results

None of the six subjects achieved the primary or secondary criterion in the 2 vs. 3 discrimination (Tab. 1, column 3). This proportion of successful guppies was significantly lower (P = 0.005) than the one found by Bisazza and colleagues (2014b). One subject expressed a significant bias for the left choice chamber. An overall analysis of the performance of the six subjects failed to reveal a significant acquisition of the 2 vs. 3 discrimination ability (47.36 ± 5.93%; t_5 = 1.091, P = 0.325). There was no overall significant improvement in the performance across the 10 days of training ($F_{9,45}$ = 1.299, P = 0.264).

2.5.3 Discussion

Both correction and non-correction procedures are currently employed in discrimination learning experiments. Because the former procedure is expected to favour fast learning (Hull & Spence, 1938), we tested whether guppies' performance in a numerical discrimination task could be improved by adopting a correction procedure. In experiment 3, we found no evidence of this improved performance. The fish trained using the correction procedure performed as poorly as the fish in experiment 1, which were trained without a correction procedure. Even in this case, we found no evidence that discrimination significantly improved over time, at least within the 120 trials administered. This experiment suggests, therefore, that the correction procedure is not responsible for the differences between this study and the previous one in which the guppies performed better (Bisazza et al., 2014b).

2.6 Experiment 4: 3D stimuli

In this experiment, we used two sets of 3D objects as stimuli.

2.6.1 Materials and methods

As in the previous experiments, we tested six new female guppies which were experimentally naïve. The stimuli were sets of yellow plastic discs, the same as in Bisazza and colleagues' (2014b) experiment. Using a needle, we fixed the discs to two green plastic panels (Fig. 9b). During the trials, the experimenter inserted these panels in the two choice chambers to present the stimuli. Using discs with five different diameters (15, 16, 17, 18, and 19 mm) and varying their position, we corrected the sets separately for cumulative surface area, overall space encompassed, and density as described for experiment 1. we added an additional fluorescent lamp above the choice chambers to light the stimuli (Fig. 8). The guillotine door was covered with opaque plastic to prevent the fish from seeing the stimuli before the experimenter completed the insertion. The monitor was not used in this experiment. Other details of apparatus and procedures were identical to experiment 1.

2.6.2 Results

One out of six subjects achieved the criterion in the 2 vs. 3 discrimination. This success rate is, again, significantly lower (P = 0.026) than the one observed by Bisazza and colleagues (2014b). Considering the secondary criterion, an additional subject achieved the 2 vs. 3 discrimination (Tab. 1, column 4). In the following 3 vs. 4 discrimination, these two subjects did not reach the primary or the secondary learning criterion (58/120 and 60/120 correct responses; all $P > 0.7$). One subject expressed a significant bias for the left choice chamber, and three subjects expressed a significant bias for the right choice chamber. An overall analysis of the performance of the six subjects revealed a significant acquisition of the 2 vs. 3 discrimination (58.02 ± 4.90%; $t_5 = 3.964$, $P = 0.011$). Even the subjects that did not reach the individual criterion tend to perform above chance level. Indeed, the individual accuracy of these four subjects ranged from 53.33 % to 56.67 %, with an average that was significantly above chance (55.00 ± 1.36 %; $t_3 = 7.325$, $P = 0.005$).

Because this is the only experiment in this study in which the subjects, overall, significantly learned the task, we compared the performance of experiment 4 with the pooled data of experiments 1, 2, and 3 in which the stimuli were presented on a monitor (Levene test for homogeneity of variances between the two groups: $F_{1,24} = 0.910$, $P = 0.350$; Boneau, 1960; de Winter, 2013). We found that guppies performed significantly better in experiment 4 than in the other three experiments of this work (independent-samples *t* test: $t_{24} = 2.727$, $P = 0.012$).

2.6.3 Discussion

In experiment 4, we used two sets of 3D objects as stimuli to test the possibility that the reduced performance shown by guppies in experiment 1 was due to difficulty with the 2D abstract stimuli, which in some cases may negatively affect discrimination performance compared to the use of 3D stimuli (Friedman et al., 2005; O’Hara et al., 2015; Truppa et al., 2015). As we expected, the use of real 3D objects improved the numerical discrimination performance of the guppies. Although this improvement was limited and only one subject reached the primary learning criterion, it seems clear that the performance of most fish tend to be above chance. Similar effects have been reported also in other species. For example, Truppa and colleagues (2015) trained Capuchin monkeys (*Sapajus spp.*) to select either the larger between two objects or the larger between two images presented with a computer monitor, finding a better discrimination performance with the former stimuli. The format of the stimuli (2D vs. 3D) appear to be an important source of variation for animal discrimination performance. This could be explained by the fact that 3D stimuli resemble the stimuli that animals normally experience and are therefore more salient (Mustafar et al., 2015). The use of 2D images is often preferred in discrimination learning experiments because it offers a better control of stimulus features. However, our results, together with those of previous works, suggest that this choice might bear the cost of reduced discrimination performance.

Despite the increased accuracy due to the 3D stimuli, the overall performance of our guppies remained very low compared with previous study (Bisazza et al., 2014b). Only two fish learned to discriminate between 2 vs. 3 objects, and none of these learned to discriminate between 3 vs. 4 objects. Therefore, even if the format of the stimuli appears to be a parameter that affects the outcome of the experiment, it can only partially explain the difference between the performances observed in our experiment 1 and in the previous experiment.

2.7 Experiment 5: extended training

In this experiment, we tested whether guppies could learn the numerical discrimination if we extended the training from 10 days to 30 (360 trials overall).

2.7.1 Materials and methods

We used three subjects from experiment 2 and three subjects from experiment 3. None of these subjects reached the learning criterion within the predetermined 120 trials. The prolonged training was performed using the modality and the apparatus described for experiments 2 and 3, respectively.

2.7.2 Results

Only one of the six subjects (originally from experiment 3) reached the secondary learning criterion after extending the training from 120 trials to 360. An overall analysis of the performance of the six subjects did not reveal a significant acquisition of the 2 vs. 3 discrimination ability (50.83 ± 3.54%; $t_5 = 0.577$, $P = 0.589$; Tab. 1, column 5), nor was there a significant improvement in the performance over time (LMM: $F_{1,4} = 0.112$, $P = 0.754$).

2.7.3 Discussion

Increasing the number of training trials from 120 to 360 did not substantially improve the performance of the guppies. Only one fish succeeded in reaching the learning criterion after extending the training to three times the original length. The overall performance of the six fish remained non-significant. Therefore, the length of the training does not seem to be the key parameter in explaining why guppies obtained a much better record in the numerical task used in the previous study (Bisazza et al., 2014b) compared with the procedure used in this study.

2.8 General comparison between the experiments

We performed a tentative analysis to compare our experiments with the previous one (Bisazza et al., 2014b). Overall, the performance of the 26 guppies of our experiments pooled together was significantly worse than the one of the guppies in the previous work (2 vs. 3 discrimination; independent-samples *t* test: $t_{32} = 16.316$, $P < 0.001$). We also run a one-way ANOVA on the pooled data of the proportion of correct responses in the 2 vs. 3 discrimination with experiment as a five-levels factor (previous experiment, our experiment 1, 2, 3 and 4). We found a significant effect of the factor experiment ($F_{4,29} = 13.977$, $P < 0.001$). A Tukey's post hoc test revealed that the performance was significantly lower in our four experiments compared to the previous experiment (experiment 1: $P < 0.001$; experiment 2: $P < 0.001$; experiment 3: $P < 0.001$; experiment 4: $P = 0.018$; Fig. 9). We performed a further comparison using the pooled data of our experiments without the 10 subjects that showed a preference bias for one of the two chambers and we found a similar result. The performance in our experiments (52.00 ± 6.51%) was significantly lower compared to that of the previous experiment ($t_{22} = 5.803$, $P < 0.001$).

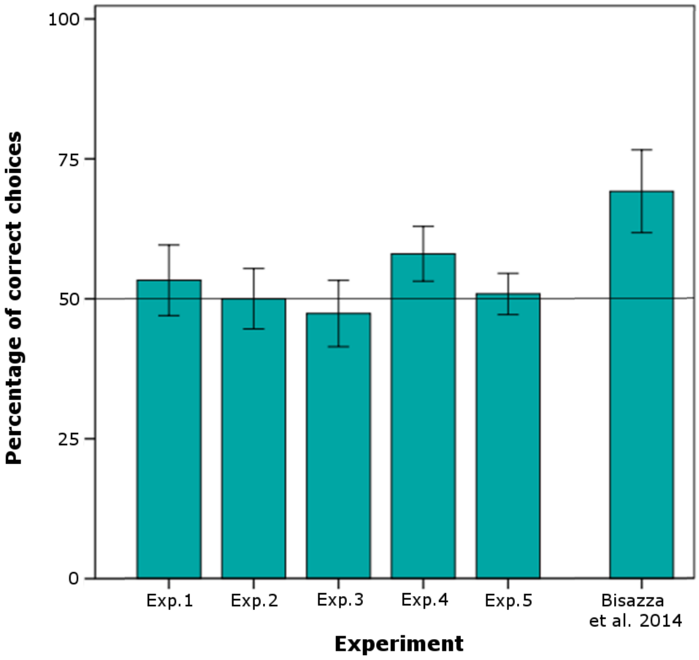


Fig. 9. Comparison of the five experiments of the present study and the results of Bisazza et al. (2014)'s study. Bars represent M ± SD percentage of correct choices in the 2 vs. 3 discrimination. Horizontal line represents chance performance

2.9 Conclusions

In this study, we provided evidence that the numerical acuity of guppies is radically different according to the training setting adopted. With the setting of the previous study by Bisazza and colleagues (2014b), five out of eight subjects achieved a 4 vs. 5 discrimination, while with the more conventional two-choice discrimination setting adopted in the present work, most of the guppies failed even an ‘easy’ 2 vs. 3 discrimination. To understand whether the task parameters could account for the difference, we varied the extension of decision time, whether or not correction was allowed, and the length of the training period. None of these variants enhanced guppies’ performance. By contrast, the use of 3D stimuli slightly improved guppies’ accuracy in numerical discrimination, but this improvement was still not enough to approximate the performance observed in the previous study. A possible limitation of our study is that we used a small number of subjects in each experiment (6 - 8 fish). Although our sample size was in line with the ones typically used in discrimination learning experiments (Schluessel et al., 2014; Prétôt et al., 2016), it is possible that this factor limited the likelihood to evidence subtle differences among the different experiments of this study. It is however unlikely that the sample size affected the general conclusion that the methods used in the two studies yield quite different results, as shown also by the overall analyses in which we pooled all our subjects (N = 26) to increase statistical power.

We detected another difference between the two studies. With the setting adopted in this work, some guppies (10 out of 26) exhibited a side bias, with four choosing the left chamber and six choosing the right chamber significantly more often. With the setting previously used by Bisazza and colleagues (2014b), none of the subject showed side bias. There are at least two possible explanations for the side bias observed in the present study. The first explanation is that the choice between the right and the left chambers was affected by the presence of spontaneous side preferences due to cerebral lateralization. Cerebral lateralization and the resulting motor and sensory asymmetries have been demonstrated in virtually all vertebrates (reviewed in Rogers & Andrew, 2002), including guppies and many other fish (reviewed in Bisazza & Brown, 2011). In most of the discrimination learning studies on fish showing lateralization, motorial asymmetries have not been reported to prevent learning (e.g., cichlids: Schluessel et al., 2014; guppies: Lucon-Xiccato & Bisazza, 2016; zebrafish: Colwill, 2005). However, in few cases, there is evidence that these motorial asymmetries may hamper learning in discrimination tasks. For example, rats showing spontaneous side bias in a T-maze showed poor learning, or even the absence of learning, when trained to choose the contralateral arm of the maze (Andrade et al., 2001). Concerning fish, Brown & Braithwaite (2005) found lower spatial learning performance in a lateralized population of the poeciliid *Brachyrhaphis episcopi* and suggested that spontaneous turning preferences interfered with the choice of the correct option. The second possible explanation is that the side bias was acquired during the training. Animals often develop side biases in discrimination tasks when the two stimuli are presented side-by-side in two fixed locations (rodents: Prusky et al., 2000; primates: Genty et al., 2004; birds: Smith & Zentall, 2016). A tendency to rely on the spatial position of the stimuli rather than on other cues may be present in guppies, too (Miletto Petrazzini et al., 2015) and might have hampered learning.

The absence of side bias in the study of Bisazza and colleagues (2014b) might be due to the fact that their setting minimized any precise left–right localization of the two stimuli by presenting them on a large plate on the bottom of the tank and changing their position every trial. Although side biases might be an important factor in our setting, it is unlikely that it alone explains the difference between the two studies. Indeed, excluding the guppies with side bias did not increase the level of the performance in the five experiments, which remained, on average, at a correct choice rate of 52%. In addition, a previous study employing a training procedure that prevented any left–right localization of the stimuli (Agrillo et al., 2014) did not obtain a numerical performance comparable to that of Bisazza and colleagues (2014b).

In sum, none of the five factors examined above appear to explain the huge difference in the performance between our study and the previous one. Such difference is likely to arise from Bisazza and colleagues' use of a setting that mimics guppies' natural foraging habitat. An interesting parallelism with our results comes from works on pigeons, which often forage by displacing leaves or gravel on the ground. Pigeons trained on concept learning using gravel of different colours hiding food –a setting that resembles their natural foraging habit– reached learning performances that were unattainable with conventional training in the Skinner box (Wright & Delius, 1994). Remarkably, in pigeons, the position of the stimuli (vertical vs. horizontal presentation) was not the key factor explaining the different performance with the two methods (Wright et al., 1988) and it would be interesting to investigate whether the same occurs in guppies.

A possible explanation of pigeons and guppy data calls into play innate predispositions to learn. Natural selection can favour the evolution of learning predispositions in one species to solve specific ecological problems (Shettleworth, 1972). As a consequence, the species shows enhanced cognitive performances in that specific context compared to other situations, and it may often outperform other species that lack the same ecological specialization. For example, food-storing birds can memorize the location of thousands of food items, but in other circumstances their mnemonic performance is similar to that other birds (Shettleworth & Hampton, 1998). Pigeons recognize rotated objects better than humans do, which might be due to the fact that objects' orientations change continuously during flight (Hollard & Delius, 1982). Based on observational studies (Dussault & Kramer, 1981; Rodd et al., 2002), we suggest that guppies might be predisposed to choose patches with a large number of objects that could potentially hide food, since the likelihood of finding food is increased in such patches.

A similar but distinct explanation for the difference between the two studies is that the naturalistic setting used by Bisazza and colleagues (2014b) measured the real cognitive repertoire of the species, while the artificial conditions of the present learning setting prevented somehow guppies from fully displaying their cognitive abilities. The experimental setting of this study resembled the classical setting of the Skinner box adopted in rodents, primates and birds (Skinner, 1933; Douglas & Whitty, 1941; Wright et al., 1988) and that is gaining popularity in fish (Gierszewski et al., 2013; Manabe et al., 2013; Proulx et al., 2014). The fish is placed in a bare environment and artificial bi-dimensional stimuli are displayed on a computer screen near the tank. To obtain the food reward, the fish is required to approach the side of the tank where the correct stimulus is presented. Since the position of the correct stimulus switch in the different trials, the fish is also required to inhibit the natural predisposition to associate food to the place where it has been previously found. All these features might favour comparison with works on other species, but might hamper learning. Both in nature and in the semi-natural conditions of the maintenance aquaria, guppies have hardly the opportunity to experience such conditions. Conversely, guppies spend considerable time moving objects in bottom to search for hidden food, a situation similar to the setting that favoured learning in guppies (Bisazza et al., 2014b). Both explanations -the naturalistic setting favours learning or the artificial setting hampers learning- appear plausible and our work cannot unfortunately provide direct support to either of them. Testing these ideas will require careful experimental investigation both in the laboratory and in the field.

The main message of this work is that the choice of the training setting sometimes affects the results of the experiment and, consequently, the estimation of the numerical discrimination abilities of the species. Although it was already known that different tests may lead to different results, such as in the case of spontaneous choice vs. training tests (Agrillo & Bisazza, 2014), our work is one of the very few studies showing that within the same type of test (training), different settings may also lead to quite different results. For fish, a similar conclusion has been recently drawn by a study of guppies' spontaneous preference for larger social groups (Chapter 4). In that study, guppies showed a higher accuracy in numerical discrimination when the experimental setting and the procedure were modified to remedy some flaws of the classical method, such as favouring the switching of the subject between the two shoals to compare them, and using stimulus fish that were well acclimated to the apparatus and thus behaved more naturally.

The problem of the experimental setting may be very relevant for studies that compare the performance of different species in the same task in order to understand the evolution of cognitive abilities. For example, one study found lower performance in numerical discrimination in zebrafish compared to that of four other fish species; however, a control experiment revealed that the zebrafish was simply less efficient in learning with the specific setting adopted in that study (Agrillo et al., 2012a). In another study, primates were thought to be unable to perform a task that cleaner fish (*Labroides dimidiatus*) perform quite well (Salwiczek et al., 2012). The setting adopted in that study matched the fish's ecology, but not that of the primates. The adoption of an experimental setting more appropriate for primates showed that they, too, were able to solve the task (Prétôt et al., 2016). To obtain an estimation of cognitive abilities that can be used in comparative research, it is therefore paramount to use a setting adequate to each species studied and, possibly, to perform multiple comparisons for the same cognitive ability.

3. Study 2: development and testing of a rapid method for measuring shoal size discrimination*

* The contents from this chapter have been published as Lucon-Xiccato, T., Dadda, M., Gatto, E., & Bisazza, A. (2017). Animal Cognition, 20:149-157.

3.1 Abstract

The shoal-choice test is a popular method to investigate quantity discrimination in social fish based on their spontaneous preference for the larger of two shoals. The shoal-choice test usually requires a long observation time (20–30 min), mainly because fish switch between the two shoals with low frequency, thus reducing the possibilities of comparison. This duration limits the use of the shoal-choice test for large-scale screenings. Here, we developed a new version of the shoal-choice test in which the subject was confined in a large transparent cylinder in the middle of the tank throughout the experiment to bound the minimum distance from the stimulus shoals and favour switching. We tested the new method by observing guppies (*Poecilia reticulata*) in a 4 vs. 6 fish discrimination (experiment 1). The new method allowed for a faster assessment of the preference for the larger shoal (<5 min), resulting in potential application for large population screenings. Guppies switched five times more frequently between the two shoals and remained close to the first chosen shoal ten times less compared to experiments with the old method. In experiment 2, we found that with the new method guppies were able to discriminate up to 5 vs. 6 fish, a discrimination that was not achieved with the classical method. This last result indicates that minor methodological

modifications can lead to very different findings in the same species and suggests that caution should be exercised when interpreting inter-specific differences in quantitative abilities.

3.2 Introduction

A diversity of fish species live most of their lives in groups. Several social groups often coexist in the same area, and they may differ from each other in the size or phenotype of the members (e.g., age, sex, colour, body size). Therefore, an individual fish has the option of choosing which group to join, a decision that may have profound effects on fitness (e.g., Lindström & Ranta, 1993; Agrillo et al., 2008a; Rosenthal & Ryan, 2005). Shoal size is one of the critical aspects that affect grouping decisions in fish, and it has been experimentally investigated in a large number of species. Because both dilution of individual predation risk and vigilance increase with increasing group size, larger groups are usually safer than smaller groups (Krause & Ruxton, 2002). As a consequence, under perceived threat, prey fish often show a preference for the largest available group (Magurran & Pitcher, 1987; Hager & Helfman, 1991; Krause et al., 1998; Hoare et al., 2004).

Preference for the larger shoal has been recently used by cognitive ethologists as one of the main tools to study numerical abilities in fish (e.g., Agrillo & Dadda, 2007; Gómez-Laplaza & Gerlai, 2011; Gómez-Laplaza & Gerlai, 2012; Bisazza et al., 2014a). Due to the practical difficulties of studying this behaviour in the wild, most experiments have measured social preferences in the laboratory using the three-chamber binary-choice test. Two groups of stimulus fish that differ in numerosity are placed in the two external chambers, and the subject is placed in the central chamber, and is free to swim. Because the central chamber is unfamiliar and potentially unsafe, the subject is expected to show antipredator behaviour and to attempt to reach the larger shoal. Shoaling preference is assessed by measuring the relative time spent by the subject in close proximity to each shoal, usually by analysing the video recording of the experiment.

With this procedure, it has been found that *Girardinus falcatus*, *Pimephales promelas* and *Pterophyllum scalare* can discriminate between shoals formed by 1 vs. 2 and 2 vs. 3 fish (Hager & Helfman, 1991; Agrillo & Dadda, 2007; Gómez-Laplaza & Gerlai, 2011). Guppies, *Poecilia reticulata*, and mosquitofish, *Gambusia holbrooki*, can distinguish subtler shoal-size differences, up to 3 vs. 4 (Agrillo et al., 2008b; Agrillo et al., 2012b). Discrimination of shoal size occurs even when the access to non-numerical cues (such as density, cumulative surface, and amount of movement) is experimentally prevented (Dadda et al., 2009), although the presence of these cues has been proven to affect discrimination performance in diverse species (e.g., Gómez-Laplaza & Gerlai, 2012).

Compared with the other tests commonly used to study numerical abilities in fish, the shoal-choice test is relatively rapid to execute (Agrillo & Bisazza, 2014). However, the time necessary to test one subject (usually 20 to 30 min; Gómez-Laplaza, 2006; Agrillo et al., 2007) and to score its performance from the recordings, makes it difficult to use the shoal-choice test for large-scale screenings (Patton & Zon, 2001) such as those required to study the genetic bases of quantification mechanisms or population differences. The length of the test might be problematic for two further reasons. The prolonged exposure to a situation perceived as dangerous is likely to cause considerable stress to the fish (Chandroo et al., 2004). Then, in natural situations, discrimination of the larger shoal usually occurs very rapidly (Krause et al., 1998), and it is possible that, with long tests, other factors affect the choice of the subject.

One of the reasons for extending observation in laboratory experiments is that the shoal-choice setting only approximately reproduces the natural situation. This may cause a large imprecision in the measure, especially at the beginning of the test, that the experimenters try to circumvent by extending the length of the observation. One of the main issues observed at the beginning of the test is the low frequency of switching between the two shoals, which arguably reduces the possibility to compare the two options and assigns a large weight to the first choice made by the subject. For example, two recent shoal-choice experiments found an average of 5.94 switches between the two shoals in 15 min and 4.23 switches in 8 min, respectively (Agrillo et al., 2012b; Lucon-Xiccato et al., 2016). Another study found an average of 5.29 min spent near the first shoal chosen before the first switch (Lucon-Xiccato & Dadda, 2017). One of the causes for the reduced switching frequency at the beginning of the test appears to be that, after it is released in the middle of the tank and took the first shoaling decision, the subject is very close to one of the two shoals and is in a relatively safe location. Even if the subject now perceives that the farther shoal is larger, it needs to cross a long distance alone to reach the better shoal. Both guppies and three-spined sticklebacks, *Gasterosteus aculeatus*, have been experimentally shown to prefer the near shoal, even if the far shoal is larger (Tegeder & Krause, 1995; Mühlhoff et al., 2011).

Here, we aimed to develop a new version of the shoal-choice test, which allows a rapid assessment of shoal size discrimination ability. In particular, we hypothesised that time of permanence near one school decreases and selectivity increases by preventing the subject from approaching the stimulus shoal too closely, and we tested if, using this expedient, the time required to assess the performance would be reduced. To bound the minimum distance from stimulus shoals, we confined the subject in a large transparent cylinder in the middle of the tank, and we took the time spent attempting to join each shoal as a measure of preference. This solution has already been used in mate choice studies to prevent the subject from approaching the stimulus fish too closely (Plath et al., 2008). we also used two novel features adopted in two recent shoal-choice experiments (Dadda et al., 2015; Lucon-Xiccato et al., 2016). To deal with the problem of pheromone released by previously tested individuals, the subject compartment received a continuous water supply from a large undisturbed population of guppies kept in seminatural conditions. To obtain well acclimated stimulus fish we permanently housed the stimulus shoals in their compartments.

In the first experiment, we measured the ability of guppies to discriminate 4 vs. 6 fish with our new method, focussing especially on the initial phase of the test. We chose the 4 vs. 6 contrast because it is challenging for guppies. Two studies found guppies were able to discriminate 4 vs. 6 fish (Dadda et al., 2015; Lucon-Xiccato et al., 2016), but two other studies reported that guppies failed to discriminate this contrast (Agrillo et al., 2012b; Bisazza et al., 2014a).

We also compared the results of experiment 1 with those obtained from three previous studies on the same guppy population that adopted two different shoal-choice methods (Agrillo et al., 2012b; Bisazza et al., 2014a; Lucon-Xiccato et al., 2016). Two early studies (Agrillo et al., 2012b; Bisazza et al., 2014a) adopted a classical shoal-choice procedure. The focal fish was released in the middle of the tank and two shoals were confined in the two lateral sectors. The third work (Lucon-Xiccato et al., 2016) was aimed at studying sex differences in shoal size discrimination, and used two of the innovations described in the present work: the water flow from a tank with a large guppy population and the stimuli permanently housed in the stimulus tank. However, in the work by Lucon-Xiccato and colleagues the subject was not confined in a large cylinder throughout all the test; the subject was kept for 2 min in a small cylinder in the middle of the tank to observe the two shoals and then released in the tank to choose the preferred option.

In a second experiment, we looked for the upper limit of shoal-size discrimination with the new method by observing guppies in more subtle discriminations (4 vs. 5 and 5 vs. 6 fish).

3.3 Materials and methods

3.3.1 Experiment 1

Subjects

The subjects were 32 adult females bred in 400-L holding tanks made of grey plastic in our laboratory at Dipartimento di Psicologia Generale (Università di Padova, Italy). v used 20 additional females of the same population as stimulus. The guppies were descendants of wild guppies caught in a high predation-risk environment in the lower Tacarigua River in Trinidad. At the time of the experiment, the subjects were 4 to 6 months old (standard length: approximately 22 mm). In the holding tanks, fish were kept in large mixed-sex groups (approximately 50 individuals per tank). We provided a gravel bottom and abundant natural plants to resemble natural conditions. We also provided water filters and 36-W fluorescent lamps (12h:12h light/dark photoperiod). Water temperature was kept constant at 26 ± 1 °C. Three times per day the guppies were fed commercial food flakes (Fioccomix, Super Hi Group, Ovada, Italy) and Artemia salina nauplii, which was the only interaction these guppies had with humans before the experiments. Each subject was used only once and, after the test, it was released in a tank and kept for breeding purposes.

Apparatus

The experimental apparatus was similar to the one adopted in previous shoal-choice experiments with guppies (Agrillo et al., 2012b; Bisazza et al., 2014a) and other species (Agrillo & Dadda, 2007; Gómez-Laplaza & Gerlai, 2011). It consisted of three adjacent glass tanks (60 × 40 × 35 cm) filled with 25 cm of water (Fig. 10). The central tank ('subject tank') housed the subject during the experiment, while the other two tanks ('stimulus tanks') housed the two stimulus shoals. The long walls and the bottom of the subject tank were covered with green plastic panels. The apparatus was placed in a completely dark room and the subject tank received light indirectly from the two stimulus compartments. Luminance measured with Gossen Mavo Monitor USB photometer was 0.85 cd/m2 in the cylinder and 105 cd/m2 in the front compartments of the stimulus tanks.

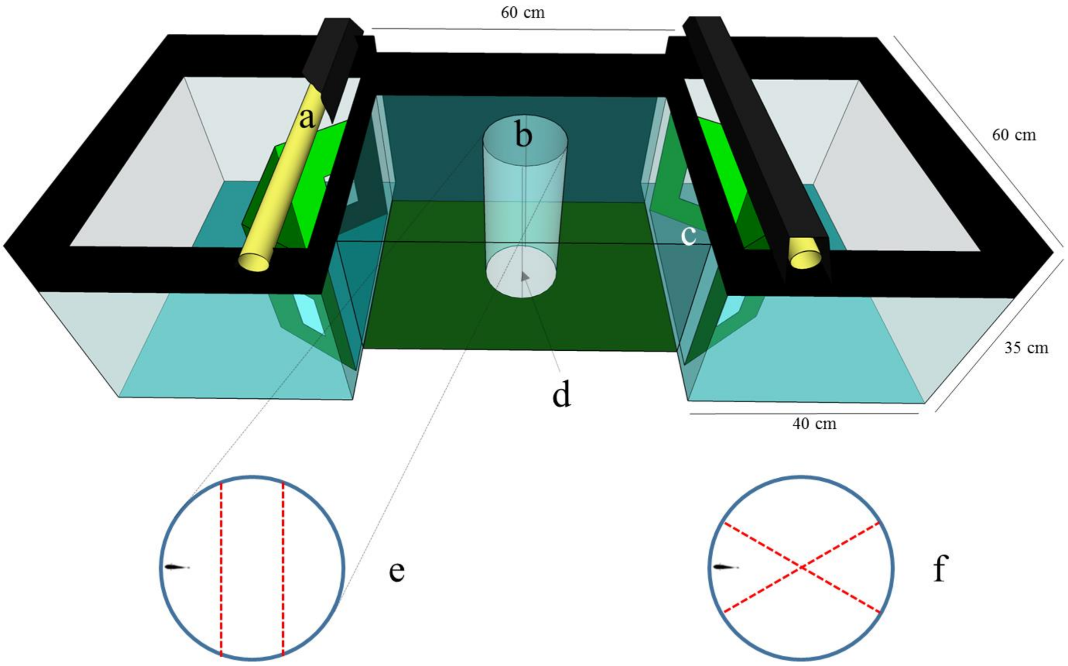


Fig. 10. Apparatus adopted in our study. The two lateral tanks housed the stimulus shoals. **a** Fluorescent lamp; **b** transparent cylinder that housed the subject; **c** front compartment to present the stimulus shoals to the subjects; **d** fissures to allow water flow into the cylinder; **e** aerial view of the cylinder with the three sectors used to measure the preference of the subject in the main analysis; **f** aerial view of the cylinder with the four sectors used to measure the preference of the subject based on its orientation.

Unlike previous studies, a transparent cylinder (height: 30 cm; diameter: 15 cm) limited the movements of the subject to the central part of the subject tank for the whole experiment and allowed the subject to reach a minimum distance of 22.5 cm from the stimulus tanks. The cylinder was made with two acetate sheets connected with paper clips. The connections were always kept on the lateral side so that they did not hamper the sight of the stimulus shoals. Under the cylinder, the bottom of the tank was made of white plastic to facilitate the tracking of the subject. To allow water flow into the cylinder, we left a 1-mm fissure in correspondence of the connections of the two acetate sheets by preventing the sheets to completely overlap (Fig. 10), and we also left a similar fissure between the cylinder and the bottom of the tank. Moreover, before each trial we lifted the cylinder to completely change the internal water.

Two additional features of our apparatus were based on two studies on shoal choice recently performed in our laboratory (Dadda et al., 2015; Lucon-Xiccato et al., 2016). The first of these features regards the water used for the subject compartment, that was supplied from a 400-L tank with the same characteristics as the holding tanks and with a large population of guppies (approximately 50 individuals, both sexes, all ages). Two pumps in this tank provided a constant 1.5-L/min flow of water to the subject tank and water in excess was drained from a hole (diameter 2 cm) on the bottom of the subject tank and pumped back to the population tank.

The second additional feature regarded stimulus tanks. We modified the stimulus tanks dividing them into a front compartment (40 × 22 cm) and a back compartment (60 × 18 cm) by means of green plastic (Fig. 10). The back compartment was provided with gravel bottom, abundant plants, water filter and heater. The front compartment was the only part of the stimulus tank visible from the subject tank. Each front compartment was illuminated by a 15-W fluorescent lamp. An opaque lid covered each lamp on three sides to prevent it from lightning the room or the subject tank. Each stimulus tank housed permanently a group of 10 female guppies. These guppies were matched for body size with the subjects. Outside the experiments, stimulus guppies could freely access both front and back compartments by two guillotine doors.

Procedure

Thirty minutes before the experiment, the required number of stimulus guppies was confined in the front compartment by closing the guillotine doors. The position of the larger shoal was randomized between the two stimulus tanks across subjects to counterbalance potential differences between the stimulus fish. The subject was netted from the holding tank, transported in a plastic jar and gently inserted in the cylinder. We recorded the behaviour of subjects for 15 min with a camera placed 70 cm above the experimental tank. The recording started after the subject was inserted in the cylinder; this was necessary to avoid familiarization with the subject tank and to study the antipredator response (i.e., preference for the larger shoal) that guppies express in novel environments. Two subjects showed freezing behaviour and did not move for the entire experiment. They were discarded and substituted.

Analysis of the recordings

The choice of the subjects was analysed from the digital recordings played back on a computer screen. In a pilot experiment with a cylinder of this size (approximately 6 guppy’s body length), we observed that guppies spent more than 80 % of their time in the two zones closest to the stimuli, swimming against the transparent cylinder in the attempt to reach one of the two shoals. This behaviour did not include the up-down movements usually reported for escaping attempts and thrashing, but rather resembled the behaviour of guppies with the old shoal-choice method when they swim against the transparent partition to reach the stimuli. However, we cannot exclude that the swimming behaviour toward the stimuli was at least partially due to the willingness to escape from the cylinder. Guppies crossed the central sector of the cylinder only to reach the opposite side and swimming toward the opposite shoal. Based on the pilot experiment, we virtually divided the section of the cylinder in three zones with equal longitudinal extension (5 cm, corresponding to more than 2 guppy's body lengths), two choice sectors facing the stimulus shoals and one central, neutral sector (Fig. 10). We obtained this division by superimposing three lines on the computer screen by means of a computer software originally developed in our laboratory ('Overlap', written in Delphi5 Borland). The fish was considered to stay in one sector when the snout was within that area, but usually the subject was completely within the choice area when trying to reach one of the two shoals. As in previous studies, video recordings were analysed using a computer program originally developed in our lab ('Ciclic Timer', written in Delphi5 Borland) by an experimenter that operated on the computer keyboard. The experimenter was blind with respect to the position of the larger shoal. The software calculated the time spent in each sector of the cylinder. As a measure of preference for the larger shoal, we computed the proportion of time spent in the sector near the larger shoal over the total time spent in the two sectors near the two shoals. We also analysed the proportion of time spent in the central sector because this measure is used as a proxy of motivation to shoal: when highly motivated to shoal, fish are expected to spend little time in the central sector (Miletto Petrazzini & Agrillo, 2016). To investigate the temporal trend of preference, data were collected for each minute of test and analysed in blocks of 5 min each. Subsequently, we repeated the analysis focussing on the first minute and the initial 3 min of test.

To validate our approach, we conducted three further analyses on a subsample of 16 randomly chosen recordings. For the first analysis, a new blind experimenter scored the recordings in order to calculate the inter-rater reliability of our measure of preference. The second analysis consisted of a more detailed behavioural examination on the initial 5 min. In this latter analysis, we measured the time spent by the subject swimming against the transparency in direction of one of the two stimulus compartment, i.e. in the area subtended by a central angle of 60° (Fig. 10). The preference index was calculated as before. Finally, we examined the number of switches between the two choice areas and the permanence of the subjects in the choice area chosen at first.

Comparison with previous methods

The three previous studies used for the comparison tested guppies from the same population tested in the present work. The recent study by Lucon-Xiccato and colleagues (2016) was performed in the same apparatus used for the present work, but without the cylinder during the test; the two early studies used an apparatus with a 60 × 36 × 35 cm subject tank (Agrillo et al. 2012b; Bisazza et al. 2014a). For the two early studies, we re-analysed the pooled data ($N = 110$) after checking the absence of a significant difference ($t_{108} = 0.943$, $P = 0.348$). For what regards the recent study by Lucon-Xiccato and colleagues (2016), data were originally scored grouped in blocks of 4 min each; we therefore re-analysed the video recordings of female guppies tested with female shoals ($N = 24$) using the procedure described for experiment 1 (main analysis in 3 blocks of 5 min each; subsequent detailed analysis of the first minute and the initial 3 min). As for the present study, analysis was performed on the initial 15 min of the trial. Preference for the larger shoal was always arcsine-square root transformed before the analysis (Sokal & Rohlf, 1995). We could not analyse the frequency of switching of the experiment by Bisazza and colleagues (2014a), because this data was not reported in the original study and we do not possess the video recordings.

3.3.2 Experiment 2

In experiment 2, we observed 24 female guppies in the discrimination of 4 vs. 5 fish and 24 female guppies in the discrimination of 5 vs. 6 fish. Apparatus and procedure were the same as experiment 1. Based on the results of experiment 1, in experiment 2 we recorded the subjects for 5 min. We also measured the number of switches between the two choice areas in a random subsample of 12 subjects for each discrimination. Two subjects were discarded and substituted because they showed freezing behaviour.

3.4 Results

3.4.1 Experiment 1

Guppies spent 83.50 ± 5.88 % (M ± SD percentage) of time in the two choice sectors facing the stimulus shoals. We compared preference for the larger shoal with chance level (50 %) with one-sample t tests. When considering the entire observation period (15 min), subjects showed a significant preference for the larger shoal (58.43 ± 18.65 %, one-sample t test: t_{31} = 2.549, P = 0.016; Fig. 11). During the first block of time (min 1-5) subjects spent significantly more time near the larger shoal (61.05 ± 23.87 %, $t31$ = 2.621, P = 0.013; Fig. 11). Conversely, the preference for the larger shoal did not significantly differ from 50 % during the remaining blocks of minutes (block 2, min 6-10: 56.74 ± 22.35 %, t_{31} = 1.671, P = 0.105; block 3, min 11-15: 57.51 ± 27.81%, t_{31} = 1.583, P = 0.124; Fig. 11). The absence of a significant preference for the larger shoal in blocks 2 and 3 could be due to the fact that guppies tended to increase the time spent in the central sector as time passed (block 1: 12.63 ± 7.50 %, block 2: 18.8 ± 7.84 %, block 3: 18.1 ± 8.05 %; repeated measures ANOVA: $F_{2,62} = 9.584$, $P < 0.001$), suggesting a decrease in motivation to shoal.

To test whether with our new method shoal size preference can be assessed also with shorter time intervals, we performed further analysis on the initials 3 min and the first minute of the test. Preference for the larger shoal was significant in the initial 3 min (61.27 ± 24.89 %; $t_{31} = 2.601$, $P = 0.014$), and even considering only the first minute (62.77 ± 30.95 %; $t_{31} = 2.410$, $P = 0.022$; Fig. 11).

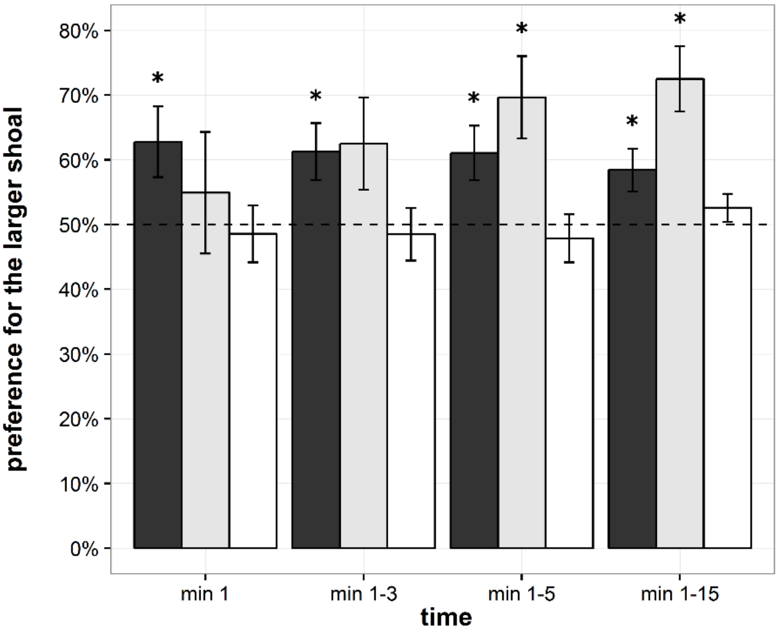


Fig. 11. Comparison of the performance (M ± SEM preference for the larger shoal in the 4 vs. 6 fish discrimination) with three different methods. Dark bars new method, experiment 1 of the present study. Grey bars method used by Lucon-Xiccato et al. (2016) that shared two features with the present work but not the use of the large cylinder. White bars classical method, pooled data from Agrillo et al. (2012b) and Bisazza et al. (2014a). Asterisks indicate that the preference is significantly above chance (50%, dashed line)

The reliability test showed that the primary measure of preference based on the position of the fish in the cylinder was robust and repeatable. The scores of the preference for the larger shoal assessed across the 15 min of test by the two experimenters were almost identical (2.63 ± 2.85 % difference between the two scores, calculated as | [score of experimenter A – score of experimenter B] | / [mean between experimenter A and experimenter B] × 100). Furthermore, there was a strong correlation between the two scorers (Pearson's $r = 0.991$, $P < 0.001$). The secondary, more refined, measure of preference based on the orientation of the subject was highly correlated with the primary measure of preference based on the position of the fish in the cylinder (Pearson's $r = 0.965$, $P < 0.001$). The number of switches between the two shoals was 38.25 ± 12.17 across the 15 min of test. Several switches occurred also in the initial 5-min block (14.13 ± 7.04). The duration of guppies' permanence in the choice area of the first chosen shoal was 31.56 ± 41.92 s.

Comparison with previous methods

With the classical method (Agrillo et al., 2012b; Bisazza et al., 2014a), guppies did not discriminate between shoals made of 4 vs. 6 fish considering the overall 15 min ($t_{109} = 1.227$, $P = 0.223$), the initial 5 min ($t_{109} = 0.614$, $P = 0.541$), the initial 3 min ($t_{109} = 0.461$, $P = 0.646$), or the first minute of test ($t_{109} = 0.362$, $P = 0.718$; Fig. 11). With the method adopted by Lucon-Xiccato and colleagues (2016) which shares two features with our new method, guppies showed a significant preference for the larger shoal in the overall 15 min and the initial 5 min of the test ($t_{23} = 4.409$, $P < 0.001$ and $t_{23} = 2.758$, $P = 0.012$, respectively), but not in the initial 3 min or in the first minute of test ($t_{23} = 1.597$, $P = 0.124$ and $t_{23} = 0.547$, $P = 0.589$, respectively; Fig. 11). When we compared the preference for the larger shoal across the three experiments, we found a significant difference in the overall 15 min (one-way ANOVA: $F_{2,163} = 8.786$, $P < 0.001$) and in the initial 5 min ($F_{2,163} = 3.705$, $P = 0.027$), but not in the 3-min and 1-min intervals ($F_{2,163} = 1.840$, $P = 0.162$; $F_{2,163} = 1.167$, $P = 0.314$, respectively). The lack of statistical difference across the three methods in the first minute and the initial 3 min of test may be partly related to the fact that the variance in the preference for the larger shoal appeared to increase as the time interval considered for the analysis decreased (Bartlett test: Lucon-Xiccato and colleagues: $P = 0.030$; classical method: $P < 0.001$; present work: $P = 0.052$; Fig. 11).

The comparison with the previous studies revealed also indirect evidence of increased switching. The frequency distribution plot of the preference for the larger shoal in the initial 5 min of the test resemble a normal distribution in our study (Fig. 12a), conversely, it is skewed in the study by Lucon-Xiccato and colleagues (2016; Fig. 12b), and it is clearly bimodal in experiments with the classical method suggesting that here fish tend to remain close to the first chosen shoal (Agrillo et al., 2012b; Bisazza et al., 2014a; Fig. 12c).

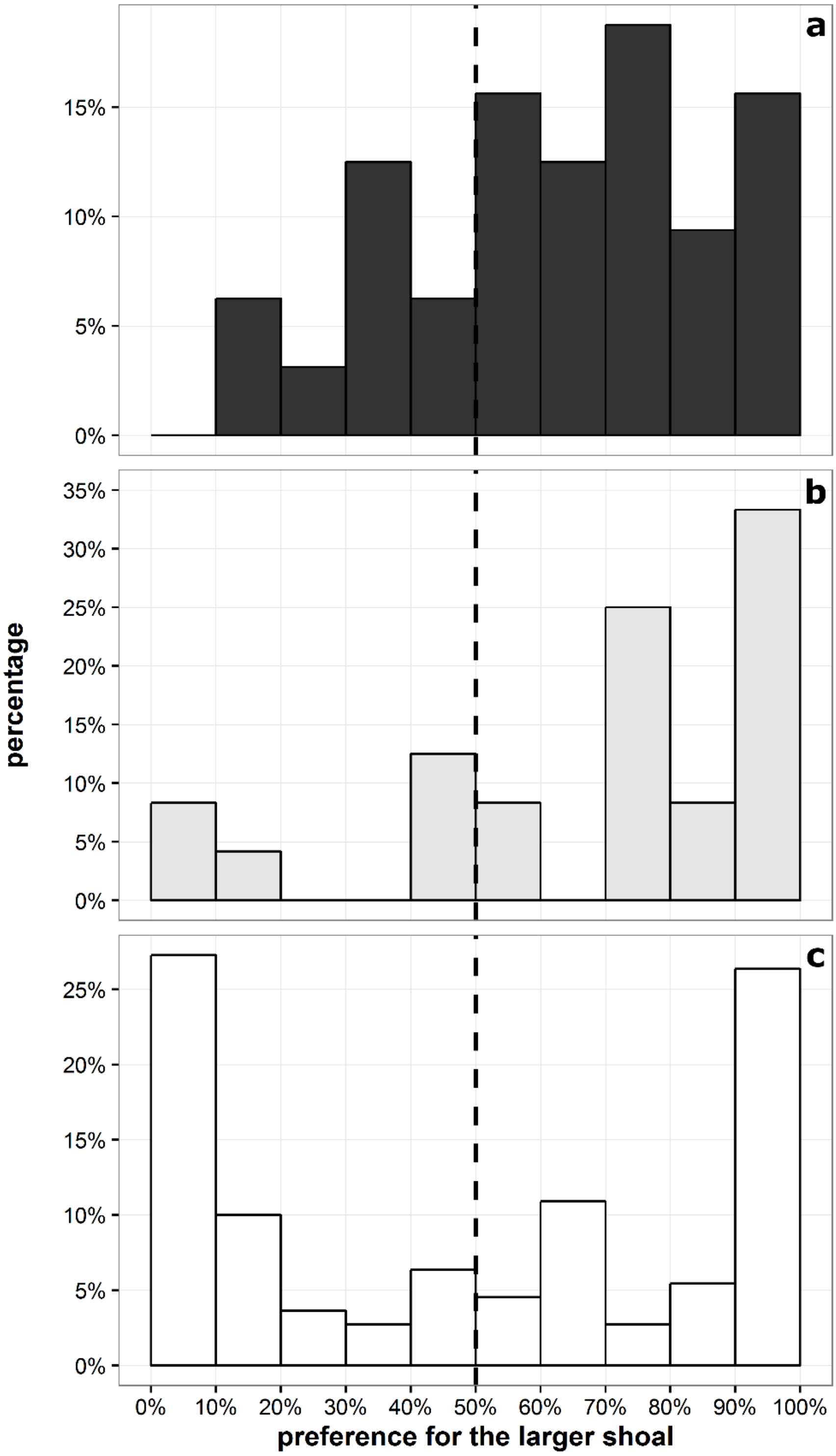


Fig. 12. Frequency distribution plot of the preference for the larger shoal (4 vs. 6 fish) in the initial 5 min assessed with **a** the new method (experiment 1), **b** the method used by Lucon-Xiccato et al. (2016) and **c** the classical method used by Agrillo et al. (2012b) and Bisazza et al. (2014a)

To confirm this increased switching, we calculated the frequency of switching per minute in our work and we compared it with that observed in previous works (Agrillo et al. 2012b; Lucon-Xiccato et al. 2016) using a one-way ANOVA. We found a significant difference between the frequency of switching across the three studies ($F_{2,53}$ = 96.59; $P <$ 0.001). A Tukey post-hoc test revealed that the frequency of switching observed with our method (2.55 ± 0.81 switches per minute), was significantly greater than the one observed with previous methods (Agrillo and colleagues: 0.40 ± 0.20 switches per minute, $P < 0.001$; Lucon-Xiccato and colleagues: 0.36 ± 0.44 switches per minute, $P < 0.001$), but there was not difference between the two previous methods ($P = 0.975$).

3.4.2 Experiment 2

In the 4 vs. 5 discrimination, guppies significantly preferred the larger shoal (62.82 ± 26.23 %; $t_{23} = 2.120$, $P = 0.045$; Fig. 13). Subjects appeared to choose the larger shoal also in the 5 vs. 6 discrimination (54.88 ± 19.01 %; Fig. 13), but here the preference was not significantly greater than chance ($t_{23} = 1.306$, $P = 0.204$). We compared the preference for the larger shoal in the initial 5 min of the experiment 1 (4 vs. 6) and the two discriminations of the experiment 2 with a one-way ANOVA. This analysis found no significant difference between the three discriminations ($F_{2,77} = 0.679$, $P = 0.510$; Fig. 13), suggesting that, although not significantly achieved, the 5 vs. 6 discrimination might represent the threshold of guppies' shoal size acuity.

Finally, we run a generalized linear model (with Poisson error distribution) to compare the number of switches during the initial 5 min of test of the three numerical discriminations tested in this work. we found that numerical discrimination significantly affected the number of switches ($\chi^2_2 = 20.319$, $P < 0.001$). A Tukey post-hoc test revealed that guppies switched between the two shoals with similar frequency in the 4 vs. 6 and the 4 vs. 5 discrimination (4 vs. 6: 14.13 ± 7.04 switches; 4 vs. 5: 11.92 ± 6.36 switches; P =

0.249); however, guppies switched more often in the 5 vs. 6 discrimination (18.92 ± 10.61 switches) than in the two easier discriminations (4 vs. 6: *P* = 0.005; 4 vs. 5: *P* < 0.001).

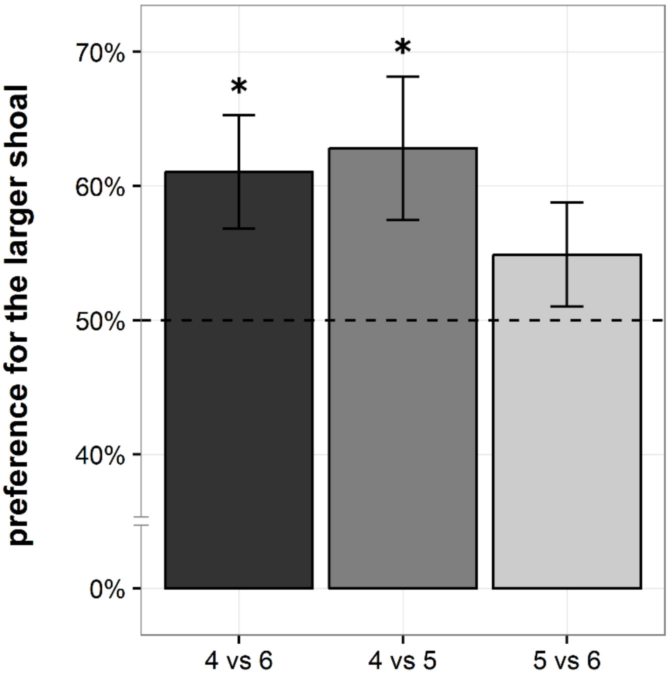


Fig. 13. Preference for the larger shoal (M ± SEM) in the three discriminations investigated with our new method (experiment 1: 4 vs. 6; experiment 2: 4 vs. 5 and 5 vs. 6). Asterisks indicate that the preference is significantly above chance (50%, dashed line)

3.5 Discussion

In the classical shoal-choice test, shoaling preference is assessed with extended periods of time (e.g., up to 30 min; Agrillo et al., 2007), mainly to counter the measurement error due to the initial low mobility of the subjects. In this study, we developed and tested a new version of the shoal-choice test, which was aimed at obtaining a rapid assessment of shoal size discrimination ability. Our new method was expected to favour more frequent switching between stimulus shoals and hence more selectivity since the beginning of the trial.

The results of experiment 1 showed that shoaling preference can be assessed very quickly. Female guppies were able to discriminate between shoals of different size after only 5 min of testing. Indeed, our data indicated that even shorter intervals can be used to assess quantity discrimination. The preference for the larger shoal was significant considering the initial 3 min of the test and even in the first minute, although variance appears to increase with decreasing time interval. Guppies tested with the old method did not significantly select the larger shoal in such reduced intervals.

Our analysis suggests that the greater efficiency of our method at the beginning of the test is related to the fact that guppies switched often between the two shoals, approximately five times more frequently than with the previous method (Agrillo et al., 2012b; Lucon-Xiccato et al., 2016). Further evidence of the same effect is that in our experiment the average time spent by subjects close to the first chosen shoal was around half a minute, which is ten times shorter than that observed in a previous experiment (Lucon-Xiccato & Dadda, 2017). The increased frequency of switching is likely to favour the comparison between the two shoals and is therefore a more accurate assessment. The same increased switching rate could in theory be obtained by greatly reducing the length of the subject tank. However, as a side effect, this would considerably reduce the distance between the two stimulus shoals with the consequence that subjects might perceive themselves as part of a single large shoal and thus reduce their selectivity. Interestingly, in experiment 2, we also found that frequency of switching increased as the ratio between the two shoals increased. It has often been assumed that animals should increase sampling to obtain high-quality information (Chittka et al., 2009; Stephens, 2008). Accordingly, guppies might need to switch more between two shoals with very similar size (5 vs. 6 fish) in order to compare the two options and eventually identify the larger shoal.

A rapid assessment of shoal-size discrimination ability can have multiple advantages. For instance, it is possible to use the new method in studies that involve a large number of subjects, such as those aimed at disclosing subtle individual differences (Cote et al., 2012) or screening a large population for genetic studies (Patton & Zon, 2001). Equally important, in nature, shoaling decisions are likely to occur very rapidly (Krause et al., 1998). A fast shoal-choice test is therefore expected to measure the cognitive processes normally activated to solve shoal-size discrimination under natural conditions. Beyond the obvious time saving for researchers, a fast test also minimises exposure of the subject to the unfamiliar environment, reducing potential stress (Cachat et al., 2010). This is very important considering that, for the subject, being enclosed in the cylinder might be *per se* more stressful than freely swimming in the apparatus.

It should be said that, despite the initial advantage of our new method, when considering intervals longer than 3 min, guppies were more accurate with the recent method adopted by Lucon-Xiccato and colleagues (2016). In this version of shoal choice, fish were free to swim in the subject tank after a 2-min acclimatisation period in a small cylinder; however, similarly to our method, water was supplied from a tank with a large guppy population, and stimuli lived in the stimulus tanks to favour habituation. These two latter features thus appear to be enough to allow a greater accuracy of guppies but do not solve the problem of initial low performance as the use of the cylinder does. Our method is perhaps more suitable to fast assessment of discrimination abilities, but the method used by Lucon-Xiccato and colleagues (2016) could be considered a valid alternative when time is not constrained or a greater accuracy level is required. Further, the method by Lucon-Xiccato and colleagues might be preferred in experiments studying ecological aspects of shoal choice or the behavioural response of the subject fish (e.g., aggression) because fish are free to swim and this might result in a more spontaneous behaviour.

The new method that we developed not only allows a fast assessment of shoal size discrimination abilities, but also appears to disclose significant discrimination for numerical ratio that normally are not achieved by fish using classical shoal-choice methods (Agrillo & Dadda, 2007; Agrillo et al., 2008a; Agrillo et al., 2012b; Gómez-Laplaza & Gerlai, 2011). In experiment 2, guppies achieved a 4 vs. 5 fish discrimination, a numerical comparison that to date guppies have been shown to solve only with extensive training procedures (Bisazza et al., 2014b). This result does not necessary indicate that with our new method the numerical accuracy of guppies is equivalent to what observed with training. Indeed, Bisazza and colleagues (2014b) prevented guppies from learning the discrimination using the non-numerical cues that covary with number (e.g., the area of the stimuli), but we did not. Our guppies might have exploited both numbers and other cues to identify the larger shoal (Gómez-Laplaza and Gerlai, 2012), and the use of multiple cues is known to affect numerical discrimination performance (Agrillo et al., 2009).

The findings of experiment 2 on the maximum numerical acuity of guppies raise the issues of the comparability of results obtained from studies that adopted different methodologies. The importance of methodological differences is a well-known issue in experiments that employ discrimination learning procedures. Frequently, the result varies depending on the type of learning mechanism involved (e.g., classical vs. operant conditioning: Brembs & Heisenberg, 2000), the type of stimuli used (e.g., real objects vs. pictures: O’Hara et al. 2015) or other features such as length of training and type of contingency (Wickens et al., 1970; Giurfa et al., 2003). Guppies have been recently found to achieve enhanced discrimination-learning performances by adopting an ‘ecological’ training method that resembles the natural foraging behaviour of this species (Bisazza et al., 2014b; Chapter 3). Our present study suggests that cognitive tests which are not based on discrimination learning procedures and exploit spontaneous behaviours are not immune to the issue of methodological differences. Indeed, in experiment 2 we showed that subtle methodological differences can affect the measurement of discrimination accuracy in guppies. Since many differences between different studies on the same species or between different species may be due to this issue, greater attention to methodological details is required in future research.

4. Study 3: quantity discrimination by treefrogs*

* The contents from this chapter have been published as Lucon-Xiccato, T., Gatto, E., & Bisazza, A. (2018). Animal Behaviour, 139, 61-69.

4.1 Abstract

To make foraging, reproductive and antipredator decisions, animals often have to discriminate discrete and continuous quantities (numbers and sizes of objects, respectively). Few studies have investigated discrete quantity discrimination in amphibians, but this has been done only in the context of prey selection. Using a species with arboreal habits, the Italian treefrog, *Hyla intermedia*, we investigated whether amphibians discriminate both discrete and continuous quantities when choosing between microhabitats. In field experiments, we showed that newly metamorphosed treefrogs exhibit a preference for microhabitats with abundant and tall grass. In the laboratory, treefrogs presented with the dichotomous choice between two sets comprising different numbers of vertical green bars (simulating grass clumps) showed a preference for the larger set and discriminated between 1 and 2 bars and between 2 and 4 bars, but not between 2 and 3 bars and between 3 and 4 bars. When presented with two bars of different size (i.e. one bar was taller and wider), treefrogs preferred the larger bar up to a 0.25 surface area ratio. Control experiments suggested that treefrogs represent numbers rather than continuous variables to discriminate between sets of bars and that they use the height but not the width of the bars to discriminate sizes. We also found evidence of a possible trade-off between speed and accuracy: individuals that chose

more quickly did not display a significant preference for the larger bar/set of bars. These findings suggest that for amphibians, as for other vertebrates, a variety of decision-making processes can rely on quantitative abilities

4.2 Introduction

Accumulating evidence suggests that animals discriminate quantities to make decisions in several contexts, such as when foraging or choosing social groups or reproductive partners (reviewed in Agrillo & Bisazza, 2014). The type of quantity to be discriminated may vary according to the context and may be either discrete (i.e. the number of objects in a group) or continuous (i.e. the size of an object). For example, mammals, birds, fish and even invertebrates can discriminate the number of food items in a patch or the size of the individual food items in order to choose the larger food quantity (Bogale et al., 2014; Cross & Jackson, 2017; Lucon-Xiccato et al., 2015). Females of many species choose between potential mates based on the area of body coloration or on the length of ornaments (Basolo, 1990; Griggio et al., 2011). Social fish choose between available shoals based on the number and the size of the individuals in each shoal (Gómez-Laplaza & Gerlai, 2011; Ranta et al., 1992).

Research on vertebrates has revealed performance similarities across species and across tasks; this has suggested the existence of a cross-modal core system for processing quantity information inherited from a common ancestor (Beran, 2008; Feigenson et al., 2004). However, most studies have focused only on mammals, birds and fish, and on only one or just a few contexts in each species; thus, more comparative research is required to evaluate the aforementioned hypothesis. Only three studies have directly investigated quantitative abilities of amphibians, and the findings from these studies provide evidence of amphibians discriminating discrete quantities in the context of choosing their prey.

Salamanders of the genus *Plethodon* can choose the larger group of live prey (fruit flies and crickets) well enough to discriminate 2 from 3 (Krusche et al., 2010; Uller et al., 2003). Similarly, oriental fire-bellied toads, *Bombina orientalis*, recognize and feed on the larger of two groups of 2 and 3 prey (Stancher et al., 2015). There is evidence that amphibians also use quantitative information in contexts other than prey selection (e.g. female mate choice: Arntzen, 1999; male calling: Gerhardt et al., 2000; Rose, 2018), but these studies usually do not directly address the cognitive abilities involved in the discrimination.

The main goal of the present study was to investigate amphibians' quantity discrimination ability in a context other than prey choice. For this, we focused on the microhabitat choices made by an arboreal anuran species, the Italian treefrog, *Hyla intermedia*. Arboreal anurans, especially the juveniles, climb vegetation upwards to evade terrestrial predators and/or to improve foraging (Stewart, 1985). It appears likely that microhabitats with more plants and larger plants will confer greater protection and contain more resources; thus, if vegetation is distributed in clumps or varies in size, treefrogs will display a preference for microhabitats with more plants or those with larger plants, as observed in fish and lizards (Bartholomew, 2012; Cooper & Whiting, 2000). In agreement with this prediction, treefrogs in nature are more common in habitats with abundant vegetation (Ildos & Ancona, 1994) and they actively prefer microhabitats with vegetation (Michaels et al., 2014), which suggests that the observed distribution might, at least in part, be due to the treefrogs' behavioural preferences.

The treefrogs' microhabitat preference might be useful to study quantity discrimination, as commonly done with spontaneous preferences in other species (reviewed in Agrillo & Bisazza, 2014). We addressed this possibility by performing four experiments in the field (experiments 1, 2, 3 and 4). In experiment 1, we determined whether newly metamorphosed Italian treefrogs living in grassland are attracted to microhabitats with abundant and tall grass, and whether they tend to climb such grass. In experiment 2, we tested whether treefrogs prefer larger grass clumps (both in number and size of leaves) using a dichotomous choice arena. These experiments showed that treefrogs exhibit a spontaneous preference for the larger available grass clump. To study quantity discrimination using this preference, it is important to control quantitative features of the stimuli, such as the area and height, and other features, such as the colour. Since this is difficult using grass stimuli, we evaluated the use of artificial stimuli: in experiments 3 and 4, we tested whether treefrogs are attracted by green bars printed on paper and whether this attraction was similar to that for real grass, respectively.

The following four experiments addressed our main objective of investigating quantity discrimination during microhabitat choice. These experiments were performed in the laboratory to ensure controlled conditions (e.g. temperature, stimuli illumination) and thus reduce variability in subjects' behaviour. In experiment 5, we tested whether treefrogs could discriminate between different numbers of same-sized printed green bars and we identified the upper limit of their capacity for discrimination by presenting choices of 1 vs. 4 bars, 2 vs. 4 bars, 2 vs. 3 bars and 3 vs. 4 bars. In experiment 6, we tested whether treefrogs could discriminate between two green bars of different sizes.

In experiments 7 and 8, we identified which object attributes are used by treefrogs to discriminate quantities. The discrimination ability determined in experiment 5 might have been brought about by the treefrogs representing the number of objects in a scene or, alternatively, by the treefrogs representing a value of a continuous variable that covaries with numerosity, such the overall surface area or the convex hull (Davis & Pérusse, 1988). For example, as the bars of experiment 5 had the same size, the set with more bars also had larger surface area (i.e. the larger stimulus is more green) and treefrogs might have based their choices solely on this continuous variable. Some studies have supported the ‘last resort’ hypothesis: that animals most often use continuous variables to discriminate between discrete quantities and that they use numerical information only when they cannot rely on the continuous variables (e.g. Vos et al., 1988). However, other studies have suggested that animals readily make spontaneous decisions based on numerosity instead of on the covarying continuous variables (Ferrigno et al., 2017). To disentangle these possibilities, in experiment 7, we observed treefrogs choosing between 2 and 4 bars with the same overall surface area (experiment 7a) or convex hull (experiment 7b).

We considered whether treefrogs could use two different attributes of the stimuli (i.e. the height and the width of the bars) when discriminating the size of the bars. If one attribute is more relevant for treefrogs' choice (e.g. taller plants might confer better protection from terrestrial predators than wider plants), they might choose based on that attribute only. It is also possible that height and width of plants covary in nature and that treefrogs have been selected to rely on one of these attributes to minimize neuronal resources required for the choice (Todd, 2001). In experiment 8, we investigated whether treefrogs were influenced more by height (experiment 8a) or by width (experiment 8b) when choosing the larger bar.

The last goal of this study was to ascertain whether trade-offs between speed and accuracy are important to treefrogs' quantity discriminations. For various cognitive tasks, gathering accurate information and comparing the available options take a considerable amount of sampling time (Chittka et al., 2009). Both long sampling times and wrong choices may be costly and may reduce fitness. It has been proposed that animals deal with this problem by trading off choice time and choice accuracy (Chittka et al., 2009). We hypothesized that treefrogs would show a trade-off between speed and accuracy when discriminating quantities, so in the laboratory experiments we measured treefrogs' choice time to investigate this hypothesis.

4.3 Methods

Subjects

We collected treefrogs 15 days after metamorphosis in northeast Italy, near to Padova (45°32’30’’N, 11°53’40”E). The population used in the study reproduces in an artificial pond. Each year, we observe a large number of egg masses (>30). Although the adults of this species are known to inhabit trees and bushes, the subjects of this study were collected while climbing grass close to the pond (Fig. 14a). We collected the frogs in the morning between 0600 and 0700 hours using a wet hand net, placing them in a plastic box with grass for transfer. We tested groups of approximately 25 frogs each day, randomly assigning each animal to the different experimental conditions. The frogs transferred to the laboratory were kept at 26°C and provided with a dish of water and spray of nebulized water to keep part of the grass moist. In the field experiments, we tested 44 frogs overall, divided in the different experiments as follow: eight in experiment 1, 12 in experiment 2, eight in experiment 3 and 16 in experiment 4. In the laboratory experiments, we tested 328 frogs overall, 96 in experiment 5, 48 in experiment 6, 48 in experiment 7 and 136 in experiment 8. Each frog was tested only once.

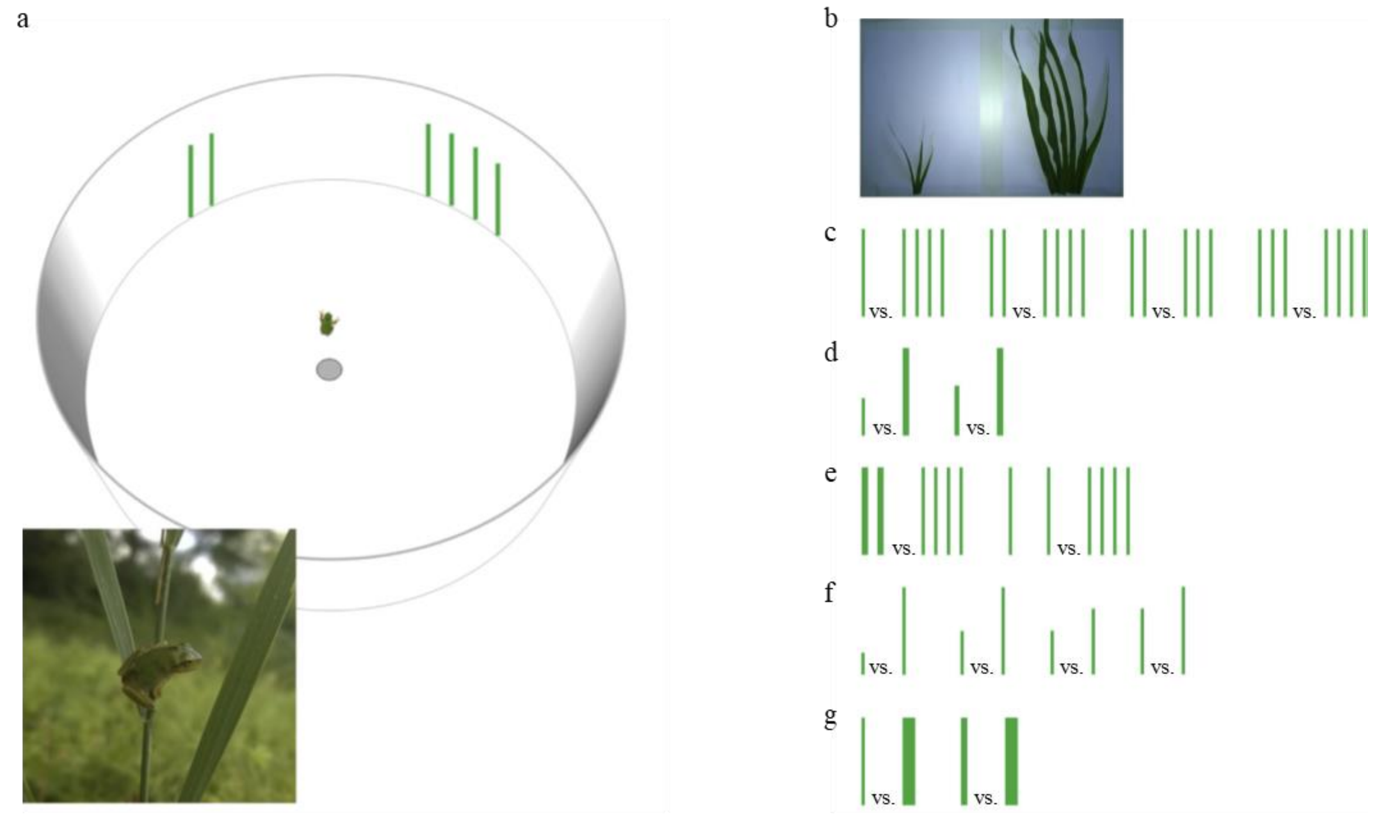


Fig. 14. (a) View from above of the arena used in experiments 2-8 and example of treefrog climbing grass in the field; stimuli used in (b) experiment 2, (c) experiment 5, (d) experiment 6, (e) experiment 7 and (f, g) experiment 8.

4.3.1 Apparatus and Procedures: Experiments in the Field

We performed experiment 1 in the same grassland where we collected the subjects to test whether treefrogs are attracted by grass and tend to climb it. We released the subjects individually into the middle of an area (diameter 100 cm) that lacked tall grass, but that was surrounded on one side by grass 25 - 35 cm in height. We constructed this area by cutting and removing the grass. We performed the experiment using four replicates of the setting. We recorded the time that the subject took to reach the tall grass, whether the subject climbed the grass, and the time to climb to a height of 15 cm.

We performed experiments 2, 3 and 4 in a building close to the field site so that fresh grass could be used as a stimulus. We conducted these experiments in a white plastic circular arena (diameter 80 cm, height 75 cm; Fig. 14a). An LED spot lamp (100W, 1000 lm, 100 angle) placed 100 cm above the middle of the arena illuminated the apparatus; this lamp was the only light source because we kept the arena in a dark room. This setting ensured absence of shadows and homogeneous illumination of the stimuli. A PVC tube (8 cm long) was connected to a hole (diameter 2.5 cm) in the middle of the floor of the arena and served as the starting point for the subjects. This tube was inclined at a 45° angle to the floor of the arena. Before each trial, the experimenter positioned the stimuli against the wall of the arena, facing the exit from the PVC tube. The stimuli were either grass clumps or green bars printed on white paper (see below). Then, the experimenter placed one, randomly selected frog in a plastic jar and transferred it inside the PVC tube for a 30 s habituation. To start the trial, the experimenter slowly injected water inside the PVC tube with a 60 cl syringe and a silicone tube connected to the bottom of the PVC tube. This caused the frog to emerge into the arena, in which it could move freely (see Supplementary video). The trial ended when the frog reached the wall of the arena and touched it, jumped on it or stayed 10 s within one body length from it. We allowed the frog 30 min to reach the wall of the arena. If the frog did not move within 15 min, we stopped the trial.

Supplementary video related to this article can be found at https://doi.org/10.1016/j.anbehav.2018.03.005

In experiment 2, we compared the choice between one microhabitat with a large grass clump and one microhabitat with a small grass clump. The stimuli were clumps of freshly collected grass leaves (Fig. 14b), attached to sheets of A4 paper using transparent tape and placed on the walls of the arena. The two clumps were 45 cm apart. We recorded whether frogs chose the large or the small clump of grass. Our operational definition of choice for one stimulus was that the frog touched the stimulus, or it stayed within one body length of the part of the wall where the stimulus was. We performed four replicates with the stimuli.

In experiment 3, we tested whether frogs are attracted by green printed bars simulating vegetation. The stimuli were 12 green bars, each 2 cm wide and 28 cm high, printed on A3 paper separated by 2 cm of empty (i.e. white) space. Treefrogs are reported to perceive colours (Gomez et al., 2010), but their exact sensitivity is unknown; thus, we tried to match as closely as possible the colour of grass in the field site. We used white paper as background to improve visibility of the bars. This stimulus consisted of alternating green and white bars, with the two colours having the same area overall. We measured whether each subject touched a green bar or a white bar first. We then recorded the time spent on the green bars and the time spent on the white bars while the subjects were climbing for 2 min or until they reached the tip of the bars. Because green and white bars occupied the same surface area, if the frogs moved randomly towards the stimuli, they would be expected to choose the green and the white bars the same number of times, and to spend an equal amount of time over the green and white bars while climbing.

In experiment 4, we compared frogs' preference for printed bars and real grass. The stimuli were a single green bar, 1.5 cm wide and 25 cm high, printed on an A4 sheet, and one grass leaf of the same size attached to an A4 sheet. We recorded whether frogs chose the bar or the grass leaf.

4.3.2 Apparatus and Procedures: Laboratory Experiments

Experiments 5, 6, 7 and 8 aimed at assessing treefrogs' quantity discrimination abilities; we performed them in the laboratory to ensure controlled conditions. The experimental apparatus and the procedures were similar to those of experiments 2, 3 and 4. The stimuli were green printed bars of different numbers and sizes according to the experiment (see below). We recorded the frog's choice of the ‘larger’ or ‘smaller’ quantity and the time it had taken to reach the stimulus after emerging from the PVC tube. We conducted the experiments between 14:00 and 20:00 hours because a preliminary study showed that frogs are more attracted by the stimuli in the afternoon. Overall, 44 frogs did not complete the trial within the predetermined time or did not touch the part of the wall where the stimuli were and were replaced. Replacement was done because frogs that did not choose between the stimuli do not provide information about discrimination ability.

In experiment 5, the stimuli were two sets with a different number of bars to study discrete quantity discrimination. Each bar was 1 cm wide and 28 cm high. Numerosity ratios and number of bars were as follow: numerosity ratio 0.25: 1 vs. 4 bars; numerosity ratio 0.5: 2 vs. 4 bars; numerosity ratio 0.67: 2 vs. 3 bars; numerosity ratio 0.75: 3 vs. 4 bars; Fig. 14c. The bars within each set were separated by a gap of 3 cm. We tested 24 frogs for each numerical ratio.

In experiment 6, we presented two bars of different size to study continuous quantity discrimination. By proportionally altering both height and width, we obtained ratios between the surface area of the two bars that corresponded to the numerical ratios significantly discriminated by the frogs in experiment 5 (ratio 0.25: a 1 × 12 cm bar vs. a 2 × 24 cm bar; ratio 0.5: a 1.5 × 16 cm bar vs. a 2 × 24 cm bar; Fig. 14d). We used the same ratios of experiment 5 to compare frogs' accuracy in discriminating discrete and continuous quantities (Lucon-Xiccato & Dadda, 2017; Lucon-Xiccato et al., 2015). We tested 24 frogs for each size ratio. From this experiment onwards, we first tested the two ratios significantly discriminated in experiment 5 to minimize the number of wild animals needed to complete the experiment (as required by law in our country). We thus only tested frogs with larger ratios if they significantly discriminated the two ratios first administered.

From the results of experiment 5 it is not possible to ascertain whether frogs discriminate the two sets of bars based on bar number or on the continuous variables covarying with bar number (Davis & Pérusse, 1988). To address this point, in experiment 7, we sequentially controlled for the two most important attributes of the bars' sets that covary with numerosity, the overall surface area (sum of the surface area of each individual bar) and the convex hull (distance between the two most external bars of a set; Davis & Pérusse, 1988). In experiment 7a (control for overall surface area), the stimuli consisted of one set of 2 bars 2 × 28 cm in size and one set of 4 bars 1 × 28 cm in size; bars within the same set were separated by a 3 cm gap (Fig. 14e). In experiment 7b (control for convex hull), the stimuli included one set of 2 bars 1 × 28 cm in size separated by a gap of 11 cm, and one set of 4 bars 1 × 28 cm separated by a gap of 3 cm (Fig. 14e). If frogs fail the discrimination in one of these two controls, we would conclude that they probably discriminate between the two sets of bars based on the continuous variable corrected for. Conversely, if frogs choose the set with more bars in both experiments 7a and 7b, this would indicate that they base their choice on the number of bars. We tested 24 frogs in experiment 7a and another 24 frogs in experiment 7b.

From the results of experiment 6 it is not possible to ascertain whether frogs based their choice on the height or on the width of the bars, because the larger bar was both taller and wider than the smaller bar. We addressed this point in experiment 8 following the strategy of experiment 7 (i.e. sequential control of the attributes). In experiment 8a, we initially presented two bars of different height, based on the quantity ratios discriminated by the frogs in experiment 5 (ratio 0.25: a 1 × 7 cm bar vs. a 1 × 28 cm bar; ratio 0.5: a 1 × 14 cm bar vs. a 1 × 28 cm bar; Fig. 14f). Since we found a significant discrimination for both ratios, we then tested frogs with the two more challenging ratios used in experiment 5 (ratio 0.67: a 1 × 14 cm bar vs. a 1 × 21 cm bar; ratio 0.75: a 1 × 21 cm bar vs. a 1 × 28 cm bar; Fig.

14f). In experiment 8b, we presented two bars of different width based on the quantity ratios discriminated by the frogs in experiment 5 (ratio 0.25: a 1 × 28 cm bar vs. a 4 × 28 cm bar; ratio 0.5: a 2 × 28 cm bar vs. a 4 × 28 cm bar; Fig. 14g). As an example, if frogs discriminate between different sized bars in experiment 8a but not in experiment 8b, then we would conclude that they base their choice on bar height rather than on bar width. We tested 24 frogs for each ratio in experiment 8a and 20 frogs for each ratio in experiment 8b.

Statistical Analysis

The statistical analysis was performed in R version 3.2.1 (The R Foundation for Statistical Computing, Vienna, Austria, http://www. r-project.org). The statistical tests were two-tailed and the significance threshold was $P = 0.05$ if not stated otherwise. To study the preference of treefrogs for a certain stimulus, we compared the observed number of subjects choosing this stimulus with the number expected by chance (50%) using chi-square tests. In experiment 3, we additionally tested the preference for the green bars by comparing the percentage of time spent over the green bars with chance (50%) using a one-sample t test. In the laboratory experiments, the analysis was initially drawn separately for the different quantity ratios (experiments 5 and 6) and control conditions (experiments 7 and 8). For the experiments in which frogs were presented with different quantity ratios (experiments 5, 6 and 8), we then performed a cumulative analysis on all the ratios using generalized linear models (GLMs) with binomial error distribution and logit link function. As a dependent variable, we used the choice of each frog (larger or smaller). We initially fitted the model with intercept only, to test whether frogs chose the larger stimulus overall, independently from the ratio; then, we fitted ratio as a factor to test for differences between the ratios. In the experiments in which frogs were observed in more than two ratios (experiment 5 and 8a), we performed a Tukey post hoc test if the factor ratio was significant; we also tested for a significant linear trend. To study speed-accuracy trade-off, we analysed frogs' choices in all the laboratory experiments using a GLM as described above. We fitted Log(choice time) as the covariate and experiment as the fixed effect. The interaction was omitted in the final model because it was not significant (Engqvist, 2005).

4.4 Results

4.4.1 Experiment 1: Attraction to Microhabitats with Grass in Nature

All eight frogs rapidly reached the tall grass (time to reach the grass: 130.38 ± 50.32 s, mean ± SD). After reaching the grass, all eight frogs rapidly climbed it up to a height of 15 cm (climbing time: 65.12 ± 48.10 s).

4.4.2 Experiment 2: Preference for Larger Grass Clumps

Ten of 12 frogs chose a stimulus. The two remaining frogs did not select any stimulus; one did not move for 15 min, at which point we stopped the trial, and the other touched the white wall of the arena. Of the 10 frogs that made a choice, nine chose the larger grass clump and one chose the smaller one. The number of frogs choosing the larger stimulus (90%) was significantly greater than chance (chi-square test: $\chi^2 = 6.400$, $P = 0.011$).

4.4.3 Experiment 3: Attraction to Green Printed Bars

All the frogs reached the stimulus paper. Seven of eight chose the green bars first (Tab. 3), a preference that was significantly greater than chance (chi-square test: $\chi^2 = 4.500$, $P = 0.034$). Overall, the frogs spent significantly more time climbing the green bars than the white bars (89.71 ± 10.20% time spent over the green bars; one-sample t test against random choice: $t_7 = 24.746$, $P < 0.0001$; Tab. 3). The only frog that initially chose a white bar spent 87% of its climbing time on the green bars.

Subject	First choice	Time climbing (s)	Time on the green bars (s)
1	Green bars	120	120
2	Green bars	120	91
3	Green bars	120	96
4	Green bars	55	52
5	Green bars	120	120
6	White bars	71	62
7	Green bars	120	120
8	Green bars	65	52

Tab. 3. Results of experiment 3.

4.4.4 Experiment 4: Preference for Green Bars versus Grass

Two frogs did not move for 15 min and were removed from the sample; the remaining 14 reached one of the stimuli. Six frogs chose the printed bar and eight chose the grass leaf (chi-square test: $\chi^2 = 0.286$, $P = 0.593$).

4.4.5 Experiment 5: Discrete Quantity Discrimination

For both the easy ratios, the number of frogs choosing the larger number of bars was higher than expected by chance (1 vs. 4: 22/24; $\chi^2 =16.667$, $P < 0.0001$; 2 vs. 4: 19/24; $\chi^2 = 8.167$, *P* = 0.004; Fig. 15a). In contrast, for the higher ratios, the number of frogs choosing the larger number of bars did not differ significantly from chance (2 vs. 3: 14/24; $\chi^2 = 0.667$, $P = 0.414$; 3 vs. 4: 13/24; $\chi^2 = 0.167$, $P = 0.683$; Fig. 15a).

When considering all the numerical ratios, the GLM showed that the number of frogs choosing the stimulus with the larger number of bars (68/96, 70.83%) was significantly greater than chance (estimate = 0.887, SE = 0.225, z = 3.952, $P < 0.0001$). The linear trend was significant ($P = 0.003$), indicating that the number of frogs choosing the larger number of bars decreased with increasing ratio between numerosities. The GLM revealed a significant difference between the ratios ($\chi^2_3 = 11.861$, $P = 0.008$). A post hoc test found that the difference was significant between the 1 vs. 4 bars and the 3 vs. 4 bars discrimination ($P = 0.040$), close to the threshold for significance between the 1 vs. 4 bars and the 2 vs. 3 bars discrimination ($P = 0.069$), and not significant between the other numerosity ratios (1 vs. 4 bars and 2 vs. 4 bars: $P = 0.628$; 2 vs. 4 bars and 2 vs. 3 bars: $P = 0.411$; 2 vs. 4 bars and 3 vs. 4 bars: $P = 0.267$; 2 vs. 3 bars and 3 vs. 4 bars: $P = 0.991$).

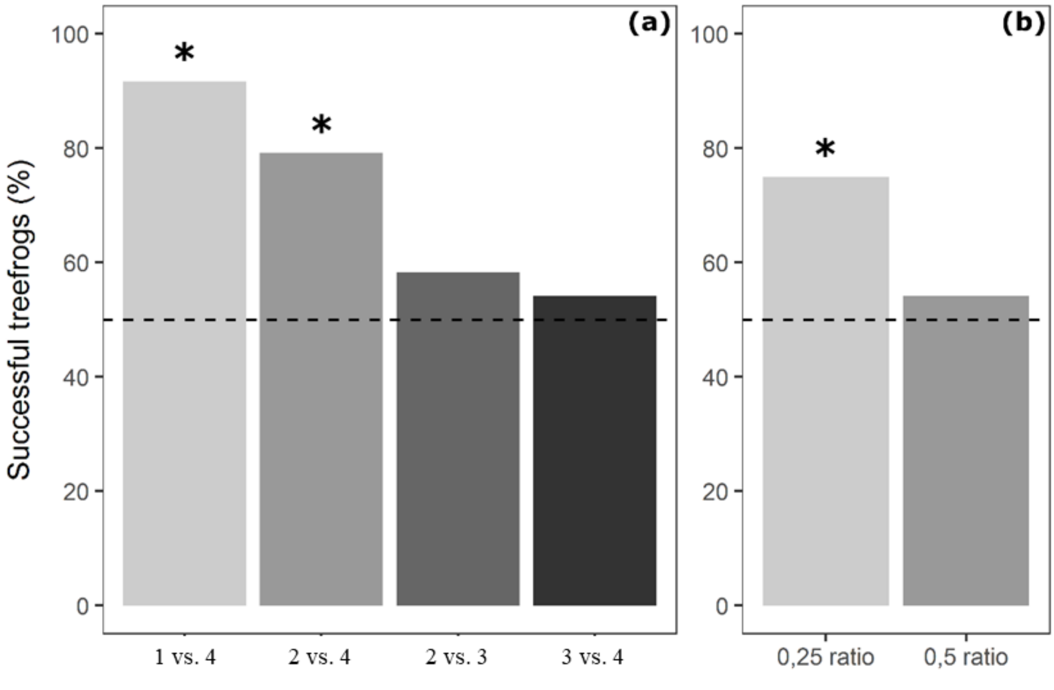


Fig. 15. Percentage of treefrogs choosing (a) the set with the larger number of bars in experiment 5 and (b) the larger bar in experiment 6. Dashed line indicates the chance level and asterisks indicate significant deviations from chance ($P < 0.05$).

4.4.6 Experiment 6: Size Discrimination

The number of frogs that chose the larger bar was significantly above chance in the 0.25 ratio (18/24; χ^2_1 = 6.000, *P* = 0.014; Fig. 15b) but not in the 0.5 ratio (13/24; χ^2 = 0.167, *P* = 0.683; Fig. 15b).

When considering both size ratios, the analysis with the GLM showed that the number of frogs choosing the stimulus with the larger bar (31/48, 64.58%) was significantly greater than chance (estimate = 0.601, SE = 0.302, z = 1.991, *P* = 0.047). The GLM did not find a significant difference between the ratios (χ^2_1 = 2.303, *P* = 0.129).

4.4.7 Experiment 7: Attributes Exploited in Discrete Quantity Discrimination

In experiment 7a (stimuli controlled for overall surface area), 17 of 24 frogs chose the stimulus with the larger number of bars; in experiment 7b (stimuli controlled for the convex hull of the bars), 20 of 24 frogs chose the stimulus with the larger number of bars. In both tests, the number of frogs that chose the stimulus with the larger numerosity was significantly greater than chance (overall surface area: χ^2_1 = 4.167, *P* = 0.041; convex hull: χ^2_1 = 10.667, *P* = 0.001; Fig. 16a), suggesting that frogs did not discriminate between the two sets of bars based on overall surface area or convex hull.

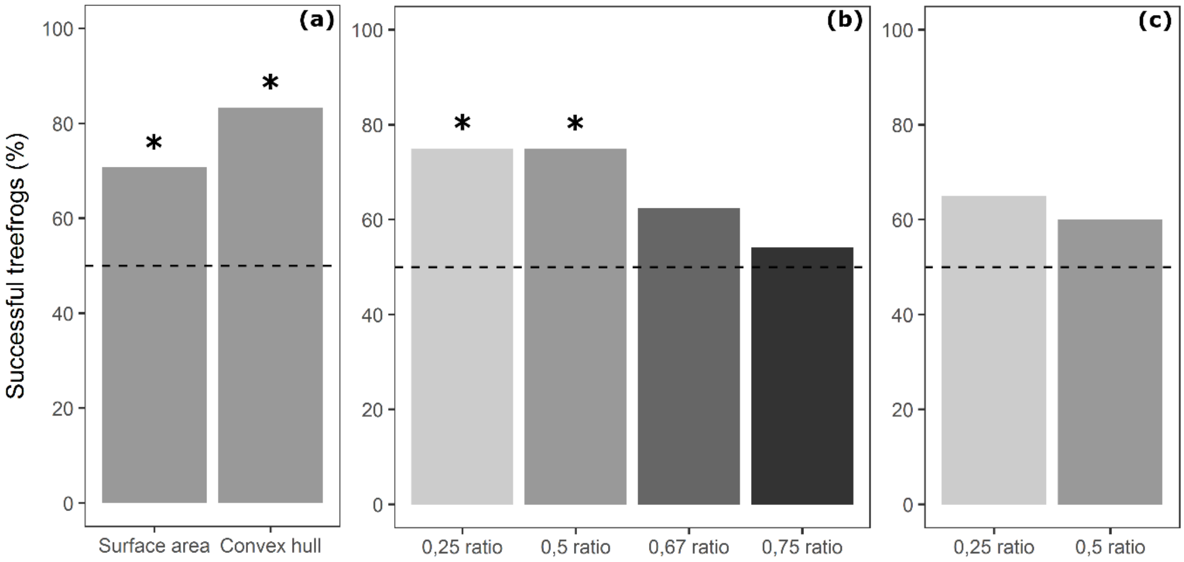


Fig. 16. Percentage of treefrogs choosing (a) the set with the larger number of bars in experiment 7, when the stimuli were corrected for overall surface area and convex hull, and the larger bar in (b) experiment 8a and (c) experiment 8b. Dashed line indicates chance level and asterisks indicate significant deviations from chance (*P* < 0.05).

4.4.8 Experiment 8: Attributes Exploited in Discriminating Size

When the two bars differed in height (experiment 8a), the number of frogs that chose the larger bar was greater than chance for the 0.25 (18/24; $\chi^2_1 = 6.000$, $P = 0.014$) and 0.5 ratios (18/24; $\chi^2_1 = 0.800$, $P = 0.014$), but not for the 0.67 (15/24; $\chi^2_1 = 1.500$, $P = 0.221$) and 0.75 ratios (13/24; $\chi^2_1 = 0.167$, $P = 0.683$; Fig. 16b). Considering all height ratios, the analysis with the GLM showed that the number of frogs choosing the taller bar (64/96, 66.67%) was significantly greater than chance (estimate = 0.693, SE = 0.217, z = 3.202, $P = 0.001$). The linear trend was close to the threshold for statistical significance ($P = 0.076$). The GLM did not find a significant difference between the ratios ($\chi^2_3 = 3.771$, $P = 0.287$).

When the two bars differed in width (experiment 8b), the number of frogs choosing the larger bar did not differ from chance (ratio 0.25: 13/20; $\chi^2_1 = 1.800$, $P = 0.180$; ratio 0.5: 12/20; $\chi^2_1 = 0.800$, $P = 0.371$; Fig. 16c); this suggests that frogs do not base their choice between different-sized bars on bar width. Similarly, the analysis on both width ratios with the GLM showed that the number of frogs choosing the wider bar (25/40, 62.50%) was not significantly greater than chance (estimate = 0.511, SE = 0.327, z = 1.564, P = 0.118). The GLM did not find a significant difference between the ratios (χ^2_1 = 0.107, P = 0.744).

4.4.9 Speed-Accuracy Trade-off

In the GLM model to study speed and accuracy, we found a negative relationship between the likelihood of choosing the larger stimulus and time taken to make the choice (χ^2_1 = 11.190, P < 0.001; Fig. 17), suggesting a speed-accuracy trade-off. There was no significant effect of experiment (χ^2_5 = 6.318, P = 0.277).

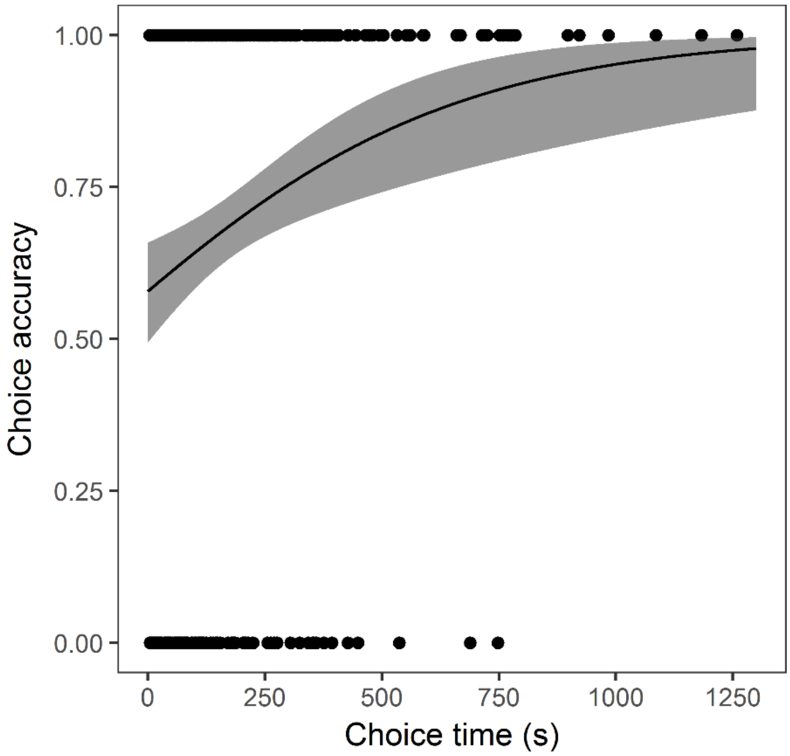


Fig. 17. Accuracy in choosing the larger quantity as a function of time taken to make the choice. Points represent the choice made by the subjects (larger or smaller quantity) versus time taken to make the choice in experiments 5, 6, 7 and 8; the line and the grey shaded area represent the speed-accuracy relationship and the confidence interval predicted by a generalized linear model with binomial error distribution and logit link function.

4.5 Discussion

Many animal species can discriminate discrete and continuous quantities in different ecological contexts (e.g. foraging, mate choice, social interactions). We have limited knowledge on how and in which contexts amphibians use quantitative information. Our experiments revealed that treefrogs rely on quantitative abilities to choose microhabitats: they showed a spontaneous preference for larger grass clumps and, by studying this behaviour in the laboratory using a dichotomous choice test, we showed that they can discriminate between numbers and heights of objects simulating vegetation.

In four experiments in the field, we investigated the possibility of studying quantity discrimination during microhabitat choice by treefrogs. Experiment 1 indicates that treefrogs are attracted to microhabitats with abundant and tall grass and that they tend to climb vegetation, as previously reported for closely related species (Ildos & Ancona, 1994; Michaels et al., 2014; Stewart, 1985). When presented with a dichotomous choice between different-sized grass clumps (experiment 2), treefrogs showed a preference for the larger one. This choice behaviour is in line with that observed in reptile and fish species (Bartholomew, 2012; Cooper & Whiting, 2000) and might be used to study quantity discrimination, provided that the stimuli can be finely controlled. In experiment 3, we found that treefrogs are also attracted by green printed bars on a white background, and that they climb them as observed for real plants. The subjects' attraction to the printed stimulus bars appears to be similar to their response to real grass (experiment 4). The green bars might be perceived similar to the grass or they might allow crypsis. More importantly for this study, the bars are stimuli that can be easily controlled and used in a laboratory setting to study quantity discrimination. The main advantage of this approach is that it is based on a spontaneous behaviour; thus, the ability shown by subjects probably resembles that expressed by the species in nature (Agrillo & Bisazza, 2014).

In the remaining experiments of this study, we focused on quantity discrimination abilities and mechanisms. In experiment 5, treefrogs presented with two sets of same-sized vertical green bars chose the set with a larger numerosity, discriminating significantly up to the 2 vs. 4 bars discrimination (0.5 numerical ratio). This discrete quantity discrimination could be achieved by representing either the number of bars or the continuous variables that covary with numerosity (Davis & Pérusse, 1988). In the two conditions of experiment 7, we separately controlled the stimuli for the two more important continuous variables that covary with numbers (overall surface area and convex hull). This did not prevent the treefrogs from identifying the set containing more bars suggesting they use numerical information to discriminate between the available options. There is an ongoing debate about the importance of numerical information in discrete quantity discrimination: some studies suggest that animals spontaneously tend to use continuous variables, and that they use numerical information as a ‘last resort’, when prevented from using continuous variables (Vos et al., 1988). Other studies align with the present study in suggesting, instead, that animals spontaneously use numerical information (Ferrigno et al., 2017). Regarding amphibians, salamanders' choice of the larger group of live prey seems to be driven by quantity of movement (Krusche et al., 2010); toads seem to spontaneously use numbers to discriminate prey groups, at least when the number of preys per group does not exceed four (Stancher et al., 2015).

In experiment 5, we also found a clear ratio effect indicating that treefrogs were more likely to choose the larger numerosity when the ratio between the number of bars was smaller. This finding supports the existence of an approximate number system for discrimination of small numerosities with an accuracy set by Weber's law (Cantlon & Brannon, 2007). When each numerosity ratio was tested separately, the maximum accuracy exhibited by the treefrogs in discrete quantity discrimination (2 vs. 4: 0.5 ratio) was lower than previously reported in other anurans when choosing prey (e.g. 2 vs. 3: 0.67 ratio; Stancher et al., 2015). There are at least four possible explanations for this result. First, there might be intraspecific differences in cognitive abilities (Clayton & Krebs, 1994; Day et al.,

1999) and different anuran species might have a different numerical ability. Second, anurans might perform differently in different tasks (e.g. in choosing between microhabitats as opposed to prey) because the different tasks are of different ecological relevance. For example, a social fish species, the guppy, *Poecilia reticulata,* shows higher discrimination abilities when choosing between shoals comprising different numbers of conspecifics (0.8 ratio) than when choosing between groups of food items (0.5 ratio; Lucon-Xiccato et al., 2015; Lucon-Xiccato et al., 2016; Chapter 4); this might be because choosing the larger social group is one of the main antipredator defences of social fish (Seghers, 1974). Third, it is possible that anurans' numerical accuracy improves with age due to experience or ontogenetic maturation of the nervous system (Bisazza et al., 2010). Accordingly, the reduced performance of our treefrogs could be because they were very young individuals. One last possibility is that anurans show different motivation in the different tasks. Treefrogs might not exhibit a preference between microhabitats with small differences in number of plants because the choice confers limited advantages. This does not exclude the possibility that treefrogs perceive the difference between stimuli. The issue of motivation is typical of procedures based on spontaneous choices; future studies should try to address it by using discrimination learning procedures (Agrillo & Bisazza, 2014).

Experiment 6 investigated discrimination of continuous quantities (i.e. surface areas). Treefrogs presented with two bars of different size (one taller and wider than the other) showed a preference for the larger one if the size ratio was 0.25, but not if the ratio was larger, suggesting that they discriminate sizes less well than numbers or that they are differentially motivated in choosing between different numbers of grass leaves and different-size leaves. In experiment 8, the heights and the widths of the bars were compared separately to see if one attribute was more important than the other. The treefrogs were accurate in discriminating height up to a ratio of 0.5 (coinciding with the most accurate discrimination in experiment 5), but they failed to discriminate bars of different width even with an easier 0.25 ratio. These results suggest that treefrogs are able to compare and discriminate continuous quantities as well as discrete quantities. Further, the results suggest that when choosing the larger of two objects, as in experiment 6, treefrogs choose the taller one but do not attend to differences in width between them. Indirect evidence of continuous quantity discrimination in amphibians has previously been provided by mate choice experiments: for example, male Andrew's toads, *Bufo andrewsi*, show a mating preference for larger females (e.g. Arntzen, 1999; Liao & Lu, 2009).

Overall, our study shows that treefrogs prefer larger clumps of vegetation and taller plants, a preference similar to that observed in other species (Bartholomew, 2012; Cooper & Whiting, 2000; Mensforth & Bull, 2008; Takahashi & Nagayama, 2016). Our study aimed to investigate the cognitive system underlying quantity discrimination and we did not address the functional significance of the behaviour that we observed. One can speculate that it might have evolved to avoid predators (Babbitt & Tanner, 1997). For a species such as the treefrog, which relies on cryptic coloration to defend itself, being in a large clump of vegetation is likely to offer greater protection. Likewise, treefrogs are arguably better protected against terrestrial predators when they climb taller plants. On the other hand, the ability to choose the larger clumps of vegetation and the taller plants might have evolved because it is advantageous during ambush feeding (Walsh & Downie, 2005), and/or because it favours homeostasis (Seebacher & Alford, 2002). In line with this idea, there is evidence that the choice of microhabitats with rich vegetation increases growth in treefrogs (Michaels et al., 2014). Testing these hypotheses will require field experiments to assess fitness advantages of microhabitat selection.

The last goal of our study was to investigate the effects of decision speed on accuracy in quantity discrimination. Decision speed is an important, but scarcely considered factor in cognitive performance (Chittka et al., 2009). Making accurate decisions often demands lengthy sampling times, which can be associated with costs. In our experiments, we found that long latencies were associated with a greater preference for the larger bar/set of bars. Previous studies have often considered similar results as an evidence of a speed-accuracy trade-off (e.g. Chang et al., 2016; Lucon-Xiccato & Bisazza, 2016). Although our study did not provide direct evidence that choice time reflects time required for cognitive processing, our finding might indeed be another example of this pervasive association. In the case of treefrogs, the cost of lengthy sampling time might be that individuals spend more time outside the cover of vegetation and consequently are exposed to potential predators for longer.

Speed-accuracy trade-offs have been reported in various decision-making processes in animals (e.g. Chittka et al., 2003; Latty & Beekman, 2011; Wang et al., 2015), but it is only in humans that there is clear evidence of this phenomenon in quantitative tasks (Moyer & Landauer, 1967). In a recent experiment on fish's shoal size discrimination abilities, we found that guppies switched more frequently between two shoals of conspecifics with increasing ratio between shoals' numerosity (Chapter 4). This finding can be interpreted as indirect evidence of a longer sampling time being needed to perform challenging quantity discriminations. Taken together, the experiment with guppies and the present study with treefrogs suggest that speed-accuracy trade-offs probably affect quantity-based decision making in nonhuman animals, too. Future research should directly study this trade-off and try to understand its causes (Chang et al., 2016; Lucon-Xiccato & Bisazza, 2016) and its ecological consequences (Chittka et al., 2003).

5. Study 4: spontaneous quantity discrimination in crickets

5.1 Abstract

Recent studies have reported that the ability to discriminate among quantities is not a prerogative of vertebrates. Ants, bees, and spiders can solve tasks in which they are required to discriminate between groups of objects. However, it is still not clear which mechanisms are used by invertebrates to discriminate between quantities, because few studies have controlled the stimuli for the continuous variables that covary with numbers. Here, I investigated the spontaneous preferences of an arthropod species, *Acheta domesticus*, for both discrete and continuous quantities when searching for a potential shelter. I exploited the natural shelter-seeking behaviour of this species by presenting sets of geometric shapes (e.g., square and rectangles) that simulated potential shelter in a dangerous context. In a dichotomous choice between sets of black shapes differing in number of items, crickets showed a preference for the set containing the larger numerosity up to 2 vs. 3 items (numerical ratio: 0.66). When stimuli were controlled for continuous variables (convex hull and cumulative surface area), crickets no longer showed a preference for larger sets of 2 vs. 3 items. When subjects were presented with a dichotomous choice between two squares differing in area, a significant number of crickets selected the larger stimuli when the size ratio was 0.25. Control experiments indicated that the results in size discrimination were poor because crickets attend to the width but not the height of a stimulus. These findings suggest

that *A. domesticus* possess a system to discriminate between quantity by summing the total area of presented items and that they rely more on some features (e.g., width) than others to make adaptive decisions.

5.2 Introduction

Animals face many situations in which being able to discriminate between quantities could have a remarkable adaptive value. Such abilities allow individuals to select the most beneficial choice among different options to maximize their fitness in relation to the ecological context (Shettleworth, 2010). For instance, animals maximize the quantity of food in relation to the cost of finding a food source by selecting the larger amount of food (Hauser et al., 2000; Lucon-Xiccato et al., 2015) or the option with minimal costs (Pantaleeva et al., 2013). Social species such as chimps, hyenas, and dogs adopt different strategies in competitive behaviour in relation to the number of potential competitors (McComb et al., 1994; Bensom-Amram et al., 2011; Bonanni et al., 2010). In an unfamiliar and dangerous environment, prey species join the larger group of conspecifics to minimize predatory risks (Hager & Helfman, 1991; Gomez-Laplaza & Gerlai, 2011; Chapter 4).

Many studies have reported that humans and other vertebrate species share two distinct non-symbolic systems to represent numerical quantity: the object-file system (OFS) and the approximate number system (ANS). The OFS is involved in the mechanism of “subitizing”: it is an object-based attentive mechanism that permits one to identify the element in a small numerosity as a unique symbol stored in working memory (Trick & Pylyshyn, 1993). This system is triggered when facing sets of small numerosity (3-4 elements). OFS accuracy is independent of the numerical ratio but depends on intrinsic memory limits and/or attentional constraints (Burr et al., 2010; Hyde, 2011). The ANS is involved in the estimation process of large numerosity (>4 items). In contrast to the OFS, the ANS does not present an absolute upper limit of elements, but the accuracy of estimated numerosity is dependent on the numerical ratio between the elements to be discriminated. The imprecision of ANS is due to the distance effect (e.g., the discrimination of 5 vs. 10 items is easier than 5 vs. 6 items) and the size effect (e.g., the discrimination of 5 vs. 6 is easier than 10 vs. 11 items; Gallistel & Gelman, 1992). Studies on numerical cognition in human adults and, partially confirmed, in infants support the idea that humans possess two distinct non-symbolic quantificational systems for representing small and large number (reviewed in vanMarle, 2015), while comparative psychologists debate on whether non-human animals possess a single or two distinct non-verbal quantification systems (Chapter 1.1.1).

In recent years, research on numerical cognition in invertebrates has raised growing interest among comparative psychologists. This interest is mainly due to the fact that invertebrates show rudimental numerical abilities that partially resemble those of vertebrates (reviewed in Pahl et al., 2013). Bees and ants use proto-counting abilities during navigation (Chittka & Geiger, 1995a; Wittlinger et al., 2006). Ladybirds, ants, beetles, and spiders assess the number of potential competitors when searching for food sources or selecting potential mates to avoid conflicts (Nelson & Jackson 2012; Hemptinne et al. 1992; Tanner 2006; Carazo et al. 2012). However, the mechanisms underlying quantity discrimination abilities in invertebrates are still far from being understood. Unfortunately, our knowledge of quantity ability mechanisms is restricted to some social insect species. Bees are able to discriminate up to 3 vs. 4 items and understand the concept of zero under control conditions (Dacke & Srinivasan, 2008; Gross et al., 2009; Howard et al., 2018) while ants were found to discriminate among nest sites by relying on the summation of light levels within the nesting entrances rather than counting the number of available cavity entrances (Franks et al., 2006).

Further studies are necessary to understand the similarities and differences in numerical cognition among invertebrate species.

In the present study, I investigated which mechanisms an invertebrate species, the arthropod *Acheta domesticus*, uses to discriminate between quantities when searching for shelter. I focused on the spontaneous preference for shelters by performing five laboratory experiments in which *A. domesticus* were presented with a dichotomous choice between geometric shapes. *A. domesticus* is a small nocturnal cricket that actively searches for borrows as refuges to escape from predators (Kieruzel, 1976). Previous studies have reported its spontaneous orientation toward dark stimuli under bright light conditions (Atkins et al., 1987; Kieruzel 1976). However, no study has investigated the preference between groups of geometric shapes with different numbers of items, even if there was evidence of spontaneous ability to discriminate the duration of male's chirps (e.g., females prefer male's chirps with a greater number of pulses per chirp; Gray, 1997). In our experiments, I positioned crickets in the centre of a white arena with dark geometric shapes (e.g., squares and rectangle) attached to the wall from which they could choose. I varied the number of shapes contained in each set and the size of shapes to investigate spontaneous preference for quantities in crickets.

I conducted a preliminary experiment to verify the spontaneous reactions to our experimental conditions and stimuli attraction. Crickets were presented with a first discrimination between a black square and a completely white paper in order to study the attraction for our stimuli. I prepared stimuli based on the information on visual acuity reported by Atkins and colleagues in similar experimental conditions (1967). Secondly, I presented a discrimination between one black square and three same-sized squares to verify the preference towards the set of figures.

In experiment 2, I investigated quantity discrimination abilities in crickets by presenting a choice between sets of same-sized squares differing in the number of items. I administered four discriminations: 1 vs. 4 (ratio 0.25); 2 vs. 4 (ratio 0.50); 2 vs. 3 (ratio 0.67); and 3 vs. 4 (ratio 0.75). I hypothesized that crickets would choose the larger stimulus because the possibility of finding better shelter increases with the number of available refuges. In experiment 3, I presented two discriminations of 2 vs. 3 items in which I controlled the stimuli for two continuous variables that covaried with numerosity: the convex hull and the cumulative surface area.

Results from experiments 2 and 3 suggested that crickets used the total area more than numerical information to discriminate between quantities. I hypothesized that crickets possess a system that permit them to estimate the total area of stimuli by summing the amount of area for each figure. Thus, in our experiments, they might have oriented towards stimuli that presented the larger total area of black. I expected that crickets would show similar accuracy when discriminating between geometrical shapes differing by size; in the last two experiments, I presented a dichotomous choice between geometrical shapes differing by size. In experiment 4, crickets were presented with a discrimination between black squares differing in size (size ratio: 0.25; 0.5; 0.67; and 0.75). Results showed that crickets were less accurate in discriminating stimuli differing by the size than by number of items. It is possible that one attribute is more relevant for crickets when discriminating the size of figures. Atkins and colleagues (1987) reported that crickets showed less responsiveness towards black figures when the height of figures was equal or larger than the width. Two species of wood crickets (e.g., *Oecanthus pellucens,* and *Nemobius sylvestris*) showed different preferences for visual cues according to their natural behaviour (e.g., high vegetation and dead life on the ground, respectively; Campman et al., 1987). In experiment 5, I presented a choice between rectangles with a size ratio of 0.67 by modifying only one dimension of the stimuli (e.g., height or width) while the other dimension was kept constant in order to investigate whether *A. domesticus* was influenced more by height or width when discriminating between stimuli.

5.3 Methods

Subjects

Adult individuals of the species *A. domesticus* (Linnaeus 1758) were purchased from a local shop (Agri Pet Garden, Padua, Italy). Crickets were maintained in the Animal Behaviour and Cognition Lab in the Department of General Psychology (University of Padua, Italy) in terraria 19 x 50 x 32 cm (width x length x height). Each terrarium was provided with wood shavings, egg boxes, a full water jar to maintain a high humidity level, and a 50-W fluorescent infra-red lamp (12:12 light/dark cycle). The temperature was kept at 25 °C ± 3. The tank was covered by a transparent plexiglass sheet with holes for ventilation. Crickets were fed with fresh salad, carrots, and commercial food pellets three times per week. The tank was cleaned every two days to avoid the proliferation of diseases and parasites.

Apparatus

The apparatus (Fig. 18a) was similar to one used recently to investigate quantity discrimination ability in treefrogs (Chapter 5). The apparatus consisted of a white plastic circular arena (Ø 80 cm; 75 cm) situated in a dark room. The light was provided by a single LED lamp (450 lumen, opening angle 100) placed 80 cm above the arena. A transparent start box 10 x 13 x 8 cm was positioned in the centre of the arena and orientated toward the stimuli.

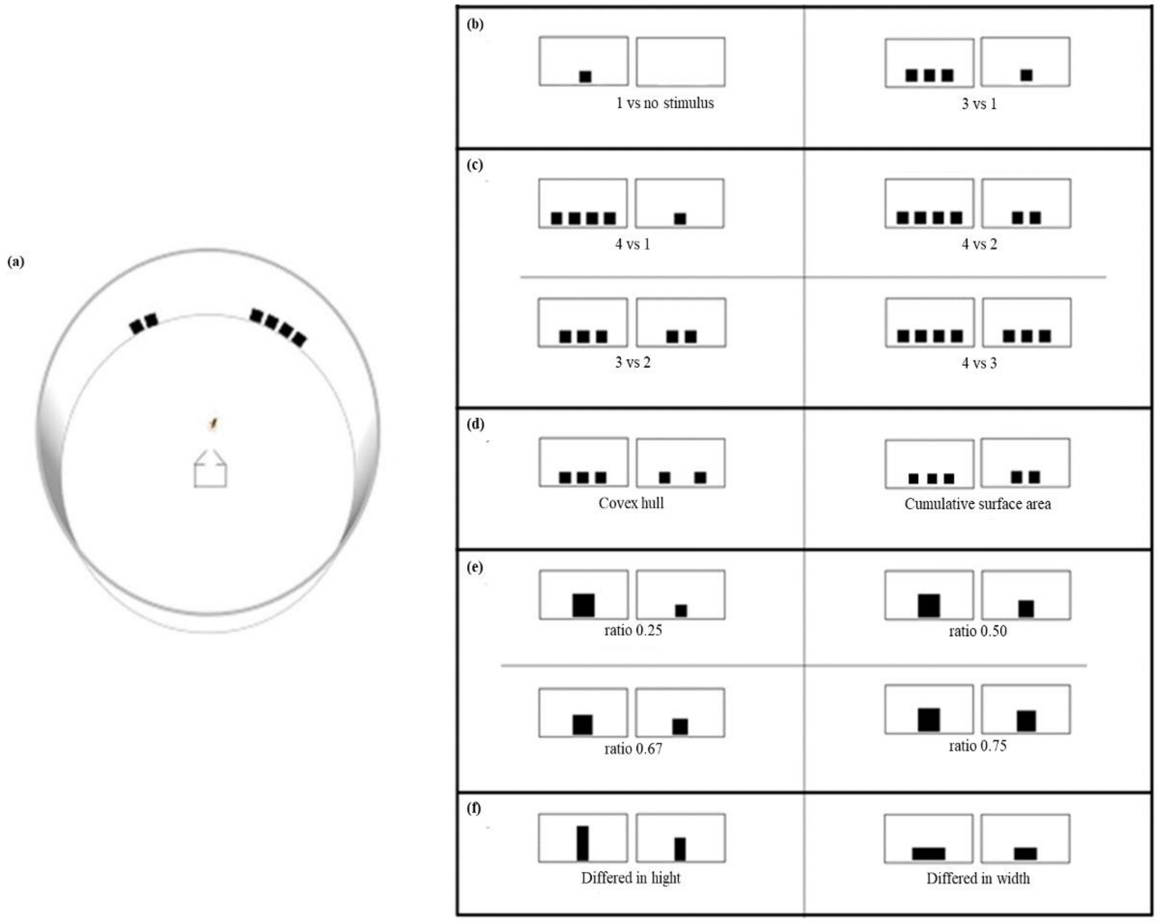


Fig. 18. (a) Top view of apparatus adopted in our study; and (b) stimuli used in experiment 1, (c) experiment 2, (d) experiment 3, (e) experiment 4, and (f) experiment 5.

The stimuli (Fig. 18 b-f; see below for details) were groups of black geometrical shapes printed on A3 white paper. Stimuli varied in number and size depending on the experiment. The centre of the printed papers formed an angle of 60° with the centre of the arena; this set up permitted crickets to see both stimuli at the same time. Previous studies reported that immobile crickets can detect different stimuli in a 180° visual field, and mobile crickets scan the environment at a visual angle of 60° either side (Weber et al., 1987). The position of stimuli in the arena was randomized for each subject by changing the left-right presentation of the two stimuli and by changing the position of presented stimuli on the wall of arena.

Procedure

I performed the experiments between 08:00 and 14:00. A single cricket was collected from the maintenance tank and gently positioned inside the start box. The trial started when the subject left the start box. I considered a choice to have occurred when the subject reached one of the two stimuli. After each test, the arena was cleaned with 99% ethanol to remove any chemical cue. Subjects that moved toward the opposite side of arena of where stimuli were presented or never moved in 20 minutes were discarded and substituted with new ones. I aimed to test 55 crickets in experiment 1, 100 crickets in experiment 2, 60 crickets in experiment 3, 100 crickets in experiment 4, and 60 crickets in experiment 5. However, I discarded 37 crickets in experiment 1, 110 in experiment 2, 20 in experiment 3, 105 in experiment 4, and 96 in experiment 5. The proportion of discarded subjects (49.53%) is common to studies based on spontaneous choice tests (Uller et al., 2003; Krusche et al., 2010). Each subject was tested once and, following the conclusion of the experiment, was released into a separate terrarium.

I performed a preliminary experiment (1) to test whether crickets were attracted to black figures with our setting. I presented a discrimination (1a) between one 7×7 cm black figure and a white paper sheet (Fig. 18b), and a second discrimination (1b) between a set of three 7×7 cm black figures vs. a single 7 x 7 cm black figure (Fig. 18b). I tested 30 subjects for discrimination 1a and 25 subjects for discrimination 1b.

Experiments 2 and 3 aimed at assessing the discrete quantity discrimination abilities of crickets. Stimuli were composed of sets of black figures with different numerosities. In experiment 2 (Fig. 18c), I presented four numerical discriminations to study the limits of quantity discrimination ability: 1 vs. 4 (numerical ratio: 0.25); 2 vs. 4 (numerical ratio: 0.50); 2 vs. 3 (numerical ratio: 0.67); and 3 vs. 4 (numerical ratio 0.75). Each item was a 7×7 cm black square printed on a white paper sheet. In each set, figures were separated by a gap of 3 cm. I set 25 subjects as the limit number for each discrimination. In experiment 3 (Fig. 18d), I separately studied the effects of two continuous variables, such as convex hull and cumulative surface area, on crickets' quantity discrimination ability. For both discriminations, I presented a dichotomous choice between 2 vs. 3 sets of same-sized squares as the limit found in experiment 2. In the convex hull discrimination (3a), crickets were presented with two 7×7 cm squares separated by 13 cm and three 7×7 cm squares separated by 3 cm. The external area occupied by each set was 27 cm. In the cumulative surface area discrimination (3b), crickets were presented with three 5.72×5.72 cm squares separated by 3 cm, and two 7×7 cm squares separated by 3 cm. The cumulative area occupied by each set was 98 cm^2. I tested 30 subjects for each discrimination.

Experiments 4 and 5 aimed at assessing size discrimination abilities of crickets by presenting discriminations between single geometrical shapes differing in size. In experiment 4 (Fig. 18e), I presented four discriminations between single black squares that differed in size proportionally to the numerical ratio used in experiment 2. For each discrimination, I considered one stimulus as the reference (dimension: 14×14 cm) and the other one was modified both in height and in width according to the size ratio: 7×7 cm (size ratio: 0.25); 9.89 ×9.89 cm (size ratio: 0.5); 11.46×11.46 cm (size ratio: 0.67); and 12.12×12.12 cm (size ratio: 0.75). I tested 25 subjects for each discrimination. In experiment 5 (Fig. 18f), I modified only one dimension of the stimuli (width or height) in order to study which attributes were more relevant in quantity discrimination. For both discriminations, I presented rectangles differing by a size ratio of 0.67 as the ratio limit found in experiment 2. In the first discrimination (5a), I presented two figures with different heights but equal widths: a 7×14 cm figure and a 7×21 cm figure. In the second discrimination (5b), I presented two figures with different widths but equal heights: a 14×7 cm figure, and a 21×7 cm. I tested 30 subjects for each discrimination.

Statistical analysis

Statistical analysis was performed in R version 3.2.0 (The R Foundation for Statistical Computing, Vienna, Austria, http://www.r-project.org). In this study, I considered the choice of the larger stimulus as the correct choice (e.g. sets of squares with the larger numerosity in experiments 2 and 3 and the larger size geometrical shapes in experiments 4 and 5). I compared the observed number of crickets that made the correct choice with the total number of crickets that chose a stimulus (binomial exact test; probability of random choices equals to 50%; Hollander & Wolfer, 1973). Statistical tests were two-tailed, and the significance threshold was $P = 0.05$.

For experiments with different quantity ratios (experiments 2 and 4), I analysed the binary choice of subjects with a general linear model (GLM) with binomial errors distribution and logit link function (Hastie & Pregibon, 1992). I initially fitted the model with intercept only to study the overall preference of crickets for larger stimuli independently from the ratio; then I fitted ratio as a factor to test differences due to numerical ratio or size ratio according to the experiment. To evaluate the differences due to the factor (ratio), I used the function ANOVA from library CAR (Fox & Weisberg, 2011). When the factor (ratio) had a significant effect on response, I conducted a pairwise comparison (Westfall et al., 2011), using the library lsmeans (Lenth, 2016).

5.4 Results

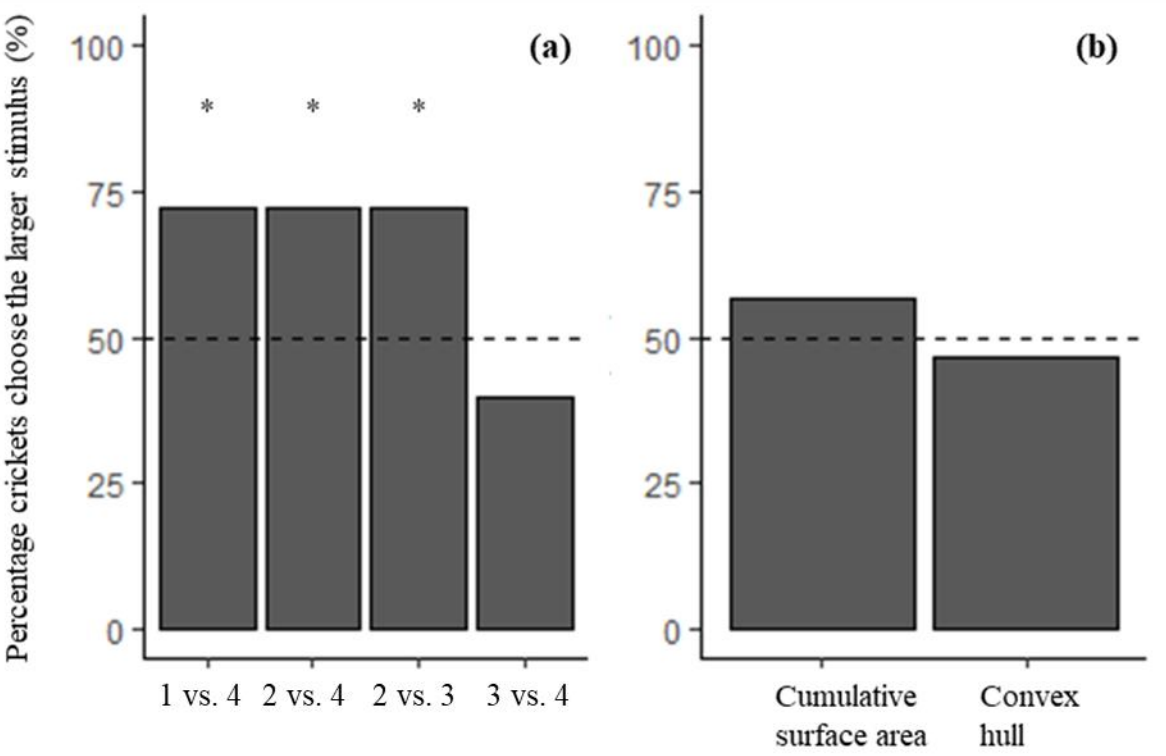
5.4.1 Experiment 1 – Spontaneous preference for dark figures

In discrimination 1a, the number of crickets (22 out of 30) that chose the stimulus containing the single black squares rather than the stimulus without a geometrical shapes was significantly above chance ($P = 0.016$); in discrimination 1b, the number of crickets (18 out of 25) that chose the sets containing the larger number of squares was significantly above chance ($P = 0.043$).

5.4.2 Experiment 2 – Quantity discrimination in crickets

In the 1 vs. 4 item discrimination, the number crickets (18 out of 25) that chose the set containing the larger number of squares was significantly above chance ($P = 0.043$). A similar result was found in the discrimination of 2 vs. 4 (18 out of 25 chose the larger set; $P = 0.043$) and in the discrimination of 2 vs. 3 (18 out of 25 chose the larger set; $P = 0.043$). In the 3 vs. 4 discrimination, the number of subjects (10 out of 25) that chose the set with the larger number of squares was not different from chance ($P = 0.424$; Fig. 19a).

Overall, crickets spontaneously orientated significantly more towards the set containing the larger number of squares (64%; estimate = 0.575; SE = 0.208; $t = 2.762$; $P < 0.006$). The GLM showed a significant difference among the numerical ratios ($\chi^2_3=8.090$; $P = 0.044$). The post hoc pairwise comparison showed that the difference between the discrimination of 3 vs. 4 and the other discriminations was significant (all $P = 0.016$), but there was not a significant differen



ce among other numerical discriminations (all $P = 1.000$).

Fig. 19. Percentage of crickets choosing (a) the set with the larger number of squares in experiment 2 and (b) the set with larger number of squares in experiment 3. Dashed lines indicate chance level and asterisks indicate significant deviation from chance ($P = 0.05$).

5.4.3 Experiment 3 – Controlling for non-continues variables

When controlling stimuli for convex hull (discrimination 3a), the number of subjects that chose the larger set (14 out of 30) was not significantly above chance ($P = 0.856$; Fig. 20b). When controlling stimuli for cumulative surface area (discrimination 3b), the number of subjects that chose the larger set (17 out of 30) was not significantly above chance ($P = 0.585$; Fig. 19b).

5.4.4 Experiment 4 – Size discrimination

In the first size discrimination (ratio 0.25), the number of crickets (19 out of 25) that chose the square with larger size was significantly above chance ($P = 0.015$), whereas in the other three size discriminations, the number of crickets did not show a preference for one of the two squares (size ratio 0.50: 12 out of 25, *P* = 1; size ratio 0.67: 15 out of 25, *P* = 0.424; size ratio 0.75: 15 out of 25, *P* = 0.424; Fig. 20a).

Overall, crickets spontaneously orientated significantly more towards the square with the larger size (61%; estimate = 0.447; SE = 0.205; $t = 2.182$; $P < 0.029$). The GLM did not show a significant difference among the size ratios (χ^2_3=4.277; $P = 0.233$).

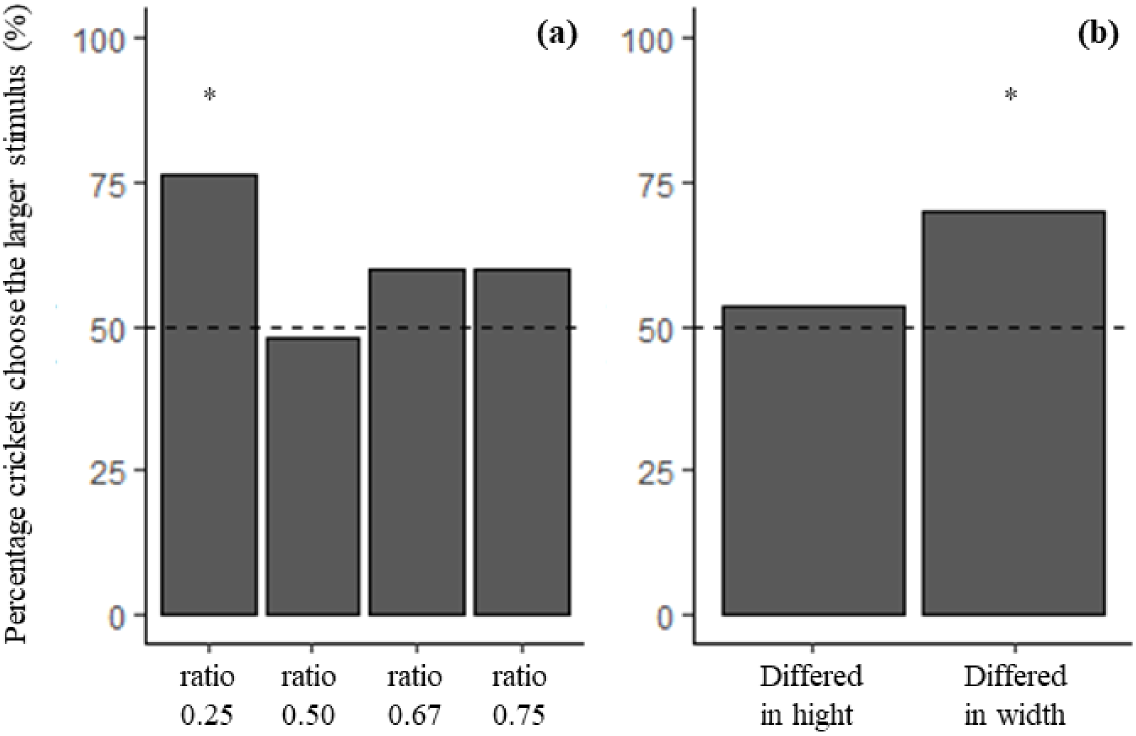


Fig. 20. Percentage of crickets choosing (a) the square with the larger size in experiment 4 and (b) the rectangle with the larger size in experiment 5. Dashed lines indicate chance level and asterisks indicate significant deviation from chance (*P* = 0.05).

5.4.5 Experiment 5 – Attributes exploited in size discrimination

When the two figures differed in height, the number of subjects that chose the higher rectangle (16 out of 30) was not significantly above chance (*P* =0. 856; Fig. 20b); whereas when stimuli differed in width, the number of subjects that chose the larger rectangle (21 out of 30) was significantly above chance (*P* = 0.043; Fig. 20b).

5.5 Discussion

Our experiments showed that *Acheta domesticus* used quantitative discrimination abilities when searching for potential refuges. *A. domesticus* showed a spontaneous preference for the sets containing the larger number of figures and a preference for the bigger figures. Our results suggested that this species makes quantity discrimination decisions without relying on numerical information.

In experiment 1, crickets showed a spontaneous preference for a set of stimuli (3 items) vs. a single figure. This indicates that crickets are attracted to the set of figures that I used as stimuli and that, therefore, I can use this method to try to study quantity discriminations in this invertebrate. In experiment 2, I focused on studying the quantity discrimination ability by presenting dichotomous choices between sets of same-sized figures differing by numerosity in function of the ratio (numerical ratio: 0.25; 0.50; 0.67; 0.75). Stimuli were black figures that simulated groups of burrows where crickets could find shelter. Results showed that crickets spontaneously chose the larger set of geometric shape in all comparisons except for the most challenging one, 3 vs. 4. The limit of 3-4 items in quantity discrimination showed by crickets was consistent with previous findings on honeybees and spiders (e.g., Chittka & Geiger, 1995a; Cross & Jackson, 2017; but see Carazo et al., 2009; Yang & Chiao 2016).

When comparing the results obtained, the mechanisms underlying quantity discrimination abilities appear similar among invertebrates (review in Pahl et al., 2013). However, it is important to note that invertebrates have been studied with very different methods and few controls for continuous variables that may influence quantity discrimination abilities. I evaluated the effect of continuous variables on crickets' decisions between set of items. In experiment 3, stimuli were separately controlled for convex hull and cumulative surface area.

Crickets did not show a preference for a set of 2 items or 3 items, as shown in experiment 2. These results suggested that cricket base quantity discrimination decisions on continuous variables, such as the total area. Regardless of the exact perceptual cue involved, non-numerical information seems to be spontaneously preferred to numbers in a quantity discrimination task, as previously observed in some vertebrate (Agrillo et al., 2008; Feigenson et al., 2002; Beran et al., 2008; Tomonaga, 2008; Krusche et al., 2010; Killian et al., 2003), and invertebrate species (Franks et al., 2006).

Results from experiments 2 and 3 suggested that crickets make quantity discrimination decisions based on the total amount of area in each stimulus. It is possible that crickets possess a mechanism that permit them to discriminate between continuous quantities (e.g., measure the total area by summing each element containing in the set). I expected that crickets would show similar or higher accuracy when discriminating between figures differing by the size. In experiment 4, I presented four dichotomous choices between figures differing by size (size ratio: 0.25; 0.50; 0.67; 0.75). Despite our previous results (experiments 2 and 3), crickets showed a preference for the larger stimulus only in the easier size discrimination (size ratio 0.25). The results found in experiments 2 and 3 suggested that crickets may use the cumulative surface area or the convex hull to discriminate between quantities. If crickets relied on such variables, when presented with a dichotomous choice between figures differing by size, I would have expected to find higher accuracy when the figures differed by the same size than when stimuli differed in numerosity. Therefore, the results from experiment 4 did not support the prediction, because crickets were less accurate in discriminating sizes than numbers.

However, there was a conflict between experiments 3 and 4. In experiment 3, I presented sets of same-sized square by varying the area or space occupied by the set. In experiment 4, larger square was both taller and wider than the smaller figures. Moreover, when I quadrupled the area of a square, the single dimension (e.g., height and width) was doubled. There is evidence regarding other crickets to suppose that one dimension may be more relevant than another. The Italian tree cricket (*Oecanthus pellucens*), living on high vegetation, prefers orientating towards vertical figures, while another wood cricket (*Nemobius sylvestris*), living on dead leaves on the ground, prefers horizontal stimuli (Campman et al., 1987). In experiment 5, I separately presented a dichotomous choice between rectangles that differed by a size ratio of 0.67 by modifying the width or the height of figures. Crickets showed a spontaneous preference for the larger rectangles when differing in width, but not when rectangles differed in height. Although the present results strongly suggest that *A. domesticus* relied on the horizontal dimension of the geometrical shapes or to the lower part of the objects, I can't state whether this is due to its habitat's features, to a limit of its visual system, or to a selective attention to the ground level.

The ability to distinguish among different quantities may have evolved to enhance the survival of organisms in different ecological contexts, such as foraging, group conflicts, parental care and predator avoidance (Shettleworth, 2010). Although this study did not completely answer which mechanisms were used by an arthropod species, *A. domesticus*, to discriminate between quantities, our results might be another example of a species that has evolved this ability. Although traditionally, quantity discrimination ability has been regarded as a higher cognitive function, many invertebrate species show a basic system to make proper decisions by discriminating quantity (Skorupski et al., 2018). It is necessary to extend our numerical cognition's knowledge of other invertebrates to understand the origin of this capacity.

6. General discussion

In the past decades, a growing number of studies have documented numerical abilities in nonhuman animals. Yet, our knowledge on this cognitive function is still far from being completely understood, and the aim of this thesis is to fill in some of the gaps in our knowledge.

In the first part of this thesis, I investigate how methodological changes influence the assessments of cognitive functions within the same species. In the first study, I adopted an operant conditioning procedure to train a fish species, the guppy (*Poecilia reticulata*), to solve a numerical discrimination task. A recent study (Bisazza et al., 2014b) found that *P. reticulata* can be trained to discriminate 4 vs. 5 objects, even when stimuli are controlled for non-numerical variables. This performance is higher than observed in some mammals and birds. In that study, guppies were required to discriminate between two patches of small objects on the bottom of the tank that they could remove to find a food reward. It is not clear whether this species possesses exceptional numerical accuracy compared with the other vertebrates or whether its remarkable performance was due to a species-specific predisposition to solve this foraging task. To disentangle this issue, I studied guppies in the same numerical discrimination as the previous study using a more conventional procedure, an operant conditioning chamber and computer-generated stimuli similar to that developed for other vertebrates and recently used to study discrimination learning in *Danio rerio* (Parker et al., 2012a, b; Proulx et al., 2014; Wang et al., 2015). The results evidenced a very low performance compared to that reported by the same species but using a more naturalistic setting. In four subsequent experiments, I manipulated four factors that could explain the differences between the two experiments, namely the decision time, the length of training, the type of training schedule (introducing a correction procedure) and the type of stimuli used (bi- versus three-dimensional). The first three factors had no effect on numerical discrimination and performance only slightly improved when using concrete objects as stimulus (the same used in previous study), instead of bi-dimensional figures displayed on the computer screen.

In the second study, I set up a modified version of a common procedure used to study spontaneous quantity ability in social fish (shoal choice test). In previous studies, guppies spontaneously preferred the largest group between shoals of 4 vs. 8 fish, but no preference was found between shoals of 4 vs. 6 fish (e.g., Agrillo et al., 2012b; Bisazza et al., 2014a). The modified version of the shoal choice test adopted in this thesis show that *P. reticulata* enhance their numerical acuity by discriminating between 4 vs. 6 fish and 4 vs. 5 fish, a comparable performance to that one assessed with an operant conditioning procedure (Bisazza et al., 2014b).

The two studies agree in demonstrating that the method has an important role on evaluating the cognitive capacity of species. In the first study, a more artificial setting led to a reduction in the numerical performance compared to that reported in a previous study (Bisazza et al., 2014b). A possible explanation could be their exceptional cognitive performance is due to an innate predisposition to solve a specific task. Natural selection can select the evolution of learning predispositions in one species to solve specific ecological problems (Shettleworth, 1972). It is possible that the setting used by Bisazza and colleagues (2014b) could favor the inherent predisposition of learning in guppies, while the same species may have more difficulty in a more artificial context. In the second study, the introduction of minor methodological changes improved the determination of the numerical accuracy of fish in discriminating the numerosity of two social groups. In addition, with the improved method, shorter time are required to assess numerical accuracy in fish. A cognitive difference within-species that seemed related to the context in which it is studied, instead, proved to be a consequence of the different numerical accuracy assessments that were carried out with different methods.

Interspecific differences in numerical cognition are generally believed to be entirely due to differences in the complexity of the neural substrate that serve these functions. My experiments show that other factors may be important. Some variation may be due to an inherent predisposition to solve specific tasks or a consequence of the different procedures used to assess numerical abilities. Interspecific comparison thus requires to identify procedures that can be adapted to different species and different contexts. For example, future studies may develop an efficient automatic operant conditioning system for low vertebrates and invertebrates similar to ones commonly used for studying cognitive functions in primates, rodents, and birds. In this way, it may be possible to directly compare the performance among distant phylogenetical species.

When attempting to reconstruct the evolution of the numerical abilities in animals I face the problem that few species have been studied and most of them belong to the same taxonomic groups. The second part of my thesis aimed to fill in this gap. In the first of the two studies, I investigated both discrete and continuous quantity discrimination ability in an amphibian (*Hyla intermedia*). In this vertebrate class, only two species have been investigated in numerical cognition research (Krusche et al., 2010; Stancher et al., 2015; Uller et al., 2003). In the second study, I investigated this capacity in the cricket *A. domesticus*), which belongs to a phylum, arthropods, in which quantity discrimination ability has been investigated in a handful of species.

In treefrogs, I investigated whether they use discrete and continuous quantity information when choosing between microhabitats. Treefrogs were presented with a dichotomous choice between sets of green bars (simulated grass clumps) that differed in the numerosity or size of the items. When presenting a dichotomous choice between two sets differing in numerosity, treefrogs showed a preference for the larger set between 1 and 2 bars and 2 and 4 bars (0.5 ratio), but not between 2 and 3 bars or 3 and 4 bars. When presented with two bars of different sizes (i.e. one bar was taller and wider), treefrogs preferred the larger bar up to a 0.25 surface area ratio. Control experiments suggested that treefrogs represent numbers rather than continuous variables to discriminate between sets of bars, and that some physical attributes of the items (height) are more relevant than others (width). Moreover, I found a possible trade-off between speed and accuracy of decisions: individuals that take more time to search for stimuli are more accurate in choosing the larger set.

In the second study, crickets were presented with a dichotomous choice between sets of black shapes that simulated potential shelters. When stimuli differed in the number of items, crickets showed a preference for the larger set of 2 vs. 3 items. When stimuli were controlled for continuous variables (e.g., convex hull and cumulative surface area), the crickets no longer showed a preference for larger sets of 2 vs. 3 items. When subjects were presented with a dichotomous choice between two squares differing in area, a significant number of crickets selected the larger stimuli only when the size ratio was 0.25. Control experiments indicated that the results in size discrimination were poor because the crickets paid attention to the width but not the height of a stimulus. These findings suggest that *A. domesticus* possess a system to discriminate between quantities by summing the total area of presented items and that they rely more on some physical features (e.g., width) than on numerical information.

The difference observed between these two species is in agreement with previous literature. Treefrogs show evidence of possessing a true numerical system as already shown for several teleost fish, mammals and birds and in one of the two previously studied anuran species (Stancher et al., 2015). On the other hand, a lack of a true numerical system in crickets is aligned with current knowledge of quantitative systems in arthropods. In fact, the species so far studied seem to share a non-numerical-based system that is used in different contexts (e.g., social interaction: Rodriguez et al., 2015; shelter choice: Franks et al., 2006). Only one species (honeybees) apparently possess a true number-based system to discriminate between quantities (reviewed in Pahl et al., 2013; Skorupski et al., 2018), which is possibly in relation of the highly sophisticated foraging behavior evolved by this species.

Studies on numerical cognition have highlighted a wide variation in results both between and within species. Despite the large amount of data that has been published in the last decade, some issues remain largely unsolved. The results of the four studies reported here contribute to the understanding of these conflicting results by assessing the importance of the methodology and by investigating less studied taxa. However, several topics still need to be investigated. Firstly, the nature of the system involved in estimating small and large numbers remain unsolved. Neurophysiological studies reported that two distinct neural circuits are active when animals discriminate between sets of small or large numbers (Nieder, 2018). Comparative behavioural studies may provide a clear understanding of this matter, but it is necessary to be careful about the methodology that is adopted to study numerical cognition across species. Secondly, human and non-human species showed similar mechanisms when estimating quantities (Beran, 2008; Feigenson et al., 2004). However, studies on animals are restricted to few taxa (e.g., primates, birds, and, recently, fish). To understand and compare the nature of numerical cognition across species, it is necessary to understand the diversity of the neural mechanisms that may be involved. Accumulating evidence on the mechanisms underlying quantity discrimination abilities in less studied species, such as amphibians and invertebrates, may greatly contribute to understand the nature and evolutionary origin of this ability. Many more studies involving new species and new contexts are necessary before the matter can be clarified.

7. Reference list

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. *Journal of Neuroscience Methods*, 234: 82-91.

Agrillo, C., & Bisazza, A. (2018). Understanding the origin of number sense: a review of fish studies. *Philosophical Transactions of the Royal Society B Biological Sciences*, 373: 20160511.

Agrillo, C., Dadda, M. (2007). Discrimination of the larger shoal in the poeciliid fish *Girardinus falcatus*. *Ethology Ecology & Evolution*, 19: 145-157.

Agrillo, C., Dadda, M., & Bisazza, A. (2007). Quantity discrimination in female mosquitofish. *Animal Cognition*, 10: 63-70.

Agrillo, C., Dadda, M., & Serena, G. (2008a). Choice of female groups by male mosquitofish (*Gambusia holbrooki*). *Ethology*, 114: 479-488.

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2008b). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Animal Cognition*, 11: 495–503.

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2009). Use of number by fish. *PLoS ONE*, 4: e4786.

Agrillo, C., Miletto Petrazzini, M.E., & Bisazza A (2014) Numerical acuity of fish is improved in the presence of moving targets, but only in the subitizing range. *Animal Cognition*, 17: 307-316.

Agrillo, C., Miletto Petrazzini, M.E., Tagliapietra, C., & Bisazza, A. (2012a). Inter-specific differences in numerical abilities among teleost fish. *Frontiers in Psychology*, 3:483.

Agrillo, C., Piffer, L., Bisazza, A., & Butterworth, B. (2012b). Evidence for two numerical systems that are similar in humans and guppies. *PLoS ONE*, 7: e31923.

Agrillo, C., Miletto Petrazzini, M.E., & Bisazza, A. (2015). At the root of math: numerical abilities in fish. In Geary, D.C., Berch, D.C., Koepke, K.M. (eds). *Mathematical Cognition and Learning*. Amsterdam, Netherlands, Elsevier, pp 1-33.

Al Aïn, S., Giret, N., Grand, M., Kreutzer, M., & Bovet, D. (2009). The discrimination of discrete and continuous amounts in African grey parrots (*Psittacus erithacus*). *Animal Cognition*, 12: 145-154.

Andrade, C., Alwarshetty, M., Sudha, S., & Chandra, J.S. (2001). Effect of innate direction bias on T-maze learning in rats: implications for research. *Journal of Neuroscience Methods*, 110: 31-35.

Andrew, R.J., & Rogers, L.J. (2002). The nature of lateralization in tetrapods. In Andrew, R.J., & Rogers, L.J. (eds). *Comparative vertebrate lateralization*, Cambridge, USD, University Press, pp 94-125.

Arntzen, J.W. (1999). Sexual selection and male mate choice in the common toad, *Bufo bufo*. *Ethology Ecology & Evolution*, 11: 407-414.

Atkins, G., Atkins, S., Schoun, D., & Stout, J.F. (1987). Scototaxis and shape discrimination in the female cricket *Acheta domesticus* in an arena and on a compensatory treadmill. *Physiological Entomology*, 12: 125-133.

Babbitt, K.J., & Tanner, G.W. (1997). Effects of cover and predator identity on predation of *Hyla squirella* tadpoles. *Journal of Herpetology*, 31: 128-130.

Baker, J.M., Morath, J., Rodzon, K.S., & Jordan, K.E. (2012). A shared system of representation governing quantity discrimination in canids. *Frontiers in Psychology*, 3: 387.

Bánszegi, O., Urrutia, A., Szenczi, P., & Hudson, R. (2016). More or less: spontaneous quantity discrimination in the domestic cat. *Animal Cognition*, doi:10.1007/s10071-016-0985-2.

Bar-Shai, N., Keasar, T., & Shmida, A. (2011). The use of numerical information by bees in foraging tasks. *Behavioral Ecology, 22: 317-325.*

Bartholomew, A. (2012). Space size relative to prey width and total cover in an area both influence the habitat choices of freshwater angelfish *Pterophyllum scalare* in mesocosms. *Marine and Freshwater Behaviour and Physiology*, 45: 29-43.

Basolo, A.L. (1990). Female preference for male sword length in the green swordtail, *Xiphophorus helleri* (Pisces: *Poeciliidae*). *Animal Behaviour*, 40: 332-338.

Benson-Amram, S., Heinen, V.K., Dryer, S.L., & Holekamp, K.E. (2011). Numerical assessment and individual call discrimination by wild spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour*, 82: 743-752.

Beran, M.J. (2008). Monkeys (*Macaca mulatta* and *Cebus apella*) track, enumerate, and compare multiple sets of moving items. *Journal of Experimental Psychology: Animal Behavior Processes*, 34: 63-74.

Beran, M.J., Evans, T.A., Leighty, K.A., Harris, E.H., & Rice, D. (2008). Summation and quantity judgments of sequentially presented sets by capuchin monkeys (*Cebus apella*). *American Journal of Primatology*, 70: 191–194.

Bisazza, A., & Brown, C. (2011). Lateralization of cognitive functions in fish. In: Brown, C., Krause, J., Laland, K.N. (eds) *Fish cognition and behaviour*, Wiley, Oxford, pp 298-324.

Bisazza, A., Butterworth, B., Piffer, L., Bahrami, B., Miletto Petrazzini, M.E., & Agrillo, C. (2014a). Collective enhancement of numerical acuity by meritocratic leadership in fish. *Scientific Reports*, 4: 4560.

Bisazza, A., Agrillo, C., & Lucon-Xiccato, T. (2014b). Extensive training extends numerical abilities of guppies. *Animal Cognition*, 17: 1413-1419.

Bisazza, A., Piffer, L., Serena, G., & Agrillo, C. (2010). Ontogeny of numerical abilities in fish. *PLoS ONE*, 5: e15516.

Bisazza, A., Tagliapietra, C., Bertolucci, C., Foà, A., & Agrillo, C. (2014c). Non-visual numerical discrimination in a blind cavefish (*Phreatichthys andruzzii*). *Journal of Experimental Biology*, 217: 1902-1909.

Bogale, B. A., Aoyama, M., & Sugita, S. (2014). Spontaneous discrimination of food quantities in the jungle crow, *Corvus macrorhynchos*. *Animal Behaviour*, 94: 73-78.

Bonanni, R., Natoli, E., Cafazzo, S., & Valsecchi, P. (2011). Free-ranging dogs assess the quantity of opponents in intergroup conflicts. *Animal cognition*, 14: 103-115.

Boneau, C.A. (1960). The effects of violations of assumptions underlying the t test. *Psychological Bulletin*, 57:49-64.

Boysen, S.T. (1997). Representation of quantities by apes. *Advances in the Study of Behavior*, 26: 435–462.

Brannon, E.M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Developmental Science*, 9: 59-64.

Brembs, B., & Heisenberg, M. (2000). The operant and the classical in conditioned orientation of *Drosophila melanogaster* at the flight simulator. *Learning & Memory*, 7: 104-115.

Brown, C., & Braithwaite, V.A. (2005). Effects of predation pressure on the cognitive ability of the poeciliid *Brachyraphis episcopi*. *Behavioral Ecology*, 16: 482-487.

Brown, C., & Laland, K.N. (2003). Social learning in fishes: a review. *Fish and Fisheries*, 4: 280-288.

Brown, C., Laland, K., & Krause, J. (eds) (2008). *Fish cognition and behavior*. New Jersey, USD, John Wiley & Sons.

Bshary, R., & Brown, C. (2014). Fish cognition. *Current Biology*, 24: 947-950.

Bshary, R., Wickler, W., & Fricke, H. (2002). Fish cognition: a primate's eye view. *Animal Cognition*, 5 : 1-13.

Buchanan, K., Burt de Perera, T., Carere, C., Carter, T., Hailey, A., Hubrecht, R., ... & Sneddon, L. (2012). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour, 83*: 301-309.

Burr, D.C., & Ross, J. (2008). A visual sense of number. *Current Biology*, 18: 425-428.

Burr, D.C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision*, 10: 1-10.

Butterworth, B., & Laurillard, D. (2010). Low numeracy and dyscalculia: identification and intervention. *ZDM*, 42: 527-539.

Cachat, J., Stewart, A., Grossman, L., Gaikwad, S., Kadri, F., Chung, K.M., … & Goodspeed, J. (2010). Measuring behavioral and endocrine responses to novelty stress in adult zebrafish. *Nature Protocols,* 5: 1786-1799.

Call, J. (2000). Estimating and operating on discrete quantities in orangutans (*Pongo pygmaeus). Journal of Comparative Psychology*, 114: 136.

Campan, R., Beugnon, G., & Lambin, M. (1987). Ontogenetic development of behaviour: The cricket visual world. *Advances in the Study of Behavior*, 17: 165-212.

Cantlon, J.F., & Brannon, E.M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, 17: 401-406.

Cantlon, J.F., & Brannon, E.M. (2007). How much does number matter to a monkey (*Macaca mulatta*)?. *Journal of Experimental Psychology: Animal Behavior Processes,* 33: 32-41.

Cantlon, J.F., Pinel, P., Dehaene, S., & Pelphrey, K.A. (2010). Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cerebral Cortex,* 21: 191-199.

Carazo, P., Font, E., Forteza-Behrendt, E., & Desfilis, E. (2009). Quantity discrimination in *Tenebrio molitor*: evidence of numerosity discrimination in an invertebrate?. *Animal Cognition*, 12: 463-470.

Carazo, P., Fernández-Perea, R., & Font, E. (2012). Quantity estimation based on numerical cues in the mealworm beetle (*Tenebrio molitor*). *Frontiers in Psychology,* 3: 502.

Chandroo, K.P., Duncan, I.J., & Moccia, R.D. (2004). Can fish suffer?: perspectives on sentience, pain, fear and stress. *Applied Animal Behaviour Science*, 86: 225-250.

Chang, C.C., Ng, P.J., & Li, D. (2016). Aggressive jumping spiders make quicker decisions for preferred prey but not at the cost of accuracy. *Behavioral Ecology,* 28: 479-484.

Chittka, L., & Geiger, K. (1995a). Can honey bees count landmarks?. *Animal Behaviour,* 49: 159-164.

Chittka, L., & Geiger, K. (1995b). Honeybee long-distance orientation in a controlled environment. *Ethology*, 99: 117-126.

Chittka, L., & Niven, J. (2009). Are bigger brains better?. *Current Biology*, 19: 995-1008.

Chittka, L., Dyer, A.G., Bock, F., & Dornhaus, A. (2003). Psychophysics: Bees trade off foraging speed for accuracy. *Nature*, 424: 388-388.

Chittka, L., Geiger, K., & Kunze, J.A.N. (1995). The influences of landmarks on distance estimation of honey bees. *Animal Behaviour*, 50: 23-31.

Chittka, L., Skorupski, P., & Raine, N.E. (2009). Speed–accuracy tradeoffs in animal decision making. *Trends in Ecology & Evolution*, 24: 400-407.

Clayton, N.S., & Krebs, J.R. (1994). Memory for spatial and object-specific cues in food-storing and non-storing birds. *Journal of Comparative Physiology A*, 174: 371-379.

Colwill, R.M., Raymond, M.P., Ferreira, L., & Escudero, H. (2005). Visual discrimination learning in zebrafish *(Danio rerio*). *Behavioural Processes,* 70: 19-31.

Cooper, W.E., & Whiting, M.J. (2000). Islands in a sea of sand: Use of Acacia trees by tree skinks in the Kalahari Desert. *Journal of Arid Environments*, 44: 373-381.

Cote, J., Fogarty, S., & Sih, A. (2012). Individual sociability and choosiness between shoal types. *Animal Behaviour*, 83: 1469-1476.

Cresswell, W. (1994). Flocking is an effective anti-predation strategy in redshanks, *Tringa totanus*. *Animal behaviour*, 47: 433-442.

Cronin, A.L. (2014). Ratio-dependent quantity discrimination in quorum sensing ants. *Animal cognition*, 17: 1261-1268.

Cross, F.R., & Jackson, R.R. (2017). Representation of different exact numbers of prey by a spider-eating predator. *Interface Focus*, 7: 20160035.

Dacke, M., & Srinivasan, M.V. (2008). Evidence for counting in insects. *Animal Cognition,* 11: 683-689.

Dadda, M., Agrillo, C., Bisazza, A., & Brown, C. (2015). Laterality enhances numerical skills in the guppy, *Poecilia reticulata*. *Frontiers in Behavioral Neuroscience,* 9: 285.

Dadda, M., Piffer, L., Agrillo, C., & Bisazza, A. (2009). Spontaneous number representation in mosquitofish. *Cognition*, 112: 343-348.

Danisman, E., Bshary, R., & Bergmüller, R. (2010). Do cleaner fish learn to feed against their preference in a reverse reward contingency task? *Animal Cognition*, 13: 41-49.

Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences*, 11: 561-579.

Day, L.B., Crews, D., & Wilczynski, W. (1999). Spatial and reversal learning in congeneric lizards with different foraging strategies. *Animal Behaviour*, 57, 393-407.

de Winter, J.C.F. (2013). Using the Student's t-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation*, 18:1-12.

Deaner, R.O., Isler, K., Burkart, J., & Van Schaik, C. (2007). Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain, Behavior, and Evolution*, 70: 115-124.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44: 1-42.

Dehaene, S. (2001). Précis of the number sense. *Mind & Language*, 16: 16-36.

Dehaene, S., & Brannon, E. (eds) (2011). *Space, time and number in the brain: Searching for the foundations of mathematical thought*. Cambridge, USD, Academic Press.

Dehaene, S., & Changeux, J.P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive neuroscience*, 5: 390-407.

Douglas, J.W.B., & Whitty, C.W.M. (1941). An investigation of number appreciation in some sub-human primates. *Journal of Comparative Psychology*, 31:129-143.

Dussault, G.V., & Kramer, D.L. (1981). Food and feeding behavior of the guppy, *Poecilia reticulata* (Pisces: *Poeciliidae*). *Canadian Journal of Zoology*, 59: 684-701.

Emmerton, J., & Delius, J. (1993). Beyond Sensation: Visual Cognition in Pigeons. In Zeigles, H.P., & Bischof, H.J. (eds). *Vision, Brain, and Behavior in Birds*, Cambridge, USD, The MIT Press, pp: 377-390.

Emmerton, J., & Renner, J.C. (2006). Scalar effects in the visual discrimination of numerosity by pigeons. *Learning & Behavior,* 34: 176-192.

Emmerton, J., & Renner, J.C. (2009). Local rather than global processing of visual arrays in numerosity discrimination by pigeons (*Columba livia*). *Animal Cognition*, 12: 511-526.

Engqvist, L. (2005). The mistreatment of covariate interaction terms in linear model analyses of behavioural and evolutionary ecology studies. *Animal Behaviour*, 70: 967-971.

Feigenson, L. (2011). Predicting sights from sounds: 6-month-olds' intermodal numerical abilities. *Journal of Experimental Child Psychology,* 110: 347-361.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object-files versus analog magnitudes. *Psychological Science,* 13: 150–156.

Feigenson, L., Dehaene, S., & Spelke, E.S. (2004). Core systems of number. *Trends in Cognitive Science,* 8: 307-314.

Féron, J., Gentaz, E., & Streri, A. (2006). Evidence of amodal representation of small numbers across visuo-tactile modalities in 5-month-old infants. *Cognitive Development*, 21: 81-92.

Ferrigno, S., Jara-Ettinger, J., Piantadosi, S.T., & Cantlon, J.F. (2017). Universal and uniquely human factors in spontaneous number perception. *Nature Communications*, 8: 13968.

Fleishman, L.J., & Endler, J.A. (2000). Some comments on visual perception and the use of video playback in animal behavior studies. *Acta Ethologica*, 3:15-27.

Friedman, A., Spetch, M.L., & Ferrey, A. (2005). Recognition by humans and pigeons of novel views of 3-D objects and their photographs. *Journal of Experimental Psychology: General,* 134: 149-162.

Fox, J., & Weisberg, S. (2011). An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Franks, N.R., Dornhaus, A., Metherell, B.G., Nelson, T.R., Lanfear, S. A., & Symes, W.S. (2006). Not everything that counts can be counted: ants use multiple metrics for a single nest trait. *Proceedings of the Royal Society of London B: Biological Sciences*, 273: 165-169.

Gallistel, C.R. (1990). *The organization of learning*. Cambridge, USD, The MIT Press.

Gallistel, C.R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition,* 44: 43–74.

Garland, A., Low, J., & Burns, K. C. (2012). Large quantity discrimination by North Island robins (*Petroica longipes*). *Animal cognition*, 15: 1129-1140.

Genty, E., Palmier, C., & Roeder, J.J. (2004). Learning to suppress responses to the larger of two rewards in two species of lemurs, *Eulemur fulvus* and *E. macaco*. *Animal Behaviour,* 67: 925-932.

Gerhardt, H.C., Roberts, J.D., Bee, M.A., & Schwartz, J.J. (2000). Call matching in the quacking frog (*Crinia georgiana*). *Behavioral Ecology and Sociobiology,* 48: 243-251.

Gierszewski, S., Bleckmann, H., & Schluessel, V. (2013). Cognitive abilities in Malawi cichlids (*Pseudotropheus sp.*): matching-to-sample and image/mirror-image discriminations. *PloS ONE,* 8: e57363

Giurfa, M., Schubert, M., Reisenman, C., Gerber, B., & Lachnit, H. (2003). The effect of cumulative experience on the use of elemental and configural visual discrimination strategies in honeybees. *Behavioural Brain Research,* 145: 161-169.

Gomez, D., Richardson, C., Lengagne, T., Derex, M., Plenet, S., Joly, P., & Théry, M. (2010). Support for a role of colour vision in mate choice in the nocturnal European treefrog (*Hyla arborea*). *Behaviour,* 147: 1753-1768.

Gómez-Laplaza, L.M. (2006). Shoal choice in juvenile angelfish (*Pterophyllum scalare*): effects of social status and activity. *Ethology Ecology & Evolution,* 18: 261-273.

Gómez-Laplaza, L.M., & Gerlai, R. (2011). Can angelfish (*Pterophyllum scalare*) count? Discrimination between different shoal sizes follows Weber's law. *Animal Cognition,* 14:1-9.

Gómez-Laplaza, L.M., & Gerlai, R. (2012). Activity counts: the effect of swimming activity on quantity discrimination in fish. *Frontiers in Psychology,* 3: 484.

Gray, D.A. (1997). Female house crickets, *Acheta domesticus*, prefer the chirps of large males. *Animal Behaviour,* 54: 1553-1562.

Griggio, M., Valera, F., Casas-Crivillé, A., Hoi, H., & Barbosa, A. (2011). White tail markings are an indicator of quality and affect mate preference in rock sparrows. *Behavioral Ecology and Sociobiology,* 65: 655-664.

Gross, H.J., Pahl, M., Si, A., Zhu, H., Tautz, J., & Zhang, S. (2009). Number-based visual generalisation in the honeybee. *PloS ONE,* 4: e4263

Hager, M.C., & Helfman, G.S. (1991). Safety in numbers: shoal size choice by minnows under predatory threat. *Behavioral Ecology and Sociobiology,* 29: 271-276.

Halberda, J., & Odic, D. (2015). The precision and internal confidence of our approximate number thoughts. In Geary, D.C., Berch, D.C., Koepke, K.M. (eds). *Mathematical Cognition and Learning*. Amsterdam, Netherlands, Elsevier, pp 305-333.

Hall, G. (2002). Associative structures in Pavlovian and instrumental conditioning. In: *Stevens' handbook of experimental psychology*. New Jersey, USD, John Wiley & Sons

Hannula-Sormunen, M.M. (2015). Spontaneous focusing on numerosity and its relation to counting and arithmetic. In: Kadosh, R.C., & Dowker, A. (eds). *The Oxford handbook of numerical cognition*. Oxford, UK, Oxford University Press, pp 275-290.

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. *Journal of Comparative Psychology*, 121: 241-249.

Hastie, T.J., & Pregibon, D. (1992). Generalized linear models. In: Chambers, J.M., & Hastie, T.J. (eds). *Statistical models in S*. London, UK, Wadsworth and Brooks/Cole, Chapter 6.

Haun, D.B., Jordan, F.M., Vallortigara, G., & Clayton, N.S. (2011). Origins of Spatial, Temporal, and Numerical Cognition: Insights from Comparative Psychology** Reprinted from *Trends in Cognitive Sciences*, 14: 552–560.

Hauser, M.D., Carey, S., & Hauser, L.B. (2000). Spontaneous number representation in semi–free–ranging rhesus monkeys. *Proceedings of the Royal Society of London B: Biological Sciences*, 267: 829-833.

Hemptinne, J.L., Dixon, A.F.G., & Coffin, J. (1992). Attack strategy of ladybird beetles (*Coccinellidae*): factors shaping their numerical response. *Oecologia*, 90: 238-245.

Hinde, R.A., (1973). Constraints on learning: An introduction to the problems. In: Hinde R.A., & Stevenson-Hinde, J. (eds) *Constraints on Learning: Limitations and Predispositions*. Oxford, UK, Academic Press.

Hoare, D.J., Couzin, I.D., Godin, J.G., Krause, J. (2004). Context-dependent group size choice in fish. *Animal Behaviour*, 67:155-164.

Hollard, V.D., & Delius, J.D. (1982). Rotational invariance in visual pattern recognition by pigeons and humans. *Science*, 218:804–806.

Hollander, M., & Wolfe, D.A. (1973), Nonparametric Statistical Methods, New York, USD, Wiley.

Hope, T., Stoianov, I., & Zorzi, M. (2010). Through neural stimulation to behavior manipulation: a novel method for analyzing dynamical cognitive models. *Cognitive Science*, 34: 406-433.

Houde, A. (ed) (1997). *Sex, color, and mate choice in guppies*. Priceton, USA, Princeton University Press.

Howard, S.R., Avarguès-Weber, A., Garcia, J.E., Greentree, A.D., & Dyer, A.G. (2018). Numerical ordering of zero in honey bees. *Science*, 360: 1124-1126.

Hull, C.L., & Spence, K.W. (1938). " Correction" vs." non-correction" method of trial-and-error learning in rats. *Journal of Comparative Psychology*, 25: 127-145.

Hyde, D.C. (2011). Two systems of non-symbolic numerical cognition. *Frontiers in Human Neuroscience*, 5: 150.

Ildos, A.S., & Ancona, N. (1994). Analysis of amphibian habitat preferences in a farmland area (Po plain, northern Italy). *Amphibia-reptilia*, 15: 307-316.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. Proceedings of the National *Academy of Sciences*, 106: 10382-10385.

Jaakkola, K., Fellner, W., Erb, L., Rodriguez, M., & Guarino, E. (2005). Understanding of the concept of numerically" less" by bottlenose dolphins (*Tursiops truncatus*). *Journal of Comparative Psychology*, 119: 296-303.

Jones, S.M., Pearson, J., DeWind, N.K., Paulsen, D., Tenekedjieva, A.M., & Brannon, E.M. (2014). Lemurs and macaques show similar numerical sensitivity. *Animal Cognition*, 17: 503-515.

Jordan, K.E., & Brannon, E.M. (2006). The multisensory representation of number in infancy. *Proceedings of the National Academy of Sciences*, 103(9), 3486-3489.

Jordan, K. E., MacLean, E. L., & Brannon, E.M. (2008). Monkeys match and tally quantities across senses. *Cognition*, 108: 617-625.

Kahneman, D., Treisman, A., & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24: 175-219.

Kalueff, A.V., Echevarria, D.J., Homechaudhuri, S., Stewart, A.M., Collier, A.D., Kaluyeva, A.A., ... & Yang, L. (2016). Zebrafish neurobehavioral phenomics for aquatic neuropharmacology and toxicology research. *Aquatic Toxicology*, 170: 297-309.

Kieruzel, M. (1976). The selection of shelter place by the house cricket. *Acta Neurobiologiae Experimentalis*, 36: 561-580.

Kilian, A., Yaman, S., von Fersen, L., & Güntürkün, O. (2003). A bottlenose dolphin discriminates visual stimuli differing in numerosity. *Animal Learning & Behavior*, 31: 133-142.

Kobayashi, T., Hiraki, K., Mugitani, R., & Hasegawa, T. (2004). Baby arithmetic: One object plus one tone. *Cognition*, 91: 23-34.

Kobayashi, T., Hiraki, K., & Hasegawa, T. (2005). Auditory–visual intermodal matching of small numerosities in 6-month-old infants. *Developmental Science*, 8: 409-419.

Krause, J., & Ruxton, G.D. (eds) (2002). *Living in groups*. Oxford, UK, Oxford University Press.

Krause, J., Godin, J.G.J., & Rubenstein, D. (1998). Group choice as a function of group size differences and assessment time in fish: the influence of species vulnerability to predation. *Ethology*, 104:68-74.

Krusche, P., Uller, C., & Dicke, U. (2010). Quantity discrimination in salamanders. *Journal of Experimental Biology*, 213: 1822-1828.

Latty, T., & Beekman, M. (2011). Speed-accuracy trade-offs during foraging decisions in the acellular slime mould *Physarum polycephalum*. *Proceedings of the Royal Society of London B: Biological Sciences*, 278: 539-545.

Liao, W.B., & Lu, X. (2009). Male mate choice in the Andrew's toad *Bufo andrewsi*: A preference for larger females. *Journal of Ethology*, 27: 413-417.

Lefebvre, L., Reader, S.M., & Sol, D. (2004). Brains, innovations and evolution in birds and primates. Brain, *Behavior and Evolution*, 63: 233-246.

Lenth, R.V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69: 1-33.

Leppik, E.E. (1953). The ability of insects to distinguish number. *The American Naturalist*, 87: 229-236.

Lindström, K., & Ranta, E. (1993). Social preferences by male guppies, *Poecilia reticulata*, based on shoal size and sex. *Animal Behaviour*, 46: 1029-1031.

List of animals by number of neurons (n.d.). In Wikipedia. Retrieved September, 10, 2018, from https://en.wikipedia.org/wiki/List_of_animals_by_number_of_neurons

Loukola, O.J., Seppänen, J.T., Krams, I., Torvinen, S.S., & Forsman, J.T. (2013). Observed fitness may affect niche overlap in competing species via selective social information use. *The American Naturalist*, 182: 474-483.

Lucon-Xiccato, T., & Bisazza, A. (2014). Discrimination reversal learning reveals greater female behavioural flexibility in guppies. *Biology Letters*, 10: 20140206.

Lucon-Xiccato, T., & Bisazza, A. (2016). Male and female guppies differ in speed but not in accuracy in visual discrimination learning. *Animal Cognition*, 19: 733-744.

Lucon-Xiccato, T., & Dadda, M. (2017). Individual guppies differ in quantity discrimination performance across antipredator and foraging contexts. *Behavioral Ecology and Sociobiology*, 71: 13.

Lucon-Xiccato, T., Dadda, M., & Bisazza, A. (2016). Sex differences in discrimination of shoal size in the guppy (*Poecilia reticulata*). *Ethology*, 122: 481-491.

Lucon-Xiccato, T., Miletto Petrazzini, M.E., Agrillo, C., & Bisazza, A. (2015). Guppies discriminate between two quantities of food items but prioritize item size over total amount. *Animal Behaviour*, 107: 183-191.

Lyon, B.E. (2003). Egg recognition and counting reduce costs of avian conspecific brood parasitism. *Nature*, 422: 495-499.

Magurran, A.E. (ed) (2005). *Evolutionary ecology: the Trinidadian guppy*. Oxford, UK, Oxford University Press on Demand.

Magurran, A.E., & Pitcher, T.J. (1987). Provenance, shoal size and the sociobiology of predator-evasion behaviour in minnow shoals. *Proceedings of the Royal Society of London B: Biological Sciences*, 229: 439-465.

Manabe, K., Dooling, R.J., & Takaku, S. (2013) An automated device for appetitive conditioning in zebrafish (*Danio rerio*). *Zebrafish*, 10 :518-523.

Matsuzawa, T. (2009). Symbolic representation of number in chimpanzees. *Current Opinion in Neurobiology*, 19: 92-98.

McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, *Panthera leo*. *Animal Behaviour*, 47: 379-387.

Mensforth, C.L., & Bull, C.M. (2008). Selection of artificial refuge structures in the Australian skink, *Egernia stokesii*. *Pacific Conservation Biology*, 14: 63-68.

Menzel, R., & Giurfa, M. (2001). Cognitive architecture of a mini-brain: the honeybee. *Trends in Cognitive Sciences,* 5: 62-71.

Menzel, R., Fuchs, J., Nadler, L., Weiss, B., Kumbischinski, N., Adebiyi, D., ... & Greggers, U. (2010). Dominance of the odometer over serial landmark learning in honeybee navigation. *Naturwissenschaften*, 97:, 763-767.

Michaels, C.J., Antwis, R.E., & Preziosi, R.F. (2014). Impact of plant cover on fitness and behavioural traits of captive red-eyed tree frogs (*Agalychnis callidryas*). *PLoS ONE,* 9: e95207.

Miletto Petrazzini, M.E., & Agrillo, C. (2016). Turning to the larger shoal: are there individual differences in small-and large-quantity discrimination of guppies?. *Ethology Ecology & Evolution*, 28: 211-220.

Miletto Petrazzini, M.E., & Wynne, C.D. (2016) What counts for dogs (*Canis lupus familiaris*) in a quantity discrimination task? *Behavioural Processes*, 122: 90-97

Miletto Petrazzini, M.E., Lucon-Xiccato, T., Agrillo, C., & Bisazza, A. (2015). Use of ordinal information by fish. *Scientific Reports*, 5: 15497.

Mirwan, H.B., & Kevan, P.G. (2015). Conditional discrimination and response chains by worker bumblebees (*Bombus impatiens Cresson,* Hymenoptera: *Apidae*). *Animal Cognition,* 18: 1143-1154.

Mou, Y. (2014). Two core systems of numerical representation in infants. *Developmental Review*, 34: 1-25.

Moyer, R.S., & Landauer, T.K. (1967). Time required for judgements of numerical inequality. *Nature*, 215: 1519-1520.

Mühlhoff, N., Stevens, J.R., & Reader, S.M. (2011). Spatial discounting of food and social rewards in guppies (*Poecilia reticulata*). *Frontiers in Psychology,* 2: 68.

Mustafar, F., De Luna, P., & Rainer, G. (2015). Enhanced visual exploration for real objects compared to pictures during free viewing in the macaque monkey. *Behavioural Processes,* 118: 8-20.

Nelson, X.J., & Jackson, R.R. (2012). The role of numerical competence in a specialized predatory strategy of an araneophagic spider. *Animal Cognition,* 15: 699-710.

Nicoletto, P.F., & Kodric-Brown, A. (1999). The use of digitally-modified videos to study the function of ornamentation and courtship in the guppy, *Poecilia reticulata*. *Environmental Biology of Fishes*, 56: 333-342.

Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. *Nature Reviews Neuroscience*, 6: 177-190.

Nieder, A. (2009). Prefrontal cortex and the evolution of symbolic reference. *Current Opinion in Neurobiology,* 19: 99-108.

Nieder, A. (2018). Evolution of cognitive and neural solutions enabling numerosity judgements: lessons from primates and corvids. *Philosophical Transactions of the Royal Society B Biological Sciences*, 373: 20160514.

Normand, E., Ban, S.D., & Boesch, C. (2009). Forest chimpanzees (*Pan troglodytes verus*) remember the location of numerous fruit trees. *Animal Cognition*, 12: 797-807.

O'Hara, M., Huber, L., & Gajdon, G.K. (2015). The advantage of objects over images in discrimination and reversal learning by kea, *Nestor notabilis*. *Animal Behaviour*, 101: 51-60.

Olson, D.J., Kamil, A.C., Balda, R.P., & Nims, P.J. (1995). Performance of four-seed caching corvid species in operant tests of nonspatial and spatial memory. *Journal of Comparative Psychology*, 109: 173-181.

Pahl, M., Si, A., & Zhang, S. (2013). Numerical cognition in bees and other insects. *Frontiers in Psychology*, 4.

Panteleeva, S., Reznikova, Z., & Vygonyailova, O. (2013). Quantity judgments in the context of risk/reward decision making in striped field mice: first “count,” then hunt. *Frontiers in Psychology*, 4.

Parducci, A., & Polt, J. (1958). Correction vs. noncorrection with changing reinforcement schedules. *Journal of Comparative and Physiological Psycholoy*, 51: 492-495.

Parker, M.O., Gaviria, J., Haigh, A., Millington, M.E., Brown, V.J., Combe, F.J., & Brennan, C.H. (2012a). Discrimination reversal and attentional sets in zebrafish (*Danio rerio*). *Behavioural Brain Research*, 232: 264-268.

Parker, M.O., Millington, M.E., Combe, F.J., & Brennan, C.H. (2012b) Development and implementation of a three-choice serial reaction time task for zebrafish (*Danio rerio*). *Behavioural Brain Research*, 227: 73-80.

Patton, E.E., & Zon, L.I. (2001). The art and design of genetic screens: zebrafish. *Nature Reviews Genetics*, 2: 956-966.

Pepperberg, I.M. (2006). Grey parrot numerical competence: a review. *Animal Cognition*, 9: 377-391

Pepperberg, I.M., & Brezinsky, M.V. (1991). Acquisition of a relative class concept by an African gray parrot (*Psittacus erithacus*): discriminations based on relative size. *Journal of Comparative Psychology*, 105: 286-294.

Piazza, M., & Izard, V. (2009). How humans count: numerosity and the parietal cortex. *The Neuroscientist*, 15: 261-273.

Plath, M., Richter, S., Tiedemann, R., & Schlupp, I. (2008). Male fish deceive competitors about mating preferences. *Current Biology*, 18: 1138-1141.

Prétôt, L., Bshary, R., & Brosnan, S.F. (2016). Comparing species decisions in a dichotomous choice task: adjusting task parameters improves performance in monkeys. *Animal Cognition*, 19: 819-834.

Proulx, M.J., Parker, M.O., Tahir, Y., & Brennan, C.H. (2014). Parallel mechanisms for visual search in zebrafish. *PLoS ONE*, 9: e111540.

Prusky, G.T., West, P.W., & Douglas, R.M. (2000). Behavioral assessment of visual acuity in mice and rats. *Vision Research*, 40: 2201-2209.

Pylyshyn, Z.W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80: 127-158.

Pylyshyn, Z.W., & Storm, R.W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3: 179-197.

Ranta, E., Juvonen, S.K., & Peuhkuri, N. (1992). Further evidence for size-assortative schooling in sticklebacks. *Journal of Fish Biology*, 41: 627-630.

Reebs, S.G. (1999). Time–place learning based on food but not on predation risk in a fish, the inanga (*Galaxias maculatus*). *Ethology*, 105: 361-371.

Reznikova, Z.I., & Ryabko, B.Y. (1993). Ants aptitude for the transmission of information on the number of objects. *Sensory systems of arthropods.–Birkhäuser Verlag, Basel*, 634-639.

Reznikova, Z.I., & Ryabko, B.Y. (1994) Experimental study of the ants' communication system, with the application of the Information Theory approach. *Memorabilia Zoologica*, 48: 219-236.

Reznikova, Z.I., & Ryabko, B.Y. (1996). Transmission of information regarding the quantitative characteristics of an object in ants. *Neuroscience and Behavioral Physiology*, 26: 397-405.

Reznikova, Z.I., & Ryabko, B.Y. (2000). Using Information Theory approach to study the communication system and numerical competence in ants. In: *From Animals to Animats 6. Proceeding of the Sixth International Conference on Simulation of Adaptive Behaviour*. Cambridge, USD, The MIT Press, 501: 506

Roberts, W.A., & Mitchell, S. (1994). Can a pigeon simultaneously process temporal and numerical information? *Journal of Experimental Psychology: Animal Behavior Processes*, 20: 66-78.

Roberts, W.A., Roberts, S., & Kit, K.A. (2002). Pigeons presented with sequences of false flashes use behavior to count but not to time. *Journal of Experimental Psychology: Animal Behavior Processes*, 28: 137-150.

Rodd, F.H., Hughes, K.A., Grether, G.F., & Baril, C.T. (2002). A possible non-sexual origin of mate preference: are male guppies mimicking fruit?. *Proceedings of the Royal Society B Biological Sciences*, 269: 475-481.

Rodríguez, R.L., Briceño, R.D., Briceno-Aguilar, E., & Höbel, G. (2015). *Nephila clavipes* spiders (Araneae: *Nephilidae*) keep track of captured prey counts: testing for a sense of numerosity in an orb-weaver. *Animal Cognition*, 18: 307-314.

Rogers, L.J., & Andrew, R. (eds) (2002). *Comparative vertebrate lateralization*. Cambridge, USD, University Press.

Rose, G.J. (2018). The numerical abilities of anurans and their neural correlates: Insights from neuroethological studies of acoustic communication. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 373: 20160512.

Rosenthal, G.G. (2000). Design considerations and techniques for constructing video stimuli. *Acta Ethologica,* 3: 49-54.

Rosenthal, G.G., & Ryan, M.J. (2005). Assortative preferences for stripes in danios. *Animal Behaviour*, 70: 1063-1066.

Salwiczek, L.H., Prétôt, L., Demarta, L., Proctor, D., Essler, J., Pinto, A.I., & Bshary, R. (2012). Adult cleaner wrasse outperform capuchin monkeys, chimpanzees and orang-utans in a complex foraging task derived from cleaner–client reef fish cooperation. *PLoS ONE,* 7: e49068.

Schluessel, V., Kraniotakes, H., & Bleckmann, H. (2014). Visual discrimination of rotated 3D objects in Malawi Cichlids (*Pseudotropheus* sp.): a first indication for form constancy in fishes. *Animal Cognition,* 17: 359-371.

Seebacher, F., & Alford, R.A. (2002). Shelter microhabitats determine body temperature and dehydration rates of a terrestrial amphibian (*Bufo marinus*). *Journal of Herpetology*, 36: 69-75.

Seghers, B.H. (1974). Schooling behavior in the guppy (*Poecilia reticulata*): An evolutionary response to predation. *Evolution,* 28: 486-489.

Sherry, D.F. (2006). Neuroecology. *Annual Reviews of Psychology*, 57:167-197.

Shettleworth, S.J. (1972). Constraints on learning. In: Lehrman, D.S., Hinde, R.A., & Shaw, E. (eds). *Advances in the study of behaviour: IV*, Oxford, UK, Academic Press pp 1-68.

Shettleworth, S.J. (ed) (2010). *Cognition, evolution, and behavior*. Oxford, UK, Academic Press.

Shettleworth, S.J., & Hampton, R.R. (1998). Adaptive specializations of spatial cognition in food storing birds? Approaches to testing a comparative hypothesis. In: Balda, R.P., Pepperberg, I.M., & Kamil, A.C. (eds). *Animal cognition in nature,* San DiegoUSD, pp 65-98.

Skinner, B.F. (1933). The rate of establishment of a discrimination. *Journal of General Psychology*, 9: 302-350.

Skorupski, P., MaBouDi, H., Dona, H.S.G., & Chittka, L. (2018). Counting insects. *Philosophical Transactions of the Royal Society B Biological Sciences*, 373: 20160513.

Smith, A.P., & Zentall, T.R. (2016). Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42: 212-200.

Sokal, R.R., & Rohlf, F.J. (eds) (1995). *Biometry: the principals and practice of statistics in biological research*. New York, USD, WH Freeman and Company.

Sovrano, V.A., Bisazza, A., & Vallortigara, G. (2003). Modularity as a fish (*Xenotoca eiseni*) views it: conjoining geometric and nongeometric information for spatial reorientation. *Journal of Experimental Psychology: Animal Behavior Processes*, 29: 199-210.

Stancher, G., Rugani, R., Regolin, L., & Vallortigara, G. (2015). Numerical discrimination by frogs (*Bombina orientalis*). *Animal Cognition*, 18: 219-229.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, 36: 97-127.

Stephens, D.W., & Krebs, J.R. (eds) (1986). Foraging theory. Princeton, USD, Academic Press.

Stephens, D.W. (2008). Decision ecology: foraging and the ecology of animal decision making. *Cognitive, Affective, & Behavioral Neuroscience*, 8: 475-484.

Stevens, J.R., Wood, J.N., & Hauser, M.D. (2007). When quantity trumps number: discrimination experiments in cotton-top tamarins (*Saguinus oedipus*) and common marmosets (*Callithrix jacchus*). *Animal Cognition*, 10: 429-437.

Stewart, M.M. (1985). Arboreal habitat use and parachuting by a subtropical forest frog. *Journal of Herpetology*, 19: 391-401.

Stoianov, I., & Zorzi, M. (2012). Emergence of a'visual number sense'in hierarchical generative models. *Nature Neuroscience*, 15: 194-196.

Takahashi, K., & Nagayama, T. (2016). Shelter preference in the Marmorkrebs (marbled crayfish). *Behaviour*, 153: 1913-1930.

Tanner, C.J. (2006). Numerical assessment affects aggression and competitive ability: a team-fighting strategy for the ant *Formica xerophila*. *Proceedings of the Royal Society of London B: Biological Sciences*, 273: 2737-2742.

Tegeder, R.W., & Krause, J. (1995). Density dependence and numerosity in fright stimulated aggregation behaviour of shoaling fish. *Philosophical Transactions of the Royal Society B Biological Sciences*, 350: 381-390.

Thornton, A., & Lukas, D. (2012). Individual variation in cognitive performance: developmental and evolutionary perspectives. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367: 2773-2783.

Todd, P.M. (2001). Fast and frugal heuristics for environmentally bounded minds. In: Selten, R., & Gigerenzer, G. (eds), *Bounded rationality: the adaptive toolbox*. Cambridge, USD, The MIT Press, pp: 51-70.

Tomonaga, M. (2008). Relative numerosity discrimination by chimpanzees (*Pan troglodytes*): Evidence for approximate numerical representations. *Animal Cognition*, 11: 43–57.

Trick, L.M., & Pylyshyn, Z.W. (1993). What enumeration studies can show us about spatial attention: evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance*, 19: 331-351.

Truppa, V., Carducci, P., Trapanese, C., & Hanus, D. (2015). Does presentation format influence visual size discrimination in tufted capuchin monkeys (*Sapajus spp.*)?. *PloS ONE,* 10: e0126001.

Uller, C., Jaeger, R., Guidry, G., & Martin, C. (2003). Salamanders (*Plethodon cinereus*) go for more: rudiments of number in an amphibian. *Animal Cognition*, 6: 105-112.

vanMarle, K. (2015). Foundations of the Formal Number Concept: How Preverbal Mechanisms Contribute to the Development of Cardinal Knowledge. In Geary, D.C., Berch, D.C., Koepke, K.M. (eds). *Mathematical Cognition and Learning*. Amsterdam, Netherlands, Elsevier, pp 175-199.

Vonk, J., & Beran, M.J. (2012). Bears ‘count’too: quantity estimation and comparison in black bears, *Ursus americanus*. *Animal Behaviour*, 84: 231-238.

Vos, P.G., Oeffelen, M.P., Tibosch, H.J., & Allik, J. (1988). Interactions between area and numerosity. *Psychological Research*, 50: 148-154.

Walsh, P.T., & Downie, J.R. (2005). The effects of shelter availability and substrate quality on behaviour and post-metamorphic growth in three species of anurans: Implications for captive breeding. *The Herpetological Journal*, 15: 245-255.

Wang, M.Y., Brennan, C.H., Lachlan, R.F., & Chittka, L. (2015) Speed–accuracy trade-offs and individually consistent decision making by individuals and dyads of zebrafish in a colour discrimination task. *Animal Behaviour*, 103: 277-283.

Weber, T., Atkins, G., Stout, J.F., & Huber, F. (1987). Female *Acheta domesticus* track acoustical and visual targets with different walking modes. *Physiological Entomology*, 12: 141-147.

West, R.E., & Young, R.J. (2002). Do domestic dogs show any evidence of being able to count?. *Animal Cognition*, 5: 183-186.

Westfall, P.H., Tobias, R.D., & Wolfinger, R.D. (2011). *Multiple comparisons and multiple tests using SAS*. SAS Institute.

Wickens, D.D., Nield, A.F., Tuber, D.S., & Wickens, C. (1970) Classically conditioned compound-element discrimination as a function of length of training, amount of testing and CS-UCS interval. *Learning and Motivation*, 1: 95-109.

Wiese, H. (ed) (2003). *Numbers, language, and the human mind*. Cambridge, UK, University Press.

Wittlinger, M., Wehner, R., & Wolf, H. (2006). The ant odometer: stepping on stilts and stumps. *Science*, 312: 1965-1967.

Wittlinger, M., Wehner, R., & Wolf, H. (2007). The desert ant odometer: a stride integrator that accounts for stride length and walking speed. *Journal of Experimental Biology*, 210: 198-207.

Wright, A.A., Cook, R.G., Rivera, J.J., Sands, S.F., & Delius, J.D. (1988). Concept learning by pigeons: matching-to-sample with trial-unique video picture stimuli. *Animal Learning & Behavior*, 16: 436-444.

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

Yang, T.I., & Chiao, C.C. (2016). Number sense and state-dependent valuation in cuttlefish. *Proceedings of the Royal Society of London B: Biological Sciences*, 283: 20161379.

Zars, T., Fischer, M., Schulz, R., & Heisenberg, M. (2000). Localization of a short-term memory in Drosophila. *Science*, 288: 672-675.

Zorzi, M., & Testolin, A. (2018). An emergentist perspective on the origin of number sense. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 373: 20170043.

Acknowledgment

I would like to spend few words to all of the kind, patient, and persistent people that supported me during this travel.

First and foremost, I have a larger debt of gratitude with my parents. They invest so much for helping me on following my career. I would never have arrived at this point without their support (e.g., I have broken three laptops during my studies).

At the same time, I owe thanks to my advisors and my colleagues. Angelo, Marco, Christian, Maria Elena, Maria, and Tyrone know how many mistakes I have done during my PhD. However, I try to remedy with hard-working and sweets. A more special thanks to Angelo for inciting me on finding my way. Secondly, I appreciate the work of all student on helping me, especially Daniele, and Basak. I would like to thank also Silvia, Alessandro, and Andrea because me, Francesco, and Martina have followed your advised and started this long travel. At least, thanks to my Finnish friend and his lovely family.

Finally, I would like to thank my friends. I have to thank all of you for your support. A special thanks to Roxanne and, especially, Elizabeth for correcting my bad English and supporting me.