https://doi.org/10.35662/unine-thesis-3149

Université de Neuchâtel
Faculté des sciences

Working Memory and Associated Cognitive Abilities in cleaner wrasse *Labroides dimidiatus*

Léonore BONIN

PhD thesis submitted to:
The Faculty of Sciences, Institute of Biology

Examination committee:
Prof. Redouan Bshary (supervisor) - University of Neuchâtel, Switzerland
Prof. Adrian Bangerter - University of Neuchâtel, Switzerland
Prof. Héctor M. Manrique - University of Zaragoza, Spain
Prof. Carel P. van Schaïk - University of Zurich, Switzerland

Defended and accepted 18.10.2024

unine
Université de Neuchâtel
Faculté des sciences

Rue Emile-Argand 11
CH-2000 Neuchâtel
doctorat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Madame Léonore BONIN

Titre :

"Working Memory and Associated Cognitive Abilities in cleaner wrasse Labroides dimidiatus"

sur le rapport des membres du jury composé comme suit :

- **Prof. Redouan Bshary,** directeur de thèse, UniNE
- **Prof. Adrian Bangerter,** UniNE
- **Prof. Carel van Schaik,** Universität Zürich, Suisse
- **Prof. Héctor Manrique,** Universidad de Zaragoza, Espagne

Neuchâtel, le 30 octobre 2024

Le Doyen, Prof. P. Brunner

Table of Contents

Résumé

Comprendre la nature de l'intelligence reste un défi majeur pour l'esprit humain. Divers types de mesures ont été employés pour déterminer ce qui constitue l'intelligence et expliquer les différences entre les espèces, ainsi qu'entre les individus d'une même espèce. Étudier la cognition, qui se définie comme l'ensemble des processus mentaux que les organismes utilisent pour percevoir, analyser et répondre adéquatement aux informations de leur environnement, permet d'évaluer et de comparer certaines unités fondamentales de ce que l'on appelle communément l'intelligence : les capacités cognitives.

Une approche commune, bien qu’imparfaite, consiste à comparer la taille relative du cerveau par rapport à celle du corps pour estimer les différences de complexité cognitive entre les espèces. Lorsque cette mesure est appliquée aux vertébrés, il apparaît que les endothermes, qui régulent leur température corporelle par des processus physiologiques internes, possèdent un cerveau proportionnellement en moyenne dix fois plus grand que les ectothermes, qui dépendent de la température extérieure. Les poissons, reptiles et amphibiens, qui composent le groupe des ectothermes, sont largement considérés comme moins intelligents que les mammifères et les oiseaux qui composent le groupe des endothermes.

Cependant, environ deux décennies de recherches sur le labre nettoyeur (*Labroides dimidiatus*), un petit poisson des récifs tropicaux, ont commencé à remettre en cause cette idée largement répandue. Des études comportementales et cognitives ont révélé que ce poisson pouvait, dans certains contextes, démontrer des capacités comparables, voire supérieures, à celles des primates. Les primates ayant un taille cérébrale relative importante, même au sein du groupe des endothermes. Ces découvertes ont incité un changement de perspective qui demande à se concentrer d’avantage sur les mécanismes par lesquels ces poissons résolvent les problèmes, plutôt que de se limiter à identifier les types de problèmes qu'ils sont capables de résoudre.

Ce cadre conceptuel a inspiré les recherches présentées dans cette thèse, qui s’appuient également sur l’approche écologique de la cognition, postulant que les capacités cognitives d'une espèce sont modulées par ses besoins écologiques.

Nous avons abordé ces questions de plusieurs manières. Tout d'abord, nous nous sommes concentrés sur la mémoire de travail, un processus cognitif clé qui sous-tend de nombreuses autres capacités, et l’avons exploré chez le poisson nettoyeur à travers divers paradigmes expérimentaux. La mémoire de travail est un type de mémoire à court terme conscient, utilisé pour combiner et manipuler des informations perceptuelles et internes, passées et présentes. Notre étude intègre différentes perspectives issues de la littérature scientifique et évalue la pertinence écologique de la mémoire de travail en tant que facteur modulateur.

Ensuite, nous avons examiné l'inférence transitive, une capacité de raisonnement par laquelle les individus déduisent des relations inconnues entre des objets ou individus à partir de celles précédemment apprises. Il s'agit d'une capacité cognitive bien étudiée, présente chez des espèces aux cerveaux plus petits et qui ne semble pas nécessaire à l'écologie du poisson nettoyeur.

Enfin, nous avons exploré la permanence de l'objet, qui se réfère à la compréhension qu'un objet continue d'exister même lorsqu'il est hors de vue. Cette capacité peut potentiellement aider à résoudre divers défis écologiques et permet d'évaluer un développement cognitif graduel, offrant ainsi des perspectives sur les différences qualitatives entre espèces, au-delà des simples différences quantitatives.

Ces trois capacités cognitives sont intimement liées, avec à la fois l'inférence transitive et la permanence de l'objet qui peuvent dépendre de la mémoire de travail. Ainsi, leur étude conjointe permet de mieux comprendre ces interrelations.

Dans ce travail de recherche, nous avons premièrement mis en évidence des échecs répétés du labre nettoyeur dans les premiers stades de test des différents paradigmes expérimentaux visant à étudier la mémoire de travail. Ceci nous a donc empêcher de tester les stades supérieurs dans les tâches visuelles autant que dans les tâches spatiales, qui auraient dû étudier certains aspects clés de la mémoire de travail reportés dans la littérature. Deuxièmement, nos recherches sur l'inférence transitive se sont aussi avérées négatives. Il paraissait aussi évident qu'un manque de mémoire était à l'origine de cet échec. De plus, en comparant nos résultats à ceux d'une autre équipe nous avons pu conclure que les poissons nettoyeurs pourraient simplement avoir besoin d'un entrainement plus étendu pour pouvoir résoudre la tâche. Ce besoin supplémentaire corrobore aussi l'approche écologique de l'inférence transitive car l'écologie du poisson nettoyeur ne semble pas nécessiter l'inférence transitive. Pour finir, nous avons démontré pour la première fois que le labre nettoyeur pouvait réussir une tâche de permanence de l'objet au stade 4 de Piaget, en retrouvant un objet caché, mais seulement après avoir été exposé au stade 3 (lorsque l'objet n'est que partiellement caché). La nature de l'objet, qui était sensée refléter des niveaux de pertinence écologique différents, n'a pas significativement impacté la performance des labres nettoyeurs. Il est à noter que nous avons obtenus la performance la plus élevée à ce jour chez un poisson pour cette tâche.

En définitive, cette thèse démontre que, grâce à un entraînement adéquat, le labre nettoyeur peut pallier certaines limitations cognitives. Nous apportons aussi des données qui supportent l'approche écologique de la cognition, même si tester directement cette hypothèse s'est avéré délicat. Finalement, cette thèse propose aussi une hypothèse pour expliquer les différences de taille cérébrale relative entre les endothermes et les ectothermes, et suggère des pistes de recherche futures. Nous proposons que la capacité à contrôler l'attention pourrait expliquer la différence entre les deux groupes de vertébrés, et des études plus poussées chez d'autres poissons et plus largement, d'autre ectothermes, devraient être effectuées.

Mots-clefs : labre nettoyeur ; cognition animale ; intelligence animale ; écologie comportementale ; approche écologique de la cognition ; mémoire de travail ; inférence transitive ; permanence de l'objet ; contrôle de l'attention.

Abstract

Understanding the nature of intelligence remains one of the most complex and enduring challenges for the human mind. Various metrics have been employed to elucidate what constitutes intelligence and to explain the variations in cognitive abilities both within and among species. The study of cognition, that is, the study of all the mental abilities that organisms use to perceive, analyze and respond to their environment, offers a pathway to measuring and comparing fundamental units that ultimately compose what we define as intelligence.

One common, albeit imperfect, approach to estimating cognitive complexity across species is by examining brain size relative to body size. When applied to vertebrates, this measure highlights a striking mean 10-fold difference between ectotherms, whose body temperature largely reflects external conditions, and endotherms, which possess internal physiological mechanisms to regulate their body temperature. Ectotherms, a group consisting of fish, reptiles, and amphibians, are generally considered less intelligent than endotherms, which include mammals and birds.

However, approximately 20 years of research on the Bluestreak cleaner wrasse (*Labroides dimidiatus*), a small reef fish, has begun to challenge this notion. Behavioral and cognitive studies have shown that this fish possesses abilities that, in some cases, rival or even surpass those of primates, which are large-brained even by endotherm standards. These key findings have prompted a shift from merely asking "what" cognitive tasks these fish can solve to exploring "how" they solve them. This is the foundation for the present work, in addition to the ecological approach to cognition, which posits that a species' cognitive toolkit is shaped by ecological demand.

We employed multiple approaches to address these questions. First, we focused on working memory (WM), a key cognitive process underpinning many other abilities, and explored its role in cleaner wrasse through various experimental paradigms. WM is a short-term memory storage system used for the conscious combination and manipulation of both perceptual and internal information, encompassing past and present experiences. Our investigation integrated different perspectives from existing literature and assessed the ecological relevance of WM as a modulating factor. Second, we examined transitive inference (TI), a reasoning ability where individuals deduce unknown relationships based on previously learned ones. TI is a well-established cognitive ability found in species with smaller brains that is not essential to cleaner wrasse ecology. Lastly, we explored object permanence (OP), which refers to the understanding that an object still exists even when put out of sight. OP has the potential to help solve a variety of ecological challenges and allows for the assessment of gradual cognitive development, providing insights into qualitative differences between species rather than mere quantitative ones. These three cognitive abilities are interconnected, with both TI and OP potentially relying, at least in part, on WM. Consequently, studying all three has provided deeper insights into their relationships.

In this thesis, we initially conducted a series of experiments to evaluate WM abilities in cleaner wrasse, though some paradigms remained incomplete due to behavioral biases. Across all paradigms, early failures prevented further exploration of key WM components in both visual and spatial tasks. These findings suggest that cleaner wrasse may rely on cognitive processes that do not require conscious recollection or control to solve problems and navigate their ecological challenges. Second, our investigation of TI yielded negative results as well, with evidence pointing to memory limitations likely impeding TI ability. In addition, comparing our results to those of another team also suggested that cleaner wrasse can succeed in the task when provided with extensive training. This necessity for additional training aligns with our hypothesis that cleaner wrasse ecology does not inherently demand TI, reinforcing the ecological approach to cognition. Lastly, we demonstrated for the first time that cleaner wrasse can successfully complete a Piagetian stage 4 task of object permanence (OP), retrieving fully hidden objects, but only after exposure to a stage 3 task (retrieving partially hidden objects). The nature of the object, which was supposed to reflect different levels of ecological relevance, did not significantly impact the performance of cleaner wrasse. Notably, this study reports the highest success rate in a fish for this task to date.

To conclude, this thesis demonstrates that, with appropriate training, cleaner wrasse can compensate for the absence of complex cognitive mechanisms. Furthermore, our data bring support to the ecological approach to cognition, though direct testing can be challenging. Finally, we offer a hypothesis for the brain-size disparity between endotherms and ectotherms and suggest future research directions. We propose attentional control as a potential distinguishing factor between these vertebrate groups and recommend broader investigations across other fish species and, more generally, other ectotherms to develop a more comprehensive understanding.

Keywords: cleaner wrasse; animal cognition; animal intelligence; behavioral ecology; ecological approach to cognition; working memory; transitive inference; object permanence; attentional control.

General Introduction

I.1. The nature of intelligence

What is intelligence? Where does it come from? Humans have been pondering these questions since ancient times. Philosophers like Aristotle and Plato were speculating on the nature of the human mind long before our modern scientific tools as they sought to connect sensory perceptions to conscious understanding. While their lines of inquiry were also impacted by the omnipresence of religion and beliefs—often linking the mind's abilities to the soul—they nevertheless established the basis for differentiating between the automatic perception of sensory inputs and their active analysis (Aristotle, De Anima, 350 BC).

In the late 1800s, Sir Francis Galton studied the heredity of human intelligence and the transmission of “talent” or “genius”, despite not giving a clear definition of what these concepts entailed (Galton, 1865). He investigated the inheritance of talent through multiple families from biographic records and argued that, although the environment could affect a person's propensity for accomplishing outstanding achievements, it played a minor role compared to heredity. While Galton’s work influenced controversial societal movements such as eugenism, it also helped lay the groundwork for the study of intelligence and psychometrics.

Building on earlier works, Alfred Binet and Theodore Simon elaborated the well-known Binet-Simon intelligence test for IQ during the early 1900s (Binet & Simon, 1905). This test, along with all of its subsequent iterations, is still widely used to classify students based on their school abilities. However, despite its success and ubiquity, this test still has limitations in measuring certain widely recognized aspects of intelligence, such as curiosity and intuition (Binet & Simon, 1905; Terman, 1950). Around the same time Binet and Simon developed their IQ test, Charles Spearman (1904) also proposed an idea of "general intelligence" where all abilities are linked by a single factor: the Spearman *g* factor. Mental abilities are positively correlated meaning that when an individual performs well on one task, he will probably perform well on another. As a result, a positive correlation, also called the positive manifold, links all cognitive abilities. Spearman explains the existence of this positive manifold by arguing that all abilities are connected by a single general factor of intelligence: the Spearman *g* factor.

Later in the 1900s, Cattell's theory of fluid and crystallized intelligence (Cattell, 1941, 1943, 1963), which should actually be called Cattell and Hebb's theory (Brown, 2016), describes two distinct forms of intelligence that are used to answer different needs. Fluid intelligence, Gf, is the ability to adapt and solve new problems using logic, whereas crystallized intelligence, Gc, reflects the ability to use knowledge and experience. For example, a person’s ability to perform successfully on a history test would require crystallized intelligence (Cattell, 1963). Additionally, Cattell argued that Gf and Gc differ in their evolution over time; specifically, a person’s Gf tends to decrease from 22 years old onward, while the evolution of Gc depends more broadly on learning opportunities (Cattell, 1963). Cattell explicitly juxtaposed his theory against that of Spearman (1904), arguing that “there is not one ‘general ability’ second-order factor, as in the Spearman-Thurstone resolution” (Cattell, 1943). However, the notions were reconciled later on with the CHC theory (Cattell, Horn, and Carroll, presented in Ortiz, 2015) where *g* affects both Gf and Gc or with some researchers claiming that Gf and *g* were the same element (review in Burkart et al., 2017). That said, despite their extensive development and refinement, none of these early theoretical frameworks of intelligence were explicitly applied to non-human animals.

From a more physiological point of view, the size and shape of a cranium were also once thought to be indicators of human intelligence. This belief stemmed from the pseudoscience of phrenology, introduced by Dr. Franz Joseph Gall in the late 1700s. Gall allegedly observed a link between the ability to memorize texts and the size of university students' eyes (Combe, 1824). This idea was further developed with claims of a positive correlation between cranial volume and cultural evolution (Hankins, 1926), as well as observed cognitive challenges among people with smaller-than-normal brains, historically referred to as 'microcephalic idiots' (Paterson, 1930, p.81). These theories were used to justify social and cultural hierarchies (Conklin, 2021) but were fortunately abandoned as society progressed into the 20th century. In contrast, a more politically and scientifically acceptable approach derived from phrenology involves using relative brain size as an indicator of intelligence, thereby reducing the emphasis on societal hierarchies based on "races" and "sex". That said, even relative brain size does not adequately explain the evolution of complex mental abilities in the human species (Henneberg, 1998). As technology has advanced, researchers have increasingly pursued more detailed investigations into the relationship between neurophysiology and intelligence. Potential factors such as neuronal count (Marhounová et al., 2019) and/or characteristics (Dicke & Roth, 2016; Goriounova et al., 2018), the characteristic of astrocytes (Oberheim et al., 2006), or even more precise measurements such as the speed of transmission (Reed et al., 2004), have all been explored. Nevertheless, despite our growing knowledge, the complete story of the origin of intelligence remains unwritten as contradictory results continue to emerge (e.g. Songthawornpong et al., 2021).

Nowadays, humans can program certain forms of intelligence, such as Artificial Intelligence (AI). Out of curiosity and for the purposes of this thesis, I prompted the AI chatbot ChatGPT to define "intelligence". ChatGPT states that intelligence is a "complex and multifaceted concept that encompasses a range of cognitive abilities, skills, and processes". ChatGPT included a cognitive component to its definition, considering "reasoning, problem-solving, abstract thinking, comprehension, learning from experience, and memory", as well as "adaptability, learning and knowledge application", as indicators of intelligence. Furthermore, ChatGPT also included an "emotional and social intelligence" aspect to its definition, and highlighted different factors that might influence one's perception of what is or is not intelligent, such as cultural influences. It concluded with the following caveat: "The understanding and measurement of intelligence continue to evolve as research in psychology, neuroscience, and related fields progresses". Overall, ChatGPT offered a comprehensive overview of current perspectives on intelligence. When asked, "Are you intelligent?", ChatGPT responded with a nuanced answer, stating that it can "simulate aspects of human intelligence" but also clarifying how its intelligence differs from human intelligence. It explained that its capabilities are "limited to processing and generating language based on patterns in my training data" and that it lacks "consciousness, self-awareness, or the ability to experience and understand the world in the way humans do". I invite curious readers to discuss intelligence with AI programs, as there is some circular irony in the process.

Despite a lack of mean to measure, compare and understand the whole complexity of human intelligence, there is a catalogue of mental abilities that can be observed and measured in individuals, allowing comparisons. All the identifiable mental abilities used to acquire, store and use the information perceived from the environment fall under the umbrella of "cognition" (Shettleworth, 2009). While often used interchangeably with the term "intelligence", cognition encompasses broader and more complex abilities than those measured via IQ tests, facilitating comparisons across species with analyses from an absolute or an evolutionary perspective.

Overall, this contributes to the broader investigation of a question that is difficult to answer: What makes a species smart?

One way to approach cognition is to link it to environmental needs. In other words, what is the environmental need filled by the presence of a certain cognitive ability? This is the logic of the ecological approach to cognition (Hutchins, 2010; Kamil, 1998; Shettleworth, 2009) that also helps identify suitable experimental designs. For example, researchers use a food caching paradigm to test for episodic-like memory in scrub-jays (Clayton & Dickinson, 1998), a species that hoards food in summer to retrieve it in winter (Vander Wall, 1990). They also test for transitive inference (i.e., the ability to infer unknown relationships between objects from known ones) in the context of a species facing complex social hierarchy challenges (Grosenick et al., 2007; Tibbetts et al., 2019; White & Gowan, 2013). I followed this logic in this present research.

Numerous recent observations are deepening our understanding and appreciation of the incredible diversity of intelligence in the natural world. For example, an orangutan was recently observed using advanced medicine to treat an open wound (Laumer et al., 2024), and an increasing number of species have been documented using tools of varying complexity (e.g., Mann & Patterson, 2013; Seed & Byrne, 2010). In addition, there is evidence of primary forms of intelligence in plants. For instance, plants can exchange volatile organic compounds to communicate with each other, helping them adapt to changing environments (reviewed by Abbas et al., 2022). Furthermore, simple forms of learning and knowledge transmission have also been found in the *Physarum polycephalum*, despite its lack of a nervous system or specialized neural cells (Boisseau et al., 2016; Vogel & Dussutour, 2016). Famously known as the “Blob”, the unicellular myxomycete found in the forest floor has shown that preliminary forms of cognitive abilities can occur through mechanisms different from brain activity (reviewed in Reid, 2023). This includes learning through changes in cell oscillation frequency (Shirakawa et al., 2011), which can be influenced by substances known to affect mammals' memory abilities (Kippenberger et al., 2023). These findings compel us to consider qualitative, rather than quantitative, differences in cognition and intelligence across organisms, investigate the modulating factors that influence them, and examine potential distinctions between performance and underlying mechanisms. This is what I did with a fish species, a small brain ectotherm species

I.2. Ectotherms and endotherms: two different groups of vertebrates.

Vertebrates can be categorized into two groups based on their ability to regulate body temperature: ectotherms and endotherms. Ectotherms rely on environmental temperature, while endotherms have internal processes to maintain their body temperature. Some ectotherms, particularly from colder environments, can maintain a constant body temperature due to their large size, a phenomenon known as gigantothermy and observed in certain marine turtles and sharks (Nakamura et al., 2017). This goes with the observation that ectothermic from cold environments tend to reach larger adult size (Walters and Hassall 2006). This phenomenon is known as the temperature-size rule (TSR, Atkinson, 1994). Bergmann's rule (see Blackburn et al., 1999 for review) applies to endotherms, stating if two species “differ only in size, the smaller species would demand a warmer climate, to the exact extent of the size difference” (Bergmann, quoted in James, 1970). Lastly, James' rule applies to variations in size within a species in relation to environmental temperature variations (James, 1970).

Another aspect that divides ectotherms and endotherms has been highlighted by Jerison (1969). He established a brain size – body size curve which shows an on average 10-fold difference in relative brain size between these two groups. One possible explanation is related to environmental factors, specifically temperature. Endothermic, mammalian brains generate their own heat (Donhoffer, 1980; Vesterdorf et al., 2011), which is essential for the functioning of neurons (Mednikova et al., 2004) and survival in colder environments. A study on modern cetaceans demonstrated a correlation between brain size and water temperature (Manger, 2006), with larger brains found in colder waters. While absolute brain mass showed low predictability, adjusting for body size using the encephalization quotient (EQ, Jerison, 1973) changed the conclusions: temperature variations explained around 70% of the EQ variation. The study also demonstrated that larger cetaceans, which are more commonly found in colder waters, have brains with greater thermogenic capacities based on various physiological measurements of brain function. Additionally, a recent study has shown that cetacean brains have superior thermogenic abilities compared to artiodactyl species (Manger et al. 2021), further emphasizing the link between environmental temperature, brain size, and thermoregulation.

Another compelling explanation for the differences in relative brain size between ectotherms and endotherms is that these differences may reflect differences in mental abilities. The logic is simple: species with bigger brains should be smarter. While this logic has sparked debate over its ability to fully capture the complexity of human intelligence (e.g. Conklin 2021), it's still worth considering why some species have larger brains compared to others if it does not result in greater cognitive abilities. Maintaining a big brain is energetically costly (e.g. Aiello et al., 2001; Heldstab et al., 2022; Martin, 1981) and thus, should bring some advantages to a species. The "expensive brain hypothesis" predicts that "having a big brain will be compensated for by the organism by reducing the energy allocated to other functions, such as digestion or locomotion" (Isler & van Schaik, 2009). Therefore, we would expect that such costs would be compensated by great benefits, which could include greater mental abilities. Some evidence supporting this has been reported. For example, male guppies with larger brains can more quickly navigate a maze to find a mate (Kotrschal et al., 2015), larger-brained carnivores more quickly open puzzle boxes (Benson-Amram et al., 2016), and species with bigger brains adapt more easily to a new environment during reintroduction events (Sol et al., 2008). In another study, guppies that were artificially selected for larger or smaller relative brain sizes demonstrated that female individuals with larger brains had an advantage in a numerical learning task (Kotrschal et al., 2013).

Given these findings, I was deeply intrigued by, and eager to explore, the question of whether relative brain size in endotherms vs. ectotherms is related to mental abilities. I was particularly interested in investigating potential key differences in mental abilities between ectotherms and endotherms, with a focus on the mechanistic aspects (the 'how') rather than performance outcomes (the 'what').

I.3. Cognition, complex cognition, and executive functions

I.3.1. Definitions

In the previous sections of this introduction, I referred to "mental abilities" or simply "abilities", but this could be more accurately characterized as "cognition" or "cognitive abilities". As discussed previously, cognition encompasses all the brain mechanisms through which an individual acquires, stores, and utilizes information from the environment (Shettleworth, 2009).

This includes processes ranging from automatic functions like perception to more complex mechanisms such as Theory of Mind, a concept that reflects the ability to attribute mental states to oneself and others (Premack & Woodruff, 1978). Various additional abilities are required for this capacity, including self-awareness (i.e., recognizing oneself as distinct from others) and perspective-taking (i.e., understanding different perspectives in a given situation).

To differentiate these various cognitive abilities, those that involve the flexible integration of multiple processes are categorized under “complex cognition” (Knauff & Wolf, 2010). For example, problem-solving, decision-making, and thinking all require interactions between more automatic processes (Knauff & Wolf, 2010). Each of these abilities requires the intervention of other abilities: the “executive functions” (EFs), also referred to as “cognitive control”. EFs refer to a group of top-down cognitive processes that “are recruited when it would be ill-advised, insufficient, or impossible to go on autopilot or rely on instinct or intuition, such as when presented with novel, unanticipated challenges” (Diamond, 2020, p.225).

The three core EFs are inhibition (i.e., the ability to resist behavioral impulses and/or control cognitive interferences), cognitive flexibility (i.e., the ability to adapt to new problems, such as by shifting mental perspectives), and working memory (i.e., a manipulative form of memory that enables goal-directed integration of past and present knowledge; Diamond, 2013, 2020). Although these three EFs are distinct, they are interconnected, such that impairment in one may influence the functioning of the others (as discussed in Diamond, 2013).

I.3.2. Neurological view

Cognitive neuroscience seeks to explain cognitive abilities by identifying their underlying brain-based mechanisms. Given their role in enabling goal-directed, flexible, and non-instinctive behaviors (Diamond, 2013, 2020), the neurological foundations of EFs have garnered significant attention from cognitive neuroscientists.

One common approach to studying EFs is task-based, where specific tasks are associated with individual EF. When performance on a task declines following cerebral lesions, it is inferred that the damaged brain region underlies the EF being tested (e.g., Gläscher et al. 2010). However, as noted by Diamond (2013), the three core EFs cannot always be fully separated, suggesting that this approach may not capture the complete picture.

A complementary approach involves neuroimaging, which provides a more comprehensive view that highlights the involvement of multiple brain areas working in concert to support EFs.

Both task-based and imaging approaches have repeatedly emphasized the importance of the prefrontal cortex (PFC) to EFs in humans (e.g. Barbey et al. 2012; Friedman & Robbins, 2022; McAlonan & Brown, 2003). In avian species, the nidopallium caudolateral (NCL) serves as the functional equivalent to the PFC, playing a key role in decision-making (Güntürkün, 2005; Herold et al., 2011). The NCL has also been linked to cognitive abilities such as spatial learning (Li et al., 2021), serial-order behavior (Johnston et al., 2020), and others (e.g., Güntürkün, 2005; Liu et al., 2018). Similarly, in teleost fish, the telencephalon is the central hub for many comparable cognitive processes (Calvo & Schluessel, 2021; Triki et al., 2023). Given the homology of these brain structures, it is reasonable to infer that EFs are localized in these respective brain areas across species. However, when considering WM as a representative EF, its localization within the brain remains elusive. Current research suggests that WM cannot be confined to a single brain structure (Brzezicka et al., 2019; Christophel et al., 2017; Collette & Van der Linden, 2002; Salehinejad et al., 2021).

I.4. WM as a key EF, underlying many capacities

I.4.1. Main models

WM is the primary mental process that allows an individual to adapt and make decisions based on previous knowledge when facing a new, unanticipated challenge (Diamond, 2013, 2020). The concept of WM was first introduced in 1960 by Miller et al., who described it as a type of memory used for quickly accessing information necessary for executing a plan (Miller et al., 1960).

Over time, the concept of WM has evolved and led to the development of multiple models and definitions. Among these, the Multicomponent model of Baddeley is the most widely accepted in human research. Initially introduced in 1986 (Baddeley, 1986) and continuously refined

Box: Aid for navigating experimental psychology literature on working memory, often referring to multiple other types of memory

Type of memory		Definition	Time interval	Source
Short-term memory (STM)	Sensory memory: Iconic (visual), echoic (auditory), haptic (touch) *	Limited-time storage of information, non-manipulative	Milliseconds to 4 seconds (echoic memory)	Darwin, Turvey, et Crowder 1972; Miller et al. 1960 ; Sperling 1960
	Working memory	Short-term memory storage used for problem-solving in the present thanks to the manipulation of temporarily activated information, time and space limited.	Few seconds to several minutes**	e.g Baddeley 1986; Cowan 1995
Long-term memory (LTM)	Explicit (declarative***): Episodic memory, including prospective memory	The ability to talk about episodes (i.e. specific events) of life, to "mentally time travel". It requires a sense of self, subjective time, and autonoetic awareness (conscious awareness that something happened to us at a specific time in the past). When the self is projected in the future to anticipate future events, we talk about prospective memory (i.e. future thinking)	Minimum of 15 seconds after acquisition of information (Atkinson & Shiffrin, 1968)	Manrique et Walker 2017; Tulving 1972, 2002
	Episodic-like memory	Due to the absence of language, it is impossible to ascertain whether autonoetic awareness accompanies recollection. Therefore, in nonhuman animals what can be reliable ascertained is whether they recall the what, where and when of a past experience, hence it is also called the what/where/when memory (WWW-memory).		(Pause et al., 2013; Suddendorf & Busby, 2003; Tulving, 1972, 2002)
	Explicit (declarative***): Semantic memory	Remembering general facts learned in the past (e.g. the country names on a map), it is the storage of factual information		Manrique et Walker 2017
	Implicit (non-declarative ***)	Memory system used in everyday skills performed automatically without conscious recall		(Manrique & Walker, 2017; Squire & Dede, 2015)

* those are the main types of sensory memory, more are referred to in Wan et al. 2020

** not clearly defined, highly variable in experimental paradigm

***Declarative or non-declarative refers to the ability to express is, using spoken or unspoken (i.e. signing or writing) languages. Hence, this terminology does not apply to animal studies.

Additional information and details can be found in (Squire & Dede, 2015)

(Baddeley & Hitch, 2019) this model defines WM as “A limited capacity system for the temporary storage and processing of information required for complex cognition” (Baddeley et al., 2021, p.10). This model was conceptualized as the authors thought about a system that should facilitate various cognitive functions through short-term memory and attention (Baddeley et al., 2021). It originally comprised three different components: the visuospatial sketchpad, the phonological loop and the central executive (Baddeley, 1986). In a recent paper, Baddeley and Hitch (2019) describe their initial model as “a three-component model comprising an attentional control system of limited capacity, the central executive, aided by two short-term storage systems, one for verbal and acoustic material, the phonological loop and its visual counterpart the visuo-spatial sketchpad” (p.93). A significant development in this model was the addition of the episodic buffer in 2000—a component proposed in response to evidence that verbal and visual systems could be simultaneously influenced during tasks like serial recall by articulatory suppression (Logie et al., 2000). This indicates that the two supposedly independent systems communicate. The episodic buffer is therefore the holding unit that integrates information from the various sub-systems and from episodic memory (as defined in Tulving, 2002), allowing for conscious access to combined data (Baddeley & Hitch, 2019; Logie et al., 2000). Further details on the evolution, unresolved questions, and methodologies behind this model can be found in several key works (Baddeley, 2012; Baddeley et al., 2021; Baddeley & Hitch, 2019).

While the aforementioned research has proved invaluable to the study of WM, this model is not applicable to non-human animals, due to its important link to language. Morrison (2005) and, more recently, Adams et al. (2018) discussed the distinctions between Baddeley’s Multicomponent model and Cowan’s Embedded-Processes model (Cowan, 1988, 1995, 2019), the latter of which is often favored in animal research due to its separation from human language. In the Embedded-Processes approach, WM is defined as “The ensemble of components of the mind that hold a limited amount of information temporarily in a heightened state of availability for use in ongoing information processing” (Cowan et al., 2020). In Morrison’s (2005) words, it is described as “the set of cognitive processes that keep mental representations in an easily accessible state”.

The central component of the Embedded-Processes model is the focus of attention (FoA), nested in an activated portion of the long-term memory (aLTM) (Cowan, 1988, 1995, 2019). This model has important limitations for the amount of information that can be processed in the aLTM, hence emphasizing the importance of attentional control (Cowan et al., 2005; Cowan & Morey, 2006). More recently, Cowan et al. (2024) explored the connection between attention and memory, as well as evidence supporting this model. Further details can be found in Chapter 1 of this thesis (i.e., Working memory in cleaner wrasse).

These two models differ on multiple aspects, some of which are presented in Table 1. Morrison (2005) argued that one of the main differences between the Multicomponent model of Baddeley and the Embedded-Processes approach of Cowan lies in how they handle different types of information. While Baddeley has separate modules for verbal and visual or spatial information (i.e., the phonological look and the visuospatial sketchpad, Baddeley, 1986; Baddeley & Hitch, 2019), Cowan’s model has a unique process where a domain-general central executive can activate information of different nature in the LTM (Cowan, 1988, 1995; Cowan et al., 2024).

The conceptual frameworks discussed here offer important insights into designing appropriate experiments, especially considering WM’s time- and capacity-limited nature, and its connection to attention reflected by sensitivity to disturbance (Carruthers, 2013). That said, the limitations of these approaches raise critical questions regarding how to interpret WM performance in animals. For instance, the impressive WM performance observed in chimpanzees (Inoue & Matsuzawa, 2007; Völter et al., 2019) has been challenged in recent publications that found previous interpretations were unsupported by evidence (Manrique & Walker, 2017; Read et al., 2022).

Table 1: Comparisons of different points of the two main WM models
Incompletely extracted and partially reworded from Baddeley et al., 2021 and Cowan et al., 2020.

Questions	Multicomponent model	Embedded-processes model
1. Definition of WM	A limited capacity system for the temporary storage and processing of information required for complex cognition	The ensemble of components of the mind that hold a limited amount of information temporarily in a heightened state of availability for use in ongoing information processing
2. Used methodologies	Experimental effects of concurrent tasks and material; neuropsychological and developmental evidence; effects of strategies; computational modelling	All methods are relevant, e.g. accuracy, reaction time, error distributions, computational modelling, electrophysiology, neuropsychology, functional magnetic resonance imaging
4. Role of attention and control	Distinction between selective attention to the external environment and internally driven attentional control. The central executive is principally concerned with internally oriented attentional control processes but also has a role in the attentional selection of perceptual information	Both are critical in shuttling items into and out of the focus of attention, allowing binding and conceptual formation to take place.
6. The role of long-term memory (LTM)	WM influenced by different LTM including semantic, episodic, and procedural, at different stages and different ways.	Critical to allow a small-capacity focus of attention to deal with a large amount of complex information through chunking and concept formation

I.4.2. Testing methodologies

The various methodologies used to study WM can be broadly classified into tasks based on (i) visual or (ii) spatial information. In some cases, both types of information are presented simultaneously (e.g. Völter et al., 2019), but such tasks are generally categorized as visual information. While an additional group of tasks involving planning also exists (Tecwyn et al., 2013; Völter & Call, 2014), these tasks often lack the methodological rigor necessary for conclusive results (Manrique & Walker 2017) and associative learning can often account for the observed outcomes of these tasks, making it difficult to isolate the role of WM.

Visual tasks have been employed extensively in WM research but are subject to significant biases. Unconscious mechanisms, such as familiarity (i.e. the ability to react to a stimulus without precise recollection of the information linked to it), may influence task performance. This can result in successful results that are attributed to WM, but in reality, they may reflect automatic, non-conscious processes (differences between familiarity and recollection in Yonelinas, 2002). In monkeys, such tasks are often referred to as "recognition tasks" (Bachevalier & Mishkin, 1986; Zola et al., 2000). In contrast, spatial tasks are considered more robust when discussing WM. There is a general consensus that "spatial familiarity" does not really exist. I would argue that further research is warranted to adequately answer this question, but these tasks have provided valuable insights into WM research. A list of commonly used tasks is presented in Table 2 with the reference and the conclusions of the study, though this list is not exhaustive. Research on WM in apes has been explored and discussed in detail by

Manrique et al., 2024; Manrique & Walker, 2017; and Read et al., 2022. Consequently, this discussion will focus on WM studies in non-primate vertebrates, as the methodologies for primates differ significantly from those of other species.

It is important to note that, despite the well-established link between WM and attention (e.g., Cowan et al. 2024), sensitivity to disturbance is often overlooked in WM studies (Table 2). WM is a type of short-term memory (STM) that is distinct from other forms of STM due to its manipulative nature. While sensory memory, a subset of STM (see box), is divided into modality-specific subsystems (e.g., visual or auditory), WM is used to combine and manipulate information. However, in some tasks (Table 2), the manipulative aspect of WM is not clearly demonstrated in our opinion.

For example, in the free Y-maze exploration task (Fontana et al., 2021), it is unclear what information is being combined or manipulated, and there does not appear to be a problem-solving component of the task. Similarly, when object permanence paradigms are restricted to Piagetian stage 4 (i.e., single visible displacement; Hoffmann et al., 2011; Piaget & Cook, 1952), such as in the studies of Fiset et al. (2003) and Triki et al. (2023), there is also a lack of information manipulation. In these cases, the subject is merely required to remember the location of a single object, which primarily involves object persistence (the understanding that an object continues to exist, Kawachi & Gyoba, 2006), only remembering the *where*. Introducing an additional element, such as hiding two objects but rewarding the retrieval of only one, could incorporate the manipulation of multiple pieces of information, thus engaging WM more fully. This criticism does not apply to stages 5 and 6 where knowledge update and/or invisible displacements are involved, and thus require more complex information manipulation (Hoffmann et al., 2011; Piaget & Cook 1952). Lastly, I want to point out the issue of the time delay. The connection between WM and LTM complicates this issue. The connection between WM and LTM complicates this issue, but I argue that researchers should pay more attention to the time delays between information acquisition and use in their paradigm. This will help to account for the impact of different delays on performance, leading to more nuanced discussions and conclusions and ultimately, a deeper understanding of WM.

I.4.3. Linked to other abilities

WM is a fundamental EF, and it is logically linked to a wide range of cognitive abilities, as evidenced by the diversity of tasks used to assess it (Table 2). For example, WM training has been associated with improvements in crystallized intelligence (Gc, Alloway & Alloway, 2009) in humans. In addition, WM and general intelligence, as reflected by the Spearman *g* factor, are tightly linked (Cowan, 2022; Frischkorn et al., 2019; reviewed and discussed in Conway et al., 2003). Ackerman et al. (2002) also found that a battery of seven WM tasks was correlated with both cognitive (e.g. verbal, numerical and spatial ability) and perceptual abilities (e.g. dual comparisons of numbers, names, naming symbols) in humans—highlighting important links and differences between WM and Gf or *g.* Furthermore, WM, sometimes assessed by the backwards Digit Span tests, is also a main predictor of Theory of Mind in middle school children (Lecce & Bianco, 2018).

In non-human species, such as mice, WM abilities have been shown to correlate with learning performance on a battery of five distinct learning tasks, suggesting that WM underpins general learning abilities (Kolata et al., 2005). As previously noted, OP is one of the cognitive abilities often used to assess WM (e.g. Doré et al., 1996; Triki et al., 2023). Other abilities, such as

Table 2: Working Memory (WM) tasks, references, conclusions and remaining questions

The ‘Conclusions’ column indicates what the authors of the respective studies stated, without indication of whether I agree with it or not. The list is not extensive, however it gives a valuable overview of the tasks that are used to test for WM in vertebrates.

Group	Task	Species	conclusions	Questions	Reference
Spatial	FMP Y-maze; look at spontaneous arms alternation; 15min free exploration recording	Zebrafish	Modification of spontaneous alternation following drugs administration indicates an enhanced WM	Manipulative aspect? Sensitivity to disturbance?	(Fontana et al., 2021)
Spatial	Delayed non-matching to position task; 20 or 90sec delay	Dog (Beagle)	A dietary supplement can have a positive effect of WM	Sensitivity to disturbance?	(Fragua et al., 2017)
Spatial (compared with visual)	Search behaviour in a task similar to “A-not-B” error task (Piagetian single invisible displacement task), same or different hiding screens	Cats and dogs	Both have WM, but dogs can update more efficiently; visual information is not used	Search behaviours more related to dogs than cats’ ecology? Sensitivity to disturbance?	(Doré et al., 1996)
Spatial	8-arms radial maze, reward in all arms, trial stops when all eaten; free or forced exploration (4 arms accessible, then 4 others after a 5min delay); reference memory task (rewards have different value)	Dogs	Dogs succeeded in WM, lower rates than rats	Time delay between reward eaten? Sensitivity to disturbance?	(Macpherson & Roberts, 2010)
Spatial	Spatial orientation in a rectangular environmental geometry (or adding a stripped wall as a spatial mapping cue in a square arena)	Mice	Mice have spatial WM, use the mapping cue only partially	Sensitivity to disturbance?	(Lee et al., 2015)
Spatial	Y-maze, spontaneous alternation or blocked arm	Mice	(methodology paper)	Sensitivity to disturbance? Spontaneous alternation, a reliable indicator?	(Kraeuter et al., 2019)
Spatial?	Object permanence (Piagetian single visible displacement, delayed before search, 0 to 240sec)	Dogs	Success remains above chance level even at 240 sec, WM abilities of dogs go up to that delay.	Sensitivity to disturbance? Manipulative aspect?	(Fiset et al., 2003)
Spatial?	Object permanence (Piagetian single visible displacement, less than 10sec)	Guppies	Success at 60% linked to bigger telencephalon size, guppies have “at least substantial WM” (p.6)	Sensitivity to disturbance? Manipulative aspect?	(Triki et al., 2023)
Visual	Delay matching-to-sample (DMTS), few seconds to few minutes	Mammals (primates, monkeys, rats), birds and bees	No major evidence of inter-species differences; performances depend on delay	Sensitivity to disturbance? Familiarity and recognition?	Review in (Lind et al., 2015)*
Visual	DMTS (3 to 5sec)	Zebrafish	Zebrafish have WM, but success is limited to certain conditions (training, type of response)	Sensitivity to disturbance? Familiarity?	(Bloch et al., 2019)

* I invite readers who want to know more about DMTS to read the review and related papers. In order to keep this table at an acceptable length, I am not adding multiple other studies on DMTS paradigm that were done prior to this review.

transitive inference (TI), thought to be prevalent in species facing complex social challenges (Tibbetts et al., 2019; White & Gowan, 2013), could potentially also rely on WM (Libben & Titone, 2008). Both OP and TI have been tested in numerous taxa (see Chapters 2 and 3, respectively), allowing for cross-species comparisons that include both ectotherms and endotherms. Therefore, we tested for the presence of OP and TI with a small-brain ectotherm species: the Bluestreak cleaner wrasse *Labroides dimidiatus*.

I.5. Why cleaner wrasse?

The cleaner wrasse *Labroides dimidiatus* has been the subject of extensive observation for the past two decades. This tropical reef fish is distributed throughout the Indo-Pacific region, ranging from the Red Sea to the eastern Australian Great Barrier Reef. Cleaner wrasse play a crucial ecological role in maintaining the health of their environment, and their presence can even increase overall fish species diversity within coral reef ecosystems (Bshary, 2003; Grutter et al., 2003). Indeed, cleaner wrasse feed on ectoparasites, dead skin, and infected tissues from other fish species, known as clients (Côté, 2000), and reduce stress levels in clients through a massaging technique, called "tactile stimulation", with their pelvic fins (Bshary et al., 2007; Ros et al., 2011; Soares et al., 2011). Cleaner wrasse also establish territories, called cleaning stations, where their cleaning interactions with clients take place (Côté, 2000).

Cleaner wrasse exhibit strategies in the wild that potentially rely on complex cognitive abilities. First, maintaining the cooperative aspect of cleaning interactions requires certain strategies. Indeed, cleaner wrasse sometimes cheat by eating the client’s protective mucus, which they prefer compared to parasites (Grutter & Bshary, 2003a). However, this cheating has consequences. Resident client species that share territory with cleaner wrasses will punish cheating cleaner wrasse by chasing them away (Bshary & Grutter, 2005; Bshary & Noë, 2003; Grutter & Bshary, 2003a). In contrast, visitor client species, which do not share the same territory, can choose to leave and seek a more cooperative cleaner wrasse (Bshary & Grutter, 2005; Bshary & Schäffer, 2002; Soares et al., 2013). Cleaner wrasses often prioritize visitor species, offering them better service to ensure they do not leave (Bshary & Grutter, 2002). Cleaner wrasse also provide superior service to predatory species (Bshary & Würth, 2001).

Other factors external to the cleaning interaction itself can influence the cooperative aspect. The quality of service actually improves when cleaners work together in pairs (Bshary et al., 2008). Image-scoring among client species increases the likelihood of cooperation (Bshary & Grutter, 2006), and an “audience effect” has been documented where cleaner wrasse modify their behavior in the presence of an audience (Pinto et al., 2011a). Cleaner wrasse can act on these effects. For example, they can manipulate their clients (Bshary & Würth, 2001) by providing tactile stimulation to resident species, in order to improve their reputation (Bshary, 2002).

In addition to their social interactions with clients, cleaner wrasse face social challenges related to their reproductive strategy. They are protogynous hermaphrodites, meaning all individuals are born female and can change into males depending on size and environmental factors (Nakashima et al., 2000). Males maintain harems of females, each of which has its own cleaning station within the male's territory (Robertson, 1972; Sakai et al., 2001). A size-based hierarchy exists, with the largest female, the alpha, potentially posing a competitive threat to the male. As a result, males spend more time interacting with the alpha female and adjust their behavior accordingly (Raihani et al., 2012).

In summary, despite having a relatively small brain, cleaner wrasses exhibit highly sophisticated intra- and inter-species behaviors, making them a fascinating species for the study of animal cognition. Research into their cognitive abilities has revealed processes that go beyond simple Pavlovian or operant conditioning (Pavlov, 2010; Skinner, 1937; Thorndike, 1927). Cleaner wrasse exhibit social learning (Truskanov et al., 2020), generalized rule learning (Wismer et al., 2016), mirror self-recognition based on a mental representation of the fish's own face (Kohda et al., 2019, 2022*a*, 2023), self-control (Aellen et al., 2021), and behavioral flexibility (Triki & Bshary, 2021). A number of these abilities relate to the ecological needs of cleaner wrasse (Gingins & Bshary, 2016).

I.6. This PhD

This PhD research builds on the work of Mélisande Aellen, a former PhD student who studied the cognitive abilities of cleaner wrasse. Aellen demonstrated that cleaner wrasse can perform on par with primates in a delayed gratification experiment (Aellen et al., 2021), generalize a concept in a match-to-sample experiment (Aellen et al., 2022), and exhibit cognitive flexibility and inhibitory control (Aellen et al., 2022), though they appear to lack general intelligence (Aellen et al., 2022). Building on these previous findings, we aimed to investigate what cognitive mechanisms might underlie this lack of general intelligence. Specifically, we sought to test for the presence of WM in cleaner wrasse, given that this EF is considered a principal predictor of *g* (e.g., Conway et al., 2003; Kolata et al., 2005). In addition, we decided to investigate two other abilities that can allow cross-species comparisons, namely TI and OP, where WM is thought to have a role, where the ecological relevancy is tested and where we addressed several methodological issues highlighted in the literature.

I.6.1. Chapter 1: Working memory in cleaner wrasse

The aim of this chapter was to investigate WM in cleaner wrasse, employing methodologies that could allow cross-species comparisons and that address the issues raised in section IV of this introduction. Additionally, we wanted to explore an ecological dimension of WM by examining whether variations in the context of experimental paradigms (i.e. abstract *vs.* ecologically relevant) would influence the performance of cleaner wrasse. While WM has been extensively discussed earlier and will be again later in this chapter, here I only provide a brief summary of our work related to it.

Our research encountered several challenges, particularly regarding the cleaner wrasse's responses to some of our paradigms, which necessitated adaptations to our experimental design. This led us to return to more basic dual-task paradigms, where, in cases of positive results, familiarity alone could have accounted for performance. Despite our series of unsuccessful paradigms (detailed in the supplementary materials), we obtained consistent findings across other designs. Our study offers novel insights into the understanding of WM and helps shed light on the brain-size gap between endotherms and ectotherms highlighted by Jerison (1969). In particular, our findings strongly suggest that cleaner wrasse do not possess WM abilities. This contrasts with previous studies that reported WM abilities in other fish species (e.g., zebrafish and guppies; Fontana et al., 2021; Triki et al., 2023). I discuss potential shortcomings in the design and interpretation of these previous works that may have led to spurious conclusions regarding WM. Given these discrepancies and the clear lack of WM observed in cleaner wrasse, we are skeptical of its presence in other fish species and recommend further investigation. About other ectotherm species, currently, no research on reptiles or amphibians exists to support or challenge our hypothesis that significant differences in brain size are either a consequence of, or a factor contributing to, a lack of WM.

As a result of the lack of WM, I would argue that species must have other strategies to overcome their environmental challenges. In the case of cleaner wrasse, I hypothesize that familiarity, episodic-like memory, and routine behaviors shaped by associative learning may serve as compensatory mechanisms. However, further studies are needed to confirm or refute this hypothesis.

I.6.2. Chapter 2: In the absence of extensive initial training, cleaner wrasse Labroides dimidiatus fail a transitive inference task (published)

The primary objective of this chapter was to test for the presence of a widely studied cognitive ability—transitive inference (TI)—in cleaner wrasse. TI presence or absence in a species where its ecological necessity is unlikely can inform our understanding of the ecological approach to cognition (Shettleworth, 2009), which justifies the presence of a cognitive ability by an environmental need. Especially that TI has also been documented in invertebrate species (Tibbetts et al., 2019), reinforcing the notion that its presence does not require a large brain. These two factors make TI an intriguing ability to examine in a small-brained ectothermic species, as both positive and negative results could provide valuable insights into the relationship between ecology, brain size, and cognition.

During the course of our study, another research group published findings on the same topic, concluding that cleaner wrasse were capable of solving a TI task (Hotta et al., 2020). Given that our research led to different conclusions, we compared methodologies and identified that the success of cleaner wrasse in solving TI tasks was largely dependent on extensive training. This suggests that external support could compensate for the lack of ecological relevance and the absence of a large brain. Furthermore, our results reinforce the notion that TI is closely linked to WM (Libben & Titone, 2008), as neither cognitive ability was observed in cleaner wrasse. These findings add to our broader understanding of the cognitive capabilities of small-brained species and the potential compensatory mechanisms they may rely on in the absence of more sophisticated cognitive functions.

I.6.3. Chapter 3: Object permanence in cleaner wrasse

This final chapter investigated OP in cleaner wrasse. Here I provide an extended definition of OP as well as details on its origin that I did not provide in the chapter directly. OP is the ability to remember the existence of an object even when out of sight. Piaget and Cook (1952, 1954) first studied OP in human infants, where they identified six stages of total OP capacity development, referred to as the Multiple Invisible displacements, that were linked to age (Piaget & Cook, 1952, 1954). The first two stages are based on reflexes and chance; children younger than four months old do not actively create schemes but will try to replicate what happens accidentally on their bodies. At stage 3, children between 4 and 8 months old are also able to incorporate external objects into their schemes, as well as recover partially hidden objects. At stage 4 (8-12 months), children begin to demonstrate intentionality by combining known action schemes (e.g., pushing an object away to later pull it closer) and actively exploring objects to understand how they can interact with them. They can also now retrieve a completely hidden object, but they always look for it at the first location, even if it is subsequently visibly moved. They can only solve a single visible displacement, commonly referred to as the "A-not-B error" or the "typical reaction" (Piaget & Cook, 1954). These two stages already require a representation persistence, meaning the object will be recognized after its temporary disappearance as equal to the object seen before the disappearance (Kawachi & Gyoba, 2006). At stages 5 and 6, children from 12 months onward can experiment within the known scheme. They can solve multiple visible displacements tasks (no more “A-not-B error”) and, at a later stage, invisible displacement (the object is hidden in the container, and the container moves to another location) and multiple invisible displacements. They repeat known schemes and create unknown endings by changing some characteristics. As miniature scientists, children try to understand the links between the object and the world. Additionally, they start to be able to have mental representations of those schemes. They no longer need to perform those activities; they can mentalize them. Imagination starts to be able to replace direct actions.

Uzgiris & Hunt (1975) later developed the Uzgiris-Hunt Ordinal Scales of Psychological Development (U-H scales). Whereas Piaget’s theory relies exclusively on the age of the subject, the U-H scale allows a more progressive change in the development of the cognitive toolkit that also gives more space for ecological factors to modulate abilities. OP belongs to the U-H scale 1, subdivided into different complexity levels. Several studies have used this approach to test for the evolution of cognitive skills in other species (see for example Hoffmann et al. 2011; Zucca et al. 2007).

The advantages of studying OP in cleaner wrasse extend beyond the use of a design that has been widely applied to other species; it also offers the flexibility to introduce variations and complexities. This approach allows us not only to determine which Piagetian stage(s) the cleaner wrasse can solve but also to incorporate manipulations addressing related questions, such as the role of memory by introducing a delay between hiding the object and releasing the cleaner wrasse. Similar to our methodology in other chapters, we also explored the ecological relevance of the task by varying the objects used. Our experimental design was informed by the findings of Aellen et al. (2022) where cleaner wrasse failure could have been due to the positioning of hiding locations, potentially leading them to adopt a simple side-preference strategy that resulted in random performance levels. Additionally, we included Piagetian stage 3 to investigate whether this intermediate step might reveal the limits of cleaner wrasse’s cognitive abilities. Our findings showed performance variations in stage 4 that were dependent on prior exposure to stage 3, underscoring the importance of training. The context (reflecting different ecological relevance) did not convincingly influence the performance of cleaner wrasse but still deserves further testing due to some biases.

The performance of cleaner wrasse in this task was similar of even higher to that of some endotherms (e.g. Caicoya et al., 2021; Huffeldt, 2020) in stage 4, but the failure in stage 5 supports earlier research on fish (Aellen et al., 2022; Guadagno & Triki, 2024; Triki et al., 2023), suggesting that only species with more complex brains, which also demonstrate WM abilities, are capable of solving the higher stages of OP (e.g., Barth et Call 2006; Pepperberg et Funk 1990; Zucca et al. 2007).

1. Chapter 1: Working Memory in cleaner wrasse *Labroides dimidiatus*

Leonore BONIN, Héctor M. MANRIQUE, Redouan BSHARY

In preparation for submission

1.1. Abstract

Working memory (WM), a short-term storage system responsible for the manipulation and integration of past knowledge with present information for goal-directed behavior, is a key executive function and a principal predictor of general intelligence. This study investigates WM in a small-brained ectothermic species, the cleaner wrasse (*Labroides dimidiatus*), which faces environmental challenges potentially requiring complex behavioral adaptations. Given that past methodological inconsistencies have led to biased conclusions across various species, we designed experiments that incorporated the critical components of WM, facilitating cross-species comparisons and accounting for potential ecological influences. Despite encountering several experimental challenges, we consistently obtained negative results across multiple experimental paradigms, thereby calling into question previous studies that reported positive WM findings in other fish species. As WM is tightly linked to other cognitive abilities and to the general intelligence factor, we propose that its absence may be a key factor underlying the significant encephalization quotient gap between ectotherm and endotherm vertebrate species.

1.2. Introduction

1.2.1. Ectotherms vs. endotherms

Among vertebrates, brain size relative to body size exhibits considerable variation (Jerison, 1969). On average, endotherms possess brains that are approximately ten times larger than those of ectotherms when corrected for body size (Jerison, 1969). Given the high energetic demands of brain tissue (Fonseca-Azevedo & Herculano-Houzel, 2012; Sukhum et al., 2016), an important question arises: what advantages do larger brains confer to endotherms? One hypothesis is that endotherms possess a more extensive cognitive toolkit compared to ectotherms. While both groups exhibit basic learning strategies, such as stimulus-response associations seen in Pavlovian conditioning and the acquisition of new behaviors following reward presentation (Pavlov, 2010; Thorndike, 1927), these strategies do not require flexible manipulation of mental objects. Whether ectotherms are capable of "complex cognition", these abilities that require executive control, such as problem-solving, decision-making, and reasoning (Knauff & Wolf, 2010), remains poorly understood and warrants further exploration.

A complementary hypothesis regarding the differences in brain size between endotherms and ectotherms relates to the sophistication of their executive functions (EFs). These higher-level cognitive skills are central to complex cognition, as they involve building intricate interactions among mental processes of varying complexity (Diamond, 2013). The three primary EFs—inhibitory control, cognitive flexibility, and working memory—“are recruited when it would be ill-advised, insufficient, or impossible to go on autopilot or rely on instinct or intuition, such as when presented with novel, unanticipated challenges” (Diamond, 2020). Although research on EFs in ectotherms is scarce, some studies suggest that fish, for example, perform well in tasks assessing cognitive flexibility as shown in reversal learning (Aellen, Burkart, et al., 2022; Parker et al., 2012) – where performance correlates positively with brain size in primates (Deaner et al., 2007) – and demonstrated proficiency in inhibitory control tasks, such as delayed gratification paradigms (Aellen et al., 2021) and detour tasks (Lucon-Xiccato & Bisazza, 2017; MacLean et al., 2014). In this study, we investigated the role that working memory (WM), a critical EF, could play in ectotherms’ cognition.

1.2.2. WM: definition and importance

WM is among the most well-studied EFs in humans, with extensive research dating back more than 80 years to Miller et al. (1960), who referred to “the memory we use for the execution of plans as a kind of quick access working memory” (p. 65). While numerous models of WM have been proposed, we focus on the Embedded-Processes model of Cowan (1988, 1995), which has undergone recent refinements (Cowan et al., 2024), and it is not reliant on language. We redirect readers to appropriate literature for descriptions and debates about other models (e.g. Baddeley & Hitch, 1974; Chai et al., 2018; Conway et al., 2003; Cowan, 1988, 1995; Cowan et al., 2024; Fuster & Bressler, 2012; Logie et al., 2021; Morrison, 2005). The Embedded-Processes model consists of a focus of attention (FoA) within a temporarily activated portion of the long-term memory (aLTM). Cowan et al. (2024) defined the FoA as a “coherent representation of several separate items or ideas guiding current thoughts and actions” and the aLTM as “the part of memory from which information is in a heightened state of accessibility”. Hence, in this model, a new stimulus can enter either consciously with the intervention of the central executive, or unconsciously into the aLTM. Part of the information related to that stimulus enters directly into the FoA. Thanks to the central executive, some knowledge stored in the LTM can be brought into the FoA to “guide focus”. By combining them, new “chunks” are created and can move from the FoA to the aLTM, and finally to the LTM where this new acquired knowledge will remain stored. In line with this model, Manrique et al. (2024) defined WM as a “brain system that provides us with temporary short-term storage and management of perceptual or other information (...), which we need for efficiently (...) carrying out, and updating, such complex cognitive tasks as mental reading, reasoning, forecasting, manipulation, (...)”. This is the definition we followed in this study. The manipulative aspect of WM distinguishes it from short-term memory.

WM is of potential importance because it has been linked to many cognitive abilities. For example, training WM can improve crystallized intelligence (Alloway & Alloway, 2009). WM is also a major predictor of *g* in mice (Kolata et al., 2005)—a species representing an endotherm with rather average brain size-body size relationship (Tartarelli & Bisconti, 2006)—and in humans (discussed in Conway et al., 2003). WM is also correlated with numerous other cognitive and perceptual abilities in humans (Ackerman et al., 2002; Engle, 2010; Unsworth et al., 2015), and the prospective aspect of WM is key to efficient tool crafting (Manrique & Walker, 2017).

At present, the only experiment that has tested the general intelligence factor *g* in fish yielded negative results (Aellen et al., 2022), suggesting that their WM might indeed be more primitive and/or not used in a domain-general way. Alternatively, ectotherms might have WM capacities similar to endotherms, but these capacities could be modular and only elicited under ecologically relevant circumstances.

1.2.3. Testing WM

There are multiple definitions of WM, as well as mixed opinions on how best to test it. Cowan et al. (2024) argue that “Discrepancies between methods or definitions often underlie discrepancies between results rather than unreliability of evidence”, highlighting the importance of a clear and common methodology. Testing WM in animals presents distinct challenges, especially in the absence of language. Tasks commonly used in human studies, such as verbal WM tasks (Hu et al., 2019), are not feasible for animals.

While non-verbal tasks have yielded interesting results, caution must be exercised due to the limitations of these methodologies, including the lack of a manipulative or problem-solving aspect. This problem is readily apparent in some spatial tasks. For example, in zebrafish, free exploration of a Y-maze and spontaneous alternation has been taken as evidence of WM (Fontana et al., 2021). Similarly, testing for the Piagetian stage 4 of object permanence abilities has also been taken as evidence of WM in cichlids (Guadagno & Triki, 2024) and in dogs (with time delay, Fiset et al., 2003) even though it is not evident that recovering the object involves the mental manipulation of temporarily stored information. This criticism does not apply to higher stages of object permanence (i.e. Piagetian stages 5 and 6) where the position of the object must be subsequently updated, thus imposing higher demands on WM. These latter stages have already been used to test for WM in corvids and psittacids (Pepperberg & Funk, 1990; Zucca et al., 2007).

Zebrafish performance on a delayed-match-to-sample task has been taken as evidence of WM (Bloch et al., 2019), and the same paradigm have also been applied to honeybees (Zhang et al., 2005) yielding negative results. However, these types of tasks rely heavily on visual perceptual information and could be seen as “recognition tasks” (Bachevalier & Mishkin, 1986; Zola et al., 2000). If so, they could be solved by relying on non-conscious processes, such as familiarity. Whether recognition of a previously presented stimulus is a reliable measure of WM is debated (Yonelinas, 2002), and it would be crucial to test whether success in a WM task depends on recollection or simply familiarity.

Recent studies have also identified methodological limitations when testing WM in nonhuman primates, such in the “limited-hold memory task” applied to measure memory of chimpanzees (Inoue & Matsuzawa, 2007). This task may instead rely on some imagery mechanisms that do not tap into the WM (Abdi et al., 2022; Manrique & Walker, 2017; S. Miller & Peacock, 1982; Özbaydar, 2012; Read et al., 2022). Many other methodological and conceptual limitations have been identified in relation to the measurement of WM in great apes, which are discussed extensively in Read et al. (2022) and Manrique et al. (2024).

Another important consideration when testing WM involves selecting the appropriate time interval separating the acquisition from the retrieval of information (Carruthers, 2013). This is not trivial, as WM is connected to other short-term storages, but also, critically, to the activation of portions of information stored in long-term memory. Further involvement of long-term memory in the resolution of a task might also depend on whether participants have time to combine the information (Morey & Cowan, 2018). Sensory memory, the non-manipulative short-time storage, is known to last from milliseconds to a maximum of 4 seconds (Darwin et al., 1972; Sperling, 1960). Previous studies often failed to sufficiently explain the specific time interval chosen to elicit WM; measurements can range from 0 sec (e.g., Fiset et al., 2003) to 15 min (Fontana et al., 2021), which can easily lead to confusion and mixed results. In spatial tasks, such as in maze experiments (Kraeuter et al., 2019), a potential shortcoming is that the time intervals used, and the number of items to be remembered, may link performance more to long-term WM (Carruthers, 2013).

The last issue we want to discuss is the involvement of attention, which is notably highlighted in Cowan's Embedded-Processes model (1988, 1995). Sensitivity to disturbance is a critical characteristic of WM testing (e.g., Angelopoulou & Drigas, 2021; Cowan, 1988; Cowan et al., 2024) that has been largely overlooked in previous publications (e.g. Doré et al., 1996; Lee et al., 2015; Macpherson & Roberts, 2010).

Despite these methodological challenges, cross-species comparisons of WM can provide valuable insights. WM may vary quantitatively rather than qualitatively among species. For example, humans can maintain 3-5 objects in their FoA, but are also able to double this number through exerting executive control over the contents of memory (Manrique et al., 2024; Read et al., 2022). In contrast, other species with lower cognitive capacities may access only one or two objects and may be less capable of combining information to form higher chunks (Cowan, 2022; Cowan et al., 2024).

1.2.4. Our study

We used the cleaner wrasse *Labroides dimidiatus* for our study as it constitutes a suitable model organism to explore the potential cognitive tool kit of ectotherms. Despite exhibiting an average brain to body size-body ratio (Chojnacka et al., 2015), cleaner wrasse possess a remarkable cognitive tool kit (Bshary & Triki, 2022) that includes basic perspective-taking (McAuliffe et al., 2021), transitive inference (under certain circumstances, Bonin & Bshary, 2023; Hotta et al., 2020), generalized rule learning (Wismer et al., 2016), chaining and configurational learning (Prat et al., 2022; Quiñones et al., 2020), true mirror self-recognition (Kohda et al., 2019, 2022a, 2023), and reversal learning (Triki & Bshary, 2021). Particularly relevant to our study is their ability to delay gratification, which is comparable to that of non-human primates (Aellen et al., 2021), despite an apparent lack of the general intelligence factor *g* (Aellen et al., 2022). Furthermore, cleaner wrasse remembered the *when* and the *what* after 2.5 - 15 minutes in a foraging task (Salwiczek & Bshary, 2011). Although this qualifies as episodic-like memory (Manrique & Walker, 2017; Pause et al., 2013; Tulving, 1972, 2002), WM could be involved here because cleaner wrasse seem to adjust and update their behavior flexibly in response to previously acquired knowledge. This prompted us to investigate cleaner wrasse WM using a dedicated experimental paradigm.

To encompass several aspects of the various definitions of WM, we designed experiments that varied in both the level of complexity and the specific context, aiming to provide a comprehensive preliminary assessment of WM in cleaner wrasse. We present the results from four experiments, but started with three others (see supplementary material for additional methodologies and explanations for untested conditions).

The first experiment, referred to as the “windows” experiment, involved placing food items inside Plexiglas rings (windows) attached to a Plexiglas plate, rendering the items invisible when the fish approached the location from the side. We also varied the ecological relevance by using a plain white plate or a fish picture to assess whether this influenced performance. Although this experiment was intended as a training step for a more complex one (see supplementary material), the data provided initial insights into the investigation of WM in cleaner wrasse.

In the second experiment, we used a methodology similar to Salwiczek & Bshary (2011), first offering a single item on the back of one plate, removing the plate, and then reintroducing the same plate alongside a second one that still had a food item. In a third experiment, we modified the methodology so that the objects were always visible to cleaner wrasse, addressing the possibility that they may have misunderstood the task by interpreting the reintroduced plate as a new object, despite its identical color.

For the fourth experiment, we designed a movable arena to test for WM in a spatial task where cleaner wrasse had to make several choices to access and eat all the hidden food items on a compartmentalized plate. In many regards, this was similar to Völter et al.'s (2019) self-ordered search task where chimpanzees had to look for food rewards in different boxes while avoiding repeating a previously searched box, relying either on spatial information or physical features of the boxes.

To ensure that our experiments did not fall into the domain of sensory memory (Darwin et al., 1972; Sperling, 1960), we set a minimum delay of 5 seconds. We set a maximum delay of 10 seconds to limit the possibility that differences in information rehearsing abilities could overshadow differences in WM (Morey & Cowan, 2018) and avoid measuring abilities strictly related to long-term memory (LTM), as previous research suggests that information can transfer into LTM within 15 to 18 seconds (Atkinson & Shiffrin, 1968; WM and LTM discussed in Rhodes & Cowan, 2018 for example). Delay times could be extended in future studies to investigate the influence of delay duration in greater detail.

Given the apparent lack of a general intelligence factor *g* in cleaner wrasse, we anticipated predominantly negative results, particularly in the more complex experiments. One of our main objectives was to assess whether the correlation between WM and *g* observed in other species holds for cleaner wrasse, with the absence of *g* implying limited WM capacity. However, cleaner wrasse frequently encounter multiple-choice scenarios in their natural environment that require the integration of previous knowledge into current decision-making. Thus, we also expected potential positive results in experiments of lower complexity. As demonstrated in rats (Bratch et al., 2016), cleaner wrasse may possess different WM sub-processes, allowing for WM measurements across diverse contexts.

1.3. Material and Methods

Individual practical and common information for the four experiments are presented in table 3.

Table 3: Information for each experiment.

*This table contains information about each experiment in this study. This information is not crucial to understanding the paradigms but are of interest to readers.

Year	Experiment	Dates	Fish	Catching dates	Prior experiment
2021	Windows experiment	7th - 10th of Oct.	18 females, analyses for 9 only due to unsatisfying data collection on the others (by the second experimenter)	Between the 18th and the 20th of September	Yes (designs presented in the supplementary material)
2021	Two plates experiment	23rd of Oct. - 4th of Nov.	Started with 17 and ended with 14 females (same as the windows experiment)	Same as the windows experiment	Same as the windows experiment
2022	Movable arena	3rd - 17th of Mar.	Started with 8 and ended with 6 males	Between 24th and 28th of January	Some yes (exposed to mirrors), no potential bias
2022	Multiple plates	2nd - 15th of Mar.	11 couples of cleaner wrasse and a single female	Between the 22nd of January and the 5th of February	Yes (inspecting laminated pictures), no potential bias
Common points	- **Location**: Mo'orea, French Polynesia - **Catching method**: scuba diving in non-protected reefs near Cook's Bay. We led cleaner wrasse into a barrier net (mesh size: 0.5cm) using handnets. Then, we placed them into zip-loc bags for transport and renewed the water after 45 minutes to ensure enough oxygen. - **Housing**: White opaque plastic aquaria, 70cm (length) x 52cm (width) x 45 cm (height), with a water height of around 30-35cm. The water was in an open circuit, closed during experiments only. - **Basic habituation**: Cleaner wrasse were not tested until 8-15 days post-capture, during which time they learned to feed on plates and became accustomed to the basic material and experimenters. **- Ethic:** The French Polynesian authorities approved the catching and the manipulation of cleaner wrasse. After the experiments, we released the cleaner wrasse at their site of capture. - **Recording:** A camera placed on the experimenter's forehead recorded each trial (except in case of technical problems). - **Analyses**: Made with R software v. 4.3.1 (R Core Team, 2022). For each model, we conducted an analysis of variance to compare two versions: one with the fish identity as a random factor for both the intercept and the slope with respect to time (trial number) and the other only with the fish identity as a random factor. We always kept the simpler versions as those were not significantly different, and they had slightly smaller AIC.				

1.3.1. Windows experiment

For this experiment, we used two different plates: (1) a plain white square (hereafter referred to as "square", 20cm x 20cm, Fig. 1a) and (2) a laminated picture of a butterflyfish (*Chaetodon ornatissimus*, hereafter referred to as "fish", 20cm length x 14cm height, Fig. 1b). We applied two conditions to both plates: (i) basic (with visible black spots marking the food items locations) and (ii) window (items are "hidden" inside Plexiglas windows of 1.5 x 1.5cm, Fig. 1). All plates contained four mashed prawn items positioned on black spots at designated locations: up, right, down, or left, which corresponded to specific areas on the fish plate (e.g., back, gills, belly, and tail) to mimic natural ecological contexts.

Each tested cleaner wrasse underwent ten trials per day for four consecutive days. Of the nine cleaner wrasse, four initially performed 20 trials using the fish plate, while five began with the square plate. The condition (basic or window) for each trial was randomized, ensuring that no more than two consecutive trials featured the same condition. Each condition was tested five times per day.

Each trial followed the same procedure: the cleaner wrasse was placed behind a transparent barrier, with the plate inserted on the opposite side of the aquarium. Once the cleaner wrasse was released, we recorded the sequence in which it ate the food items, and counted the number of mistakes (i.e. the number of times it picked at an already-eaten location). Based on observations during data collection, we also extracted additional information from video footage, including: the total time taken to eat from all locations (if achieved, hereafter called "duration"), instances of "inspection" (i.e., the cleaner swimming around the plate, sometimes visually examining the locations without picking on the black spot), and instances of "leaving" (when the fish moved at least 5 cm away from the plate, marked by a line at the aquarium's base). Both inspection and leaving behaviors, which were grouped under the categorical variable "looks", allowed the fish to re-examine (i.e. look again) the plate, potentially preventing mistakes. These three variables (number of mistakes, duration, and looks) were then used as three distinct response variables to estimate foraging efficiency. Plates were left in the aquarium for a maximum of one minute, but data were only extracted for the first 30 seconds, as eating the four items should occur fairly quickly.

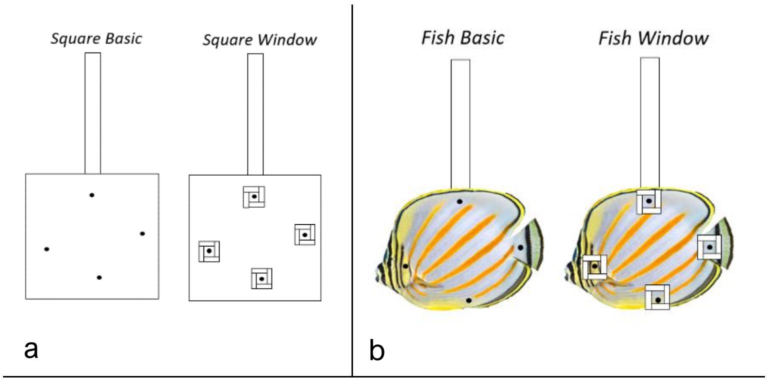


Figure 1: White square (a) and butterflyfish (b) testing plates
We tested cleaner wrasse on four plates: Square or Fish, each in Basic or Window conditions. The black spots held mashed prawn item. Each window was 0.5cm deep.

Note that the day after we finished collecting the data, to verify that what we termed "mistakes" were indeed mistakes, we conducted two post-experimental trials using red-colored prawn items and removed the black spots from both plates, ensuring the items were fully visible to us. Each fish underwent two supplementary trials (one with the square plate and one with the fish plate), counterbalancing the starting plate across subjects.

Statistical analyses

We categorized our measurements into three response variables: the duration of time (maximum of 30 seconds) taken to eat the food items from all four locations, the number of mistakes (repeated location), and the number of "looks". We aimed to assess the effects of ecological relevance (represented by the plate: square or fish), visual information (or complexity, represented by the condition: basic or window), and learning (examined through performance trends across trials). We began with the most complex model, including all possible interaction terms, and performed model selection to identify the best fit.

Discarded trials

Trials in which subjects failed to eat items from all locations within 30 seconds were excluded from the analysis. Additionally, one trial on Day 2 (8th trial) was omitted due to a camera malfunction that prevented data extraction, and one trial for Fish E (Day 1, 10th trial) was excluded because the plate was removed prematurely, preventing access to all locations.

Duration data

Despite applying various transformations to our duration data, the distribution remained problematic, with none yielding homogenous variances. The best-fitting model was a linear mixed-effects model applied to log-transformed duration data, with a transformation parameter of -1.5. The explanatory variables included the plate, condition, and trial number, along with all interaction terms. The identity of cleaner wrasse subjects was treated as a random factor, and heteroscedasticity was modelled between the “basic square” condition and other combinations using the VarFunc class from the *nlme* package (Pinheiro et al., 2007) in R. We simplified the model by retaining only the plate, condition, and trial number as fixed effects.

Mistakes and looks data
As both response variables represented count data, we used generalized linear mixed-effects models with a Poisson distribution. The same methodological approach as used for the duration analysis was applied, leading to similarly simplified models.

Models diagnoses and post-hoc analyses
Model diagnostics were performed using the *DHARMa* package (Hartig, 2022) in R, where we examined outliers and checked for model assumptions. We retained all outliers in our analyses as there was no valid justification for their removal. Overdispersion was checked when required. For post-hoc analyses, we employed least-square means methods from the *emmeans* package (Lenth, 2023) in R.

1.3.2. Two plates experiment

We conducted 200 trials over 13 days. On the first and last days, only five trials were conducted per cleaner wrasse, with day 1 serving as a training day. From day 2 to day 4, we carried out ten trials per day. On day 2, we experimented with an alternative method to facilitate manipulations, but subsequently reverted to the initial methodology. After day 4, we increased the trial count to 20 per day per cleaner wrasse.

The experimental paradigm was designed to minimize disturbance by dividing the aquarium into three compartments using transparent barriers (Fig. 2). The cleaner wrasse was placed in one of the side compartments, with the first inspection occurring in the central compartment and the second in the far compartment. During the first inspection, the cleaner wrasse inspected one of two plates (either grey with a red pattern or white with a green and yellow pattern, Fig. 2) and ate a food item (mashed prawn item) placed behind the plate, where the pattern was replicated to provide post-choice feedback.

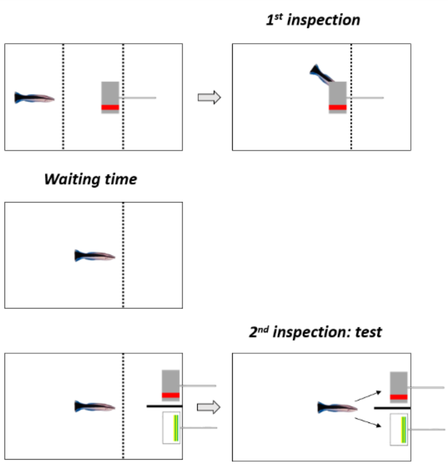


Figure 2: Experimental design
The figure illustrates the structure of a trial. Dotted lines represent see-through barriers. Each trial involved two inspections, with the cleaner wrasse expected to make a choice during the second inspection based on the plate presented during the first inspection and the assigned treatment. The waiting time between inspections ranged from 5 to 10 seconds.

Following the removal of the first plate, the cleaner wrasse was presented with two additional plates after a waiting period of 5 to 10 seconds: one was identical to the plate from the first inspection, and the other was novel. Half of the fish were assigned to the "same" treatment, where they would find a food item behind the identical plate, while the other half were assigned to the "other" treatment, where the food item was placed behind the different plate.

The plate presented during the first inspection was randomized, and no more than two consecutive trials featured the same plate, preventing the development of color preferences.

During the second inspection, cleaner wrasse were allowed to inspect both plates, regardless of the outcome of the first inspection, as this approach has proven effective in facilitating learning in associative learning tasks. The red or yellow-green plates were presented randomly in each trial, and we ensured that no plate was used for more than two consecutive trials to avoid the development of color preference.

Statistical analyses

Due to variations in training and methodology adjustments, data analysis for our main research questions commenced from the third day of testing, beginning with trial 16. Data from day 2 were analyzed separately.

We performed a binomial generalized linear mixed-effects model, using the trial outcome (0: failure, 1: success) as the response variable, with treatment ("same" or "other"), trial number (from 16 to 200), and their interaction as explanatory variables. The identity of the cleaner wrasse was included as a random factor. Model assumptions were checked graphically using the diagnostic tools available in the *DHARMa* package (Hartig, 2022) in R.

Post-hoc analyses were conducted using the least-square means method from the *emmeans* package (Lenth, 2023) and simple slope effect calculations from the *reghelper* package (Hughes & Beiner, 2023) in R. The same statistical methods were applied to analyze data from day 2 separately, allowing for comparison between days 2 and 3.

1.3.3. Working memory box: the movable arena

We provided extensive training over 15 days to habituate the cleaner wrasse to the complexity of the experimental setup. The training was conducted in intermediary steps: (a) picking a food item placed in the center of a dark circle drawn on a white plate; (b) replacing the dark circle with a piece of pipe (3 cm in diameter and depth); (c) inserting the pipe through a plate to create an entrance, leading the cleaner wrasse through it with food as a reward; (d) adding a piece of latex glove (the prevented re-entrance from the other side) to the prototype; (e) utilizing the final compartment plate with food items positioned at their final locations (preventing the cleaner wrasse from changing compartments from the rear). Several reward plates were used to guide the fish through the arena, particularly through the corridors. When the cleaner wrasse exhibited fear of the setup, the arena was left in the aquarium with reward plates scattered around to encourage inspection. This process was repeated until the cleaner wrasse swam comfortably and swiftly through the corridors and through the entire arena.

Trials started after subjects were habituated to other elements of the experimental setup, such as swimming through doors, adapting to the arena's movement (which caused significant water displacement), and familiarized with different swimming routes within the arena.

Procedure

Each trial began when the cleaner wrasse entered the arena (details provided in Fig. 3a–d) and concluded once all food items were eaten. Trials were terminated as a form of punishment if the cleaner wrasse attempted to bypass the task by moving around the arena to access food with minimal effort. Data were retained if the fish initiated cheating but ceased as soon as the experimenter began to intervene. During each trial, the cleaner wrasse made repeated choices, visiting compartments to obtain food rewards before returning to the original zone via the corresponding corridor to make subsequent choices.

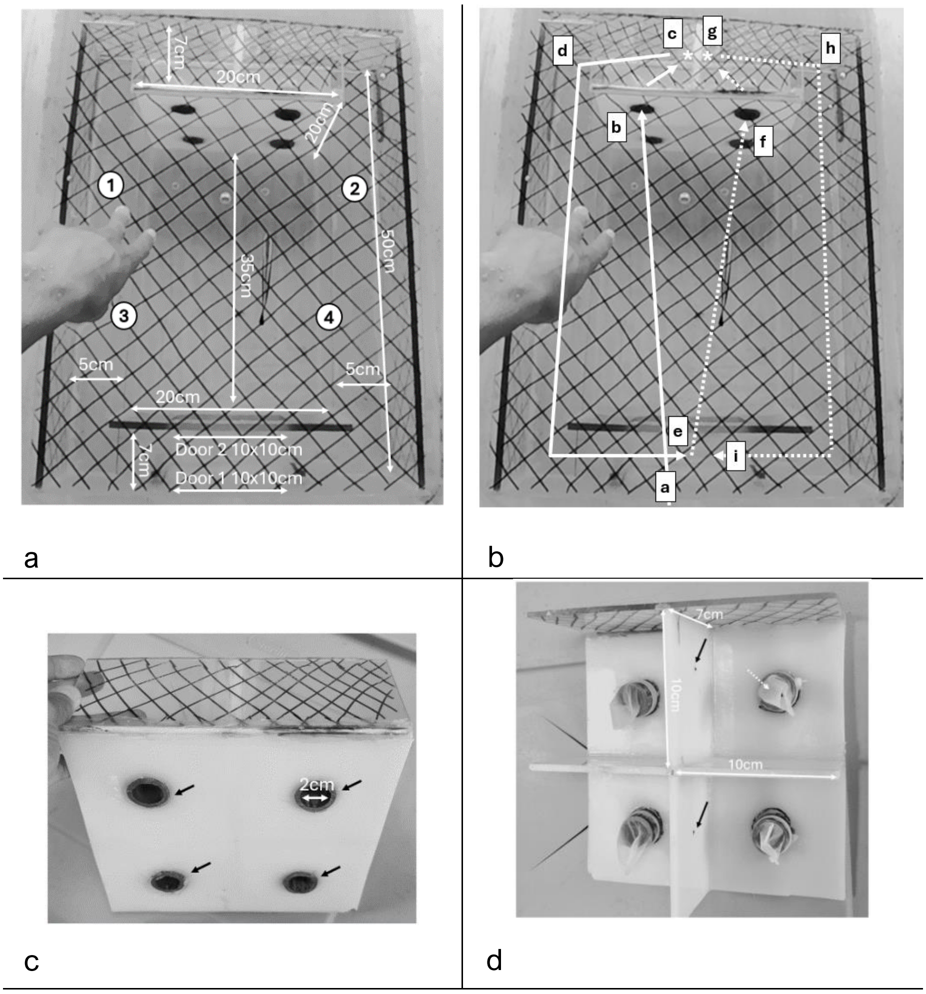


Figure 3: Details of the working memory testing arena

a. Arena design

The main door, "Door 1," was manually operated by the experimenter, and cleaner wrasse entered the arena through this door to begin the trial. "Door 2" remained open and helped confine the cleaner wrasse to a specific area before each choice. Corridors 1, 2, 3, and 4 are depicted, with corridors 3 and 4 positioned beneath 1 and 2, respectively. The top part of the arena was not fixed to facilitate placement in the aquarium. Transparent sections of the arena (e.g., the top part and external side walls) are dashed, and internal corridor walls are blurred to help the cleaner wrasse discern boundaries while maintaining sufficient lighting.

b. Movements of cleaner wrasse in the arena.

Solid white arrows indicate the sequence of the first choice, while dashed arrows represent a second choice. Letters correspond to key steps: (a) The cleaner enters the arena and passes through Doors 1 and 2; (b) The cleaner selects and enters the upper-left compartment; (c) The cleaner wrasse eats the food item (*) and exits the compartment; (d) The cleaner uses the corridor to return to the starting zone; (e) The cleaner re-enters through Door 2 to make another selection. Letters (f) to (i) mirror steps (b) to (e), representing additional choices.

c. Front view of the compartment plate

Black arrows indicate the entrances to the four compartments.

d. Rear view of the compartment plate

The cleaner wrasse could only access the four compartments by returning through Door 2 via the corridors, as the back of the compartment plate was blocked by an external wall. A food item was placed on the internal wall of each compartment (black arrows), and a latex glove piece was added to prevent the cleaner wrasse from exiting the compartment through its entrance (white-dashed arrow).

Data

Due to external events causing behavioral disturbances, 50 trials per fish were examined for analyses, and despite careful preparation and extensive training, not all of them were usable. We only included trials where intervention was not required or where the cleaner wrasse made an initial error (e.g., revisiting a previously visited compartment) prior to cheating. One cleaner wrasse was excluded from analysis for spending excessive time in the corridors, leading to delays between choices that exceeded acceptable limits (up to several minutes).

Since we wanted to see how many choices cleaner wrasse could make correctly before the first mistake, we still considered a trial where the cleaner cheated on the third choice to see whether they could make two correct choices in a row. Ultimately, data from five cleaner wrasse were analyzed, with a variable number of usable trials for each individual.

Statistical analyses

We compared the probabilities of correctly selecting 2, 3, or 4 compartments consecutively by chance with the actual performance of cleaner wrasse using a Wilcoxon signed-rank test.

1.3.4. Multiple plates

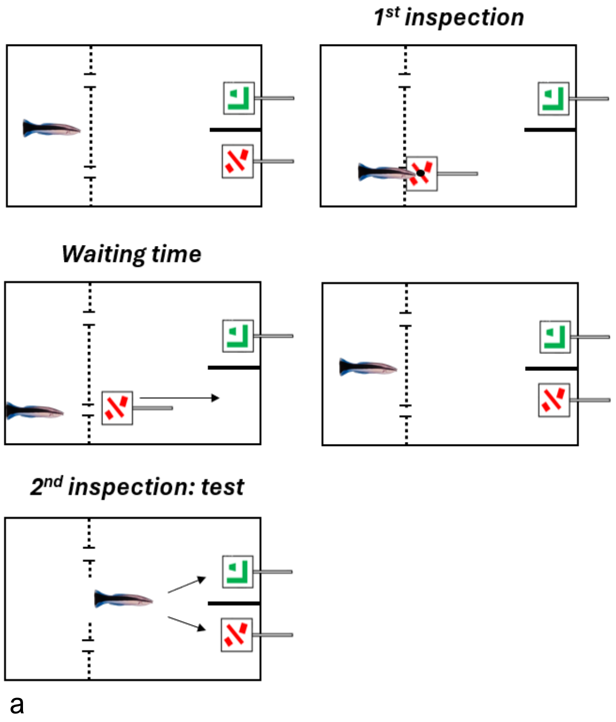
This experiment was initially designed to test different levels of complexity by sequentially adding more plates during inspections. The aim was to assess how much information cleaner wrasse could process in their WM and whether they prioritized visual (plate color) or spatial (plate position: right or left) information, in addition to modifying the methodology of plate presentation to cleaner wrasse from the "two plates" experiment. However, since subjects failed immediately, further testing was not conducted (i.e., we did not increase the number of plates per presentation). This explains why we utilized four distinct plates, organized into six pairs, each presented twice daily to the cleaner wrasse, resulting in 12 daily trials for four days. Since subjects did not fully acclimate to the setup, we transformed the first day of data collection into a training day, excluding these data from our analyses. On the fifth day, we altered the methodology to sequential manipulation (hence resembling "two plates" experiment again), whereby the first plate was presented to the cleaner wrasse in its compartment and then placed at the back of the tank alongside the second plate.

We counterbalanced which plate was shown first in each trial (2 trials per pair per day, one of the plates per trial). When positioning the plates, the plate to be retrieved (different from the first inspection plate) was alternated between the right and left sides of the tank.

Procedure

To begin a trial, the two plates were placed at opposite sides at the rear of the tank (Fig. 4a). We brought one of the plates (predefined for that trial) forward in a straight line towards the see-through barrier, rotated it to expose the hidden food item, and allowed the cleaner wrasse to eat it through a hole in the barrier. The plate was then turned back around (concealing the dark spot where the food had been placed) and moved to the back of the aquarium. The pair of plate was selected amond four possible plates (Fig. 4b).

We measured the time interval between the end of the first plate inspection and the moment the barrier door was opened for the second inspection. Using the central door of the barrier, we ensured that the cleaner wrasse started from a standardized location. The choice of the cleaner wrasse was noted. If it selected the same plate as in the first inspection, the trial was considered a failure, and both plates were removed before the cleaner wrasse could access the food item on the other plate.



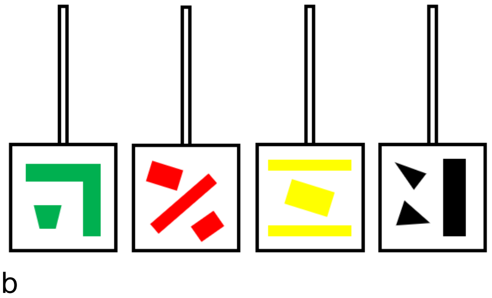


Figure 4: Experimental design
a. Representation of a trial.
Dotted lines represent the see-through barrier, and the interruptions indicate two holes. For each trial, we inserted the initial plate in the water, turned it around while moving toward the barrier to expose the food item hidden in the back (black spot). After allowing the cleaner wrasse to eat, we returned the plate to the back of the aquarium. Following a waiting period of under 10 seconds, the barrier door was opened, and we noted the choice of the cleaner wrasse.
b. The four different plates.
Four different plates, each 4 cm x 4 cm with colored patterns (red, yellow, green, or black), were organized into six pairs.

Statistical analyses

To account for variables that were unavoidable during data collection (e.g., the experimenter, the sex of the cleaner wrasse, the specific pair of plates, the trial number, and the time interval between inspections—below or above 10 seconds), we employed a binomial generalized linear mixed-effects model. The choice of the trial (0: failure, 1: success) was the response variable, and all variables mentioned were treated as covariates. The identity of the cleaner wrasse was included as a random factor. After none of the variables proved significant, we employed a null model to test the significance of the overall success rate. To specifically assess WM, we conducted a separate analysis for trials with inspection delays of less than 10 seconds.

For all models, we compared two versions using analysis of variance: one model with the cleaner wrasse identity as a random factor for both the intercept and slope with respect to time (trial number), and the other model with cleaner wrasse identity as a random factor for the intercept only. The simpler models were retained, as they showed no significant difference and had slightly lower Akaike Information Criterion (AIC) values.

We graphically checked the model assumptions using the diagnostic tools from the *DHARMa* package (Hartig, 2022) in R. The significance of the variables was tested using Type II Wald Chi-square tests, and post-hoc analyses were conducted using the least-square means method from the *emmeans* package (Lenth, 2023).

1.4. Results

1.4.1. Overtraining: Windows experiment

One trial was not analyzed for all cleaner wrasse as we had no recordings. One additional trial was removed for cleaner wrasse E due to a manipulation error. Among the 350 trials left for analyses, cleaner wrasse did not eat at all locations after 30 seconds in 19 of them. Among these, two were in the basic condition vs. 17 in window condition, and nine were with the fish plate vs. 10 with the square plate. Also, these trials were spread among eight of the nine cleaner wrasse; C and G had one trial, E and F had two trials, D, K, and M had three trials, and N had four trials.

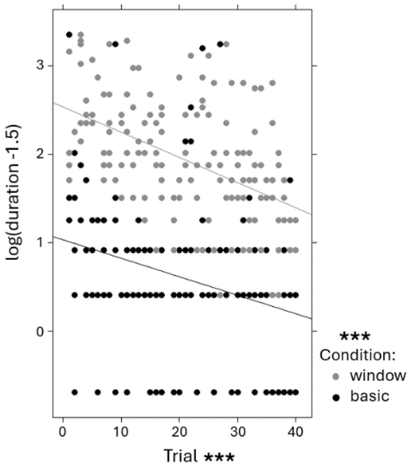


Figure 5: Effect of time and condition on the duration to eat the four items. The graph shows an interaction between "trial" and "condition" that was not included in the final model, but we used it for graphical representation purposes. The y-axis represents the log-transformed duration, and the x-axis shows the trial number. The colour code indicates the condition. The stars indicate the significance of the associated variable.
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Duration

With the full model (i.e., with all possible interactions), the condition:trial number was only marginally significant (Type II Wald Chisquare tests, condition:trial, Chisq = 2.92, df = 1, p-val = 0.09). We first simplified the model by keeping only that interaction along with the other covariates to see whether the tendency would be confirmed. As it came out non-significant (Type II Wald Chisquare tests, condition:trial: chisq = 2.22, df = 1, p-val = 0.14), we finally removed it as well and kept the simpler model with the three separated covariates.

Note that we decided to keep the two detected outliers as those were the two trials in which the maximum duration was 30 seconds. The immediate shorter durations were 28 seconds; we saw no reason to discard them.

Cleaner wrasse took 2.55 more time to eat the four items in the windows compared to the basic condition (least-square means, basic: predicted mean = 3.31sec, SE = 0.15 and windows: predicted mean = 8.45sec, SE = 0.54, significance from Type II Wald Chisquare tests, condition: chisq = 354.95, df = 1, p < 2e-16, Fig. 5). In both conditions, the minimum and maximum time needed to eat all items were almost equal (basic: minimum = 2 sec and maximum = 30 sec; window: minimum = 3 sec and maximum = 30 sec).

Also, cleaner wrasse were almost twice as fast to eat the four items on trial 40 compared to the first trial (least-square means, trial 1: predicted mean = 7.43sec, SE = 0.55 and trial 40: predicted mean = 3.69 sec, SE = 0.20, significance from Type II Wald Chisquare tests, trial: chisq = 73.12, df = 1, p-val < 2e-16, Fig. 5). Finally, the plate was only marginally significant (Type II Wald Chisquare tests, chisq = 2.77, df = 1, p-val = 0.096).

Mistakes

None of the interaction terms was significant in the full model (Type II Wald Chisquare tests, Plate:Condition: chisq = 2.65, df = 1, p-val = 0.10, Plate:Trial: chisq = 0.60, df = 1, p-val = 0.44, Condition:Trial: chisq = 0.05, df = 1, p-val = 0.81, Plate:Condition:Trial: chisq = 1.36, df = 1, p-val = 0.24).

After the model was simplified, we saw that cleaner wrasse made 1.5 times more mistakes (i.e., repeating a location) in the square compared to the fish plate (least-square means, square:

predicted mean = 0.36, SE = 0.05 and fish: predicted mean = 0.24, SE = 0.04, significance from Type II Wald Chisquare tests, plate: chisq = 4.22, df = 1, p-val = 0.04, Fig. 6). They also made around 1.7 times fewer mistakes in basic compared to windows condition (least-square means, basic: predicted mean = 0.23, SE = 0.04 and windows: predicted mean = 0.38, SE = 0.06, significance from Type II Wald Chisquare tests, condition: chisq = 6.84, df = 1, p-val = 0.009, Fig. 6).

Lastly, the number of mistakes was 3.4 times smaller at trial 40 compared to trial 1 (least-square means, trial 1: predicted mean = 0.55, SE = 0.1 and trial 40: predicted mean = 0.16, SE = 0.04, significance from Type II Wald Chisquare tests, trial: chisq = 13.69, df = 1, p-val < 0.001, Fig. 6).

During the two post-experimental verifying trials, we observed that they would pick again on the location if they had left some prawn there.

Looking

About the interaction terms in the full model, only the interaction between plate and condition came out significant (Type II Wald Chisquare tests, chisq = 5.15, df = 1, p-val = 0.023).

In this case, the interaction between plate and condition was significant (Type II Wald Chisquare tests, chisq = 5.10, df = 1, p-val = 0.02). When tested with the fish plate, cleaner wrasse looked 4.2 times less in basic than in window condition (least-square means, basic: predicted mean = 0.66, SE = 0.09 and windows: predicted mean = 2.73, SE = 0.20, z-ratio = -9.70, p-val < 0.0001, Fig. 7). When tested with the square plate, cleaner wrasse looked only 2.6 times less in basic than in windows condition (least-square means, basic: predicted mean = 0.91, SE = 0.10 and window: predicted mean = 2.41, SE = 0.18, z-ratio = -7.21, p-val < 0.0001, Fig. 7).

Overall, cleaner fish had to look 1.55 times more during the first trial compared to trial 40 (least-

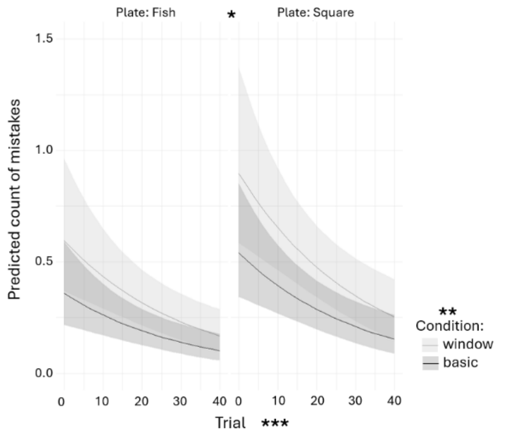


Figure 6: Effect of plate, condition, and time on the number of mistakes.
The graph shows an interaction between "plate, "condition", and "trial" that was not included in the final model, but we used it for graphical representation purposes. The y-axis is the predicted mistakes count, and the x-axis displays the trial number. Each panel represents a plate, and the colour indicates the condition. The stars indicate the significance of the associated variable.
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

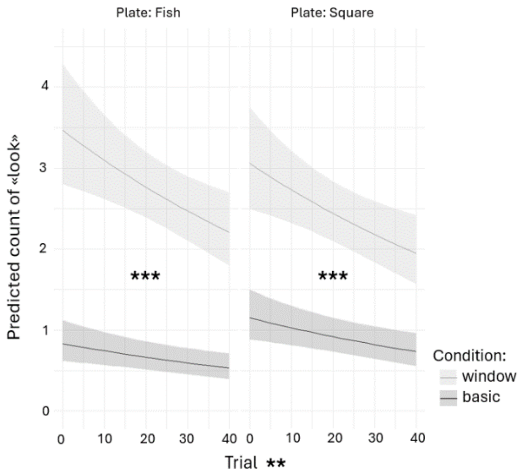


Figure 7: Effect of plate, condition, and time on the number of times cleaner looked around.
The graph shows an interaction between "plate, "condition", and "trial" that was not included in the final model; only plate:condition was, but we used it for graphical representation purposes.
The y-axis is the predicted count of times cleaners looked around (*i.e.*, opportunities to see the items again), and the x-axis is the trial number. Each panel represents a plate, and the colour code indicates the condition. The significance of the difference between both condition lines is given on each panel, and the significance of the covariate "Trial" is shown next to the x-axis label.
Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

square means, trial 1: predicted mean = 1.77, SE = 0.16 and trial 40: predicted mean = 1.77, SE = 0.11, significance from Type II Wald Chisquare tests, trial: chisq = 9.04, df = 1, p-val = 0.003, Fig. 7).

1.4.2. Two plates experiment

We found a significant interaction between the treatment and time (Fig. 8). In the "same" treatment, the success rate of cleaner fish increased by 6.3% over time, but it was insignificant (Simple slope effect, slope = 1.001, t-value = 0.132). In the "other" treatment, it decreased by 13.4% (Simple slope effect, slope = 0.997, t-value = 0.008).

Overall, both treatments brought success rates significantly different from the 50% chance (Least-square means, same: predicted p of success = 47%, z-ratio = 2.41, p-val = 0.03 and other: predicted p of success = 53.6%, z-ratio = -2.4, p-val = 0.03).

During day 2 only, the two treatments were significantly different from each other as well (least-square means, same: predicted p or success = 67.5%, SE = 0.052 and other: predicted p of success = 44.4%, SE = 0.052, z-ratio = 2.99, p-val = 0.003) but only "other" treatment differed from chance (least-square means, other: z-ratio = 3.06, p-val = 0.004 and same: z-ratio = -1.05, p-val = 0.293).

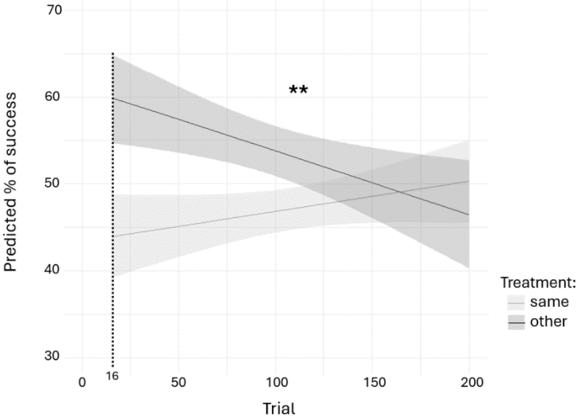


Figure 8: Different effects of time on the proportion of success depending on the treatment
The y-axis is the predicted proportion (in percentages) of success, the x-axis shows the trial number, and the colours represent the treatments. In the "same" treatment, cleaner wrasse had to choose the plate that was the same as the one from the first inspection during the second inspection, while it was the other in the "other" treatment. The stars represent the significance of the interaction in the model.
Significance codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The comparison between day 2 and day 3 did not indicate any difference (Type II Wald chisquare tests, day: chisq = 0.304, df = 1, p-val = 0.581 and treatment:day: chisq = 0.690, df = 1, p-val = 0.406).

1.4.3. Working memory box: the movable arena

Cleaner wrasse failed to perform 2, 3 or 4 correct choices in a row above chance (Wilcoxon signed rank exact tests, 2 choices: V = 6 and p-val = 0.81, 3 choices: V = 5 and p-val = 0.63, 4 choices: V = 3 and p-val = 0.31, Fig.9).

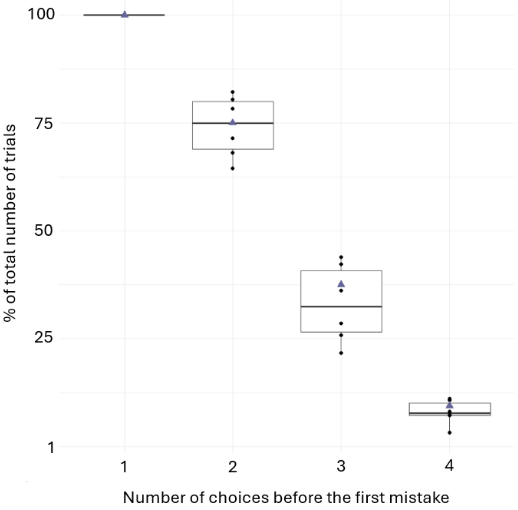


Figure 9: Percentages of the number of trials in which cleaner wrasse succeeded in making 1, 2, 3, or 4 correct choices in a row.
The y-axis is the percentage of the number of trials, calculated on the usable trials only. The x-axis is the number of correct choices cleaner wrasse made before their first mistake. Because the first choice can never be a mistake, they succeeded in 100% of trials. Triangles indicate the random mean of success.

1.4.4. Multiple Plates

None of the covariates we studied had a significant effect on the choice of cleaner wrasse (Type II Wald chisquare tests on binomial glmer, experimenter: chisq = 0.064, df = 1, p-val = 0.800; sex of fish: chisq = 0.098, df = 1, p-val = 0.755; pair: chisq = 2.723, df = 5, p-val = 0.743; trial: chisq = 0.838, df =1, p-val =0.360 and time between inspection: chisq = 0.138, df = 1, p-val = 0.710).

The null model indicated that cleaner wrasse did not visit the correct plate above chance (least-square means, predicted mean = 52.4%, SE = 0.015, z-ratio = 1.566, p-val = 0.117, Fig. 10).

When we separated the trials for which the time between inspections was below (1038 / 1102 = 94.2% of all trials) or above (64 / 1102 = 5.8% of trials) 10 seconds, the result was the same for both (least-square means, below 10 seconds: predicted mean = 52.3%, SE = 0.016, z-ratio, p-val = 0.136 and above 10 seconds: predicted mean = 50.6%, SE = 0.080, z-ratio = 0.073, p-val = 0.942).

On the last day, we changed to a sequential methodology. On this day, cleaner wrasse succeeded above chance (least-square means, predicted mean = 56.9%, SE = 0.0298, z-ratio = 2.280, p-val = 0.023). With the trials with a time between inspections below 10 seconds only (241 / 276 = 87.3% of trials), the success remained only marginally significants (least-square means, predicted mean = 56.0%, SE = 0.032, z-ratio = 1.864, p-val = 0.062).

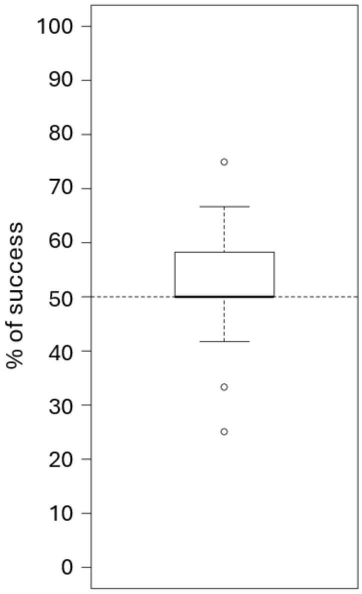


Figure 10: Percentage of correct choices
Boxplot showing the overall percentage of success (y-axis) of cleaner wrasse during the task. The dashed horizontal line indicates the random threshold of 50%.

1.5. Discussion & Conclusion

This study comprises four distinct experiments that pioneer the investigation of WM in cleaner wrasse. Overall, the performance of cleaners was modest, yielding weak positive results in two visual tasks and negative results in the other tasks. The absence of stronger positive evidence prevented further testing aimed at ruling out alternative problem-solving mechanisms based on familiarity (i.e. the ability to react to a stimulus without explicit recollection of the information associated to it; see the difference between familiarity and recollection in Yonelinas, 2002). Similarly, the lack of positive evidence also made it unnecessary to test for sensitivity to disturbance or storage capacity of WM (e.g., number of items that can be held and manipulated simultaneously), two key aspects often associated with WM capacity and largely overlooked in the litterature. We recognize that absence of evidence does not imply evidence of absence, but consistent results across the four different experiments make us skeptical about the existence of WM in cleaner wrasse. As previous laboratory experiments with cleaner wrasse highlighted the use of strategies that capitalized on previous knowledge and experiences (Binning et al., 2017; Bshary & Grutter, 2006; Grutter, 2004; Pinto et al., 2011a). Because the deployment of such strategies seems to require agile behavioral adjustments, relying on executive control and adaptability, we expected WM to be recruited (Baddeley & Hitch, 2019; Cowan, 1995; Diamond, 2013).

We provide detailed discussions of each experiment in the supplementary material, including details of earlier setups that were confusing for cleaner wrasse.

1.5.1. Link to ecology

We strived to design experiments that could reflect the ecology of cleaner fish, with the goal of engaging their WM abilities, if such abilities exist. For example, when cleaner wrasse clean their clients, they inspect different surface locations (*e.g.,* mouth, gills, tail). For the sake of efficiency, it would be expected that cleaner wrasse remember and avoid revisiting previously inspected locations (i.e., they optimize their foraging strategy). Avoiding previous locations should reflect the ability to store and update information, a key indicator of WM. Our “windows” experiment and the “movable arena experiment” aimed to elicit this ability. We found that cleaner wrasse were more efficient when windows were overlaid on an image of a client fish (here, a butterflyfish) *versus* overlaid on a plain white square plate. The picture of a client fish was intended to give additional ecological relevance to the experimental setup, and our results seem to validate this explanation. That said, we cannot rule out that simple differences in coloring may affect the behavior of cleaner wrasse, acting as a cue. Future experiments should address this possibility.

The other two experiments where we used pairs of plates were also intended to simulate cleaner wrasse's ecology. Because cleaners have multiple repeated interactions per day (Grutter, 1996a), we hypothesized that they would likely remember recent clients to avoid inspecting the same individual again, thereby optimizing their foraging behavior. Indeed, a previous study found that cleaner wrasse could adapt their foraging choices based on information they acquired up to several minutes before, avoiding “client plates” that they had inspected within the time windows (2.5 to 12.5 min) set by the experimenter (Salwiczek & Bshary, 2011). In other words, cleaner wrasse could remember "*when* they interacted with *what* after a single event" (Salwiczek & Bshary, 2011). Although the possession of something akin to episodic-like memory would account for these results (Pause et al., 2013; Tulving, 2002), it would be reasonable to expect that cleaner wrasse would succeed when the time intervals separating acquisition and retention is shorter, provided they possess something akin to WM. Lack of success in this new condition reinforces the idea that time is a critical factor for eliciting WM and that the separation between LTM and WM is a valid one, even if one is to accept that WM could operate with information reactivated from LTM (Cowan, 2008, 2019; Norris, 2017).

1.5.2. The lack of WM in cleaner wrasse, and other ectotherms?

We did not anticipate that cleaners would perform so poorly in all four experimental tasks, but the lack of positive evidence for WM fits the apparent absence of a general intelligence factor *g* reported for this species (Aellen et al., 2022). In mammals, individual performance in WM tasks is a reliable predictor of performance in multiple other cognitive tasks, indicating that the presence of WM is a precondition for the presence of the *g* factor (Ackerman et al., 2002; discussed and reviewed in Conway et al., 2003; Engle et al., 1999; Shipstead et al., 2014). Consequently, we did not expect cleaner wrasse performance would be comparable to that of other species famous for their ingenuity (e.g., primates or corvids). We also expected that cleaners would be more easily distracted due to low attentional control (discussed in Carruthers, 2013 and Manrique & Walker, 2017). This is not surprising given the apparent gap in executive capacity between humans and other species (Carruthers, 2013; Manrique et al., 2024; Manrique & Walker, 2017). However, failing to find any significant evidence for WM in cleaner wrasse came as a surprise.

A question arising from this study is whether the absence of WM is specific to cleaner wrasse, or is a key difference that separates the cognition of ectotherms and endotherms (Jerison, 1969). To answer this question, we need to conduct additional studies with other ectotherm species. This should include new experimental paradigms that allow for cross-species comparisons and test for central cognitive abilities linked to WM, such as resistance to interference and the storage and manipulation of information in relation to problem-solving (Carruthers, 2013; Cowan et al., 2024).

The consensus in the literature is that WM is a system that keeps recently acquired information in an active state, allowing its combination with knowledge previously acquired (e.g., LTM) to assist prospective adaptive action (e.g., Baddeley & Hitch, 1974; Cowan, 1995, 2019). With this definition in mind, we consider the extant literature regarding WM in fish to be very incipient. Studies to-date have investigated the natural alternation rate by zebrafish in a Y-maze during a free exploration task (Fontana et al., 2021) or showed zebrafish succeeding in a delayed match-to-sample experiment (Bloch et al., 2019). Furthermore, both cichlids and guppies selected for a bigger telencephalon size showed significant, though modest, success in the Piagetian stage 4 of object permanence (i.e., recovering a hidden object, Piaget & Cook, 1952), which the authors directly associated with WM (Guadagno & Triki, 2024; Triki et al., 2023). However, previous studies that used mazes to measure WM have been criticized for not considering whether the time interval separation between acquisition and retrieval is within a range compatible with WM (Carruthers, 2013). In addition, the manipulative and problem-solving aspect of WM is barely detectable in the task, especially when it is about purely explorative behavior, as in Fontana et al. (2021). Second, match-to-sample tasks are visual tasks where processes like familiarity could satisfactorily account for the observed results, without invoking higher-level executive control (e.g. Aggleton et al., 1986; Cowan, 2019; Yonelinas, 2002). That is why they are often considered "recognition tasks" (Bachevalier &

Mishkin, 1986; Zola et al., 2000). Lastly, we question the direct association between positive evidence for object permanence and WM, arguing that only later stages of object permanence (i.e., Piagetian stages 5 and 6) can reliably be linked to WM, as the combination and manipulation of information are absent in previous stages. In Piagetian stage 4 (single visible displacement task, Piaget & Cook, 1952), the subject must only retrieve a hidden object. Therefore, we would argue that there is no necessity for the combination or manipulation of previously acquired information. In contrast, later stages (5 and 6, with invisible displacement, Piaget & Cook, 1952) almost certainly engage WM, as they require varying degrees of mental computation to track the trajectory of an unseen moving object.

In conclusion, our comprehensive series of experiments thoroughly demonstrate a lack of WM ability in cleaner wrasse, which contrasts with evidence for WM in endotherm vertebrates, including corvids and psittacids (Pepperberg & Funk, 1990; Zucca et al., 2007), mice (Kolata et al., 2005), rats (Bratch et al., 2016) and chimpanzees (e.g. Völter et al., 2019, debates in Manrique et al., 2024; Read et al., 2022). However, interestingly, the other two executive functions (EFs)—cognitive flexibility (Aellen et al., 2022; Parker et al., 2012) and inhibitory control (Aellen et al., 2021; Lucon-Xiccato & Bisazza, 2017; Sovrano et al., 2018)—have been found in fish. This begs the question: Are WM and the associated *g* factor, but not other EFs, the basis for differences in relative brain size between endotherms and ectotherms (Jerison, 1969)?

1.6. Acknowledgements

The staff of the Richard B. Gump research station in Mo’orea, French Polynesia. S. Duc, E. de Sousa, and R. Bonvini for helping during data collection. Prof. K. Zuberbüler for his help in conceptualizing the movable arena. C. S. Clements for his help in improving the writing.

1.7. Supplementary material

1.7.1. Methods

This supplementary document provides a detailed account of the various experimental designs we initially developed to assess WM abilities in cleaner wrasse. Throughout the testing phase, we faced several challenges, which prompted us to adapt our methods. Here, we describe both the intellectual approach and the methodological adjustments that were made, ensuring that readers gain a comprehensive understanding of how these factors shaped the experiments presented in the main text. All the designs discussed here were implemented during the fieldwork conducted in 2021 in Moorea, French Polynesia.

The compartmentalized plate

The compartmentalized plate was the first design we conceived. We conducted the training in two phases to acclimate the cleaner wrasse to this setup (Fig. S1). First, we randomly placed the compartmentalized plate into the tank, with food items in all four compartments, allowing the wrasse to explore and habituate to it without exhibiting stress. Second, we closed the back of the compartmentalized plate, requiring cleaner wrasse to enter from the sides, which was intended as the final experimental condition.

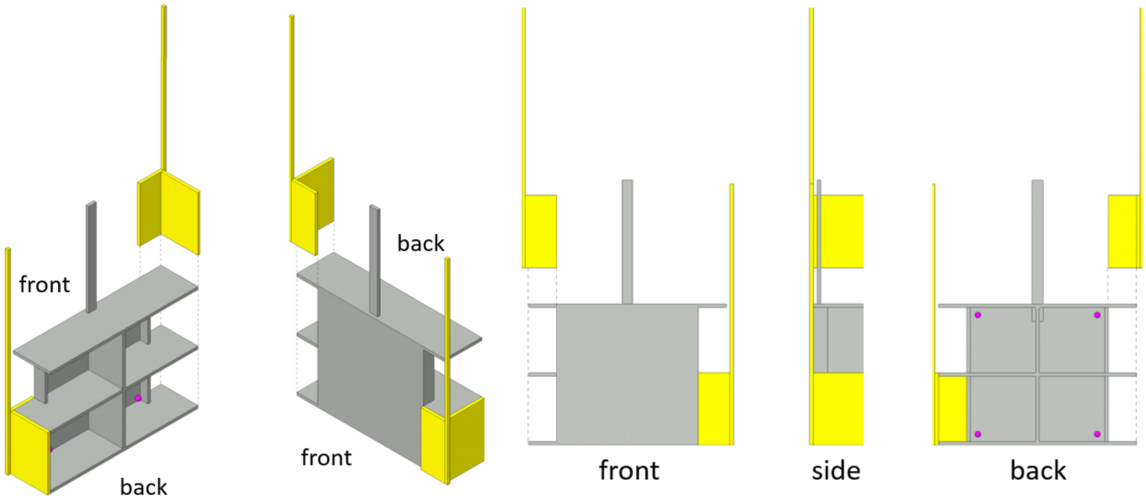


Figure S1: The compartmentalized plate
The compartmentalized plate (20 cm x 20 cm) consisted of four equal-sized compartments. Constructed entirely from white Plexiglas. A food item (represented by a pink ball) was placed in each compartment. We placed the plate with its back part against it to force cleaner wrasse to swim around to the front to change compartments. The doors (yellow structures) allowed us to selectively open or close specific compartments. We planned to use several variations of the plate design: one without colors (as shown here), another with each compartment marked by a different color, and a fish-shaped plate.

We encountered several issues, some of which were material-based and fixable, such as needing to adjust the plate multiple times to ensure the wrasse could only access the compartments from the sides. However, other problems were more critical, as they involved cleaner wrasse behavior. For example, we observed that cleaner wrasse subjects tended to remain inside the first visited compartment, which we initially suspected could be related to stress. However, their overall behavior did not support this idea (normal swimming when outside the compartment).

While subjects occasionally moved to the adjacent compartment directly above or below, they rarely switched to the opposite side. As a result, we concluded that this setup was unsuitable for addressing our research question and decided to transition to a smaller plate design to reduce the stress and confusion observed among subjects.

“First attempt”

To address the limitations of the compartmentalized plate, we developed a new setup (Fig. S2) designed to test the cleaner wrasse under various conditions (i.e., complexities, Fig. S2a) and contexts (i.e., plate types, Fig. S2b). The idea was to present the cleaner wrasse with the different plate-condition combinations and observe their inspection behavior. We counted repeated visits to the same location as errors and aimed to assess whether condition and/or context influenced their success in eating all items.

In the basic condition, all items were visible, and we expected no errors from the cleaner wrasse. In the "window" condition, items were hidden when the cleaner wrasse approached from the side, meaning that any mistakes could suggest a lack of WM, as they would need to recall previous locations. Lastly, the "hidden" condition increased the complexity further, as the food item was entirely hidden. If the cleaner wrasse made no mistake here, we would check for routine-based behavior (e.g., inspecting locations in the same sequence every time). If no such routines were found, it would suggest that the cleaner wrasse indeed possess WM.

We conducted five trials per cleaner wrasse but encountered an issue with the hidden condition: the cleaner wrasse either could not access the reward or could see it, thus violating the goal of the setup.

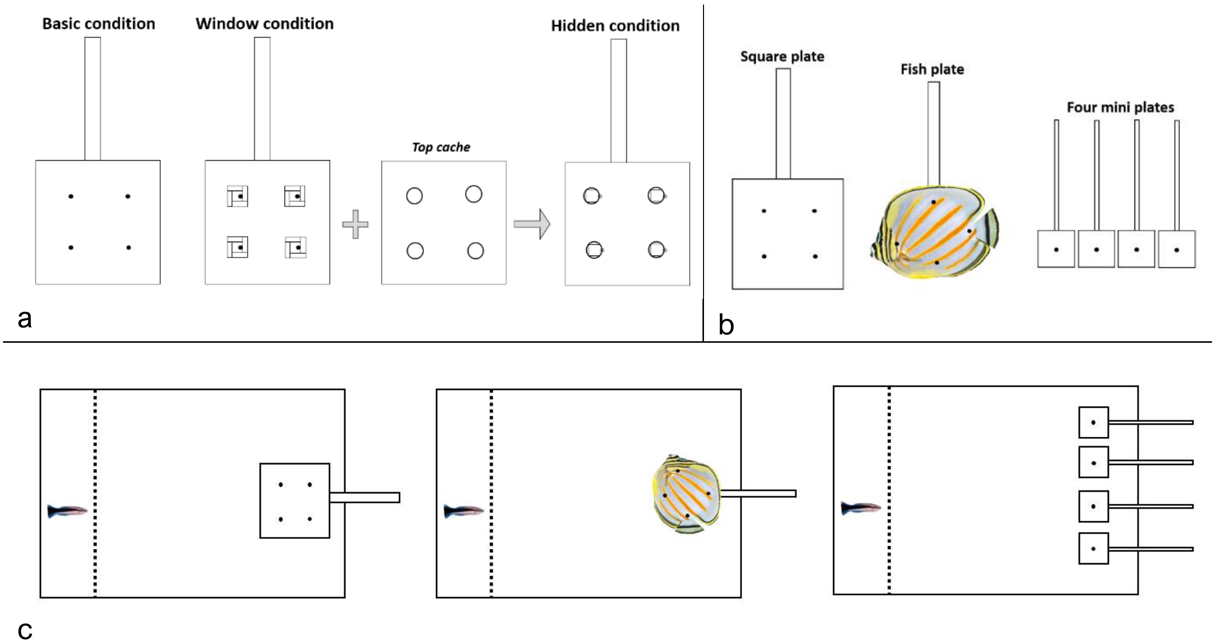


Figure S2: “First attempt” design.
We used different experimental conditions (a) and plates (b), all with a similar setup (c). Food items were placed on designated black spots. There were three different conditions, each representing a different level of complexity (a): basic (items placed directly on black spots, the simplest), window (items placed within a shallow 0.5 cm Plexiglas window, moderate complexity), and hidden (a Plexiglas cover with holes added over the window plate, concealing the items, the most complex). We planned to test these using three different plate designs (b): a square plate, a fish-shaped plate, and four smaller plates. The experimental procedure involved placing the plate in the tank while the cleaner wrasse observed from behind a transparent barrier (c). Once released, the fish was allowed to explore the plate, with repeated visits to the same location being counted as mistakes. The trial ended when all locations had been visited.

“Holes plates”
To overcome the issues encountered with the hidden condition in the previous design, we modified the setup by enlarging the holes on the cache to make it easier for the cleaner wrasse to access the food (Fig. S3a-b). To avoid the food being visible, we used rubber bands around each hole and placed the food reward behind them, requiring the cleaner wrasse to actively enter the holes. We also re-arranged the food items to mimic the layout of the fish plate.

The concept of this design was to allow the cleaner wrasse to inspect two locations and eat the food items during a first round of inspection (Fig S3b-d). Afterward, we separated the cleaner wrasse from the plate, opened the remaining locations, and observed which ones it inspected in a second round (all locations were then available). Any repeated visit to the same location counted as a mistake. By randomizing the accessible locations during the first inspection, we minimized the possibility of the cleaner wrasse to rely solely on routine behaviors.

Two methods were tested during the inspections (Fig. S3c-d). The first method (Fig. S3c) was problematic due to delays between the last food item eaten and the second inspection. We aimed for a 5-10 second interval but struggled with separating the cleaner wrasse quickly enough without causing stress. Without this flaw, any negative results would have been reliable since cleaner wrasse always had visual access to the plate. The second method (Fig. S3d) streamlined the process, but we still couldn't consistently meet the desired time interval without stressing the fish, and some cleaner wrasse managed to see the food items, biasing the setup.

An additional simplified attempt involved closing three locations during the first inspection and opening just one during the second. However, manipulation challenges persisted, and we abandoned the setup.

Ultimately, we opted to return to a simpler design, comparing a “basic” and “window” condition from the first attempt setup (Fig. S2), as we thought it would already be a starting point. This led us to the "windows experiment," as described in the main study, followed by the "two plates" experiment, a design better suited for cleaner wrasse.

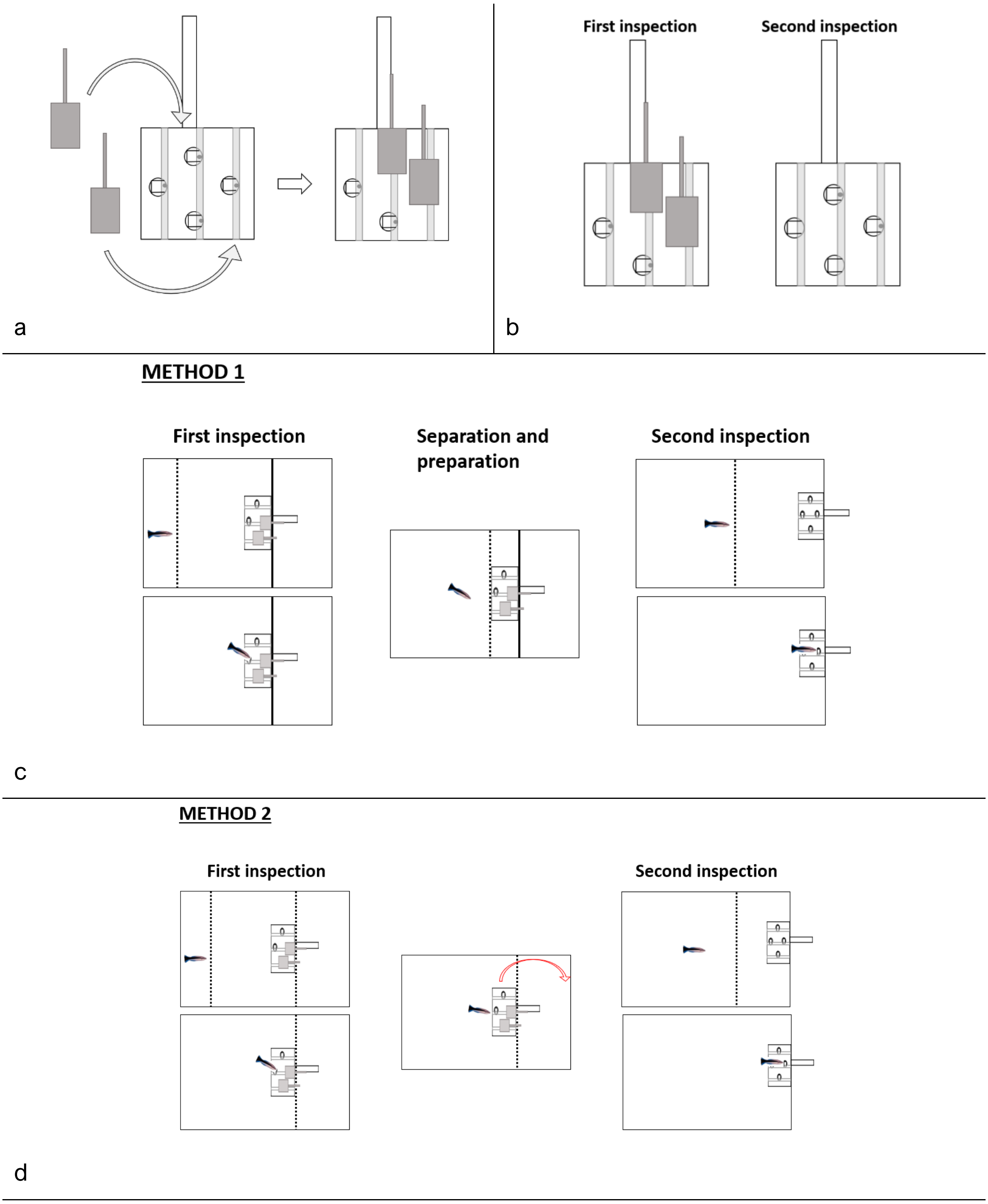


Figure S3: The "holes plate" design.
The holes plate (a) and how we intended to use the different inspections (b). The plate was built using grey caches to block selected locations (a-b), and food items were placed on black spots behind rubber bands (clear grey vertical bands, a-b).
Two methods were tested to organize the inspections (c-d). For the first method (c), cleaner wrasse were allowed to feed at two locations during the first inspection. Afterward, they were gently separated from the plate using a transparent barrier (dotted line) and the plate was moved to the back of the aquarium before the cleaner wrasse was released for the second inspection. In the second method (d), aimed at reducing manipulation time, instead of removing the cleaner wrasse from the plate between inspections, the plate itself was moved to another section of the aquarium to facilitate the process.

1.7.2. Discussion

In this section, we present detailed discussion for each experiment that we fully carried out and analyzed (presented in the main core of this article).

Windows experiment

Initially designed as a training step for another experiment (see the previous section for details), the cleaner wrasse were "overtrained" to locate four food items on a plate. Over time, it became evident that this could serve as a useful first step in examining whether cleaner wrasse could efficiently find food items when they were not always directly visible. If the presence of windows impacted performance, it would suggest a potential limitation in their WM. The central hypothesis was that both the amount of visual information (conditions: basic or window) and ecological factors (plates: square or fish) could influence the foraging efficiency of cleaner wrasse. Additionally, a learning component was expected.

A first key observation was that cleaner wrasse failed to eat within 30 seconds almost exclusively on the window condition. This already demonstrated that reducing visual information significantly lowered their foraging efficiency. However, there was no discernible difference between the plates, which undermined the hypothesis that ecological factors, such as the fish-shaped plate, would enhance performance.

Originally, foraging efficiency was measured by counting mistakes (i.e., revisiting locations where the food item had already been eaten). However, it became apparent that cleaner wrasse exhibited behaviors such as visually inspecting (either directly or by circling the plate) before picking. As a result, the mistake count alone did not accurately capture their behavior, prompting us to add additional metrics to estimate foraging efficiency more comprehensively. While one could argue that the duration and the number of visual inspections (referred to as "looks") are related (i.e., the more they inspect, the longer it takes) but some inspections were brief, while others lasted several seconds. Thus, both variables complemented each other without being redundant.

The three response variables—duration, mistake count, and the number of looks—yielded complementary findings. The amount of visual information (basic or window condition) influenced the duration of foraging, but there was no effect of the ecological relevance (plate: fish or square). Both visual information and ecological relevance independently affected the number of mistakes. Finally, for the number of looks, an interaction between visual information and ecological relevance was observed. A learning component was also detected across all response variables.

Post-experimental verification trials suggested that what was initially classified as a "mistake" might have been cleaner wrasse simply finishing an item they had only partially eaten, even if we were careful in the size of items, trying to make it impossible to leave a piece off it. Nevertheless, the overall findings remained consistent—the error rate was higher in the window condition.

In conclusion, both the availability of visual information and the ecological relevance of the plate significantly influenced cleaner wrasse's foraging efficiency. The importance of visual cues suggests that cleaner wrasse may lack WM, as they required full visual access to effectively locate all four food items. This limitation suggests that they may not integrate past and present information to guide goal-directed behavior.

As for the difference between the plates, caution is needed when interpreting these results. The patterns and colors on the fish plate could have acted as cues without necessarily being linked to ecological relevance. While we did not explicitly test for this explanation, cleaner wrasse could also rely on color and patterns in nature, making it reasonable to infer some ecological relevance. However, this interpretation should be approached with care.

Finally, learning may play a role in compensating for cleaner wrasse's potential lack of WM. This possibility is discussed in detail in the main text of this article.

Two plates experiment

We observed that cleaner wrasse exhibited a natural inclination to inspect a plate they had not previously examined when given the choice, rather than returning to the same one. Although the overall mean success rate for this behavior (“other” treatment) was only 53.6%, limiting the strength of our conclusions, this tendency aligns with what we would expect based on their efficient foraging strategies in the wild, where cleaner wrasse likely avoiding cleaning successively the same individual. These findings are consistent with previous research suggesting that cleaner wrasse adjust their choices based on previous interactions (Bshary & Triki, 2022; Salwiczek & Bshary, 2011). Finally, comparing the results from day 2 with those from day 3 yielded no significant differences, and variation in our methodology did not appear to affect cleaner wrasse performance in a meaningful way. Despite the more pronounced differences between treatments on day 2, this did not translate into measurable improvements in subsequent trials (Salwiczek & Bshary, 2011). Hence, there is no evidence that the methodology variation affected cleaner wrasse performance.

Interestingly, during the first days of the experiment, the success rate for the “other” treatment was around 60%, but we noticed a surprising "unlearning" effect over time, where performance declined instead of improving. This was unexpected, as we had anticipated that cleaner wrasse would become more proficient at the task with repeated trials. The training method we employed, which involved allowing cleaner wrasse to inspect both plates regardless of their initial choice, may have inadvertently caused this decrease in performance. It is possible that this approach reduced the motivation of the cleaner wrasse by making the cost of inspecting both plates lower than that of solving the task, leading them to gradually adopt a more random inspection strategy. Interestingly, this method has been successful in previous studies involving simple associative learning, where it accelerated the learning process for cleaner wrasse (Redouan Bshary pers. obs.). However, in this more complex task—where they needed to update information about the plates after each trial—the same method might have led to confusion, causing cleaner wrasse to abandon efforts to understand the task and simply go around the barrier separating the plates if they initially picked the wrong one.

Multiple plates

This experiment was a critical step in our investigation of WM abilities in cleaner wrasse. This method differed from our previous “two plates” experiment in that it did not involve temporarily removing the plate, which could unintentionally test for some aspect of object permanence (i.e., the ability to remember that an object still exists when put out of sight, Piaget & Cook, 1952, 1954). We were expecting positive outcomes since the cleaner wrasse had continuous visual access to the plate, and the time intervals between inspections were kept short. However, their consistent failure in this task further supports the hypothesis that cleaner wrasse may not possess WM.

Interestingly, our sequential methodology yielded some notable results, or at least strong tendencies, which were somewhat unexpected given our earlier findings. Nevertheless, the overall success rate of subjects remained low and further decreased in trials with a true WM delay (i.e., less than 10 seconds between inspections). Furthermore, a sequence effect must be considered; on day five, we altered the methodology for all the cleaner wrasse, meaning that by then, cleaner wrasse had already completed 60 trials (12 of which were conducted during the training day) that followed a similar logic. This prior experience could have influenced their behavior in the later trials, making it difficult to draw definitive conclusions about their WM abilities based solely on this experiment.

Working memory box: the movable arena

The failure of the cleaner wrasse in our paradigm could suggest two possibilities: (a) cleaner wrasse lack WM (as in previous setups) or (b) WM might not be detectable using this specific novel setup due to the complexity of it which may have prevented an accurate assessment of WM capacity. Given these uncertainties, we wanted to explore whether introducing color cues to the compartments could aid cleaner wrasse in solving the task. Unfortunately, due to time constraints during the field season, we were unable to collect data on this variation.

Despite these methodological challenges—partially stemming from the limited time available for data collection—the movable arena remains a promising paradigm for investigating WM. One of its key advantages is that, when cleaner wrasse are adequately trained, the time interval between choices can be easily manipulated by adjusting the length of the testing area. Additionally, the setup requires minimal intervention from the experimenter, in contrast to paradigms that involve the physical movement of plates or objects within the water, which is critical for minimizing disturbance during WM tests.

The flexibility of this paradigm also offers other advantages, as it allows for modifications such as adding colors or other types of information to the compartments, thus enabling the modification of various test conditions. This methodological flexibility offers intriguing possibilities for investigating different cognitive dimensions of WM in cleaner wrasse.

2. Chapter 2: In the absence of extensive initial training, cleaner wrasse *Labroides dimidiatus* fail a transitive inference task.

Leonore Bonin[1]*, Redouan Bshary[1]

[1] Behavioural Ecology Laboratory, Biology Institute, Faculty of Science, University of Neuchâtel, Neuchâtel, Switzerland

*Corresponding author
E-mail: leonore.bonin@unine.ch

Published in 2023:
Bonin, L., & Bshary, R. (2023). In the absence of extensive initial training, cleaner wrasse Labroides dimidiatus fail a transitive inference task. *PloS One*, *18*(6), e0287402.

2.1. Abstract

Transitive inference (TI) is a reasoning capacity that allows individuals to deduce unknown pair relationships from previous knowledge of other pair relationships. Its occurrence in a wide range of animals, including insects, has been linked to their ecological needs. Thus, TI should be absent in species that do not rely on such inferences in their natural lives. We hypothesized that the latter applies to the cleaner wrasse *Labroides dimidiatus* and tested this with 19 individuals using a five-term series (A > B > C > D > E) experiment. Cleaners first learned to prefer a food-rewarding plate (+) over a non-rewarding plate (-) in four plate pairs that imply a hierarchy from plate A to plate E (A+B-, B+C-, C+D-, D+E-), with the learning order counterbalanced between subjects. We then tested for spontaneous preferences in the unknown pairs BD (transitive inference task) and AE (as a control for anchors), interspersed between trials involving a mix of all known adjacent pairs. The cleaners systematically preferred A over E and showed good performance for A+B- and D+E- trials. Conversely, cleaners did not prefer B over D. These results were unaffected by the reinforcement history, but the order of learning of the different pairs of plates had a main impact on the remembrance of the initial training pairs. Overall, cleaners performed randomly in B+C- and C+D- trials. Thus, a memory constraint may have prevented subjects from applying TI. Indeed, a parallel study on cleaner wrasse provided positive evidence for TI but was achieved following extensive training on the non-adjacent pairs which may have over-ridden the ecological relevance of the task.

2.2. Introduction

Animal cognition involves a suite of mechanisms through which an individual acquires, stores, and uses information from the environment (Shettleworth, 2009). Many learned abilities among animals may be explained with Pavlovian or operant conditioning, where individuals either learn to associate a primarily involuntary action with an external stimulus (Pavlov, 2010) or associate a voluntary action with a consequence (Thorndike, 1927). However, it remains a major challenge to identify more complex cognitive processes, such as transitive inference (TI), in non-human animals. TI is defined as the ability to deduce the relative value of two objects that were never observed together—and thus not subject to potential reinforcement learning—from other previously learned pair relationships. For instance, in a list of three objects: A, B, and C; if A is greater than B, and B is greater than C, one can deduce that A is greater than C.

Despite its supposed cognitive complexity, evidence of TI has been suggested in a wide variety of non-human vertebrates, including birds (*e.g.*, jays Bond et al., 2003; Paz-y-Miño C et al., 2004); crows (Lazareva et al., 2004; Lazareva & Wasserman, 2006); pigeons (Couvillon & Bitterman, 1992; T. R. Zentall et al., 2019); geese (Weiß et al., 2010)), mammals (e.g., chimpanzees (Gillan, 1981); macaques (Jensen et al., 2016); lemurs (Tromp et al., 2015); mice (Silverman et al., 2015), and fish (*e.g.*, brook trout (White & Gowan, 2013); cichlids (Grosenick et al., 2007); cleaner wrasse, Hotta et al. 2020), as well as some invertebrates (*e.g.,* paper wasps, Tibbetts et al. 2019). The occurrence of this capacity has been largely linked to specific aspects of social structure (Bond et al., 2003; MacLean et al., 2008; Paz-y-Miño C et al., 2004). For example, in large and complex social groups, TI may be used to minimize competitive interactions and thus the cost of fighting and injuries (Bond et al., 2010; Grosenick et al., 2007). This is achieved because TI allows an individual to deduce dominance relationships from observations and avoid fighting stronger individuals.

The ecological approach to cognition (Hutchins, 2010; Kamil, 1998; Shettleworth, 2009) predicts a tight link between ecological needs and performance in cognitive tasks. The positive results from studies involving social dominances in fishes (Grosenick et al., 2007; White & Gowan, 2013) and paper wasps (Tibbetts et al., 2019) support the ecological approach to cognition by showing a link between the ecological need and the presence of TI in species that lack highly derived brain structures like a mammalian neocortex or an avian neostriatum. More recently, intraspecific differences observed in TI capacity that are directly linked to sex and rank further demonstrate the important modulation of TI by social need (Daisley et al., 2021). The ecological approach to cognition also predicts that in the absence of a need to represent hierarchical orders in the brain, TI should be absent, again irrespective of brain anatomy. This is exemplified by a study yielding negative results for honeybees (Benard & Giurfa, 2004) compared to positive results in paper wasps (Tibbetts et al., 2019). However, few negative results are ever reported, and hence more studies on TI are needed on species where the ecological approach predicts failure.

Here, we tested the presence or absence of TI in the bluestreak cleaner wrasse *Labroides dimidiatus.* This protogynous species has a size-based social hierarchy (Robertson, 1972), where the largest individual in a group is a male that has a harem of smaller females. Males have a higher reproductive output than females, and sex change is socially suppressed via aggression (Robertson, 1972) from the male to the targeted female. While this has yet to be studied in female – female cleaner wrasse interactions, size-based hierarchies are apparently stabilized by individuals suppressing the growth of subordinate individuals through the threat of aggression (Booth, 1995; Buston, 2003). Cleaner wrasse harems are of small size (typically

3-6 adult females in (Robertson, 1972) (mean 3.5 females per harem, Sakai & Kohda, 2001). Thus, cleaner wrasse moving between harems can use absolute information (size) to assess the relative strength of a few conspecifics without the need for TI. Similarly, interactions with clients allow direct assessment of client size, mucus quality, and parasite load, so that TI is not needed to establish a hierarchy of client quality as a food patch. Therefore, if the presence of TI is driven by ecological needs (Hutchins, 2010; Kamil, 1998; Shettleworth, 2009), we hypothesize that we should not find it in cleaners. Note that our interpretation of the cleaner wrasse social system differs from that of colleagues who independently and almost in parallel to us, also studied TI in cleaner wrasse (Hotta et al., 2020). While we do not see an ecological need for TI in cleaner wrasse, they have a complex interspecific social life, with about 2000 interactions per day with ‘client’ fishes (Grutter, 1996b). Clients visit to have ectoparasites removed, but cleaners prefer to eat client mucus (Grutter & Bshary, 2003b). The resulting conflict of interest has apparently led to the evolution of high strategic sophistication in cleaners (Grutter & Bshary, 2003b; Salwiczek et al., 2012) allowing them to match or even outperform mammals in various tasks (Aellen et al., 2021; Salwiczek et al., 2012; Triki et al., 2022). Also, cleaners have an impressive cognitive tool kit for an ectotherm vertebrate that includes generalized rule learning (Wismer et al., 2016), long-term memory retention of single events (Triki et al., 2019), and mirror self-recognition (Kohda et al., 2019, 2022b). Thus, we decided to test a “smart” species that we predict, based on our understanding of its ecological needs, should not possess TI mechanism. This will help elucidate whether TI could be an emergent property of a complex social life (see Oliveira et Bshary 2021 for the similarities between intra- and interspecific social interactions) that involves cooperation, defection, and competition and thus warrants social competence (Taborsky & Oliveira, 2012).

To test for the presence of TI in cleaner wrasse, we used a standard experimental setup that is based on a five-term series test where the individual must learn and understand the transitive link between five objects (e.g., A > B > C > D > E). Previous studies investigating TI varied the number of objects from three to seven (Gazes et al., 2017; Taborsky & Oliveira, 2012; Tibbetts et al., 2019). Three objects (A > B > C, testing AC pair (Gazes et al., 2017)) are not sufficient for assessing TI presence/absence with this paradigm, since success in the AC test can be explained as a simple associative learning component (*i.e.,* A was always rewarded and C never, so I choose A). Even with five objects and the crucial test B vs D, researchers have reported various potential alternative explanations for transitive-like responses (Allen, 2006; Vasconcelos, 2008). These explanations are linked to different types of learning, such as reinforcement history, value transfer mechanisms, and absolute knowledge. For example, TI experiments may yield a systematic difference in reinforcement history between BC and DE combinations. This is because B is losing against the most dominant object (A) and its value needs to be re-established, while E is never rewarded and hence learning D > E is relatively straightforward and achievable in fewer trials than B > C. These differences between pairs may cause individuals to favor B over D simply because B has been rewarded more often until the learning criterion had been reached, without any implication of transitive inference. To control for these alternative explanations/confounding variables, we explicitly tested for effects of the order of pair presentations and the number of trials for each object pair on individual performance in the experimental trials. Note that at the time of data collection, we were unaware that the parallel study obtained positive evidence for TI, using the same general paradigm (see below) but different training methods (Hotta et al., 2020). In light of diverging results, we dedicate an important part of the discussion to compare the methods.

2.3. Material and Methods

The study was conducted in Australia, at the Lizard Island Research Station, from February to March 2020. Due to the loss of one individual during the training phase, a total of 19 fish were tested. All individuals were released at the site of capture after the experiment. The manipulations were approved by Animal Ethics Committee (AEC), permit number CA 2019/11/1336.

2.3.1. Training

The training phase aimed to make fish understand the relative value of each plate compared to the others. To do this, 4 pairs of plates (hereafter referred to as the training pairs) were presented: A+B-, B+C-, C+D, D+E-, where plus and minus indicate the reinforced and the non-reinforced choice, respectively. On the reinforced plate, a small food item of mashed prawn was put on the back of the plate. The non-reinforced plate did not contain any food. Cleaners had to choose between the two simultaneously presented plates, which were separated by an opaque barrier (Fig. 11). A choice was scored once the cleaner went behind one of the plates. The cleaners could hence always inspect the back of the chosen plate, which showed the same color pattern as the front to allow a post-choice feedback. Thus, they could eat the food item if the choice had been correct, and afterward still access the non-rewarding plate to see that there was no food. In contrast, the correct plate was removed and hence inaccessible if the choice had been wrong. In order to prevent any systematic preferences in color patterns and/or biases due to the sequence of presentation of pairs of plates, we counterbalanced these variables across subjects (Fig. 12).

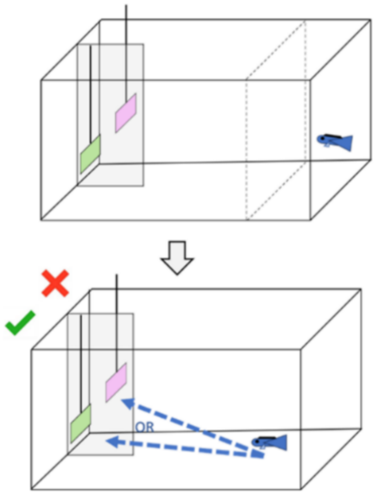


Fig 11. Experimental setup
The fish is put behind a see-through barrier (dotted lines] while the plates are inserted and separated by an opaque barrier (full line and grey rectangle). If the fish goes to the right one, he gets the reward and can inspect the other plate (green "V") whereas both were removed directly if he chose the wrong one (red cross). After the training, we performed the transitive inference task to test whether fish can infer the value of plates from new combinations (i.e. test combinations): B+D- and A+E-. During the task, the training combinations were randomly exposed to the fish and the test combinations were also randomly presented once in a while.

To facilitate learning, subjects could inspect both plates during the first five trials for each pair of plates, irrespective of their initial choice. With this addition, we ensured early exposure to a rewarding plate that had just before been the non-rewarding plate in the previous combination. Subjects were considered to have learned the correct choice if they reached one of the following criteria: 3 times at least 7 correct choices out of 10 trials, 2 times at least 8 correct choices out of 10 trials, or 10 correct out of 10 trials. Once subjects had subsequently reached learning criteria for all four combinations, we ran a single set of four trials in which each pair was presented once. We conducted 10 trials per fish per day.

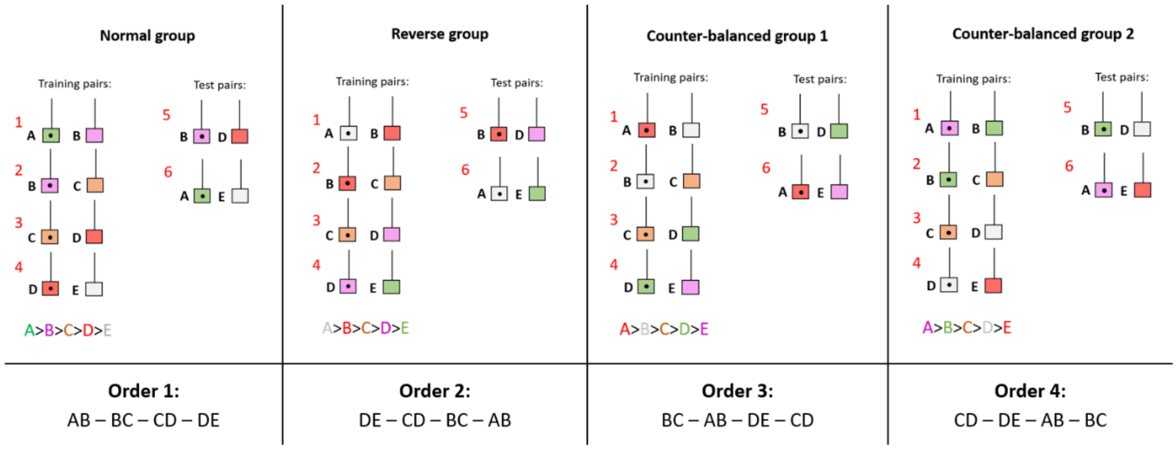


Fig 12. Representation of the four experimental colors and order of learning groups

The 20 initial fish were equally divided between four experimental groups that differed with respect to the color rank of the plates: Normal, reverse, counter-balanced 1 and 2. Within each group, training pairs are learned in 4 different orders. Each color group included the 4 order of learning groups. After one loss, the counter-balanced group 2 had only 4 fish and the order 1 was not represented anymore.

2.3.2. Transitive inference task

Prior to the task, four reminding trials were performed in order to expose individuals to each training pair again right before the task. For the task, we tested the subjects' spontaneous preferences when confronted with two novel combinations: plate A versus plate E, and plate B versus plate D. As A had previously only been used as a rewarding stimulus, and E only as a non-rewarding stimulus, we expected a clear preference for plate A when presented with E. In contrast, both plate B and plate D had been rewarding in one plate combination, and non-rewarding in the second plate combination. Thus, only if subjects had formed a hierarchical representation of the five plates, the use of transitive inference would make them spontaneously prefer plate B over plate D. Otherwise, plate choice was expected to be random. Subjects were tested with both AE and BD combinations three times each.

To avoid habituation to such conditions, the six test trials were interspersed in between a sequence of learned plate pairs (AB, BC, CD, DE) where the correct choice yielded the standard reward of a prawn item. More precisely, a single test trial was included in a sequence of 4 training pairs trials. Thus, a total of 30 trials were conducted to obtain our six experimental trials. Moreover, the plate pair presented in the trial before a test was varied as much as possible within individuals and counterbalanced across individuals.

2.3.3. Statistical analyses

All the analyses were carried out on R v.4.0.2.

To check whether the fish performed above chance for the entire set of training pairs during the task, we first calculated the proportion of successful trials for each individual. The % values were then used for a generalized linear mixed effect model that compared observed values to the expected binomial distribution (50% success) based on the null hypothesis that no learning had taken place. We included the fish ID as random variable to avoid pseudo-replication. The significance of the intercept was the indicator of the overall learning of the training pairs or not.

While every fish passed the learning criterion for every training pair and was hence ready for testing, we also analyzed whether all the training pairs were equally remembered during the actual TI task. We performed a Friedman rank sum test on the count of success per training pair, considering the fish identity. We then conducted pairwise comparisons between all training pairs using the Wilcoxon rank sum test with continuity correction. We built different binomial generalized linear mixed-effects models to test the different possible explanatory variables. The choice at each trial was the response variable, and the training pair was always included. Because the amount of data was not sufficient to include both the color groups and the order of learning groups plus all the interaction terms, we had to perform separate models. After model selection using variances analyses, and because of the experimental design, the order of learning groups as well as the interaction term with the training pair were kept in the model while the color group was not.

Finally, to check which training pair was remembered above chance, we performed individual Wilcoxon rank sum tests with continuity corrections for each training pair after graphically checking that the data were homogeneously distributed around the median. The same analyses were performed on each training pair for each order of learning and repeated separately with the fish having an overall memory above 65% on the training pairs.

To compare the test pairs AE and BD, we started by investigating the data of the first trial of the task only, using a two-tailed binomial test for each pairing. When considering the three trials of the task, we applied the same method as we used on the training pairs to see whether the fish succeeded above chance and to investigate the potential roles of the different variables. In this case, the final model included only the test pair as an explanatory variable, since the amount of data did not allow to run the model including the variables considered in our other analyses (i.e., color and order of learning).

To evaluate the extent to which reinforcement history during the initial training phase may have affected choices in the BD test pair, we calculated the ratio between the number of times that B or D plate was rewarded versus non-rewarded: for B the number of succeeded BC trials over failed AB trials, for D the number of succeeded DE trials over failed CD trials. The difference between these “Rb” and “Rd” was calculated and analyzed in a non-parametric Kruskal-Wallis test (because of the small sample size) to see whether there was a systematic difference in reinforcement history between B and D, or a difference between each group of learning order. Furthermore, to test whether variation in the ratio difference could be correlated to a difference in individual success in the BD task, we used a Spearman correlation test on the ratio difference and the counts of success for BD.

All the data and R codes are publicly accessible on figshare (https://figshare.com/projects/No_evidence_for_transitive_Inference_in_cleaner_wrasse_data/158132).

2.4. Results

2.4.1. Training pairs

Overall, the fish remembered the training pairs above chance, succeeding on average in 61.4% of trials (null binomial glmer, intercept: $p < 0.0001$). However, the training pairs were not all equally remembered (Friedman chi-squared = 10.34, df = 3, p = 0.02, Fig. 13a). AB (median = 5 with se = 0.359) was remembered significantly better than BC (median = 2 with se = 0.413, pairwise Wilcoxon rank sum test with Bonferroni correction, p = 0.008), and DE (median = 4 with se = 0.350) tended to be remembered better than BC (p = 0.07). The anchoring pairs (AB and DE) were remembered above chance levels (Wilcoxon signed rank test with continuity correction, respectively V = 143 with p = 0.001 and V=117 with p=0.01, Fig. 13a), while the performances in trials involving the two middle pairs (BC and CD) did not differ significantly from chance (respectively V = 50.5 with p = 0.37 and V = 72 with p = 0.51). A closer inspection of the most successful fish (overall performance above 65%, 6 individuals) showed similar tendencies, with the middle pairs (BC and CD) not equally remembered (Friedman rank sum test, chi-squared = 11.509 of 3 df, p = 0.009). Due to the limited sample size of this subset of

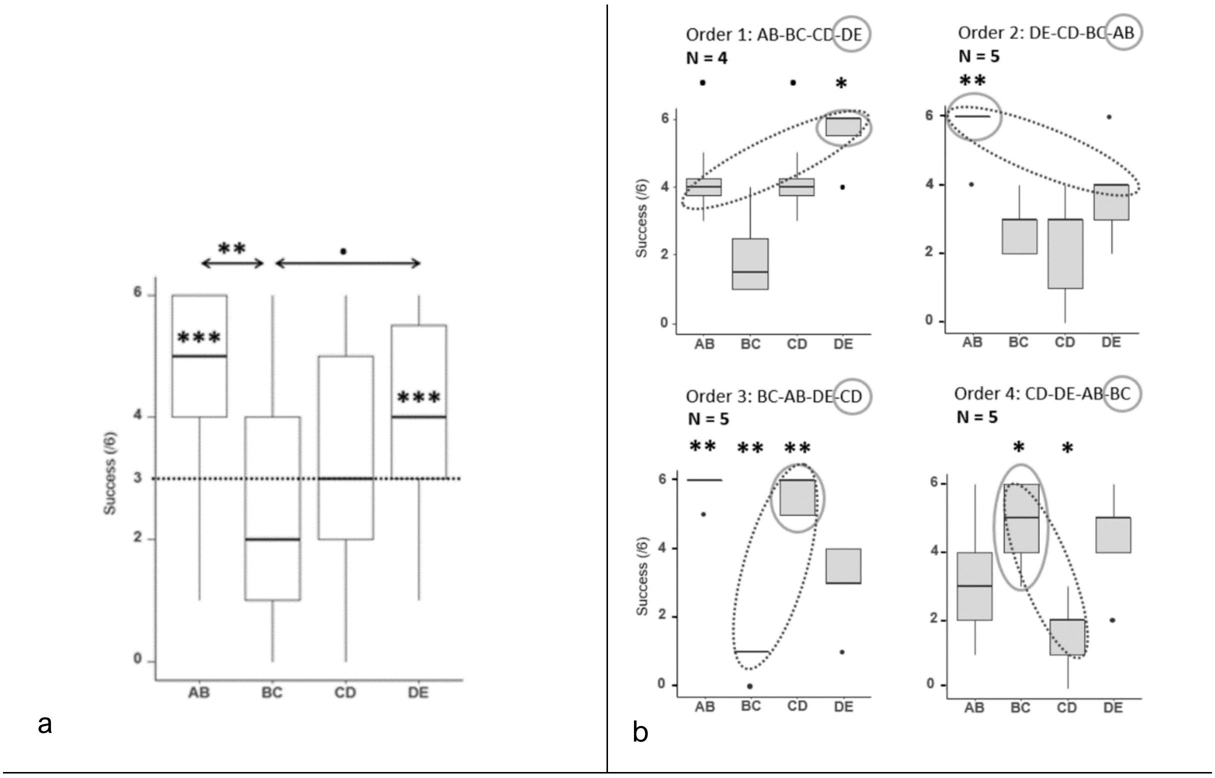


Fig 13. Graphical representations of the count of choices for each training pair globally (a) and for each order of learning separately (b) during the task.

The a boxplot shows the number of correct choices fish made during the task. Each test pair was presented 6 times. The large horizontal lines indicate the medians, boxes are limited by the upper and the lower quartiles, and the vertical lines indicate the deciles. On the A graph, the large black dotted line indicates the random percentage of success of 50%. The stars inside the boxes show the significance of the difference from the random threshold.

Significance code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

The b boxplots contain the same information for each order of learning separately. In addition, the clear grey full-lined circles point the pair lastly learned and the darker dotted circles show the group of the pairs learned in the first and the last position.

fish, the assumption for the Wilcoxon rank sum tests was not met (homogeneity around the median) and no further analyses were conducted.

We also found that there was a significant effect of learning order on remembrance of the training pairs during the task (anova on binomial glmer, Chisq = 103.8 of 9 df, $p < 0.005$, Fig. 13b). All groups were able to reach the chance level and above for AB and DE (mostly at chance for DE), while middle pairs (BC and CD) were only remembered above chance when learned in the last position. For example, BC was the last pair to be learned for group 4 and they remembered it in 80% of trials (Wilcoxon rank sum test, W = 22.5, p = 0.006, Fig. 13b), whereas the three other groups show performances at or lower than chance level.

2.4.2. Test pairs

For the AE pair on the first trial of the task only, 19 of 19 subjects chose the A plate (two-sided binomial test: probability of success = 1 and $p < 0.0001$, Fig. 14a). In contrast, only 8 of 19 cleaners chose B over D (two-sided binomial test, probability of success = 0.42 and p = 0.65, Fig. 14a).

To compare our results to the study of Hotta et al. 2020, we checked whether the fish who learned the training pair in the rank-descending order (from highest to lowest) as their subjects had done (our order 1: AB, BC, CD, DE) would show similar performances. Among these 4 individuals, two chose B during the first trial, compared to 4 of 4 in their study (Note that Hotta et al. erroneously reported 3 of 4 correct first choices; the archived data show 4 of 4).

Considering all 3 trials per task, the results remain unchanged. For the AE pair, fish significantly preferred plate A over plate E (mean of success = 2.6, Wilcoxon signed rank test with continuity correction, V = 186.5, $p < 0.001$), while they did not solve the BD pair above chance (mean of success = 1.47, V = 92, p = 0.92). The four fish trained in a similar order to those from Hotta et al. (2020) made five correct choices out of twelve trials. This strongly contrasts with the findings of Hotta et al. (2020), which reported a strong preference for D (our B) in the BD pair in all subjects, with a minimum of 10 of 12 correct choices per fish.

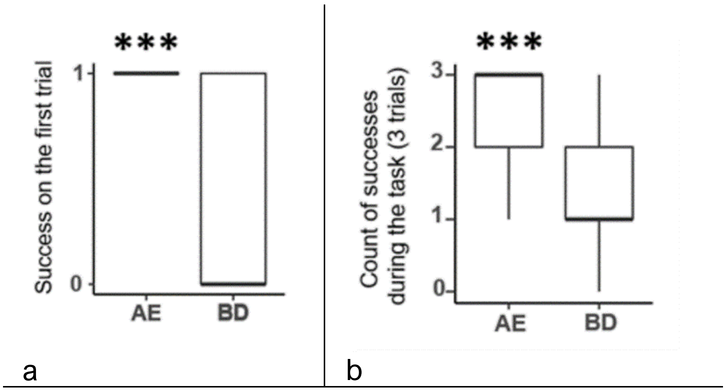


Fig 14. Graphical representations of the success of all fish for each test pair separately including on the first trial of the task (a) and including the 3 trials (b).
The stars above the boxes show the significance of the difference from it. The large horizontal lines indicate the medians, boxes are limited by the upper and the lower quartiles and the vertical lines indicate the deciles.
Significance code: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Because of the difference among groups of learning order for the training pairs, even if the learning order appeared non-significant overall in the analyses of the test pairs, we checked for the group that remembered the training pair BC the most (73.3% of all trials were correct during the task). It is also the group that learned BC last. In this group of fish (N = 5, learning order 4), 3 of 5 fish chose B over D in the 3 trials. The 2 last fish succeeded only in the last trial and one of these successfully remembered the training pairs 42% of the time (highlighting a more general weakness in the remembrance of the pairs). Note also that DE was equally remembered in this group (around 75%, does not appear significantly different from the random threshold, probably due to the small sample size). Due to the small sample size (N = 5), no subsequent analyses were carried out for these groups.

In our final analysis, we evaluated the reinforcement history of plates B and D. The differences between Rb and Rd ratios were not similar among the four groups of learning order (Kruskal-Wallis, Chisq = 9.95, p = 0.02, Fig. 15). The first group significantly differed from group 2 and 4 (pairwise comparisons using Fisher's least significant difference, for both: difference = 10.5, p = 0.01, Fig. 15) and tended to differ from group 3 in the same way (Fisher's least significant difference, difference = 7.5, p = 0.06, Fig. 15).

Finally, the individual differences in the reinforcement history of B and D did not correlate with the success in the BD task (Spearman's correlation test, rho = -0.05, p = 0.8).

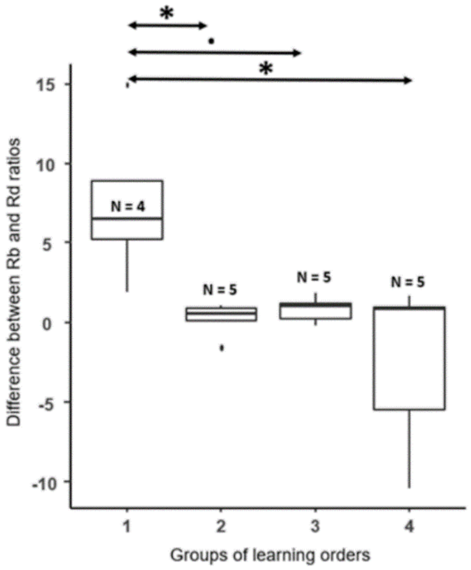


Fig 15. Graphical representations of the difference between Rb and Rd ratios in each of the four groups of learning orders.

Rb and Rd are the ratio between the number of times that B or D were rewarded and non-rewarded during the training. The Y axis is the calculated difference between those 2 ratios.

The large horizontal lines indicate the medians, boxes are limited by the upper and the lower quartiles and the vertical lines indicate the deciles. N indicates the number of fish for each group.

The stars above the boxes show the significance of the comparison between the boxes indicated by the arrows.

Significance code : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

2.5. Discussion

An increasing number of studies have found evidence for TI among birds (e.g., jays (Bond et al., 2003; Paz-y-Miño C et al., 2004); crows (Lazareva et al., 2004; Lazareva & Wasserman, 2006); pigeons (Couvillon & Bitterman, 1992; T. R. Zentall et al., 2019); geese (Weiß et al., 2010)), mammals (e.g., chimpanzees (Gillan, 1981); macaques (Jensen et al., 2016); lemurs (Tromp et al., 2015); mice (Silverman et al., 2015)), and fishes (e.g., brook trout (White & Gowan, 2013); cichlids (Grosenick et al., 2007); cleaner wrasses (Hotta et al., 2020)) that is predicted to derive from the ecological needs of the species of interest. However, potential alternative explanations associated with positive findings, as well as a paucity of published negative results (but see Benard et Giurfa 2004), make it important to also test species where the ecological approach to cognition predicts that they should lack TI. In this study, we tested whether cleaner fish use TI to organize information in their brain. We hypothesized that cleaners are lacking ecological conditions where TI would be a useful cognitive tool and thus should lack this cognitive ability according to the ecological approach to cognition. Our view contrasts with that of Hotta et al. (2020) who proposed that the ecology of cleaner wrasse predicts the presence of TI. Readers need to decide whether they would expect TI to be present in a protogynous species living in harems that very rarely contain more than six females, with a size-based hierarchy. While they found positive evidence for TI in cleaner fish in their almost simultaneously conducted study (Hotta et al., 2020), we found no evidence for it in our study. A large part of our discussion therefore focuses on potential explanations for the diverging results. As we see it, differences in training provide the most likely explanation.

Importantly, failure to show TI in an experiment could be due to a number of factors, which may or may not be mutually exclusive. One potential explanation would be that the species tested lacks the cognitive process. Alternatively, an inability to memorize the relationship in every training pair may prevent the subject from exhibiting their capacity for TI. The latter is evident in our results, as cleaners lacked the remembrance of the middle pairs BC and CD that were presented between the actual TI trials. This was not expected, since all fish successfully reached the learning criterion during initial training for each pair. Thus, cleaners could not properly remember all four pairs. Instead, our results highlight two well-known effects in TI studies that contribute to this middle pairs outcome: (1) the recency effect (i.e., an item is favored if it was seen just before) and (2) the anchoring effect (i.e., the first and the last pair of the list are better remembered than the middle pairs). The memory performances on our training pairs were highly impacted by the recency effect, as the better-remembered pair was the one that was learned the last. Moreover, the observed U-shaped pattern of the remembrance of each training pair, in addition to the failure to remember the middle pairs above chance, are indicative of the anchoring effect. This U-shaped pattern has been previously reported in both unsuccessful (Benard & Giurfa, 2004) and successful (Bryant & Trabasso, 1971; Daisley et al., 2010; Mikolasch et al., 2013; Woocher et al., 1978) TI experiments. In previous successful studies (Bryant & Trabasso, 1971; Daisley et al., 2010; Mikolasch et al., 2013; Woocher et al., 1978), the middle pairs were still remembered above chance, unlike in the present study. The U-shaped pattern provides a potential explanation for lower preferences in the BD pair compared to the AE pair (Von Fersen et al., 1991) and could be explained by the serial position effect (SPE), which predicts that performances on further objects in a series are higher compared to the performances on neighboring objects (Bryant & Trabasso, 1971; Lazareva et al., 2020; McGonigle & Chalmers, 1984; Vasconcelos, 2008, p. 200; Woocher et al., 1978; Wu & Levy, 2001). In our case, the U-shaped pattern could not be found when the results from the 4 groups of learning order were looked at independently. Only the combination of the results from all of these groups showed the U-shaped pattern. When we looked at each order of learning independently, we could see again the anchoring effect and clearly observe the recency effect. Indeed, by comparing orders of learning 1 and 2 to orders 3 and 4, we saw that the difference between the first and last learned pairs was reduced when those were the anchoring pairs AB and DE. We observed that the better-remembered training pair for each group of learning order was the pair that was learned the last, which is a clear sign of the recency effect. These effects had the same importance for the 6 fish with an overall memory equal to or above 65%, which indicates that the failure in BC and CD was not influenced by the overall success in the memorization of the training pairs. These memory differences between the anchoring and the middle pairs can explain both the precision with which subjects preferred to approach A over E and the lack of preference in BD trials.

Several other potentially confounding variables reported in previous studies were absent in our study, and our sample size allowed us to control for different factors that could have biased the results. Our color groups avoided a particular value attribution to a specific plate (as in Grosenick et al., 2007 with artificial dominance). Also, using 4 groups of counterbalanced learning order made it unnecessary to have a “bias reversal procedure” that consists of over-reinforcing D over B. Then, based on reinforcement history, the individual would significantly choose D in the BD task (Lazareva et al., 2004; Lazareva & Wasserman, 2006). Hotta et al. (2020) used the same sequence of training for their four subjects. Our group of four subjects that experienced the same order of learning as the fish in the study by Hotta et al. (Hotta et al., 2020) still showed considerably different results. Moreover, in our study, the pair of plates presented in the trial before a test was varied as much as possible within individuals and counterbalanced across individuals to avoid a systematic bias that could result from a recency effect (Bond et al., 2003), which could, for example, influence an individual to choose B over D if exposed to B+C- in the previous trial. Furthermore, it appears that the success/failure of our fish was not affected by differences in the reinforcement history of the various plate combinations, which have been suggested as means to yield transitive-like responses in TI tasks (Couvillon & Bitterman, 1992; Wynne et al., 1992). In contrast, reinforcement history could explain the success of three out of four fish used by Hotta et al. (2020). In summary, it appears that no systematic bias could have produced positive results with our study design; poor memorization of the middle pairs (BC and CD) emerges as the main reason for the cleaners’ poor performance. The reason for this failure is still interesting as it sticks to ecological relevancy, so we discuss it later.

2.5.1. Comparison to Hotta et al. 2020

When comparing our study to the work of Hotta et al. (2020), there are several methodological differences that may have contributed to our divergent findings (presented in Table 4). One potential explanation, which we consider unlikely, is that the positive results reported by Hotta et al. (2020) were due to a combination of chance and small sample size. The authors only tested four cleaners compared to our 19, but all these four cleaners consistently preferred the B plate over the D plate from trial 1 onwards over a series of 12 trials (note that the authors erroneously report only 3 of 4 correct first choices; their published data show 4 of 4 correct choices). Thus, while their sample size was much smaller than ours, we consider it parsimonious to conclude that the differences in performance are due to differences in methods.

A major difference between our two studies involves the number of training pairs. Hotta et. al. (2020), as well as other studies (Jensen et al., 2016; Weiß et al., 2010), added non-adjacent pairs to their training (i.e., AC and CE). These combinations may have helped to consolidate the subjects' memory of the BC and CD combinations. One obvious prerequisite for the ability to demonstrate TI is that subjects memorize the ranks within all pairs so that they can deduce the ranking within novel pairs. In our series of 30 experimental trials, this prerequisite was fulfilled only for the anchoring pairs AB and DE. By adding those non-adjacent pairs, subjects obtained complete information about the rank position of plate C; it is non-rewarding when paired with A or B, and it is rewarding when paired with D and E. This extended knowledge could be necessary for the fish to establish the link between the different plates and be able to infer the value of B over D, potentially by helping to the establishment of a hierarchical list of the plates.

Table 4: Main differences in methods and results between Hotta et al. 2020 and the present study

Difference		Hotta et al. 2020	This study
Methods	Plates	Different colors	Colors and pattern differences
	Reward	Smeared prawn on the front, get one if succeeded or none if failed	One prawn item hidden in the back, get it if succeeded or none if failed
	Time between trials	1 or 2min between the trials, 2 or 3h between the sessions	20 to 50min between trials
	Sample size	N = 4	N = 19
	Training	- Phase 1: A-B+, B-C+, A-C+, C-D+, D-E+, C-E+ Criteria: 2 x 5 or 6/6 - Phase 2: A-B+, B-C+, C-D+, D-E+ Criteria: 1 x 5 or 6/6 - Phase 3: Sessions of 8 trials, each pair presented twice. Criteria: at least 1/2 for each pair per session and over 6/8 in two consecutive sessions	- 5 trials to present each new pair. - Learning: A+B-, B+C-, C+D-, D+E- Criteria: 3 x 7/10; 2 x 8,9 or 10/10, 9 or 10/10 + 7/10 - Reminding trials: 4 trials, each pair presented once before testing
	Test	3 x 4 trials for BD	3 x BD and 3 x AE, randomly presented within training pairs trials [10 trials per day, 1 test pair presented per 5 trials]
	Counterbalancing	BD colors	Colors, order of learning, which training plate was presented prior to a test pair
Results	Training pairs	All remembered equally	Not remembered equally
	Test	B-D+ succeeded	B+D- failed. A+E- succeeded
	Reinforcement history	Could explain the transitive response for 3 out of 4 fish	No global difference for B and D but variations depending on the group of learning order. Not correlated with BD success

Note: the list is not exhaustive, only the bigger ones were listed. Readers are advised to refer to Hotta et al. 2020 to get more details on terminology or else.

An individual then has absolute knowledge about the plates located in both parts of the final list (i.e., A, B, C and C, D, E) that they can use to infer the new BD relationship. Without this supplementary knowledge, it is possible that the list could not be established, so the different pairs would remain as unlinked objects, preventing TI to be applied. The fact that some studies have yielded positive results using a similar experimental design to that used here—without the non-adjacent pairs training—could be a sign that the establishment of the list is already a cognitively demanding step that necessitates extended training for fish compared to some other species.

An interesting middle ground between these two cleaner studies could be to present only the four adjacent pairs but with additional rounds of BC and CD trials until all correct choices are properly memorized. On the other hand, we think that it is an important result of our study that cleaners find it challenging to memorize the middle pairs. This result raises an interesting question—what ecological conditions would yield enough training for memorization of frequently occurring pairs before TI can be used to spontaneously resolve a unique new combination? The advantage of TI is that it is a capacity that allows one to get maximum information with minimum learning. Our negative results suggest that cleaners cannot apply TI without extensive prior training in the four combinations, which negates its major advantage. While the positive results by Hotta et al. (2020) seem to suggest that cleaners can apply TI provided that they receive extensive training, we reiterate that we consider their additional training on the AC and CE combinations as a potentially confounding factor. In any case, we consider the fact that our cleaners failed to remember all combinations after they all succeeded the initial learning also points at the ecological irrelevancy of the TI task. Nevertheless, we also acknowledge that replicating the social paradigm used in other species of fishes (Grosenick et al., 2007; White & Gowan, 2013) could help to clarify whether cleaners could succeed more easily (i.e., without an additional non-adjacent pair training) in this type of TI task compared to that utilized in a foraging context. The cognitive process of TI might still be present in cleaners, even if our study did not find evidence for it in this context.

2.5.2. Reinforcement history

It has been noted several times in the literature that a difference in reinforcement history could lead to a transitive-like response in TI tasks (Couvillon & Bitterman, 1992; Wynne et al., 1992). For example, fish subjected to five-item tasks in our study might on average need more reinforcement trials to learn BC than DE because B only loses against a plate with a pure positive association (plate A), while D only wins against a plate with a pure negative association (plate E). More reinforcement trials on B could thus potentially lead to a preference for B over D. That said, we did not find such inequality in the number of reinforcement trials (except for one group of learning order) in our study. Moreover, individual differences in reinforcement history ratios did not correlate with a preference for B or D.

2.5.3. Transitive inference tests

The types of evidence needed to adequately test for the presence of TI continue to be debated (Allen, 2006; Galizio et al., 2017; Guez & Audley, 2013), but the generally accepted paradigm typically involves information on social dominance—where few observations of pairs fighting may reveal information of relative strength without subjects receiving any rewards (Grosenick et al., 2007; White & Gowan, 2013). In contrast, our paradigm can be criticized on the basis that a plate's 'state' is 1 (food) or 0 (no food) depending on the plate it is paired with. Such a presentation does not match on some continuous scale. Unfortunately, experiments involving reinforcement learning cannot use a continuous scale (like fighting ability) as subjects would obtain absolute information on food reward (e.g., A offers more food than B, which offers more food than C, etc.) so that TI is not needed for subjects to prefer B over D. Nevertheless, we note that the food paradigm has been used on various species (Benard & Giurfa, 2004; Davis, 1992; Gillan, 1981; Zentall et al., 2019); mix with social hierarchy (Tromp et al., 2015), often yielding positive results. These differences are interesting and interpretable. The comparison between paper wasps and bees (Benard & Giurfa, 2004; Tibbetts et al., 2019), as well as within-species variations in chicks (Daisley et al., 2021), fits the ecological approach to cognition, and we argue that our negative results do so as well. Cleaner fish obtain absolute information about client value as a food source in nature, and their social hierarchies in each harem are based on relative body size (i.e., also based on absolute information). Thus, we would argue that cleaners should not need to employ transitive inference under natural conditions.

2.5.4. Conclusion and outlook

The ecological approach to cognition has helped to identify various high performances in species that are phylogenetically distant from humans, as demonstrated in specific tasks that correspond to relevant ecological pressures. A classic example involves the selective spatial memory abilities of Clarke's nutcrackers compared to closely related species that rely less on seed caching (Joseph, 1995). The ecological approach has also helped to identify suitable experimental paradigms to test for advanced cognitive processes, for example by using food caching to test for episodic-like memory in scrub-jays (Clayton & Dickinson, 1998). The cleaner wrasse *L. dimidiatus* provides another suitable species for the ecological approach as their foraging conflict with clients has apparently selected for the ability to adjust to context-specific situations—i.e., whether the interaction is observed (Pinto et al., 2011b), whether they clean alone or with a partner (Bshary et al., 2008), their own physiological state (Binning et al., 2017), client species features (Roche et al., 2021) or fish densities (Triki et al., 2019, 2022). This fine-tuned strategic sophistication raises questions regarding the underlying cognitive processes that can explain these behaviors. As it stands, cleaners perform well in tasks testing for two executive functions, namely self-control (Aellen et al., 2021) and flexibility (Triki & Bshary, 2019). While these functions are clearly useful for their interactions with clients, transitive inference is not needed as far as we can judge. Given this, we would argue that designing studies where failure is predicted is an important complement to test the ecological approach to cognition. This is because the ecological approach implicitly assumes that brains function in rather modular ways. Thus, positive results where the ecological approach predicts failure would identify aspects of brain functioning that are more general-purpose tools. Our study suggests that cleaner fish would not use the general-purpose tool of transitive inference in their daily life, though our data do not allow us to distinguish whether this is due to memory constraints or to the absence of the TI process. We also propose that our data can be reconciled with the results of Hotta et al. (2020) in that only additional training and performance control before the crucial test may yield positive results, albeit with the caveat that associative learning cannot be excluded as a simpler mechanism. It would be interesting to test larger-brained endotherm vertebrate species that also do not need TI in their daily lives to see whether this cognitive tool might eventually emerge as a by-product of larger computing power.

2.6. Acknowledgments

The staff from the Lizard Island Research Station. N. Truskanov and Z. Yanai for helping with the setup and manipulations, the statistician R. Slobodeanu for his advice and answers, and C. Clements for his review of the manuscript.

3. Chapter 3: Cleaner wrasse can reach Piagetian stage 4 of object permanence when primarily exposed to stage 3

Leonore BONIN*, Sophie DUC, Redouan BSHARY

* corresponding author

leonorebonin@gmail.com

ORCID Leonore Bonin: 0000-0001-9349-0616

Affiliation (same for all authors):

University of Neuchâtel, Biology Institute, Neuchâtel, Switzerland

Submitted to Animal Cognition

Keywords
Object permanence – ecological cognition – fish cognition – Piagetian stage 4 – cleaner wrasse – animal cognition

3.1. Abstract

Object permanence (OP) is the ability to retain a stable representation of an object, even if it is partially or completely out of sight. Despite its seeming simplicity, this concept has posed challenges for many species, with factors like life history and development being proposed as potential explanations. We tested for OP in the cleaner wrasse *Labroides dimidiatus*, a species that exhibits complex ecological strategies in the wild, who has been tested in a variety of cognitive experiments, and could help shed light on differences in brain size and cognitive abilities between ectotherms and endotherms. In two separate series of experiments, we investigated the potential role of ecological relevancy and the complexity of the condition. To test for OP, we used a food reward, a conspecific, or a model predator that we presented to cleaner wrasse Piagetian stage 3 or 4 and added a delay or a supplementary cognitive charge with an "A-not-B error" paradigm. Our findings suggest that cleaner wrasse possess some primary forms of OP that depend on training and are thus comparable in performance to some larger-brained endotherm species. Furthermore, ecological relevancy did not clearly influence cleaner wrasse performance. Our findings provide the first evidence of OP in cleaner wrasse; however, we discuss methodological improvements that could help minimize potential biases in future studies.

3.2. Introduction

Fifty-five years ago, Jerison (1969) highlighted an average 10-fold difference in the relative brain size of ectotherms and endotherms, which he referred to as “lower vertebrates” and “high vertebrates”. Given that brain tissue is energetically costly (Fonseca-Azevedo & Herculano-Houzel, 2012; Sukhum et al., 2016), it raises questions about the benefits of a bigger brain in endotherms. One hypothesis is that endotherms have a more extensive cognitive toolkit than ectotherms. The underlying logic entails that a small brain does not have enough neuronal tissue to host the same abilities as a bigger brain. However, past studies on fish cognition did not show a clear gap in the cognitive toolkit compared to endotherms. Indeed, researchers have found cognitive abilities that go beyond Pavlovian or operant conditioning (Pavlov, 2010; Skinner, 1937), such as social learning and traditions (Helfman & Schultz, 1984; Truskanov et al., 2020), transitive inference (under specific conditions; Bonin & Bshary, 2023; Hotta et al., 2020), generalized rule learning (Wismer et al., 2016), and even mirror self-recognition based on a mental representation of the own face (Kohda et al., 2019, 2022a, 2023).

Studying human infant development can help understand the link between the presence/absence of cognitive ability and the amount of brain tissue. Following this logic, Piaget and Cook (Piaget & Cook, 1952, 1954) established six different stages in development, clearly defined by age, that led to the full capacity of object permanence. Object permanence (OP) is the ability to remember that an object still exists even when placed out of sight (Piaget & Cook, 1952, 1954). The complete object permanence capacity, also called the multiple invisible displacements (Piagetian stage 6, Piaget & Cook, 1952, 1954), can be reached by children from 12 months old onward. Building upon the work of Piaget & Cook 1952, 1954), Uzgiris & Hunt (1975) established the Uzgiris-Hunt Ordinal Scales of Psychological Development (UH scale), which seeks to leave space to factors other than age, such as environmental influences, to impact the presence/absence of a cognitive ability. This vision allows a more tempered vision of the development of cognitive skills, and instead of being present or absent, variation in complexity can be found.

This is what we were interested in testing within a well-studied species of fish: the cleaner wrasse *Labroides dimidiatus*. Current evidence suggests that only endotherms possess the most complex form of OP, including some species of “smart” birds (corvids and psittacids; (Hoffmann et al., 2011; Pepperberg & Funk, 1990; Ujfalussy et al., 2013; Zucca et al., 2007), great apes (Barth & Call, 2006; Beran & Minahan, 2000; Call, 2001; De Blois et al., 1998), several monkey and macaque species (Anderson, 2012; Majecka & Pietraszewski, 2018), and elephants (Miller, 2019). Many species, such as captive horned puffins (Huffeldt, 2020), racoons (Colbourne, 2018), dwarf goats (Nawroth et al., 2015), and coal and great tits (Marhounová et al., 2017), did not succeed above stage 4, while others needed extensive training to succeed (glaucous-wigged gulls; Obozova et al., 2011, and pigeons; Zentall & Raley, 2018). Furthermore, some species reached stage 4 but were never tested for stages 5 and 6, such as giraffes (with a 30 seconds delay; Caicoya et al., 2021; Caicoya et al., 2019), European bison and forest buffaloes (Caicoya et al., 2021), and guppies (when selected for bigger telencephalon; Triki et al., 2023). Other species also exhibited intermediate results. Pigeons needed extensive training or a delay to reach stage 4 (Zentall & Raley, 2018) while bottlenose dolphins exhibited contradictory results, one of the differences being the type of response that was measured (Jaakkola et al., 2010; Johnson et al., 2015), which raised the question of whether OP tasks can be appropriately adapted for underwater animals. Note that we detected few variations in the definition of the different stages in the pre-cited studies and based our classification on the definition of Piaget & Cook (1952) (see Jaakkola, 2014 and Pepperberg, 2015 for criticisms and thoughts about OP testing).

While OP has been tested repeatedly in endotherms, tests of OP among ectotherms remain limited. A few studies have been conducted using fish, but these simply tested whether subjects can swim to a reinforced stimulus that is moved out of sight into one of two possible hiding spots (stage 4, with minimal time delay between hiding and testing). Under such conditions, cleaner wrasse (Aellen, Burkart, et al., 2022), as well as cichlids (Guadagno & Triki, 2024), performed at chance levels. In guppies, subjects selected for a bigger relative telencephalon size performed above chance albeit with low precision (60% correct choices), while subjects selected for a relatively smaller telencephalon performed at chance level (Triki et al., 2023). In conclusion, current evidence indicates that fishes have poor skills at OP.

Here, we provide a more systematic study on OP in the reef fish *Labroides dimidiatus*, using the same methodology as Hoffmann et al. (2011) on carrion crows and Zucca et al. (2007) on Eurasian jays, and applying the tasks designed by Uzgiris and Hunt (1975). We chose *L. dimidiatus* because it is a cognitively well-studied species (Bshary & Triki, 2022). Given that cleaners failed at stage 4 OP in a previous study (Aellen et al., 2022), we made various adjustments to our experiments. First, we changed the configuration of the hiding spots from going left/right to placing compartments into opposite corners of the aquarium. We made this change to minimize the possibility that chance level would result from individuals choosing a cheap strategy, such as side preference, instead of actively trying to solve the paradigm. Second, we started on a lower complexity level, i.e., offering half-hidden objects of interest (Piagetian stage 3). Only individuals that pass this level would be kept for testing a completely hidden object (Piagetian stage 4), and only performance above chance would lead to further testing with a longer delay between hiding and testing, as well as a test of the “A-not-B” error (multiple visible displacements, Piagetian stage 5). Third, we considered that variation in the ecological value of the task may affect performance.

In Aellen et al.’s (2022) previous experiment that produced negative results, the object was a plate with a food reward. In nature, cleaners obtain their food from interactions with so-called “client” reef fishes (Côté, 2000). Clients visit cleaner wrasse to have ectoparasites removed (Côté, 2000). Thus, it appears possible that moving a plate with food out of sight mimics a client not being interested in a cleaning interaction, and hence of no current importance for the cleaner fish subject in the OP task. In contrast, other objects should retain relevance even when moved out of sight, such as a conspecific. Cleaner wrasse are protogynous hermaphrodites, living in harems with a size-based hierarchy, where the largest individual is the male (Robertson, 1972; Sakai et al., 2001). We hypothesized that a female should find the experimental presentation of a new, smaller female highly relevant and remember her existence even if out of sight—even if it would not necessarily result in an active “looking for her” type of behaviour in nature. The other object of high ecological relevance is a potential predator. As it stands, cleaners are not entirely immune to predation risk (Bshary & Noë, 2003; Grutter, 2004). Thus, a sit-and-wait predator like a lizardfish moving out of sight may remain close and, hence, a potential threat in a complex three-dimensional habitat like a coral reef. Cleaner wrasse should therefore be cautious or avoid the place where the predator disappeared if possible.

In the first set of experiments, we tested cleaners for different complexity levels, comparing their performance when exposed to (1) a food reward plate, (2) a smaller conspecific or (3) a 3D model of a predator (lizardfish). While we had no specific predictions regarding the maximal complexity level that cleaners can still solve, we expected that both the conspecific and the predator would be more salient stimuli than the food plate, and hence yield relatively higher accuracy within a given complexity level. Cleaner wrasse solved Piagetian stage 4 (object completely hidden) when the time interval between hiding and testing was short and the type of object had no effect. A possible explanation was that cleaners already appeared to habituate to the conspecific and predator model during stage 3 testing. Therefore, we conducted a second set of experiments with new cleaner wrasse starting with stage 4, and again using a food plate, a smaller conspecific, and a live lizardfish. Comparing the results of both experiments led to the conclusion that prior exposure to Piagetian stage 3 is mandatory for cleaner wrasse to solve Piagetian stage 4, and that the ecological relevance of stimuli should be reconsidered.

3.3. Material and Methods

3.3.1. Individuals, catching and housing

The first experiment was conducted in Moorea, French Polynesia, between September 30–October 27, 2021. The second was conducted on Lizard Island, Queensland, Australia, between April 24–26, 2023. In both experiments, we tested 18 females *Labroides dimidiatus*. In Moorea, cleaner wrasse were caught via SCUBA on reefs outside marine reserves between Opunohu Bay and Irihoriu Pass. In Australia, cleaner wrasse were caught via SCUBA on the reefs surrounding Lizard Island. Hand nets were used to lead cleaner wrasse into a barrier net (mesh size: 0.5cm), then transferred it to a zip-loc bag for transport. Water was changed after 45 minutes maximum to ensure individuals had enough oxygen. Upon arrival to the wet lab at each locale, cleaner wrasse were kept individually in aquaria with open-circuit water (and oxygen stones in 2023). Water flow was continuous, except during the experimental manipulations.

3.3.2. Habituation

In 2021, we acclimatized and habituated fish for 15 days before commencing the experiments. We ensured that all fish could feed (i.e., mashed prawn from the Plexiglas plate) and were adequately adapting to captivity. Once cleaner wrasse were habituated, we gradually introduced them to the different elements of the experimental design to ensure they were not stressed by these elements during the experiment. One female had to be excluded from the experiment in 2021 due to stress behaviours and a lack of participation. Cleaner wrasse used in 2023 had participated in previous, unrelated experiments, and thus, had to be habituated to fewer elements. Conspecifics that were used as subject stimuli were habituated to spend time in containers using multiple food rewards. This also helped conspecifics avoid repeated stress associated with removal and reintroduction into containers.

Prior to testing, we used the routine of the future trials to ensure minimal stress to subjects in the future and make sure that they could enter the corners to seek an object.

3.3.3. Experiments

Schema with detailed information on the setups are shown in Fig.16.
The general principle of our experiments relied on showing a whole object to the tested cleaner wrasse located in a central compartment of the aquarium, then hiding it behind a corner and observing whether the cleaner wrasse would seek it again (or willingly avoid it, depending on the condition) after it was freed. We applied this principle to 3 ecological contexts at 4 different levels of complexity. The ecological contexts were a) Foraging, b) Social, and c) Predatory, and the complexity levels were: i) object partially hidden (stage 3), ii) Completely hidden (stage 4), iii) 10-second delay (with a completely hidden object, stage 4), and iv) "A-not-B" error (stage 5). We organized the cleaner wrasse into three groups of six individuals to counterbalance the order in which we would test the ecological context. The tests were performed sequentially for the four complexity levels, with the idea that a cleaner wrasse would be tested at a higher complexity level only if it passed the lower complexity level.

We performed every trial with the following six steps: 1) placing the cleaner wrasse in the central compartment, 2) placing the two corners at the predefined location, 3) showing the object to the cleaner wrasse (ensuring its attention), 4) hiding (partially or completely) the object in the corner previously chosen (when necessary, we slightly moved the corner to allow the object to pass the entrance), 5) releasing the tested cleaner wrasse (by removing the central compartment), and 6) noting its choice.

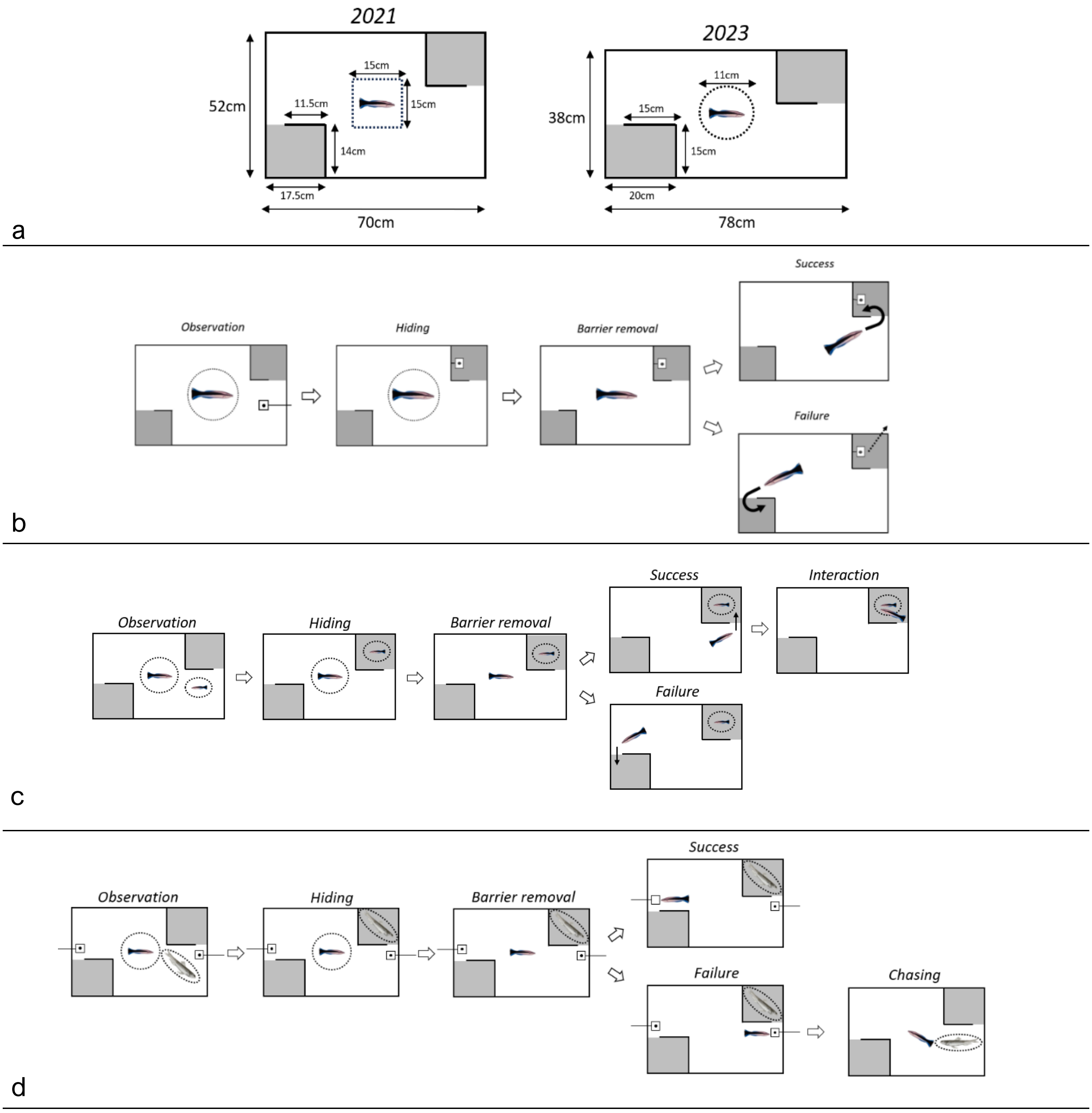


Fig. 16: Experimental design and material
a. Measurements in 2021 and 2023
The corners were sometimes moved during the manipulations when the entrance was not large enough for the object to enter. Those were put back in place before the release of the cleaner wrasse.
We repeated the manipulations in three contexts: **b. Foraging, c. Social, and d. Predatory**.
In all cases, the same phases were involved: Observation (the fish looks at the object, we get its attention), hiding (we move the object into the hiding corner), release (we free the cleaner wrasse from the central compartment), and choice of the cleaner wrasse. Note that cleaner wrasse were habituated to all the components of the design except the predators.
In the foraging context, we used a white Plexiglas plate of 10x3cm (2021) or grey plates of 4x4cm (2023) with a single black spot in the middle on which we placed a food item.
During the social context, we contained the conspecific in a pot of 11cm (diameter) x 8cm (height, 2021) or 10.5cm (diameter) x 10cm (height, 2023). In 2021, the pot had another compartment inside, 7cm (length) x 5.5cm (width) x 5cm(height), that could be opened or closed (to confine the conspecific in a restricted space when needed).
During the predatory context, we used a 3D-printed model of a lizard fish in 2021 that was 10cm long. In 2023, we had a living lizard fish of 11cm contained in a pot of 14cm (length) x 9cm (width) x 11cm (height). The plates with the rewards were 3cm x 3cm.

In the social context, we placed the smaller conspecific in the container (in 2021: 4.9cm, young adult, minimum size difference with tested cleaner wrasse: 0.4cm, maximum: 2.7cm; in 2023: 5.8cm, young adult, minimum size difference with tested cleaner wrasse: 0.4cm, maximum: 2.1cm) and allowed a maximum of 10 seconds of interaction between the tested cleaner wrasse and its conspecific.

In the predatory context, we started by hiding the predator. Then, we added a reward plate with one food reward in front of the entrance of each corner and made sure the cleaner wrasse saw both. Additionally, when the cleaner wrasse failed and picked up the food item near the predator model, we chased it with the model to mimic a consequence and avoid the motivation to drop (*i.e.,* cleaner wrasse getting accustomed to the predator).

In the partially hidden condition, half of the object was sticking out of the corner and thus was always visible to the cleaner wrasse. In the predatory context, the head of the 3D-printed lizardfish model was visible. For the 10-second delay condition, we waited 10 seconds after we hid the object before we freed the cleaner wrasse from the central compartment. The "A-not-B" error consisted of three sub-trials in a row for one trial. After hiding the object twice in a row in the same place, we changed the corner on the third sub-trial. That is when we noted the choice of the cleaner wrasse.

Cleaner wrasse succeeded in the foraging and social contexts by going into the corner that had the object. It was the opposite in the predator context, since they were supposed to go eat at the entrance of the corner where the predator was not hidden. In the foraging context, the reward was the food item on the hidden plate (the cleaner wrasse had nothing in case of failure). In the social context, we allowed a maximum of 10 seconds of "interaction" between the cleaner wrasse and its conspecific (no physical contact was possible as the smaller female was in the container). Lastly, in the predatory context, both reward plates contained a food item, and the predator model chased the cleaner wrasse when it ate near the predator's corner. We counted the choice as soon as the cleaner wrasse could see inside the corners (when the object was completely hidden) or when it got near the partially hidden object. When the cleaner wrasse did not choose a corner directly, we waited a maximum of 30 seconds before indicating "no choice".

In 2021, the partially hidden (stage 3) condition was tested with three trials daily, each day in a different context (*i.e.*, foraging, social, predatory). For the "completely hidden" condition (stage 4), we did five trials per context and added another five trials for the foraging condition during another day at the end. For the 10-second delay condition, we tested only the foraging context (as we observed the motivation had dropped in the other contexts) and did five daily trials for two days. Finally, we made five trials for the "A-not-B" error, also only in the foraging context. We divided our cleaner wrasse into two groups, each tested by a different experimenter (both female). We recorded every trial using a DJI camera placed on the experimenter's forehead.

For both experiments, we organized the cleaner wrasse into three groups of six individuals to counterbalance the order in which we tested them for the different contexts. The position of the two corners (up right, up left, bottom right, bottom left) and the one used to hide the object were counterbalanced and randomized between trials for all cleaner wrasse to avoid side preference to develop as much as possible.

3.3.4. Statistical analyses

Statistical analyses were done with R software v. 4.3.1 (R Core Team, 2022).

The data and R codes are accessible on Figshare via the following link: (https://figshare.com/projects/Object_Permanence_in_Cleaner_wrasse_DATA_CODE/22120).

Due to the sequential order of testing and the logic underlying it (i.e., if the cleaner wrasse succeeds on one step, it can move on to the next), we analyzed the two levels of complexity left (i.e., half and completely hidden) separately. We wanted to test the potential effect of the context (i.e., categorical explanatory variable with three levels: foraging, social, and predatory). Due to behavioral observations during our experiments, we decided to test a potential effect of time reflected by the trial number—not "day", since it is associated with "context" in a randomized way in our design. We used a binomial Generalized Linear Mixed Effect Model, with the result of the trial as the response variable (0 for failure or 1 for success). In the "partially hidden" complexity level (Piagetian stage 3), we could not incorporate the trial and the interaction within the context of the model (nor as a random factor) because of a complete separation problem in the social context. Therefore, we ran three distinct McNemar tests for each context, comparing the first and the third trials. In the "completely hidden" complexity (also 2023 data), we included the context and the trial as explanatory variables and the interaction term. We removed the interaction term and the "trial" variable from the model as nothing was significant about "trial". For 2021 data, the day was also added as a random factor to account for the second day of foraging trials that we added afterwards. We tested whether cleaner wrasse succeeded above chance with these final models, compared the context pairwise using the least-square means as post-hoc analyses, and adjusted the p-values for the multiple comparisons using the Holm correction.

To investigate a potential learning effect in the "10-second delay" condition, tested only in the foraging context, we first analyzed our data using a Generalized Linear Mixed Effect Model (GLMM), with the result of the trial as the response variable (0 for failure or 1 for success), the "day", "trial", and their interaction term as explanatory variables, and "fish ID" as a random factor. Because none of them was significant, we used the corresponding null GLMM, from which we extracted the test of comparison with 50% of the global proportion of success (test the hypothesis "intercept=0" on the logit scale).

Fits of all models were assessed visually using *the DHARMa* package (Hartig 2022) in R.

3.3.5. Ethical Note

For 2021, the French Polynesian authorities approved the study and the catching of cleaner fish (11626/VP/DRM). In 2023, the Queensland DAF ethical permits committee approved the fish and the experimentations under permit CA 2023/03/1708.

Being caught and kept in aquaria can be a stressful experience for fish, but they all adapted well to captivity and human presence. At the end of the studies, we released them at their capture site to minimize the environmental impact.

3.4. Results

3.4.1. Experiment 1

Piagetian stage 3: Object is “partially hidden”

The success rate of the cleaner wrasse subjects evolved differently through trials depending on the context (Type II Wald chisquare tests on binomial GLMER, chisq = 10.04, df = 2, p-val = 0.007). There was no difference between the first and the last trial for the foraging and the social context (McNemar’s Chi-squared with continuity correction, foraging context: McNemar’s chi-squared = 0, df = 1, p-val = 1 and social context: McNemar’s chi-squared = 0.5, df = 1, p-val = 0.48), and subject performance decreased between trials in the predatory context (McNemar’s Chi-squared with continuity correction, McNemar’s chi-squared = 5.14, df = 1, p-val = 0.02).

Cleaner wrasse succeeded above chance in every context (least-square means, foraging context: emmean = 94.80%, z.ratio = 4.26, p-val = 0.0001; social context: emmean = 97.57%, z.ratio = 4.26, p-val = 0.0001, and predatory context: emmean = 80.84%, z.ratio = 3.00, p-val = 0.003). In addition, the context significantly modulated the success of cleaner wrasse (Type II Wald chisquare tests on binomial GLMER, chisq = 9.39, df = 2, p-val = 0.009). Post-hoc pairwise comparisons showed that cleaner wrasse succeeded better in social than in predatory context (least-square means, social: predicted p of success = 97.57% and predatory: predicted p of success = 80.84%, z-ratio = -2.66, p-val = 0.021) and that they tended to succeed better in foraging compared to predatory context (least-square means, foraging: predicted p of success = 94.80% and predatory: predicted p of success = 80.84%, z-ratio = 2.19, p-val = 0.072), but there was no difference between the foraging and the social contexts (least-square means, foraging: predicted p of success = 94.80% and social: predicted p of success = 97.57%, z-ratio = -0.86, p-val = 0.668). Habituation drove these differences, or "unlearning effect", that we observed in the predatory context as cleaner wrasse succeeded equally in all contexts on the first trial (Foraging: 16/17, social: 17/17, predatory: 16/17).

Piagetian stage 4: Object is “completely hidden”

We had a very small number of trials where cleaner wrasse did not make a choice (i.e. did not enter a corner, foraging and social contexts: 0.02% of trials).

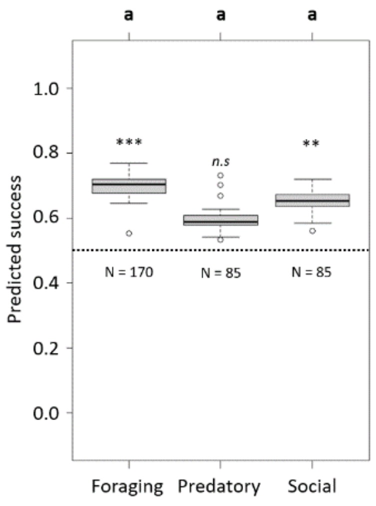


Figure 17: Representation of the post-hoc analyses
This shows the predicted values of success of cleaner wrasse in the three different contexts (i.e., foraging, predatory and social) in 2021 (Piagetian stage 4). The dotted line indicates the random threshold of 50%. The compact letter display indicates the difference between contexts.
Significance code: < 0.001 ***, < 0.01 **, < 0.05 *, > 0.05 *n.s.* N indicates the number of trials kept in the context.

Neither trial nor the interaction with the context were significant (Fig. 17, Type II Wald chisquare tests on GLMER, trial: chisq = 2.79, df=2, p-val = 0.594 and context:trial: chisq = 9.69, df=8, p-val = 0.287). After simplification, the final model found no effect of context on the probability of success (Type II Wald chisquare tests on GLMER, chisq = 2.47, df = 2, p-val = 0.291). Cleaner wrasse succeeded significantly above chance in the foraging context (least-square means, predicted probability of success, p = 69.39%, z-ratio = 4.59, p-val < 0.0001) and in the social context (least-square means, predicted probability of success, p = 67.38%, z-ratio = 2.84, p-val = 0.009) but not in the predatory context (least-square means, predicted probability of success, p = 58.52%, z-ratio = 1.42, p-val = 0.155).

Higher complexity levels

In the “10 seconds delay” stage 4 condition, cleaner wrasse did not succeed above chance (estimated proportion of success = 51.76%, Z-test = 0.46, p-val = 0.65).

For the “A-not-B error” testing (Piagetian stage 5), cleaner wrasse exhibited almost systemic side preference and insufficient success in the two initial sub-trials, preventing proper testing on the third sub-trial. Consequently, we did not further analyse the data, and concluded that cleaner wrasse failed the task.

3.4.2. Experiment 2

Cleaner wrasse did not make a choice in 24% trials for the foraging context and 37% of trials in social context, which were discarded to avoid biasing the results. We found no evidence for a learning component through trials, neither overall (Type II Wald chisquare tests, Chisq = 1.78, df = 2, p-val = 0.410) or considering the contexts separately (Type II Wald chisquare tests, Chisq = 4.70, df = 4, p-val = 0.319). The resulting simplified model yielded no significant differences between the contexts (*i.e.*, foraging, social, or predatory) (Fig. 18, Type II Wald chi-square tests, Chisq = 1.12, df = 2, p-val = 0.571) and the performances of cleaner wrasse were at chance level in all cases (Fig. 18, Least-square means, foraging: predicted probability of success, p = 56.09%, z.ratio = 0.74, p-val = 0.947; social: predicted probability of success, p = 59.05%, z.ratio = 1.00, p-val = 0.947; predatory: predicted probability of success, p = 48.10%, z.ratio = -0.27, p-val = 0.947).

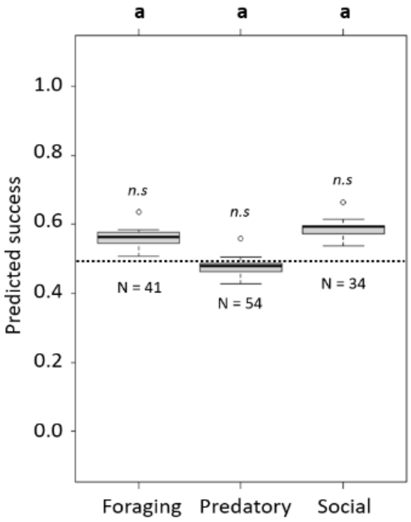


Figure 18: Representation of the post-hoc analyses
This shows the predicted values of success of cleaner wrasse in the three different contexts (*i.e.*, foraging, predatory and social) in 2023 (Piagetian stage 4). We removed trials where cleaner wrasse made no choice. The dotted line indicates the random threshold of 50%. The compact letter display indicates the difference between contexts. Significance code: > 0.05 *n.s.* N indicates the number of trials kept in the context.

3.5. Discussion & Conclusion

We studied the performance of cleaner wrasse (*Labroides dimidiatus*) in OP tasks, varying task complexity (i.e., the Piagetian stage) and the context (i.e., type of object). We report the best performance of a fish species in a Piagetian stage 4 task to date. We found no evidence for stage 5 (multiple visible displacements), and a delay between hiding and testing at level 4 also led to failure. Furthermore, cleaner wrasse needed trials on Piagetian stage 3 (partially hidden object) as training in order to succeed at stage 4. Finally, contrary to our predictions, the object type had no apparent effect on cleaner wrasse performance. Below, we discuss these findings in more detail.

The literature on tests for OP in fish is sparse, with an exclusive focus on immediate performance when an object associated with a food reward is completely hidden in one of two possible places (Piagetian stage 4). Cleaner wrasse and cichlids failed at this task (Aellen, Burkart, et al., 2022; Guadagno & Triki, 2024). Only guppies with a large telencephalon could reach around 60% success in an OP task (Triki et al., 2023). Hence, the present study shows a higher success rate in an OP task based on retrieving food (69%) than what has been reported for fish until now.

It is difficult to compare tests of OP performance in cleaner wrasse to those of other species due to differences in experimental methodologies (e.g. individual-based approach with criteria to reach, leading to higher success rates at the population level (e.g., Yao & Garcia-Pelegrin, 2024). However, when compared to species that were recently tested using similar methodologies, the performance of cleaner wrasse is notable. For example, at Piagetian stage 4, the European bison succeeded at a rate of 68%, giraffes at 72% (Caicoya et al., 2021), horses at 62% (Trösch et al., 2020), racoons at 85% (Colbourne, 2018), and hand-reared coal tits at 71% (Marhounová et al., 2017). Interestingly, cleaner wrasse also outperformed some endotherms in that stage. Indeed, forest buffaloes (Caicoya et al., 2021), great tits (Marhounová et al., 2017) and pigeons (before training, Zentall & Raley, 2018) failed to succeed in a Piagetian stage 4 task above chance.

Furthermore, our study provides the first evidence for a fish that delaying the trial after hiding the item leads to chance level. This contrasts with results in pigeons, where adding a delay significantly increases the success rate (Zentall & Raley, 2018). We are also the first to test a multiple visible displacement task (Piagetian stage 5, test for the "A-not-B" error) in a fish species.

Thus, there is currently no evidence that fishes possess higher levels of OP. This is an important result as higher levels of OP would also provide evidence for working memory, a key executive function (Diamond, 2013). Some researchers in the past have associated Piagetian stage 4 of OP with working memory (Guadagno & Triki, 2024; Triki et al., 2023), a type of short-term storage used for the manipulation and combination of past and present information in order to solve a problem in the present (e.g. Baddeley & Hitch, 2019; Cowan, 2022). In our opinion, one should carefully consider this direct association. The critical aspect of our parsimonious approach lies in the manipulative aspect of working memory. Indeed, in a Piagetian stage 4 task, there is not much information to manipulate or combine. One single object is used and hidden at a location. The individual could simply remember "*go there*", which does not require a combination of information.

On the contrary, in multiple visible displacements, or single and multiple invisible displacement(s) tasks (Piagetian stages 5 and 6, Piaget & Cook, 1952), the individual has to maintain different information linked to the different positions occupied by the object or the combination of cover (the material used to hide the object) and location (see Hoffmann et al., 2011; Yao & Garcia-Pelegrin, 2024 for detailed examples of testing stages 5 and 6). To assess working memory via a similar paradigm to a Piagetian stage 4, one could hide two different objects in two different locations where only one is the correct choice (i.e., only one would lead to a reward). To succeed in the task, the tested subject would have to remember and combine two pieces of information (*what* and *where*). This debate is essential as working memory is a clear predictor of other cognitive abilities in humans (Alloway & Alloway, 2009; Conway et al., 2003; Engle et al., 1999), including general intelligence (Conway et al., 2003; Kolata et al., 2005; Shipstead et al., 2014), and we have found no evidence for neither working memory (Bonin et al. in prep.) nor general intelligence (Aellen et al., 2022) in cleaner wrasse.

An important finding of our study was that cleaners apparently needed to be tested on Piagetian stage 3 to perform above chance at stage 4. Cleaner wrasse solved the Piagetian stage 3 task between 81% and 98% for the three contexts, and their performances dropped with increasing complexity levels, which we expected. While it needs to be acknowledged that the two sets of experiments (experiment 1: test stage 3 then stage 4; experiment 2: stage 4 only) were conducted at two different locations in two different years, we consider it unlikely that a fundamental process like OP would be present in one location but absent in the other within the same species.

Across our trials, cleaner wrasse exhibiting “no choice” were more frequent when subjects had not been tested in stage 3. Additionally, we trained cleaner wrasse to enter the hiding corners before testing them. Clearly it was insufficient and suggests that "looking for" an object is probably not part of the cleaner wrasse's natural behaviour. Nonetheless, it can become part of the behavioural repertoire when cleaner wrasse are repeatedly exposed to the logic. We found similar importance for training when we compared the results of our study on transitive inference (Bonin & Bshary, 2023) to those of our colleagues (Hotta et al., 2020).

Consequently, we hypothesise that stage 3 testing is mandatory for cleaner wrasse to understand what the task was generally about. Thus, cleaner wrasse are similar to some other species that also need training. In the case of guppies selected for large telencephalons, the number of trials apparently positively impacted the performance (Triki et al., 2023) and pigeons were tested with a minimum of 30 trials with a close-to-maximal success rate reached after 120 trials (Zentall & Raley, 2018). For higher Piagetian stages, dogs show greater success rates in invisible displacement tasks if they underwent visible displacement tasks before (Gagnon & Doré, 1992) and bottlenose dolphins may need to be introduced to objects moving inside other objects (Jaakkola et al., 2010; Johnson et al., 2015; Pérez-Manrique & Gomila, 2018). In contrast, a stage 3 before a stage 4 task did not affect the performances of Eurasian jays (Zucca et al., 2007), and captive horned puffins even failed in a stage 4 task after succeeding in a stage 3 task (Huffeldt, 2020).

In our more complicated tests, adding the 10-second delay on stage 4 was too challenging for our cleaner wrasse. Previous work suggests that pigeons, which did not exhibit high-level OP (Reid, 1996), could benefit from a delay that seemed to help them make better choices (Zentall & Raley, 2018). The authors propose that the delay may help pigeons go against impulsive choice. However, in our case, the delay (10 seconds vs. 5 seconds for Zentall & Raley, 2018) led to worse choices, supporting past findings about a negative link between delay and performance (Thorndike, 1911). Lastly, a proper test of the "A-not-B error" at stage 5 requires solid performance at stage 4, as individuals need to succeed in the two sub-trials to be tested on the third one. Therefore, the former is not feasible with cleaner wrasse.

Clearly, more ectotherm species should be tested for OP to know whether a general limit is reached at stage 4, which could highlight a shared mechanism at higher stages of OP that would rely on brain tissue. However, the link to ecology must be also considered. In the past, some studies have suggested that ecological needs, such as food hiding and retrieving in some bird species (Marhounová et al., 2017; Pollok et al., 2000), could directly define OP abilities in a species. Based on such an ecological approach to cognition (Hutchins, 2010; Kamil, 1998; Shettleworth, 2009), as well as previous results on our study species (Aellen et al., 2022), we had expected that cleaners would perform poorly when the object is food but better when the object is a smaller conspecific or a potential predator. These predictions were not confirmed; if anything, the cleaner wrasse's performance was the highest in the foraging task. We provide two potential explanations. First, the foraging context was the most straightforward experimental design with a simple reward plate to move around, while the two other contexts needed more material to move around. Consequently, there was more movement, and thus potentially more confusion and disturbance for the cleaner wrasse. Second, cleaner wrasse may have learned during stage 3 that they could not directly interact with the smaller conspecific and that the lizardfish was just a model, and therefore the salience of these two stimuli was already diminished when we moved on to stage 4. In contrast, food is always rewarding. It was to test for this second potential explanation that we repeated the experiment with new naïve cleaner wrasse, starting directly at level 4 to avoid the habituation problem. We also acknowledge that a client swimming out of sight does not necessarily go out of reach in a complex habitat like a coral reef, and consequently, a food plate may have been a salient stimulus of ecological relevance as well. In line with this thought, we note that cleaner wrasse almost always inspected a hiding place in the first series of experiments, indicating that their motivation to participate was generally high across the three stimuli.

We experimentally demonstrated that cleaner wrasse were able to reach Piagetian's stage 4, which proved to be the upper limit—or even unattainable—for some other species with larger brains (Caicoya et al., 2021; Colbourne, 2018; Dumas & Wilkie, 1995; Marhounová et al., 2017). That said, the observed variation in endotherms could be due to ecological needs and ecological relevance of the task. More studies on ectotherm vertebrates are crucial to establish whether ectotherms have or have not evolved higher levels of OP and high precision in OP tasks. If future studies yield similar results to ours, then ectotherm vertebrates may indeed face constraints on OP abilities. To better facilitate cross-species comparisons, one could consider using a cup-like design, where food is hidden under a cup, facilitating further manipulations with the hiding object. Also, one could further test the evolution of OP ability in cleaner wrasse by comparing their success through much longer testing sessions.

3.6. Acknowledgements

The Swiss National Science Foundation for fundings, the staff from the Richard B. Gump Research station in Mo’orea, French Polynesia and from the Lizard Island Research Station, Australia. B. Hatzinger, R. Snyder, C. Lebet and D. Berger for their precious help in catching a lizard fish, L. Pessina and M. Amann for catching the cleaner wrasse, M. Amann for her help caring for and habituate cleaner wrasse, R. Slobodeanu for his statistical expertise, and C.S. Clements for his precious help in improving the writing.

General Discussion

The primary objective of this thesis was to investigate the mechanistic bases the cognitive toolkit of cleaner wrasse, with a focus on understanding *how* they solve problems, rather than merely assessing "*what*" problems they can solve (i.e., cognitive mechanisms *vs.* cognitive performance). We also incorporated an ecological component into our investigations, hypothesizing that the presence or absence of cognitive abilities in cleaner wrasse could depend on the ecological relevance of the tasks or of the cognitive ability itself. The goal was to determine whether cognitive mechanisms could shed light on what could a relatively big brain do compared to the relatively smaller brain of a species like the cleaner wrasse.

This general discussion will: (i) provide a summary of the main findings and their broader implications; (ii) propose potential explanations and outline avenues for future investigations; (iii) offer insights into how cleaner wrasse navigate and overcome ecological challenges; and (iv) present a line of reasoning that explore the relationship between working memory, general intelligence, and relative brain size. I will finish with a general conclusion of the current work.

D.1. Summary of findings and implications

This work is divided into three main chapters: (1) *Working Memory*, (2) *Transitive Inference*, and (3) *Object Permanence*. Below is a concise summary of the findings and their implications.

D.1.1. Chapter 1 – Working Memory (WM)

In this chapter, we observed consistent negative results across four experimental paradigms. We presented and detailed the challenges we faced during experiments, which prompted several adjustments to our methodology. Although we were not able to test for critical aspects of the definition of WM, consistent negative results obtained in the early stages of testing allowed us to confidently rule out the presence of WM in cleaner wrasse.

Two of our four paradigms were based on visual information, both of which yielded weak and inconsistent positive results. Cleaner wrasse exhibited complete failure in the two remaining paradigms, which were based on spatial tasks. Our findings support the idea that spatial tasks are more complex to solve and do not rely on some forms of spatial familiarity, in contrast to visual-based tasks, which can also be referred to as "recognition tasks" (Bachevalier & Mishkin, 1986; Zola et al., 2000). However, direct comparisons between the two designs are limited, as the spatial tasks involved four locations, while visual tasks used only two plates.

Cleaner wrasse's overall failure in these paradigms contradicts our initial hypothesis, which predicted that their ecological demands would require some level of WM. However, repeated negative results led us to consider alternative strategies that might explain how cleaner wrasse address their ecological challenges. Furthermore, our findings also led us to consider whether WM, as a core executive function, could help explain differences in brain size between ectothermic and endothermic vertebrates.

D.1.2. Chapter 2 – Transitive inference (TI)

In this chapter, we found that cleaner wrasse could not solve a non-social TI task. This is consistent with cleaner wrasse ecology, where neither the social nor the foraging challenges require TI, and supports the ecological approach to cognition (Hutchins, 2010; Robertson, 1972; Shettleworth, 2009). TI has been observed in some small brain species, including invertebrates (Tibbetts et al., 2019), and thus our findings reinforce the notion that TI is more closely tied to ecological relevance than to brain size.

A previous study found positive evidence for TI in cleaner wrasse (Hotta et al., 2020), but our negative results call for alternative potential explanations. We suggest that cleaner wrasse may not possess a natural ability for TI, but extensive training and additional information, as given by Hotta et al. (2020) (e.g., non-adjacent pairs training) allow them to solve the task. This would suggest that cleaner wrasse can compensate for a lack of inherent cognitive ability when given numerous opportunities to learn. Furthermore, in this chapter, we noted that "(…) the establishment of the list is already a cognitively demanding step that necessitates extended training for fish compared to some other species". This notion is of particular interest given the lack of WM we observed in clear wrasse. Establishing such a list (e.g., A > B > C > D > E), which is essential to show TI ability, likely depends on WM as it requires combining knowledge about several pairs to infer relationships between new pairs (e.g., B > D) (Libben & Titone, 2008).

The fact that training helps overcome the lack of ecological relevance underscores the capacity of cleaner wrasse to adapt through learning, effectively gaining new abilities through repeated exposure. Furthermore, as TI has been found in species with similar or smaller brains (e.g., Tibbetts et al., 2019; White & Gowan, 2013), our findings suggest a balance between ecological relevance and brain size where smaller-brained species can solve tasks if they are highly relevant. Conversely, whether large-brained species can solve tasks that are irrelevant to them deserves further investigation. That said, previous studies have found that cleaner wrasse and primates perform similarly in a delay-of-gratification paradigm (Aellen et al., 2021), which is highly relevant for cleaner wrasse (e.g., Bshary, 2002; Bshary & Würth, 2001). This highlights how ecological relevance can compensate for a lack of brain tissue or complexity.

D.1.3. Chapter 3 – Object Permanence (OP)

This chapter examined how cleaner wrasse perceive and respond when objects of interest are put out of sight. Cleaner wrasse failed this task in the past (Aellen et al., 2022), so we refined the methodological approach and tested whether they could solve OP tasks depending on the ecological relevance of the context. As in Chapter 2, our findings revealed that cleaner wrasse required extensive training to understand the task, indicating that they may not possess an innate ability to "look for" objects. We again argue that this seems consistent with their ecological demands. We also highlight similar observations documented in other species (e.g., Triki et al., 2023; Zentall & Raley, 2018), and underly that cleaner wrasse's performance was comparable to, or even superior to, that of some endothermic vertebrates (e.g. Huffeldt, 2020; Zentall & Raley, 2018). On another hand, as the performance of cleaner wrasse was not stable enough, we could not test properly for the "A-not-B" error that would have accounted for a stage 5.

Another finding is that the ecological approach did not influence the results. This came as a surprise, especially concerning the predatory context in which cleaner wrasse did not perform above chance. We used lizard fish as the model predator in our experiments because previous observations suggested that they elicit strong aversive reactions among cleaner wrasse (Redouan Bshary, pers. obs.). However, this was not the case in our experiments—neither with a 3D-printed model nor with a living lizard fish. We cannot rule out the possibility that the cleaner wrasse also understood that the lizardfish was in the container and therefore posed no threat, despite the limited number of trials intended to prevent such habituation. On the other hand, it might also indicate that cleaner wrasse does not naturally focus on whether a predator is hidden, which may be due to their unique relationship with predators. Unlike other species, cleaner wrasse can also clean their predators (Grutter, 2004), potentially complicating the dynamics of predator avoidance.

While experimental biases complicate some of our conclusions, our findings provide important methodological insights and demonstrate the first positive evidence of OP in cleaner wrasse. Our results, along with existing literature on OP in other species (e.g., Caicoya et al., 2021; Zentall & Raley, 2018), suggest that the presence of OP is not solely determined by brain size.

D.1.4. Interconnectedness of the chapters

The three chapters of this thesis are interconnected through their examination of cognitive mechanisms that facilitate success across various tasks. WM, a core executive function (Diamond, 2013, 2020), is linked to numerous cognitive abilities in both human and non-human species (Ackerman et al., 2002; Alloway & Alloway, 2009; Conway et al., 2003; Engle, 2010; Kolata et al., 2005; Manrique & Walker, 2017; Shipstead et al., 2014; Unsworth et al., 2015). TI, by contrast, is a reasoning ability often used to resolve social challenges by enabling individuals to infer their social rank within a hierarchy without engaging in direct comparisons between all pairs of individuals. The connection between WM and TI stems from the necessity to update previously learned information about certain pairs with new knowledge about others (Libben & Titone, 2008). Indeed, as previously mentioned, individuals must compile a list of all the different objects based on the relative values observed during the presentation of pairs. To achieve this, the knowledge of each learned pair must be processed in such a way that each object can be 1) extracted from the pairs it was associated with during the learning phase, and 2) assigned a unique relative value, enabling the construction of a transitive hierarchy among all objects.

OP, meanwhile, could help address various ecological challenges, such as foraging, mating, and predator evasion. We propose that the connection between OP and WM emerges only after stage 4. This is because earlier stages do not involve the manipulation of information. In stage 4, the individual is required merely to return to the location where the object was last seen before its disappearance. However, beginning with stage 5, the individual must update its knowledge and counteract the “typical reaction” assessed in the “A-not-B” error task (Piaget, 1954; Piaget & Cook, 1952). Accordingly, our findings showed consistent results between WM and OP studies, with both yielding negative outcomes, as the cleaner wrasse could not be tested beyond stage 4.

D.1.5. Potential experimental improvements in future experiments

Future studies on WM in non-human species should consider the types of information stored and combined. Wong et al. (2008) found that the number of categories did not impact young adults' memory capacity, but the presence of specific stimuli (here, human faces) did affect memory load capacity. This suggests that controlling and testing different types of stimuli could improve our understanding of how information is processed and stored in memory. Similarly, Cowan (2017) discussed the development of WM capacity in human infants and found that the amount of information about the object stored varies with age. This supposes that at early stages, the more similar the objects are, the less they can be remembered distinctively. To my knowledge, such distinctions or analyses remain absent in non-human animal studies.

Additionally, future work should control for the impact of learning on memory capacity (Cowan, 2017). For example, when cleaner wrasse are trained to associate a food item with a black spot, they may create a new single object, or "chunk," that they subsequently process as a single piece of information (e.g., “food item”; Cowan, 2001; Cowan et al., 2024). Researchers should carefully consider how learning may alter necessary information storage and design tasks accordingly.

This approach would also necessitate modifications to the OP paradigm. For example, hiding two different objects (from the same or different categories) in two different locations would require subjects to bind the "*what*" with the "*where*," providing further insight into information processing. Although language is thought to be essential for linking "*what*" and "*where*" (Washburn et al., 2007), which would obviously cause negative results with fish, cleaner wrasse may have to remember "*which client is where*" in the wild, making this investigation particularly informative.

Ultimately, we chose to run our experiments with a restricted number of trials because we expected that learning effects would bias our results, especially in contexts that incorporated an ecological component. Indeed, we would argue that training should naturally occur in the wild if the paradigm is ecologically relevant. However, we acknowledge that experimental tasks in our study were likely more challenging for cleaner wrasse than those used for other species in past studies (e.g., Zentall & Raley, 2018). Therefore, it would be valuable to compare cleaner wrasse performance to other species when subjected to a comparable number of trials.

D.2. A lack of attentional control as a potential explanation?

D.2.1. Why is attentional control a potential explanation?

The process of WM has sometimes been divided into three distinct subsystems, each with a unique function and, consequently, corresponding testing methods (Beaman, 2010). One of these subsystems, known as the attentional subprocess, may be absent or underdeveloped in cleaner wrasse and could explain their repeated failure across several WM paradigms. Numerous studies demonstrate that attention is fundamental to WM (e.g., discussed in Cowan et al., 2024). For instance, D’esposito and Postle (2015) noted that “information is encoded into working memory by allocating attention to internal representations”. Similarly, Van Ede and Nobre (2023) emphasized “the importance of selective attention in selecting and prioritizing the relevant contents within working memory”. Conway et al. (2003) argued that “the basic processes that contribute to WM span tasks suggest that they critically tap an executive attention-control process,” while Kane et al. (2001) highlighted that “it is the individual differences in the controlled-attention component of WM that are responsible for the correlations among WM span and complex cognition measures”. However, the concept of attention is broad and thus requires further precision.

Van Ede et Nobre (2023) defined two attention processes: (1) outside-in, equivalent to Baddeley’s (1986) concept of perceptual selective attention, which selects external sensory information for internal processing in WM, and (2) inside-out, also known as internal selective attention (e.g., Ester & Nouri, 2023), which selects WM content to guide behavior. Several experimenters have observed the importance of ensuring that cleaner wrasse pay attention to the correct object during data collection, so that they can have a chance of solving the task. In the context of Van Ede et Nobre's (2023) framework, this means that experimenters can control the outside-in attention. Previous work has demonstrated that experimentally inducing outside-in attention increases performance in WM tasks among humans (Kong & Fougnie, 2019), and outside-in attention is also found in ectotherm vertebrates (review in Krauzlis et al., 2018). This suggests that outside-in attention may be controlled or influenced externally by experimenters (or more broadly, the environment) but could be limited in its conscious control. Also, the ability to direct inside-out attention may be lacking in cleaner wrasse. In humans and other vertebrates, attentional control over internal information is essential for solving complex problems (i.e., inside-out selective attention, Kane et al., 2001; Van Ede & Nobre, 2023), but I could not find any study that investigated it in ectotherms.

Cleaner wrasse may instead rely on environmental cues to unconsciously trigger the retrieval of relevant information, rather than consciously directing their attention to specific stored knowledge. For example, when a person loses their keys, they may consciously trace back through past events to locate them (example given by H. M. Manrique, as relevant example to explain when WM is used in everyday life, private conversation). Cleaner wrasse, on the other hand, would not actively recall where they left their keys; instead, an environmental trigger, such as seeing a specific stimulus, might evoke a memory about where they left them. This compelling example highlights a key distinction that may exist between small and large brains: the capacity to consciously act on knowledge. Cleaner wrasse and other small-brained species might rely on mechanisms such as familiarity—responding to a stimulus without explicit recollection of the information associated with it (see Yonelinas, 2002, for the distinction between familiarity and recollection).

As an aside, I think that a similar phenomenon occurs with outside-in selective attention. While cleaner wrasse may be able to maintain attention on an object, if their focus is disrupted—whether by natural factors (e.g., a time delay) or artificial ones (e.g., an external disturbance)—they seem unable to consciously refocus on the object. Using Diamond's (2013) words, I would say that cleaner wrasse likely lack inhibitory control over their attention. Support for this idea can be found in Aellen et al. (2021) who found that allowing the fish to self-distract from the reward did not improve their performance, in contrast to findings from primates and infants (Evans & Beran, 2007; Raghunathan et al., 2023). Thus, in cleaner wrasse, self-distraction may act as a disruption that breaks their focus of attention. While one might expect their performance to worsen with self-distraction, it's difficult to determine how the fish actually perceived the distracting objects. For instance, the authors noted that "individuals did not manipulate the distractor objects", which is surprising given that the distractors could have triggered the cleaner wrasse's interest. Indeed, the distractors included a surgeon fish laminated picture (a client fish species), a cleaner wrasse laminated picture (a potential competitor), and a plate resembling a foraging plate.

Another key finding supporting the hypothesis that small-brained species may differ due to a lack of attentional control is the strong correlation between WM and fluid intelligence, both of which are tied to the need for attentional control (Engle, 2018; Engle et al., 1999). Fluid intelligence (Gf), as defined by Cattell (1941, 1943), is the ability to adapt and solve new problems using logic. While Cattell's theory may not directly apply to non-human animals, its conceptual links to theories that are applicable to other species, such as Spearman's (1904) theory of general intelligence, are of interest. Spearman's (1904) theory posits that all cognitive abilities are positively correlated, resulting in a single general correlation coefficient known as the Spearman *g* factor, which estimates "general intelligence". Past studies found that WM is the main predictor for *g* (Engle et al., 1999; Takeuchi et al., 2018), even if some call for more tempered conclusions (review in Conway et al., 2003; Frischkorn et al., 2019). Attentional control is key to WM, WM is key to *g* (or fluid intelligence), and *g* has never been documented in an ectotherm vertebrate (e.g., in cleaner wrasse in Aellen et al., 2022). Hence, many findings support this new perspective on the evolution of larger brains.

D.2.2. Potentially challenging previous findings

While the hypothesis that cleaner wrasse lack attentional control aligns well with multiple other findings, some prior studies on executive functions seem to challenge this notion. For example, inhibition is another executive function (Diamond, 2013) that can alter attentional control. A previous study found that cleaner wrasse showed inhibition abilities in a detour task paradigm (Aellen et al., 2022). In this context, cleaner wrasse suppress the urge to approach the goal (the food reward) in order to navigate around an obstacle—a transparent cylinder that blocks access to the goal. This seems to require the ability to shift attention away from the goal to the obstacle. Researchers also prevented the development of routine behaviors by sometimes placing a plate in front of the obstacle. In another study, researchers found that cleaner wrasse were able to resist the urge to immediately eat a food reward in order to have more food later (Aellen et al., 2021), suggesting that cleaner wrasse can decide to focus on future outcomes as well.

However, a closer look at the nuances of inhibitory control can reconcile these findings with my hypothesis involving attentional control. Diamond (2013) distinguishes between self-control (behavioral inhibition) and interference control (cognitive inhibition and selective attention). A key aspect for my hypothesis is that Diamond categorizes the inhibitory component of the delay gratification paradigm as self-control, which relates to behavioral inhibition, rather than interference control, which is associated with selective attention. The two can be moderately correlated but remain separate mechanisms (Diamond, 2013; Wolff et al., 2016). This distinction allows for the possibility that cleaner wrasse possess self-control without interference control. Their success in behavioral inhibition tasks (e.g., detour tasks, delay-of-gratification; Aellen et al. 2021, 2022) may not depend on their ability to selectively control attention, which is the key mechanism behind interference control. It is important to note that the “A-not-B” error task also assesses a form of self-control, as it involves resisting the impulse associated with the “typical reaction” described by Piaget and Cook (1952). However, our inability to adequately test for this limits my conclusions regarding self-control in our study.

D.2.3. How to test my hypothesis?

To test the hypothesis that cleaner wrasse lack attentional control, several established paradigms from human and other animals cognitive research could be adapted. The Stroop task, retro-cues task, and antisaccade task have increased our understanding of how to prime, modulate, and use selective attention to increase success on WM-related tasks (discussion and review in van Ede et Nobre 2023; Engle 2018; Matzel et Kolata 2009).

Stroop Task Adaptation: The original Stroop task (Stroop, 1935) measures cognitive control by having participants name the color in which a word is written while ignoring the word itself. The interference between the two mechanisms (i.e., ignoring the word and naming the color) is now called the Stroop interference and is taken as a measure of cognitive control (e.g., Wolff et al., 2016). While verbal tasks are unsuitable for cleaner wrasse, an adapted version could involve visual or symbolic cues. Beran et al. (2007) adapted this task for chimpanzees using symbols, and cleaner wrasse have been shown to discriminate symbols in a match-to-sample paradigm (Aellen et al., 2022). The auditory equivalent of the Stroop task, known as the Simon task (Simon & Rudell, 1967), relies on spatial control and has been successfully applied to pigeons (Mueller, 2023; Urcuioli et al., 2005) after being adapted for visual cues. This approach offers a promising avenue for research on ectothermic species, such as cleaner wrasse. For a review of the methodological variations of the Stroop task, see MacLeod (1991).

Antisaccade Task: The antisaccade task (Hallett, 1978; Kane et al., 2001) is a non-verbal measure of attentional control. For example, in Kane et al. (2001), participants had to look at the center of a screen that showed two boxes. Then, one of the boxes flickers unpredictably. Then, there are two conditions: (1) in the prosaccade condition, the flickering indicates which box will reveal a cue to remember; (2) in the antisaccade condition, the participant must look at the box that did not flicker to obtain the cue. Thus, in the antisaccade condition, participants must control their attention to keep it on the non-flickering box. This type of task could be adapted for cleaner wrasse.

Pre-Cue/Retro-Cue Task: This paradigm artificially controls the attention of individuals using a pre-cue/retro-cue in a task (Griffin & Nobre, 2003; Sperling, 1960). In this task, individuals will be given a cue before (pre-cue) or after (retro-cue) a stimuli presentation that should guide their attention toward the relevant part of that stimulus, held in WM (Griffin & Nobre, 2003). Consequently, testing this with cleaner wrasse could help determine whether, with help to control the focus of their attention, they could solve a WM paradigm.

This list is not exhaustive, and while there is debate over the accuracy of these tasks (Draheim et al., 2021), they offer a starting point for exploring how attention functions in non-human species, particularly those with smaller brains. Such investigations could improve our understanding of whether control of attention is the key to the absence of some cognitive mechanisms in these species.

D.3. How do cleaner wrasse face ecological challenges?

Having proposed a hypothesis that could help explain the absence of WM (and hence, cognitive abilities that would rely on it) in cleaner wrasse, I now need to explore how these fish may compensate for this deficit to thrive in their environment. According to van Ede et Nobre (2023), WM should serve adaptability in a dynamic, changing environment. It is also strongly future-directed. Hence, what could be cleaner wrasse's alternative strategies?

My intellectual approach is guided by the consistent effect of training/learning highlighted in our results. In the wild, cleaner wrasse engage in around 2000 cleaning interactions per day (Grutter, 1996) which provides ample opportunities to learn through trial and error. In addition, associative learning is important in this species. Experimenters exploit associative learning during lab experiments for habituation and training, and cleaner wrasse use it in the wild in their foraging behaviour (e.g., Grutter & Bshary, 2003a). Their numerous interactions, combined with associative learning, could allow cleaner wrasse to compensate for the lack of complex cognitive mechanisms. I could easily imagine a balance between the cognitive toolkit, the number of learning opportunities, and the ecological needs where a species facing challenges with numerous learning opportunities does not require complex cognitive mechanisms. On the other hand, a species that has limited opportunities to learn will need a more complex cognitive mechanism to face the challenge.

So how could a potential balance between the cognition, the opportunities and the needs be linked to the lack of attentional control? More training/opportunities to learn means more possibilities to randomly change the object of attention, and if it brings greater success, the cleaner wrasse just has to repeat the pattern. Thanks to associative learning, cleaner wrasse could create foraging routines, such as "start at the tail, move to the back, then to the belly, and finish with the mouth". Once established, this type of repetitive strategy requires minimal cognitive effort while maximizing foraging efficiency. Such a behavioral pattern repetition strategy is a "cognitively cheap" way to handle daily challenges and could explain why cleaner wrasse perform well in the wild despite their apparent lack of complex cognitive mechanisms.

Moreover, there may be an element of inter-generational transmission in cleaner wrasse behavior. Juveniles frequently occupy the same cleaning stations as adults (pers. obs.), potentially learning routines and behavioral patterns through observation and mimicry, which could also be tested in lab experiments or observed in the wild. Studies have shown that juvenile cleaner wrasse can learn the consequences of cheating by observing adult behavior (Truskanov et al., 2020). This opens up avenues for further exploration of whether ecological knowledge and behavioral strategies could be passed down across generations in cleaner wrasse populations.

Lastly, we previously discussed that cleaner wrasse could also rely on familiarity to react appropriately to a stimulus. This again could be due to earlier learning processes (Yonelinas, 2002).

In summary, cleaner wrasse may use simple associative learning mechanisms that lead to a strategy of behavioral pattern repetition, familiarity, and possible inter-generational knowledge transfer (which is a complex cognitive mechanism in itself) to compensate for a lack of cognitive abilities such as WM. This strategy is sufficient for a species living in an environment with abundant learning opportunities but may be inadequate for species that face more unpredictable or novel challenges.

D.4. WM, *g*, and the relative brain size curve

The link between WM and *g* lies in the role of WM in producing goal-directed behaviors and adapting to environmental changes by integrating past knowledge with present information (e.g., Conway et al., 2003; Cowan, 2022; Engle et al., 1999; Frischkorn et al., 2019). Moreover, studies investigating the correlation between brain size and *g*, particularly in the context of sex differences in humans, have reported significant correlations between the two (Nyborg, 2005; van der Linden et al., 2017). The correlation coefficient for this relationship in humans typically ranges from 0.2 to 0.4 (Bouchard, 2014).

In non-human primates, absolute brain size has been found to be a better predictor of performance across a battery of cognitive tests (Deaner et al., 2007). Similarly, in rodents, there is substantial evidence supporting the presence of a general intelligence factor, which also correlates with brain size, particularly in rats (Anderson, 1993; Burkart et al., 2017). Further, some studies have examined the correlation between the size of specific brain regions and WM abilities, as highlighted in a neuroimaging meta-analysis by Yuan and Raz (2014). For a comprehensive review of *g* in non-human animals, see Burkart et al. (2017).

These findings lead to an intriguing question: what is the significance of the evolution of *g* and WM in relation to the average tenfold difference in brain size between ectothermic and endothermic vertebrates? While this remains an open question, the repeated correlations found between *g* and brain size, as well as between WM and *g*, suggest that these cognitive processes may require a large quantity of neuronal tissue. It is important to note that the evolution of general intelligence is still a subject of ongoing research (discussed in Burkart et al., 2017; Ramus, 2017 offers a contrasting view). Additionally, given that WM and other executive functions are not entirely independent (as discussed by Diamond, 2013), this question may also be relevant to other executive functions.

The fact that researchers have yet to precisely localize WM within the brain (Brzezicka et al., 2019; Christophel et al., 2017; Collette & Van der Linden, 2002; Salehinejad et al., 2021) highlights the vast potential for future research, from behavioral to neurophysiological fields.

General conclusion

This thesis offers valuable insights into the cognitive toolkit of cleaner wrasse, enhances our understanding of the relationship between a key executive function and other cognitive abilities, provides additional data to study cognition from an ecological approach, and proposes a hypothesis to explain an evolutionary gap between endotherm and ectotherm vertebrates. Furthermore, this research opens up new avenues for future investigations, including:

(i) Control of attention and alternative strategies
This research direction involves further exploring cleaner wrasse's attentional control abilities using classic attention-related tasks and revisiting their WM with attention manipulation techniques. Moreover, researchers could investigate the potential use of alternative strategies, such as associative learning, familiarity, routine behaviors, and social transmission, as ways cleaner wrasse may compensate for the absence of complex cognitive processes in overcoming ecological challenges.

(ii) Comparisons with other ectotherms
Another promising direction is to extend these studies to other ectotherm vertebrates. Similar experiments could be conducted with species such as lizards, which are just beginning to be studied for their cognitive abilities (Font, 2019; Noble et al., 2014), and other species of fish and amphibians. This would help determine whether the patterns observed in cleaner wrasse are representative of broader taxonomic trends among ectotherms.

This thesis thus lays the groundwork for further exploration of cognitive abilities across species, emphasizing the need to consider both evolutionary and ecological factors in understanding cognition.

Acknowledgements

I want to start by thanking Redouan Bshary for his trust, for giving me this opportunity, and for his benevolence throughout the challenges I faced during this PhD.

To Héctor M. Manrique, for being present, helpful and caring.

Werner Müller for his precious help in all the excursions and numerous coffees, discussions and laughs we shared.

Klaus Zuberbüler for the interesting conversations around methodologies.

Emily Genty and Sergio Rasmann for leading the doctoral program.

Radu Slobodeanu for his expertise in statistical analyses

Romain Bessire for multimedia-related advices and expertise.

To the sitel members who helped me numerous times save my computer that always threatened me to die at critical times.

To the secretariat: Simon Rowell, Roberta Ventura, Natacha Schneiter and Joanne Ren, thanks for your help and presence.

The “Bureau presse et promotion”, especially Nando Luginbühl, Mario Cafiso, Julie Mégevand, Dilyana Petrova, as well as Gaëlle Liechti for organising MT180 and creating an incredible friendly ambience and allowing access to multiple wonderful courses and coaching. To my fellows MT180 competitors who brought a lot of happiness at the end of this route: Felix Delerm, Amanda Jousset, Astrid Pilottin, Christophe Hurter, Maroussia Nicolet-dit-Felix, Sophie Guermann, and Valérie Rufenacht. To Séverine Marquand-Koller who transformed our stage presence and was very caring and empathetic.

To my other experts during the “Faunistique des vertébrés” excursions: Valéry Uldry and Ludovic Genoud, Robin Arnoux, Maxime Chèvre, Christophe Jaberg. I really enjoyed discovering all these species with you.

To my fellow PhD colleagues with whom I shared multiple years, including the very beginning: Yasmin Emery, aka “meuf” and Thierry Kuhn, aka “gerdy”.

To my younger colleagues: Marine Amann and Letizia Pessina, the fish girls, you are both incredible. To the monkey girls: Maria Granell-Ruiz and Josefien Tankink, you’re amazing.

To the master students I shared time and projects with: Sophie Duc, Richard Bonvini, Leda Ramelli, and Edson de Sousa, thanks for all the shared moments.

To the staff of both research stations that hosted me: Lizard Island Research Station in Queensland, Australia, and Richard B. Gump in Mo’orea, French Polynesia.

To all my friends wo supported me through this process: Elodie and Erwan, Adeline, Méli, Aurélye, Camille, Patrick, Alida, Romain, Nishan, Loic, Luni-Li, Caroline, Aïcha and Mattieu, Caro and Yan, Dominique.

To my family: My mom, Odile, who supported me, from finding the master program at unine to helping me and encouraging me in my hardest moments; my brother, Manu, who sent me funny memes to make time funnier and heard my complaints and Solene who was there for him when he was there for me.

To my cats, Yoda-Mr Monkey boy and TiBouli-Mr Puffle boy who always brought a smile on my face.

To my man, my partner, Cody, I would probably not have made it without you. Thanks for your love, support, and help.

To all the cleaner wrasse, my little swimmers, thanks for all the things you taught me. I wish I could have brought you all back in the ocean.

References

R.1. General Introduction

Abbas, F., O'Neill Rothenberg, D., Zhou, Y., Ke, Y., & Wang, H.-C. (2022). Volatile organic compounds as mediators of plant communication and adaptation to climate change. *Physiologia Plantarum*, 174(6), e13840. https://doi.org/10.1111/ppl.13840

Ackerman, P. L., Beier, M. E., & Boyle, M. D. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, 131(4), 567–589. https://doi.org/10.1037/0096-3445.131.4.567

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of Working Memory: Differences in Definition, Degree of Modularity, Role of Attention, and Purpose. *Language, Speech, and Hearing Services in Schools*, 49(3), 340–355. https://doi.org/10.1044/2018_LSHSS-17-0114

Aellen, M., Burkart, J. M., & Bshary, R. (2022). No evidence for general intelligence in a fish. *Ethology*, 128(5), 424–436. https://doi.org/10.1111/eth.13275

Aellen, M., Dufour, V., & Bshary, R. (2021). Cleaner fish and other wrasse match primates in their ability to delay gratification. *Animal Behaviour*, 176, 125–143. https://doi.org/10.1016/j.anbehav.2021.04.002

Aellen, M., Siebeck, U. E., & Bshary, R. (2022). Cleaner wrasse Labroides dimidiatus perform above chance in a "matching-to-sample" experiment. *PloS One*, 17(1), e0262351. https://doi.org/10.1371/journal.pone.0262351

Aiello, L. C., Bates, N., & Joffe, T. (2001). In defense of the Expensive Tissue Hypothesis. In D. Falk & K. R. Gibson (Eds.), Evolutionary Anatomy of the Primate Cerebral Cortex (pp. 57–78). *Cambridge University Press*. https://doi.org/10.1017/CBO9780511897085.006

Alloway, T., & Alloway, R. (2009). The efficacy of working memory training in improving crystallized intelligence. *Nature Precedings*. https://doi.org/10.1038/npre.2009.3697.1

Atkinson, D. (1994). Temperature and Organism Size—A Biological Law for Ectotherms? In M. Begon & A. H. Fitter (Eds.), Advances in Ecological Research (Vol. 25, pp. 1–58). *Academic Press*. https://doi.org/10.1016/S0065-2504(08)60212-3

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and Its Control Processes. https://escholarship.org/content/qt5kd4s4j3/qt5kd4s4j3.pdf

Bachevalier, J., & Mishkin, M. (1986). Visual recognition impairment follows ventromedial but not dorsolateral prefrontal lesions in monkeys. *Behavioural Brain Research*, 20(3), 249–261. https://doi.org/10.1016/0166-4328(86)90225-1

Baddeley, A. (1986). Working memory (pp. xi, 289). *Clarendon Press/Oxford University Press.*

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, 63(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D., Hitch, G., & Allen, R. (2021). A multicomponent model of working memory. In *Working Memory: State of the Science*, 10–43.

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, 112, 91–106. https://doi.org/10.1016/j.cortex.2018.05.015

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., & Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain*, 135(4), 1154–1164. https://doi.org/10.1093/brain/aws021

Barth, J., & Call, J. (2006). Tracking the displacement of objects: A series of tasks with great apes (Pan troglodytes, Pan paniscus, Gorilla gorilla, and Pongo pygmaeus) and young children (Homo sapiens). *Journal of Experimental Psychology: Animal Behavior Processes*, 32(3), 239–252. https://doi.org/10.1037/0097-7403.32.3.239

Benson-Amram, S., Dantzer, B., Stricker, G., Swanson, E. M., & Holekamp, K. E. (2016). Brain size predicts problem-solving ability in mammalian carnivores. *Proceedings of the National Academy of Sciences of the United States of America*, 113(9), 2532–2537. https://doi.org/10.1073/pnas.1505913113

Binet, A., & Simon, Th. (1905). Application of the new methods to the diagnosis of the intellectual level among normal and subnormal children in institutions and in the primary schools. *L'Année Psychologique*, 245–236. https://doi.org/10.1037/11069-003

Blackburn, T. M., Gaston, K. J., & Loder, N. (1999). Geographic gradients in body size: A clarification of Bergmann's rule. *Diversity and Distributions*, 5(4), 165–174. https://doi.org/10.1046/j.1472-4642.1999.00046.x

Bloch, S., Froc, C., Pontiggia, A., & Yamamoto, K. (2019). Existence of working memory in teleosts: Establishment of the delayed matching-to-sample task in adult zebrafish. *Behavioural Brain Research*, 370, 111924. https://doi.org/10.1016/j.bbr.2019.111924

Boisseau, R. P., Vogel, D., & Dussutour, A. (2016). Habituation in non-neural organisms: Evidence from slime moulds. *Proceedings of the Royal Society B: Biological Sciences*, 283(1829), 20160446. https://doi.org/10.1098/rspb.2016.0446

Brown, R. E. (2016). Hebb and Cattell: The genesis of the theory of fluid and crystallized intelligence. *Frontiers in Human Neuroscience*, 10. https://doi.org/10.3389/fnhum.2016.00606

Brzezicka, A., Kamiński, J., Reed, C. M., Chung, J. M., Mamelak, A. N., & Rutishauser, U. (2019). Working Memory Load-related Theta Power Decreases in Dorsolateral Prefrontal Cortex Predict Individual Differences in Performance. *Journal of Cognitive Neuroscience*, 31(9), 1290–1307. https://doi.org/10.1162/jocn_a_01417

Bshary, R. (2002). Biting cleaner fish use altruism to deceive image–scoring client reef fish. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1505), 2087–2093. https://doi.org/10.1098/rspb.2002.2084

Bshary, R. (2003). The cleaner wrasse, Labroides dimidiatus, is a key organism for reef fish diversity at Ras Mohammed National Park, Egypt. *Journal of Animal Ecology*, 72(1), 169–176. https://doi.org/10.1046/j.1365-2656.2003.00683.x

Bshary, R., & Grutter, A. S. (2002). Experimental evidence that partner choice is a driving force in the payoff distribution among cooperators or mutualists: The cleaner fish case. *Ecology Letters*, 5(1), 130–136. https://doi.org/10.1046/j.1461-0248.2002.00295.x

Bshary, R., & Grutter, A. S. (2005). Punishment and partner switching cause cooperative behaviour in a cleaning mutualism. *Biology Letters*, 1(4), 396–399.

Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441(7096), 975–978. https://doi.org/10.1038/nature04755

Bshary, R., Grutter, A. S., Willener, A. S. T., & Leimar, O. (2008). Pairs of cooperating cleaner fish provide better service quality than singletons. *Nature*, 455(7215), Article 7215. https://doi.org/10.1038/nature07184

Bshary, R., & Noë, R. (2003). Biological markets: The ubiquitous influence of partner choice on the dynamics of cleaner fish–client reef fish interactions.

Bshary, R., Oliveira, R. F., Oliveira, T. S., & Canário, A. V. (2007). Do cleaning organisms reduce the stress response of client reef fish*? Frontiers in Zoology*, 4(1), 21. https://doi.org/10.1186/1742-9994-4-21

Bshary, R., & Schäffer, D. (2002). Choosy reef fish select cleaner fish that provide high-quality service. *Animal Behaviour*, 63(3), 557–564. https://doi.org/10.1006/anbe.2001.1923

Bshary, R., & Würth, M. (2001). Cleaner fish Labroides dimidiatus manipulate client reef fish by providing tactile stimulation. *Proceedings of the Royal Society B: Biological Sciences*, 268(1475), 1495–1501. https://doi.org/10.1098/rspb.2001.1495

Burkart, J. M., Schubiger, M. N., & Van Schaik, C. P. (2017). The evolution of general intelligence. *Behavioral and Brain Sciences*, 40, e195. https://doi.org/10.1017/S0140525X16000959

Caicoya, A. L., Amici, F., Ensenyat, C., & Colell, M. (2021). Comparative cognition in three understudied ungulate species: European bison, forest buffalos and giraffes. *Frontiers in Zoology*, 18(1), 30. https://doi.org/10.1186/s12983-021-00417-w

Calvo, R., & Schluessel, V. (2021). Neural substrates involved in the cognitive information processing in teleost fish. *Animal Cognition*, 24(5), 923–946. https://doi.org/10.1007/s10071-021-01514-3

Carruthers, P. (2013). Evolution of working memory. *Proceedings of the National Academy of Sciences*, 110 (supplement_2), 10371–10378. https://doi.org/10.1073/pnas.1301195110

Cattell, R. B. (1941). Some theoretical issues in adult intelligence. *Psychological Bulletin*, 38(7), 592.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40(3), 153–193. https://doi.org/10.1037/h0059973

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1–22. https://doi.org/10.1037/h0046743

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J.-D. (2017). The Distributed Nature of Working Memory. *Trends in Cognitive Sciences*, 21(2), 111–124. https://doi.org/10.1016/j.tics.2016.12.007

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395(6699), 272–274. https://doi.org/10.1038/26216

Collette, F., & Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience & Biobehavioral Reviews*, 26(2), 105–125. https://doi.org/10.1016/S0149-7634(01)00063-X

Combe, G. (1824). *Outlines of Phrenology*.

Conklin, A. L. (2021). Phrenology as global science, as race science. *Global Intellectual History*, 6(2), 105–114. https://doi.org/10.1080/23801883.2021.1892909

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552. https://doi.org/10.1016/j.tics.2003.10.005

Côté, I. (2000). Evolution and ecology of cleaning symbioses in the sea. *Oceanography and Marine Biology.*

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104(2), 163–191. https://doi.org/10.1037/0033-2909.104.2.163

Cowan, N. (1995). Attention and memory: An integrated framework (pp. xv, 321). *Oxford University Press.*

Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin*, 145(8), 822–847. https://doi.org/10.1037/bul0000199

Cowan, N. (2022). Working Memory Development: A 50-Year Assessment of Research and Underlying Theories. *Cognition*, 224, 105075. https://doi.org/10.1016/j.cognition.2022.105075

Cowan, N., Bao, C., Bishop-Chrzanowski, B. M., Costa, A. N., Greene, N. R., Guitard, D., Li, C., Musich, M. L., & Ünal, Z. E. (2024). The Relation Between Attention and Memory. *Annual Review of Psychology*, 75(1), 183–214. https://doi.org/10.1146/annurev-psych-040723-012736

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the Capacity of Attention: Its Estimation and Its Role in Working Memory and Cognitive Aptitudes. *Cognitive Psychology*, 51(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

Cowan, N., & Morey, C. C. (2006). Visual working memory depends on attentional filtering. *Trends in Cognitive Sciences*, 10(4), 139–141. https://doi.org/10.1016/j.tics.2006.02.001

Cowan, N., Morey, C. C., & Naveh-Benjamin, M. (2020). An embedded-processes approach to working memory. *Working Memory: The State of the Science*, 44.

Darwin, C. J., Turvey, M. T., & Crowder, R. G. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, 3(2), 255–267.

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A. (2020). Chapter 19—Executive functions. In A. Gallagher, C. Bulteau, D. Cohen, & J. L. Michaud (Eds.), *Handbook of Clinical Neurology* (Vol. 173, pp. 225–240). Elsevier. https://doi.org/10.1016/B978-0-444-64150-2.00020-4

Dicke, U., & Roth, G. (2016). Neuronal factors determining high intelligence. Philosophical *Transactions of the Royal Society B: Biological Sciences*, 371(1685), 20150180. https://doi.org/10.1098/rstb.2015.0180

Donhoffer, S. (1980). Homeothermia of the brain. Cerebral blood flow, metabolic rate, and brain temperature in the cold. The possible role of neuroglia. *Akadémiai Kiadó*. https://real-eod.mtak.hu/15254/

Doré, F. Y., Fiset, S., Goulet, S., Dumas, M.-C., & Gagnon, S. (1996). Search behavior in cats and dogs: Interspecific differences in working memory and spatial cognition. *Animal Learning & Behavior*, 24(2), 142–149. https://doi.org/10.3758/BF03198962

Fiset, S., Beaulieu, C., & Landry, F. (2003). Duration of dogs' (Canis familiaris) working memory in search for disappearing objects. *Animal Cognition*, 6(1), 1–10. https://doi.org/10.1007/s10071-002-0157-4

Fontana, B. D., Gibbon, A. J., Cleal, M., Sudwarts, A., Pritchett, D., Miletto Petrazzini, M. E., Brennan, C. H., & Parker, M. O. (2021). Moderate early life stress improves adult zebrafish (Danio rerio) working memory but does not affect social and anxiety-like responses. *Developmental Psychobiology*, 63(1), 54–64. https://doi.org/10.1002/dev.21986

Fragua, V., Lepoudère, A., Leray, V., Baron, C., Araujo, J. A., Nguyen, P., & Milgram, N. W. (2017). Effects of dietary supplementation with a mixed blueberry and grape extract on working memory in aged beagle dogs. *Journal of Nutritional Science*, 6, e35. https://doi.org/10.1017/jns.2017.33

Friedman, N. P., & Robbins, T. W. (2022). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*, 47(1), 72–89. https://doi.org/10.1038/s41386-021-01132-0

Frischkorn, G. T., Schubert, A.-L., & Hagemann, D. (2019). Processing speed, working memory, and executive functions: Independent or inter-related predictors of general intelligence. *Intelligence*, 75, 95–110. https://doi.org/10.1016/j.intell.2019.05.003

Galton, F. (1865). Hereditary talent and character. *Macmillan's Magazine.*

Gingins, S., & Bshary, R. (2016). The cleaner wrasse outperforms other labrids in ecologically relevant contexts, but not in spatial discrimination. *Animal Behaviour*, 115, 145–155. https://doi.org/10.1016/j.anbehav.2016.02.022

Gläscher, J., Rudrauf, D., Colom, R., Paul, L. K., Tranel, D., Damasio, H., & Adolphs, R. (2010). Distributed neural system for general intelligence revealed by lesion mapping. *Proceedings of the National Academy of Sciences*, 107(10), 4705–4709. https://doi.org/10.1073/pnas.0910397107

Goriounova, N. A., Heyer, D. B., Wilbers, R., Verhoog, M. B., Giugliano, M., Verbist, C., Obermayer, J., Kerkhofs, A., Smeding, H., Verberne, M., Idema, S., Baayen, J. C., Pieneman, A. W., de Kock, C. P., Klein, M., & Mansvelder, H. D. (2018). Large and fast human pyramidal neurons associate with intelligence. *eLife*, 7, e41714. https://doi.org/10.7554/eLife.41714

Grosenick, L., Clement, T. S., & Fernald, R. D. (2007). Fish can infer social rank by observation alone. *Nature*, 445(7126), 429–432.

Grutter, A. S., & Bshary, R. (2003). Cleaner wrasse prefer client mucus: Support for partner control mechanisms in cleaning interactions. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270 (suppl_2). https://doi.org/10.1098/rsbl.2003.0077

Grutter, A. S., Murphy, J. M., & Choat, J. H. (2003). Cleaner Fish Drives Local Fish Diversity on Coral Reefs. *Current Biology*, 13(1), 64–67. https://doi.org/10.1016/S0960-9822(02)01393-3

Guadagno, A., & Triki, Z. (2024). Executive functions and brain morphology of male and female dominant and subordinate cichlid fish. *Brain and Behavior*, 14(5), e3484. https://doi.org/10.1002/brb3.3484

Güntürkün, O. (2005). The avian 'prefrontal cortex' and cognition. *Current Opinion in Neurobiology*, 15(6), 686–693. https://doi.org/10.1016/j.conb.2005.10.003

Hankins, F. H. (with Osmania University & Digital Library Of India) (1926). The Racial Basis Of Civilization. *Alfred. A. Knopf, Inc*.

Heldstab, S. A., Isler, K., Graber, S. M., Schuppli, C., & van Schaik, C. P. (2022). The economics of brain size evolution in vertebrates. *Current Biology*, 32(12), R697–R708. https://doi.org/10.1016/j.cub.2022.04.096

Helfman, G. S., & Schultz, E. T. (1984). Social transmission of behavioural traditions in a coral reef fish. *Animal Behaviour*, 32(2), 379–384. https://doi.org/10.1016/S0003-3472(84)80272-9

Henneberg, M. (1998). Evolution of the Human Brain: Is Bigger Better? *Clinical and Experimental Pharmacology and Physiology*, 25(9), 745–749. https://doi.org/10.1111/j.1440-1681.1998.tb02289.x

Herold, C., Palomero-Gallagher, N., Hellmann, B., Kröner, S., Theiss, C., Güntürkün, O., & Zilles, K. (2011). The receptor architecture of the pigeons' nidopallium caudolaterale: An avian analogue to the mammalian prefrontal cortex. *Brain Structure and Function*, 216(3), 239–254. https://doi.org/10.1007/s00429-011-0301-5

Hoffmann, A., Rüttler, V., & Nieder, A. (2011). Ontogeny of object permanence and object tracking in the carrion crow, Corvus corone. *Animal Behaviour*, 82(2), 359–367. https://doi.org/10.1016/j.anbehav.2011.05.012

Hotta, T., Ueno, K., Hataji, Y., Kuroshima, H., Fujita, K., & Kohda, M. (2020). Transitive inference in cleaner wrasses (Labroides dimidiatus). *PloS One*, 15(8), e0237817.

Huffeldt, N. P. (2020). Performance of horned puffins (Fratercula corniculata) on an object permanence task. *Behavioural Processes*, 181, 104274. https://doi.org/10.1016/j.beproc.2020.104274

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, 2(4), 705–715. https://doi.org/10.1111/j.1756-8765.2010.01089.x

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, 17(23), R1004–R1005. https://doi.org/10.1016/j.cub.2007.10.027

Isler, K., & van Schaik, C. P. (2009). The Expensive Brain: A framework for explaining evolutionary changes in brain size. *Journal of Human Evolution*, 57(4), 392–400. https://doi.org/10.1016/j.jhevol.2009.04.009

James, F. C. (1970). Geographic Size Variation in Birds and Its Relationship to Climate. *Ecology*, 51(3), 365–390. https://doi.org/10.2307/1935374

Jerison, H. J. (1973). Evolution of The Brain and Intelligence. *Elsevier*.

Jerison, H. J. (1969). Brain Evolution and Dinosaur Brains. *The American Naturalist*, 103(934), 575–588. https://doi.org/10.1086/282627

Johnston, M., Porter, B., & Colombo, M. (2020). Nidopallium caudolaterale neuronal responses during serial-order behaviour in pigeons. *Behavioural Brain Research*, 378, 112269. https://doi.org/10.1016/j.bbr.2019.112269

Kamil, A. C. (1998). On the proper definition of cognitive ethology. In *Animal cognition in nature* (pp. 1–28).

Kawachi, Y., & Gyoba, J. (2006). A new response-time measure of object persistence in the tunnel effect. *Acta Psychologica*, 123(1–2), 73–90. https://doi.org/10.1016/j.actpsy.2006.04.003

Kippenberger, S., Pipa, G., Steinhorst, K., Zöller, N., Kleemann, J., Özistanbullu, D., Kaufmann, R., & Scheller, B. (2023). Learning in the Single-Cell Organism Physarum polycephalum: Effect of Propofol. *International Journal of Molecular Sciences*, 24(7), Article 7. https://doi.org/10.3390/ijms24076287

Knauff, M., & Wolf, A. G. (2010). Complex cognition: The science of human reasoning, problem-solving, and decision-making. *Cognitive Processing*, 11(2), 99–102. https://doi.org/10.1007/s10339-010-0362-z

Kohda, M., Bshary, R., Kubo, N., Awata, S., Sowersby, W., Kawasaka, K., Kobayashi, T., & Sogawa, S. (2023). Cleaner fish recognize self in a mirror via self-face recognition like humans. *Proceedings of the National Academy of Sciences*, 120(7), e2208420120.

Kohda, M., Hotta, T., Takeyama, T., Awata, S., Tanaka, H., Asai, J., & Jordan, A. L. (2019). If a fish can pass the mark test, what are the implications for consciousness and self-awareness testing in animals? *PloS Biology*, 17(2), e3000021.

Kohda, M., Sogawa, S., Jordan, A. L., Kubo, N., Awata, S., Satoh, S., Kobayashi, T., Fujita, A., & Bshary, R. (2022). Further evidence for the capacity of mirror self-recognition in cleaner fish and the significance of ecologically relevant marks. *PloS Biology*, 20(2), e3001529.

Kolata, S., Light, K., Townsend, D. A., Hale, G., Grossman, H. C., & Matzel, L. D. (2005). Variations in working memory capacity predict individual differences in general learning abilities among genetically diverse mice. *Neurobiology of Learning and Memory*, 84(3), 241–246. https://doi.org/10.1016/j.nlm.2005.07.006

Kotrschal, A., Corral-Lopez, A., Amcoff, M., & Kolm, N. (2015). A larger brain confers a benefit in a spatial mate search learning task in male guppies. *Behavioral Ecology*, 26(2), 527–532. https://doi.org/10.1093/beheco/aru227

Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I., Immler, S., Maklakov, A. A., & Kolm, N. (2013). Artificial Selection on Relative Brain Size in the Guppy Reveals Costs and Benefits of Evolving a Larger Brain. *Current Biology*, 23(2), 168–171. https://doi.org/10.1016/j.cub.2012.11.058

Kraeuter, A.-K., Guest, P. C., & Sarnyai, Z. (2019). The Y-Maze for Assessment of Spatial Working and Reference Memory in Mice. In *P. C. Guest (Ed.), Pre-Clinical Models* (Vol. 1916, pp. 105–111). Springer New York. https://doi.org/10.1007/978-1-4939-8994-2_10

Laumer, I. B., Rahman, A., Rahmaeti, T., Azhari, U., Hermansyah, Atmoko, S. S. U., & Schuppli, C. (2024). Active self-treatment of a facial wound with a biologically active plant by a male Sumatran orangutan. *Scientific Reports*, 14(1), 8932. https://doi.org/10.1038/s41598-024-58988-7

Lecce, S., & Bianco, F. (2018). Working memory predicts changes in children's theory of mind during middle childhood: A training study. *Cognitive Development*, 47, 71–81. https://doi.org/10.1016/j.cogdev.2018.04.002

Lee, S. A., Tucci, V., Sovrano, V. A., & Vallortigara, G. (2015). Working memory and reference memory tests of spatial navigation in mice (Mus musculus). *Journal of Comparative Psychology*, 129(2), 189–197. https://doi.org/10.1037/a0039129

Li, M.-M., Fan, J.-T., Cheng, S.-G., Yang, L.-F., Yang, L., Wang, L.-F., Shang, Z.-G., & Wan, H. (2021). Enhanced hippocampus-nidopallium caudolaterale connectivity during route formation in goal-directed spatial learning of pigeons. *Animals*, 11(7), 2003.

Libben, M., & Titone, D. (2008). The role of awareness and working memory in human transitive inference. *Behavioural Processes*, 77(1), 43–54. https://doi.org/10.1016/j.beproc.2007.06.006

Lind, J., Enquist, M., & Ghirlanda, S. (2015). Animal memory: A review of delayed matching-to-sample data. *Behavioural Processes*, 117, 52–58. https://doi.org/10.1016/j.beproc.2014.11.019

Liu, X., Zhao, K., Wang, D., Ping, Y., & Wan, H. (2018). Goal-directed behavior elevates gamma oscillations in nidopallium caudolaterale of pigeon. *Brain Research Bulletin*, 137, 10–16. https://doi.org/10.1016/j.brainresbull.2017.10.013

Logie, R. H., Del Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual Similarity Effects in Immediate Verbal Serial Recall. *The Quarterly Journal of Experimental Psychology Section A*, 53(3), 626–646. https://doi.org/10.1080/713755916

Macpherson, K., & Roberts, W. A. (2010). Spatial memory in dogs (Canis familiaris) on a radial maze. *Journal of Comparative Psychology*, 124(1), 47–56. https://doi.org/10.1037/a0018084

Manger, P. (2006). An examination of cetacean brain structure with a novel hypothesis correlating thermogenesis to the evolution of a big brain. *Biological Reviews of the Cambridge Philosophical Society*, 81, 293–338. https://doi.org/10.1017/S1464793106007019

Manger, P. R., Patzke, N., Spocter, M. A., Bhagwandin, A., Karlsson, K. Æ., Bertelsen, M. F., Alagaili, A. N., Bennett, N. C., Mohammed, O. B., Herculano-Houzel, S., Hof, P. R., & Fuxe, K. (2021). Amplification of potential thermogenetic mechanisms in cetacean brains compared to artiodactyl brains. *Scientific Reports*, 11(1), 5486. https://doi.org/10.1038/s41598-021-84762-0

Mann, J., & Patterson, E. M. (2013). Tool use by aquatic animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1630), 20120424. https://doi.org/10.1098/rstb.2012.0424

Manrique, H. M., Read, D. W., & Walker, M. J. (2024). On some statistical and cerebral aspects of the limits of working memory capacity in anthropoid primates, with particular reference to Pan and Homo, and their significance for human evolution. *Neuroscience & Biobehavioral Reviews*, 158, 105543. https://doi.org/10.1016/j.neubiorev.2024.105543

Manrique, H. M., & Walker, M. J. (2017). Early Evolution of Human Memory: Great Apes, Tool-making, and Cognition. *Springer International Publishing*. https://doi.org/10.1007/978-3-319-64447-9

Marhounová, L., Kotrschal, A., Kverková, K., Kolm, N., & Němec, P. (2019). Artificial selection on brain size leads to matching changes in overall number of neurons. *Evolution*, 73(9), 2003–2012. https://doi.org/10.1111/evo.13805

Martin, R. D. (1981). Relative brain size and basal metabolic rate in terrestrial vertebrates. *Nature*, 293(5827), 57–60.

McAlonan, K., & Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioural Brain Research*, 146(1), 97–103. https://doi.org/10.1016/j.bbr.2003.09.019

Mednikova, Y. S., Pasikova, N. V., & Kopytova, F. V. (2004). Effects of temperature on the spike activity of cortical neurons in guinea pigs. *Neuroscience and Behavioral Physiology*, 34, 459–465.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). Plans and the structure of behavior (pp. x, 222). *Henry Holt and Co*. https://doi.org/10.1037/10039-000

Morrison, R. G. (2005). Thinking in working memory. *Cambridge Handbook of Thinking and Reasoning*, 457–473.

Nakamura, I., Matsumoto, R., & Sato, K. (2017). Gigantothermy: A strategy toward cold environment.

Nakashima, Y., Sakai, Y., Karino, K., & Kuwamura, T. (2000). Female-female spawning and sex change in a haremic coral-reef fish, Labroides dimidiatus. *Zoological Science*, 17(7), 967–970.

Oberheim, N. A., Wang, X., Goldman, S., & Nedergaard, M. (2006). Astrocytic complexity distinguishes the human brain. *Trends in Neurosciences*, 29(10), 547–553. https://doi.org/10.1016/j.tins.2006.08.004

Ortiz, S. O. (2015). CHC Theory of Intelligence. In S. Goldstein, D. Princiotta, & J. A. Naglieri (Eds.), *Handbook of Intelligence* (pp. 209–227). Springer New York. https://doi.org/10.1007/978-1-4939-1562-0_15

Paterson, D. G. (1930). Physique and Intellect. *The Century CO*. http://archive.org/details/physiqueandintel001131mbp

Pause, B. M., Zlomuzica, A., Kinugawa, K., Mariani, J., Pietrowsky, R., & Dere, E. (2013). Perspectives on Episodic-Like and Episodic Memory. *Frontiers in Behavioral Neuroscience*, 7. https://doi.org/10.3389/fnbeh.2013.00033

Pavlov, P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, 17(3), 136.

Pepperberg, I. M., & Funk, M. S. (1990). Object permanence in four species of psittacine birds: An African Grey parrot (Psittacus erithacus), an Illiger mini macaw (Ara maracana), a parakeet (Melopsittacus undulatus), and a cockatiel (Nymphicus hollandicus). *Animal Learning & Behavior*, 18(1), 97–108. https://doi.org/10.3758/BF03205244

Piaget, J., & Cook, M. (1952). The origins of intelligence in children (Vol. 8). *International Universities Press New York.*

Piaget, J., & Cook, M. T. (1954). The development of object concept. In *The Construction of Reality in the Child. Basic Books.*

Pinto, A., Oates, J., Grutter, A., & Bshary, R. (2011). Cleaner Wrasses Labroides dimidiatus Are More Cooperative in the Presence of an Audience. *Current Biology*, 21(13), 1140–1144. https://doi.org/10.1016/j.cub.2011.05.021

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526.

Raihani, N. J., Pinto, A. I., Grutter, A. S., Wismer, S., & Bshary, R. (2012). Male cleaner wrasses adjust punishment of female partners according to the stakes. *Proceedings of the Royal Society B: Biological Sciences*, 279(1727), 365–370. https://doi.org/10.1098/rspb.2011.0690

Read, D. W., Manrique, H. M., & Walker, M. J. (2022). On the working memory of humans and great apes: Strikingly similar or remarkably different? *Neuroscience & Biobehavioral Reviews*, 134, 104496. https://doi.org/10.1016/j.neubiorev.2021.12.019

Reed, T. E., Vernon, P. A., & Johnson, A. M. (2004). Confirmation of correlation between brain nerve conduction velocity and intelligence level in normal adults. *Intelligence*, 32(6), 563–572. https://doi.org/10.1016/j.intell.2004.07.004

Reid, C. R. (2023). Thoughts from the forest floor: A review of cognition in the slime mould Physarum polycephalum. *Animal Cognition*, 26(6), 1783–1797. https://doi.org/10.1007/s10071-023-01782-1

Robertson, D. R. (1972). Social control of sex reversal in a coral-reef fish. *Science*, 177(4053), 1007–1009.

Ros, A. F. H., Lusa, J., Meyer, M., Soares, M., Oliveira, R. F., Brossard, M., & Bshary, R. (2011). Does access to the bluestreak cleaner wrasse Labroides dimidiatus affect indicators of stress and health in resident reef fishes in the Red Sea? *Hormones and Behavior*, 59(1), 151–158. https://doi.org/10.1016/j.yhbeh.2010.11.006

Sakai, Y., Kohda, M., & Kuwamura, T. (2001). Effect of changing harem on timing of sex change in female cleaner fish Labroides dimidiatus. *Animal Behaviour*, 62(2), 251–257.

Salehinejad, M. A., Ghanavati, E., Rashid, M. H. A., & Nitsche, M. A. (2021). Hot and cold executive functions in the brain: A prefrontal-cingular network. *Brain and Neuroscience Advances*, 5, 23982128211007769. https://doi.org/10.1177/23982128211007769

Seed, A., & Byrne, R. (2010). Animal Tool-Use. *Current Biology*, 20(23), R1032–R1039. https://doi.org/10.1016/j.cub.2010.09.042

Shettleworth, S. J. (2009). Cognition, evolution, and behavior. *Oxford university press.*

Shirakawa, T., Gunji, Y.-P., & Miyake, Y. (2011). An associative learning experiment using the plasmodium of Physarum polycephalum. *Nano Communication Networks*, 2(2), 99–105. https://doi.org/10.1016/j.nancom.2011.05.002

Skinner, B. F. (1937). Two Types of Conditioned Reflex: A Reply to Konorski and Miller. *The Journal of General Psychology*, 16(1), 272–279. https://doi.org/10.1080/00221309.1937.9917951

Soares, M. C., Cardoso, S. C., Nicolet, K. J., Côté, I. M., & Bshary, R. (2013). Indo-Pacific parrotfish exert partner choice in interactions with cleanerfish but Caribbean parrotfish do not. *Animal Behaviour*, 86(3), 611–615. https://doi.org/10.1016/j.anbehav.2013.06.017

Soares, M. C., Oliveira, R. F., Ros, A. F. H., Grutter, A. S., & Bshary, R. (2011). Tactile stimulation lowers stress in fish. *Nature Communications*, 2(1), 534. https://doi.org/10.1038/ncomms1547

Sol, D., Bacher, S., Reader, S. M., & Lefebvre, L. (2008). Brain Size Predicts the Success of Mammal Species Introduced into Novel Environments. *The American Naturalist*, 172(S1), S63–S71. https://doi.org/10.1086/588304

Songthawornpong, N., Teasdale, T. W., Olesen, M. V., & Pakkenberg, B. (2021). Is There a Correlation Between the Number of Brain Cells and IQ? *Cerebral Cortex*, 31(1), 650–657. https://doi.org/10.1093/cercor/bhaa249

Spearman, C. (1904). "General Intelligence," Objectively Determined And Measured. *American Journal of Psychology*, 15(2), 201–293.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74(11), 1.

Squire, L. R., & Dede, A. J. (2015). Conscious and unconscious memory systems. *Cold Spring Harbor Perspectives in Biology*, 7(3), a021667.

Suddendorf, T., & Busby, J. (2003). Mental time travel in animals? *Trends in Cognitive Sciences*, 7(9), 391–396. https://doi.org/10.1016/S1364-6613(03)00187-6

Tecwyn, E. C., Thorpe, S. K., & Chappell, J. (2013). A novel test of planning ability: Great apes can plan step-by-step but not in advance of action. *Behavioural Processes*, 100, 174–184.

Terman, L. M. (1950). The Binet-Simon Intelligence Scale. In W. Dennis (Ed.), *Readings in general psychology*. (pp. 331–337). Prentice-Hall, Inc. https://doi.org/10.1037/11352-047

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, 39(1/4), 212–222.

Tibbetts, E. A., Agudelo, J., Pandit, S., & Riojas, J. (2019). Transitive inference in Polistes paper wasps. *Biology Letters*, 15(5), 20190015.

Triki, Z., & Bshary, R. (2021). Sex differences in the cognitive abilities of a sex-changing fish species Labroides dimidiatus. *Royal Society Open Science*, 8(7), 210239. https://doi.org/10.1098/rsos.210239

Triki, Z., Fong, S., Amcoff, M., Vàsquez-Nilsson, S., & Kolm, N. (2023). Experimental expansion of relative telencephalon size improves the main executive function abilities in guppy. *PNAS Nexus*, 2(6), pgad129. https://doi.org/10.1093/pnasnexus/pgad129

Truskanov, N., Emery, Y., & Bshary, R. (2020). Juvenile cleaner fish can socially learn the consequences of cheating. *Nature Communications*, 11(1), 1159. https://doi.org/10.1038/s41467-020-14712-3

Tulving, E. (1972). Episodic and semantic memory. In Organization of memory (pp. xiii, 423–xiii, 423). *Academic Press*.

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, 53(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Uzgiris, I. C., & Hunt, J. (1975). Assessment in infancy: Ordinal scales of psychological development.

Vander Wall, S. B. (1990). Food hoarding in animals. *University of Chicago Press.*

Vesterdorf, K., Blache, D., & Maloney, S. K. (2011). The cranial arterio-venous temperature difference is related to respiratory evaporative heat loss in a panting species, the sheep (Ovis aries). *Journal of Comparative Physiology B*, 181, 277–288.

Vogel, D., & Dussutour, A. (2016). Direct transfer of learned behaviour via cell fusion in non-neural organisms. *Proceedings of the Royal Society B: Biological Sciences*, 283(1845), 20162382. https://doi.org/10.1098/rspb.2016.2382

Völter, C. J., & Call, J. (2014). Younger apes and human children plan their moves in a maze task. *Cognition*, 130(2), 186–203.

Völter, C. J., Mundry, R., Call, J., & Seed, A. M. (2019). Chimpanzees flexibly update working memory contents and show susceptibility to distraction in the self-ordered search task. *Proceedings of the Royal Society B: Biological Sciences,* 286(1907), 20190715. https://doi.org/10.1098/rspb.2019.0715

Walters, R. J., & Hassall, M. (2006). The Temperature-Size Rule in Ectotherms: May a General Explanation Exist after All? *The American Naturalist*, 167(4), 510–523. https://doi.org/10.1086/501029

Wan, C., Cai, P., Wang, M., Qian, Y., Huang, W., & Chen, X. (2020). Artificial Sensory Memory. *Advanced Materials*, 32(15), 1902434. https://doi.org/10.1002/adma.201902434

Warner, R. (1988). Traditionality of mating-site preferences in a coral reef fish. *Nature*, 335, 719–721. https://doi.org/10.1038/335719a0

White, S. L., & Gowan, C. (2013). Brook trout use individual recognition and transitive inference to determine social rank. *Behavioral Ecology*, 24(1), 63–69.

Wismer, S., Grutter, A., & Bshary, R. (2016). Generalized rule application in bluestreak cleaner wrasse (Labroides dimidiatus): Using predator species as social tools to reduce punishment. *Animal Cognition*, 19(4), 769–778. https://doi.org/10.1007/s10071-016-0975-4

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, 46(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Zola, S. M., Squire, L. R., Teng, E., Stefanacci, L., Buffalo, E. A., & Clark, R. E. (2000). Impaired recognition memory in monkeys after damage limited to the hippocampal region. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 20(1), 451–463. https://doi.org/10.1523/JNEUROSCI.20-01-00451.2000

Zucca, P., Milos, N., & Vallortigara, G. (2007). Piagetian object permanence and its development in Eurasian jays (Garrulus glandarius). *Animal Cognition*, 10(2), 243–258. https://doi.org/10.1007/s10071-006-0063-2

R.2. Chapter 1

Abdi, S. K., Syed, A. A., & Batool, S. (2022). Eidetic Image Psychology: A Review of Literature. *Pakistan Journal of Applied Psychology*, 2(2), 125–136.

Ackerman, P. L., Beier, M. E., & Boyle, M. D. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, 131(4), 567–589. https://doi.org/10.1037/0096-3445.131.4.567

Aellen, M., Burkart, J. M., & Bshary, R. (2022). No evidence for general intelligence in a fish. *Ethology*, 128(5), 424–436. https://doi.org/10.1111/eth.13275

Aellen, M., Dufour, V., & Bshary, R. (2021). Cleaner fish and other wrasse match primates in their ability to delay gratification. *Animal Behaviour*, 176, 125–143. https://doi.org/10.1016/j.anbehav.2021.04.002

Aggleton, J. P., Hunt, P. R., & Rawlins, J. N. (1986). The effects of hippocampal lesions upon spatial and non-spatial tests of working memory. *Behavioural Brain Research*, 19(2), 133–146. https://doi.org/10.1016/0166-4328(86)90011-2

Alloway, T., & Alloway, R. (2009). The efficacy of working memory training in improving crystallized intelligence. *Nature Precedings.* https://doi.org/10.1038/npre.2009.3697.1

Angelopoulou, E., & Drigas, A. (2021). Working memory, attention and their relationship: A theoretical overview. *Research, Society and Development*, 10, e46410515288. https://doi.org/10.33448/rsd-v10i5.15288

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and Its Control Processes. https://escholarship.org/content/qt5kd4s4j3/qt5kd4s4j3.pdf

Bachevalier, J., & Mishkin, M. (1986). Visual recognition impairment follows ventromedial but not dorsolateral prefrontal lesions in monkeys. *Behavioural Brain Research*, 20(3), 249–261. https://doi.org/10.1016/0166-4328(86)90225-1

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), Psychology of Learning and Motivation (Vol. 8, pp. 47–89). *Academic Press*. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, 112, 91–106. https://doi.org/10.1016/j.cortex.2018.05.015

Binning, S. A., Rey, O., Wismer, S., Triki, Z., Glauser, G., Soares, M. C., & Bshary, R. (2017). Reputation management promotes strategic adjustment of service quality in cleaner wrasse. *Scientific Reports*, 7(1), 8425. https://doi.org/10.1038/s41598-017-07128-5

Bloch, S., Froc, C., Pontiggia, A., & Yamamoto, K. (2019). Existence of working memory in teleosts: Establishment of the delayed matching-to-sample task in adult zebrafish. *Behavioural Brain Research*, 370, 111924. https://doi.org/10.1016/j.bbr.2019.111924

Bonin, L., & Bshary, R. (2023). In the absence of extensive initial training, cleaner wrasse Labroides dimidiatus fail a transitive inference task. *PloS One*, 18(6), e0287402.

Bratch, A., Kann, S., Cain, J. A., Wu, J.-E., Rivera-Reyes, N., Dalecki, S., Arman, D., Dunn, A., Cooper, S., Corbin, H. E., Doyle, A. R., Pizzo, M. J., Smith, A. E., & Crystal, J. D. (2016). Working Memory Systems in the Rat. *Current Biology*, 26(3), 351–355. https://doi.org/10.1016/j.cub.2015.11.068

Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441(7096), 975–978. https://doi.org/10.1038/nature04755

Bshary, R., & Triki, Z. (2022). Fish ecology and cognition: Insights from studies on wild and wild-caught teleost fishes. *Current Opinion in Behavioral Sciences*, 46, 101174. https://doi.org/10.1016/j.cobeha.2022.101174

Carruthers, P. (2013). Evolution of working memory. *Proceedings of the National Academy of Sciences*, 110 (supplement_2), 10371–10378. https://doi.org/10.1073/pnas.1301195110

Chai, W. J., Abd Hamid, A. I., & Abdullah, J. M. (2018). Working Memory From the Psychological and Neurosciences Perspectives: A Review. *Frontiers in Psychology*, 9. https://doi.org/10.3389/fpsyg.2018.00401

Chojnacka, D., Isler, K., Barski, J. J., & Bshary, R. (2015). Relative brain and brain part sizes provide only limited evidence that machiavellian behaviour in cleaner wrasse is cognitively demanding. *PloS One*, 10(8), e0135373.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552. https://doi.org/10.1016/j.tics.2003.10.005

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104(2), 163–191. https://doi.org/10.1037/0033-2909.104.2.163

Cowan, N. (1995). Attention and memory: An integrated framework (pp. xv, 321). *Oxford University Press.*

Cowan, N. (2008). Chapter 20 What are the differences between long-term, short-term, and working memory? In *Progress in Brain Research* (Vol. 169, pp. 323–338). Elsevier. https://doi.org/10.1016/S0079-6123(07)00020-9

Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin*, 145(8), 822–847. https://doi.org/10.1037/bul0000199

Cowan, N. (2022). Working Memory Development: A 50-Year Assessment of Research and Underlying Theories. *Cognition*, 224, 105075. https://doi.org/10.1016/j.cognition.2022.105075

Cowan, N., Bao, C., Bishop-Chrzanowski, B. M., Costa, A. N., Greene, N. R., Guitard, D., Li, C., Musich, M. L., & Ünal, Z. E. (2024). The Relation Between Attention and Memory. *Annual Review of Psychology*, 75(1), 183–214. https://doi.org/10.1146/annurev-psych-040723-012736

Darwin, C. J., Turvey, M. T., & Crowder, R. G. (1972). An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, 3(2), 255–267.

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(2), 115–124. https://doi.org/10.1159/000102973

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A. (2020). Chapter 19—Executive functions. In A. Gallagher, C. Bulteau, D. Cohen, & J. L. Michaud (Eds.), *Handbook of Clinical Neurology* (Vol. 173, pp. 225–240). Elsevier. https://doi.org/10.1016/B978-0-444-64150-2.00020-4

Doré, F. Y., Fiset, S., Goulet, S., Dumas, M.-C., & Gagnon, S. (1996). Search behavior in cats and dogs: Interspecific differences in working memory and spatial cognition. *Animal Learning & Behavior*, 24(2), 142–149. https://doi.org/10.3758/BF03198962

Engle, R. W. (2010). Role of Working-Memory Capacity in Cognitive Control. *Current Anthropology*, 51(S1), S17–S26. https://doi.org/10.1086/650572

Engle, R. W., Laughlin, J. E., Tuholski, S. W., & Conway, A. R. A. (1999). Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *Journal of Experimental Psychology:General*, 128(3), 309–331.

Fiset, S., Beaulieu, C., & Landry, F. (2003). Duration of dogs' (Canis familiaris) working memory in search for disappearing objects. *Animal Cognition*, 6(1), 1–10. https://doi.org/10.1007/s10071-002-0157-4

Fonseca-Azevedo, K., & Herculano-Houzel, S. (2012). Metabolic constraint imposes tradeoff between body size and number of brain neurons in human evolution. *Proceedings of the National Academy of Sciences*, 109(45), 18571–18576.https://doi.org/10.1073/pnas.1206390109

Fontana, B. D., Gibbon, A. J., Cleal, M., Sudwarts, A., Pritchett, D., Miletto Petrazzini, M. E., Brennan, C. H., & Parker, M. O. (2021). Moderate early life stress improves adult zebrafish (Danio rerio) working memory but does not affect social and anxiety-like responses. *Developmental Psychobiology*, 63(1), 54–64. https://doi.org/10.1002/dev.21986

Fuster, J. M., & Bressler, S. L. (2012). Cognit activation: A mechanism enabling temporal integration in working memory. *Trends in Cognitive Sciences*, 16(4), 207–218. https://doi.org/10.1016/j.tics.2012.03.005

Grutter, A. (1996). Parasite removal rates by the cleaner wrasse Labroides dimidiatus. *Marine Ecology Progress Series*, 130, 61–70. https://doi.org/10.3354/meps130061

Grutter, A. S. (2004). Cleaner Fish Use Tactile Dancing Behavior as a Preconflict Management Strategy. *Current Biology*, 14(12), 1080–1083. https://doi.org/10.1016/j.cub.2004.05.048

Guadagno, A., & Triki, Z. (2024). Executive functions and brain morphology of male and female dominant and subordinate cichlid fish. *Brain and Behavior*, 14(5), e3484. https://doi.org/10.1002/brb3.3484

Hartig, F. (2022). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed). https://rdrr.io/cran/DHARMa/

Hotta, T., Ueno, K., Hataji, Y., Kuroshima, H., Fujita, K., & Kohda, M. (2020). Transitive inference in cleaner wrasses (Labroides dimidiatus). *PloS One*, 15(8), e0237817.

Hu, Z., Barkley, C. M., Marino, S. E., Wang, C., Rajan, A., Bo, K., Samuel, I. B. H., & Ding, M. (2019). Working Memory Capacity Is Negatively Associated with Memory Load Modulation of Alpha Oscillations in Retention of Verbal Working Memory. *Journal of Cognitive Neuroscience*, 31(12), 1933–1945. https://doi.org/10.1162/jocn_a_01461

Hughes, J., & Beiner, D. (2023). reghelper: Helper Functions for Regression Analysis. R Package Version 1.1.2. https://CRAN.R-project.org/package=reghelper

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, 17(23), R1004–R1005. https://doi.org/10.1016/j.cub.2007.10.027

Jerison, H. J. (1969). Brain Evolution and Dinosaur Brains. *The American Naturalist*, 103(934), 575–588. https://doi.org/10.1086/282627

Knauff, M., & Wolf, A. G. (2010). Complex cognition: The science of human reasoning, problem-solving, and decision-making. *Cognitive Processing*, 11(2), 99–102. https://doi.org/10.1007/s10339-010-0362-z

Kohda, M., Bshary, R., Kubo, N., Awata, S., Sowersby, W., Kawasaka, K., Kobayashi, T., & Sogawa, S. (2023). Cleaner fish recognize self in a mirror via self-face recognition like humans. *Proceedings of the National Academy of Sciences*, 120(7), e2208420120.

Kohda, M., Hotta, T., Takeyama, T., Awata, S., Tanaka, H., Asai, J., & Jordan, A. L. (2019). If a fish can pass the mark test, what are the implications for consciousness and self-awareness testing in animals? *PloS Biology*, 17(2), e3000021.

Kohda, M., Sogawa, S., Jordan, A. L., Kubo, N., Awata, S., Satoh, S., Kobayashi, T., Fujita, A., & Bshary, R. (2022). Further evidence for the capacity of mirror self-recognition in cleaner fish and the significance of ecologically relevant marks. *PloS Biology*, 20(2), e3001529.

Kolata, S., Light, K., Townsend, D. A., Hale, G., Grossman, H. C., & Matzel, L. D. (2005). Variations in working memory capacity predict individual differences in general learning abilities among genetically diverse mice. *Neurobiology of Learning and Memory*, 84(3), 241–246. https://doi.org/10.1016/j.nlm.2005.07.006

Kraeuter, A.-K., Guest, P. C., & Sarnyai, Z. (2019). The Y-Maze for Assessment of Spatial Working and Reference Memory in Mice. In P. C. Guest (Ed.), Pre-Clinical Models (Vol. 1916, pp. 105–111). *Springer New York*. https://doi.org/10.1007/978-1-4939-8994-2_10

Lee, S. A., Tucci, V., Sovrano, V. A., & Vallortigara, G. (2015). Working memory and reference memory tests of spatial navigation in mice (Mus musculus). *Journal of Comparative Psychology*, 129(2), 189–197. https://doi.org/10.1037/a0039129

Lenth, R. V. (2023). emmeans: Estimated marginal means, aka least-square means. *R Package Version.*

Logie, R., Camos, V., & Cowan, N. (2021). Working Memory: The state of the science. *Oxford University Press.*

Lucon-Xiccato, T., & Bisazza, A. (2017). Complex maze learning by fish. *Animal Behaviour*, 125, 69–75. https://doi.org/10.1016/j.anbehav.2016.12.022

MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., Aureli, F., Baker, J. M., Bania, A. E., Barnard, A. M., Boogert, N. J., Brannon, E. M., Bray, E. E., Bray, J., Brent, L. J. N., Burkart, J. M., Call, J., Cantlon, J. F., Cheke, L. G., … Zhao, Y. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences*, 111(20). https://doi.org/10.1073/pnas.1323533111

Macpherson, K., & Roberts, W. A. (2010). Spatial memory in dogs (Canis familiaris) on a radial maze. *Journal of Comparative Psychology*, 124(1), 47–56. https://doi.org/10.1037/a0018084

Manrique, H. M., Read, D. W., & Walker, M. J. (2024). On some statistical and cerebral aspects of the limits of working memory capacity in anthropoid primates, with particular reference to Pan and Homo, and their significance for human evolution. *Neuroscience & Biobehavioral Reviews*, 158, 105543. https://doi.org/10.1016/j.neubiorev.2024.105543

Manrique, H. M., & Walker, M. J. (2017). Early Evolution of Human Memory: Great Apes, Tool-making, and Cognition. *Springer International Publishing*. https://doi.org/10.1007/978-3-319-64447-9

McAuliffe, K., Drayton, L. A., Royka, A., Aellen, M., Santos, L. R., & Bshary, R. (2021). Cleaner fish are sensitive to what their partners can and cannot see. *Communications Biology*, 4(1), 1127.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). Plans and the structure of behavior (pp. x, 222). *Henry Holt and Co*. https://doi.org/10.1037/10039-000

Miller, S., & Peacock, R. (1982). Evidence for the Uniqueness of Eidetic Imagery. *Perceptual and Motor Skills*, 55(3_suppl), 1219–1233. https://doi.org/10.2466/pms.1982.55.3f.1219

Morey, C. C., & Cowan, N. (2018). Can we distinguish three maintenance processes in working memory? *Annals of the New York Academy of Sciences*, 1424(1), 45–51. https://doi.org/10.1111/nyas.13925

Morrison, R. G. (2005). Thinking in working memory. *Cambridge Handbook of Thinking and Reasoning*, 457–473.

Norris, D. (2017). Short-term memory and long-term memory are still different*. Psychological Bulletin*, 143(9), 992–1009. https://doi.org/10.1037/bul0000108

Özbaydar, B. (2012). Methodological problems in the study of Eidetic Imagery. *Psikoloji Çalışmaları*, 4, 41–54.

Parker, M. O., Gaviria, J., Haigh, A., Millington, M. E., Brown, V. J., Combe, F. J., & Brennan, C. H. (2012). Discrimination reversal and attentional sets in zebrafish (Danio rerio). *Behavioural Brain Research*, 232(1), 264–268. https://doi.org/10.1016/j.bbr.2012.04.035

Pause, B. M., Zlomuzica, A., Kinugawa, K., Mariani, J., Pietrowsky, R., & Dere, E. (2013). Perspectives on Episodic-Like and Episodic Memory. *Frontiers in Behavioral Neuroscience*, 7. https://doi.org/10.3389/fnbeh.2013.00033

Pavlov, P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, 17(3), 136.

Pepperberg, I. M., & Funk, M. S. (1990). Object permanence in four species of psittacine birds: An African Grey parrot (Psittacus erithacus), an Illiger mini macaw (Ara maracana), a parakeet (Melopsittacus undulatus), and a cockatiel (Nymphicus hollandicus). *Animal Learning & Behavior*, 18(1), 97–108. https://doi.org/10.3758/BF03205244

Piaget, J., & Cook, M. (1952). The origins of intelligence in children (Vol. 8). *International Universities Press New York.*

Piaget, J., & Cook, M. T. (1954). The development of object concept. In *The Construction of Reality in the Child. Basic Books*.

Pinto, A., Oates, J., Grutter, A., & Bshary, R. (2011). Cleaner Wrasses Labroides dimidiatus Are More Cooperative in the Presence of an Audience. *Current Biology*, 21(13), 1140–1144. https://doi.org/10.1016/j.cub.2011.05.021

Prat, Y., Bshary, R., & Lotem, A. (2022). Modelling how cleaner fish approach an ephemeral reward task demonstrates a role for ecologically tuned chunking in the evolution of advanced cognition. *PloS Biology*, 20(1), e3001519.

Quiñones, A. E., Leimar, O., Lotem, A., & Bshary, R. (2020). Reinforcement Learning Theory Reveals the Cognitive Requirements for Solving the Cleaner Fish Market Task. *The American Naturalist*, 195(4), 664–677. https://doi.org/10.1086/707519

R Core Team. (2022). R: A Language and Environment for Statistical Computing | *BibSonomy*. https://www.bibsonomy.org/bibtex/7469ffee3b07f9167cf47e7555041ee7

Read, D. W., Manrique, H. M., & Walker, M. J. (2022). On the working memory of humans and great apes: Strikingly similar or remarkably different? *Neuroscience & Biobehavioral Reviews*, 134, 104496. https://doi.org/10.1016/j.neubiorev.2021.12.019

Rhodes, S., & Cowan, N. (2018). Attention in working memory: Attention is needed but it yearns to be free. *Annals of the New York Academy of Sciences*, 1424(1), 52–63. https://doi.org/10.1111/nyas.13652

Salwiczek, L., & Bshary, R. (2011). Cleaner Wrasses Keep Track of the ’When’ and ’What’ in a Foraging Task. *Ethology*, 117(11), 939–948.

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, 72, 116–141. https://doi.org/10.1016/j.jml.2014.01.004

Sovrano, V. A., Baratti, G., & Potrich, D. (2018). A Detour Task in Four Species of Fishes. *Frontiers in Psychology*, 9. https://doi.org/10.3389/fpsyg.2018.02341

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74(11), 1.

Sukhum, K. V., Freiler, M. K., Wang, R., & Carlson, B. A. (2016). The costs of a big brain: Extreme encephalization results in higher energetic demand and reduced hypoxia tolerance in weakly electric African fishes. *Proceedings of the Royal Society B: Biological Sciences*, 283(1845), 20162157. https://doi.org/10.1098/rspb.2016.2157

Tartarelli, G., & Bisconti, M. (2006). Trajectories and Constraints in Brain Evolution in Primates and Cetaceans. *Human Evolution*, 21(3–4), 275–287. https://doi.org/10.1007/s11598-006-9027-4

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, 39(1/4), 212–222.

Triki, Z., & Bshary, R. (2021). Sex differences in the cognitive abilities of a sex-changing fish species Labroides dimidiatus. *Royal Society Open Science*, 8(7), 210239. https://doi.org/10.1098/rsos.210239

Triki, Z., Fong, S., Amcoff, M., Vàsquez-Nilsson, S., & Kolm, N. (2023). Experimental expansion of relative telencephalon size improves the main executive function abilities in guppy. *PNAS Nexus*, 2(6), pgad129. https://doi.org/10.1093/pnasnexus/pgad129

Tulving, E. (1972). Episodic and semantic memory. In Organization of memory (pp. xiii, 423–xiii, 423). *Academic Press.*

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, 53(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2015). Working Memory Delay Activity Predicts Individual Differences in Cognitive Abilities. *Journal of Cognitive Neuroscience*, 27(5), 853–865. https://doi.org/10.1162/jocn_a_00765

Völter, C. J., Mundry, R., Call, J., & Seed, A. M. (2019). Chimpanzees flexibly update working memory contents and show susceptibility to distraction in the self-ordered search task. *Proceedings of the Royal Society B: Biological Sciences*, 286(1907), 20190715. https://doi.org/10.1098/rspb.2019.0715

Wismer, S., Grutter, A., & Bshary, R. (2016). Generalized rule application in bluestreak cleaner wrasse (Labroides dimidiatus): Using predator species as social tools to reduce punishment. *Animal Cognition*, 19(4), 769–778. https://doi.org/10.1007/s10071-016-0975-4

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, 46(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Zhang, S., Bock, F., Si, A., Tautz, J., & Srinivasan, M. V. (2005). Visual working memory in decision making by honey bees. *Proceedings of the National Academy of Sciences*, 102(14), 5250–5255. https://doi.org/10.1073/pnas.0501440102

Zola, S. M., Squire, L. R., Teng, E., Stefanacci, L., Buffalo, E. A., & Clark, R. E. (2000). Impaired recognition memory in monkeys after damage limited to the hippocampal region. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 20(1), 451–463. https://doi.org/10.1523/JNEUROSCI.20-01-00451.2000

Zucca, P., Milos, N., & Vallortigara, G. (2007). Piagetian object permanence and its development in Eurasian jays (Garrulus glandarius). *Animal Cognition*, 10(2), 243–258. https://doi.org/10.1007/s10071-006-0063-2

R.3. Chapter 2

Aellen, M., Dufour, V., & Bshary, R. (2021). Cleaner fish and other wrasse match primates in their ability to delay gratification. *Animal Behaviour*, 176, 125–143. https://doi.org/10.1016/j.anbehav.2021.04.002

Allen, C. (2006). Transitive inference in animals: Reasoning or conditioned associations? In S. Hurley & M. Nudds (Eds.), Rational Animals? (pp. 175–186). *Oxford University Press*. https://doi.org/10.1093/acprof:oso/9780198528272.003.0007

Benard, J., & Giurfa, M. (2004). A Test of Transitive Inferences in Free-Flying Honeybees: Unsuccessful Performance Due to Memory Constraints. *Learning & Memory*, 11(3), 328–336. https://doi.org/10.1101/lm.72204

Binning, S. A., Rey, O., Wismer, S., Triki, Z., Glauser, G., Soares, M. C., & Bshary, R. (2017). Reputation management promotes strategic adjustment of service quality in cleaner wrasse. *Scientific Reports*, 7(1), 8425. https://doi.org/10.1038/s41598-017-07128-5

Bond, A. B., Kamil, A. C., & Balda, R. P. (2003). Social complexity and transitive inference in corvids. *Animal Behaviour*, 65(3), 479–487. https://doi.org/10.1006/anbe.2003.2101

Bond, A. B., Wei, C. A., & Kamil, A. C. (2010). Cognitive representation in transitive inference: A comparison of four corvid species. *Behavioural Processes*, 85(3), 283–292. https://doi.org/10.1016/j.beproc.2010.08.003

Booth, D. J. (1995). Juvenile Groups in a Coral-Reef Damselfish: Density-Dependent Effects on Individual Fitness and Population Demography. *Ecology*, 76(1), 91–106. https://doi.org/10.2307/1940634

Bryant, P. E., & Trabasso, T. (1971). Transitive Inferences and Memory in Young Children. *Nature*, 232(5311), 456–458. https://doi.org/10.1038/232456a0

Bshary, R., Grutter, A. S., Willener, A. S. T., & Leimar, O. (2008). Pairs of cooperating cleaner fish provide better service quality than singletons. *Nature*, 455(7215), Article 7215. https://doi.org/10.1038/nature07184

Buston, P. (2003). Size and growth modification in clownfish. *Nature*, 424(6945). https://doi.org/10.1038/424145a

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395(6699), 272–274. https://doi.org/10.1038/26216

Couvillon, P., & Bitterman, M. (1992). A Conventional Conditioning Analysis of "Transitive Inference" in Pigeons. *Journal of Experimental Psychology Animal Behavior Processes*, 18, 308–310. https://doi.org/10.1037/0097-7403.18.3.308

Daisley, J. N., Vallortigara, G., & Regolin, L. (2010). Logic in an asymmetrical (social) brain: Transitive inference in the young domestic chick. *Social Neuroscience*, 5(3), 309–319. https://doi.org/10.1080/17470910903529795

Daisley, J. N., Vallortigara, G., & Regolin, L. (2021). Low-rank Gallus gallus domesticus chicks are better at transitive inference reasoning. *Communications Biology*, 4(1), Article 1. https://doi.org/10.1038/s42003-021-02855-y

Davis, H. (1992). Transitive inference in rats (Rattus norvegicus). *Journal of Comparative Psychology* (Washington, D.C. : 1983), 106, 342–349. https://doi.org/10.1037//0735-7036.106.4.342

Galizio, A., Doughty, A. H., Williams, D. C., & Saunders, K. J. (2017). Understanding behavior under nonverbal transitive-inference procedures: Stimulus-control-topography analyses. *Behavioural Processes*, 140, 202–215. https://doi.org/10.1016/j.beproc.2017.05.010

Gazes, R. P., Hampton, R. R., & Lourenco, S. F. (2017). Transitive inference of social dominance by human infants. *Developmental Science*, 20(2). https://doi.org/10.1111/desc.12367

Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. Journal of Experimental Psychology: *Animal Behavior Processes*, 7(2), 150.

Grosenick, L., Clement, T. S., & Fernald, R. D. (2007). Fish can infer social rank by observation alone. *Nature*, 445(7126), 429–432.

Grutter, A. (1996). Parasite removal rates by the cleaner wrasse Labroides dimidiatus. *Marine Ecology Progress Series*, 130, 61–70. https://doi.org/10.3354/meps130061

Grutter, A. S., & Bshary, R. (2003). Cleaner wrasse prefer client mucus: Support for partner control mechanisms in cleaning interactions. *Proceedings of the Royal Society of London*. Series B: Biological Sciences, 270(suppl_2). https://doi.org/10.1098/rsbl.2003.0077

Guez, D., & Audley, C. (2013). Transitive or not: A critical appraisal of transitive inference in animals. *Ethology*, 119(9), 703–726.

Hotta, T., Ueno, K., Hataji, Y., Kuroshima, H., Fujita, K., & Kohda, M. (2020). Transitive inference in cleaner wrasses (Labroides dimidiatus). *PloS One*, 15(8), e0237817.

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, 2(4), 705–715. https://doi.org/10.1111/j.1756-8765.2010.01089.x

Jensen, G., Alkan, Y., Munoz, F., Ferrera, V. P., & Terrace, H. S. (2016). Transitive inference in humans and rhesus macaques after massed training of the last two list items. *bioRxiv*, 055335.

Kamil, A. C. (1998). On the proper definition of cognitive ethology. *Animal cognition in nature* (pp. 1–28). Elsevier.

Kohda, M., Hotta, T., Takeyama, T., Awata, S., Tanaka, H., Asai, J., & Jordan, A. L. (2019). If a fish can pass the mark test, what are the implications for consciousness and self-awareness testing in animals? *PLoS Biology*, 17(2), e3000021.

Kohda, M., Sogawa, S., Jordan, A. L., Kubo, N., Awata, S., Satoh, S., Kobayashi, T., Fujita, A., & Bshary, R. (2022). Further evidence for the capacity of mirror self-recognition in cleaner fish and the significance of ecologically relevant marks. *PLoS Biology*, 20(2), e3001529.

Lazareva, O. F., Paxton Gazes, R., Elkins, Z., & Hampton, R. (2020). Associative models fail to characterize transitive inference performance in rhesus monkeys (Macaca mulatta). *Learning & Behavior*, 48(1), 135–148.

Lazareva, O. F., Smirnova, A. A., Bagozkaja, M. S., Zorina, Z. A., Rayevsky, V. V., & Wasserman, E. A. (2004). Transitive responding in hooded crows requires linearly ordered stimuli. *Journal of the Experimental Analysis of Behavior*, 82(1), 1–19.

Lazareva, O. F., & Wasserman, E. A. (2006). Effect of stimulus orderability and reinforcement history on transitive responding in pigeons. *Behavioural Processes*, 72(2), 161–172.

MacLean, E. L., Merritt, D. J., & Brannon, E. M. (2008). Social complexity predicts transitive reasoning in prosimian primates. *Animal Behaviour*, 76(2), 479–486.

McGonigle, B., & Chalmers, M. (1984). The selective impact of question form and input mode on the symbolic distance effect in children. *Journal of Experimental Child Psychology,* 37(3), 525–554.

Mikolasch, S., Kotrschal, K., & Schloegl, C. (2013). Transitive inference in jackdaws (Corvus monedula). *Behavioural Processes*, 92, 113–117.

Oliveira, R. F., & Bshary, R. (2021). Expanding the concept of social behavior to interspecific interactions. *Ethology*, 127(10), 758–773. https://doi.org/10.1111/eth.13194

Pavlov, P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, 17(3), 136.

Paz-y-Miño C, G., Bond, A. B., Kamil, A. C., & Balda, R. P. (2004). Pinyon jays use transitive inference to predict social dominance. *Nature*, 430(7001), 778–781. https://doi.org/10.1038/nature02723

Pinto, A., Oates, J., Grutter, A., & Bshary, R. (2011). Cleaner wrasses Labroides dimidiatus are more cooperative in the presence of an audience. *Current Biology*, 21(13), 1140–1144.

Robertson, D. R. (1972). Social control of sex reversal in a coral-reef fish. *Science*, 177(4053), 1007–1009.

Roche, D. G., Jornod, M., Douet, V., Grutter, A. S., & Bshary, R. (2021). Client fish traits underlying variation in service quality in a marine cleaning mutualism. *Animal Behaviour*, 175, 137–151.

Romonchuk Wayne Joseph. (1995). *The role of memory in cache-recovery in Clark's nutcrackers and pinyon jays*.

Sakai, Y., & Kohda, M. (2001). Spawning timing of the cleaner wrasse, Labroides dimidiatus, on a warm temperate rocky shore. *Ichthyological Research*, 48, 23–30.

Salwiczek, L. H., Prétôt, L., Demarta, L., Proctor, D., Essler, J., Pinto, A. I., Wismer, S., Stoinski, T., Brosnan, S. F., & 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(11), e49068.

Shettleworth, S. J. (2009). Cognition, evolution, and behavior. *Oxford university press.*

Silverman, J. L., Gastrell, P. T., Karras, M. N., Solomon, M., & Crawley, J. N. (2015). Cognitive abilities on transitive inference using a novel touchscreen technology for mice. *Cerebral Cortex*, 25(5), 1133–1142.

Taborsky, B., & Oliveira, R. F. (2012). Social competence: An evolutionary approach. *Trends in Ecology & Evolution*, 27(12), 679–688. https://doi.org/10.1016/j.tree.2012.09.003

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology*, 39(1/4), 212–222.

Tibbetts, E. A., Agudelo, J., Pandit, S., & Riojas, J. (2019). Transitive inference in Polistes paper wasps. *Biology Letters*, 15(5), 20190015.

Triki, Z., & Bshary, R. (2019). Fluctuations in coral reef fish densities after environmental disturbances on the northern Great Barrier Reef. *PeerJ*, 7, e6720.

Triki, Z., Li Richter, X.-Y., Demairé, C., Kurokawa, S., & Bshary, R. (2022). Marine cleaning mutualism defies standard logic of supply and demand. *The American Naturalist*, 199(4), 455–467.

Triki, Z., Wismer, S., Rey, O., Ann Binning, S., Levorato, E., & Bshary, R. (2019). Biological market effects predict cleaner fish strategic sophistication. *Behavioral Ecology*, 30(6), 1548–1557.

Tromp, D., Meunier, H., & Roeder, J. J. (2015). Transitive inference in two lemur species (Eulemur macaco and Eulemur fulvus). *American Journal of Primatology*, 77(3), 338–345.

Vasconcelos, M. (2008). Transitive inference in non-human animals: An empirical and theoretical analysis. *Behavioural Processes*, 78(3), 313–334.

Von Fersen, L., Wynne, C. D., Delius, J. D., & Staddon, J. E. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 17(3), 334.

Weiß, B. M., Kehmeier, S., & Schloegl, C. (2010). Transitive inference in free-living greylag geese, Anser anser. *Animal Behaviour*, 79(6), 1277–1283.

White, S. L., & Gowan, C. (2013). Brook trout use individual recognition and transitive inference to determine social rank. *Behavioral Ecology*, 24(1), 63–69.

Wismer, S., Grutter, A., & Bshary, R. (2016). Generalized rule application in bluestreak cleaner wrasse (Labroides dimidiatus): Using predator species as social tools to reduce punishment. *Animal Cognition*, 19(4), 769–778. https://doi.org/10.1007/s10071-016-0975-4

Woocher, F. D., Glass, A. L., & Holyoak, K. J. (1978). Positional discriminability in linear orderings. *Memory & Cognition*, 6(2), 165–173.

Wu, X., & Levy, W. B. (2001). Simulating symbolic distance effects in the transitive inference problem. *Neurocomputing*, 38, 1603–1610.

Wynne, C. D., Von Fersen, L., & Staddon, J. E. (1992). Pigeons' inferences are transitive and the outcome of elementary conditioning principles: A response.

Zentall, T. R., Peng, D., & Miles, L. (2019). Transitive inference in pigeons may result from differential tendencies to reject the test stimuli acquired during training. *Animal Cognition*, 22(5), 619–624.

R.4. Chapter 3

Aellen, M., Burkart, J. M., & Bshary, R. (2022). No evidence for general intelligence in a fish. *Ethology*, 128(5), 424–436. https://doi.org/10.1111/eth.13275

Alloway, T., & Alloway, R. (2009). The efficacy of working memory training in improving crystallized intelligence. *Nature Precedings*. https://doi.org/10.1038/npre.2009.3697.1

Anderson, M. R. (2012). Comprehension of object permanence and single transposition in gibbons. *Behaviour*, 149(3–4), 441–459.

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, 112, 91–106. https://doi.org/10.1016/j.cortex.2018.05.015

Barth, J., & Call, J. (2006). Tracking the displacement of objects: A series of tasks with great apes (Pan troglodytes, Pan paniscus, Gorilla gorilla, and Pongo pygmaeus) and young children (Homo sapiens). Journal of Experimental Psychology. *Animal Behavior Processes*, 32(3), 239–252. https://doi.org/10.1037/0097-7403.32.3.239

Beran, M. J., & Minahan, M. F. (2000). Monitoring spatial transpositions by bonobos (Pan paniscus) and Chimpanzees (P. troglodytes). *International Journal of Comparative Psychology*, 13(1–2), 1–15.

Bonin, L., & Bshary, R. (2023). In the absence of extensive initial training, cleaner wrasse Labroides dimidiatus fail a transitive inference task. *Plos One*, 18(6), e0287402.

Bshary, R., & Noë, R. (2003). Biological markets: The ubiquitous influence of partner choice on the dynamics of cleaner fish–client reef fish interactions.

Bshary, R., & Triki, Z. (2022). Fish ecology and cognition: Insights from studies on wild and wild-caught teleost fishes. *Current Opinion in Behavioral Sciences*, 46, 101174. https://doi.org/10.1016/j.cobeha.2022.101174

Caicoya, Á. L., Amici, F., Ensenyat, C., & Colell, M. (2019). Object permanence in Giraffa camelopardalis: First steps in giraffes' physical cognition. *Journal of Comparative Psychology*, 133(2), 207.

Caicoya, A. L., Amici, F., Ensenyat, C., & Colell, M. (2021). Comparative cognition in three understudied ungulate species: European bison, forest buffalos and giraffes. *Frontiers in Zoology*, 18(1), 30. https://doi.org/10.1186/s12983-021-00417-w

Call, J. (2001). Object permanence in orangutans (Pongo pygmaeus), chimpanzees (Pan troglodytes), and children (Homo sapiens). *Journal of Comparative Psychology*, 115(2), 159.

Colbourne, J. A. (2018). Object permanence in the racoon (Procyon lotor).

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552. https://doi.org/10.1016/j.tics.2003.10.005

Côté, I. (2000). Evolution and ecology of cleaning symbioses in the sea. *Oceanography and Marine Biology.*

Cowan, N. (2022). Working Memory Development: A 50-Year Assessment of Research and Underlying Theories. *Cognition*, 224, 105075. https://doi.org/10.1016/j.cognition.2022.105075

De Blois, S. T., Novak, M. A., & Bond, M. (1998). Object permanence in orangutans (Pongo pygmaeus) and squirrel monkeys (Saimiri sciureus). *Journal of Comparative Psychology*, 112(2), 137.

Diamond, A. (2013). Executive Functions*.* *Annual Review of Psychology*, 64(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dumas, C., & Wilkie, D. M. (1995). Object permanence in ring doves (Streptopelia risoria). *Journal of Comparative Psychology*, 109(2), 142–150. https://doi.org/10.1037/0735-7036.109.2.142

Engle, R. W., Laughlin, J. E., Tuholski, S. W., & Conway, A. R. A. (1999). Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *Journal of Experimental Psychology: General*, 128(3), 309–331.

Fonseca-Azevedo, K., & Herculano-Houzel, S. (2012). Metabolic constraint imposes tradeoff between body size and number of brain neurons in human evolution. *Proceedings of the National Academy of Sciences*, 109(45), 18571–18576. https://doi.org/10.1073/pnas.1206390109

Gagnon, S., & Doré, F. Y. (1992). Search behavior in various breeds of adult dogs (Canis familiaris): Object permanence and olfactory cues. *Journal of Comparative Psychology*, 106(1), 58.

Grutter, A. S. (2004). Cleaner Fish Use Tactile Dancing Behavior as a Preconflict Management Strategy. *Current Biology*, 14(12), 1080–1083. https://doi.org/10.1016/j.cub.2004.05.048

Guadagno, A., & Triki, Z. (2024). Executive functions and brain morphology of male and female dominant and subordinate cichlid fish*.* *Brain and Behavior*, 14(5), e3484. https://doi.org/10.1002/brb3.3484

Helfman, G. S., & Schultz, E. T. (1984). Social transmission of behavioural traditions in a coral reef fish. *Animal Behaviour*, 32(2), 379–384. https://doi.org/10.1016/S0003-3472(84)80272-9

Hoffmann, A., Rüttler, V., & Nieder, A. (2011). Ontogeny of object permanence and object tracking in the carrion crow, Corvus corone. *Animal Behaviour*, 82(2), 359–367. https://doi.org/10.1016/j.anbehav.2011.05.012

Hotta, T., Ueno, K., Hataji, Y., Kuroshima, H., Fujita, K., & Kohda, M. (2020). Transitive inference in cleaner wrasses (Labroides dimidiatus). *PloS One*, 15(8), e0237817.

Huffeldt, N. P. (2020). Performance of horned puffins (Fratercula corniculata) on an object permanence task. *Behavioural Processes*, 181, 104274. https://doi.org/10.1016/j.beproc.2020.104274

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, 2(4), 705–715. https://doi.org/10.1111/j.1756-8765.2010.01089.x

Jaakkola, K. (2014). Do animals understand invisible displacement? A critical review. *Journal of Comparative Psychology*, 128(3), 225–239. https://doi.org/10.1037/a0035675

Jaakkola, K., Guarino, E., Rodriguez, M., Erb, L., & Trone, M. (2010). What do dolphins (Tursiops truncatus) understand about hidden objects? *Animal Cognition*, 13(1), 103–120. https://doi.org/10.1007/s10071-009-0250-z

Johnson, C. M., Sullivan, J., Buck, C. L., Trexel, J., & Scarpuzzi, M. (2015). Visible and invisible displacement with dynamic visual occlusion in bottlenose dolphins (Tursiops spp). *Animal Cognition*, 18(1), 179–193. https://doi.org/10.1007/s10071-014-0788-2

Kamil, A. C. (1998). On the proper definition of cognitive ethology. *In Animal cognition in nature* (pp. 1–28). Elsevier.

Kohda, M., Bshary, R., Kubo, N., Awata, S., Sowersby, W., Kawasaka, K., Kobayashi, T., & Sogawa, S. (2023). Cleaner fish recognize self in a mirror via self-face recognition like humans. *Proceedings of the National Academy of Sciences*, 120(7), e2208420120.

Kohda, M., Hotta, T., Takeyama, T., Awata, S., Tanaka, H., Asai, J., & Jordan, A. L. (2019). If a fish can pass the mark test, what are the implications for consciousness and self-awareness testing in animals? *PLoS Biology*, 17(2), e3000021.

Kohda, M., Sogawa, S., Jordan, A. L., Kubo, N., Awata, S., Satoh, S., Kobayashi, T., Fujita, A., & Bshary, R. (2022). Further evidence for the capacity of mirror self-recognition in cleaner fish and the significance of ecologically relevant marks. *PLoS Biology*, 20(2), e3001529.

Kolata, S., Light, K., Townsend, D. A., Hale, G., Grossman, H. C., & Matzel, L. D. (2005). Variations in working memory capacity predict individual differences in general learning abilities among genetically diverse mice. *Neurobiology of Learning and Memory*, 84(3), 241–246. https://doi.org/10.1016/j.nlm.2005.07.006

Majecka, K., & Pietraszewski, D. (2018). Where's the cookie? The ability of monkeys to track object transpositions. *Animal Cognition*, 21(4), 603–611. https://doi.org/10.1007/s10071-018-1195-x

Marhounová, L., Frynta, D., Fuchs, R., & Landová, E. (2017). Object permanence in the food-storing coal tit (Periparus ater) and the non-storing great tit (Parus major): Is the mental representation required? *Journal of Comparative Psychology* (Washington, D.C.: 1983), 131(2), 115–127. https://doi.org/10.1037/com0000061

Miller, D. (2019). Object Permanence in Asian Elephants (Elephas maximus). https://academicworks.cuny.edu/hc_sas_etds/406/

Nawroth, C., von Borell, E., & Langbein, J. (2015). Object permanence in the dwarf goat (Capra aegagrus hircus): Perseveration errors and the tracking of complex movements of hidden objects. *Applied Animal Behaviour Science*, 167, 20–26.

Obozova, T. A., Smirnova, A. A., & Zorina, Z. A. (2011). Observational Learning in a Glaucous-winged Gull Natural Colony. *International Journal of Comparative Psychology*, 24(2). https://doi.org/10.46867/IJCP.2011.24.02.02

Pavlov, P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, 17(3), 136.

Pepperberg, I. M. (2015). Reply to Jaakkola (2014): "Do animals understand invisible displacement? A critical review". *Journal of Comparative Psychology*, 129(2), 198–201.

Pepperberg, I. M., & Funk, M. S. (1990). Object permanence in four species of psittacine birds: An African Grey parrot (Psittacus erithacus), an Illiger mini macaw (Ara maracana), a parakeet (Melopsittacus undulatus), and a cockatiel (Nymphicus hollandicus). *Animal Learning & Behavior*, 18(1), 97–108. https://doi.org/10.3758/BF03205244

Pérez-Manrique, A., & Gomila, A. (2018). The role of previous visual experience in the acquisition of object permanence skills in bottlenose dolphins: A pilot study*. Tesis Psicológica*, 13(2), 118–132.

Piaget, J., & Cook, M. (1952). The origins of intelligence in children (Vol. 8). *International Universities Press New York.*

Piaget, J., & Cook, M. T. (1954). The development of object concept. In *The Construction of Reality in the Child. Basic Books.*

Pollok, B., Prior, H., & Güntürkün, O. (2000). Development of object permanence in food-storing magpies (Pica pica). Journal of Comparative Psychology, 114(2), 148–157. https://doi.org/10.1037/0735-7036.114.2.148

Reid, S. L. (1996). Search for hidden objects by pigeons: Place learning vs" object permanence". University of Ottawa (Canada).

Robertson, D. R. (1972). Social control of sex reversal in a coral-reef fish. *Science*, 177(4053), 1007–1009.

Sakai, Y., Kohda, M., & Kuwamura, T. (2001). Effect of changing harem on timing of sex change in female cleaner fish Labroides dimidiatus. *Animal Behaviour*, 62(2), 251–257.

Shettleworth, S. J. (2009). Cognition, evolution, and behavior. *Oxford university press.*

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, 72, 116–141. https://doi.org/10.1016/j.jml.2014.01.004

Skinner, B. F. (1937). Two Types of Conditioned Reflex: A Reply to Konorski and Miller. *The Journal of General Psychology*, 16(1), 272–279. https://doi.org/10.1080/00221309.1937.9917951

Sukhum, K. V., Freiler, M. K., Wang, R., & Carlson, B. A. (2016). The costs of a big brain: Extreme encephalization results in higher energetic demand and reduced hypoxia tolerance in weakly electric African fishes. *Proceedings of the Royal Society B: Biological Sciences*, 283(1845), 20162157. https://doi.org/10.1098/rspb.2016.2157

Thorndike, E. L. (1911). Animal intelligence: Experimental studies (pp. viii, 297). *Macmillan Press*. https://doi.org/10.5962/bhl.title.55072

Triki, Z., Fong, S., Amcoff, M., Vàsquez-Nilsson, S., & Kolm, N. (2023). Experimental expansion of relative telencephalon size improves the main executive function abilities in guppy. *PNAS Nexus*, 2(6), pgad129. https://doi.org/10.1093/pnasnexus/pgad129

Trösch, M., Flamand, A., Chasles, M., Nowak, R., Calandreau, L., & Lansade, L. (2020). Horses Solve Visible but Not Invisible Displacement Tasks in an Object Permanence Paradigm. *Frontiers in Psychology*, 11. https://doi.org/10.3389/fpsyg.2020.562989

Truskanov, N., Emery, Y., & Bshary, R. (2020). Juvenile cleaner fish can socially learn the consequences of cheating. *Nature Communications*, 11(1), 1159. https://doi.org/10.1038/s41467-020-14712-3

Ujfalussy, D. J., Miklósi, Á., & Bugnyar, T. (2013). Ontogeny of object permanence in a non-storing corvid species, the jackdaw (Corvus monedula). *Animal Cognition,* 16(3), 405–416. https://doi.org/10.1007/s10071-012-0581-z

Uzgiris, I. C., & Hunt, J. (1975). Assessment in infancy: Ordinal scales of psychological development.

Wismer, S., Grutter, A., & Bshary, R. (2016). Generalized rule application in bluestreak cleaner wrasse (Labroides dimidiatus): Using predator species as social tools to reduce punishment. *Animal Cognition*, 19(4), 769–778. https://doi.org/10.1007/s10071-016-0975-4

Yao, R., & Garcia-Pelegrin, E. (2024). Oriental pied hornbills (Anthracoceros albirostris) solve invisible displacement tasks in a test of Piagetian object permanence. *Biology Letters*, 20(1), 20230547. https://doi.org/10.1098/rsbl.2023.0547

Zentall, T., & Raley, O. (2018). Object Permanence in the Pigeon (Columba livia): Insertion of a Delay Prior to Choice Facilitates Visibleand Invisible-Displacement Accuracy. *Journal of Comparative Psychology*, 133. https://doi.org/10.1037/com0000134

Zucca, P., Milos, N., & Vallortigara, G. (2007). Piagetian object permanence and its development in Eurasian jays (Garrulus glandarius). *Animal Cognition*, 10(2), 243–258. https://doi.org/10.1007/s10071-006-0063-2

R.5. General discussion and conclusion

Ackerman, P. L., Beier, M. E., & Boyle, M. D. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, 131(4), 567–589. https://doi.org/10.1037/0096-3445.131.4.567

Aellen, M., Burkart, J. M., & Bshary, R. (2022). No evidence for general intelligence in a fish. *Ethology,* 128(5), 424–436. https://doi.org/10.1111/eth.13275

Aellen, M., Dufour, V., & Bshary, R. (2021). Cleaner fish and other wrasse match primates in their ability to delay gratification. *Animal Behaviour*, 176, 125–143. https://doi.org/10.1016/j.anbehav.2021.04.002

Aellen, M., Siebeck, U. E., & Bshary, R. (2022). Cleaner wrasse Labroides dimidiatus perform above chance in a “matching-to-sample” experiment. *PloS One*, 17(1), e0262351. https://doi.org/10.1371/journal.pone.0262351

Alloway, T., & Alloway, R. (2009). The efficacy of working memory training in improving crystallized intelligence. *Nature Precedings*. https://doi.org/10.1038/npre.2009.3697.1

Anderson, B. (1993). Evidence from the rat for a general factor that underlies cognitive performance and that relates to brain size: Intelligence? *Neuroscience Letters*, 153(1), 98–102. https://doi.org/10.1016/0304-3940(93)90086-Z

Bachevalier, J., & Mishkin, M. (1986). Visual recognition impairment follows ventromedial but not dorsolateral prefrontal lesions in monkeys. *Behavioural Brain Research*, 20(3), 249–261. https://doi.org/10.1016/0166-4328(86)90225-1

Baddeley, A. (1986). Working memory (pp. xi, 289). *Clarendon Press/Oxford University Press.*

Beaman, C. P. (2010). Working Memory and Working Attention: What Could Possibly Evolve? *Current Anthropology*, 51(S1), S27–S38. https://doi.org/10.1086/650297

Beran, M. J., Washburn, D. A., & Rumbaugh, D. M. (2007). A Stroop-Like Effect in Color-Naming of Color-Word Lexigrams by a Chimpanzee (Pan Troglodyte). *The Journal of General Psychology*, 134(2), 217–228. https://doi.org/10.3200/GENP.134.2.217-228

Bouchard, T. J. (2014). Genes, Evolution and Intelligence. *Behavior Genetics*, 44(6), 549–577. https://doi.org/10.1007/s10519-014-9646-x

Brzezicka, A., Kamiński, J., Reed, C. M., Chung, J. M., Mamelak, A. N., & Rutishauser, U. (2019). Working Memory Load-related Theta Power Decreases in Dorsolateral Prefrontal Cortex Predict Individual Differences in Performance. *Journal of Cognitive Neuroscience*, 31(9), 1290–1307. https://doi.org/10.1162/jocn_a_01417

Bshary, R. (2002). Biting cleaner fish use altruism to deceive image–scoring client reef fish. *Proceedings of the Royal Society of London*. Series B: Biological Sciences, 269(1505), 2087–2093. https://doi.org/10.1098/rspb.2002.2084

Bshary, R., & Würth, M. (2001). Cleaner fish Labroides dimidiatus manipulate client reef fish by providing tactile stimulation. *Proceedings of the Royal Society B*: Biological Sciences, 268(1475), 1495–1501. https://doi.org/10.1098/rspb.2001.1495

Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017). The evolution of general intelligence. *Behavioral and Brain Sciences*, 40, e195.

Caicoya, A. L., Amici, F., Ensenyat, C., & Colell, M. (2021). Comparative cognition in three understudied ungulate species: European bison, forest buffalos and giraffes. *Frontiers in Zoology*, 18(1), 30. https://doi.org/10.1186/s12983-021-00417-w

Cattell, R. B. (1941). Some theoretical issues in adult intelligence. *Psychological Bulletin*, 38(7), 592.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40(3), 153–193. https://doi.org/10.1037/h0059973

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J.-D. (2017). The Distributed Nature of Working Memory. *Trends in Cognitive Sciences*, 21(2), 111–124. https://doi.org/10.1016/j.tics.2016.12.007

Collette, F., & Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience & Biobehavioral Reviews*, 26(2), 105–125. https://doi.org/10.1016/S0149-7634(01)00063-X

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552. https://doi.org/10.1016/j.tics.2003.10.005

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, 24(1), 154–176. https://doi.org/10.1017/S0140525X0161392X

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, 24(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cowan, N. (2022). Working Memory Development: A 50-Year Assessment of Research and Underlying Theories. *Cognition*, 224, 105075. https://doi.org/10.1016/j.cognition.2022.105075

Cowan, N., Bao, C., Bishop-Chrzanowski, B. M., Costa, A. N., Greene, N. R., Guitard, D., Li, C., Musich, M. L., & Ünal, Z. E. (2024). The Relation Between Attention and Memory. *Annual Review of Psychology*, 75(1), 183–214. https://doi.org/10.1146/annurev-psych-040723-012736

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(2), 115–124. https://doi.org/10.1159/000102973

D'Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working Memory. *Annual Review of Psychology*, 66(1), 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A. (2020). Chapter 19—Executive functions. In A. Gallagher, C. Bulteau, D. Cohen, & J. L. Michaud (Eds.), *Handbook of Clinical Neurology* (Vol. 173, pp. 225–240). Elsevier. https://doi.org/10.1016/B978-0-444-64150-2.00020-4

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2021). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology*: General, 150(2), 242–275. https://doi.org/10.1037/xge0000783

Engle, R. W. (2010). Role of Working-Memory Capacity in Cognitive Control. *Current Anthropology*, 51(S1), S17–S26. https://doi.org/10.1086/650572

Engle, R. W. (2018). Working Memory and Executive Attention: A Revisit. *Perspectives on Psychological Science*, 13(2), 190–193. https://doi.org/10.1177/1745691617720478

Engle, R. W., Laughlin, J. E., Tuholski, S. W., & Conway, A. R. A. (1999). Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *Journal of Experimental Psychology:General*, 128(3), 309–331.

Ester, E. F., & Nouri, A. (2023). Internal selective attention is delayed by competition between endogenous and exogenous factors. *iScience*, 26(7). https://doi.org/10.1016/j.isci.2023.107259

Evans, T. A., & Beran, M. J. (2007). Chimpanzees use self-distraction to cope with impulsivity. *Biology Letters*, 3(6), 599–602. https://doi.org/10.1098/rsbl.2007.0399

Font, E. (2019). Rapid learning of a spatial memory task in a lacertid lizard (Podarcis liolepis). *Behavioural Processes*, 169, 103963.

Frischkorn, G. T., Schubert, A.-L., & Hagemann, D. (2019). Processing speed, working memory, and executive functions: Independent or inter-related predictors of general intelligence*. Intelligence*, 75, 95–110. https://doi.org/10.1016/j.intell.2019.05.003

Griffin, I. C., & Nobre, A. C. (2003). Orienting Attention to Locations in Internal Representations. *Journal of Cognitive Neuroscience*, 15(8), 1176–1194. https://doi.org/10.1162/089892903322598139

Grutter, A. (1996). Parasite removal rates by the cleaner wrasse Labroides dimidiatus. *Marine Ecology Progress Series*, 130, 61–70. https://doi.org/10.3354/meps130061

Grutter, A. S. (2004). Cleaner Fish Use Tactile Dancing Behavior as a Preconflict Management Strategy. *Current Biology*, 14(12), 1080–1083. https://doi.org/10.1016/j.cub.2004.05.048

Grutter, A. S., & Bshary, R. (2003). Cleaner wrasse prefer client mucus: Support for partner control mechanisms in cleaning interactions. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(suppl_2). https://doi.org/10.1098/rsbl.2003.0077

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, 18(10), 1279–1296.

Hotta, T., Ueno, K., Hataji, Y., Kuroshima, H., Fujita, K., & Kohda, M. (2020). Transitive inference in cleaner wrasses (Labroides dimidiatus). *PloS One*, 15(8), e0237817.

Huffeldt, N. P. (2020). Performance of horned puffins (Fratercula corniculata) on an object permanence task. *Behavioural Processes*, 181, 104274. https://doi.org/10.1016/j.beproc.2020.104274

Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science*, 2(4), 705–715. https://doi.org/10.1111/j.1756-8765.2010.01089.x

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130(2), 169.

Kolata, S., Light, K., Townsend, D. A., Hale, G., Grossman, H. C., & Matzel, L. D. (2005). Variations in working memory capacity predict individual differences in general learning abilities among genetically diverse mice. *Neurobiology of Learning and Memory*, 84(3), 241–246. https://doi.org/10.1016/j.nlm.2005.07.006

Kong, G., & Fougnie, D. (2019). Visual search within working memory. Journal of Experimental Psychology. *General*, 148(10), 1688–1700. https://doi.org/10.1037/xge0000555

Krauzlis, R. J., Bogadhi, A. R., Herman, J. P., & Bollimunta, A. (2018). Selective attention without a neocortex. *Cortex*; a Journal Devoted to the Study of the Nervous System and Behavior, 102, 161–175. https://doi.org/10.1016/j.cortex.2017.08.026

Libben, M., & Titone, D. (2008). The role of awareness and working memory in human transitive inference. *Behavioural Processes*, 77(1), 43–54. https://doi.org/10.1016/j.beproc.2007.06.006

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163.

Manrique, H. M., & Walker, M. J. (2017). Early Evolution of Human Memory: Great Apes, Tool-making, and Cognition. *Springer International Publishing*. https://doi.org/10.1007/978-3-319-64447-9

Mueller, P. (2023). Bird of a feather: Exploring social facilitation effects on learning and suboptimal choice in pigeons (Columba livia). *University of Kentucky Libraries*. https://doi.org/10.13023/ETD.2023.177

Noble, D. W. A., Byrne, R. W., & Whiting, M. J. (2014). Age-dependent social learning in a lizard. *Biology Letters*, 10(7), 20140430. https://doi.org/10.1098/rsbl.2014.0430

Nyborg, H. (2005). Sex-related differences in general intelligence g, brain size, and social status. *Personality and Individual Differences*, 39(3), 497–509. https://doi.org/10.1016/j.paid.2004.12.011

Piaget, J. (1954). *The Construction of Reality in the Child.*

Piaget, J., & Cook, M. (1952). The origins of intelligence in children (Vol. 8). *International Universities Press New York*.

Raghunathan, R. S., Musci, R. J., Knudsen, N., & Johnson, S. B. (2023). What children do while they wait: The role of self-control strategies in delaying gratification. *Journal of Experimental Child Psychology*, 226, 105576. https://doi.org/10.1016/j.jecp.2022.105576

Ramus, F. (2017). General intelligence is an emerging property, not an evolutionary puzzle. *Behavioral and Brain Sciences*, 40, e217. https://doi.org/10.1017/S0140525X1600176X

Robertson, D. R. (1972). Social control of sex reversal in a coral-reef fish. *Science*, 177(4053), 1007–1009.

Salehinejad, M. A., Ghanavati, E., Rashid, M. H. A., & Nitsche, M. A. (2021). Hot and cold executive functions in the brain: A prefrontal-cingular network. *Brain and Neuroscience Advances*, 5, 23982128211007769. https://doi.org/10.1177/23982128211007769

Shettleworth, S. J. (2009). Cognition, evolution, and behavior. *Oxford university press*.

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, 72, 116–141. https://doi.org/10.1016/j.jml.2014.01.004

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *The Journal of Applied Psychology*, 51(3), 300–304. https://doi.org/10.1037/h0020586

Spearman, C. (1904). “General Intelligence,” Objectively Determined And Measured. *American Journal of Psychology*, 15(2), 201–293.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74(11), 1.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. J*ournal of Experimental Psychology*, 18(6), 643.

Takeuchi, H., Taki, Y., Nouchi, R., Yokoyama, R., Kotozaki, Y., Nakagawa, S., Sekiguchi, A., Iizuka, K., Hanawa, S., Araki, T., Miyauchi, C. M., Sakaki, K., Sassa, Y., Nozawa, T., Ikeda, S., Yokota, S., Daniele, M., & Kawashima, R. (2018). General intelligence is associated with working memory-related brain activity: New evidence from a large sample study. *Brain Structure and Function*, 223(9), 4243–4258. https://doi.org/10.1007/s00429-018-1747-5

Tibbetts, E. A., Agudelo, J., Pandit, S., & Riojas, J. (2019). Transitive inference in Polistes paper wasps. *Biology Letters*, 15(5), 20190015.

Triki, Z., Fong, S., Amcoff, M., Vàsquez-Nilsson, S., & Kolm, N. (2023). Experimental expansion of relative telencephalon size improves the main executive function abilities in guppy. *PNAS Nexus*, 2(6), pgad129. https://doi.org/10.1093/pnasnexus/pgad129

Truskanov, N., Emery, Y., & Bshary, R. (2020). Juvenile cleaner fish can socially learn the consequences of cheating. *Nature Communications*, 11(1), 1159. https://doi.org/10.1038/s41467-020-14712-3

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2015). Working Memory Delay Activity Predicts Individual Differences in Cognitive Abilities. *Journal of Cognitive Neuroscience*, 27(5), 853–865. https://doi.org/10.1162/jocn_a_00765

Urcuioli, P. J., Vu, K.-P. L., & Proctor, R. W. (2005). A Simon Effect in Pigeons. *Journal of Experimental Psychology: General*, 134(1), 93–107. https://doi.org/10.1037/0096-3445.134.1.93

van der Linden, D., Dunkel, C. S., & Madison, G. (2017). Sex differences in brain size and general intelligence (g). *Intelligence*, 63, 78–88. https://doi.org/10.1016/j.intell.2017.04.007

Van Ede, F., & Nobre, A. C. (2023). Turning Attention Inside Out: How Working Memory Serves Behavior. *Annual Review of Psychology*, 74(1), 137–165. https://doi.org/10.1146/annurev-psych-021422-041757

Washburn, D. A., Gulledge, J. P., James, F., & Rumbaugh, D. M. (2007). A Species Difference in Visuospatial Working Memory:Does Language Link “What” with “Where”? *International Journal of Comparative Psychology*, 20(1). https://doi.org/10.46867/IJCP.2007.20.01.02

White, S. L., & Gowan, C. (2013). Brook trout use individual recognition and transitive inference to determine social rank. *Behavioral Ecology*, 24(1), 63–69.

Wolff, M., Krönke, K.-M., & Goschke, T. (2016). Trait self-control is predicted by how reward associations modulate Stroop interference. *Psychological Research*, 80(6), 944–951. https://doi.org/10.1007/s00426-015-0707-4

Wong, J. H., Peterson, M. S., & Thompson, J. C. (2008). Visual working memory capacity for objects from different categories: A face-specific maintenance effect. *Cognition*, 108(3), 719–731. https://doi.org/10.1016/j.cognition.2008.06.006

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, 46(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yuan, P., & Raz, N. (2014). Prefrontal cortex and executive functions in healthy adults: A meta-analysis of structural neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 42, 180–192. https://doi.org/10.1016/j.neubiorev.2014.02.005

Zentall, T., & Raley, O. (2018). Object Permanence in the Pigeon (Columba livia): Insertion of a Delay Prior to Choice Facilitates Visibleand Invisible-Displacement Accuracy. *Journal of Comparative Psychology*, 133. https://doi.org/10.1037/com0000134

Zola, S. M., Squire, L. R., Teng, E., Stefanacci, L., Buffalo, E. A., & Clark, R. E. (2000). Impaired recognition memory in monkeys after damage limited to the hippocampal region. *The Journal of Neuroscience*, 20(1), 451–463. https://doi.org/10.1523/JNEUROSCI.20-01-00451.2000