

1 **No evidence for general intelligence in a fish**

2 **Mélisande Aellen ^{1*}, Judith M. Burkart ², Redouan Bshary ¹**

3 **Author summary**

4 ¹ *University of Neuchâtel, Department of Behavioural Ecology, Rue Emile-Argand 11, 2000*
5 *Neuchâtel, Switzerland.*

6 ² *University of Zürich, Anthropological Institut and Museum, Winterthurerstrasse 190, 8057*
7 *Zürich, Switzerland.*

8 **Author contributions**

9 All authors conceived the project and designed the experiments. M. A. collected the data
10 (with the help of Yasmin Emery; see acknowledgements). M. A. performed the analyses and
11 all authors analysed the results. All authors wrote the manuscript and revised it.

12 **Declaration of interests**

13 The authors declare no conflict of interest.

14 **Abstract**

15 **Differences in human general intelligence or reasoning ability can be quantified with the**
16 **psychometric factor g , because individual performance across cognitive tasks is positively**
17 **correlated. g also emerges in mammals and birds, is correlated with brain size and may**
18 **similarly reflect general reasoning ability and behavioural flexibility in these species. To**
19 **exclude the alternative that these positive cross-correlations may merely reflect the general**
20 **biological quality of an organism or an inevitable by-product of having brains it is paramount**
21 **to provide solid evidence for the absence of g in at least some species. Here, we show that**
22 **wild-caught cleaner fish *Labroides dimidiatus*, a fish species otherwise known for its highly**
23 **sophisticated social behaviour, completely lacks g when tested on ecologically non-relevant**
24 **tasks. Moreover, performance in these experiments was not or negatively correlated with**
25 **an ecologically relevant task, and in none of the tasks did fish caught from a high population**
26 **density site outperform fish from a low-density site. g is thus unlikely a default result of how**
27 **brains are designed, and not an automatic consequence of variation in social complexity.**
28 **Rather, the results may reflect that g requires a minimal brain size, and thus explain the**
29 **conundrum why the average mammal or bird has a roughly 10 times larger brain relative to**
30 **body size than ectotherms. Ectotherm brains and cognition may therefore be organized in**
31 **fundamentally different ways compared to endotherms.**

32 **Keywords:** general intelligence; *Labroides dimidiatus*; ectotherm vertebrates; cognition; brain
33 evolution

34 **Introduction**

35 Fish do many apparently smart things (1,2). Yet, being ectotherm vertebrates, they have on
36 average ten times smaller brains corrected for body size compared to endotherm vertebrates
37 (3,4). Size differences are even more pronounced if only the pallial part of the forebrain is
38 considered, i.e. the part that is responsible for more complex cognitive functions (5,6). So
39 what can mammals and birds do with big, physiologically expensive brains (7,8) that reptiles,
40 amphibians and fishes cannot? A possible hypothesis is that only endotherms have general
41 intelligence. In humans, general intelligence in its broad definition involves various domains
42 of cognition, such as reasoning, planning, problem-solving as well as learning from experience,
43 that are centrally integrated to allow for more flexible behaviours, to understand abstract
44 concepts, and to have conceptual thoughts (9). A good indicator of the presence of general
45 intelligence in humans is the psychometric factor g . g results from the positive manifold, i.e.
46 the well-established finding that in humans, individual performance across tasks testing
47 different domains is correlated (10–15). Factor analyses of performance across such tasks will
48 thus result in a first factor on which all tasks load positively, and this general factor is referred
49 to as g . g has been demonstrated for a variety of mammals in controlled laboratory
50 experiments, including average brained species such as mice (16). Evidence for g is less clear
51 in birds (17) (but see (18)) but those studies have been conducted on small numbers and/or
52 in the field, potentially creating biased data as mostly motivated and/or bold individuals will
53 participate under these conditions (17).

54 Using g as indicator of general intelligence species other than humans has been criticised with
55 the argument that the positive manifold might be a pure side-effect. For instance, the positive
56 manifold may simply reflect variation in low-level biological properties, due to ontogenetic

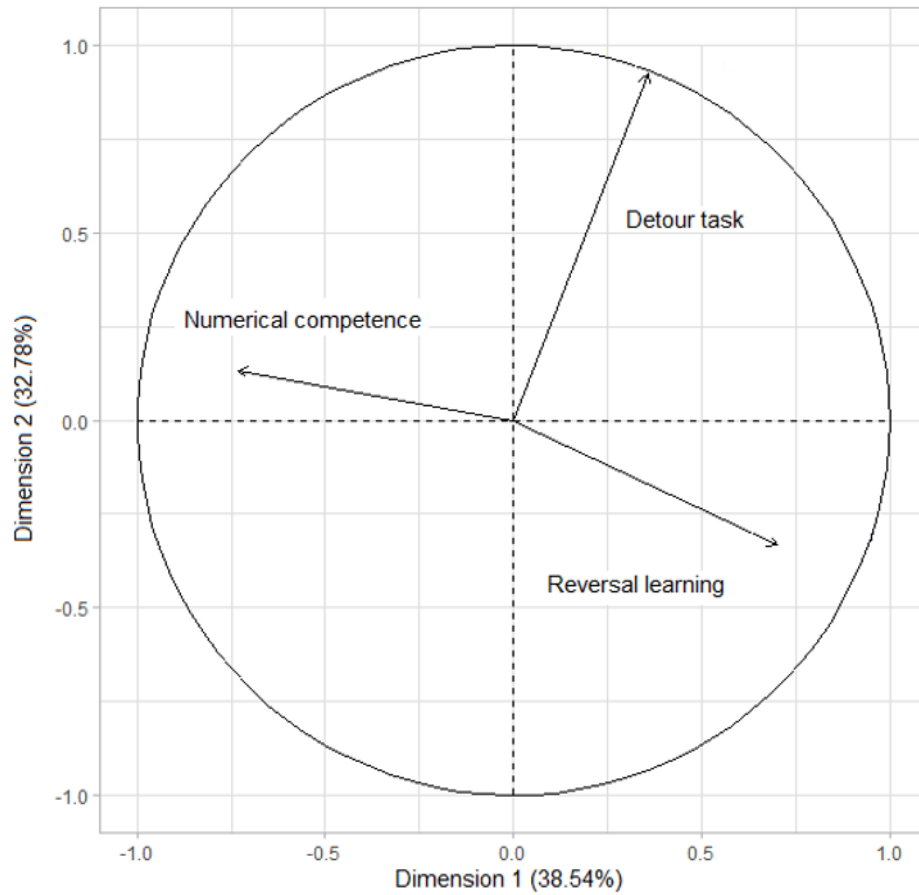
57 disturbances, or genetic load (i.e. the accumulation of deleterious, pleiotropic mutations, e.g.
58 (19,20)). Individuals with less disturbances, or less genetic load, may more fully express their
59 growth potential, which may also lead to better myelination of the nervous system, which
60 ultimately will operate smoother and faster across domains (21). Moreover, a positive
61 manifold can be an artefact of how brains are generally organized. Thomson (22) pointed out
62 already in 1916 that a positive manifold can arise in the absence of general intelligence due to
63 between-task neural overlap (see also (19,23–28)).

64 Thus, our aim was to simultaneously test two closely linked hypotheses by using cleaner fish
65 *Labroides dimidiatus* as study species. First, is general intelligence restricted to endotherms?
66 Second, does a g – factor mandatorily arise whenever a large enough, unbiased sample of
67 animals is tested under appropriate conditions? *L. dimidiatus* is a particularly suitable study
68 species as there is plenty of evidence suggesting that its cognitive performance is rather
69 outstanding for an ectotherm vertebrate. It engages in interactions with client fish that visit
70 to have ectoparasites removed while cleaners prefer to eat client mucus (29). Most likely due
71 to this conflict of interest, cleaners show high strategic sophistication in ecologically relevant
72 tasks, even outperforming primates (30). Furthermore, cleaners are able of generalised rule
73 learning (31), and apparently even pass the mirror test (32). Moreover, as variation in
74 intraspecific social complexity affects performance in g tasks in birds (18) and as it is well
75 established that fish densities affect cleaner fish expression of strategic sophistication in
76 ecologically relevant social tasks (33,34), we compared 80 individuals from two sites. One site
77 harbours high cleaner densities and the other one low densities (3.3 versus 0.4 cleaners /
78 100m², see SI, S2 Fig), which correlate strongly with large client densities (33). If cleaners from
79 the high-density site perform better also in g tasks, this would increase the likelihood that we

80 would find a positive manifold. Taken together, according to current knowledge we maximised
81 the chance to find g in an ectotherm vertebrate, which would support the notion that g is an
82 artefact and refute the hypothesis that g requires large brains.

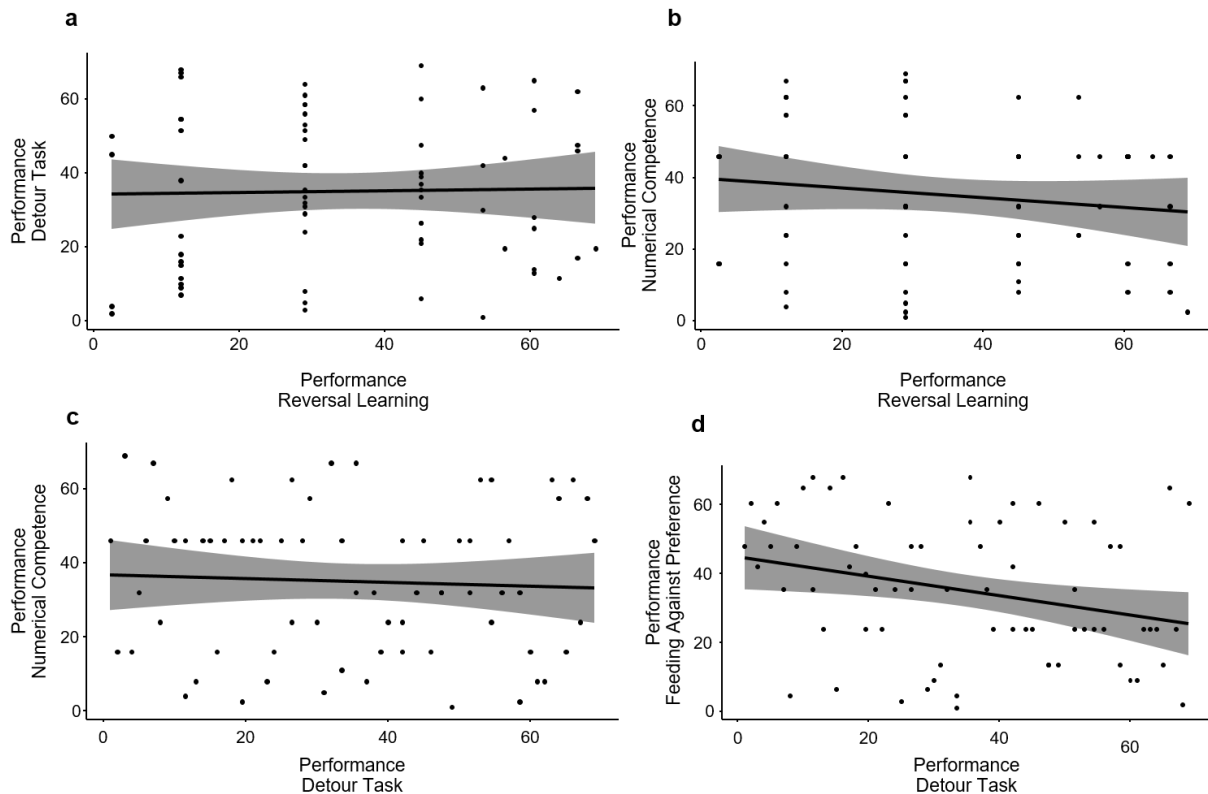
83 **Results**

84 We conducted four laboratory experiments (Fig 4) that have also been used on mammals
85 before (16). The four experiments tested for different cognitive domains: flexibility (reversal
86 learning; abbreviated 'RL'), self-control (detour task, abbreviated 'DT'), numerical competence
87 (dot number task, abbreviated 'NC') and working memory (object permanence). As tests for g
88 have to warrant performance above chance without ceiling effects, we had to omit the object
89 permanence task (presented in SI, S5 Fig) but could keep the other three tasks. Of the 80
90 subjects, we had to remove eleven individuals because they did not participate in at least one
91 task. With the remaining 69 individuals, we used a principal component analysis to test for the
92 presence of a positive manifold (Fig 1). Dimension 1 explained 38.5 % of the variance in
93 performance, with an Eigenvalue of 1.16. Two tasks loaded positively (RL and DT), whereas
94 the other one loaded negatively (NC). Dimensions 2 and 3 explained rather similar amounts of
95 variance (32.8 and 28.7 %; Eigenvalues 0.98 and 0.86). The results thus revealed no evidence
96 of a psychometric g in *L. dimidiatus*. Indeed, correlations of individual performance across
97 tasks support the view that the performances in the three tasks are independent of each other
98 (Spearman-Rank correlations, all $n = 69$, RL-DT: $r = 0.026$; RL-NC: $r = -0.137$; DT-NC: $r = -0.051$;
99 all NS, Fig 2a-c). Also, separate analyses for cleaners from high- and low-density sites did not
100 reveal any evidence for g (SI, S3 Fig).



101

102 **Fig 1. No evidence for g in a fish.** Principal Component Analysis (PCA) of the three cognitive
103 tasks. Dimension 1 (explaining 38.54 % of the variance in performance) and dimension 2
104 (explaining 32.78 % of the variance) are represented. The results for each task are represented
105 as vectors.

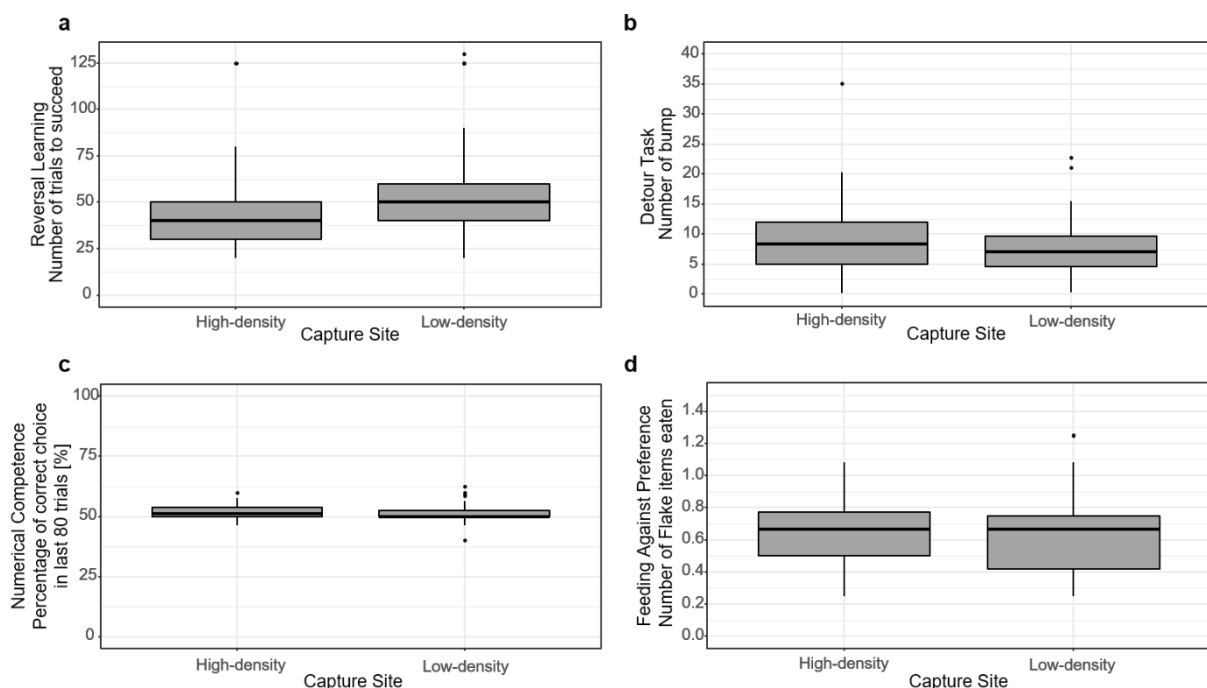


106

107 **Fig 2. Lack of positive correlations across cognitive tasks.** Pairwise correlations between the
108 ranks of individual performances in four different cognitive tasks (rank 1 being highest
109 performance, rank 69 being lowest performance). Individual performances did not
110 significantly correlate between any two of the three *g* tasks (a) RL – DT, Spearman-Rho $r =$
111 0.026 , $p = 0.8$; b) RL – NC, Spearman-Rho $r = -0.137$, $p = 0.3$; c) DT – NC, Spearman-Rho $r = -$
112 0.051 , $p = 0.7$). The only significant correlation was negative rather than positive, and found
113 between DT and FAP (d) Spearman-Rho $r = -0.283$, $p = 0.02$ ($p = 0.03$ with Bonferroni
114 correction)), a *g* task and an ecologically relevant task that are both testing for inhibition.

115 In addition to the *g* tasks, we conducted one ecologically relevant experiment, i.e. the
116 individuals' ability to feed against preference (abbreviated FAP) in order to prolong
117 interactions with clients and hence obtain more food. Like the detour task, the experiment
118 measures self-control, allowing us to ask how performance in an ecological task relates to
119 performance in an abstract task within a single domain. We found that DT was negatively
120 correlated with FAP (Spearman-Rho $r = -0.283$, $p = 0.018$, Fig 2d). While not expected, the
121 result suggests that positive correlations in performance may even be absent within domains.

122 Comparisons of cleaners from high- and low-density sites revealed no consistent differences
123 in the performance in each task separately (Fig 3). Differences were far from being significant
124 for all tasks (RL (panel a), DT (panel b), NC (panel c), and FAP (panel d) (Wilcoxon-tests, all $p >$
125 0.13). There was also no difference in cleaner body length between sites (Wilcoxon-test, $p =$
126 0.24; SI, S4 Fig), suggesting no systematic differences in age that could potentially have
127 confounded the results.



128

129 **Fig 3. Fish densities do not affect performance.** Box plots showing median, interquartiles and
130 95 % range of individual performances split by site of capture (high-density and low-density
131 site). a: number of trials to complete reversal learning task; b: mean number of head bumps
132 into Plexiglas separation in the detour task; c: % correct choices in the numerical competence
133 task; d: mean number of flakes eaten before a prawn item (i.e. highly preferred food) in the
134 feeding against preference task.

135 Discussion

136 The lack of correlated performance, and thus g in this fish species has far-reaching
137 implications. First, given that we tested a large random sample of cleaner fish under controlled
138 laboratory conditions, we provide the strongest evidence as yet that a positive manifold does
139 not emerge as a default in brains. Instead, our results support the notion that a major

140 consequence of the differences in brain size between mammals and fish is that mammals have
141 some system-level cognitive abilities that can be used across tasks tapping into different
142 cognitive domains, while the fish brain is organised in a more modular way. A more modular
143 organisation does apparently not prevent the emergence of complex cognitive processes like
144 transitive inference, generalised rule learning and even mirror self-recognition (31,32,35), i.e.
145 processes that go beyond conditioning. Thus, a presence/absence of these cognitive processes
146 cannot explain the ten-fold difference in relative brain size and the even larger difference with
147 respect to the pallial brain part (5,6) between endotherm and ectotherm vertebrate clades.
148 Also, the cleaners' performance in at least two *g* tasks (reversal learning and detour) was very
149 good, and there is evidence for high numerical competence as well as object permanence in
150 other fishes (2,35), suggesting that single cognitive tasks used to test for general intelligence
151 are not per se too difficult for ectotherms. In contrast, the transition from a modular
152 organisation to more general-purpose intelligence may have required a massive increase in
153 brain size.

154 A second major insight from our results is that environmental complexity, as indicated by
155 cleaner and other fish densities, largely fails to affect performance of cleaner fish in *g* tasks, in
156 contrast to results from Australian magpies (18). Therefore, it appears that previously
157 documented differences in strategic sophistication in ecologically relevant tasks between
158 cleaner fish individuals from high-density and low-density sites (33,34) are due to specific
159 experience effects rather than cognitive abilities per se.

160 Currently, most discussions on brain evolution focus on the relative importance of social
161 versus environmental complexity. While this discussion remains key to explain variation in size
162 and structure within clades, the fundamental divide between endotherm and ectotherm

163 vertebrates brain sizes may be largely driven by brains being more general-purpose machines
164 in the former versus specialised learning machines in the latter.

165 **Materials and Methods**

166 **Cleaner fish: *Labroides dimidiatus***

167 The cleaner wrasse, *Labroides dimidiatus*, is a protogynous fish and lives in a small territory
168 called cleaning station (36). It lives in harems and males can have up to five females in their
169 territory comprising several cleaning stations (one per female) (37). The species is widespread
170 in the Indopacific ocean and can also be found in the Red Sea (38). It feeds on the surface of
171 other reef fish called clients by removing ectoparasites from them. Cleaner fish have around
172 2000 interactions per day (39). As cleaner fish prefer to eat mucus over ectoparasites (29) this
173 creates a conflict of interest between cleaner and client over what the cleaner should eat. As
174 a consequence, cleaners need to eat against preference in order to cooperate, and hence to
175 avoid that clients respond with evasive actions like chasing the cleaner or switching to a
176 different partner for their next interaction (40). Therefore, the experiment in which cleaners
177 needed to feed against preference reflects high ecological relevance (41).

178 **Capture, individuals housing, and acclimation**

179 The study was conducted at the Lizard Island Research Station, Great Barrier Reef, Australia
180 in February - May 2018 and 2019. By finding pairs of cleaners and avoiding the larger
181 individual, 80 female cleaner fish were caught with a barrier net (2 m long, 1.5 m high, mesh
182 size 0.5 cm; 40 fish each year) and hand-nets on nearby reefs. 40 individuals were from a high-
183 density client fish area (Birds Islet crest; 20 each year; SI, S1 Fig), and another forty were from
184 a low-density client fish area (Birds Islet lagoon; 20 each year; SI, S1 Fig). 30 m long transects

185 covering a width of 5m (for detailed methods see (42)) revealed that mean cleaner fish
186 density, which is highly correlated with the density of large clients (33), differed by a factor 8
187 (3.3 versus 0.4 individuals / 100m²; Wilcoxon-test, m = 6, n = 5; p = 0.012; SI, S2 Fig). Fish were
188 housed individually in glass aquaria (62 x 27 x 38 cm). Each year, the forty individuals were
189 caught at the beginning of the field trip and split into two experimental cohorts of 20
190 individuals each, which were tested simultaneously. Cleaner fish were acclimatized for at least
191 twelve days before being subjected to five different experiments (experimental cohorts 1 and
192 3) or for at least 44 days (experimental cohorts 2 and 4). Cleaner fish were acclimated to feed
193 on Plexiglas plates, mimicking client fish in the captive environment. We provided mashed
194 prawn as food that we smeared on the Plexiglas plate. When fish were well accustomed to
195 their feeding plate (which took two to three days), they were trained to eat small pieces of
196 mashed prawn placed on dots drawn on a new feeding plate. Once fish were eating invariably
197 well on the feeding plate with and without dots, they were habituated to the different plates
198 and barriers that we used during all the different experiments (Fig 4). Most notably, fish were
199 exposed to a barrier that divided the aquarium into the holding compartment and the
200 experimental compartment. A door (dimension 7 x 18 cm) in the barrier could be opened so
201 that the cleaner fish could swim through. Cleaners thus had to be habituated to swimming
202 through the door.

203 **General experimental procedures**

204 During all experiments, subjects were guided into the holding compartment before each trial.
205 This allowed the experimenter to subsequently to set up the trial (i.e. plates with or without
206 food, partitions, barriers) without the fish intervening. Once a trial was properly set up, the
207 door was opened so that the cleaner fish could swim through and make its choices. During

208 simultaneous choice tasks (reversal learning and numerical competence tasks), an opaque
209 separation (10 cm wide) between the two plates helped to define a subject's choice: when the
210 subject's head passed the imagined line perpendicular to the start of the opaque separation
211 (Fig 4a,b,e), we scored that the subject had chosen the plate inside the compartment. If the
212 choice was wrong, the experimenter removed the correct plate immediately. If the choice was
213 correct, we allowed subjects to still inspect the wrong plate after having finished eating the
214 reward before removing both plates simultaneously.

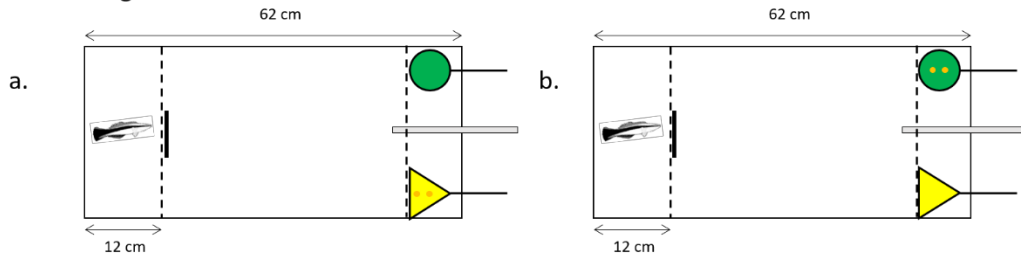
215 In all tasks, the experimenter tested each fish once before moving on to the next fish. Once
216 each fish had been exposed to one trial the next round of trials began, leading to an intertrial
217 interval of about 30 min. Twenty trials could be completed within a day, starting at 6:00 and
218 finishing at 17:00.

219 Fish conducted the experiments in their home aquarium. All experiments were video recorded
220 with a GoPro mounted on the forehead of the experimenter. All fish from 2018 and 18 fish
221 from 2019 were released at their respective sites of capture after experiments had been
222 completed. Another 22 fish were used in another project that included brain analyses. The
223 project was approved by the Animal Ethics Committee of the Queensland government (DAFF;
224 AEC Application Reference Number CA 2018/01/1155).

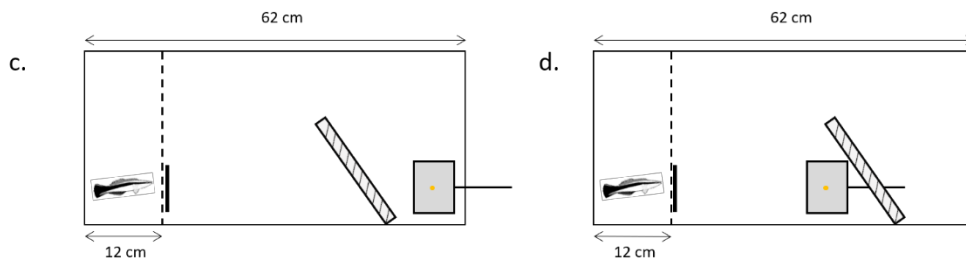
225 Before and in between the experiments, the fish were fed ad libitum every day, introducing
226 the feeding plate in the morning and removing it at the end of the day. During experimental
227 days, fish had to obtain food from making the correct choices as long as trials were conducted.
228 They were fed ad libitum at the end of the day after the trials had been completed. One day
229 off was kept between each experiment.

230 *Experimental setups*

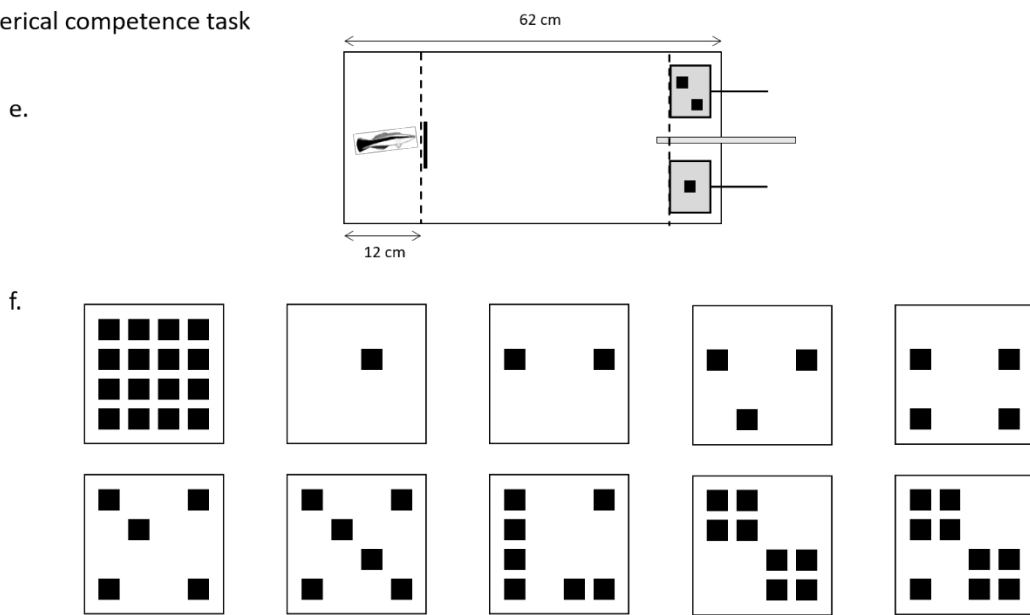
Reversal learning task



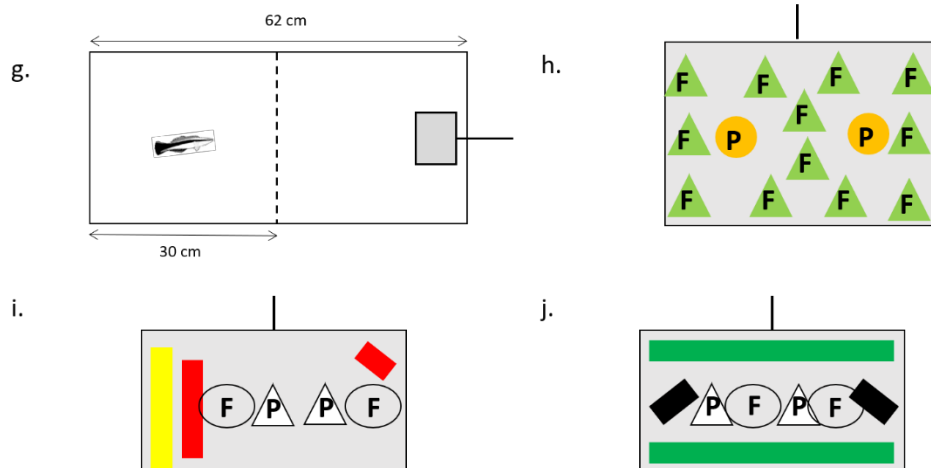
Detour task



Numerical competence task



Feeding against preference task



232 **Fig 4. Spatially explicit experimental setups of the three *g* tasks (a – f), and experimental**
233 **setup of the ecologically relevant task, the ability to feed against preference in order to**
234 **increase food intake (g – j).** Dashed lines: the transparent barrier that separated the holding
235 compartment from the experimental compartment (with the short black line indicating the
236 door through which fish could cross). Thicker grey and hatched structures represent opaque
237 and transparent barriers. Orange dots on the plates show the location of food rewards
238 (mashed prawn items). Panel a. shows the initial associative learning task, upon its completion
239 we started the reversal learning task by changing the role of the two plates (panel b.). Panels
240 c. and d. show the positions of the reward plate in the detour task, where only trials with the
241 reward plate behind the barrier were analysed. Panel e. shows the experimental setup for the
242 numerical competence task. In panel f., the plate to the upper left indicates the 16 possible
243 positions for the back squares, while the other plates give examples for plates displaying 1-9
244 squares. g. general setup. The dashed line shows the see-through barrier that kept fish in the
245 holding compartment while the experimenter placed a plate with food items in the
246 experimental compartment. h. the plate used to train fish that eating flake items is allowed
247 while eating a prawn item leads to the immediate removal of the plate. F: flake item. P: prawn
248 item. i. and j. the two plates used alternately in the actual experiment, offering 2 flake and
249 2 prawn items.

250 **Specific experimental protocols**

251 *1. Reversal learning task: learning – flexibility (Fig 4a,b)*

252 In reversal learning tasks, subjects first learn to associate one of two stimuli with a reward.
253 After this initial association learning, the contingencies are reversed and the other stimulus is
254 rewarded. This task thus measures how flexibly individuals can adjust to the new reward
255 contingencies. For this task, we offered two plates simultaneously: a yellow triangle shape
256 Plexiglas plate (8 cm wide and 7.5 cm height) was placed on the left side of the aquarium with
257 two pieces of prawn located on the back. On the other side, a green round shape Plexiglas
258 plate (8 cm wide and 7.5 cm height) without food was positioned. The positions of the two
259 plates remained constant during a first training phase, testing for initial learning and during
260 the reversal learning phase. For each trial, fish were first placed behind a transparent grid
261 barrier with a transparent door on one side of the aquarium, 12 cm from the aquarium wall
262 to form the waiting area. The plates were then introduced on the other side (with the handle

263 leaning against the aquarium wall providing stability), and only then the door was lifted and
264 the fish could make its choice by approaching one of the 2 plates.

265 To facilitate learning the initial association, we put two pieces of prawn on the back of the
266 rewarding plate and left both plates in the aquarium. Initially, individuals differed with
267 respect to how fast they would approach both plates. We hence varied the duration of first
268 trials such that all individuals would inspect both plates and eat the food, yielding trial
269 durations between five and 30 min. within two days, trial durations were down to 30 – 60
270 seconds. Such time intervals allowed the fish to eat the two items and to confirm that the
271 other plate did not offer any food. This training phase consisted of two exposures per day over
272 ten consecutive days. We then tested the fish for a significant preference for the rewarding
273 plate, offering only one item per trial, conducting 2 sessions of ten trials each per day. The
274 criteria for success were either 10/10 or 9/10 correct choices in a session, twice 8/10 in two
275 consecutive sessions, or at least 7/10 in three consecutive sessions. Once an individual had
276 reached criteria for the initial preference, we reversed the role of the two plates. Thus, two
277 pieces of mashed prawn were now placed behind the green round shape Plexiglas plate on
278 the right side of the aquarium. Each fish performed 20 trials per day and was tested the same
279 way according to the same learning criteria as described for the initial learning. We initially
280 ran trials for up to five days. If a subject had not succeeded yet within these 100 trials, we
281 conducted an extra five trials in which we prevented subjects to swim to the yellow plate by
282 inserting a see-through barrier in front of it. These extra trials either ensured that the
283 individuals were exposed to feeding off the green round plate, or they showed that some
284 individuals simply refused to approach that plate. We considered that these latter individuals
285 were rather afraid of the plate and hence their performance could not be interpreted as a

286 failure to learn. We therefore removed them from the data set. In contrast, we exposed the
287 subjects that had eaten off the green round plate to another 20 trials on the 6th day to see if
288 after these extra five trials they could reach the criterion or not. The number of trials needed
289 to reach learning criterion in this reversal learning task was used as measure of behavioural
290 flexibility for the statistical analysis.

291 2. *Detour task: inhibitory control (Fig 4c,d)*

292 We tested if cleaners were able to swim around an obstacle to get a food reward. This task
293 permitted to measure inhibitory control as well as spatial problem-solving. On the day prior
294 to the first test we familiarised subjects with the anthracite Plexiglas plate (10 cm x 5 cm) that
295 offered a visible piece of mashed prawn on its front side. We also acclimated the fish to the
296 obstacle by inserting it in the aquarium for 45 minutes twice during the day (once on the right-
297 hand side and once on the left-hand side). Given that cleaners had never been tested in a
298 detour task, we did not know what level of performance we could expect. We therefore
299 decided a priori to start with a simple task and to increase difficulty on consecutive days as
300 long as subjects readily managed to access the food reward. On the first day, we placed a
301 transparent obstacle in form of a plate, made visible by drawing a grid of black lines (1 – 1.5
302 cm apart) onto it, in front of a food plate. The obstacle was placed perpendicular to the
303 aquarium side wall. It was 19 cm wide, leaving subjects 5 cm to swim round it to get to the
304 food plate 15 cm behind the obstacle during test trials. In 50% of trials, the reward plate was
305 placed in front of the obstacle. During the total of 10 morning trials, the plate, obstacle and
306 door were always on the right side from the cleaners' perspective. In the 10 afternoon trials,
307 all equipment was moved to the left side from the cleaners' perspective. As cleaners accessed
308 the food plate without any problems, we counterbalanced the position of equipment within

309 sessions on the second day, conducting no more than two consecutive trials on any side. As
310 subjects continued to perform well, we placed the obstacle at a 45 degrees angle (Fig 4c,d),
311 further increasing the inhibition requirement because fish had to swim away from the food in
312 order to access it. We only used data from day 3 for our analyses. Subjects had a maximum of
313 60 secs to reach the reward plate. While we measured time to complete a trial, that measure
314 may not well reflect cognitive performance as it is influenced by swimming speed, which may
315 vary according to body size and/or motivation. Instead, we recorded the number of head
316 bumps against the obstacle in each trial as a measure of a cleaner's ability of self-control. Only
317 trials in which the plate was behind the obstacle were analysed; trials in which the plate was
318 in front of the obstacle only served to prevent the development of route routines. A total of
319 20 trials (10 experimental and 10 controls) were conducted, with an inter-trial interval of
320 about 30 min.

321 *3. Numerical competence task: quantitative reasoning (Fig 4e,f)*

322 This task tested quantitative reasoning in fish in a general form. In each trial, subjects were
323 presented two white plates (7.4 cm x 7.4 cm) with differing numbers of black squares on them
324 (Fig 4e). Each square was 11 mm² in size, and the number of squares on a plate varied between
325 1 to 9. On the day prior to the experiment, we acclimated the fish to the plate first by smearing
326 mashed prawn on a version displaying 10 squares. We later conducted two presentations
327 where we placed 2 prawn items on the back of the plate. During experiments, we used in total
328 20 different combinations of square numbers: 5:1, 6:3, 6:2, 6:4, 4:3, 3:2, 2:1, 4:1, 5:2, 3:1, 5:3,
329 7:3, 7:4, 7:2, 9:3, 8:3, 8:4, 8:5, 9:4, and 9:5. Each combination was presented once per day,
330 over a total of 8 days (yielding 160 trials in total), with the order of presentation randomised
331 between days. There were in total 16 potential positions for a square on a plate (Fig 4f). We

332 randomised the positions of squares for each number, removing any configurational cues for
333 cleaners to make choices. Across the 8 days, the position (left or right) of the two plates in
334 each combination was counterbalanced. The plate with the greater number of squares
335 invariably offered two food items on its back, while the other plate contained inaccessible
336 food on its back. Thus, some plates (those with 2-6 squares) sometimes yielded food and
337 sometimes they did not, depending on which plate they were paired with. As a consequence,
338 only the learning of a general rule based on numeric competences would yield performance
339 above chance levels (“always choose the plate that shows more squares”). Previous research
340 on cleaner fish has shown that this species has numerical competence, being able to learn to
341 prefer one plate over another one based on the number of black squares (rather than spread
342 or total black surface area) (43).

343 4. *Feeding against preference task: inhibitory control (ecologically relevant task) (Fig 4g-*
344 *j)*

345 The first three tasks presented cognitive challenges in abstract experimental setups, testing
346 for cognitive skills in the absence of ecologically relevant contexts. In contrast, the ability of
347 feeding against preference is of high ecological relevance. In natural interactions with client
348 fish, cleaners need to largely feed on less preferred ectoparasites (gnathiid isopodes) instead
349 of preferred client mucus so that it pays clients to visit cleaners (29). Our task mimicked
350 natural interactions, replacing clients, parasites and mucus with a plate, preferred prawn
351 items and less preferred flake items. There were three flake and three prawn items on the
352 plate. A trial continued as long as cleaners ate items, while the experimenter removed the
353 plate as soon as a cleaner ate a preferred prawn item. The food-maximising strategy was
354 hence to eat against preference. Therefore, the feeding against preference task is
355 conceptually linked to the Detour task as it measures inhibitory control. We included the task

356 to test whether there was any correlation between ecological and non-ecological tasks within
357 a single cognitive domain.

358 In order to prepare the fish for the experiment, we first gave them fish flakes (a mixture of
359 20% fish flakes and 80% mashed prawn to make the food stick to plates) on feeding plates
360 instead of pure mashed prawn to familiarise subjects with the new type of food. In the next
361 step, we conducted six training trials using a 15 x 10 cm plate with 14 items displayed on it (12
362 flake and 2 prawn items (44); Fig 4h). During a single trial, the plate was removed for 30 s as
363 soon as the cleaner ate the first prawn item, and then reintroduced to allow the cleaner a
364 second feeding bout until it ate the second prawn item. With this design, all subjects ate flake
365 items and could hence potentially learn that feeding on a prawn item leads to the removal of
366 the plate, while eating a flake item has no negative consequences. The intertrial interval was
367 of one hour.

368 The next day, we conducted the experiment. We exposed cleaners 12 times to a single plate
369 (10 cm x 6 cm) with two flake and two prawn items (Fig 4i,j). For a different study on reputation
370 management, we also conducted trials involving the simultaneous presentation of two plates,
371 but those data were not relevant for the current analyses. We took the mean number of flake
372 items eaten over the 12 trials as a measure of inhibitory control for the analysis.

373 **Data Analysis and Supplementary Results**

374 In each experiment, we ranked each fish by its performance according to the criteria specified
375 in each section above (SI, S1 Table). From the eighty fish caught for this study, we were able
376 to run the analysis with 69 individuals. A total of four individuals died before completing all
377 tasks, while seven failed to participate in at least one task. Of the 69 individuals, 36 were from

378 the high-density site (Birds Islet crest), and 33 were from the low-density site (Birds Islet
379 lagoon).

380 We first examined whether there was sufficient interindividual variation in the performance
381 of the cleaners in each task using descriptive statistics (means with standard deviation, and
382 maximum – minimum values; S1 Table). For the ND task, we also verified that cleaners as a
383 group performed above chance using a non-parametrical Wilcoxon signed-rank test in IBM[®]
384 SPSS[®] Statistics version 25.0.0.1 (Wilcoxon-test, $p = 0.001$). In all three g tasks, we found
385 substantial variation in performance and neither floor nor ceiling effects so that the three data
386 sets were included in a principal component analysis (PCA; using the packages FactoMineR
387 and missMDA). The key question was whether individual performance in each task was loading
388 positively on the first PC factor. In a second step, we calculated correlations in individual
389 performance across all possible pairs of the three tasks, using Spearman-Rho correlations as
390 presented in the main text.

391 In order to check the robustness of our results, we investigated whether non-cognitive
392 variables such as the site of capture, the year, the experimental set, as well as body length
393 could have had an effect on the variation found in the PCA using a linear model (package
394 'lme4') (45). Moreover, we analysed the individuals from the different capture sites
395 separately. The PCA analysis with the individuals from the high-density site explained 41.1 %
396 of the variance in performance in dimension 1, with an Eigenvalue of 1.23. Two tasks loaded
397 positively (RL and DT), whereas NC loaded negatively. No significant correlations were found
398 between the three different tasks (Spearman-Rank correlations, $n = 36$, RL-DT: $r = 0.179$; RL-
399 NC: $r = -0.086$; DT-NC: $r = -0.072$; S3a Fig). In the PCA analysis with the individuals from the
400 low-density site, dimension 1 explained 41.12 % of the variance in performance with an

401 Eigenvalue of 1.23. Two tasks loaded positively (DT and NC), whereas RL loaded negatively.
402 The correlations of individuals performance across tasks were invariable slightly negative and
403 non-significant (Spearman-Rank correlations, $n = 33$, RL-DT: $r = -0.095$; RL-NC: $r = -0.217$; DT-
404 NC: $r = -0.010$; S3b Fig). No difference was found in the body length between the high and the
405 low-density captures sites (Wilcoxon-test, $p = 0.12$; S4 Fig).

406 To test the extent to which abstract and ecological tests yield similar performances, we
407 correlated individual performances in the detour task and the feeding against preference task,
408 using a Spearman-Rho correlation. Thus, we calculated in total four correlations with our data.
409 As a consequence, we used Bonferroni correction to calculate a new $\alpha' = 0.0125$ (α/number
410 of tasks = $0.05/4$) to control for finding a spurious result. Finally, we used Wilcoxon- tests to
411 evaluate whether performance in any one test differed systematically as a function of site of
412 capture. Again, we used Bonferroni correction to control for multiple testing which resulted
413 in an $\alpha' = 0.0125$ ($\alpha/\text{number of tasks} = 0.05/4$. The Spearman-Rho correlation test, the PCA,
414 the linear models as well as the Wilcoxon-Test were carried out in R Rstudio © (R Version
415 1.3.1093, © 2009-2019 RStudio, PBC).

416 **Data archiving**

417 All data, meta-data as well as the R codes can be found on
418 <https://figshare.com/s/5ba3e10c1501ee1219ec>.

419 **Acknowledgements**

420 We are grateful to Lizard Island Research Station and their amazing staff. We thank in
421 particular Carel van Schaik for stimulating discussions. We thank Yasmin Emery for her
422 precious help collecting some data during the first two sets in the first year. We thank Radu

423 Slobodeanu for statistical help and support. The study was supported by the Swiss National
424 Science Foundation (grant number: 310030B_173334/1 to RB).

425 **References**

- 426 1. Brown C. Fish intelligence, sentience and ethics. *Anim Cogn*. 2015 Jan;18(1):1–17.
- 427 2. Bshary R, Gingers S, Vail AL. Social cognition in fishes. *Trends Cogn Sci*. 2014 Sep;18(9):465–71.
- 428 3. Jerison HJ. *Evolution of the Brain and Intelligence*. Academic Press New York. 1973;482.
- 429 4. Tsuboi M, van der Bijl W, Kopperud BT, Erritzøe J, Voje KL, Kotrschal A, et al. Breakdown of
430 brain–body allometry and the encephalization of birds and mammals. *Nat Ecol Evol*. 2018
431 Sep;2(9):1492–500.
- 432 5. Finger TE, Yamamoto N, Karten HJ, Hof PR. Evolution of the forebrain - revisiting the pallium:
433 Editorial. *J Comp Neurol*. 2013 Nov;521(16):3601–3.
- 434 6. Karten HJ. Vertebrate brains and evolutionary connectomics: on the origins of the mammalian
435 ‘neocortex’. *Philos Trans R Soc B Biol Sci*. 2015 Dec 19;370(1684):20150060.
- 436 7. Herculano-Houzel S. The remarkable, yet not extraordinary, human brain as a scaled-up
437 primate brain and its associated cost. *Proc Natl Acad Sci*. 2012 Jun
438 26;109(Supplement_1):10661–8.
- 439 8. Olkowitz S, Kocourek M, Lučan RK, Porteš M, Fitch WT, Herculano-Houzel S, et al. Birds have
440 primate-like numbers of neurons in the forebrain. *Proc Natl Acad Sci*. 2016 Jun
441 28;113(26):7255–60.
- 442 9. Gottfredson LS. Why g matters: The complexity of everyday life. *Intelligence*. 1997
443 Jan;24(1):79–132.
- 444 10. Carroll JB. *Human cognitive abilities: a survey of factor-analytic studies*. Cambridge ; New York:
445 Cambridge University Press; 1993. 819 p.
- 446 11. Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. *Nat Rev*
447 *Neurosci*. 2010 Mar 1;11(3):201–11.
- 448 12. Jensen AR. *The G Factor: The Science of Mental Ability*. 1999;8.
- 449 13. Jensen AR, Weng L-J. What is a good g? *Intelligence*. 1994 May;18(3):231–58.
- 450 14. Nisbett RE, Aronson J, Blair C, Dickens W, Flynn J, Halpern DF, et al. *Intelligence: New findings*
451 *and theoretical developments*. *Am Psychol*. 2012;67(2):130–59.
- 452 15. Spearman C. ‘General Intelligence,’ Objectively Determined and Measured. *Am J Psychol*. 1904
453 Apr;15(2):201.
- 454 16. Burkart JM, Schubiger MN, van Schaik CP. The evolution of general intelligence. *Behav Brain Sci*
455 [Internet]. 2017 [cited 2019 Oct 23];40. Available from:
456 https://www.cambridge.org/core/product/identifier/S0140525X16000959/type/journal_article
- 457 17. Searcy WA, Nowicki S. Birdsong learning, avian cognition and the evolution of language. *Anim*
458 *Behav*. 2019 May;151:217–27.

- 459 18. Ashton BJ, Ridley AR, Edwards EK, Thornton A. Cognitive performance is linked to group size
460 and affects fitness in Australian magpies. *Nature*. 2018 Feb 7;554(7692):364–7.
- 461 19. Arslan RC, von Borell CJ, Ostner J, Penke L. Negative results are needed to show the specific
462 value of a cultural explanation for *g*. *Behav Brain Sci*. 2017;40:e198.
- 463 20. Penke L, Denissen JJA, Miller GF. The evolutionary genetics of personality. *Eur J Personal*. 2007
464 Aug;21(5):549–87.
- 465 21. Matzel L, Townsend D, Grossman H, Han Y, Hale G, Zappulla M, et al. Exploration in outbred
466 mice covaries with general learning abilities irrespective of stress reactivity, emotionality, and
467 physical attributes. *Neurobiol Learn Mem*. 2006 Sep;86(2):228–40.
- 468 22. Thomson GH. A hierarchy without a general factor. *Br J Psychol*. 1916;8(3):271.
- 469 23. Bartholomew DJ, Deary IJ, Lawn M. A new lease of life for Thomson’s bonds model of
470 intelligence. *Psychol Rev*. 2009 Jul;116(3):567–79.
- 471 24. Bartholomew DJ, Allerhand M, Deary IJ. Measuring mental capacity: Thomson’s Bonds model
472 and Spearman’s *g*-model compared. *Intelligence*. 2013 Jul;41(4):222–33.
- 473 25. Kovacs K, Conway ARA. Process Overlap Theory: A Unified Account of the General Factor of
474 Intelligence. *Psychol Inq*. 2016 Jul 2;27(3):151–77.
- 475 26. Kovacs K, Conway ARA. What Is IQ? Life Beyond “General Intelligence”. *Curr Dir Psychol Sci*.
476 2019 Apr;28(2):189–94.
- 477 27. Rabaglia CD, Marcus GF, Lane SP. What can individual differences tell us about the
478 specialization of function? *Cogn Neuropsychol*. 2011 Jun;28(3–4):288–303.
- 479 28. Van Der Maas HLJ, Dolan CV, Grasman RPPP, Wicherts JM, Huizenga HM, Raijmakers MEJ. A
480 dynamical model of general intelligence: The positive manifold of intelligence by mutualism.
481 *Psychol Rev*. 2006;113(4):842–61.
- 482 29. Grutter AS, Bshary R. Cleaner wrasse prefer client mucus: support for partner control
483 mechanisms in cleaning interactions. *Proc R Soc B Biol Sci*. 2003 Nov 7;270(Suppl_2):S242–4.
- 484 30. Salwiczek LH, Prétôt L, Demarta L, Proctor D, Essler J, Pinto AI, et al. Adult Cleaner Wrasse
485 Outperform Capuchin Monkeys, Chimpanzees and Orang-utans in a Complex Foraging Task
486 Derived from Cleaner – Client Reef Fish Cooperation. *PLoS ONE*. 2012 Nov 21;7(11):e49068.
- 487 31. Wismer S, Grutter A, Bshary R. Generalized rule application in bluestreak cleaner wrasse
488 (*Labroides dimidiatus*): using predator species as social tools to reduce punishment. *Anim Cogn*.
489 2016 Jul;19(4):769–78.
- 490 32. Kohda M, Hotta T, Takeyama T, Awata S, Tanaka H, Asai J, et al. If a fish can pass the mark test,
491 what are the implications for consciousness and self-awareness testing in animals? de Waal
492 FBM, editor. *PLOS Biol*. 2019 Feb 7;17(2):e3000021.
- 493 33. Triki Z, Wismer S, Rey O, Ann Binning S, Levorato E, Bshary R. Biological market effects predict
494 cleaner fish strategic sophistication. Candolin U, editor. *Behav Ecol*. 2019 Nov 8;30(6):1548–57.

- 495 34. Wismer S, Pinto AI, Triki Z, Grutter AS, Roche DG, Bshary R. Cue-based decision rules of cleaner
496 fish in a biological market task. *Anim Behav.* 2019 Dec;158:249–60.
- 497 35. Grosenick L, Clement TS, Fernald RD. Fish can infer social rank by observation alone. *Nature.*
498 2007 Jan 25;445(7126):429–32.
- 499 36. Côté IM. Evolution and ecology of cleaning symbioses in the sea. *Oceanogr Mar Biol Annu Rev*
500 [Internet]. 2000 [cited 2012 Aug 15];38. Available from:
501 <http://www.vliz.be/imis/imis.php?refid=3753>
- 502 37. Robertson DR. Social Control of Sex Reversal in a Coral-Reef Fish. *Science.* 1972 Sep
503 15;177(4053):1007–9.
- 504 38. Randall JE. A Review of the Labrid Fish Genus *Labroides*, with Descriptions of Two New Species
505 and Notes on Ecology. 1958 Oct [cited 2013 Oct 21]; Available from:
506 <http://scholarspace.manoa.hawaii.edu/handle/10125/7923>
- 507 39. Grutter A. Parasite removal rates by the cleaner wrasse *Labroides dimidiatus*. *Mar Ecol Prog*
508 *Ser.* 1996;130:61–70.
- 509 40. Bshary R, Grutter AS. Asymmetric cheating opportunities and partner control in a cleaner fish
510 mutualism. *Anim Behav.* 2002 Mar;63(3):547–55.
- 511 41. Gingins S, Werminghausen J, Johnstone RA, Grutter AS, Bshary R. Power and temptation cause
512 shifts between exploitation and cooperation in a cleaner wrasse mutualism. *Proc R Soc B Biol*
513 *Sci.* 2013 Apr 24;280(1761):20130553–20130553.
- 514 42. Wismer S, Pinto AI, Vail AL, Grutter AS, Bshary R. Variation in Cleaner Wrasse Cooperation and
515 Cognition: Influence of the Developmental Environment? Schneider J, editor. *Ethology.* 2014
516 Jun;120(6):519–31.
- 517 43. Triki Z, Bshary R. Cleaner fish *Labroides dimidiatus* discriminate numbers but fail a mental
518 number line test. *Anim Cogn* [Internet]. 2017 Nov 13 [cited 2017 Dec 7]; Available from:
519 <http://link.springer.com/10.1007/s10071-017-1143-1>
- 520 44. Bshary R, Grutter AS. Punishment and partner switching cause cooperative behaviour in a
521 cleaning mutualism. *Biol Lett.* 2005 Dec 22;1(4):396–9.
- 522 45. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using **lme4**. *J Stat*
523 *Softw* [Internet]. 2015 [cited 2018 Oct 29];67(1). Available from:
524 <http://www.jstatsoft.org/v67/i01/>
525