

1   **Prepared as a Short Communication**

2   **Low Repeatability of Aversive Learning in Zebrafish (*Danio***  
3   ***rerio*)**

4  
5   Dominic Mason<sup>1</sup>, Susanne Zajitschek<sup>1,2</sup> Hamza Anwer<sup>1</sup>, Rose E O'Dea<sup>1</sup>, Daniel Hesselson<sup>3,4</sup>,  
6   Shinichi Nakagawa<sup>1</sup>

7   <sup>1</sup>Evolution and Ecology Research Centre, School of Biological and Environmental Sciences,  
8   University of New South Wales, Sydney 2052 NSW, Australia.

9   <sup>2</sup>School of Biological and Environmental Sciences, Liverpool John Moores University,  
10   Liverpool L3 3AF, United Kingdom.

11   <sup>3</sup>Diabetes and Metabolism Division, Garvan Institute of Medical Research, Sydney, NSW,  
12   Australia.

13   <sup>4</sup>St Vincent's Clinical School, UNSW, Australia, Sydney, NSW, Australia.

14   Author for correspondence: D. Mason

15   **e-mail:** [dpmason91@gmail.com](mailto:dpmason91@gmail.com)

16   **Keywords.** Repeatability, Learning, Zebrafish

17   **Ethics.** All procedures approved by the Garvan Animal Ethics Committee (ARA 18\_18) as  
18   noted in Methods.

19   **Data accessibility.** All data and code can be accessed at the Open Science Framework  
20   : (<https://osf.io/t95v3/>).

21   **Author Contributions. Dominic Mason:** Methodology, Data Curation, Formal Analysis,  
22   Investigation, Writing – original draft preparation, Writing – review and editing,

23   Visualisation. **Susanna Zajitschek:** conceptualisation, methodology, Data Curation, Writing

24 – review and editing, Supervision. **Hamza Anwer:** Writing – review and editing. **Rose**  
25 **O’Dea:** Data Curation, Writing – review and editing, Visualisation. **Daniel Hesselton:**  
26 Resources, Writing – review and editing. **Shinichi Nakagawa:** Conceptualization,  
27 Methodology, Software, Formal Analysis, Resources, Data Curation, Writing – review and  
28 editing, Supervision, Funding acquisition  
29 **Competing interests.** No competing interests.  
30 **Funding.** This research was funded through an Australian Research Council Discovery grant  
31 (DP180100818) awarded to S. Nakagawa

## 32 **Summary Statement**

33 Zebrafish were trained to avoid colour stimuli through negative associations with a mild  
34 electric shock. Very low repeatability estimates suggest individuals were not distinguishable  
35 by their aversive learning abilities.

## 36 **Abstract**

37 Aversive learning – avoiding certain situations based on negative experiences – can  
38 profoundly increase fitness in animal species, yet no studies have systematically quantified its  
39 repeatability. Therefore, we assessed the repeatability of aversive learning by conditioning  
40 approximately 100 zebrafish (*Danio rerio*) to avoid a colour cue associated with a mild  
41 electric shock. Across eight different colour conditions zebrafish did not show consistent  
42 individual differences in aversive learning ( $R = 0.04$ ). Within conditions, when zebrafish  
43 were conditioned to the same colour, blue conditioning was more repeatable than green  
44 conditioning ( $R = 0.15$  and  $R = 0.02$ ). Overall, aversive learning responses of zebrafish were  
45 weak and variable. We speculate that the effect of aversive learning might have been too  
46 weak to quantify consistent individual differences, or directional selection might have eroded  
47 additive genetic variance. We also discuss how confounded repeatability assays and  
48 publication bias could have inflated estimates of repeatability in the literature.

## 49 **Introduction**

50 Animals use the cognitive process of learning, which can be defined as a change in behaviour  
51 due to past experience, to respond to the environment (Kawecki, 2010). Learning has a  
52 profound influence on survival and reproductive success (Krebs & Davies, 1987; Skinner,  
53 1984), and has been studied in a wide range of taxa. For example, individual learning speed  
54 has been correlated with foraging performance in bees (Raine & Chittka, 2008) and  
55 grasshoppers (Pasquier & Grüter, 2016); and greater cognitive capacity has been linked to  
56 higher reproductive success in magpies (Ashton et al., 2018) and male robins (Shaw et al.,  
57 2019), as well as to healthier body condition in wild primates (Huebner et al., 2018).

58

59 Animals learn through association, which is reinforced differently by positive and negative  
60 experiences (appetitive and aversive learning, respectively). Appetitive learning takes place  
61 when individuals associate a stimulus with a ‘positive’ event, usually a food reward stimulus,  
62 whereas in aversive learning the association is with a ‘negative’ event, usually a fear inducing  
63 stimulus. Failing to learn from positive experiences (appetitive learning) prevents a potential  
64 benefit (i.e., a minor opportunity cost). Failing to learn from negative experiences may yield  
65 an immediate fatal cost. Therefore, both types of learning can increase lifetime fitness and  
66 drive natural selection, but appetitive learning may be under weaker selection than aversive  
67 learning.

68

69 For traits to evolve they need heritable variation that can be subject to selection. For labile  
70 traits (i.e., traits expressed more than once over a lifetime) the consistency of individual  
71 differences in trait expression indicates potential heritability. The common approach to  
72 quantify consistent individual differences in eco-evolutionary studies is estimating the  
73 statistical index ‘repeatability’ ( $R$ ; otherwise known as the ‘intra-class correlation coefficient’

74 or ICC; Lessells & Boag, 1987; Nakagawa & Schielzeth, 2010). Repeatability partitions  
75 variance into within-individual (residual) and between-individual components. Biologically,  
76 the repeatability of a trait indicates the amount of observed variance which is due to  
77 individuals sustaining trait differences between each other (Nakagawa & Schielzeth, 2010),  
78 but estimates can be inflated by measurement errors and experimental confounds (Dohm,  
79 2002; Niemelä & Dingemanse, 2017).

80

81 Generally, behavioural traits are moderately repeatable ( $R = 0.34$ ; Bell et al., 2009; cf.  
82 Holtmann et al., 2017), with cognitive behavioural traits showing somewhat lower  
83 repeatability ( $R = 0.15 - 0.28$ ; Cauchoux et al. 2018). Our understanding of how natural  
84 selection shapes the evolution of cognitive traits remains poor (Boogert et al., 2018). Despite  
85 the extensive literature on aversive learning, no published study has comprehensively  
86 quantified its repeatability (but note Cauchoux et al. (2018) includes three unpublished studies  
87 with some measures of aversive learning). To reduce this knowledge gap, we quantify the  
88 repeatability of aversive learning behaviour in zebrafish (*Danio rerio*), a popular model  
89 organism in cognitive science (Gerlai, 2016; Norton & Bally-Cuif, 2010). Zebrafish exhibit a  
90 range of distinct behaviours that can be measured in previously established assays  
91 (Fangmeier et al., 2018; Meshalkina et al., 2017).

92

93 Here, we use an avoidance conditioning assay — associating a visual cue with a mild electric  
94 shock (see Fig. 1A-E) — to thoroughly assess the repeatability of colour preferences and  
95 aversive learning in both male and female zebrafish. We expect individuals to consistently  
96 differ in their aversive learning speeds (i.e., separation of better and worse learners), but do  
97 not have particular expectations for sex differences. We estimate repeatabilities in two  
98 different ways. First, we examine repeatability across different colour pairs (four different

pairs with eight possible combinations: 8 measurements per individual; Fig. 1F). Given the estimates for appetitive learning summarised in Cauchoix et al. (2018), we predict a low to moderate repeatability. Second, we test whether repeatability is increased in a constant learning environment by using just one colour pair (both combinations of green and blue; 3 repeated measurements per individual for each colour; Fig. 1F). For both types of repeatability measurements (within and across learning environment), we also quantify colour preferences and their repeatabilities, to give a comparator in individual differences that can be compared to aversive learning.

## **Materials and Methods**

### **Zebrafish population**

Adult wildtype zebrafish were bred on the 24<sup>th</sup> of January 2019 (5 months old at the commencement of experiments) and maintained at the Garvan Institute of Medical Research in Sydney, Australia. The wildtype stock was derived from of a mixture of Tübingen long fin, AB and other unidentified strains to increase genetic diversity, which had been interbred for 8-10 generations. Fish were housed in 3.5L Tecniplast ZebTEC tanks (maximum of 24 fish per 3.5L tank) under standard laboratory conditions (~28°C; ~pH 7.5; ~1000 µs conductivity; 12/12h from 7:30 light/dark rotation) and fed live *Artemia salina* nauplii twice a day and commercially available fish food once per day (O.range GROW-L).

We marked juvenile fish for individual identification at around 90 days post-fertilisation with coloured tags (red, brown, purple, black, white, yellow, orange, pink, or green). For marking, fish were anaesthetised in a tricaine solution (4.2ml of 0.4% in 100ml of system water) for 20 seconds before being injected with Visible Implant Elastomer tags (VIE, Northwest Marine

Technologies, Inc.; Shaw Island, Washington, United States). We injected fish twice (unless one mark was blank), one on either side of the dorsal fin (Hohn & Petrie-Hanson, 2013). Among these marked fish, we used a total of 103 zebrafish with approximately equal sex ratios kept in 4 tanks of 24 individuals (12 males, 12 females) for both experiments. At any one time during the experiments, the same 96 fish were used, but to compensate for death, illness or experimenter error, seven fish were replaced by seven new fish over the three-month study. Due to incomplete data for zebrafish size (described below) the across conditions and within conditions analyses included 93 and 94 zebrafish, respectively. The Garvan Animal Ethics Committee approved all procedures described above and experiments described below (ARA 18\_18). Further, Garvan veterinarians oversaw fish welfare associated with aversive learning prior to our pilot tests.

## **Experimental Design**

### *Aversive Learning Assay*

We used an avoidance conditioning method to quantify aversive learning in a simple, automated assay (Brock et al., 2017; Fontana et al., 2019). We ran all assays using four Zantiks AD units (Zantiks Ltd., Cambridge, UK; see <https://osf.io/t95v3/> for further details). The units employed infrared tracking using an integrated computer to record fish movement and collect data. In the assay, a visual cue (colour or pattern) was associated with a negative stimulus (brief mild electric shock; 7V DC 80ms), which motivated fish to avoid the associated visual cue. We then measured the extent of avoidance (i.e., time spent away from the cue associated with an electric shock) compared to the baseline preference to quantify aversive learning (learning response). We based our initial assay parameters (e.g., the acclimation period, voltage, etc) on previous research (Brock et al., 2017), and subsequently modified the parameters based on the outcomes of pilot tests.

148

149 Before each assay we individually placed fish into one of four lanes within rectangular tanks  
150 (see Fig. 1A). For the assay, we exposed the fish to four stages; (i) Habituation: we  
151 habituated the fish to isolation in a novel environment over a 30-minute acclimation period  
152 (Fig. 1B); (ii) Baseline: the tank was visually split into two even zones via the colour  
153 displaying screen at the bottom of the tank (Fig. 1C). One of these two colours would later  
154 become conditioned with the mild electric shock (CS+), the other colour remained  
155 unconditioned (CS-). Here, the position of the colours (left or right) automatically switched  
156 every five minutes for a 30-minute period, and we recorded zebrafish preference for the CS+  
157 to obtain a baseline preference before conditioning; (iii) Conditioning: first, the CS+ (visual  
158 cue associated with shock) was displayed across the entire screen for 1.5 seconds then  
159 immediately afterwards paired with the US (mild electric shock) to condition the fish to an  
160 aversive experience. Second, the CS- (visual cue not associated with shock) covered the  
161 screen for 8.5 seconds (Fig. 1D). This phase was repeated nine times, sufficient for fish  
162 learning to avoid the CS+; and (iv) Probe: akin to the baseline period, the tank was split into  
163 two even zones (left or right) depicted by different visual cues. We tracked fish movement  
164 and recorded fish preference for the visual cue associated with the shock (CS+) over 5  
165 minutes. During this time, the visual cues switched every minute (see Fig. 1E). We used only  
166 2 minutes out of the 5-minute probe time since we determined a clear decrease in learning  
167 response in our observations. This probe length is similar to other studies: Brock et al. (2017)  
168 used a 2-minute probe, and Fontana et al. (2019) used a 1 minute probe. Probe CS+  
169 preference was used in comparison to baseline CS+ preferences to quantify learning.

170



## *Experimental Conditions*

We used a range of colour conditions to test aversive learning. Each condition was comprised of two visual cues, one aversive and one control (CS+ paired with CS-) (Fig. 1F). We selected different colour combinations to use as visual cues for the zebrafish, which had either been worked in pre-existing assays or were reported to evoke a clear colour preference (Brock et al., 2017; Roy et al., 2019). As a result, we chose seven colours (green, blue, grey, orange, magenta, red, yellow) and 1 pattern (check; hereafter, this pattern is also referred to as a 'colour' with the others). We used four visual cue combinations ('Check/Grey', 'Green/Blue', 'Red/Yellow', 'Magenta/Orange') and their reverse ('Grey/Check', 'Blue/Green', 'Yellow/Red', 'Orange/Magenta') for a total of eight conditions. For example, the 'Check/Grey' condition used check pattern as the CS+ (cue associated with shock) and grey colour as the CS- (control cue); the 'Grey/Check' condition used grey colour as the CS+ and check pattern as the CS-, and so on.

Prior to the experiment, we assigned fish into quartets (four fish that underwent trials within the same unit/assay tank simultaneously) that systematically rotated between trials. The balanced design accounted for three potential confounding variables: the time of day (quartet rotated), Zantiks unit (quartet rotated), and lane position (individual within quartet rotated). We estimated repeatability in two different situations (across conditions and within a single condition). Across conditions, we ensured fish experienced trials from all four colour pairs before subjecting them to their exact reverse four conditions (with trials conducted over four weeks in June and July 2019). We included this form of reverse learning to negate memory of the CS+ colour between trials, which may impact both baseline and probe colour preference. Within conditions, each zebrafish underwent trials in the 'Blue/Green' and 'Green/Blue' conditions a further two times (over two weeks in September 2019).

196

### 197 *Fish Size Measuring*

198 We took photos of each fish approximately one week after across-conditions trials and  
199 another set of photos approximately one week after within-conditions trials. We captured top  
200 down photos of live fish and measured fish in ImageJ (Schindelin et al., 2015). We used fish  
201 length (standard length) and width (at widest part of body) to calculate the ellipsoid size of  
202 the fish by using  $\pi \left( \left( \frac{\text{fish length}}{2} \right) \times \left( \frac{\text{fish width}}{2} \right) \right)$ . This controlled for a potential size effect  
203 resulting from loss of penetrance and effectiveness of the mild electric shock due to larger  
204 body size.

205

### 206 **Data Processing and Analysis**

207 All data processing and analyses were conducted in the *R* computing environment (version  
208 4.0.2; R Core Team, 2019). Linear mixed models were run using the *lme4* package (version  
209 1.1.21; Bates et al., 2014) in conjunction with the *lmerTest* package (version 3.1.2;  
210 Kuznetsova, Brockhoff, & Christensen, 2017) that provides Satterthwaite's degrees of  
211 freedom correction. We obtained repeatability values via the *rptR* package (version 0.9.22;  
212 Stoffel et al., 2017) that uses the *lme4* package to run mixed models. Based on visual  
213 assessments of residual distributions, assumptions of normality and constant variance were  
214 not clearly violated. The Zantiks units recorded time spent in each CS zone, total distance  
215 travelled and how often fish changed zones. All code, and the raw and processed data, are  
216 available at: <https://osf.io/t95v3/>. We deemed our results statistically significant at the alpha  
217 = 0.05 level (or when 95% confidence intervals did not overlap zero).

218

### Quantifying Aversive Learning

We determined learning by the difference in time that fish spent in the CS+ before and after the aversive experience. To analyse learning across all the sessions included in this study, we used the time difference (*'difference' = time spent in the CS+ during baseline – time spent in the CS+ during probe*) as the response variable in a linear mixed-effects model (LMM) via the *lmer* function in the *lme4* package. We fitted individual 'fish ID' (96 levels) and 'experimental condition ID' (8 levels, see Fig. 1F) as random effects in the model. Also, we included the following fixed effects: (1) 'sex' (female or male) to investigate sex differences in learning, (2) 'day' since first trial, to account for time effects of sequential days on learning or learning via repeated trials (e.g., 1 being the first day and 8 being the 7<sup>th</sup> day from the first), (3) 'fish size' to control for fish's response to conditioning which might be size dependent due to potential differences in body penetrance of a mild shock, (4) 'learning' (initial and reverse) to find if learning was affected when the CS $\pm$  of a condition were switched in successive trials. Note that we z-transformed the fixed effects 'day' and 'fish size' to make the intercept meaningful and slope estimates comparable (Schielzeth, 2010).

### Quantifying the Repeatability of Aversive Learning

We obtained enhanced agreement repeatability (hereafter referred to as repeatability) estimates by incorporating statistically significant fixed effects from the model and retaining their variance in the denominator (Nakagawa & Schielzeth, 2010). We only fitted the random effect 'fish ID' and included 'sex' as a fixed effect. The R package *rptR* computes repeatability values using the within and between individual variance in linear mixed models fitted with restricted maximum likelihoods (Nakagawa & Schielzeth, 2010). Using *rptR*, we obtained standard errors and 95% confidence intervals (CIs) through parametric bootstrapping, with each model set to 10,000 bootstrap samples. Following Bell (2009) and

Wolak (2012), we categorised our repeatability results into low ( $<0.2$ ), moderate ( $>0.2$  -  $<0.4$ ) and high ( $>0.4$ ).

### *Colour Preference and Repeatability*

An underlying assumption of our aversive learning assay was that zebrafish can discriminate between different colours. Therefore, from the baseline period (prior to aversive conditioning), we quantified underlying colour preferences (tendency to associate more heavily with one colour in a pair), and the consistency of individual differences in colour preference (i.e., repeatability of colour preference).

In each condition, preference for one colour was only compared to the other paired colour (e.g., preference for red is only relative to preference for yellow; see Fig. 1F). Given we examined relative colour preference, preferences for either colour in a condition were the inverses of each other. Hence, to be able to determine colour preference for each colour, we grouped conditions of matching colours into four groups for analysis (e.g., Group 1, 'Red/Yellow' & 'Yellow/Red'; Group 2, 'Green/Blue' & 'Blue/Green'; Group 3, 'Check/Grey' & 'Grey/Check'; Group 4, 'Orange/Magenta' & 'Magenta/Orange').

To analyse relative colour preference, we ran LMMs for each group of colours using across conditions data. We used baseline colour preference as the response variable 'baseline' for these models. We fitted the random effect 'fish ID' in the models (Group 1 & 4, 97 levels; Group 2 & 3, 98 levels; levels differ because one fish died prior to completing all conditions). Further, we fitted the following fixed effects: (1) 'day' (days since first trial) to control for potential colour preference change with time, (2) 'sex' (male or female) to account for sex differences and (3) 'learning' (initial and reverse) to see the effect of reverse learning on

colour preference. To determine the repeatability of colour preference, we used *rptR* mixed-effects models with the response variable ‘baseline’ to generate repeatability estimates. We did not find any fixed effects to be statistically significant, as such, they were excluded, and the colour preference models were fit with the random effect ‘fish ID’.

## Results & Discussion

We found negligible repeatability in aversive learning across the eight different conditions / colours ( $R = 0.04$ , 95% CI [0.001 - 0.097], Fig. 2A), despite individuals being able to discriminate between colours (as measured by moderate to high repeatabilities for colour preferences; Grey:  $R = 0.45$ , 95% CI [0.276 - 0.607]; Green:  $R = 0.45$ , 95% CI [0.278 - 0.604]; Red:  $R = 0.43$ , 95% CI [0.250 - 0.584]; Orange:  $R = 0.46$ ; 95% CI [0.283 - 0.605] Fig. 2B). Within conditions, we found very low repeatability in one condition (‘Green/Blue’ Fig. 2A;  $R = 0.02$ , 95% CI [0 - 0.153]), and low repeatability in the other (‘Blue/Green’ Fig. 2A;  $R = 0.15$ , 95% CI [0.023 - 0.278]). Therefore, the substantial variation in aversive learning we observed (as in Fig. 3A) was most likely driven by current (intrinsic or extrinsic) environmental factors, rather than additive genetic variance or canalized developmental differences (cf. Sznajder, Sabelis, & Egas, 2012).

Zebrafish showed strong relative colour preference in all four conditions (see Fig. 3B). We found preference for grey, green, red and orange, over check, blue, yellow and magenta, respectively. The strongest relative colour preference was found for red and orange, providing further evidence that zebrafish prefer colours with longer wavelengths (Roy et al., 2019). We did not find any statistically significant sex difference in colour preference, aversive learning, and their repeatability estimates (see Supplementary Information Table S1-2 and Fig. S1). Of relevance, a previous meta-analysis of repeatability for behavioural traits

found males tend to be more repeatable than females (Bell et al., 2009), but reported this result to be inconclusive.

Our findings of low repeatability for aversive learning are surprising, given low to moderate repeatability of behaviour and cognition reported in two meta-analyses. For general behaviour, Bell et al. (2009) reported an average repeatability of  $R = 0.34$ . For cognitive performance, Cauchoux et al. (2018) found  $R = 0.15-0.28$ , mostly based on temporal repeatability estimates from appetitive learning trials. Below we discuss four potential reasons why zebrafish in our experiment showed much less consistent individual differences in aversive learning compared to the previous estimates from Cauchoux et al. (2018) and Bell et al. (2009).

First, while zebrafish did demonstrate aversive learning, the average effect was small, and in many trials, individuals did not seem to avoid the negative stimulus. On average, individuals spent just 4-6 fewer seconds per minute in the negatively associated colour following conditioning (across conditions: female average = 3.89 seconds per min,  $SE = 1.05$ ,  $t_{33} = 3.65$ ,  $P < 0.001$ ; male average = 5.64 seconds per min,  $SE = 0.94$ ,  $t_{22} = 5.21$ ,  $P < 0.001$ ; Fig. 3A). The small effect could be caused by individuals not learning or quickly forgetting. It is also possible that learning performance would be greater at the group level; zebrafish are a shoaling species and learning may have evolved to depend on group dynamics. When tested individually, zebrafish display more variable behaviour and are more prone to stress (Pagnussat et al., 2013). Low repeatability could therefore be caused by zebrafish being largely insensitive or unresponsive to the conditioning when housed individually (i.e., poor aversive learners, a weak assay or anxious fish with impeded movement). However, the fact that there was a population shift in the direction of aversive learning raises the question of

319 why individuals who learnt in one trial did not maintain their performance across trials; if a  
320 particular subset of zebrafish had consistently learnt, or failed to learn, then we would have  
321 detected higher repeatability. Further, while the behaviour change following aversive  
322 conditioning was modest, zebrafish learnt much faster (in 1.5 minutes) compared to previous  
323 assays with appetitive training (e.g., over 20 days; Brock et al., 2017). As far as we are  
324 aware, no studies have investigated a relationship between the strength of associative learning  
325 and the magnitude of repeatability. Furthermore, it should be noted that our experiment only  
326 considered visual cues, but in the wild, fish often use chemical cues to detect danger (Brown,  
327 2003). Although technically more challenging, aversive learning using different sensory cues  
328 other than visual cues should be considered in the future.

329  
330 Second, past selection pressures on our study population may have eroded additive genetic  
331 variance associated with aversive learning, which was not restored in the intervening  
332 generations. In the wild, aversive learning could be under strong selection (e.g., to learn to  
333 evade predators), and individuals could be selected to learn from negative experiences as  
334 quickly as possible. Indeed, aversive learning could be under stronger selection than  
335 appetitive learning, as mortality costs of negative experiences can easily exceed opportunity  
336 costs of missing positive experiences. Stronger selective pressures could explain why we  
337 found substantially lower repeatability for aversive learning compared with previous results  
338 for appetitive learning. In a similar vein, traits more closely associated with fitness (e.g.,  
339 aversive learning) tend to not be as heritable (thus, repeatable; cf. Dohm, 2002) than less  
340 fitness related traits (e.g., appetitive learning; Merilä & Sheldon, 2000). However, we cannot  
341 be sure that the performance of zebrafish in our laboratory assay accurately captures their  
342 ability to aversively learn in their natural habitat.

343

344 Third, some of the repeatability values in the meta-analyses by Cauchoux et al. (2018) and  
345 Bell et al. (2009) may have been overestimated. An inflated repeatability estimate, also  
346 known as ‘pseudo-repeatability’, is the result of within-individual variation being erroneously  
347 accredited to differences between individuals (Niemelä & Dingemanse, 2017; Westneat et al.,  
348 2011). Pseudo-repeatability occurs when the conditions between measurements are too  
349 similar (e.g., environmental conditions are unchanged or intervals between measurements are  
350 too short) and might explain why we found higher repeatability when zebrafish were  
351 measured repeatedly within a single condition (‘Blue/Green’;  $R = 0.15$ ), compared to across  
352 eight separate conditions (although no inflation was seen in ‘Green/Blue’). Indeed, Cauchoux  
353 et al. (2018) and Bell et al. (2009) included studies with testing conditions which did not  
354 change over the course of a study, similar to our within-condition estimates. Further, most  
355 studies in both meta-analyses had relatively short intervals between measurements (most  
356 intervals were under a week in Cauchoux et al. (2018), and almost all were under a year in  
357 Bell et al. (2009). The short intervals between measurements reported in Bell et al. (2009)  
358 were significantly associated with higher repeatability values, consistent with pseudo-  
359 repeatability. Relevantly, two recent studies on birdsong reported that associative learning  
360 among individuals was not repeatable between years, indicating that estimates obtained over  
361 short intervals may not be a true reflection of consistent individual differences defined in  
362 animal personality (Soha et al., 2019; Zsebök et al., 2017).

363

364 Fourth, publication bias might have contributed to an inflation of the overall repeatability  
365 estimates in the published literature (cf. Parker et al., 2016). The average repeatability of 0.34  
366 reporting by Bell et al. (2009) was based on a meta-analysis of published studies. Cauchoux et



al. (2018) included many more unpublished datasets ( $n = 38$ ) compared to published datasets ( $n = 6$ ); they mentioned that their unpublished datasets produced, overall, a lower repeatability estimate than that of the published studies. This finding is consistent with the pattern that larger effect sizes are more likely to be published. Recent studies are increasingly reporting non-significant and low repeatability (e.g., Reichert et al., 2020; Vernouillet & Kelly, 2020). Therefore, an updated future meta-analysis may reveal a lower overall repeatability estimate in behaviour.

In conclusion, zebrafish did not show clear consistent between-individual differences in aversive learning. The low repeatability could potentially indicate that strong past selection pressure has almost driven aversive learning to fixation, because of the vital importance of learning to avoid danger. Alternatively, low repeatability may be due to the small effect of fish learning to avoid the stimuli. In addition, published repeatability estimates could be inflated by within-individual variance frequently being measured as between-individual differences (i.e., ‘pseudo-repeatability’), and by publication bias. We contend that these issues can be diminished in future behavioural research by controlling for confounding effects and reporting every estimate of behavioural traits, whether repeatable or not.

## Acknowledgements

We are grateful for the staff at the Biological Testing Facility, Garvan Institute of Medical Research for their support and husbandry of zebrafish. This study was supported by ARC (Australian Research Council) Discovery grant (DP180100818). No competing interests declared.

## References

- Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*, 554(7692), 364–367. <https://doi.org/10.1038/nature25503>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *ArXiv Preprint ArXiv:1406.5823*.
- Bell, A. M., Hankison, S. J., Laskowski, K. L. (2009). The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, 77(4), 771–783. <https://doi.org/10.1016/j.anbehav.2008.12.022>
- Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1756). <https://doi.org/10.1098/rstb.2017.0280>
- Brock, A. J., Sudwarts, A., Daggett, J., Parker, M. O., & Brennan, C. H. (2017). A fully automated computer based Skinner box for testing learning and memory in zebrafish. *BioRxiv*, 110478. <https://doi.org/10.1101/110478>
- Brown, G. E. (2003). Learning about danger: Chemical alarm cues and local risk assessment in prey fishes. *Fish and Fisheries*, 4(3), 227–234. <https://doi.org/10.1046/j.1467-2979.2003.00132.x>
- Cauchoux, M., Chow, P. K. Y. K. Y., Van Horik, J. O. O., Atance, C. M. M., Barbeau, E. J. J., Barragan-Jason, G., Bize, P., Boussard, A., Buechel, S. D. D., Cabirol, A., Cauchard, L., Claidière, N., Dalesman, S., Devaud, J. M., Didic, M., Doligez, B., Fagot, J., Fichtel, C., Henke-Von Der Malsburg, J., ... Morand-Ferron, J. (2018). The repeatability of cognitive performance: A meta-analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1756). <https://doi.org/10.1098/rstb.2017.0281>
- Dohm, M. R. R. (2002). Repeatability estimates do not always set an upper limit to heritability. *Functional Ecology*, 16(2), 273–280. <https://doi.org/10.1046/j.1365-2435.2002.00621.x>
- Fangmeier, M. L., Noble, D. W. A., O'Dea, R. E., Usui, T., Lagisz, M., Hesselson, D., & Nakagawa, S. (2018). Computer Animation Technology in Behavioral Sciences: A Sequential, Automatic, and High-Throughput Approach to Quantifying Personality in Zebrafish (*Danio rerio*). *Zebrafish*, 15(2), 206–210. <https://doi.org/10.1089/zeb.2017.1532>
- Fontana, B. D., Cleal, M., Clay, J., & Parker, M. O. (2019). Zebrafish (*Danio rerio*)

423 behavioral laterality predicts increased short-term avoidance memory but not stress-  
 424 reactivity responses. *BioRxiv*, 565309. <https://doi.org/10.1101/565309>

425 Gerlai, R. (2016). Learning and memory in zebrafish (*Danio rerio*). *Methods in Cell Biology*,  
 426 134, 551–586. <https://doi.org/10.1016/bs.mcb.2016.02.005>

427 Hohn, C., & Petrie-Hanson, L. (2013). Evaluation of visible implant elastomer tags in  
 428 zebrafish (*Danio rerio*). *Biology Open*, 2(12), 1397–1401.  
 429 <https://doi.org/10.1242/bio.20136460>

430 Holtmann, B., Lagisz, M., & Nakagawa, S. (2017). Metabolic rates, and not hormone levels,  
 431 are a likely mediator of between-individual differences in behaviour: a meta-analysis.  
 432 *Functional Ecology*, 31(3), 685–696. <https://doi.org/10.1111/1365-2435.12779>

433 Huebner, F., Fichtel, C., & Kappeler, P. M. (2018). Linking cognition with fitness in a wild  
 434 primate: Fitness correlates of problem-solving performance and spatial learning ability.  
 435 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1756).  
 436 <https://doi.org/10.1098/rstb.2017.0295>

437 Kawecki, T. J. (2010). Evolutionary ecology of learning: Insights from fruit flies. *Population*  
 438 *Ecology*, 52(1), 15–25. <https://doi.org/10.1007/s10144-009-0174-0>

439 Krebs, J. R., & Davies, N. B. (1987). An introduction to behavioural ecology, 2nd ed. In *An*  
 440 *introduction to behavioural ecology*, 2nd ed. Sinauer Associates.

441 Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in  
 442 Linear Mixed Effects Models . *Journal of Statistical Software*, 82(13).  
 443 <https://doi.org/10.18637/jss.v082.i13>

444 Lessells, C. M., & Boag, P. T. (1987). Unrepeatable Repeatabilities: A Common Mistake.  
 445 *The Auk*, 104(1), 116–121. <https://doi.org/10.2307/4087240>

446 Merilä, J., & Sheldon, B. C. (2000). Lifetime reproductive success and heritability in nature.  
 447 *American Naturalist*, 155(3), 307–310. <https://doi.org/10.2307/3078867>

448 Meshalkina, D. A., Kizlyk, M. N., Kysil, E. V., Collier, A. D., Echevarria, D. J., Abreu, M.  
 449 S., Barcellos, L. J. G., Song, C., & Kalueff, A. V. (2017). Understanding zebrafish  
 450 cognition. In *Behavioural Processes* (Vol. 141, pp. 229–241).  
 451 <https://doi.org/10.1016/j.beproc.2016.11.020>

452 Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A  
 453 practical guide for biologists. In *Biological Reviews* (Vol. 85, Issue 4, pp. 935–956).  
 454 <https://doi.org/10.1111/j.1469-185X.2010.00141.x>

455 Niemelä, P. T. T., & Dingemanse, N. J. J. (2017). Individual versus pseudo-repeatability in  
 456 behaviour: Lessons from translocation experiments in a wild insect. *Journal of Animal*

*Ecology*, 86(5), 1033–1043. <https://doi.org/10.1111/1365-2656.12688>

Norton, W., & Bally-Cuif, L. (2010). Adult zebrafish as a model organism for behavioural genetics. In *BMC Neuroscience* (Vol. 11). <https://doi.org/10.1186/1471-2202-11-90>

Pagnussat, N., Piato, A. L., Schaefer, I. C., Blank, M., Tamborski, A. R., Guerim, L. D., Bonan, C. D., Vianna, M. R. M., & Lara, D. R. (2013). One for all and all for one: The importance of shoaling on behavioral and stress responses in zebrafish. *Zebrafish*, 10(3), 338–342. <https://doi.org/10.1089/zeb.2013.0867>

Parker, T. H., Forstmeier, W., Koricheva, J., Fidler, F., Hadfield, J. D., Chee, Y. E., Kelly, C. D., Gurevitch, J., & Nakagawa, S. (2016). Transparency in Ecology and Evolution: Real Problems, Real Solutions. In *Trends in Ecology and Evolution* (Vol. 31, Issue 9, pp. 711–719). <https://doi.org/10.1016/j.tree.2016.07.002>

Pasquier, G., & Grüter, C. (2016). Individual learning performance and exploratory activity are linked to colony foraging success in a mass-recruiting ant. *Behavioral Ecology*, 1(1), arw079. <https://doi.org/10.1093/beheco/arw079>

Raine, N. E., & Chittka, L. (2008). The correlation of learning speed and natural foraging success in bumble-bees. *Proceedings of the Royal Society B: Biological Sciences*, 275(1636), 803–808. <https://doi.org/10.1098/rspb.2007.1652>

Reichert, M. S., Crofts, S. J., Davidson, G. L., Firth, J. A., Kulahci, I. G., Quinn, J. L., & Quinn, J. L. (2020). *Multiple factors affect discrimination learning performance , but not between-individual variation , in wild mixed-species flocks of birds.*

Roy, T., Suriyampola, P. S. P. S., Flores, J., López, M., Hickey, C., Bhat, A., & Martins, E. P. E. P. (2019). *Color preferences affect learning in zebrafish, Danio rerio.* 9(1), 14531. <https://doi.org/10.1038/s41598-019-51145-5>

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1(2), 103–113. <http://doi.wiley.com/10.1111/j.2041-210X.2010.00012.x>

Schindelin, J., Rueden, C. T., Hiner, M. C., & Eliceiri, K. W. (2015). The ImageJ ecosystem: An open platform for biomedical image analysis. *Molecular Reproduction and Development*, 82(7–8), 518–529. <https://doi.org/10.1002/mrd.22489>

Shaw, R. C., Mackinlay, R. D., Clayton, N. S., Burns, K. C., Shaw, R. C., Mackinlay, R. D., Clayton, N. S., & Burns, K. C. (2019). Memory Performance Influences Male Reproductive Success in a Wild Bird Memory Performance Influences Male Reproductive Success in a Wild Bird. *Current Biology*, 29(9), 1498-1502.e3. <https://doi.org/10.1016/j.cub.2019.03.027>

- Skinner, B. F. (1984). THE EVOLUTION OF BEHAVIOR. *Journal of the Experimental Analysis of Behavior*, 41(2), 217–221. <https://doi.org/10.1901/jeab.1984.41-217>
- Soha, J. A., Peters, S., Anderson, R. C., Searcy, W. A., & Nowicki, S. (2019). Performance on tests of cognitive ability is not repeatable across years in a songbird. *Animal Behaviour*. <https://doi.org/10.1016/j.anbehav.2019.09.020>
- Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2017). rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 8(11), 1639–1644. <https://doi.org/10.1111/2041-210X.12797>
- Sznajder, B., Sabelis, M. W., & Egas, M. (2012). How Adaptive Learning Affects Evolution: Reviewing Theory on the Baldwin Effect. *Evolutionary Biology*, 39(3), 301–310. <https://doi.org/10.1007/s11692-011-9155-2>
- Team, R. C. (2017). *R: A Language and Environment for Statistical Computing* (1.2.5042). <http://www.r-project.org/>
- Vernouillet, A., & Kelly, D. M. (2020). Individual exploratory responses are not repeatable across time or context for four species of food-storing corvid. *Scientific Reports*, 10(1), 1–11. <https://doi.org/10.1038/s41598-019-56138-y>
- Westneat, D. F., Hatch, M. I., Wetzel, D. P., & Ensminger, A. L. (2011). Individual variation in parental care reaction norms: Integration of Personality and plasticity. *American Naturalist*, 178(5), 652–667. <https://doi.org/10.1086/662173>
- Wolak, M. E. M. E. E., Fairbairn, D. J. D. J. J., & Paulsen, Y. R. Y. R. R. (2012). Guidelines for estimating repeatability. *Methods in Ecology and Evolution*, 3(1), 129–137. <https://doi.org/10.1111/j.2041-210X.2011.00125.x>
- Zsebők, S., Herczeg, G., Blázi, G., Laczi, M., Nagy, G., Szász, E., Markó, G., Török, J., & Garamszegi, L. Z. (2017). Short- and long-term repeatability and pseudo-repeatability of bird song: sensitivity of signals to varying environments. *Behavioral Ecology and Sociobiology*, 71(10). <https://doi.org/10.1007/s00265-017-2379-0>

## Figure Legends

### Figure 1

**Colour conditions and aversive learning assay:** (A) zebrafish are placed in the experimental tanks and (B) acclimated to the novel environment for 30-minutes; (C) in a 30-minute baseline period, initial CS± preference is established; (D) during the conditioning

phase, fish are presented the CS+, then immediately subjected to a mild electric shock; and (E) in a 5-minute probe phase, learning is determined by fish spending less time in the CS+ when compared to the baseline. (F) Each condition is a combination of two visual cues (zones), one conditioned to a mild electric shock (CS+), the other is not (CS-). Across conditions eight colour conditions and eight sessions (each session is represented by a white box). Within conditions: two colour conditions and four sessions (in addition to two sessions in Experiment 1).

## **Figure 2**

**Repeatability of aversive learning and colour preference in zebrafish.** Points and whiskers represent means and 95% confidence intervals via parametric bootstrapping. (A) Zebrafish show somewhat consistent individual differences in aversive learning within the 'Blue/Green' pair, but not within the 'Green/Blue' pair or across all colour combinations. (B) Zebrafish show consistent individual differences in colour preferences (variation depicted in Fig. 3B).

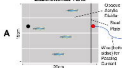
## **Figure 3**

**Violin plots for aversive learning and colour preferences.** Smaller coloured points depict individual trials. Larger black points and error bars depict means and standard deviations of observations. (A) The top panel shows means and variation in aversive learning, split by sex (female = purple; male = green) when all the session data is combined. Points above the line at zero depict trials in which zebrafish spent less time in the aversive stimulus colour in the probe period (the colour associated with an electric shock) relative to the baseline period (i.e., aversive learning). (B) The bottom panel shows the tendency of zebrafish to favour one colour in a pair during the baseline period (i.e., before administration of electric shocks). The

547 dashed horizontal line at 0.5 represents no colour preference (i.e., spending 30 seconds in  
548 each colour zone).

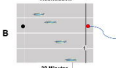
# Avoidance Conditioning Assay

## Experimental Tank



Start of Assay

## Habituation



30 Minutes

**C**



Baseline

x6



30 Minutes

**D**



Conditioning

x9



3.5 Minutes

**E**



Probe

x5



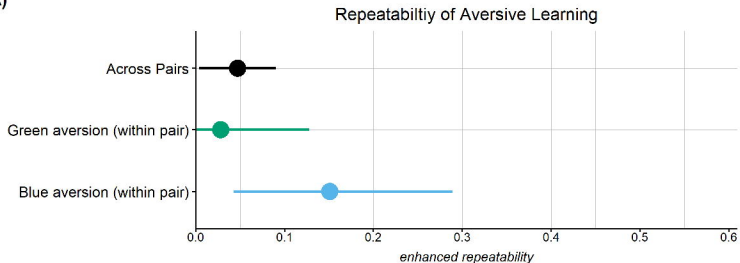
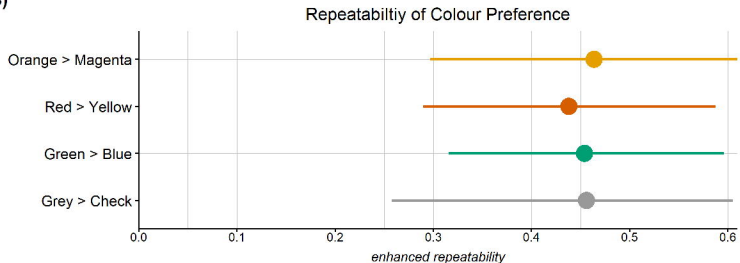
1 Minute

## Experimental Conditions

**F**

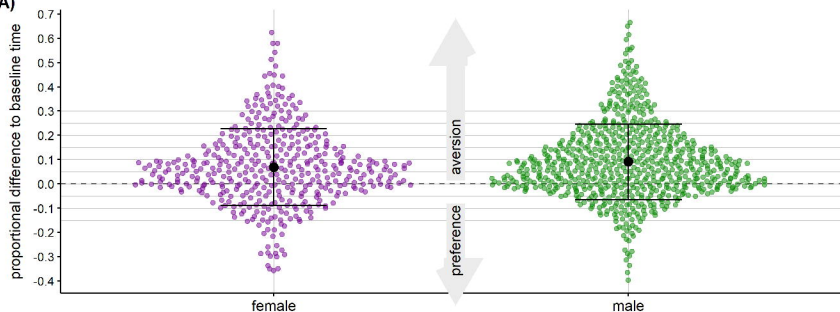




**(A)****(B)**

## Aversive Learning

(A)



## Baseline Colour Preferences

(B)

