SUPPLEMENTARY INFORMATION

Ocean acidification does not impair the behaviour of coral reef fishes

Timothy D. Clark[1*], Graham D. Raby[2], Dominique G. Roche[3,4,5], Sandra A. Binning[4,5], Ben Speers-Roesch[6], Fredrik Jutfelt[7], Josefin Sundin[7,8,9*]

[1] *School of Life and Environmental Sciences, Deakin University, Geelong, VIC, Australia 3216.*

[2] *Great Lakes Institute for Environmental Research, University of Windsor, 2601 Union St., Windsor, ON, N9B3P4, Canada.*

[3] *Fish Ecology and Conservation Physiology Laboratory, Department of Biology, Carleton University, 1125 Colonel By Dr., Ottawa, ON, K1S5B6, Canada.*

[4] *Institut de Biologie, Éco-Éthologie, Université de Neuchâtel, Rue Emilie-Argand 11, 2000, Neuchâtel, Switzerland.*

[5] *Département de sciences biologiques, Université de Montréal, Montréal, Québec, Canada.*

[6] *Department of Biological Sciences, University of New Brunswick, Saint John, New Brunswick, Canada.*

[7] *Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway.*

[8] *Department of Neuroscience, Uppsala University, Uppsala, Sweden.*

[9] *Department of Aquatic Resources, Swedish University of Agricultural Sciences, Drottningholm, Sweden.*

* Corresponding authors: Clark: +61 3 9244 6035; t.clark@deakin.edu.au and Sundin +46 730 302270; josefin@teamsundin.se; josefin.sundin@slu.se

Supplementary Methods

Experiments were conducted in August-September 2014, April-July 2015, and January 2016. Predator cue avoidance and activity trials were conducted at the Lizard Island Research Station (LIRS) in 2014 and 2016, and at the Australian Institute of Marine Science (AIMS) in Townsville in 2015. Behavioural lateralisation trials were conducted at LIRS in 2014 and at AIMS in 2015.

Response to predator chemical cues

Lizard Island Research Station 2014

The setup for the two-channel choice flume followed established protocols[45] and consisted of a primary header tank, which drained into two secondary header tanks, each of which drained into one side of the choice flume. The primary header tank contained 60 L of flow-through (12 L min^{-1}) ambient seawater (~23ºC, ~410 µatm CO_2). These conditions were maintained in the primary header tank when running trials on control fish, whereas CO_2 was increased to ~1,000 µatm in this tank (AT-Control System) when running the CO_2 treatment fish. pH levels were monitored by the AT-Control System at all times, and additional direct CO_2 measurements were performed at least twice per day using a Vaisala CO_2 meter (as described in the methods section of the main paper).

The secondary header tanks (60 L) each received water from the primary header tank at 6 L min^{-1} and were used to manipulate predator chemical cues in the water flowing into the choice flume. While one of these tanks was always kept predator-free and thus contained identical water as that of the primary header tank, the other secondary header tank housed four individual predatory fish (*Cephalopholis cyanostigma*) to add predator chemical cues to the water. The four *C. cyanostigma* that were used in each trial (i.e., in one header tank) had an individual body mass of 170 - 350 g. The fish in the high CO_2 group had been acclimated to the CO_2 treatment for 5 - 16 days prior to commencement of experiments, while control fish had been held for 4 - 16 days.

The choice flume was a custom-built, larger version of a two-channel choice flume used in previous studies[46]. Detailed information on the design and function of two-channel choice flumes is given elsewhere[45]. Briefly, the flume was made of an insulated polystyrene tank (L × W × H = 580 × 260 × 280 mm; water depth 80 mm) sectioned lengthways by a central PVC wall that ran half the length of the flume. The downstream half of the flume was the choice arena (180 × 260 mm) where the individual test fish were positioned. Flow from each secondary header tank (6 L min^{-1}) entered opposite sides (channels) of the flume at the

upstream end and ran in parallel through collimators before reaching the choice arena where the central PVC wall ended. The flows remained separate and laminar through the choice arena (confirmed by routine validation tests using food dye; https://youtu.be/EStc2VcGPCY) despite there being no physical barrier, and thus fish were free to move throughout the choice arena so that side (cue) preferences could be determined.

Australian Institute of Marine Science 2015

We initially tested flumes built to the exact specifications (L × W × H = 270 × 45 × 45 mm) of those used in previous papers examining ocean acidification and fish behaviour/chemoreception with larval and juvenile coral reef fishes (e.g.,[4,5,9,25]). They were tapered at the downstream end, which was blocked off (to prevent the fish from escaping) with 0.3 mm mesh, the channels were left open and accessible to the fish (following Munday et al. (2009)[9]; Dixson et al. (2010)[4]), and no baffles or collimators were used at the anterior end of the flume. We were unable to achieve laminar flow using this set-up; both incoming streams of water thoroughly mixed throughout the flume (https://youtu.be/jrtyc-rLGWc?t=705). Thus, we modified the flume design by added several layers of fine mesh at the inlets of both channels to ensure water was baffled to produce an even flow (see https://youtu.be/jrtyc-rLGWc?t=292). In order to ensure that we were carefully following previous protocols (e.g., Dixson et al. (2010)[4]), inflow for each channel was initially set to 100 mL min^{-1} (3.6 mm s^{-1}; controlled with flow-meters attached in-line to the hose going to each flume channel). Once again, achieving laminar flows was impossible at this speed, so trials were carried out at 135 mL min^{-1} (4.9 mm s^{-1}). However, to follow established guidelines for choice flume design[45], we built and used two new flumes for the chemical cue trials rather than use the ones described above. These flumes were similar in dimension (same width, depth, and water speed) as the flumes described above, but instead the downstream end was flat (with three holes through which water drained; total length 235 mm), and there 0.8 mm of mesh were present at both the upstream and downstream ends of the choice arena (90 mm long) that prevented the fish from entering either of the channels (see https://youtu.be/jrtyc-rLGWc?t=3).

To produce predator chemical cues, we used four flagtail grouper (*Cephalopholis urodeta*, wild caught, two dedicated to each treatment), a coral reef mesopredator that is sympatric with *A. polyacanthus* in nature and in the same genus as the predator used at Lizard Island in 2014 (see above) and 2016 (see below). The *C. urodeta* in the control treatment were 50.2 and 51.6 g (total = 101.8 g), while those kept in elevated CO_2 were 33.4 and 78.3 g (total = 111.7 g),

and were each fed 2 - 3 (0.2 - 0.4 g each) freshly sacrificed (by cerebral percussion) *A. polyacanthus* per day.

Both groupers from a matching treatment (i.e., control predators used for control trials and CO_2 acclimated predators used in CO_2 trials) were moved from their acclimation tank into the predator header tank (~20 L) in the morning of choice flume trials 90 minutes before the start of flume trials (trials were run for the control and CO_2 treatments on alternating days). The same header tank was used to house predators for flume trials throughout experiments, while the other header tank was kept as a predator-free tank. The two header tanks gravity-fed the two choice flumes and were constantly flushed with water from either a control or CO_2 sump at ~0.4 L min^{-1}, enough to ensure the header tanks remained full (with excess water spilling down the drain through a stand pipe) to maintain constant head pressure while also supplying one channel (at 135 mL min^{-1}) in each of the two flumes. The hoses that connected the header tanks to the flumes first passed through a valve system that allowed us to switch the predator chemical cue water from one channel to another (in both flumes) without any disturbance to the fish (visual screens were installed around the flumes.

After their use in the choice flume, the test fish were either transferred to a separate holding tank to keep them separated from fish yet to be used in trials, or they were sacrificed (cerebral percussion) and used to feed the two groupers not being used in that day's trials. Experiments took 4 d, and control and high CO_2 trials were performed on alternating days. At the end of all experiments, the *C. urodeta* were donated to a public display aquarium at AIMS.

Lizard Island Research Station 2016

Five two-channel choice flumes were used in parallel: one had total dimensions of 610 × 200 mm and choice arena dimensions of 200 × 200 mm with flow rates of 4 L min^{-1} per side and water depth of 100 mm; two had total dimensions of 290 × 93 and choice arena dimensions of 93 × 93 mm with flow rates of 675 ml min^{-1} per side and a water depth of 37 mm; and two had total dimensions of 235 × 45 mm and choice arena dimensions of 90 × 45 mm with flow rates of 135 ml min^{-1} per side and water depth of 22 mm (flow rates maintained with flow meters, as for AIMS 2015 described above). The setup of the primary header tank and secondary header tanks followed the description above for LIRS 2014, with the exception that the ambient seawater temperature was 29.5 ± 1ºC. Control CO_2 was ~520 µatm at this time of year. These conditions were maintained in the primary header tank when running trials on control fish, whereas CO_2 was increased to ~1,000 µatm in this tank (AT-Control System) when running the CO_2 treatment fish. pH and CO_2 were monitored as described above.

The secondary header tank used to manipulate predator cues contained seven or eight *C. cyanostigma* with an individual body mass of 170 - 350 g in each if the predator cue trials. Predators in the control treatment had been held for 1 - 17 days prior to commencement of experiments, while predators in the high CO_2 treatment had been acclimated for 4 - 18 days.

Laminar flow and mixing characteristics within the choice arenas of the flumes were tested regularly using food dye and were always found to be stable (e.g.: https://www.youtube.com/watch?v=jrtyc-rLGWc&t=9s).

Following the trial, the fish were measured for standard length (SL) and transferred to a new holding tank under their respective acclimation conditions.

Activity levels

Lizard Island Research Station 2014

All tanks were positioned within a large reservoir bath (L × W × H = 1,000 × 500 × 500 mm; water depth 70 mm) containing 35 L of water. The tanks had white sheeting placed between them to prevent interaction between fish in adjacent tanks.

Behavioural lateralisation

Lizard Island Research Station 2014

The central channel of the T-maze was 23 × 5 cm (L × W) and water depth was 5 cm. Each end of the central channel was blocked by a perpendicular opaque barrier (9 × 4 cm; L × H) set back by 4 cm, which prompted the fish to turn left or right around the barrier once it reached the end of the channel.

Australian Institute of Marine Science 2015

The central channel of the T-maze was 15 × 3 cm (L × W) and water depth was 3.5 cm. Each end of the central channel was blocked by a perpendicular opaque barrier (5 × 4 cm; L × H) set back by 4 cm.

In light of the study by Chung et al. (2014)[29], where it was reported that fish exposed to elevated CO_2 suffered impaired retinal function, as well as a study by Ferrari et al. (2012)[26] reporting that the response to visual cues may be impaired by CO_2, one end of the double T-maze was modified to examine the vision/cognition of fish exposed to elevated CO_2. Specifically, the barrier at one end of the central channel was offset by 5 mm to create a situation where the path around the barrier was shorter if the fish turned left compared with turning right. We predicted that control fish would display a left-shifted lateralisation

preference at the manipulated end of the T-maze, whereas fish under high CO_2 would not perceive the shorter distance afforded by the offset barrier and would thus maintain the same lateralisation pattern as they displayed at the opposite end of the T-maze. After each run the fish was measured for standard length to the nearest mm.

DATA ANALYSIS, STATISTICS AND BOOTSTRAPPING

General

Time spent in predator cue, and activity levels, were quantified for each minute of the fish's behavioural trial using tracking software, which meant many repeat observations for each individual. This design should, in theory, allow us to detect important interactions (e.g., across species), whereby an effect of acclimation treatment (control vs. elevated CO_2) occurs at the start of a behavioural trial but disappears thereafter, or vice versa. However, three limitations prevented the application of this approach. First, we had no *a priori* expectation that the effect of time would be linear across the trial, and visualization of the data confirmed that trends over time followed a variety of non-linear patterns. The use of smoothers in generalized additive mixed models (GAMMs; mgcv package in R[57]) is ideal for modelling non-linear temporal trends, so we explored their use here. A second problem was that the data were temporally autocorrelated, regardless of the use of individual fish ID as a random effect, therefore violating the assumption of independence (e.g., a fish's time spent in the predator cue in minute 4 was highly correlated with time spent in minute 5, but less so for minute 10). The use of a temporal autocorrelation structure was effective for addressing the independence problem based on autocorrelation function plots. However, the third, and most challenging problem with including an effect of time was that these data were censored between 0 and 60 s min^{-1} for the activity data, and 0 and 100% time spent in the predator cue for the flume data. The minute-by-minute data were bimodal around the minimum and maximum values (see Extended Data Fig. 3 for an example), not conforming to any distribution readily available for use in GAMMs (with the mgcv package in R). When those data were used in exploratory GAMMs (with a normal distribution), the residuals were widely and non-normally distributed, and the resulting models had very low R^2 values (<0.05). Thus, for simplicity, we took a mean across the entire trial for each fish (for choice flume and activity data; see below), which resulted in data being normally distributed and without autocorrelated repeated measurements, allowing us to use general linear models.

Response to predator chemical cues

Videos were analysed with ViewPoint tracking software (Zebralab, Lyon, France), which tracked fish movements in the choice flumes and quantified how much time per minute a fish spent on each side of the flume. In the 18 min trials in 2014 and 2015, we averaged 6 min prior to the cue side-switch and 6 min after the switch (i.e., 12-min means). In the 80 min trials in 2016, we averaged 16 min prior to the cue switch and 16 min after the cue switch (32-min means). These means were used in subsequent analyses.

To account for overrepresentation of *A. polyacanthus* in our dataset (n = 194 individuals across 2015 and 2016), we used the RAND() function followed by ascending sorting in Microsoft Excel to randomly select 120 individuals (60 control, 60 high CO_2) for inclusion in the bootstrapping procedure. This ensured that the sample size for *A. polyacanthus* closely matched the representation of the next most abundant species in the dataset (*D. aruanus*: 61 control, 60 high CO_2).

To enable robust comparisons of our data with previous publications, bootstrapping simulations (10,000 bootstraps with replacement) were conducted in R using the subsequent dataset of n = 247 control and n = 239 high CO_2 fish. Different bootstrapping simulations were run using samples of n = 10, n = 20 and n = 60 fish per treatment group, and frequency histograms were constructed of the mean percent time in predator cue for each of the treatment groups and each sample size scenario (Fig. 3a-c). Raw data from Welch et al. (2014)[25] presented in Extended Data Fig. 4 were obtained from: https://research.jcu.edu.au/researchdata/default/detail/eeed61ab874dcc1596ea4a19d9e0e468/.

Activity levels

Videos were analysed with ViewPoint tracking software (Zebralab, Lyon, France), which tracked fish movements in the behavioural arena and categorized the fish as inactive (< 0.5 SL s^{-1}) or active (> 0.5 SL s^{-1}). Data from the start of each trial were influenced by the setup process, thus we analysed a total of 27 min in 2014, 87 min in 2015, and 37 min in 2016.

Behavioural lateralisation

Testing for lateralisation is not straightforward because it involves multiple binomial experiments with structure. Since a binomial trial (a decision to turn left or right) is repeated typically 10 times per individual in the detour test, trials conducted on multiple individuals are non-independent. This data structure is different from a simpler situation, such as in a coin toss

or a trial involving a decision to turn left or right, repeated multiple times by a single individual. Here, each trial can be treated as independent.

Roche et al. (2019)[52] have developed and examined the sensitivity of two statistical approaches for testing population- and individual-level lateralisation. See Roche et al. (2019)[52] for a description of issues with the statistical approaches used by previous studies to assess lateralisation. A test for detecting lateralisation at the population level requires examining the mean lateralisation score across all individuals in the sample since population-level lateralisation is present when a group of individuals collectively exhibits a side-bias. Specifically, this test involves assessing whether the mean number of turns to the right (or to the left) is significantly different from 5, given 10 trials per individual. In contrast, a test for detecting individual-level lateralisation requires examining the sample variance since individual-level lateralisation is present when more individuals exhibit a side-bias than expected by chance (irrespective of whether it is to the left or to the right). For explanations and examples of these two concepts, see Bisazza et al. (1997)[48], Domenici et al. (2012)[53] and Roche et al. (2013)[54].

We tested population-level lateralisation with a generalized linear random-effects model (GLMM with glmer function in R) that sets the intercept equal to the grand mean of the data[52]. This model is specified as:

```
g <- glmer(X~1+(1|ind), data=dat, family="binomial")
```

where 'dat' is a data frame containing three columns: individual ID ('ind'); trial number (1 to 10); and turning side (column 'X') as right (1) or left (0). Specifying the error family as "binomial" sets the probability of success (*P*) to 0.5. Here, a right turn is considered a success. One *expects* a mean number of 5 turns to the right across all individuals if $P = 0.5$ (i.e., there is a 50% chance of turning left or right). We tested whether the *observed* mean number of turns to the right differs from the *expected* mean under a binomial error distribution with $P = 0.5$ by extracting the p-value for the model intercept as such:

```
summary(g)$coefficients[4]
```

If $P < 0.05$, there is statistical evidence that the population is lateralised (i.e., has a side-bias).

We tested individual-level lateralisation with a chi-square test comparing the observed variance (numerator) to the expected variance (denominator) assuming a normal approximation to the binomial distribution[52]. This is analogous to testing for overdispersion (i.e., are there

more observations in the tail ends of the distribution than expected by chance), and is achieved as such:

```
chi_sq=((N-1)*var(X1)/(n*.5*.5))
pchisq(chi_sq,df=(N-1),lower.tail=F)
```

where *n* is the total number of individuals tested; n is the number of trials per individual (i.e., 10); and X1 is a column in a data frame corresponding to the number of right (or left) turns per individual (each row of the data frame represents one individual).

If $P < 0.05$, there is statistical evidence that some individuals are lateralised (i.e., more individuals have an extreme lateralisation score than expected by chance based on a normal approximation to the binomial distribution with $P = 0.5$).

Supplementary Notes

Note on replication studies

Our study is, to a large extent, a replication of previous work published in the field. Two main types of replication studies exist: direct replications and conceptual replications, both of which are designed to evaluate the robustness of scientific findings[58-60]. A direct replication is "a study that attempts to recreate the critical elements (e.g., samples, procedures, and measures) of an original study where those elements are understood according to a theoretical commitment based on the current understanding of the phenomenon under study, reflecting current beliefs about what is needed to produce a finding"[60,61]. As such, a direct replication "does not have to duplicate all aspects of an original study. Rather it must only duplicate those elements that are believed necessary for producing the original effect"[60]. A conceptual replication "tests an extension of the theory to a new context"; for example, "whether an effect extends to a different population given theoretical reasons to assume it will be either significantly weaker or stronger in different groups"[60]. Some researchers believe that conceptual replications are more informative than direct replications because they give better evidence for the generalisability of an effect[62]. In the case of our study, we performed both a direct replication (for the species previously tested (*Acanthochromis polyacanthus, Pomacentrus amboinensis, Pomacentrus moluccensis, Dascyllus aruanus*) in the same way from the same populations), and a conceptual replication (for species not previously tested (*Chromis atripectoralis, Dischistodus perspicillatus*) as well as tested using a different variant of the detour test).

Note on life stages and definition of larval versus juvenile fishes

The majority of coral reef fishes (including five of six of our study species) have a planktonic larval stage, which can range from several days to several months depending on the species[63]. Seven to 39 days is typical for the damselfish family[63,64]. The spiny chromis (*Acanthochromis polyacanthus*) used here and in previous studies is among the few known coral reef fishes lacking a pelagic larval stage[63].

Light traps, which are used to collect wild fish in many CO_2 studies, typically catch late-stage (i.e., settlement stage) larvae just prior to settling on the reef, and thus, at the ecological transition from a pelagic to benthic mode of life[63]. In many species, settlement also coincides with a range of morphological changes that prepare a fish for benthic reef life, a phase known as metamorphosis[63,65]. McCormick et al. (2002)[65] state that "the term "metamorphosis" is used to encompass the changes in structure and function that occur as a fish takes on its juvenile form, which often coincides with settlement". Although this transition period can be long for some species, McCormick et al. (2002)[65] note that *Pomacentrus amboinensis*, one of the species that we studied, does not undergo significant structural changes following settlement and, in fact, settles directly into its adult environment. In both *P. amboinensis* and *P. moluccensis*, McCormick et al. (2002)[65] also found that pigment changes associated with metamorphosis occurred within 6h of capture with light traps[65]. As most studies keep light-trapped fish in the laboratory under CO_2 treatment conditions for a minimum of 4 days, they can no longer be referred to as larvae. Indeed, Allan et al. (2013)[66] state in their methods that they used "newly metamorphosed individuals of the common damselfish *Pomacentrus amboinensis*". Similarly, Lönnstedt et al. (2013)[67] specify that for their light-trapped *P. amboinensis*, "all fish were collected at the end of their larval phase." McCormick et al. (2013)[68] report collecting "settlement stage larvae" in light traps which were "reared for four days" in the treatments before testing. Ferrari et al. (2011)[69], who worked on *P. amboinensis* and *P. moluccensis* specify that "light traps collect these fish at or immediately before their arrival on the reef at the end of the planktonic larval stage", and subsequently refer to these fish as "young fish" or "juveniles" rather than larvae. Hence, for many CO_2 papers, including ours, it is more accurate to refer to the fish as juveniles rather than larvae.

Note on the importance of inter-individual variation

Inter-individual variation enables the persistence of populations and species and is a fundamental biological phenomenon acted upon by selection[70]. Inter-individual variation is

evident for most traits measured in wild animals[71,72]. The amount of variation among individuals depends on the trait(s) being measured. For example, in the study of animal physiology, sample sizes of $n = 6 - 8$ are typical for some variables because there is low inherent among-individual variability under common conditions (e.g., for resting heart rate, circulating levels of Na+, Cl- ions, blood pH). However, behaviours are inherently more variable, not only among individuals but also within individuals – although in most species, within-individual consistency in behaviours can typically be found[73,74]. Quantifying, appreciating, and understanding inter-individual variability is useful for understanding ecological and evolutionary processes, especially considering animal responses to environmental change in marine ecosystems such as ocean acidification. In our study, we observed substantial variation in how individual fish behaved in the test arenas (e.g., amount of time active, time spent in predator cue). This observed variation is in stark contrast to some previous findings in this field of research reporting negligible among-individual variation in the percentage of time fish spent in the side of a choice flume containing a predator chemical cue[4,5,16,17]. Our results, bootstrapping simulations, and the broader literature on inter-individual variation[73,75-77], help to highlight that results showing negligible inter-individual variation in behaviour should be interpreted with caution.

Further details, Fig. 3

Frequency outputs from bootstrapping simulations are presented for 10,000 bootstraps per scenario of the mean percent time in predator cue using 4-min means from the present study. Sampling was performed with replacement.

Note that Gould et al. (2015)[31] included a mean sample size around 40 and thus does not fit cleanly on any of the figure panels, but is included on panel e for comparative purposes. The extremely high variance in Welch et al. (2014)[25] (panel f) was caused by an exceedingly high proportion of control individuals reported to have spent 0% of their time in the conspecific chemical alarm cue (grey solid bars in Extended Data Fig. 4a) and an equally high proportion of high CO_2 individuals reported to have spent 100% of their time in the cue (blue solid bars in Extended Data Fig. 4b). The present study did not detect any such individual-level bimodality at 0% and 100% (open bars in Extended Data Figs. 4a and b).

Supplementary Information References

57 Wood, S. (2017). mgcv. Available from https://cran.rproject.org/web/packages/mgcv/index.html.

58 Kelly, C. D. Replicating empirical research in behavioral ecology: how and why it should be done but rarely ever is. *Quart. Rev. Biol.* **8**, 221-236 (2006).

59 Nakagawa, S. & Parker, T. H. Replicating research in ecology and evolution: feasibility, incentives, and the cost-benefit conundrum. *BMC Biol.* **13**: 88 (2015).

60 Zwaan, R. A., Etz, A., Lucas, R. E. & Donnellan, M. B. Making replication mainstream. *Behav. Brain Sci.* **41**, e120 (2018).

61 Nosek, B. A. & Errington, T. M. Making sense of replications. *eLife* **6**, e23383 (2017).

62 Yong, E. Bad copy. *Nature* **485**, 298-300 (2012).

63 Leis, J. M. & McCormick, M. I. In: *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem* (ed Sale, P.) 171-199. (Elsevier Academic Press, San Diego, 2002).

64 Wellington, G. M. & Victor, B. C. Planktonic larval duration of one hundred species of Pacific and Atlantic damselfishes (Pomacentridae). *Mar. Biol.* **101**, 557-567 (1989).

65 McCormick, M. I. Makey, L. & Dufour, V. Comparative study of metamorphosis in tropical reef fishes. *Mar. Biol.* **141**, 841-853 (2002).

66 Allan, B. J., Domenici, P., McCormick, M. I., Watson, S. A., & Munday, P. L. Elevated CO2 affects predator-prey interactions through altered performance. *PLoS One* **8**: e58520 (2013).

67 Lönnstedt, O. M., Munday, P. L., McCormick, M. I., Ferrari, M. C. O. & Chivers, D. P. Ocean acidification and responses to predators: can sensory redundancy reduce the apparent impacts of elevated CO_2 on fish? *Ecol. Evol.* **3**, 3565-3575 (2013). doi:10.1002/ece3.684

68 McCormick, M. I., Watson, S. A. & Munday, P. L. Ocean acidification reverses competition for space as habitats degrade. *Sci. Rep.* **3**, 3280 (2013).

69 Ferrari, M. C. O. et al. Intrageneric variation in antipredator responses of coral reef fishes affected by ocean acidification: implications for climate change projections on marine communities. *Global Change Biol.* **17**, 2980-2986 (2011). doi:10.1111/j.1365-2486.2011.02439.x

70 Forsman, A. & Wennersten, L. Inter-individual variation promotes ecological success of populations and species: evidence from experimental and comparative studies. *Ecography* **39**, 630-648 (2016).

71 Bennett, A. F. In *New Directions in Ecological Physiology* (eds Feder, M., Bennett, A., Burggren, W. & Huey, R.) 147-169. Cambridge University Press, Cambridge, 1987).

72 Réale, D., Reader, S. M., Sol, D., McDougall, P. T. & Dingemanse, N. J. Integrating animal temperament within ecology and evolution. *Biol. Rev.* **82**, 291-318 (2007).

73 Bell, A. M., Hankison, S. J. & Laskowski, K. L. The repeatability of behaviour: a meta-analysis. *Anim. Behav.* **77**, 771-783 (2009).

74 Roche, D. G., Careau, V. & Binning, S. A. Demystifying animal 'personality' (or not): why individual variation matters to experimental biologists. *J. Exp. Biol.* **219**, 3832-3843 (2016).

75 Dingemanse, N. J., Kazem, A. J., Réale, D., & Wright, J. Behavioural reaction norms: animal personality meets individual plasticity. *Trends Ecol. Evol.* **25**, 81-89 (2010).

76 Careau, V., Thomas, D., Humphries, M. M., & Réale, D. Energy metabolism and animal personality. *Oikos* **117**, 641-653 (2008).

77 Stamps, J. & Groothuis, T. G. The development of animal personality: relevance, concepts and perspectives. *Biol. Rev.* **85**, 301-325 (2010).