Galhardo et al. [1] present the results of an experiment designed to test how the social environment modulates the expression of two personality traits, exploration–avoidance and neophobia. The authors highlight that novel object (NO) tests are frequently used to measure the personality traits exploration–avoidance and neophobia, but that in most cases animals are tested alone, and as such the social environment is ignored. The authors presented an object to 12 male cichlids (Oreochromis mossambicus) when socially isolated, in the presence of a familiar conspecific female, and in the presence of an unfamiliar female. Males spent less time freezing (which I assume measured neophobia) and more time ‘exploring’ when in the presence of the familiar female. However, an individual’s behaviour in one context only correlated with behaviour in another context in one case: between the percentages of time spent exploring the object in the isolated and familiar female treatments. The authors conclude that ‘the lack of correlations across the three treatments for neophobia and exploratory behaviour indicates a lack of consistency in these measures, which are usually taken as measures of a temperament trait’. While I applaud the authors for attempting to address this overlooked aspect of animal personality, the study [1] suffers from several conceptual and methodological flaws. I thus feel their conclusions should be interpreted with extreme caution.

The authors assumed that the behaviours that they measured were indicative of the personality traits neophobia and exploration; however, it is unclear that they were measuring these traits. The authors conclude that the NO test ‘lacks convergent validity across social contexts’ [1]. Convergent validity investigates whether individuals’ behaviour correlates between two different tests designed to measure the same trait [2]. By completing the same test in multiple contexts, the authors are attempting to measure consistency (or reliability/repeatability, though it is more typical to do this in one context before extending to multiple contexts [3]). The authors rightly conclude elsewhere that their NO test lacks consistency [1], and thus show no evidence of having actually measured a personality trait (demonstrating consistency is standard in personality studies [3]). It is therefore premature to conclude that personality traits are modulated by social context [1]; the most the authors can conclude is that behaviour changes depending on the social context, which is not novel and should even be expected (for example, caching behaviour in birds [4]). Whether the documented behaviours reflect personality traits remains to be confirmed.

Further, the authors used only one ‘novel’ object for all three treatments. Both exploration and neophobia rely on novelty as part of their face validity [5]. After the first trial, the object will no longer be novel and individuals are likely to change their response to it through time. This would be controlled for in the current experimental design if there were no differences in rates of habituation to the experimental stimulus, however, differences in habituation rates should be expected a priori given evidence from other species [6,7], and it is not reported that this did not occur in the study [1].
Finally, from a methodological perspective, laboratory studies should perform appropriate controls. In this study, only in the unfamiliar female treatment is the tank next to the male disturbed before the presentation of the object (when the familiar female was replaced with the unfamiliar female) [1]. The males could have decreased their activity in response to the disturbance in the nearby tank before the NO was presented, or in response to the behaviour of the unfamiliar female (if her behaviour was different after capture/handling), and not have changed their exploration or neophobia in response to the change in social context per se. Further, differences in the males’ behaviour before and after the presentation of the NO in each treatment would present a more appropriate unit of analysis. For example, in a laboratory study of boldness in sticklebacks (*Gasterosteus aculeatus*) that compared behaviour before a treatment to behaviour during a treatment, there were no significant differences in the individuals’ behaviour in the treatment and the control. The authors concluded that the treatment was not measuring the targeted personality trait and a different trait was being measured [8].

Understanding the effect of the social context is an important future direction for personality research, however it is not clear whether Galhardo et al. [1] (i) measured a personality trait, and (ii) whether the individuals’ observed behavioural differences were owing to the social context rather than the experimental design. Care should be taken when designing personality studies [3] so that strong conclusions can be drawn.

Many thanks to LARG for helpful discussions.

References