

FIGURE 2

- Use same colours in figure A for the groups, this was done correctly in figure B
- I am still somewhat puzzled by the use of the ANOVA in the top graphs. What does this analysis add to the data compared to the lower two ANOVA's? I feel that a t-test would suffice to assess differences within a condition (predator, non-predator and non-stimulus). In the bottom two graphs the comparison is made between conditions within one fish type (sender and receiver). One could combine all of this into a single graph by using capital letters and lowercase letters to distinguish between two ANOVA's
- I am not sure an ANOVA can be performed over the data in figure 2B. Having only 3 data points seems too few for a reliable data distribution analysis. Perhaps assess the data with a non-parametric analysis and compare the outcome?

VIDEO ANALYSIS

- The authors explain that they assessed video for the full 60 (55 without the first 5 min) minutes. However, why did they not repeat the procedure and split their data in time-frames of 10 minutes?
- The authors mention in their rebuttal that the first 5 minutes were deleted from the video? Why not compare this to the overall average to see if it indeed shows different behavior? Please include the information on excluding the first 5 min of the video in the video analysis section.
- In addition, I noticed that it was not noted how many people analysed the video. Was the video analysed by a single person? Was the analysis done blind (i.e. without knowing what group was being assessed?). Please include.

CORTISOL

- Please provide a reference for the minimal weight of tissue needed for cortisol extraction and method.
- Why did the authors not exclude fish that were too small or opted to pool all fish in groups of two (perhaps based on weight – heaviest + lightest, so all pooled had a similar weight). I am not sure the pooling of some fish and not others is an appropriate solution. At least indicate which samples were pooled.

NUMBER OF EXPERIMENTAL REPLICATES

- The Manuscript states that each experiment was conducted in triplicate, yet I have trouble finding where the results of these replicates are reported? FIGURE 2 only shows 7-8 data points per group for cortisol. Even with pooling a few fish into a single data point, there seems to be data missing (3 replicates, 10 fish per experiment = 30 total per group as is indicated by the authors in section 2.4). Did the authors observe differences between replicates (did they test this)? Have they pooled all samples into a large group and if so, why was this done? Where did the missing samples go? The raw data sheet shows 45 data points analysed for cortisol out of 240 fish used.... This seems extremely low.
- FIGURE 2B also shows error bars, suggesting this is where the triplicates come into play (a single data point per replicate)? How does the cortisol data then relate the behavioural data?

- FIGURE 2B How did the authors come to 3-4 data points per group? There were only three replicates per group, so where would a 4th data point come from? The RAW data pdf shows 4 data points for all groups (4 replicates per group?).

FEEDING FISH

- Feeding was not done ad libitum. Ad libitum feeding would implicate 'free-feeding' as is seen in rats and mice, where food is available 24h per day. Feeding here, was done till satiation (as much food as fish can eat at a given feeding moment).

ADDITIONAL CONTROLS

- Why were these groups included? They are not really compared to other groups in the Manuscript. What information do they provide to the reader? I found 'no changes in behaviour' meaning they are equal to the behaviour seen in the no stimulus fish RF group? Similarly, a cortisol value is given, but what does that mean? Cortisol seems much lower than the values found in any of the other groups analysed. Thus there seems to be an effect on cortisol when fish can see, but not 'smell' or 'interact' with them?