November 1 2016

RE: COMMENTS ON RESPONSES TO 20 QUERIES REGARDING LÖNNSTEDT AND EKLÖV 2016 *SCIENCE*

To:

Prof. Eva Åkesson	Vice-Chancellor, Uppsala University
Prof. Anders Malmberg	Deputy Vice-Chancellor, Uppsala University
Prof. Johan Tysk	Vice-Rector, Faculty of Science & Technology, Uppsala University
Prof. Lars Tranvik	Head, Department of Ecology and Genetics, Uppsala University
Per Andersson	Utbildningsledare, Uppsala University
Lena Berke	Chairman, Central Ethical Review Board
Jörgen Svidén	Chief Secretary, Central Ethical Review Board

Please find below our comments (in **blue**) regarding the responses provided by Prof. Eklöv and Dr. Lönnstedt (in **orange**) to the 20 queries that we initially raised after the *Science* paper was published (in black). We feel compelled to reply because the authors' answers are largely inadequate, do not clarify matters and fail to address many of the questions we posed.

We wish to reiterate that these queries primarily relate to flaws in the *Science* paper that are identifiable by any well-trained, objective reader, and hence do not contain any of our direct allegations of research misconduct. The allegations were listed in point form (from A to K) in the document "Timeline Lönnstedt and Eklov.docx" (sent to Per Andersson at Uppsala University on July 19, 2016). Nonetheless, many of the issues discussed in this document relate to our allegations of misconduct, as existing evidence indicates that the authors' responses are inaccurate.

The original document with the queries (Lonnstedt_Eklov_Science_2016_queries.pdf) was sent by the signatories of this document to the authors (Lönnstedt and Eklöv) on June 16, 2016, and to the office of the Vice-Chancellor of Uppsala University on June 20, 2016. Much confusion and misinformation surrounds the subsequent handling of this document, so we wish to highlight the chain of events following its original submission.

The report of the investigators appointed by Uppsala University (UFV 2016/1074) states that we (the signatories of this document) had been provided with answers to our 20 queries at the time the report was filed on August 31, 2016. This is incorrect.

The authors, Eklöv and Lönnstedt, declined to send their responses to us when asked over email by Dr Fredrik Jutfelt on July 11, 2016. Per Andersson (administrator for errand UFV 2016/1074 at Uppsala University) informed Jutfelt that he did not have access to the authors' responses since he was not part of the investigative committee (email correspondence). However, Prof Per Jensen (a member of the investigative committee) stated in an email that all correspondence and documentation regarding the investigation had been sent to the investigative committee via Per Andersson. In frustration, we sent a document titled "Request for access to document containing answers to 20 queries regarding Lönnstedt and Eklöv 2016 *Science*" to Uppsala University and to the Swedish Central Ethical Review Board (CERB) on October 3, 2016. We were provided with the authors' responses via email by Eva Kaaman Modig (administrative secretary, CERB) the same day as our request was submitted. Eva Kaaman's email stated that this document was sent to Uppsala University on July 27, 2016, almost three months before we received a copy. We would also like to point out that this

document is not listed in the ärendekort (the errand card for case UFV 2016/1074), where all documentation regarding this case should be listed.

Yours sincerely,

Dominique Roche, Ph.D., Institute of Biology, Université de Neuchâtel
Fredrik Jutfelt, Ph.D., Department of Biology, Norwegian University of Science and Technology
Josefin Sundin, Ph.D., Department of Neuroscience, Uppsala University
Graham Raby, Ph.D., Great Lakes Institute for Environmental Research, University of Windsor
Ben Speers-Roesch, Ph.D., Department of Biology, University of New Brunswick
Sandra Binning, Ph.D., Institute of Biology, Université de Neuchâtel
Timothy Clark, Ph.D., Institute for Marine and Antarctic Studies, University of Tasmania

---

Dear Oona and Peter,

We read your recent Science paper on microplastics with great interest (Lönnstedt and Eklöv, June 03 2016). None of us work on microplastics but we all have a keen interest in experimental studies testing the effects of environmental stressors and pollutants on fish behaviour and physiology. While we find the topic of the study very timely and important, we are hoping that you can clarify several issues which we found concerning. Our main concerns relate to missing data (stated in the paper as being available in the supplementary materials and in an institutional repository), inconsistencies in the sample sizes reported and the microplastic concentrations used, and issues with the statistical design and analyses. We have listed detailed comments and questions below, many of which could affect the validity of the results presented in the paper. Given the implications of your paper for conservation and policy, we feel it is essential that it meets the highest reproducibility and transparency standards so as to set an example for future research in this field of study.

Thank you in advance for taking the time to answer our queries.

Best regards,

Dominique Roche, Institute of Biology, Université de Neuchâtel
Graham Raby, Great Lakes Institute for Environmental Research, University of Windsor
Ben Speers-Roesch, Department of Biology, University of New Brunswick
Sandra Binning, Institute of Biology, Université de Neuchâtel
Josefin Sundin, Department of Neuroscience, Uppsala University
Fredrik Jutfelt, Norwegian University of Science and Technology
Timothy Clark, Institute for Marine and Antarctic Studies, University of Tasmania

Methods and results:

Inconsistencies in the methods/results:

1) There appears to be a large discrepancy between the concentration of microplastics in the average and high microplastic treatments and the number of particles that were ingested by larval fish.

In the average and high treatments, 58-60 eggs were exposed to a concentration of 10,000 or 80,000 particles per m^3, respectively. It seems from the methods in the supplementary material that hatched larvae were subsequently transferred to new 1000ml aquaria at a density of 15-25 larvae per aquarium.

The average microplastics treatment contained a concentration of 10,000 particles per m^3, so 10 particles per 1 L aquarium. Since there were 15-25 larvae per aquarium, the mean number of particles ingested by fish should range between 0.4 (in the case of there being 15 larvae) and 0.7 particles (in the case of there being 25 larvae), assuming that all particles were ingested. However, the value presented in the paper is 1.4 ± 0.35 particles (from the text, it is unclear whether ± 0.35 represents a standard deviation or a standard error). To reach this value, the number of particles in the 1 L aquariums would have had to range between 21 and 35. Given that 60-70% of the particles sank to the bottom (as stated in the paper) and would not have been viewed by the perch as planktonic food, the number of particles in each 1 L aquarium would have probably had to be between 53 and 117. These required microplastic concentrations are 5 to 11 times higher than the value of 10 particles/aquarium that you have specified in the paper. Can you please explain this large disparity?

The high microplastics treatment contained a concentration of 80,000 particles per m^3, so 80 particles per 1 L aquarium. Since there were 15-25 larvae per aquarium, the mean number of particles ingested by fish should, in this case, range between 3.2 and 5.3 particles. However, the value presented in the paper is 7.15 ± 1.2 particles (again, it is unclear whether ± 1.2 is a standard deviation or a standard error). To reach this value, the number of particles in the 1 L aquariums would have had to be between 107 and 179, assuming that every particle present in the aquarium was ingested by the larvae. As noted above, the fact that most of the particles were not suspended in the water column means that the required number of particles would have had to be 268 to 597 per aquarium, which is 3 to 7 times higher than the value of 80 particles/aquarium specified in the paper.

The photo in Fig 3J shows a perch larva containing 19 particles in its digestive tract. Was this fish an outlier? The paper states that "a subset of randomly chosen individuals" were photographed for stomach contents. Can you clarify how many fish were randomly selected and photographed and please include these photographs with the raw data accompanying the paper as per Science's data accessibility policy (please see comment #20 on data availability)?

1) Response from PE and OL: This is something other people have had queries about and we understand that this must be explained in greater detail. With respect to the ingested microplastic particles we randomly sampled 20 fish from each treatment and counted the number of plastic particles in their stomachs. These fish were sampled from fifteen different tanks (five tanks per treatment with four fish collected from each tank). So with regards to number of ingested particles: out of 80 total fish in the average treatment (taken from the predator exposure setup tanks that usually contained closer to 15 fish per tank), we randomly sampled 20 fish (25%) for gut contents. One of these randomly sampled fish had ingested 6 particles and 2 had ingested 3 particles (which is approximately 1/4 of the total microplastics for average concentrations in only 4% of the total number of fish in average concentrations!). For the high treatment this ranged from between 2 to >20 particles. So there is a large spread in number of ingested particles in each fish (the values reported are standard error). This seems to be tied to personality as our more recent studies suggest that bolder individuals tend

to rapidly ingest a greater number of particles. Science has received a file which contains pictures of fish that have ingested microplastic particles.

1) Our response: As you state, other people have ongoing concerns regarding the discrepancy between the particle concentrations reportedly used in your experiment and the number of particles ingested by fish (http://science.sciencemag.org/content/352/6290/1213.e-letters). This clearly shows that the *Science* paper and supplementary materials did not satisfactorily explain important details of the experiment, and these concerns remain unaddressed. Moreover, your response above is insufficient to clarify many of the issues that have been raised regarding the *Science* paper, as detailed below:

- Given the information you provided in your response above, the mean number of plastic particles ingested by the fish in the average treatment was 0.6 (± 0.5 SEM). **This is different from the result provided in the paper**, where it is stated "*fish from the average microplastics treatment consumed 1.4 ± 0.35 plastic particles*". Your response above therefore further adds to our confusion and concern.

- Microplastic particles must be suspended in the water column to resemble plankton in order to be eaten by the perch larvae. Given that 60-70% of the microplastic particles sank to the bottom of the tanks (as stated in the *Science* paper), the tanks in the "average concentration treatment" contained a maximum of 4 suspended particles to be shared across 15 fish (this is the number you state in your response above, but refer to query #2 below). **How is it possible that one fish in the "average concentration treatment" ate 6 particles when there were only 4 particles available?** Surely, this means that the other 14 fish in the beaker ate zero particles. It appears from your response above that the sampling was not random if you preferentially sampled individuals that consumed the most (or all) microplastic particles.

- The updated supplementary material that you sent to *Science*, which was available online on July 7, 2016 (approximately 1 month after the paper was first published), contains 35 photos of larval fish (http://science.sciencemag.org/content/sci/suppl/2016/06/01/352.6290.1213.DC1/Lonnstedt.SM.pdf). However, given the information you provide in your response above, there should be 60 photos in total (20 photos for each of the three treatments). **Where are these additional photos?**

- There are no captions or metadata associated with the photos you have sent to *Science*, **which makes their interpretation impossible**.

- The data supplied to *Science* as supplementary material and allegedly "*archived at the research database at Uppsala University*" should contain the date, tank number, treatment, and number of particles ingested by each fish in order to be interpretable. **However, none of these necessary data are included** (refer to the document "Appendix 29", sent to Per Andersson at Uppsala University on July 19, 2016, and pasted at the end of this document, where all outstanding data are listed). *Science*'s data policy states: "*After publication, all data and materials necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of Science. All computer codes involved in the creation or analysis of data must also be available to any reader of Science. After publication, all reasonable requests for data or materials must be fulfilled. […] Unreasonable restrictions on data or material*

availability may preclude publication". The policy can be accessed here: http://www.sciencemag.org/authors/science-editorial-policies.

- What exactly are you referring to by "*taken from the predator exposure setup tanks*"? **This comment does not align with any treatment that is described in the *Science* paper.**

2) Sample sizes are reported as N, which is the symbol typically used to indicate total sample size (across all treatment groups). What does N represent in the paper – is it the total sample size (N) or the sample size per treatment group (typically denoted as n)? Even if N denotes the sample size per treatment group, we are having difficulties reconciling the sample sizes presented in the paper with the number of larvae that are said to have been exposed.

For example, the sample size for the behavioural trials is indicated as N=36 in the paper. Assuming this means n = 36 per treatment (i.e. n = 12 per cue injected, including the conspecific cue, heterospecific cue, and water control), this implies that 108 fish were used per treatment for this experiment (3*36=108 since 3 fish were used in each trial). If n=36 per cue in each treatment, this would suggest that 324 fish were used per treatment (three cues per treatment: 36*3=108, 3 fish per trial: 108*3=324).

The sample size for the predator trials is indicated as N=45-47: was that per treatment or were ~15 fish used per treatment?

The paper states that N=20 fish were used for growth and stomach analyses. Was this per treatment (i.e. n = 20 per treatment) or in total (i.e. n = 6 to 7 per treatment)?

Adding up these numbers gives the following potential number of fish used per treatment:

Low-end estimate: 130 fish used per treatment (108 fish were used in behavioural trials, 15 were used in predation trials and 7 were used for growth and stomach analyses; 108+15+7=130). 130 fish/treatment * 3 treatments = total of 390 fish used in the study.

Mid-range estimate: 160 fish used per treatment (108 fish were used in behavioural trials, 45 were used in predation trials and 7 were used for growth and stomach analyses; 108+45+7=160). 160 fish/treatment * 3 treatments = total of 480 fish used in the study.

High-end estimate: 173 fish used per treatment (108 fish were used in behavioural trials, 45 were used in predation trials and 20 were used for growth and stomach analyses, 108+45+20=173). 173 fish/treatment * 3 treatments = total of 519 fish used in the study.

All of these numbers are higher than the total number of larvae exposed per treatment (stated as 75-125 in the paper). How do you reconcile these differences in actual versus stated sample sizes?

2) Response from PE and OL: There seems to be a misunderstanding here with regards to number of fish used and we are happy to clear it up for you. The behaviour (N=36; which we meant as sample size per treatment group) was first recorded for fish from each treatment and then 12 of these 36 fish were exposed to one of the three stimulus cues (that is N=12 for stimulus cues). This gives a total of 108 fish in the behavioural studies. In predator exposure studies 45-47 individual fish were exposed to predators from each study. Then 20 randomly

selected individuals were photographed for size and stomach analysis. This gives a total of 173-175 larvae for each treatment. The number of fertilized eggs per treatment (as stated in the study) was 290-300 (five tanks per treatment with 58-60 eggs per tank). In the high microplastics treatment 19% of these eggs didn't hatch out but there was still plenty of larvae left to cover the number of fish needed in the experiment.

2) Our response: The numbers you present do not add up:

- **N=36 per treatment group does not equate to 108 fish used in the behavioural trials as you state in your response**. The *Science* paper states that behavioural assays were conducted using three fish at a time, with a single focal individual being observed, and that no individual was used more than once. The experiments were performed as 3 min observation followed by a cue injection, and then another 3 min observation. This means that only one focal fish could be watched throughout the 6 min observation period, and the other two individuals could no longer be used because they had already experienced the cue injection. Hence, based on the information you have provided, N=36 per treatment means you used 324 fish in total for the behavioural trials across all three treatments (3*36*3=324).

- The number of larvae exposed per treatment is stated as 75-125 in the paper (based on reported densities of 15-25 larvae per aquaria, 5 aquaria per treatment gives 15*5=75 to 25*5=125 larvae per treatment). However, here you say that the sample size was "*173-175 larvae for each treatment*". **Hence, even if disregarding our point above, your own calculation above sums up to many more fish than were actually used according to the *Science* paper**. Further, in your response to our query #1 you wrote that there was a total of 80 fish in the average treatment. **We do not understand how you could use twice as many fish in your experiments as what you had available**, given each individual was used only once, as stated in the supplementary material.

- You say in your response that the number of eggs per tank was 58-60 and that there were five tanks per treatment. Therefore, the total number of eggs required for the experiment was 870-900 (58*5*3=870, 60*5*3=900). This number far exceeds the actual number of fertilized eggs obtained from egg strands that Sundin collected for Lönnstedt. On 4 May, 2015, Lönnstedt told Sundin that 100 eggs were fertilized from two out of the four egg strands collected (Appendix 25 related to the document "Timeline Lonnstedt and Eklov.pdf"). Since all four egg strands were approximately the same size, and undoubtedly contained a similar number of fertilized eggs, there would have been approximately 200 fertilized eggs in total. **Therefore, there is at least a 4-fold discrepancy in the number of eggs that were available (~200 eggs) and the sample size reported in the *Science* paper (>870 eggs)**.

3) The paper states that water replacements were conducted daily to remove uneaten food but without removing the 0.09 mm plastic particles: "Each day, a water change (ca. 300-400 mL) was carried out which also ensured that any food not ingested was removed from the aquarium. Care was taken to ensure that no microplastic particles were sifted out during the water change." This seems like a fairly challenging procedure as 30-40% of particles were suspended in the aquarium (as stated in the paper) and food would be interspersed with the particles due to mixing by the airstones. Can you please provide details on how the water and uneaten food were drained without removing any microplastics?

3) Response from PE and OL: The water change was easily done by turning off the airstones and allowing microplastic particles to sink towards the bottom half of the tank. Water could then be carefully sifted out from the top part of the tank. When there was leftover Artemia on the floor of the tank these were removed using a fine suction tube (it works like a tiny vacuum cleaner). The 'slush' was then looked at in a microscope and any microplastic particles accidentally removed were returned to the aquaria.

3) Our response: The statement that water replacements were done daily is incorrect. Lönnstedt was away from the research station for two out of the 10 days during which she conducted the microplastics experiments. During these two days, none of the other researchers present at the station were asked by Lönnstedt to perform water changes. Moreover, the procedures you describe above (siphoning 'slush' from all of the experimental beakers and carefully searching for 0.09 mm microplastics under a microscope) must have taken a considerable amount of time. Thus, your response raises serious questions, particularly because neither Jutfelt nor Sundin observed Lönnstedt performing any water changes or microscopy to sort 'slush' despite sharing the laboratory space with her throughout the course of the microplastics experiment.

4) The paper's conclusion states "Here it appears that larvae preferentially feed on plastic particles." However, standard or small strain artemia nauplii (420-480 µm) can only be eaten by perch larvae ~2 to 3 days after hatching (Vlavonou et al 1999, Cuvier-Peres and Kestemont 2002). As such, the perch larvae could have ingested microplastics (90 µm) during this period, but not food, potentially biasing your result. Can you indicate what strain of artemia was used to feed the perch as this is not indicated in the supplementary material? The paper also states that the larvae in the high treatment ate only microplastics, not artemia. How is it that these survived without food for 14 days post hatching, given that you state (in the supplementary materials) that the yolk sac in this species is completely absorbed within 5 days after hatching?

4) Response from PE and OL: As is stated in both the main text and supplementary materials it only appears as larvae preferential1y feed on microplastics. We can of course not state this with entire certainty, perch may have fed on Artemia previously and passed this and then have eaten rnicroplastics just prior to being photographed. However, the stomach content pictures of control larvae suggested they had ingested only Artemia and perch in high microplastics treatments only had ingested microplastics. When larvae started feeding Oona had brought with her a small homemade Artemia hatcher which allowed her to hatch out appropriately sized food particles. This is the same one we have used for experiments at Uppsala University and it works remarkably wel1.

4) Our response: Several outstanding questions remain:

- **How is it possible for the fish in the high microplastics treatment to have survived 14 days post-hatch without any food**? Is there any evidence in the published or grey literature that perch larvae can survive 14 days post-hatch without food at the experimental temperature used in the experiment (~15-20˚C)?

- **It is incorrect that Lönnstedt used a homemade artemia hatcher at the research station to obtain the food used for her experiments**. She used artemia that the other researchers at the station (Rafael Augusto, Mário Cunha, Josefin Sundin) were producing for their own experiments in the hatchery room. Thus, your response

contrasts with the observations of three eyewitnesses. However, regardless of the strain of artemia used, the fact remains that perch larvae can only start feeding on artemia ~2 to 3 days after hatching (Vlavonou et al 1999, Cuvier-Peres and Kestemont 2002). This means that the perch larvae were only able to ingest microplastics (90 µm) but not food (420-480 µm) during this period, which has major implications for the results reported in the paper.

5) The paper states that "standardized protocols" were used for conducting behavioural observations. However, the two papers cited are studies conducted by the first author and the methods are, in fact, not common in the wider behavioural literature. As stated in the supplementary methods, the experimenter simultaneously recorded three behavioural measures on one focal individual (with two other fish present in the tank) during trials that lasted 6 min 15 s (3 min + 15 sec + 3 min). The behaviours measured were (i) activity level (ii) total distance moved, and (iii) the amount of time focal fish were immobile in the water column. Based on our experience, it doesn't seem feasible that an observer could reliably and simultaneously measure so many variables during ~ 6 min without filming the trials. It seems likely that the measurement accuracy would have been greatly improved by filming the trials and analyzing videos. How could the experimenter keep track of the focal individual (among the other two fish in the aquarium) while simultaneously writing notes and manipulating equipment to inject an alarm cue into the tank? Were any of these trials filmed, and if not, why?

5) Response from PE and OL: This is not entirely true and may be more a matter of difference of opinion. If you look into the wider behavioural literature (especially on that of chemical alarm cue research with Brian Wisenden, Grant Brown, Douglas Chivers and Maud Ferrari as well as that of coral reef behaviourists such as Mark McConnick, Phil Munday and Danielle Dixson) you can see that the protocols used in the current study are fairly standard procedure. And due to the space limitation in a Science report we chose to cite two of the first author's studies as these were already needed in the paper. We understand that to someone that is used to filming all their trials it may seem like an impossible task to look at more than one behavioural measure at a time. However, where the first author was trained the more senior behaviourists can collect up to 6 behavioural variables at any one time, while injecting alarm cues and this being done while SCUBA diving! The first author would not dare such a feat but at least three variables are possible to observe with a structured protocol. However, filming is something we may consider doing in the future.

5) Our response: This response does not satisfactorily address our concerns. **How exactly did Lönnstedt measure and record three behaviours simultaneously, during two distinct time periods for each trial, while injecting cues with a syringe, while never taking her eye off the focal fish that was swimming with two other individuals?** The authors have already made the following statement in response to Uppsala University's preliminary investigation:

"About laboratory notebooks: At JCU, where Oona is trained and worked for 10 years before she came to UU, laboratory notebooks was not used in the same manner as in Sweden (since most experiments were made by divers under the water everything was written down on waterproof paper and this was later typed in the computer), so a aquaria-journals/lab-book is unfortunately not a routine that Oona has with her so there is none available. But of course Oona was there and did all the experiments as stated."

Thus, it appears that Lönnstedt did not use a notebook during the experiments because she was not trained to use a notebook during her PhD at James Cook University. Notably, this is the same training environment where "*senior behaviourists can collect up to 6 behavioural variables at any one time, while injecting alarm cues and this being done while SCUBA diving!*". **Where did Lönnstedt record the data she collected during the trials if not in a notebook or on paper?** Based on your response, she either made a mental note of what she observed, or scored behaviours directly in an electronic spreadsheet on her laptop, while not taking her eye off the fish. Lönnstedt subsequently claimed that her computer and all her external hard-drive backups were stolen when she left them in an unlocked car.

6) On top of the logistical difficulties associated with our above points, the paper also states that the behavioural experiments were performed blind (i.e. the experimenter was unaware of the microplastic treatment groups when performing the experiments) – this is stated at least twice in the supplementary material. In addition to of all the other complications associated with measuring behavioural data with the naked eye, we do not understand how blind data recording was achieved in this study. How many people were involved in performing the experiments?

6) Response from PE and OL: This is something one must master when not filming the trials. We use a coding system which was taught during Oona's postgraduate work. All test tanks and syringes have a code which corresponds to the treatment. However, the observer is unaware of which treatment at the time of testing as there is only a code present.

6) Our response: This response does not clarify our concern. The answer given above pertains to the chemical alarm cue treatment administered to the fish from the different microplastic treatment groups. Thus, there are two levels of blinding needed for this experiment: 1) blinding to the chemical alarm cue treatment and 2) blinding to the microplastic treatment. Our initial question concerns the supposed blinding of the microplastic treatment as stated in the supplementary material, which remains unanswered. **Blinding seems impossible given that Lönnstedt was working alone and hence was the sole person setting up the treatment tanks, performing alleged daily water changes, fetching the focal animals from the treatment tanks, conducting the behavioural observations, organising the data, and running the statistical analyses.** Thus, Lönnstedt could not have been blind to the microplastics treatment of each focal fish when performing the behavioural trials, even if she claims that she was blind to the chemical cues being injected.

Although we did not ask about cue injection with syringes, this too seems impossible for one person to achieve in a blinded manner – **a coding system cannot work effectively when the same person formulates and assigns the codes, fills the syringes with the treatment water, performs the injections shortly afterwards, and observes the behaviours**.

7) We do not understand the distinction between two of the behavioural variables that were described in the 'Behaviour' section of the supplementary methods. It seems to us that a fish's activity level (measured as the total number of lines crossed on a 5 x 5 mm grid present at the bottom of the tank) and total distance moved (measured as the total distance that fish swam during the 3-min observation period; in mm, assessed using the grid) are the same thing, only expressed in different units. If this is the case, why are both of these measures needed and why are the values presented in Fig.1A not correlated to the values presented in

Fig. 1B? Shouldn't values in Fig. 1B be the exact same values as in Fig. 1A, except multiplied by five?

7) Response from PE and OL: We had a reviewer ask us the same thing and he/she agreed with the point we made below. A fish can cross a line back and forth multiple times without actually swimming very far (they can just go back and forth across a line while sitting in the same spot). Distance moved is an important variable to see how far a fish actually swims. This is why we and our editor decided on keeping both variables in the manuscript.

7) Our response: Since a reviewer pointed this out during the review process, this explanation should have been included in the supplementary material. Based on your explanation above, active and inactive fish can cross the same number of lines. Therefore, what is the ecological relevance of the metric "number of lines crossed"?

8) Differences in distance moved/lines crossed (Fig. 1 A, B) seem to be due to differences in the amount of time fish spent inactive (Fig. 1 C). Fish in the high plastic concentration treatment spent more time inactive than in the other two treatments. If mean distance moved/number of lines crossed is calculated using a time value excluding time inactive, then there would appear to be no difference between treatments. Do you agree with this interpretation? This will be easy to check once the data are available (see comment #20 below).

8) Response from PE and OL: We're not really sure what you mean here. Fish in high microplastics treatments were less active than the other fish in all variables tested. And the raw data has been sent to Science.

8) Our response: The data to answer our query were not available at the time because 1) we were not sent the data despite several email requests (we were sent data from another paper and ANOVA tables with statistical outputs); 2) the servers at Uppsala University were apparently malfunctioning when Lönnstedt tried to 'archive' the data (see below); and 3) according to Eklöv and Lönnstedt, *Science* failed to upload the raw data on its website (refer to document UFV 2016/1074; *Preliminary investigation concerning the alleged scientific misconduct*).

Note that making a backup of your data on an institutional server is not the same as archiving data on a public repository (e.g. Dryad, figshare, Zenodo). ***Science*'s data policy requires that all the data necessary to reproduce your analyses be publicly accessible**, not simply backed up on a hard drive or a server at Uppsala University. "Data must be available either in the main text or Supplementary Material, archived in an approved database (with an accession number included in the acknowledgements)."
See: http://www.sciencemag.org/authors/science-editorial-policies.

It seems highly unlikely that the servers at Uppsala University were continuously malfunctioning between May 2015 and June 2016. In your response to Uppsala University's preliminary investigation (UFV 2016/1074), you state: *"There is an automatic syncing of your computer when you log on to the university network"*. Hence, you imply that the University network was malfunctioning between May 2015, when the mircoplastics experiment was completed, and June 2016, when Lönnstedt's computer and backups were allegedly stolen. Can you produce documentation to support this?

Using the partial raw dataset that was provided by the authors (1 month after it was first requested), **we have performed the analyses to answer the question we asked**. If time inactive is deducted from total time, and distance moved is calculated using the time that was left for the fish to move, i.e., using distance moved per second, and number of lines crossed per second as response variables (as is common in behavioural research), **there is no effect of microplastics treatment** (ANOVA: distance moved s^{-1}, treatment: P=0.791, lines crossed s^{-1}, treatment: P=0.896). **These results contradict the findings you present in the Science paper.** See box plots of the data at the end of this document and the dataframe and R script used for these analyses: https://figshare.com/s/ad983e9ca17d344a81c5.

9) The paper states that hatching of *P. fluviatilis* occurs in 1-3 weeks depending on temperature. It also states that hatching success was determined over 3 weeks, suggesting that the water temperature in the experimental aquaria was low. Please provide data documenting the temperature and salinity recordings in the experimental setup - the data provided in the paper are for mean monthly temperatures of the Baltic Sea and Lake Bäste. We are concerned about this because there is no mention of temperature regulation, such as a heating and/or a cooling system in the aquaria. Without temperature control, the water in the 1L aquaria would quickly equilibrate with the room temperature. This is potentially problematic if the objective of your study was to mimic the natural environment. There is also a lack of information on how perch eggs were kept before the exposure treatment started. The references cited on rearing methods are on cod and turbot, not perch. There is no information on when the eggs were collected and when the exposure started; how long were the eggs kept at the field station before they were placed in the microplastic treatments?

9) Response from PE and OL: Hatching normally occurs within 1-3 weeks, and in the current experiment they hatched quite rapidly, some within <3 days of being placed in the treatments. However, hatching was monitored for the entire 3 weeks after the larvae were first brought to the station to give the larvae in all treatments enough time to hatch (when eggs were found dead these were of course removed). The values for the Baltic Sea were taken by Josefin Sundin in your group so we're sure she is happy to provide you with these. We took measurements from the beakers that contained Lake Baste water. Unfortunately we can't provide you with this data as Oonas computer was stolen along with her backups three weeks back. Perch eggs were collected at the end of April and brought back to the research station where they were kept in flow through Lake Baste water. At the start of May they were then sampled under microscopes (to see which were successfully fertilized) and transferred into their different experimental containers. Containers were all exposed to the same environmental conditions and although there was no temperature regulation the measurements taken suggest there were no extreme anomalies in temperature and/or salinity of the aquariums.

9) Our response: Several aspects of this query remain unanswered. Additionally, the *Science* paper and associated supplementary material lack critical details, including the exact dates when the experiment was executed. We know, however (and have provided ample evidence to Uppsala University's investigative committee and to the Swedish Central Ethical Review Board), that **Lönnstedt and Eklöv's microplastics experiment was conducted over a period of only 10 days between May 5 and 15, 2015**. It is therefore impossible that eggs were maintained for a minimum of 3 days with behavioural observations subsequently performed on 10-day old larvae, as the authors claim. Moreover, predator-induced mortality, growth rates and stomach analyses could not have been assessed on 2-week old larvae, and the number of successful hatching events could not have been monitored over a 3-week

period (as stated in the paper). Following the reported exposure periods of 13 days (3+10) to 3 weeks, the experiments themselves would have taken at least an additional 6-7 days, given that the predator-induced mortality experiment alone must have required 4-5 days (10 tanks were allegedly used simultaneously, as stated in point #12 below; the reported N in the paper was 45-47, and each set of replicates lasted 24 h). In addition to this experiment, and the apparent pilot-study using both perch and flounder, Lönnstedt simultaneously had to perform behavioural trials, water changes, and general maintenance of the exposure setup. **All available evidence indicates that it is impossible for all of these tasks to have been performed by one person over the timelines specified by the authors**.

We did not ask for temperature data from the Baltic Sea. We know that these data were provided by Sundin and we indeed have access to them. We asked for the water chemistry data from the daily measurements in the containers in which the microplastics experiment took place. You write in your response: "*We took measurements from the beakers*". However, Sundin and Jutfelt never saw Lönnstedt performing any measurements of temperature or salinity in the beakers despite sharing the experimental room with her at the Ar research station. Further, **none of the other researchers at the station were asked by Lönnstedt to monitor temperature and salinity in the microplastics experiment while she was away from the station. If Lönnstedt took measurements from the beakers as she claims, why did she request these data from Sundin nine months after the experiments concluded** (point 12 in "Timeline Lönnstedt and Eklöv.pdf")?

We are also confused that you state in your reply that you took "*measurements from the beakers*", when you have previously stated that the beakers depicted in several photos (provided to Uppsala University) do not represent the experiment described in the *Science* paper.

You wrote "*Unfortunately we can't provide you with this data as Oonas computer was stolen along with her backups*", but this does not answer why the data were not included in the supplementary material in the first place, rather than irrelevant measurements from the Baltic Sea and Lake Bäste (see further below). An incomplete version of the raw behavioural data was uploaded to *Science*'s website after repeated requests; where did these data come from if Lönnstedt's computer and all backups had disappeared?

You write: "*there was no temperature regulation*" in the room where the experiments were conducted. Why is this information omitted from the *Science* paper and associated supplementary material? This is particularly important because you state in the supplementary material that "*water quality measurements […] were monitored daily and followed natural conditions*" but you fail to present these data and instead state reference values from Lake Bästeträsk (mean ± SD April: 8.8 ± 0.47 ˚C, May: 10.9 ± 0.93 ˚C) and the Baltic Sea (mean ± SD April: 6.4 ± 0.85 ˚C, May: 10.8 ± 1.5 ˚C). **It is impossible that the closed-system, air-exposed beakers used in the microplastics experiment followed the natural temperature conditions of these water bodies.** Rather, an experiment without continuous water flow and no temperature regulation would quickly reach the temperature of the experimental room (air temperature), which in this case was ~15-20˚C.

10) The supplementary methods state that "Behavioral observations were conducted on 10-day old larvae in glass aquaria (70 x 35 x 30 mm)."; that "Each test aquarium […] was supplied with aerated estuarine water (Fig. S1)."; and that "prior to experimentation, three fish were placed in each experimental aquarium, fed ad libitum and left to acclimate for 1 hr."

Given the size of the aquaria, we estimate that a total amount of ~30 mL water could fit (assuming a 3 cm water height; information not provided in the paper). Given the small amount of water and the 1 hr acclimation period, during which time the fish were fed, oxygen levels would have dropped and the water would have equilibrated with the room temperature. Differences in temperature and hypoxia are well known to have a strong negative influence on fish activity and behaviour. Were any measurements taken to assess how these conditions changed in the aquaria during trials? If the aquaria contained an airstone to oxygenate the water (this information is not provided), how was this done without disturbing the fish in the small aquaria and not obstructing the experimenter's view due to bubbles on the surface of the water?

10) Response from PE and OL: This is another unfortunate misunderstanding, partly due to our not writing it out properly. First of all, behaviours were tested on fish that were 10 days or younger and since there were airlines present during the acclimation period (which was up to 1 hour) the tanks oxygen levels would not have dropped that much. The airlines were simply a tube with a hole in it running along the top of all test tanks, it ensured a very small but steady amount of 02 seeped into the aquaria. This was of course carefully removed during the observations as bubbles tend to interfere with visually viewing animals especially in such small places (we did not draw this in the picture of the tank set up as we had already mentioned that aerated water was supplied in the text).

10) Our response: In your response to Uppsala University's preliminary investigation (sent to us by Per Andersson), you state: *"The scientific method we have chosen to use is described very clearly in the SM to the Science and anybody is welcome to replicate the study*" Yet, here, you acknowledge that you '*have not written [your methods] out properly*'.

Your response fails to address our critical point about the water in the experimental tanks warming to room temperature throughout the trial; the experiments were not performed under the thermal conditions given in the supplementary material. This is critically important because the health and survival of larval fish are very temperature sensitive.

If your response outlines the actual procedures you employed, why have you not published an eLetter or an addendum with this information alongside the *Science* paper? For example, in the paper and the supplementary material, you state that behavioural observations were conducted on 10-day old larvae; however, here you indicate that behaviour was assessed on fish that were 10 days or younger. Similarly, the paper states that the acclimation time was one hour, not 'up to one hour'.

Missing information:

11) The microplastics used in the study were obtained from a manufacturer (Polysciences, Inc.) and therefore did not contain pollutants accumulated in the wild. In the paper, you state that "This [the results] suggests that polystyrene particles may be chemically affecting larvae in both average and high concentrations, as exposure potentially reduces hatching rates of fertilized *P. fluviatilis* eggs." However, nowhere in the paper or the supplementary methods do you discuss the specific chemicals in polystyrene that may leach into the water and/or enter the body of fishes. Could you please provide this information?

11) Response from PE and OL: The goal of our paper was not to discuss the chemical composition of polystyrene. If reviewers and/or Science editors had wanted this information

during the revision stage we would have been happy to provide it. Now, if you're interested in specific chemicals associated with polystyrene we would suggest you read up on Tommy Cedervalls research as he works on toxic effects of nanoplastic polystyrene particles.

11) Our response: The *Science* paper states *"To assess direct chemical effects of polystyrene microplastics on fish, we collected fertilized egg strands of P. fluviatilis from natural populations in the Baltic Sea and placed them in 1000-ml glass aquaria that contained one of the three microplastic concentrations and filtered estuarine water"*. The paper then states *"Overall, successful hatching rates of fish were significantly related to microplastic concentration [analysis of variance (ANOVA): $F_{2,12} = 19.4$, $P = 0.0002$]"*. Therefore, **understanding chemical leaching from the microplastics appears to have been a central objective of the study**, which is why we expected the authors would have an idea of the factors that were responsible for their reported findings.

We refer you to **similar concerns expressed by other researchers**:
http://www.rheothing.com/2016/06/throwing-baby-fish-out-with-ocean-water.html
http://science.sciencemag.org/content/352/6290/1213.e-letters

12) How many pike were used in the predation trials? The paper states that juvenile pike were 31 ± 1.5 mm total length but the sample size is not provided. The text also does not specify what this measure of variation in size represents (is 1.5 a standard deviation or a standard error?). Were individual pike used more than once? Irrespective of whether pike were re-used, Pike ID should have been included as a random factor in the model since each predation trial included several perch and to control for differences in pike motivation.

Overall, very little information is given on how this experiment was performed. How was the level of satiation in the pike controlled for if the same individual was used for several trials in a row? Did you run several trials simultaneously (if so, how many were run simultaneously)? There is no information provided on the collection or housing of pike before the trials commenced. Were pike kept in aquaria for a similar time period as the perch? The glass mesocosms measured 300mm x 200mm x 150 mm (9,000 mL): what was the water level in these aquaria and how was the vegetation cover (i.e. refugia) standardized across mesocosms if several were used in parallel? This information is essential to so the reader can understand your experimental design and to allow others to replicate your study.

12) Response from PE and OL: The length of pike is in mm TL± SE. We are unable to provide you with Pike IDs as the first authors computer has been stolen. Our reviewers and science editors thought the information provided was enough to reproduce the experiment and we are sorry to see that you disagree. Pike were collected during the spring and kept in flow through Baste Lake water. They were fed ad lib with Artemia from the large hatching room at the research station. What we can say about this part of the experiment from memory is that several (-10) predation tanks were set up simultaneously in the different areas utilized by Oona. Sand covered the bottom of the floor and there was a similar amount of vegetation, small rocks and shelters in all the tanks used (e.g. mesocosms were set up as natural environments).

12) Our response: You have not answered our question. How many pike were used in the predation experiment? How was each pike's level of satiation controlled for? Given the number of available pike (~5; see below) and the reported sample size for the predation experiments, individual pike must have been used more than once. This has critical

implications for how your data should be analysed (e.g., pike ID and satiation should have been controlled for, and the alleged theft of the first author's computer cannot be blamed for this shortcoming given that the data were analysed many months earlier).

Your response seems to imply that at least 10 pike were used in the experiment. However, **this is impossible since Sundin herself collected the pike from the wild and therefore knows the sample size better than anyone**: 19 pike were collected, approximately 14 were used by Sundin for her own (unrelated) experiments, which left approximately 5 pike for Lönnstedt to use in her experiment. Additionally, **neither Jutfelt's photos, nor those taken by the *Universen* photographer Mikael Wallerstedt, show the setup that you describe above**. None of the researchers who were at the Ar research station in May 2015 have any recollection of such an experimental setup. The same is true for the "other" microplastics exposure setup that the authors claim is different from the one depicted in all the available photographic evidence. The paper also states that predation trials were performed 14 days after eggs had hatched; this is also impossible, as Lönnstedt had already left the research station at this point (and no one was asked to do such experiments for her).

You write *"Our reviewers and science editors thought the information provided was enough to reproduce the experiment and we are sorry to see that you disagree"*. We refer you to Smith (2006) and references therein, where the following is stated: *"At the BMJ we did several studies where we inserted major errors into papers that we then sent to many reviewers[3,4]. Nobody ever spotted all of the errors. Some reviewers did not spot any, and most reviewers spotted only about a quarter. Peer review sometimes picks up fraud by chance, but generally it is not a reliable method for detecting fraud because it works on trust"*.

Smith, R. 2006. Peer review: a flawed process at the heart of science and journals. Journal of the Royal Society of Medicine, 99: 178-182.

13) "After 2 weeks 20 randomly chosen *P. fluviatilis* were removed from the treatments and photographed against a scale for length measurements". Does this mean 20 individuals in total, divided by three treatment groups, so 6 to 7 individuals per group? How were these fish selected among the five replicates per treatment? Again, this is essential for the reader to understand what was done and to allow replication of the study.

13) Response from PE and OL: Let us explain. After ~2 weeks, 20 perch from each treatment were removed for measurements. This means that 20 individuals from each of the three treatments were selected randomly from the different tanks used in the tretaments.

13) Our response: This information is incorrect since Lönnstedt only spent 10 days at the Ar research station during the microplastics experiment. Lönnstedt borrowed a camera from Mário Cunha for these photographs, which were taken around May 12, 2015. This was approximately one week after the microplastics exposure had commenced, not two weeks as stated in the paper, supplementary material, and in your response to our queries. Reiterating our comment above, the document you sent to *Science* contains 35 photos of larval fish, so where are the other 25 photos that you claim were taken?

14) We know from experience that animal ethics approval to conduct predation trials with live fish can be very difficult to obtain. The numbers of the approved animal ethics protocols are not given in the paper but acknowledgements state that "All work reported herein was conducted in accordance with the guidelines for the care and use of animals in research of the

Swedish Board of Agriculture with approval by the Uppsala University Ethics Committee." Can you please provide the numbers for these approved protocols?

14) Response from PE and OL: Josefin Sundin repeatedly contacted our ethics department during 2015 and 2016 claiming we had no permit to do the research. She was then specifically asked to contact Peter or Oona from several people to receive a copy of the ethics. But this was not done.

14) Our response: You have not answered our question: we asked for the number of your animal ethics approval, which is not provided in the *Science* paper.

In similarity to your response to Uppsala University's preliminary investigation and in your comments to *Retraction Watch*, you appear to be accusing Sundin of 'being out to get you'. In your reply to the preliminary investigation (sent to us by Per Andersson), you describe Sundin's behaviour as "*deeply unethical*", "*highly distasteful*" and as having "*nothing to do with research criticisms*". Very much to the contrary, Sundin and Jutfelt are concerned with the major disparity between their personal account of your study and its description in the *Science* paper. These discrepancies were uncovered only after the *Science* paper was published. Understandably, Sundin feels an ethical and moral responsibility to report these discrepancies to the scientific community.

Here, you claim that Sundin repeatedly contacted "*our ethics department*" but provide no evidence of this and no contact details of the 'ethics department' or name(s) of the person(s) she would have contacted to verify this statement.

The reason that we (the seven signatories of this letter) asked for the ethical permit number is because **the permit was issued after Lönnstedt performed the microplastics experiment at the Ar research station.** This information can be verified by the date of issue on the ethical permit by veterinarian Anne Halldén Waldemarson (anne.waldemarson@neuro.uu.se) at Uppsala University.

Statistical analyses:

Response from PE and OL: We have already gone through Science rigorous peer review (that is several reviewers, statisticians and Science editors have gone through our statistics), and we are not going to go through another statistical analysis with your group. We appreciate your thoughts on the statistical design but are unable to incorporate it into an already published article.

Our response: There is ample evidence in the published literature that peer-review and publication in scientific journals do not preclude articles from containing errors, fraudulent data and inadequate statistical analyses (discussed above). For example, *Science* published your article without ensuring that your data were included in the supplementary material and/or that an accession number to the data in an online repository was provided in the acknowledgements (http://www.sciencemag.org/authors/science-editorial-policies). It is not inconceivable, therefore, that your paper would contain other errors despite having been assessed by reviewers and editors.

***Science*'s policy states:** "Authors should present results in complete and transparent fashion so that stated conclusions are backed by appropriate statistical evaluation and limitations of the study are frankly discussed." As such, **we would appreciate responses to our queries**.

15) Nested 2-factor analyses (MANOVAs and ANOVAs) were used to test for treatment and tank effects on behavioural measures. Tank effects should in fact be controlled for by including tank as a random factor nested in treatment following a mixed-effects modelling approach. It is incorrect to treat tank as a fixed factor because it artificially inflates degrees of freedom in the model (and therefore considerably increases the risk of Type I error). The methods in the supplementary material state: "Since our statistical analysis showed that there was no significant effect of exposure tank on the behaviour of individual fish ($P = 0.49$), we can with confidence say that there was no individual variation within the exposure aquaria that affected the behavioural results (e.g. a potential tank effect). Thus we treat each individual fish as an independent sample, which is often the case in behavioural studies (12, 22, 23, 33-35)." We are confused by this statement. Does that mean that one-way ANOVAs were carried out in the end? Tank should certainly not be removed from the analyses (otherwise resulting in pseudoreplication) and should be treated as a random, not a fixed factor. It does not appear that tank was included as a factor in the analysis of egg hatching success. Please clarify why the analyses were carried out as such.

15) Response from PE and OL: This statement was only added to appease a reviewer. Nested 2-factor MANOVAs and ANOVAs were employed.

15) Our response: You have not answered our question. We cannot check this ourselves since tank number (in addition to many other pieces of information) is still missing from the partial dataset uploaded in the new version of the supplementary material.

16) A total of ~ 35 tests (although this is hard to tell from the methods) were carried out between all the MANOVAs, ANOVAs and post-hoc tests performed. This is a considerable number of tests which greatly increases the probability of Type I errors. Why were MANOVAs carried out if each response variable was examined independently afterwards? The supplementary material states that: "(i) total distance moved (mm); (ii) activity level […]; and (iii) total time inactive […] were compared among treatments using a MANOVA approach as the three response variables were not independent from each other". This appears to be in contradiction with the distinct ANOVA tests that are presented in the paper. Could you please clarify the logic behind the statistical approach used? Given the large number of tests that were conducted, why was the alpha value for the study not corrected? For example, a Šidák correction would be suitable (as opposed to a more conservative Bonferroni correction).

16) Response from PE and OL: We start with MANOVAs to account for any effect of the behavioural measurements being non independent from one another. This is commonly done in these types of behavioural studies.

16) Our response: Several of us are behavioural ecologists by training and publish in journals such as *Animal Behaviour*, *Behavioural Ecology*, *Behavioural Ecology and Sociobiology* and *Ethology*. We can attest that discounting exceedingly high risks of Type I errors is **not** "*commonly done in these types of behavioural studies*". *Science*'s editorial policy states that: "Adjustments made to alpha levels (e.g., Bonferroni correction) or other

procedure used to account for multiple testing (e.g., false discovery rate control) should be reported." See: http://www.sciencemag.org/authors/science-editorial-policies.

17) Assumptions of normality and homoscedasticity were apparently checked for all tests (but see the next comment). What about the effect of (univariate) outliers, which are typically examined using leverage plots? Outliers can have important effects on the robustness of model outputs. In addition to verifying the assumptions of univariate tests, were the assumptions of MANOVA verified? There is no mention of such diagnostics in the paper.

17) Response from PE and OL: We found that assumptions in MANOVA and ANOVA were met using normality and homoscedacity.

17) Our response: As stated above, normality and homoscedasticity are not the only assumptions of these tests. **Our question was: were the other assumptions of these and other tests verified, and if so, how?** For example, were the assumptions of the Cox's proportional hazard model examined? The supplementary material does not state whether these assumptions were tested and met. *Science*'s editorial policy states: "Sufficient information should be supplied to allow readers to judge whether any assumptions necessary for the validity of statistical approaches (e.g., data are normally distributed, survival data are consistent with proportional hazards in a Cox regression model) have been verified." See: http://www.sciencemag.org/authors/science-editorial-policies.

18) Statistical analyses were carried out to test for differences among groups in several variables that are inherently non-normally distributed because they are bound by 0 and 1 (e.g. the proportion of successful hatching events, the proportional change in activity, area use, and freezing behaviour among treatment groups). Because they are non-normally distributed, proportions have traditionally been arcsine sqrt transformed before analysis, although the recommendation is now to use a logit transformation (Warton and Hui 2011). We were very surprised to read that "assumptions of homogeneity of variance and normality were examined with residual analysis" and that only "time inactive did not follow parametric assumptions and was log10(x+1) transformed." We would expect many more of the models to violate assumptions of normality and homoscedasticity given the nature of the response variables. Can you confirm that this was not the case? Why was a log10(x+1) transformation used instead of an arcsine sqrt or logit transformation in the case of time inactive?

18) Response from PE and OL: We found that only time inactive required a transformation and this was corrected using a log 1O(x+ 1) transformation.

18) Our response: You have not answered our question. Several of the variables you examined are inherently non-normally distributed. Therefore, why were they not dealt with accordingly?

Unfortunately, your assertion that "*only time inactive required a transformation*" cannot be verified because 1) you did not adhere to *Science*'s data archiving policies when your paper was published, 2) Lönnstedt's computer and all backups containing the data were allegedly stolen after the paper was published, 3) you do not state what statistical software was used for the analyses in the paper (other than for the survival analysis), and 4) residuals of your models cannot be examined because you have not archived the script to reproduce your analyses as per *Science*'s editorial policy: "All computer codes involved in the creation or analysis of data must also be available to any reader of Science." See:

http://www.sciencemag.org/authors/science-editorial-policies.

19) The survival analysis must take into account the autocorrelation (i.e. non-independence) of data points resulting from the fact that 4-6 larvae were tested in the same trial. This requires a mixed-effects modelling approach, which can be implemented for survival analysis. Why was this not done? In the current analysis, the observations are treated as independent, which is not the case. Were the assumptions of the Cox's proportional hazards model verified? If so, how was this done? Were the assumptions met?

19) Response from PE and OL: This was, again, approved by all our reviewers and Science statisticians and therefore we decided to use this analysis.

19) Our response: The fact that these analyses were approved by reviewers does not mean that they are correct. As stated above (point #12 and under the heading "statistical analyses"), reviewers and editors commonly miss problems with statistical analyses because they may not be familiar with specific tests or methods.

While your responses to our statistical questions are unsatisfactory, **the statistical issues we raised are only secondary to the much more serious concerns of scientific misconduct.**

Data availability:

20) In accordance with Science's data policy (http://www.sciencemag.org/authors/science-editorial-policies), you state in the acknowledgements that: "The data reported in this paper are archived at the research database at Uppsala University and are also included in the online supplementary materials." However, no such data are available in the supplementary materials, which contain only supplementary methods, figures and references. We were also unable to find Uppsala University's research database online. Can you please provide the link to this database? To our knowledge, the only website supported by Uppsala University where data could potentially have been deposited is DiVA. However, DiVA is a server for publications by researchers affiliated with Uppsala University, not a data repository. The only file associated with Lönnstedt and Eklöv 2016 is DATASET01.pdf: http://uu.diva-portal.org/smash/get/diva2:922441/DATASET01.pdf. This pdf file contains three bar plots with the caption "Data S2. Raw data for the behavioral change in response to predators." As we are sure you are aware, bar plots are not raw data. We would be grateful if you could archive the raw data used to produce the figures and statistical results presented in the paper. Suitable data repositories you may wish to use include Dryad (http://datadryad.org/), figshare (https://figshare.com/) or Zenodo (http://zenodo.org/).

20) Response from PE and OL: All the raw data had been tabulated as a Supplementary file (as per other recent Science studies). There seems to have been a misunderstanding here as Science approved this at first and it was only very recently that our editor told us that they also needed raw data files. Before this mix-up was noticed Oona's computer bag containing her computer, backups, keys etc was stolen under very unfortunate circumstances (and the department server had not backed up the files), so we have had to go through old emails to locate the raw data. Luckily we could salvage most of the files and these have now been uploaded to Science.

20) Our response: Dr Graham Raby notified you on June 3, 2016, that there were no data in the supplementary material on the *Science* website. **Raby's request to obtain these data was**

ignored by Lönnstedt until Dr Sacha Vignieri (senior editor at *Science*) emailed you on June 13, 2016, requesting the data**. Within 24 h, Lönnstedt's computer and all backups were allegedly stolen.** Thus**, the requests for raw data came before the apparent theft** of the computer and all backups. **The file that was eventually uploaded to *Science*'s website still lacks most of the data** necessary to reproduce the results presented in the *Science* paper (see Appendix 29 to the document "Timeline Lonnstedt and Eklov.pdf", included on p. 20 below).

As a result, **your paper does not comply with *Science*'s editorial policy**, which states that: "After publication, all data and materials necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of Science."
See: http://www.sciencemag.org/authors/science-editorial-policies.

References listed in the in original queries document

Cuvier-Péres A and Kestemont P (2001) Development of some digestive enzymes in Eurasian perch larvae *Perca fluviatilis*. Fish Physiology and Biochemistry 24: 279-285.

Vlavonou RS, Masson G and Moreteau JC (1999) Growth of *Perca fluviatilis* larvae fed with *Artemia* spp. nauplii and the effects of initial starvation. Journal of Applied Ichthyology 15: 29-33.

Warton DI and Hui FKC (2011) The arcsine is asinine: the analysis of proportions in ecology. Ecology 92: 3-10.

N.B.: The data listed in Appendix 29 are still missing as of November 1 2016.

Appendix 29

List of data still missing after a partial dataset had been uploaded as supplementary material, 5 weeks after the data had been requested (requested on June 3, uploaded on July 7). The full dataset needed to reproduce the analyses in the paper is still missing. Data currently missing as of July 18 2016 include:

- **The raw data for the analysis of fish responses to olfactory cues**
  No data are provided for this analysis. Missing data include: distance moved, number of lines crossed, time inactive, exposure treatment, exposure tank, type of cue injected, date and time.

- **The raw data for the analysis of plastic particles ingested by fish**
  No data are provided for this analysis. Missing data include: count data of the number of ingested plastic particles per animal, exposure treatment, exposure tank, date and time.

- **The raw data for the analysis of successful hatching events**
  The table provided in the supplementary material "Raw data S1" contains the proportion of eggs that hatched per exposure tank. However, the table should contain count data of the number of successful hatching events per day in each treatment

tank, over the three-week period of the experiment. The paper states that "Hatching events (defined as the number of larvae that successfully hatched out of eggs over a 3-week period) were monitored daily".

- **Exposure tank, date, and time for the analysis of distance moved by fish**
  The table in the supplementary material "Raw data S2" requires additional information specifying each animal's exposure tank, and the date and time of each observation.

- **Predator ID, trial number, exposure tank, date and time for the analysis of predator induced mortality**
  The data provided in the supplementary material "Raw data S3" require additional information specifying the predator ID, trial number (this information is critical because several animals were tested in the same tank; i.e. observations are non-independent), the exposure tank of each test animal, and the date and time of each observation.

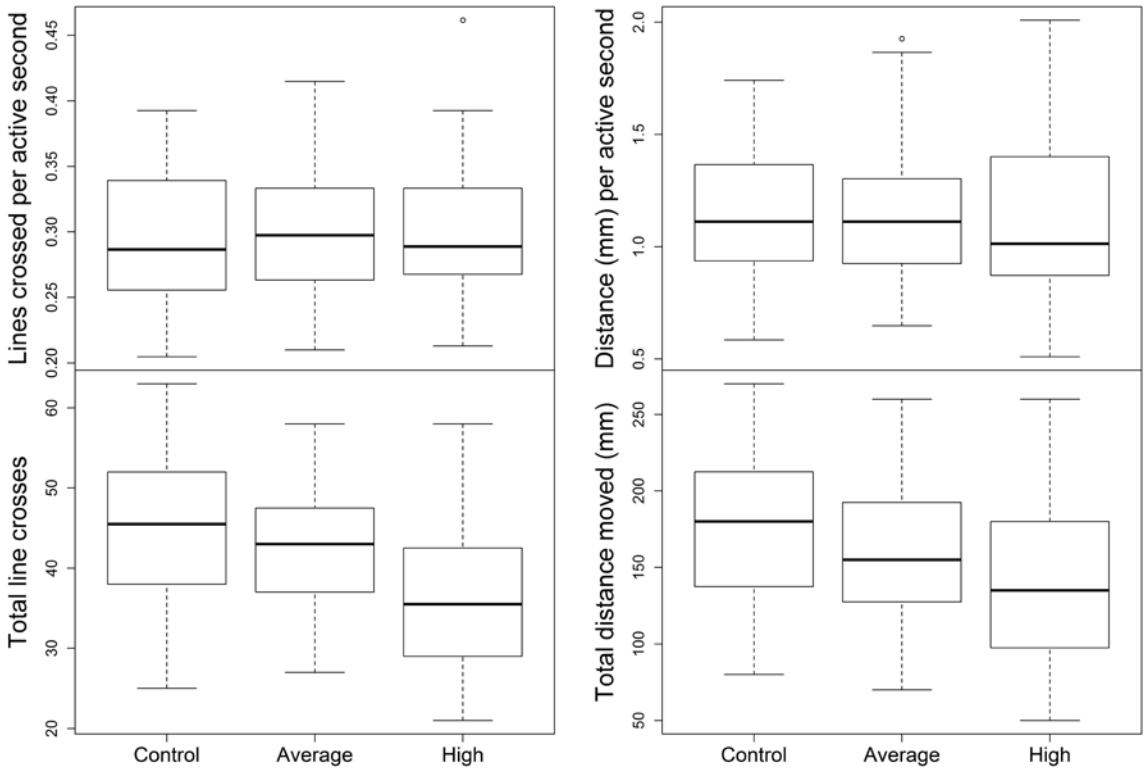
- **Date and exposure tank for the analysis of fish size**
  The data provided in the supplementary material "Raw data S4" " require additional information specifying the exposure tank of each test animal, and the date at which length measurements were taken.

- **Figure captions for the photos provided in the updated supplementary material**
  What treatment and exposure tank did each fish come from? When were the photos taken?

Discrepancies in the metadata associated with the "Raw data" tables:

- The legend for the table "Raw data S2" indicates "Behaviour of 5-10 day old *Perca fluviatilis*". The paper states that these experiments were performed on ten-day old *Perca fluviatilis*.
- The legend for the table "Raw data S3" indicates "Predator-induced mortality of 1-2 week old *Perca fluviatilis*". The paper states that these experiments were performed on two-week old *Perca fluviatilis*.
- One of the column headers in table "Raw data S3" is "Time to Death (min)". The paper states that time to death was measured in hours.

Box-plots related to query #8. Number of lines crossed and distance moved of perch larvae after being adjusted (top panels) or not adjusted (bottom panels) for the time spent inactive. Adjusting for the time fish spent inactive in the experiment eliminates any treatment effect of microplastics on the number of lines crossed and distance moved.



Total number of lines crossed (i.e. unadjusted for time inactive)
ANOVA: treatment **P=0.001**

Number of lines crossed s^{-1} (i.e. adjusted for time inactive)
ANOVA: treatment **P=0.896**

Total distance moved (mm) (i.e. unadjusted for time inactive)
ANOVA: treatment **P=0.007**

Distance moved (mm) s^{-1} (i.e. adjusted for time inactive)
ANOVA: treatment **P=0.791**