Interactive comment on “Changing mineralogical properties of shells may help minimize the impact of hypoxia-induced metabolic depression on calcification” by Jonathan Y. S. Leung and Napo K. M. Cheung

Jonathan Y. S. Leung and Napo K. M. Cheung
jonathan_0919@hotmail.com

Received and published: 2 June 2017

COMMENT 1: I have read the manuscript by Leung and Cheung on the ability to form calcium carbonate of a tubeworm species under anoxic experimental conditions. After careful assessment, I have to advise to reject this manuscript from publication in Biogeosciences. The experimental design is flawed with only two times two treatments (normoxia versus hypoxia and stress versus no-stress) to allow for a sound analysis of the effect of O2 and stress on tube-formation.
RESPONSE 1: We completely disagree that 2 x 2 factorial design is an experimental flaw. To evaluate the effects of hypoxia and shell damage, two groups for each factor are statistically enough. Thus, our full factorial design can provide a sound interpretation for the effects of hypoxia, shell damage and their interaction. In fact, 2 x 2 factorial design is extremely common in ecological studies (e.g. Connell and Russell, 2010; Mukherjee et al., 2013; Ghedini et al., 2015).

COMMENT 2: It is unclear how many polychaetes were incubated for each treatment, how many survived the experimental period, and exactly how many measurements were done on the formed tubes/ how many analyses were done to determine the respiration, clearance rate, etc.

RESPONSE 2: For incubation, we have 10 individuals per bottle (Ln 89) and 3 replicate bottles per DO level per context (Ln 92). For shell hardness and elasticity, we have 5 fragments from 5 individuals per DO level per context (Ln 100-101). We have 3 replicates for calcite to aragonite ratio (Ln 108-109), magnesium to calcium ratio (Ln 117) and amorphous calcium carbonate (Ln 124-125). We have 5 replicate syringes per DO level per context for respiration rate (Ln 129-130) and 5 replicate bottles per DO level per context for clearance rate (Ln 136). The number of replicates for each parameter has been clearly written. We can add survival rate as supplementary information in the revision as it is not very relevant to the hypothesis of this study.

COMMENT 3: It is a pity that tube length was only assessed once (at the end of the 3-week incubation period) and not throughout the experiment. This would have allowed to estimate whether the worms did not produce their tubes only in e.g. the first days of the experimental period and whether, e.g. absence of an essential nutrient may have hampered growth by these organisms in this experimental setup.

RESPONSE 3: We measured tube length three times (before, in the mid of and after the exposure), but the overall shell growth rate has the greatest interpretive value and should be presented. This study does not aim to examine whether the individuals pro-
duce tubes on the first few days or not, which has limited scientific value. Nutrition is not a problem as the food, which was provided daily (Ln 92-93), is optimal for supporting tube growth (Leung and Cheung, 2017).

COMMENT 4: There is no real control group followed to see whether the handling of the specimens in the microcentrifuge tubes affected the worms’ functioning.

RESPONSE 4: We are not interested in studying the effect of containers on physiological functioning. Microcentrifuge tubes are used to allow the individuals to grow in the upright position and avoid individual interaction (Leung and Cheung, 2017).

COMMENT 5: It is unclear how long the incubation were to determine the oxygen uptake (respiration) and whether or not the (under normoxic conditions) O2 levels already decreased towards 0 within the first period of the incubation.

RESPONSE 5: The incubation time was 1 hour (Ln 131). The speculation that “whether or not the (under normoxic conditions) O2 levels already decreased towards 0 within the first period of the incubation” is ungrounded. The animals were incubated inside a sealed syringe for a single time only; therefore, there is no first or second period of incubation. From normoxia to almost zero DO concentration in 1 hour, the respiration rate needs to be \( \sim 42 \, \mu g\, O_2\, ind^{-1}\, hr^{-1} \), which is much higher than what we reported (\( \sim 26 \, \mu g\, O_2\, ind^{-1}\, hr^{-1} \)).

COMMENT 6: Table 1 lists results with averages and their standard error instead the standard deviation. This would only make sense if very many measurements were performed, but that is impossible to tell from this manuscript.

RESPONSE 6: The choice of standard error (of mean) over standard deviation is required by the statistical context. The core message of Table 1 is how hypoxia and shell damage affect physiological performance and shell geochemical properties. This can only be shown by inferential statistics that highlight and quantify the difference between treatment and control groups, i.e. estimated central tendency and the error of the es-
The use of plain descriptive statistics, such as SD, is not justifiable in this case. If the readers (out of curiosity or whatever reasons) would like to know how the individuals from the same treatment group differ from each other, SD can be easily calculated by multiplying the provided SE by the square root of sample size (see RESPONSE 2). In addition, the statement “This (SE) would only make sense if very many measurements were performed” is logically and mathematically incorrect. Given that SE = SD/sqrt(N), where N is the sample size, SE will approach zero if N is a very large number. On top of that, one has to be very careful with inferential statistics that come from a very large sample size as tiny difference between the estimates of central tendency of treatment groups will manifest as “significant difference” in statistical sense, but not necessarily in biological sense. Therefore, showing SE can be misleading if N is very large.

COMMENT 7: It is unclear how measuring DO under hypoxic conditions with this experimental design can result in accurate respiration rates and how they can be compared to rates under normoxia.

RESPONSE 7: This experimental design is widely used for comparison across DO treatments (e.g. Zhao et al., 2011; Leung et al., 2013).

COMMENT 8: The discussion contains many over-interpretations and conclusions have little relation to the results.

RESPONSE 8: This comment is ambiguous. The discussion and conclusion are based on our results and hypotheses.

OVERALL RESPONSE: While we are pleased to accept comments, it is also our responsibility to validate them. In this case, many details are substantially overlooked by this reviewer (e.g. COMMENTS 2 & 5) and some comments are subjective without clear reasons (e.g. COMMENTS 1, 7 & 8). If rejection is based on the careless assessment, the quality and objectivity of this review become questionable.
References


