Dear editor and reviewers,

I would like to thank you for all your remarks. You made me think more deeply about the article and also about the way how to present it to the audience. I am very grateful for that and I hope, that my answers to your questions and comments will satisfy you.

**Reviewer: Aleksandra Kilian**

**Basic reporting**

Please, add more sufficient background about previous studies in which head width was used to determine larval instars.

* *Reply: I addressed this issue in this version of article.*

Therer are many studies where the haed width was measured and difference of head width among larval instars are commonly used. So, it is not a new character for larval determination,I suggest to underline a practical aspect of this character in forensic entomology as well as the importance of other morphological features in determining of larval instars in legal investigation.

* *Reply: Yes, there is a number of studies, often descriptions of larvae or larval instars, but these descriptions are based on few specimens. They did not cover the variation in size based characters and that is why I am referring only to those studies, which were focused on revealing this variation. I tried to explain it better in the new version and focus more on practical aspects as you suggested, so I hope that you will be satisfied with that.*

**Experimental design**

See comments on the attached annotated manuscript

**Validity of the findings**

See comments on the attached annotated manuscript

**Comments for the author**

See comments on the attached annotated manuscript

**Comments inside of the manuscript**

Is it one model or more

*Reply: It is one model for each stage.*

There are many previous studies on larval morphology with head width measurements and it is clear that one of good characters for determining larval instars is head width ,especially in groups with poor chaetotaxy. It would be good to underline this well-known rule and support references. And it would be good better underline that for this species, there are other good characters, especially brown spot on head and chaetotoaxy**.**

*Reply: I was referring to your article, but I explicitly added this information into the new version.*

Table 2 show that maximum head width L1 is bigger than minimum HWL2, similarly, maximum HWL2 is bigger than min HWL3 .the reason why should be better explained by author because this results show that head width is NOT good or enough for determining larval instar in this species. Mean value is , in determinig specimens of larva, practically useless.

*Reply: This is why I strongly suggest in this article to use several independent characters for instar determination and not relying on few dependent characteristics like size of head, pronotum or body. All these characters are highly correlated and therefore they do not provide stronger evidence. Using variety of independent characters is what I suggested. Overlap in the size between instar is natural and we expected to see some extreme values, but the main take home message was that these extremes are rare and the majority of same age larvae were of similar size (see standard deviation for each instar at Table 2).*

I think author should more underline importance of Thermal model than head width. What’s more, author does not provide detail information in which points, exactly, head width have been measured. It should be given in methods (the widest point, below antenna ?).

*Reply: I would like to report both phenomena, because they are important for practical use of this species. For the exact point of measurement see Fig 2.*

It is valuable that author provided a large data on head width, which enable better understanding changes among instars although this character is probably overrated.

*Reply: In the case of* S. watsoni *I would argue that number of characters for the instar determination is very limited and every additional character is potentially valuable. Especially when**they are independent of others.*

"Therefore identification of this species in every stage of development is not an issue." I can’t agree with this statement. Description of larval stages has the same main purpose as description new species\_ to analble otherpeople to discriminate, determine. In most examples, onlu detailed observations make possible of determination of larval instars.

*Reply: I am speaking about identification of this particular species across all stages in contrast to others where only third instar is described and we have no DNA information at all. Purpose of this sentence is to show that there is enough information about this issue and there is no reason for me to address it in this article.*

" a new method for identifying larval stages based on combination of morphological features mentioned by (Kilian & Mądra, 2015) and the size based characters " I can’t find in result or discussion exact explanation, summary of this method and providing combination of size based and morphological characters.

*Reply: Summary of basic principles of methodology and combination of characters was added into discussion.*

"...we had plenty of characters to choose from." Plenty of data? Pictures? if characters, it is unclear what exactly, why author chose only bead width

*Reply: The main aim was to explain that we chose to use head width as our main character for instar determination because head is changing its size after each molting and it is also very heavily sclerotized structure so our measurements were much more consistent than if we would be using other, less sclerotized parts of body.*

**"**...omitted these parts, and also the body length, as good characters for instar determination.**"** It is not good reason for omitting characters.

*Reply: All these characters are correlated in each individual; therefore we were looking for character, which could give us more accurate and consistent results.*

**"**...camera had low resolution and those parts would be very challenging to measure accurately.**"** Low resolution of camera is also not reasonable and scientific explanation of omitting some characters. Try to explain it in more scientific way.

*Reply: Addressed!*

"...good enough to allow precise measurements of the head width. Those pictures covered all three larval instars (L1 = 591, L2 = 500 and L3 = 640 pictures). The bias in number of pictures between different stages was caused by difference in the duration of development of these instars..." See also comment below. Please, explain /resolve overlaping measurement of head width of instars and how you check when larva with head width 350 mm, is second or first. And how, that’s why, use this character in legal investigation.

*Reply: We recorded all larvae continuously so we were able to say exactly in which stage the larva was regardless of its head width.*

**"** on continual observation of individuals from egg until pupation **"** The idea is good but practically, is it possible to observe individuals when you have hundred larvae in one petri dish? Maybe this is a reason of overlapping of head width and should be repeated/or larvae with border head width should be examined better (chaetotaxy, other measurements under light microscope) and then decided to which instars they belong. Author write about in 259-263 verse but I do not if he used this methods to check his own larvae with border measurements.

*Reply: As I mentioned in the methodology. Larvae of* S. watsoni *were separated right after eclosion and bred individualy, therefore we had exact information about their stage and no additional measurement or checking was necessary.*

"Thus chaetotaxy can be used for the discrimination of the first and second instar larvae." Better is using brown spot of head - is well visible in second and third instars. Did author observe this character during his study?

*Reply: I did notice it, but I was unaware of its usefulness for instar determination. Thank you for pointing that out. I added it to the text.*

### Reviewer 1

#### Basic reporting

The writing in English was remarkably clear (although in some instances did not conform to usual practice).   
The background for how the work fits into a broader field of knowledge is not sufficient. The author "... proposes head width as a new character for larval instar determination ..." (line 18-20), but this is not a new character, unless the author is proposing it only for S. watsoni. Dyar's law was proposed in 1890 (H. G. Dyar, 1890. The number of molts of lepidopterous larvae. Psyche 4: 420-422), suggesting that head width increases by a ratio which is constant for a species. Recognition of instars by head width is commonly done by entomology graduate students, and published literature has revealed that Dyar's law is often useful, although exceptions to the constancy of the ratio are not uncommon. The author's argument that head width has advantages as a character is intelligent and valid, but not novel.

*Reply: I see that I did not phrase this correctly. My claim was not meant to be general and I am referring only to head width of* S. watsoni*. I corrected it in the updated text. I also agree that novelty of this finding was overstated. So, I rewrite it to better describe what was actually done and what the value of those findings is.*

An important finding is that chaetotaxy of the 1st instar is rudimentary compared to other instars. Is this a new finding? This should be presented clearly in the context of the field of knowledge of chaetotaxy of larvae of this species and group (the family or Coleoptera in general).

* *Reply: I explained it better in the updated version of the text, so it will hopefully not cause any confusion.*

The relatively low optimal temperature for development of this species is surprising to me, although it may be consistent with temperature in its geographic range. I would appreciate discussion of optima for species having similar ranges to S. watsoni.

* *Reply: To be clear. The term "optimum" refers to the range of temperatures where the rate of development is high, but mortality is not. In this regard is the proposed range of temperatures not so surprising especially because S. watsoni lives in leaf litter where the temperatures are not very high throughout the year. I discussed this with other species from the same subfamily (Cholevinae), with similar geographic range and similar habitat requirements. So, I am little bit surprised with your request, but to improve the quality of the discussion I included also Thanatophilus mutilatus.*

I am uncertain that the paper is a sufficient "unit of publication," because of the extremely high mortality (>80%) for 3rd instar larvae and for the pupae. Due to the challenges of the study, it may be rightfully judged that presentation of adequate data for a "validation-ready" model for forensics should be a separate study, but I suggest that at least some preliminary experimental data should be presented in two important areas, the possible cause of the mortality and a comparison of the development of isolated larvae with development under natural conditions.

* *Reply: I added information about development of* S. watsoni *in a group as a contrast to development of an isolated individual and elaborated more about the cause of mortality.*

Insight into a possible cause of the mortality is provided in lines 222-229, suggesting that adults could provide bacteria that increase "... the quality of the food source." This is both fascinating and disturbing, since the number of larval instars is increased by prolonged development due to poor nutrition in some coleopterans.

* *Reply: Number of larval instars is stabile for this species and we did not observed any case with more than three larval instars. This was also shown by Kilian & Mądra (2015) and is belived to be generaly thrue for the most of Staphylinoidea (see Beutel & Leshen (2005): Handbook of Zoology - Coleoptera (vol. 1)). The length of development in our breeding experiment also should not be affected because we observed similar times in our observation study, where the whole population thrived. Also Kilian & Mądra (2015) reported similar length of development of their specimens (see the updated text).*

At a minimum, I would think that the typical range of head capsule widths for mature larvae in the field (the cast larval head capsules are likely to be present near the pupae) should be reported.

* *Reply: Observations of larvae outside of laboratory are very rare and I did not have the pleasure of finding them in field (mature or other vice). Also I am not sure if the cast of larval head capsule could be used for identification, because there is possibility of misidentifying it with other two species of genus* Sciodrepoides*, which larvae are still un-described. Therefore your proposition would be nice, but is not doable right now.*

Another valuable addition could be preliminary investigation of pre-treatment of the food with short term feeding by adult males (used to avoid oviposition).

* *Mortality occurred during the third instar, which is last larval instar before pupation. According to literature and our observations, this instar is the most demanding and risky for larvae to finish. Therefore there is a higher level of mortality in this instar, but this is normal even under natural conditions. Weak individuals are not able to undergo the pupation. Larvae that failed to pupate could prolong their development, but we did not use the data from larvae with unusually long duration of instars from our models. This procedure should be sufficient to keep our data reliable despite the increased mortality.*

#### Experimental design

The research question is inadequately or incorrectly defined. It think that the research is the investigation of methods that may bring us closer to development of a model of development of this beetle for forensics. The present paper clearly does not present a model which is ready for forensic application, and this has to be made clear in the abstract and elsewhere.

*Reply: I would like to object that the model in general is a useful tool, but tool never the less. Further in my rebuttal I will explain how these models are used by field forensic entomologist and why they cannot be used without previous validation for local conditions. So,*

The investigation was conducted to a high technical standard and described with sufficient information to be reproducible, and is reported with high ethical standards. However, I suggest that the typical handling of larvae for photo documentation be described more fully.

* *Reply: Thank you for spotting this problem. It was corrected in the updated text.*

Since the usefulness of photo documentation for estimation of head width was a major finding of the paper, it would have been interesting to have a comparison showing how similar these estimates are to those provided by a micrometer scale in the eyepiece of a microscope to measure head width of larvae that have been immobilized by chilling or other methods.

*Reply: It would be interesting indeed, but measuring such small and living animals under binocular microscope is not possible. At least not with appropriate precision or without applying knock out methods, which are very invasive and could influence the rate of development.*

None of the experimental temperatures were low enough to prevent development, and therefore, the developmental threshold has not been established.

*Reply: Lower developmental threshold is a value that is estimated mathematically from developmental data and usually it is not necessary to prove it empirically. However I additionally tested the lower temperature (12°C), but breeding did not occurred, which supports my original finding.*

#### Validity of the findings

The conclusions about data are valid. In my opinion, however, it must be made clear that the model is unlikely to provide useful predictions of development in the field. I base my opinion on high mortality data and on my belief that it is very likely that development was abnormally slow. The methodology must be improved, if the model is going to have a reasonable chance of validation. Any predictive model requires validation before its predictions can be used with confidence, and in this case, I assume that validation would be comparison of predictions with data on presence of developmental stages at intervals after infestation of food under natural conditions.

*Reply: Remark about need for validation is just, but we did not attempt to validate our model for a reason. Thermal summation models and their parameters are only a measuring tool and as such they have to be validate for conditions where they will be used. This is due to a fact that Tmin and SET parameters differ between distant populations (see Kypiatkov and Lopatinka 2010, Entomological Review, vol.90, No.2, pp.163-184), therefore before using any species for purposes of legal investigation it is common practice to validate the model for each region/state.*