

This study is an interesting and valuable piece of research exploring the demographic and productivity consequences of heatwave-like conditions. The methods and analyses appear comprehensive and sound. I think the emphasis on the utility of microcolonies for experimental manipulations is valuable and well-justified here. I think you could highlight this even more in the discussion. There are some examples of using “colony fragments” in ant research that could be good support for this if you can track down any relevant references.

The conclusions are well-reasoned and the authors do a good job of avoiding over-reaching. A major piece that I think deserves a bit more acknowledgment is the fact that ambient heatwave temps do not necessarily equal in-nest temps. I would recommend highlighting this more as deserving future study—i.e., how much do *B. terrestris* nest temps vary and what is the variation with ambient temps? Perhaps there is already some info out there on this, though I suspect not much. Where possible, I think it would be good to include data on the extent to which *Bombus* more broadly can buffer against environmental temps, whether passively through nest architecture or actively through thermoregulatory behaviors. I think *Bombus* are an interesting test case for heatwave effects because the cavities they nest in can be so variable, and likely have variable thermal properties. Much of this is beyond the scope of your study, but I think if you could touch on some of these topics a bit more it would go a long way toward situating these findings within their ecological context.

Apart from that, my comments are minor and I think this study makes a valuable contribution to the literature. I hope these ideas will help improve the paper, which I believe is nearly ready for publication.

Raw data:

More descriptive metadata is needed to interpret the raw data. This info could go in the “description” section on Mendeley Data. For example, it would be helpful to list a description of each file and how to interpret the columns in that file. For example, in *Behaviour_Data.csv*, what do “Fanning1”, “Fanning2”, etc., refer to, and what is the one that’s just “Fanning”? I can probably guess these are different scan samples but including a brief description is a good practice for making this dataset usable in future research. Also, in “*Weight_data.csv*”, what are the units here? It would be helpful to go through each file and see what extra information would be necessary for interpretation.

Line Comments

Lines 83-84: I think it would be really valuable to cite/discuss studies describing effects of heatwaves on other non-pollinator organisms, especially plants and other animals.

Lines 88-89: I don’t agree that because bumble-bees are cold-adapted they are necessarily more vulnerable to current and future climatic changes. In fact, tropical species may be most vulnerable to climatic changes due to narrower tolerance breadths. I would rephrase or cut this sentence.

Lines 105-106: It seems important to mention that elevated temperatures can also accelerate development and therefore impact adult body size. Okay, I now see you get to this in the discussion. I think it could also merit a brief mention in the intro.

Lines 111-112: I'd recommend rephrasing as "Bumblebees may also be exposed to heat stress during adulthood."

Line 217: Is drone production measured as the total number of drones at the end of the four weeks? Or some other way? Is drone mortality within the 4 weeks accounted for, or not? Please clarify.

Lines 267-70: Please also include the stats on the comparison between the control and the extreme treatment heatwave.

Lines 311-13: I think this phrasing is a little confusing/misleading. I would recommend just stating that there were not significant differences in brood incubation.

Lines 316-342: This first paragraph is really lengthy/detailed. Instead, can you just summarize the results from your study and their broad significance, and then break off the comparisons to other studies into a second paragraph?

Line 322: I recommend removing "even below the thermal threshold" since this is redundant with "sublethal."

Line 330: What are the observed differences in mortality rates that you refer to here? Differences between the last two studies you mentioned? Or differences between these studies and your study? Either way, I don't think this sentence is adding much so you could just cut it.

Lines 333-338: I think it's important to mention here that these are ambient air temperatures, while the temperatures experienced by bumblebees in the nest are likely to be much lower. As you mention, we unfortunately don't have much data on in-nest temperatures, but I think it's important to acknowledge that these 50 degree heatwaves probably won't translate to 50 degrees experienced by the brood.

Lines 355: Can you be more specific with what you mean by "activated"?