Dear Elena Salillas,

Please, find attached a revised version of our manuscript “Cognitive entrainment to isochronous rhythms is independent of both sensory modality and top-down attention” by Diana Cutanda, Daniel Sanabria and Ángel Correa.

We would like to thank you for considering our manuscript for publication and once again, we would like to thank the Reviewers for their final comments and suggestions. We detail our responses to all Reviewers’ comments in the pages that follow. Changes in the manuscript from the previous version are highlighted in yellow. We would also want to sincerely apology for the delay on the resubmission.

Thank you in advance for your consideration,
Ángel Correa on behalf of all authors.
1) Methodology and design

- Outcome measures

Mean RT is the only index of performance speed used here. Did the authors consider assessing (intra-individual) RT variability? Momentary boosts and/or entrainment of attentional functions might be particularly reflected in a (diminished) intra-individual variability of performance, not only in the RT mean score. That is, one would predict that performance becomes more stable (i.e., better controlled) rather than simply faster as a result of successful rhythmic entrainment. The authors may want to have a look at our recent work on measuring RT variability using percentile-point functions in this regard (Steinborn, Langner, & Huestegge, 2017). However, as the use of mean RT as performance measure is a standard in the field, I do not demand further analysis but would appreciate if the authors could touch on the aspect of performance stability/consistency after entrainment in the discussion.

We agree with the reviewer that it would be interesting to analyze the RT variability, since it could provide a stronger support regarding our results. We have now included in our manuscript the convenience to assess this matter in the future (p. 23).

- Foreperiod duration

The foreperiod duration is optimal to enable fast responding. One reason for the lack of effect in condition with low levels of anisochrony could be that there was simply not enough “room” for improvement imposed by the rhythmic prime presented before the foreperiod (i.e., due to a ceiling effect). This could be tested by comparing a short- and a long-foreperiod condition. One could predict that the relative benefits are greater in the long (vs. the short) constant-foreperiod condition, despite the greater distance from the rhythmic prime. Maybe auditory-rhythm superiority could be revealed by using such a (possibly more sensitive) condition with a long foreperiod. I suggest discussing this issue in a revised version of the manuscript.

The inclusion of two different foreperiod conditions (short and long) would result in a foreperiod effect (faster responses after the longest than after shorter foreperiods; Sanabria et al., 2011). In order to avoid this effect it would be necessary to also include catch trials. Following previous studies, catch trials would compromise between 12.5% (De la Rosa et al., 2012) and 33% of the total trials (Sanabria et al., 2011). Since our design included seven anisochrony conditions, the number of trials may not have been enough to perform reliable analyses with the inclusion of two different foreperiods and catch trials, especially after excluding those trials with premature and late responses and also those in which wrong responses were given for the n-back task.

We have now included this explanation in the manuscript (p. 8).

Temporal resolution of visual stimulus presentation. The visual stimuli were presented on a 60-Hz computer monitor, allowing the experimenter to differentiate time windows
of no less than ca. 16.7 ms. Thus, the manipulation of low levels of anisochrony (0, 10 or 20 ms) might in reality not have worked as precisely as planned and programmed, leading to a substantial number of trials with equal, rather than different, degrees of anisochrony. This temporally “noisy” set-up, in turn, may be another reason why there were not even numerical (let alone statistically significant) differences for these low levels of anisochrony, as compared to the 0-deviation isochrony condition, in the visual version of the task.

We thank the reviewer for this point. We have double-checked the monitor we used and realized that we made a mistake when describing the methods. The visual stimuli were indeed presented on a 100-Hz monitor, so we have now changed it on the manuscript (p 7). In any case, there were no differences in RTs not only between the 0 ms and the 20 ms, but also between the 0 ms and 50 ms conditions in neither the auditory or the visual modalities. Thus, it is unlikely that any differences would be found with smaller anisochronies.

- **Line comments**

  **page 5, lines 1-9**

  A lack of interference by the n-back task could also mean that this task is generally not effective to induce dual-task costs. A stronger argument at this point would be to provide a showcase where this variant of the n-back task effectively produced interference with a primary task, and then to argue that the same (effective) n-back task is not capable to abolish any benefit obtained by auditory or visual rhythms in a foreperiod situation as used here.

  We understand the Reviewers concerns regarding the effectiveness of the n-back task to interfere with the primary task. We decided to use this task after the results reported in our previous paper (Cutanda, Correa, & Sanabria, 2015). In that study, we first used a Sternberg type memory task and compared participants’ RTs to the target after either isochronic or anisochronic sequences. We did not find differences in the entrainment effect between sequences, and the overall RTs remained the same regardless the manipulation of the memory load. The substitution of the Sternberg type task, (which is purely a memory recovery task) for the n-back task (which requires also memory update) resulted in an overall increase of the RT for the 2-back condition, showing that the manipulation was effective: This argument is now included in page 5. Same results were found in the current study, with overall RTs increasing with the memory load.

  Moreover, previous studies have found that temporal processing of intervals and temporal orienting (temporal preparation driven by symbolic cues) are impaired when concurrently performed with tasks involving update of information (Rammsayer & Ulrich, 2011; Capizzi, Correa, & Sanabria, 2013; Capizzi et al., 2012).

  **page 8, line 15**

  **How was the start of the foreperiod announced in each condition? (visual vs. auditory)**

  The start of the foreperiod was not announced, since it could have interfered with the rhythmic entrainment. However, there were not intervals longer than 500 ms, that is, the
foreperiod was always longer than any of the intervals between the stimuli and, thus, participants were able to tell when the sequences had finished. They could also listen/see the target as many times as necessary prior to the experiment, to make sure that they recognized it.

**page 12, discussion**

The term "entrainment window" could be explained a bit more clearly, as the reference for assessing differences might be unclear to the reader. What’s more, how is the presence of “entrainment” established at all? In my view, entrainment would indicate the benefits of a given rhythmic conditions against a control condition where no entrainment is provided. I would appreciate a clarification at this point.

We agree with the Reviewer that the term “entrainment window” could be clarified. We have now described it as “a peak of the entrainment effect with faster reaction times” (p. 12)

As for the establishment of the presence of entrainment, rhythmic entrainment is defined in the manuscript as synchronization with external rhythms resulting in behavioural benefits such as reductions in the RTs (p. 3). These reductions in RTs can be observed between different anisochrony conditions. In our experiments we wanted to study the entrainment from a continuous perspective, opposite to the more common experiments were only isochronic vs anisochronic conditions are studied. From a dichotomic perspective, the anisochronic condition would constitute a control condition. Due to the anisochrony manipulation in our study, we cannot establish a “pure” control condition. However, the longer deviations of anisochrony work as a control condition when compared against the pure isochronic condition (0 ms).

**page 14, procedure**

A figure or visual analogy would be appreciated, if possible, so that the reader can instantly understand how the dual-task setting was actually implemented.

A schematic representation of one trial from Experiment 2 has been included (p. 15).

**page 17, Figure 4**

Little information is provided about the n-back performance. I wonder whether this information would be relevant for the interpretation of the findings.

Analysis on the accuracy performance on the n-back task is now reported, showing that the manipulation was effective, with higher accuracy for the 0-back than for the 2-back conditions. (p. 16). We have now included a figure of the n-back accuracy (p. 17)

**page 20, line 15**

Why not add the % of premature responses to the figures (but see Steinborn et al., 2008, Exp. 3, and for a review)?
The requested figures are now included as Supplementary material.