Author’s response to reviews

Title: Six weeks of balance or power training induce no generalizable improvements in balance performance in healthy young adults

Authors:

Louis-Solal Giboin (louis-solal.giboin@uni-konstanz.de)

Markus Gruber (m.gruber@uni-konstanz.de)

Andreas Kramer (andreas.kramer@uni-konstanz.de)

Version: 1 Date: 22 Jul 2019

Author’s response to reviews:

#Dear editor, dear reviewers,

thank you very much for your detailed comments and suggestions. They have helped us to improve on the overall quality of the manuscript. We carefully revised the manuscript according to your suggestions (changes are marked in red). You will find our responses to your comments written below:

Lucy Parrington, PhD (Reviewer 1):
The authors describe a study aimed at evaluating two different 6 week training interventions (balance v. plyometric) to improve balance and reduce fall risk on an untrained balance task. They found that neither intervention improved performance or acquisition on the untrained task, and conclude that the findings relate to task-specificity (or lack thereof), and indicate the need for more effective fall prevention programs.
The manuscript is generally well written. I feel the study in itself is interesting with regards to the intervention, assessing learning/ acquisition and task specificity.
#Thank you for this positive assessment!

But, I have a few concerns with this being framed around fall risk and balance assessment, and discussion on interventions for at risk populations.

1. The population tested is not an at risk population (young adults). How does testing this population relate to testing at risk for falls populations? Results are not generalizable outside of the population tested.
#We agree that the results of this study cannot directly be transferred to an at-risk population. However,
this does not change the rationale of the study (assessing whether strength or balance training improves the acquisition or performance in an untrained task). The idea behind testing a young healthy population instead of at-risk population is to better control the effects of balance training. Indeed, in at-risk populations, balance training must be seen more as a multi-modal training than a pure balance training (i.e. a training that can improve strength, endurance, power, fear of falling etc.). In this situation (i.e. where all parameters change, e.g. strength, power…), it is difficult to assess the effect of the skill training on untrained skills. We revised the manuscript (limitation section) to make it clearer that the results of the study have implications for fall prevention programs, but that studies in specific at-risk populations are necessary to confirm our results in those populations.

2. Given any balance deficit in this population is likely subtle in comparison to at-risk populations, could the lack of findings here be confounded/ceiling effect of running an intervention on a more fit population?

#Thank you for raising this point. Although this is a possibility we cannot completely exclude, the clear learning effects in the trained tasks as well as in the untrained tasks do not support this notion.

3. The trials were performed for 10 seconds. How do the authors know this test is capturing balance? Has this test been validated against other measures of balance or postural control? What are the sensitivity/specificity of the 2 systems used for testing balance? How well do they distinguish between people with and without balance deficit?

#We do not suggest that these tests (or any other test) can capture "balance". We and other research groups recently showed that the effects of balance training are very task-specific and that there is no single test or test battery that can assess balance (see Giboin et al., 2015 HMS or Donath et al., 2017 Sports Med). What we actually try to convey throughout the present study (and several past studies) is that the concept of balance as an ability should be replaced by a concept of balance as a set of skills that cannot be used interchangeably. We included the relevant literature in the introduction of the manuscript.

4. Trip and falls risk are often related to tripping, turning and episodes of freezing of gait in at-risk populations (i.e. Parkinson’s). Gait and balance are controlled differently. How does the testing of balance here, relate to fall risk?

#You are right. We revised the manuscript to make sure that the results have to be tested in at-risk populations to confirm our results in these populations.

5. Please explicitly outline what fixed and random effects, and what interactions were tested in the model, and if any covariates were tested within the model.

#You are right, sorry for this oversight. Please note that due to your comment on the normality, we have now slightly changed analyses. For the test on the balance performance: the random effects were the subjects, and fixed effects were groups and trials. We specifically tested an interaction between groups and trials. For the trials Pre-training, we have a random intercept by subject and a random slope by subject over trials (and their covariance was included in the model). However, we could not add random slopes by subjects for the post-training data since this addition prevented the convergence of the model. In this case, the model calculated correlations between fixed estimates only. For the tests related to the transfer: the random and fixed effects were similar to above. However, this time we could add a random slope by subject (the model could converge). It must be noted that for the test of transfer of performance we tested previously the averaged performance, which is less powerful than testing the 10 trials clustered by subject. Therefore, in the model we now take into account all the trials. Overall, this did not change the results. We included this information in the statistics subsection of the methods section and changed the results section accordingly.
6. Looking at the plot of the points of sensoboard time data, data seem to be centralized around 0-2.5s, with less people achieving 5s or higher. Did the authors assess normality of the data? Please present information relating to this.

This is a very good point, and we thank you for raising it! Indeed, with Q-Q plots we can see that the distribution is not normal. A square root transformation helped the Pre values of both the sensoboard and tilt board to reach a normal distribution. For the Post values, the transformation is helpful for the tilt board but less for the sensoboard. In any cases, the AIC was much lower for the models with the transformed data, indicating a better fit. Therefore, all our analyses were performed again with the squared root transformed data. It must be noted that results remained globally unchanged, only an interaction could be seen in the Pre-training for the sensoboard task which stemmed from a lower slope for the group Power compared to the group Balance. This point is shortly discussed in the Discussion section.

Minor issues:

1) Please remove colloquialisms (i.e. 'double edged sword')
   #You are right, we rephrased this sentence.

2) Can the authors please rationalize their approach of counterbalancing between subjects and not fully randomizing the order?
   #In training studies with a low to medium number of participants, randomization can lead to groups that already differ pre training, which considerably makes the interpretation of the results more difficult. We have added this point to the methods section.

3) Though the explanation in the methods is good, a diagram showing the progression of testing an intervention (flow chart) would be helpful for readers to get a visual on what was run.
   #We agree and have added a flow chart (new figure 1).

4) The final group composition demographics and the ANOVAs, could this be in a table? Help to de-clutter the paragraph.
   #As requested, we have placed the group composition in a table. We kept the ANOVA results in the text, as it seemed easier to understand (especially due to the presence of 1 post-hoc test).

5) Rather than F score throughout within the results section, a table with effect size coefficients (beta) for each effect and interaction within the model and confidence intervals would be preferable.
   #This is a good point. We have kept the F and p-values in the result section for readers not too accustomed to the present statistics, and have added a table (Table 2) with the principal model outputs.

Thomas Muehlbauer, PhD (Reviewer 2):

Title
Please include information on the investigated population (i.e., healthy young adults).
#This is a good point. As suggested, we included this information.

Abstract
Please include information on the investigated population (i.e., age, sex, training status) and training frequency and training duration per session.
#You are right, we added this information to the abstract.

Please indicate whether the trained/practiced tasks were also tested or did you performed transfer tests only?
#The progression of some tasks during the training was monitored, but they were not included in the test battery pre and post training. To avoid confusion, we removed the phrase about the improvement in the trained tasks from the abstract.

Did you detect a group by time interaction?
#Thank you for pointing this. We added the information that there was no interaction effect.

Please add information on practical relevance (i.e., an effect size measure) of your findings.
#The main object of our study was the interaction effects in the transfer and the learning rate in the untrained tasks, which were both non-significant. Therefore, we think that adding effect sizes of the interactions may not add much practical information and make this abstract section busy. The estimates of the models are now displayed in Table 2 (such estimates can be seen as effect sizes).

Introduction
Page 4, line 32: Explain the "task-specificity principle of balance training" a bit more in detail.
#We added some examples and details on this principle and the research it is based on.

Page 5, line 15: I was wondering whether the transfer depends on "incorporating many balance tasks" instead of "incorporating different types of balance (i.e., static, dynamic, proactive, reactive)".
#You are right, this is a possibility, although we did not find any training studies supporting this notion. We rephrased this sentence, it now reads "incorporating many different balance tasks".

Page 6, lines 24-39: Besides correlation analyses, there are several studies [2-4] available that showed transfer of balance training to muscle strength performances and vice versa, that is indicative of an cause and effect relationship. Please elaborate a bit more on that issue and describe corresponding literature.
#You are right, and we thank you for raising this point (and for providing the literature) as it shows that our introduction was not clear enough. The rationale of our paper is mainly to extend our previous work (Giboin et al. 2019 "Motor learning of a dynamic balance task: Influence of lower limb power and prior balance practice", JSMS), where we have shown an association between power and learning rate of a balance task, but no effect of one session of varied balance training on the learning rate of a new balance task. Therefore, here, the novelty of the study was to elucidate whether longer balance or power training can affect the learning rate of a new balance task. We have included two of the articles you proposed in the introduction and in general, added more references (we have not cited the paper from Ganacher et al., 2010, since there, it is not clear whether the tested balance task was different from the trained balance task).

Concerning your hypothesis: Why do you think that six weeks of training are enough to get transfer effects? Is there any hint for this training duration in the literature?
#You are right that this training duration might be too short to induce transfer after balance training. However, many balance training programs in clinics and rehabilitation facilities are limited to several weeks, which is why we chose this duration. In the discussion, we had already included the point that transfer might need much longer training durations.
Please state which of the two training regimen will show better transfer and indicate why?
#As the training regimens are quite different and could have induced transfer via different mechanisms, this is a two-tailed hypothesis favoring neither of the two types of training.

Methods
Page 6, lines 56: Please state what the "benefit and risks of the present investigation" are?
#We removed this phrase from the manuscript, as it does not add any valuable information. The risks and benefits are the usual ones when partaking in such a study: primarily a risk of injury due to falls, and muscle soreness. The benefits are the positive effects of the training (i.e., conditioning).

Please provide a figure that shows a subject performing the two tested balance tasks.
#As suggested, we have added a figure (new figure 2).

Please indicate which type of balance (static, dynamic, proactive, reactive) is represented by the two tested tasks.
#As our research supports the concept of balance as a set of skills instead of an ability with several subcategories, we do not believe that this classification is helpful. Instead, it promotes the image of four different types of balance, whereas our research indicates that there are as many subcategories as there are balance tasks. For example, we showed that training on a tilt board in antero-posterior direction induces no transfer on the same tasks with the same tilt board, only in medio-lateral direction (Giboin et al. 2015, Task-specificity of balance training, HMS). For the readers that still want to classify these balance tasks, we believe that the task description provides ample material for doing so.

How did you increase training intensity for the plyometrics.
#We tried to adhere as much as possible to the training described by Adams et al., 1992 (Journal of Applied Sport Science Research). Therefore, we modulated the training load of plyometrics mainly by changing the volume. For the drop jumps, intensity was raised by increasing the height of the jumps. For the other plyometrics exercises, subjects were instructed to perform them with the maximal intensity possible. We suppose that intensity was raised throughout the 6 weeks by their increased capacity to perform the exercises, but we had no mean to control that.

Please add a table that shows the 6-weeks power training program per session (sets, reps, weight, duration, exercises, no. of jumps or ground contacts, rests etc.).
#As suggested, we have added a table containing this information.

Please describe why the demands to perform the tasks during balance training differed from those when performing the two tested (untrained) tasks? There should be unique differences between the trained and tested (untrained) balance task if the goal is to show performance transfer.
#From a fall prevention point of view, it can be argued that it does not matter how different the tasks are, as the balance tasks potentially causing a fall can't be foreseen. In addition, the aforementioned concept of task-specificity suggests that it only matters if the task has been trained or not, we showed in a review that this is independent of how similar the untrained task is (Kümmel et al 2016, "Specificity of Balance Training in Healthy Individuals: A Systematic Review and Meta-Analysis "Sports Medicine).

How did you ensure balance training progression and/or increased balance training intensity?
#We ensured a training progression by changing the difficulty of the exercises and by modulating volume (see Table 2).
Consider providing an effect size measure to see whether your outcomes are of practical relevance. #See response above (abstract). The model estimates given in table 2 can be considered as effect sizes.

Results
Did you observe any baseline balance performance differences between the three groups? #You are right, this is an important point to consider. There were no differences between groups (apart from the learning slope of the Power group in the tilt-board task, which we discuss in the discussion section), see second paragraph of the results section.

Discussion
Please consider the following alternative explanation for the observed balance training effects: Shea and colleagues [1,6,7] showed that in accordance to the Hikosaka model [5] the transfer of learning (to different effectors, loads, and spatial requirements) is large for early in practice (after a single training session) but small for late in practice (after several sessions/days). Thus, you can argue that after six weeks of training (12 sessions in total) the participants achieved an increased level of movement automation (i.e., late in practice or high levels of movement experience) and the control of the trained motor task becomes more specific leaving less room for transfer.

On the other hand (and this could be investigated in a separate study), during stages of less movement automation (i.e., early in practice or low levels of movement experience), the control of a practiced motor task is relatively unspecific. That is muscle selection, computation, and the muscle activation is not effectively developed and therefore coded on a rather abstract level. As a consequence, the execution of movements with different task characteristics (e.g., untrained balance tasks) can easily be performed.

#We agree that this is a possibility, especially considering the amount of evidences brought by studies done with upper body visuomotor tasks (with for example the publication you mentioned, but also others e.g. Braun et al., 2009, Curr Biol, with a different type of “transfer” effect). However, we think that such transfer effects, and the underlying theories (e.g. the Hikosaka model or the model developed by Wolpert and co-authors), may not necessarily be directly applicable to balance tasks. Indeed, upper body visuo motor tasks and balance tasks are most probably related to different neural correlates, and therefore possibly different capacities of transfer/specificity (this point is shortly developed in Giboin et al., 2019 NeuroImage, in Press). Interestingly, in a previous study we found no transfer after one practice session of varied balance tasks (Giboin et al. 2019 "Motor learning of a dynamic balance task: Influence of lower limb power and prior balance practice", JSMS), which again tends to demonstrate the difficulty to induce transfer in balance performance with only balance training. This also support the idea that upper body visuo motor tasks and balance tasks may not share the same properties in regard to transfer capabilities. This information is included at the beginning of the discussion ("The present study expands on previous results where one varied practice session with several balance tasks and devices had no effect on the acquisition or retention of an untrained balance task compared to a control group (18).").

In the discussion section there is a lack in describing potential physiological and mechanical mechanisms that might be responsible for the assumed transfer effects of balance and power training to the two untrained balance tasks. Please add.

#In the introduction we shortly explained why we could expect generalization effects from both training types. Since we found no generalization effects, we preferred to discuss why no effects could be seen. In the second paragraph of the discussion, we proposed a short explanation of why we could
not see any effects. We have added a short explanation in the discussion regarding the lack of transfer effects from the balance training. Regarding power, we also have some potential explanation about the lack of effect. However, it is true that we have not delved so much in tentative physiological explanation of the presence/absence of effect in regard to balance performance, since we think that it would remain very speculative due to the lack of mechanistic research on this precise topic.