We have carefully read and considered both the editors’ and reviewers’ comments. Below, please find our detailed answers to each comment. We revised the manuscript to be in line with all modifications. All modifications and revised sections are highlighted in green in the original manuscript.

|  |  |  |
| --- | --- | --- |
| **EDITORS** | | **AUTHORS** |
| **Editorial Comments** | | |
| **01.** | Changes to be made by the Author(s): 1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version. | We thank the editorial team for this advice. We proofread the manuscript and revised spelling and grammar to our best knowledge. |
| **REVIEWER #1** | | **AUTHORS** |
| **Major Concerns** | | |
| **01.** | Are the title and abstract appropriate for this methods article? Yes. Although the term "gap" is used throughout the paper without providing an operational definition or even a reference to delta (besides in the figure) | Thank you for this comment. We agree and, thus, rephrased sections in the Summary (Line 33) and in the Abstract (Lines 44, 51, 54). Also, we added a sentence to the Introduction to clarify the meaning of “gap” (Line 73) and to the Statistical Analysis section to specify the definition and to refer to delta (Line 302, 350). |
| **02.** | Are there any other potential applications for the method/protocol the authors could discuss? This hard to agree with based on their research design and non-significant findings. | Thank you for indicating this issue. We agree with the statement, that it is hard to suggest other applications of the presented method/protocol, mostly because it is the first study to investigate the effect of a video on the accuracy of self-reported moderate-to-vigorous physical activity and final results remain to be presented. However, as stated in Summary, Abstract, and Discussion the subject of our manuscript is the whole trial whereas the video as the specific intervention content may be replaceable. As protocol steps work as intended and lead to data eligible to answer the research question (independent of whether results are significant or not), we conclude that the trial’s steps are reasonable. |
| **03.** | Do you think the steps listed in the procedure would lead to the described outcome? Yes. if powered correctly. The video addition to these procedures is interesting and warrants further investigation. | Thank you for this comment. This relates to our answer to comment number 02. We agree that further investigation is needed, also to provide reliable assumptions for a priori power analysis. |
| **04.** | Are any important steps missing from the procedure? The statistical analysis section is a bit thin. No mention of delta, or an a priori alpha (which leads to a wishy-washy result), what about effect size? | In line with JoVE’s scope and manuscript guidelines, the subject of the manuscript is method and design with emphasis on how to apply all steps relevant to replication. We based our manuscript structure on examples and templates provided by JoVE’s editorial team. Alpha as well as assumed deltas are presented in the Power Calculation section (page 4, Protocol section 2.3). The “wishy-washy” result (including deltas) refers to preliminary results of interim analysis as indicated throughout the manuscript. As agreed upon with JoVE’s editorial team, preliminary results are sufficient to give readers an idea of how results may look like. |
| **05.** | Are the anticipated results reasonable, and if so, are they useful to readers? No. They lost power along the way which may have contributed to their lack of findings. I would try again. Across several clusters. | This relates to our answer to comment number 04. Results refer to preliminary interim analysis. As we understand JoVE’s aims and scope, they are supposed to illustrate an example. Hypothesis testing is not the focus of the manuscript. Nevertheless, an analysis across several clusters surely is a good idea for future studies with respective sample sizes. |
| **06.** | Are any important references missing and are the included references useful? I believe the requisite referenced information in the decision making of several steps of the procedures requires attention. | We referenced all information that we found necessary for the reader to comprehend our decision. However, some aspects of study design are decisions to be made considering feasibility and effective trial conduction without being based on specific literature. If there are aspects that need to be referenced, please designate them specifically. |
| **Minor Concerns** | | |
| **01.** | Reference page needs attention. Caps for article titles? Inconsistent. Check throughout ref list. | Thank you for pointing this out. We revised the reference list accordingly (Line 497). |
| **REVIEWER #2** | | **AUTHORS** |
| **Major Concerns** | | |
| **01.** | My main concern is why the authors present the results of the interim analysis instead of waiting for the final results. | Thank you for this comment. We agree that final results would have been more informative. However, in line with JoVE’s scope and manuscript guidelines, the subject of the manuscript is method and design with emphasis on how to apply all steps relevant to replication. As we understand JoVE’s aims and scope, hypothesis testing is not the focus of the manuscript. As agreed upon with JoVE’s editorial team, preliminary results are sufficient to give readers an idea of how results may look like. |
| **02.** | Some other outcomes might be taken into account in additional analyses: the whole time of PA (including light PA); the proportion of moderate and vigorous PA. | Thank you for this advice. We agree that there are more outcomes worth to be investigated. We decided to focus on moderate-to-vigorous PA (MVPA) because this outcome is of main importance in the context of cardiovascular health (as indicated in the Introduction line 61). Nevertheless, light PA may be interesting to investigate because on the one hand individuals may overestimate it similar to how they overestimate MVPA. On the other hand they may have miss-classified light PA as moderate PA which could lead to lower light PA in the video group compared to controls. As JoVE manuscripts do not focus on outcome analyses, we decided to retain these ideas for future studies. |
| **03.** | I suggest the authors should discuss what the importance of more accurate self-reports of PA is. E.g., for purposes of confounder adjustment, correct ordinal information might be sufficient. | Thank you for this comment. In response, we added a sentence to the Introduction to point out research areas that may profit from higher accuracy (line 97). |
| **Minor Concerns** | | |
| **01.** | The power analysis cannot be changed any more but I am astonished that the authors assumed a gap of 90 minutes for the control group. I guess most people are far from having 90 minutes of MVPA, so a gap of 90 minutes is even more unlikely. | We thank the reviewer for pointing this out. As indicated in the manuscript (line 170), we based our assumptions on data of a comparable sample from a previous study of ours. Please see Baumann et al. *Scand J Med Sci Sports.* 2018;28:1056–1063 for information on sample and study context. Accelerometry-derived MVPA per day is very similar in both studies. The gap we found in the previous study amounted to 90 minutes per day which appears to be very much. But from our experience, the IPAQ is not the most user-friendly questionnaire. The item format makes it hard for respondents to accurately assess their PA. Apart from the difficulty to remember daily activities during the last seven days they need to remember the exact days, count the days, judge the intensity of the activities, count the amounts of time per day being active before breaking it down to an average. However, we agree with the reviewer that our assumption of the gap in controls obviously was set too high as indicated by our preliminary results. |
| **REVIEWER #3** | | **AUTHORS** |
| **Major Concerns** | | |
| **01.** | Line 145: I think that given you are assessing differences a Bland Altman plot should be used. | We thank the reviewer for this advice. We agree and, thus, added Bland Altman plots for both control group and video group to the Results section. |
| **02.** | Line 131: Why use the IPAQ? It may be useful to employ a questionnaire that records the types of activity given the evidence of accelerometers' inability to capture certain types of data (e.g., resistance training, inclined walking/running, swimming, biking, etc.) | Thank you for pointing this out. We chose the IPAQ because it is one of the most widely used PA questionnaires. Assessment of PA intensity levels seems inevitable in cardiovascular health research (as indicated in line 61). Thus, it seems reasonable to create strategies that might help to overcome limitations of the IPAQ. Our video is one approach to do this. Nevertheless, we agree with the reviewer that accelerometers may miss certain activities (as indicated in line 70) which adds to the gap (as discussed in line 411). Thus, comparing accelerometry with data derived from a questionnaire that records types of activity is an intriguing idea for future research. |
| **Minor Concerns** | | |
| **01.** | Line 86: Can you explain briefly the rationale for investigating whether effects differ according to participant's physical constitution? (please elaborate on the brief reference on line 66) | Thank you for this reasonable question. We are interested in results among overweight and obese persons because we assume that the numbers may be a little different compared to normal-weight persons. First, the gap among controls may be higher because they should be more exerted performing PA (and report this correctly in the questionnaire), whereas, to our knowledge, accelerometer counts do not account for relative exertion, thus, accelerometers may under-estimate PA intensity among overweight and obese persons. Second, the video may affect responses differently as obese persons may not identify with the model and his performance. We added a sentence in the manuscript to elaborate this issue a little more (line 66). |
| **02.** | Line 111: I agree these are important cues. What is the reason for highlighting these specific cues for intensity. Other cues may include mental focus, sweating (in some populations), how muscles feel, temperature. | We agree with the reviewer that there are more cues for intensity than those we addressed, such as those mentioned by the reviewer. We decided to address heart rate, breathing, and ability to talk because they are easy to understand and salient to individuals during their performance. In the attempt to “keep it simple” for respondents, we decided to keep the amount of cues to a minimum. |
| **03.** | Line 113: Using the treadmill is a nice way of depicting intensity in a controlled environment. However, it's possible that accurately identifying intensity is more difficult in other activities such as strength training, intermittent sport. Perhaps other activity intensity examples may be appropriate or at least mentioned. | We agree with the reviewer that the treadmill in a fitness center is a controlled environment and also may be a bit artificial. We decided to use this environment mainly for financial reasons. Shooting a video in more than one setting would have multiplied production costs. To accurately identify intensities of activities other than running absolutely would be more difficult. The problem with specific activity examples remains that the same activity may be moderate for one person whereas for another it may be vigorous. However, we mentioned other activity examples in the video as depicted in Figure 1. |
| **04.** | Line 200: How do you ensure the participant wears the device during all waking hours? Are prompts sent? When analyzing the data is a minimum wear time cut-off used- be specific (I see that now in your discussion line 373, please move to methods)? | Thank you for indicating this important issue. We pointed out the importance of wear time towards participants during recruitment but we did not send prompts. We agree that a cut-off should be applied in the analysis. Thus, we added step 6.2 to the Statistical Analysis section, line 299 and adjusted the results in line with the adjusted analysis sample (lines 314, 320, 326, 327, 351, 395 as well as Table 1 and Figure 3. However, since a cut-off value should be chosen depending on the specific research question, we decided to additionally keep the discussion of this issue in the original section (line 383-396). |
| **05.** | Line 262: Why vertical axis and not vector magnitude? | Thank you for pointing this out. We agree that using vector magnitude may be appropriate but we wanted to stay in line with previous publications for reasons of comparability. Nevertheless, we are thankful for the advice and are going to consider using vector magnitude in future analysis. |
| **06.** | Line 271: Please specify the population this is specific to, this would vary depending on level of mobility, age, etc. | We agree and revised this section accordingly (line 273). |
| **REVIEWER’S COMMENTS FROM ATTACHED PDF** | | **AUTHORS** |
|  | In addition to the comments listed directly in the email, we also received a pdf with further comments. This person stated: “I am not sure if I am the best person to review this paper. I am unfamiliar with the formatting of this paper. I have never reviewed a non-traditional journal article like this read my review through this lens. If presenting your data like this is typical for this journal then please take or leave any of my irrelevant comments.”  Considering some comments’ lack of application to JoVE’s specific journal format we agree with this statement. Thus, we did not take comments into account we believe to be irrelevant but indicated “**not applicable**” instead. | |
| **01.** | Line 52: I would place the hypothesis before the method sentences. | Thank you for this advice. We agree and moved the hypothesis (line 45). |
| **02.** | Line 55: Did you measure this? Then be more humble | This statement refers to our believe that a randomized controlled trial to test the effect of a video the way it is described in the manuscript should be applicable for other populations and contexts. |
| **03.** | Line 71: Could add our work here. | We believe the authors are blinded to reviewers’ names. |
| **04.** | Line 72: That is a big jump between sentences. No justiifcation as to why? | The reasons behind that are explained subsequently. The sentence was supposed to build a junction to the next section. We agree that this might appear a bit abrupt. Thus, we moved the sentence to the next section (line 76) |
| **05.** | Line 82: Why is this the target population? | In this sentence we describe the person in the video. We aimed at showing a person that is similar to our target group. We indicate the target group in section 3.3. |
| **06.** | Line 91: Says who? Place this after the results, if warranted. | **not applicable** (We structured the manuscript to our best knowledge of JoVE’s manuscript format and respective author instructions. As we did not receive editorial comments on moving contents to other sections of the manuscript we decided to adhere to the current version.) |
| **07.** | Line 93: This sentence too. Why here? Unjustified. | **not applicable** (We structured the manuscript to our best knowledge of JoVE’s manuscript format and respective author instructions. As we did not receive editorial comments on moving contents to other sections of the manuscript we decided to adhere to the current version.) |
| **08.** | Line 96: I am unfamiliar with this type of method presentation. Perhaps you need a reviewer more familiar with this journal. | **not applicable** |
| **09.** | Line 121: What do these subjective clarifiers mean? | The person that is shown in the video is described so readers may get an idea of the video content and which cues may have been salient to participants. |
| **10.** | Line 148: refs? | **not applicable** (Researchers who apply our method may do this among other population samples or in other study contexts than ours. Appropriate assumptions on mean differences in controls and intervention group are not fixed but need to be made according to the target population. The literature we based our specific assumptions on, is cited in section 2.3, line 170). |
| **11.** | Line 182: One word | Thank you for this advice. We corrected spelling as suggested. |
| **12.** | Line 213: Do you need references for all of this decision making? I think so. | **not applicable** |
| **13.** | Line 297: I would discuss delta here. | We agree. Revised as described in our answer to comment number 1 of reviewer #1. |
| **14.** | Line 298: alpha? effect size? | Revised as described in our answer to comment number 4 of reviewer #1. |
| **15.** | Line 313: Why only half? Do you disentangle this in the results? Why report on the other half? | In the manuscript, results of an interim analysis of a randomized controlled trial to test the effect of a video are presented. Thus, there is one group that received the video and one group that did not receive the video. Descriptives of both groups are reported. However, as our phrasing appears to be difficult to understand, we revised this section (line 320). |
| **16.** | Line 316: T value? Not significant. Call it like it is. | We thank the reviewer for this comment. As stated throughout the manuscript, we presented preliminary results of an interim analysis. To our best knowledge, our phrasing is appropriate to report results of an interim analysis. The presented results do not refer to hypotheses testing but to verification of early stopping of the study. Apart from that, we added the t-value to the manuscript (line 327). |
| **17.** | Line 336: These formulae could be better defined in the method section. | **not applicable** (According to JoVE’s format we believe that the method section, i.e. protocol, would not be the right place for this information.) |
| **18.** | Line 344: Again. Why are you introducing these concepts in the figure legends and not the method/results text?  I am not sure if I am the best person to review this paper. I am unfamiliar with the formatting of this paper. I have never reviewed a non-traditional journal article like this read my review through this lens. If presenting your data like this is typical for this journal then please take or leave any of my irrelevant comments. | **not applicable** (We structured the manuscript to our best knowledge of JoVE’s manuscript format and respective author instructions. As we did not receive editorial comments on moving contents to other sections of the manuscript we decided to adhere to the current version.) |
| **19.** | Line 364: Cant you just take a ratio if this occurs? | Thank you for this comment. To our knowledge, it is common practice to use cut-off values for accelerometer wear time. However, using a ratio is an interesting idea we are going to consider in our main analysis. |
| **20.** | Line 369: Your numbers are getting lower and lower, have you thought this may be the reason you did not reach significance? | Thank you for this comment. As stated above, we presented preliminary results of an interim analysis (n=131) as agreed upon with JoVE’s editorial team. Our results did not reach significance because they refer to an interim analysis. Apart from that, participants’ adherence to accelerometer wear time surely is an issue when investigating differences between self-reports and accelerometry data. Thus, we are discussing this in the manuscript. |
| **21.** | Line 378: Can you relate any of your suggestions to previous literature? Surely you are not the first researchers to make these suggestions. | We believe our suggestions refer to basic statistical knowledge, such as considering adjustment. for potential confounders, which is why we did not present references. |
| **22.** | Line 384: Which you did not accomplish? Where is your acceptance of the null hypothesis explanation? | **not applicable** (see above comments and remarks) |
| **23.** | Line 392: This paragraph is just a re-iteration of the introduction with no explanation of the findings, relation to literature, or practical application of the null findings.  My advice....conduct the full study and report those results. This paper does not add to the literature. Maybe once, you have found significant improvement in your video model, will others attempt to replicate your method for other populations. | **not applicable** (see above comments and remarks) |
| **24.** | Line 474: Caps for article titles? Inconsistent. Check throughout ref list. | Thank you for pointing this out. We revised the reference accordingly. |