Peer Community In
Health & Movement Sciences

Is effort evaluation domain-specific or general?

Boris Cheval *based on peer reviews by* ***James Steele****,* ***Ines Pfeffer*** *and 1 anonymous reviewer*

Wanja Wolff, Johanna Stähler, Julia Schüler, Maik Bieleke (2024) On the specifics of valuing effort: a developmental and a formalized perspective on preferences for mental and physical effort. PsyArXiv, ver. 3, peer-reviewed and recommended by Peer Community in Health and Movement Sciences. https://doi.org/10.31234/osf.io/ycvxw

Submitted: 06 September 2023, Recommended: 05 July 2024

Cite this recommendation as:
Cheval, B. (2024) Is effort evaluation domain-specific or general?. *Peer Community in Health and Movement Sciences*, 100041. 10.24072/pci.healthmovsci.100041

Published: 05 July 2024

The law of least effort suggests that, certis paribus, people tend to exert as little effort as possible when engaged in a goal-directed task (Cheval & Boisgontier, 2021). At the same time, however, large interindividual differences in the processing of effort have been observed, suggesting that effort per se can sometimes be valued positively (Inzlicht et al., 2018). However, until the present study by Wolff et al. (2024), all previous studies had largely ignored whether these individual differences in the valuation of effort might depend on the context (mental versus physical), i.e., in layman's terms, we do not know whether people value any effort or whether these preferences are specific to the mental and/or physical domain. The aim of the present study (Wolff et al., 2024) was to answer this question on the basis of two independent studies.

Study 1 (N = 39) used a binary decision task to measure preferences for allocating mental versus physical effort and showed that people differ markedly in their preferred allocation of effort. Crucially, a disposition to value mental effort (as assessed by the Need for Cognition Scale) was associated with a higher preference for mental effort, whereas a disposition to value physical effort (as assessed by the recently developed Value of Physical Effort Scale) was associated with a preference for physical effort.

Study 2 (N = 300 students) confirmed the robustness of the findings and showed that the tendency to value mental effort was associated with better grades in math (but showed no evidence of such an association in sport), whereas the tendency to value physical effort was associated with better grades in sport (but showed no evidence of such an association in math). Furthermore, the study extended these findings by showing that valuing physical effort was associated with less boredom in sports, whereas valuing mental effort was associated with less boredom in math.

In summary, the results of this research provide the first evidence suggesting that the valuation of effort is domain-specific rather than general. This finding paves the way for future research aimed at improving our understanding of the valuation of physical or mental effort. This article makes an important contribution to the knowledge of the key issues surrounding whether effort valuation is domain-specific or general.

Since all reviewers have indicated that they are satisfied with the authors' revision, which accurately and comprehensively addresses the reviewers' and my comments, it is my pleasure to recommend this preprint.

References

Cheval B, Boisgontier MP. The theory of effort minimization in physical activity. *Exerc Sport Sci Rev.* 2021;49(3):168-178. https://doi.org/10.1249/JES.0000000000000252

Inzlicht M, Shenhav A, Olivola CY. The effort paradox: effort is both costly and valued. *Trends Cogn Sci.* 2018;22(4):337-349. https://doi.org/10.1016/j.tics.2018.01.007

Wolff W, Stähler J, Schüler J, Bieleke M. On the specifics of valuing effort: a developmental and a formalized perspective on preferences for mental and physical effort. *PsyArXiv*, version 3. Peer-reviewed and recommended by Peer Community in Health and Movement Sciences.
https://doi.org/10.31234/osf.io/ycvxw

Reviews

Evaluation round #2

DOI or URL of the preprint: https://doi.org/10.31234/osf.io/ycvxw
Version of the preprint: 2

Authors' reply, 14 June 2024

Dear Dr. Cheval,

Again, thank you very much for your careful handling of our manuscript. We are pleased to know that the reviewers found our previous revision satisfactory and we appreciate the added comments by reviewer #3 which we have perceived as constructive and helpful. We have now addressed these comments and feel this has further strengthened the paper. Uploaded to this text you will find the detailed point-by-point letter & we have updated the preprint on PsyArxiv accordingly.

With best regards,
Wanja Wolff

Download author's reply

Decision by Boris Cheval, posted 21 May 2024, validated 21 May 2024

Request for revision of the paper *On the specifics of valuing effort: a developmental and a formalized perspective on preferences for mental and physical effort*

Dear Authors,

The first 2 reviewers involved in the initial evaluation of the paper were convinced by your response. This was also my case.

However, due to a potential COI, we have decided to ask for an additional reviewer for this additional step. This is to objectively ensure that our review process is aligned with the gold standard procedure to guarantee an unbiased decision.

This new reviewer has raised some very interesting points that I feel need to be addressed before I can recommend publication. Therefore, I ask for a final effort on your part to address the issues raised by this reviewer.

Boris Cheval

Reviewed by James Steele, 16 May 2024

Apologies for my tardiness with reviewing your revisions. It wasn't clear at first whether I was permitted to or not due to possible COI (for clarity for any readers of this review, the lead author of the manuscript is a recommender handling a manuscript of my own which is under review with PCI Health and Movement Sciences), then marking season hit, and then I picked up COVID again. Bum luck.

Anyway, thank you for your point by point responses. I have read through them carefully, checked the online materials and I'm happy at this stage that my main points have been addressed or appropriately responded to and I don't wish to hold this up any further. So I'm happy to suggest acceptance.

Thanks again,

James Steele

Reviewed by anonymous reviewer 1, 13 May 2024

I have carefully reviewed the manuscript revised by the authors, as well as their response to my comments and questions. I have also read the responses to the recommender and Reviewer 1. I am satisfied with the overall document. I would like to once again commend the authors for their manuscript, which contributes important information to the literature on the subject of effort, particularly concerning the distinction that can be made between preference for cognitive or physical effort.

Title and abstract

Does the title clearly reflect the content of the article? Yes

Does the abstract present the main findings of the study? Yes

Introduction

Are the research questions/hypotheses/predictions clearly presented? Yes

Does the introduction build on relevant research in the field? Yes

Materials and methods

Are the methods and analyses sufficiently detailed to allow replication by other researchers? Yes

Are the methods and statistical analyses appropriate and well described? Yes

Results

In the case of negative results, is there a statistical power analysis (or an adequate Bayesian analysis or equivalence testing)? Yes

Are the results described and interpreted correctly? Yes

Discussion

Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? Yes

Are the conclusions adequately supported by the results (without overstating the implications of the findings)? Yes

Reviewed by Ines Pfeffer, 18 May 2024

Thank you for having the opportunity to read and review this interesting paper. In this study the domain specificity of valuing effort (cognitive vs. physical) was examined based on two studies. Study 1 employed an decision task to measure preferred allocation of cognitive versus physical effort. This measure was shown to be associated with the dispositions of valuing cognitive and physical effort respectively. In the second study the authors showed that the disposition to value cognitive effort was linked to grades in mathematics but not sports. In contrast, valuing physical effort was linked to better grades in sports but not mathematics. Taken together, the results provide first results for the domain specificity of valuing effort.

Abstract

Line 24: I think that the sample in study 2 with a mean age of 15.25 years (± 1.57) should not be characterized as schoolchildren. I think that adolescents would be the better characterization.

Introduction

Lines 33-35: I think it could be helpful to refer to and discuss to the Theory of effort minimization in physical activity by Cheval an colleagues in this context. This theory argues that that individuals have an automatic attraction toward effort minimization. Furthermore, this automatic attraction needs to be overridden by controlled processes to become physically active. In how far are the assumptions of this theory relevant for the concept of valuing physical effort?

Line 56: I was wondering if people perceive the effort invested in intrinsically motivated (self-determined) physical or cognitive behaviors as less aversive than in extrinsically motivated (non-self-determined) behaviors? Is valuing the effort invested in an activity associated with the type of motivation? What if learning the language is intrinsically motivated, feels pleasant and is associated with enjoyment?

Line 143-144:

Methods

Line 156: Please provide the median for the income of the participants.

Line 181ff: Please provide examples for the cognitive and physical activities that were used in the task to assess preferred effort allocations (e.g. Activity A and Activity B in Figure 1)

Line 214f: The German grade system is an ordinal scale and therefore parametric statistics such as Pearson correlations might not be suitable.

Results

Lines 249-252: How can the differences in the mean values between the two samples be explained? Please discuss in the discussion section.

Line 266: In the description of Figure 2 it is stated that "For study 1 the answer scales ranged from 1-7, whereas for study 2, they ranged 266 from 1-5." Why was the answering scale different between the two studies?

Figure 4: In the labeling of the x-axis the abbreviation ME (mental effort) should be changed in CE (cognitive effort).

Line 301ff: As mentioned above, grades hold an ordinal measurement level. Is a linear regression analysis adequate to examine the hypothesis?

Evaluation round #1

DOI or URL of the preprint: https://doi.org/10.31234/osf.io/ycvxw
Version of the preprint: 1

Authors' reply, 10 April 2024

Download author's reply

Decision by Boris Cheval, posted 13 October 2023, validated 13 October 2023

Request for revision of the paper *On the specifics of valuing effort: a developmental and a formalized perspective on preferences for mental and physical effort*,

I have now received two reviews of your manuscript *On the specifics of valuing effort: a developmental and a formalized perspective on preferences for mental and physical effort*, and I have read the paper carefully myself. As you can see from the notes below, the reviewers, who are highly qualified with respect to the topic of the paper, highlight particular strengths and note the importance of your work in helping us better understand some of the key questions revolving around whether effort valuations is domain-specific or domain-general. I fully agree with the reviewers that this is a very interesting question that has not been fully addressed in the current literature.

However, they also raised some questions and concerns that need to be addressed before I can recommend your paper for PCI Health & Movement Science. I will not repeat all of the comments, but I agree with most of them. In particular, I concur with the following points that I believe need to be carefully addressed to improve the overall quality of the manuscript:

(1) Add a definition of effort.

(2) Justify some inferences (e.g., the relationship between engaging in more or less effortful activities and the cost of effort, the use of school grades and sports performance as proxies for actual behavior).

(3) Explain why motivation to engage in the task and perceived task difficulty were not included, despite the current study's reliance on motivational intensity theory, in which these two variables are critical in explaining effort (dis)engagement. If relevant, discuss the potential implications for data interpretation of not including motivation and perceived difficulty.

(4) Review the data and the R script. This includes documentation of the data set (codebook). Some questions for my part:

- I wonder how the results of Study 1 and Study 2 can be integrated, since the samples are drastically different (e.g., differences in the mean scores on the scale and their reliabilities). Moreover, since the results of Study 2 are, in my opinion, rather inaccurate proxies for the valuation of effort (effort and performance can be drastically different, especially in math and sports), I wonder whether the second study adds more "value" than "harm" to the current article. But of course, this is a decision for the authors. I just raise this point from my outside perspective.
- One last question/comment: Do you think that, all things being equal (i.e., the level of expected reward associated with the behavioral alternatives is identical or no reward is expected), people may choose to **repeatedly** engage in the more effortful behavioral option? I would appreciate the authors' thoughts on this comment. Thank you again for submitting your manuscript to PCI Health & Movement Science. I look forward to receiving your revision.

Reviewed by James Steele, 09 October 2023

Thank you for the opportunity to review this manuscript and apologies for my tardiness with submitting my comments. I enjoyed reading the studies reported here a lot given my interest in whether or not effort related phenomena are general or domain specific (such as their psychophysics) and hope my comments are useful.

James Steele

General comments

In general I found the manuscript to be very well written, succinct and clear. In some places though, particularly in the methods, I felt there could be more detail provided which could be added to the supplementary materials so as to keep the manuscript itself succinct. The title and abstract reflect the content of the article well. The introduction does a good job of presenting the theoretical rationale for the study and the research question is clear, though you could more explicitly state you're a priori hypotheses, though these are deducible to an extent from the introductory discussion of theory and prior empirical work. As noted, there are elements of the methods I think could be expanded upon for clarity and I note these more specifically below along with some questions from myself. The results are clearly reported and visually presented, though I will suggest some alternative approaches to presentation that might aid interpretation. The discussion is appropriate and clear, relating the findings back to prior hypotheses. The scope of the findings and limitations are also clearly discussed.

Specific comments

Principle or law of least effort - In mentioning the principle/law of least effort I would try to be consistent in what word you use throughout. To be honest I would lean towards principle as the existence of something like the 'effort paradox' implies that simple effort minimisation is not law-like.

Mental or cognitive - Minor point, and depending on how mental is defined might just be semantic, but 'cognitive' might be a better word... Bruya and Tang discuss this in their interpretive analysis of Kahneman's Attention and Effort - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6136270/ - In fact, cognitive might be better and align with your use of the Need for Cognition Scale.

Effort definition – You don't define effort at any point during the manuscript. Given the variability of definitions across the literature it might be worth offering the definition you are employing in this work. Also, whether or not you are referring to effort or its phenomenology.

At the beginning of the 3rd paragraph of the introduction I would add "e.g.," to the examples in parentheses.

Sample size – As the study was not pre-registered for particular model specifications and to test certain hypotheses for particular parameters I would remove the following comment:

"The achieved sample sizes exceeds the required threshold for conducting the planned statistical procedures. For example, our sample is sufficiently large to provide high power to quantify significance of small effects in regression analyses."

I would instead be explicit in stating that the study was exploratory, though you may have had theoretically driven hypotheses to examine, and that sample size was necessarily resource constrained (either by time, funds, access to participants etc.).

McDonald's omega – For those not familiar, I would specify that this is what is being reported when you refer to the different scales reliabilities. (Also it's just an observation, but I wonder why these scales haven't been explored in the context of Item Response Theory... for example, it would be interesting to explore the fit of the partial credit model to them)

Study 1 multinomial model - I had a quick read of the papers cited as I am not wholly familiar with the ring task... It wasn't clear to me why participants should be classified and then this be the variable modelled - classification of a continuous variable inevitably loses information, some people might be just on the edge

of a category whilst others are more clearly within it, so why not model the resultant angles directly? Or the coordinates for PE and ME? Some further justification of this choice might be useful (and perhaps exploring different model specifications and the sensitivity of your substantive conclusions to them as this is exploratory research). If you opt to stick with the multinomial model as your main one though I'd probably suggest reporting the results as probabilities for each category as opposed to odds ratios as everyone tends to find the latter less interpretable as far as I know. Or if you went for the continuous model of the angle you could report predicted angles and their CIs for varying levels of each predictor and then post modelling interpret what categories they tend to fall in.

Dataset - Mind you, in thinking about the angle model I was trying to figure out how the pmeo variable in the dataset (which I assume is the angle from working through your analysis script) was calculated from the rest of the data but it appears this has been processed prior to uploading the data. I think the accompanying dataset may need a data dictionary to explain what each variable is to enhance reproducibility… also, for each of the presented 24 choices would be good to know so I would upload that to the supplementary material or add the PE and ME to each choice in the data. At the moment I assume it is just 1 or 2 depending on which option they chose.

Study 2 multivariate model - I got the impression that the grade outcomes, and the boring-ness responses, were separate substantive hypotheses. So, I don't think it is necessarily needed to model it as a multivariate outcome. It might be worth clarifying why this approach was taken by relating it back to your substantive hypotheses more clearly. Also, both outcomes seem to me to be ordinal in nature (I assume for the grade not really knowing whether any particular measurement models are employed for German school grades) and so an ordered logit or probit model might be more appropriate. You could again present predicted probabilities over the range of predictor values for each ordered category.

Descriptive statistics and t-test – I wonder about the appropriateness of the t-test between scales. It's not clear to me that these scales are necessarily comparable in anything other than an operational sense. Also, even if we grant that the operations are similarly capturing the value of effort in the same scale for each domain, the difference seems small even though significant. I am also not sure that you really need to report this given the correlations and network model.

Limitations – To me it is a very important point when you note "… school grades are a proxy of actual behavior, as the effort invested into school tasks does not necessarily lead to better grade.". Some people may have low ability but still value the effort required for those tasks and so put in a lot yet still produce low grades. This links to some extent to the manner in which you conceptualise effort though, and perhaps a reason why offering your concept definition is valuable.

Domain specificity of effort psychophysics? – A thought crossed my mind that I just thought I'd try to flesh out and share. Feel free to ignore or if you think relevant maybe discuss.

The interpretation of these results to some extent may differ depending on whether the psychophysics of effort is general or domain specific. If general (and assuming strong identity of phenomenology), then results such as yours suggest people may well value effort in different domains differently. But if the psychophysical relationship were domain specific also (differing in the strength of its identity perhaps across domains and at different levels of actual effort), then apparent differences in valuation might be confounded by this. For example, let's say that at an equivalent level of effort (I am assuming my own conceptualisation of actual effort here - https://psyarxiv.com/kbyhm) across both a physical and cognitive task is attempted but a person perceives it differently in either domain. Or that different tasks requiring different actual effort are

perceived similarly effortful. It then becomes difficult to tease apart whether their valuation from a binary decision task is due to the differences in actual effort required, or what they perceive it to be. Of course, this may not be an issue for interpretation of your results specifically as you have used hypothetical tasks and we could perhaps assume that participants imagined tasks based on your instructions and thus remembered the associated phenomenology, or if they had not experienced such tasks they instead attempted to forecast what it might feel like. So, their valuation would be assumed to be based on how effortful they thought it might feel. This is where I like your suggestion of examining this with real tasks. In this it might be possible to try to present choices where the actual effort required for the tasks presented is known and on the same scales for physical and cognitive (though, whilst easy for physical it's much harder for cognitive as I have found e.g., https://psyarxiv.com/6pvht), and you could ask participants to both forecast how much effort they thought they would perceive the task would require and also ask them to report it when completing the chose task.

Code – I would double check your code script as I noticed at least one typo (line 101). I also can't seem to reproduce your network graphs (I seem to also get three communities for study 1). There are also some arguments that seem unnecessary in certain parts (e.g., in the raincloud plot scale_X_discrete colors).

Reviewed by anonymous reviewer 1, 07 October 2023

In this article, the authors explored the preference individuals have for exerting mental and/or physical effort and the potential impact this can have on their performance (as reflected in academic results), and the boredom experienced in these two types of activities. To do this, they conducted two studies. The first was conducted online and aimed to demonstrate that there is a distinction in preference between mental and physical effort. This was achieved by using the ring task and two scales. Following this initial study, the authors applied their findings to a real-world context: school. Students filled out these scales, and correlations were drawn with their performance in mathematics and physical education (serving to establish a connection between mental and physical effort). They also reported their usual level of boredom in these two academic activities.

I truly enjoyed reading this article. It provides a missing piece of information in the literature, namely that the two types of efforts can be perceived and evaluated differently depending on the individual. The two studies presented in this article are an important first step toward delving deeper into the exploration of this dichotomy. I also greatly appreciated the outlined limitations of these two studies, showing that the authors do not make claims beyond what the data indicates. This is pleasing.

However, I have several comments to make regarding this article. You will find them below, with the first section containing general comments, the second section containing specific comments, and the final section including a few remarks concerning the data and the R script.

General Comments

1) I am somewhat surprised not to see a measure of motivation, particularly in the second study. Indeed, a person is more inclined to invest effort when the motivation to perform the task is high. On page 15, the authors mention a grain that could be a bit finer to measure the value of effort in more specific domains. Does this not refer to the motivation to engage in this activity? Aren't preference for an activity and the motivation to undertake it two sides of the same coin?

I noticed that the authors used the Motivational Intensity Theory in their reasoning (references 1 and 9). The Motivational Intensity Theory (Richter et al., 2016) suggests that potential motivation is the limit of the effort one invests in a task. Thus, effort can only be invested when there is a motivation to exert it. In the same vein, the authors' first sentence in the abstract is that "Effort is instrumental for goal pursuit". This indicates that there is motivation involved in achieving a goal. I understand that these two studies were not designed to measure motivation, however, I believe that a paragraph in the discussion, in the implications section for

instance, would provide significant additional insight. Another option could be to discuss this point in the limitations / perspectives section as it was not measured, but is highly relevant for future research.

2) Following on from the first comment, the Motivational Intensity Theory (Richter et al., 2016) predicts that the effort invested in a task is a function of the perceived difficulty of that task, with potential motivation being the limit (the maximum one invests). How can this statement be reconciled with the fact that here there is a preference for investing effort in one domain rather than another? Is the perceived difficulty different?

Specific Comments

1) Page 4: An inference is made between the preference for engaging in more or less physical/mental effort and the cost of that same effort (”a preference for less mental and/or physical effort will be a proxy that the corresponding effort is costly...”). This is a logical inference made here and at first glance seems valid. However, to the best of my knowledge, the link between the two has not been demonstrated, and no citation is given for this sentence. Therefore, this inference should be included in the limitations of the article since the connection between the two does not seem to be demonstrated at the current time, or a reference should be added to support this statement

2) Page 5: The authors chose to conduct the first study online via MTurk. The justification for this choice is not provided in the body of the text. Is it to have more participants more quickly? There are also other platforms besides MTurk for conducting online studies (e.g., Prolific, CrowdFlower, etc.). Some of them seem to yield higher-quality data (Peer et al., 2017, 2021). Can the authors justify the choice of MTurk? The study here is relatively simple and therefore probably minimally impacted by the quality of the respondents, especially with the verification questions posed to participants.

3) Page 5, it is indicated that there are 37% female in the first study, and 62% in the second. Please also indicate the number of individuals who chose not to respond, as well as the number of ”other” gender responses (1 in the first study I believe, and several in the second). Also, indicate the rate of missing data. I also wanted to know if it was sex at birth or gender that was asked of the participant. I assume it is the gender with which the participants identify, given the dataset and the statement on pages 6 and 7 about additional measures. If this is the case, the word ”female” (sex) should be replaced with ”women” (gender).

4) Page 5: it is mentioned that the VoPE scale contains 4 items. In the dataset available on OSF, I can see 10 columns named vope or voes. Is this the same scale? Why is there a discrepancy between the 4 mentioned in the text and the 10 present in the dataset? If it is indeed the same scale, why administer all 10 items of the scale instead of just the 4 of interest?

5) Page 5, following up on the previous question: I am wondering about the difference in the number of items between the two scales, particularly when it comes to the clusters that can be found later in the article. For the NfC scale, the authors observe two clusters, but only one for the VoPE. Since the VoPE only has 4 items (used), it is more challenging to find multiple clusters. Isn’t this a limitation to the cluster analysis? Without a counterargument, I believe that this is a limitation to be noticed.

6) Page 6 and 7: The authors explain on page 6 that the middle of the scale is 50 (and therefore the center of the circle has coordinates (50,50)). I think it would be interesting for the reader to see the value of 50 appear on both the x-axis and y-axis in Figures 1B and 3.

7) Page 7: Figure 1A is too small; once printed, it is no longer legible. Please enlarge it (perhaps by rotating it 90° and extending it across the width of the page).

8) Page 8: It is not indicated whether the authors checked for outliers. I assume there are none in this case. Can they confirm this?

9) Page 8: In the dataset for study 2, I noticed that some data were missing for certain participants. This is typically observed when working with children, not an issue for me. This information is indirectly reported on page 8 during the t-test between NfC and VoPE with n = 284. However, I believe this information should appear somewhere, possibly detailed in supplementary data.

10) Page 8: Could you please add the effect sizes of the t-tests? This allows researchers who wish to conduct

meta-analyses to do so more quickly, in addition to indicating the strength of the difference to the readers.

11) Page 8: The authors state that the correlation between VoPE and NfC is weak and therefore physical effort is, on average, evaluated differently from mental effort. I understand what the authors are trying to convey. However, since the correlations are significant, even if it's weak, it means that a higher score on one scale is associated with a higher score on the other. This sentence needs to be revised. Moreover, this raises the question of the threshold from which the authors would indicate that the two scales vary simultaneously (e.g., r = .20 or r = .30). This value would inevitably be arbitrary.

12) Page 9: Is it possible to create clusters with more distinguishable colors? It is mentioned on page 8 that there are three clusters, two of which come from the NfC scale. However, the color does not allow for easy distinction between the two clusters, especially in the case of study 2. Finally, regarding the graph, the scale is referred to as NfC in the body of the text, but as ndfc in the figure. I suggest that the authors harmonize the two or explain the difference.

13) Page 10: Figure 3. This is just a suggestion. While reading (on paper), I wrote the percentages indicated in the text on the figure to better remember them. It might be interesting to include them directly for the reader, for instance in the blue part of the figure.

14) Page 11: The authors mention that they measured gender, but it is not indicated whether this had an impact on the results. Does gender have an effect on the correlations or on the perceived boredom? Moreover, the chosen school subjects (mathematics and sports) are seen as more masculine or at least preferred by men (Eccles et al., 1993). Could this have an impact?

15) Page 13 at the top: The score on the VoPE also predicts boredom in mathematics, even though this is to a lesser extent. Given the huge expertise in boredom of some authors of this article, it might be interesting to mention and briefly discuss this. It would also be interesting to put these results in perspective with those of study 1, particularly with figure 3 where we can see that there are very few people who only maximize physical effort, but those who do maximize this effort also seem to maximize mental effort.

16) Page 13: "Accordingly, researchers should be specific about the kind of effort they are addressing in their research". I couldn't be more in agreement with this statement!! This sentence could maybe be in a better position in the conclusion as a final take-home message.

Comments on the Data and the R Script

1) In the data from study 2, I noticed two errors that seem to be a typo. Could you please check the other data and redo the analyses correcting for these errors? I imagine that this should not fundamentally change the results since only two data points appear to be erroneous.

A. Participant 83 has a value of 6 for item 1 of the NfC scale, while the German version of this scale indicates on page 7 that the scale ranges from 1 to 5.

B. Participant 78 has a value of -5 for item 4 of the NfC scale.

2) A read-me file should be added with the data to explain the variable names. The data files are currently hard to understand, making it difficult to replicate the analyses. For instance, I do not understand what the columns pmeo, gspa, or gsma represent.

3) Page 5: Regarding the age, it is correctly indicated in the article, but in the provided R script, it says "strange answers", and the results for nationality and level are not indicated either. This information is just for you, to harmonize.

4) Page 5: It is mentioned that the materials, including the questionnaires, are available on OSF, however, I could not find the file containing the questionnaires. Could you please add it, or remove the private option on OSF as the sentence indicates that they are available? If not, please adjust the statement.

As I mentioned at the beginning of the review, I truly enjoyed reading this article. I commend the authors for the work they have done. I am confident that this article will have a significant impact on research regarding

effort.

Bibliography

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and Gender Differences in Children's Self- and Task Perceptions during Elementary School. Child Development, 64(3), 830. https://doi.org/10.2307/1131221

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk□: Alternative platforms for crowdsourcing behavioral research. Journal of Experimental Social Psychology, 70, 153□163. https://doi.org/10.1016/j.jesp.2017.01.006

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2021). Data quality of platforms and panels for online behavioral research. Behavior Research Methods, 54(4), 1643□1662. https://doi.org/10.3758/s13428-021-01694-3

Richter, M., Gendolla, G. H. E., & Wright, R. A. (2016). Three Decades of Research on Motivational Intensity Theory. In Advances in Motivation Science (Vol. 3, p. 149□186). Elsevier. https://doi.org/10.1016/bs.adms.2016.02.001