RESPONSE LETTER

Dear Dr Cheval,

we want to thank you and the reviewers for your positive evaluation of our manuscript and the valuable feedback. Please find our responses to your and the reviewers' comments below.

Best regards,

Maik Bieleke, Johanna Stähler, Wanja Wolff, and Julia Schüler

Dear Authors,

Thank you for your submission titled "Development and validation of the Value of Physical Effort (VoPE) scale". The study establishes the basic psychometric properties of the VoPE and tests associations with measures of sports and exercise behavior.

I am pleased to inform you that the review process for your manuscript has been completed. In my own reading of the manuscript (and echoed by the comments from the two reviewers), this a very well-done study with several strengths, the methods and analysis are rigorously conducted, and the presentation of the results is transparent. I found the points raised by the reviewers important to address before I can recommend publication. I agree with all of them and would like to highlight two in particular:

1) A critical aspect that needs further elaboration is the process used to arrive at the final 4-item version of the scale. The current description lacks sufficient detail, making it challenging for readers to fully understand the methods and rationale underlying the selection of the four items. In my opinion, the manuscript would be strengthened by providing more detailed information on the procedural aspects of scale development.

We agree and added this information to our manuscript. Please see our comments to both reviewers for details.

2) Another important consideration is the "dual nature" of effort, at least in the context of physical effort. Based on the existing literature, it can be argued that effort can be perceived as both aversive and positive, depending on various factors such as individual differences, context, and prior experience. The manuscript would benefit from a more nuanced discussion that incorporates this point.

This is an important point, and we addressed it by answering the corresponding reviewers comments. For instance, we added links to the theory of effort minimization in physical activity (see Comment 1 by Silvio Maltagliati) and provide a more nuanced discussion of what contributes to the valuation of effort (see Comment 2 by Erik Bijleveld).

Therefore, I encourage the authors to address all of the reviewers' suggestions and incorporate them into their manuscript.

Best regards,

Boris Cheval

Review by Silvio Maltagliati, 18 Apr 2024 12:52

Overview and general recommendation:

I have read this manuscript with the greatest interest and I feel it stands out by its sophisticated methods, the transparency of reported results, as well as by the quality of its figures. I also have a few comments and I hope they will help to improve the overall quality of the manuscript.

Thank you very much for your kind words! We carefully addressed all of your comments and suggestions, and we hope you agree that doing so has improved the manuscript.

Introduction:

1- The Introduction section is straightforward and nicely expands on the fact that despite its seemingly aversive nature, effort can also be valued, at least by some individuals. At least one question came to my mind. How would the tendency to value physical effort fit other theoretical models and would potentially, enhance their predictive value? For example, could the value assigned to physical effort moderate the link between intention and physical activity behavior (see some work my research team is conducting at the moment: : https://doi.org/10.51224/SRXIV.375) or could it be rather be involved in mediational pathways between other motivational factors and physical activity? I have the feeling this point could either be addressed in the Introduction section or in the Discussion section and would facilitate the theoretical integration of this construct within other motivational models.

Thank you very much for drawing out attention to this research! This is highly relevant for our work and we agree that locating the VoPE scale in these theories is promising. We now discuss this in our manuscript as follows:

"On a related note, accounting for individual or contextual differences in the value of physical effort might help to advance theoretical models of physical activity behavior, such as the theory of effort minimization in physical activity (Cheval & Boisgontier, 2021). In line with predictions derived from this theory, a stronger inclination to approach (vs. avoid) physical effort has already been shown to foster the link between the intention to be physically active and physical activity (Maltagliati et al., 2024). It is plausible that people who assign greater value to physical effort also tend to approach effort more readily. Testing ideas like this and locating the value of physical effort in theoretical models is thus a promising endeavor for future research."

Methods and results:

1- I think that some important points are missing regarding the development of the scale. For example, how were items developed (e.g., number of authors involved, discussion between them), were more items initially developed (before items reduction)? Also, I wonder whether retaining four items just because they were positively framed items makes complete sense from a psychometric perspective – using reversed items could also be relevant and could had further enhanced the reliability of the scale? I think that these methodological steps should be clarified and/or identified as missing in the Discussion section. I think this article could be of interest: https://doi.org/10.3389/fpubh.2018.00149

Thank you for this making this important point. We added several details about how we constructed the scale. We hope that this clarifies your questions about our approach.

"In Study 1, the VoPE scale was administered as part of a general survey. To construct the VoPE scale, we adapted items from the Need for Cognition Scale (NfCS; Cacioppo et al., 1984) — a well-established instrument that measures "an individual's tendency to engage in and enjoy effortful cognitive endeavors" (p. 306). We jointly reviewed, discussed, and ranked the 34 NfCS items regarding their suitability for adoption to the domain of sports and exercise. We ended up with a set of ten items that provided the best fit in terms of adoptability (5 positively and 5 negatively framed items) and administered them in all three studies. We translated the VoPE scale from German (Study 1) to English (Studies 2 and 3) with the assistance of AI technology (www.deepl.com; see Table 1 for both language versions). After data collection, we reduced the items to the four items that make up the final VoPE scale (see Table 2).

First, we excluded the five negatively framed items because they formed a separate community in network analyses, likely reflecting a common method artifact that plagues mixed framing scales (Lindwall et al., 2012). The NfCS has also been criticized for introducing artificial factors through its negatively framed items (Zhang et al., 2016), which we aimed to avoid. Second, we removed one of the positively framed item which, in retrospection, referred to an artificial situation that might elicit ambivalent responses: "When I know that a sports activity is going to involve physical effort, I look forward to it, even if I do not know what exactly it involves." Despite these exclusions, all ten items are available in the published dataset on OSF for transparency. Participants indicated their agreement with each item on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree), and their responses were averaged into an overall VoPE score."

2- Another important limitation stems from the fact that the scale seemed to be developed in German originally, but then MTurk samples were composed of US citizens. How was the translation of the scale from German to English done? Ensuring cross-cultural equivalence appears as an important step to ensure comparability across nationally diverse samples. More details could be provided here and would deserve to be discussed later in the manuscript.

Thank you for making us aware of this point. We used a pragmatic approach and translated the items with the assistance of AI technology (www.deepl.com). That being said, the focus of our work was not on performing cross-cultural comparisons between studies. Differences between our samples beyond cultural background would render such comparisons difficult anyway (e.g., paid crowd-workers vs. ordinary participants, prescreening of physical activity in Study 3). However, we provide both the English and the German wording of the items in the manuscript, allowing other researchers to evaluate the quality of our translation and to modify them as they see fit. We added the information about the translation process to the manuscript.

"We translated the VoPE scale from German (Study 1) to English (Studies 2 and 3) with the assistance of AI technology (www.deepl.com; see Table 1 for both language versions)."

2- The Methods and Results sections are very dense, because of the diversity of measures/analyses that were conducted. I am not an expert in some of these statistical approaches (e.g., machine learning, network analyses), but they seem relevant and adequately described. One point I struggled to follow was the number of participants that were retained for each analysis, as well as the Study it corresponds to. Could authors please consider adding a

Table (or a Figure) summarizing the number of participants that were included for each analysis, as well as whether participants from Studies 1, 2 and 3 were gathered or not? For example, it was unclear to me which sample was used for machine learning analyses – was the full sample used there?

Thank you for this suggestion! We added the info about studies and relevant samples sizes in the headers and in the main text throughout the methods and results section. Additionally, we added this info to the tables and figures. We hope you agree that the relevant information is now readily accessible.

2- Also, I would be very curious to see whether the VoPE scale score correlates with some sociodemographic determinants (e.g., with age or education level)? Also, have authors considered adjusting the models predicting physical activity levels for some of these relevant sociodemographic factors?

Thank you for bringing up this point. Indeed, we observed significant pooled correlations of VoPE score with all sociodemographic variables and added this information to the manuscript.

"Finally, we observed significant pooled correlations of VoPE score with age, r = −0.09, 95% CI [−.14, −.04], z = 3.41, p < .001, gender (0 = female, 1 = male), r = 0.22, 95% CI [.17, .27], z = 8.19, p < .001, income, r = 0.19, 95% CI [.13, .25], z = 6.41, p < .001, education, r = 0.09, 95% CI [.03, .14], z = 12.31, p = .005, and subjective SES, r = 0.19, 95% CI [.13, .24], z = 6.21, p < .001. Most individual correlations were in line with these meta-analytic findings, except non-significant links between VoPE and age in Study 1 and between VoPE and education in Study 2 (see OSF for detailed results)."

We re-analyzed our forecasting models for physical activity (Table 4 in the manuscript) by adding all these demographic factors as control variables. This had negligible effects on our results. The only change in significance was that the VoPE score no longer was a significant predictor of mild physical exercise measured with the GSLTPAQ, β = 0.13, SE = 0.09, p = 0.138, 95% CI [-0.04; 0.30]. However, the link was already weak in the original analysis, which is a plausible finding (see your comment below). We now report these additional robustness checks in the manuscript.

"Adjusting for sociodemographic variables (age, gender, income, education, employment, and subjective socio-economic status) had negligible effects on our results. The only change in terms of significance was that the VoPE score no longer predicted mild physical exercise measured with the GSLTPAQ, β = 0.13, SE = 0.09, p = 0.138, 95% CI [−0.04; 0.30]."

3- The network analyses are fascinating and I was glad to learn more about that. However, I also think that a "basic" correlation table could also be added (maybe in the Supplementary material) to facilitate comparisons with existing literature (e.g., with the Physical Effort Scale, Cheval et al., 2023)?

This is an excellent suggestion. We added the correlation table in OSF and point to it in the manuscript.

"To facilitate comparisons with the existing literature, we additionally determined bivariate correlations between all scales and provide them on OSF."

Discussion

I have really enjoyed reading the Discussion section and I have a few comments/questions.

1- Authors state that "Boredom is a value-based experience (Martarelli et al., 2023), which makes it plausible that people get bored by sports and exercise because they ascribe little value to the physical effort involved". From my naive understanding, boredom constitutes an unpleasant experience in which people perceive time as passing slowly, and feeling restless, trapped unchallenged, and perceiving the situation as meaninglessness (see https://doi.org/10.1177/10888683211010297 for a review). Following this definition, can allocating a high effort to a certain task lead to boredom (i.e., little challenge for example)? From my opinion, assigning a low value to effort may rather overlap with effort-based amotivation (e.g., 10.1111/j.1559-1816.1999.tb00122.x). Maybe that a stronger rationale or a more nuanced explanation could be proposed here to discuss the link between the VoPE score and boredom in sports.

That's an important comment, thanks for clarifying this point! Indeed, we should have been more precise here. Generally speaking, if individuals ascribe low levels of value to physical effort, they should indeed find physically effortful activities boring. Conversely, if individuals ascribe high levels of value to physical effort, they should find activities boring that require little physical effort. So you are right – people can be bored even when they ascribe high value to effort, and our statement only holds under the additional assumption that sports and exercise activities are typically effortful. Therefore, we revised the statement as follows:

"Boredom is a value-based experience (Martarelli et al., 2023), and most theories of boredom assume that low levels of value give rise to boredom (for an overview, see Bieleke et al., 2024). This implies that people should get bored by effortful sports and exercise activities when they ascribe little value to the physical effort that is commonly involved in these activities."

Regarding amotivation, we actually can check your intuition using our data in Study 3. We found a small negative correlation between amotivation and value of effort ($r = -0.18$) and a more substantive correlation between boredom and value of effort ($r = -0.46$). The psychometric network analysis further suggests a robust and replicable link only between boredom and VoPE but not between amotivation and VoPE (see Figure 3c). Taken together, our data supports your intuition that amotivation is linked to the value of physical effort. However, at least in our data, it seems that the link to boredom is more pronounced. We feel that this discussion is rather specific and might distract readers from the essential points of the discussion. Therefore, we did not include it in the manuscript.

2- Authors highlight that "the VoPE scale explained little variance of mild forms of exercise and activities, which might be due to its focus on sports rather than on physical activity." This result is not only consistent with the findings regarding the PES (Cheval et al., 2023), but it also suggests that beyond a certain level of perceived effort (e.g., standing, walking), the valuation of physical effort may not be as relevant as to predict more vigorous activities. Maybe that finer-grained measures (e.g., accelerometry, ecological momentary assessment) are in fact needed to better investigate this question and overcome social desirability biases that might blur these

associations. Overall, I think that the self-reported nature of the physical activity data should be mentioned as a limitation from the current findings.

We agree that other measures of physical activity might complement the present approach nicely. We now address this in the limitation section of our discussion as follows:

"In addition, VoPE scores might be correlated with indicators of sports and exercise behavior or physical activity that are measured by devices such as sports and fitness watches. For instance, collecting accelerometer data and conducting ecological momentary assessments could provide valuable insights into the weak link between VoPE scores and self-reported mild forms of exercise and physical activity observed in this study."

3- Finally, I think that readers (be they researchers or practitioners) may enjoy finding the full version of the scale somewhere, with a manual providing all the instructions and scoring procedure. Maybe that authors could consider adding this manual as a supplementary material?

This is a great idea! We added the scale as separate file to OSF.

Again, I want to congratulate authors for their hard work and I wish them the best of luck with their ongoing and future projects.

Thank you very much for your positive evaluation and your valuable feedback!

Silvio Maltagliati

Review by Erik Bijleveld, 31 May 2024 15:00

I have read and understood the manuscript titled "Development and validation of the Value of Physical Effort (VoPE) scale", which was submitted to PCI Health & Movement Sciences. The paper reports three studies that assess the psychometric properties of a new 4-item scale to measure the value of physical effort. I should note that I am not an expert on psychometrics and scale development, so I cannot assess whether the authors empirical work meets all relevant methodological standards. That said, I can see that this new scale could be useful to the research community, and -- to the best of my assessment -- the empirical work appears thorough. I also appreciate the completeness and clarity of the OSF repository. I have no objections to this work being published, but I do have two comments:

1) Physical exercise is associated with many different potential rewards. For some people, exercise may come with social rewards (e.g., having pleasant social interactions during or after the exercise; the pleasure of scoring a goal or a point) or mastery-related rewards (i.e., the pleasant experience of getting better at something). For others, exercise is experienced as progress towards a valued goal (e.g., related to health, fitness, appearance). Still others reward themselves after exercise (e.g., by allowing themselves to use their favorite bath salt). I agree that the prospects of obtaining tangible rewards may often be neglible (p. 2), but my point is that non-tangible rewards may of course have substantial subjective value as well. Against this background, there seem to be two possibilities:

-- A) Physical effort is aversive, but it can readily be offset by all the rewards mentioned above. This is why people may choose to engage in exercise, and this is why they may enjoy activities that involve exercise (e.g., gym visits). However, the effort itself remains aversive. People engage in sports and exercise despite the physical effort.

-- B) Physical effort is, for some people, pleasant in itself. This is indeed akin to what Cacioppo argued for mental effort: Some people like to think hard. By extension, some people may just like the exercise hard. People engage in sports and exercise because of the physical effort.

In the introduction, the authors seem to double down on mechanism (B). But can they really exclude A? In my view, these two perspectives are hard to disentangle. More concretely, if someone scores high on the VoPE, can the authors be really sure this person enjoys effort per se? Or do they just enjoy going to the gym, for example, because this allows them to also be around their friends? I would welcome a more nuanced perspective about this issue: when people score high on the VoPE, what exactly do they enjoy or value, and how do the authors know this?

Note: I put some thought in this issue myself, and -- for mental effort -- my co-authors and I came to the opposite conclusion (written up in David et al., 2022, which the authors cite; we will post an update in June 2024). But as mentioned, I think the two options are hard to dissociate definitively, and I think there is ample room for different intepretations, especially since the authors' rationale specifically focuses on physical effort.

Thank you very much for sharing your insights on this issue! We agree that this is an extremely delicate question, and we readily admit that we do not have a definite answer either. In our view, effort has an obvious instrumental value that comes from both tangible and intangible rewards, and many people engage in sports and exercise because of these

rewards. If effort is repeatedly coupled with these rewards, it might be valued on its own. This is a core assumption of the learned industriousness theory and one of the main reasons for us to believe that people might ascribe value to physical effort itself.

To make this reasoning clear, we revised our manuscript in both the introduction and in the discussion. In the introduction, we now emphasize the role of tangible and non-tangible rewards and put the value of effort more into perspective. In the discussion, we expand on the difficulties of pinpointing the exact source of the value people ascribe to effort (i.e., your options A and B). We think that our description is now more balanced and we hope you agree!

Introduction:

"Every year, millions of people invest substantial resources (e.g., money, time, risk of injury) to participate in exhausting endurance events with negligible prospects of obtaining tangible rewards in return (Maxcy et al., 2019). And while there surely are non-tangible rewards to justify the effort (e.g., better fitness, social engagement), these people still seem to attach levels of value to the experience of physical effort that are difficult to reconcile with a purely instrumental perspective on effort. Put differently, they might engage in sports and exercise not despite the physical effort it requires but because of it (Wolff, Hirsch, et al., 2021), suggesting that they value effort to some degree."

Discussion:

"Finally, our results suggest that viewing effort solely as costly and explaining its expenditure in purely instrumental terms might be an overly narrow perspective. Our results suggest that some people value the physical effort involved in sports and exercise to some degree, possibly because effort is commonly followed by rewards, making it a secondary reinforcer (Eisenberger, 1992). This is in line with research showing that consistently rewarding effort can lead people to choose more effortful activities (e.g., Clay et al., 2022), although valuing effort for its own sake still remains the exception rather than the rule (e.g., David et al., 2022; Wolff et al., 2023). What contributes to the costs and value people associate with effort thus remains an open question, and the VoPE scale hopefully assists researchers in disentangling these contributions further."

2) I think it would be worthwhile to know a bit more about how the authors went from 10 items (which they apparently deemed relevant a priori) to 4 items. Specifically, by removing these 6 items, does the measure still capture the full breadth of the construct? I am not saying it doesn't, but the paper would be stronger if the authors more clearly justify the choices to keep vs. delete items.

This is an important point, thanks for bringing it up! We expanded our explanation of how we constructed the scale, and provide more details on our rationale for shortening the scale. We hope that this addresses your questions.

"In Study 1, the VoPE scale was administered as part of a general survey. To construct the VoPE scale, we adapted items from the Need for Cognition Scale (NfCS; Cacioppo et al., 1984) — a well-established instrument that measures "an individual's tendency to engage in and enjoy effortful cognitive endeavors" (p. 306). We jointly reviewed, discussed, and ranked the 34 NfCS items regarding their suitability for adoption to the domain of sports and exercise. We ended up with a set of ten items that provided the best fit in terms of adoptability (5 positively and 5

negatively framed items) and administered them in all three studies. We translated the VoPE scale from German (Study 1) to English (Studies 2 and 3) with the assistance of AI technology (www.deepl.com; see Table 1 for both language versions). After data collection, we reduced the items to the four items that make up the final VoPE scale (see Table 2).

First, we excluded the five negatively framed items because they formed a separate community in network analyses, likely reflecting a common method artifact that plagues mixed framing scales (Lindwall et al., 2012). The NfCS has also been criticized for introducing artificial factors through its negatively framed items (Zhang et al., 2016), which we aimed to avoid. Second, we removed one of the positively framed item which, in retrospection, referred to an artificial situation that might elicit ambivalent responses: “When I know that a sports activity is going to involve physical effort, I look forward to it, even if I do not know what exactly it involves.” Despite these exclusions, all ten items are available in the published dataset on OSF for transparency. Participants indicated their agreement with each item on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree), and their responses were averaged into an overall VoPE score.”