



Rating of Perceived Effort: Methodological Concerns and Future Directions

Israel Halperin^{1,2} · Aviv Emanuel³

© Springer Nature Switzerland AG 2019

Abstract

Rating of perceived effort (RPE) scales are the most frequently used single-item scales in exercise science. They offer an easy and useful way to monitor and prescribe exercise intensity. However, RPE scales suffer from methodological limitations stemming from multiple perceived effort definitions and measurement strategies. In the present review, we attend these issues by covering (1) two popular perceived effort definitions, (2) the terms included within these definitions and the reasons they can impede validity, (3) the problems associated with using different effort scales and instructions, and (4) measuring perceived effort from specific body parts and the body as a whole. We pose that the large number of interactions between definitions, scales, instructions and applications strategies, threatens measurement validity of RPE. We suggest two strategies to overcome these limitations: (1) to reinforce consistency by narrowing the number of definitions of perceived effort, the number of terms included within them, and the number of scales and instructions used. (2) Rather than measuring solely RPE as commonly done, exercise sciences will benefit from incorporating other single-item scales that measure affect, fatigue and discomfort, among others. By following these two recommendations, we expect the field will increase measurement validity and become more comprehensive.

Key Points

In exercise science there are numerous definitions and measurement instruments of perceived effort. This leads to confusion and hinders measurement validity.

To solve this problem, there is a need to clarify and narrow the number of definitions and terms included within the perceived effort construct.

There is a need to measure other perceptions and emotions more frequently to improve the field's precision and comprehensiveness.

1 Introduction

Ratings of perceived effort (RPE) are the most used single-item scales in exercise and sport sciences.¹ RPE scales have a range of benefits, including relationships with physiological and performance measures within and between training sessions [1–5]. Hence, RPE scales can serve as a surrogate measure of physiological indices, and assist in quantifying exercise intensity and load. RPE scales can also be used to regulate exercise intensity [6–9]. By pre-setting specific RPE targets, and instructing people to remain within their boundaries during exercise, people can effectively reach or avoid certain intensities. These benefits of RPE scales persist across different populations and exercise modalities [5, 6, 8, 9].

While the advantages of RPE scales are clear, they suffer from certain methodological limitations, most of which stem from multiple definitions of perceived effort (PE). Indeed, debates about PE definitions and its neurophysiological underpinning are still ongoing [10–13]. The literature is flooded with definitions that fall under the same umbrella of PE, although

✉ Israel Halperin
ihalperin@tauex.tau.ac.il

¹ School of Public Health, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel

² Sylvan Adams Sports Institute, Tel Aviv University, Tel-Aviv, Israel

³ School of Psychological Sciences, Tel-Aviv University, Tel-Aviv, Israel

¹ Despite the call to distinguish between effort and exertion [8], in this article we use the terms synonymously. We address this distinction in Sect. 2.

they may be addressing different perceptions [10, 11, 13–16]. Moreover, a growing number of RPE scales and instructions are being used by different laboratories [2, 4, 17–19]. RPE scales are also applied to rate effort in specific body parts or in the body as a whole [20–23]. The number of interactions between definitions, scales, instructions and administration strategies, all attempting to capture the same perception, are exceptionally large. This poses a threat to measurement validity of RPE scales. By measurement validity we refer to the degree to which the observed scores (RPE) meaningfully capture the ideas contained in the corresponding concepts (PE construct) [24].

Here we build and expand on recent RPE articles [10, 12, 15], and cover (1) common PE definitions, (2) how the terms included within these definitions can hinder the scales' validity, (3) the problems that can arise from using different effort scales and instructions, and (4) possible shortcomings of measuring PE in specific body parts and in the whole body. We conclude with recommendations on how to improve PE measurements, and discuss the benefits of incorporating measurements of other constructs more often.

2 Definitions

We present two common definitions of PE.² The differences between them partly stem from different perspectives on PE's neurophysiological underpinnings. Since this article focuses on methodological aspects of PE measurement, we only briefly discuss these mechanistic underpinnings as they pertain to PE definitions. We refer the reader to other articles covering the neurophysiological basis of PE [10, 12, 26].

A prominent figure in the RPE literature is Gunnar Borg. Borg developed several scales, including the 6–20 RPE Scale [27], the Borg category-ratio (CR) 10 Scale [28], and Borg centiMax Scale (CR100) [29]. In his book, Borg [30] viewed effort as physiological and psychological signals integrated into a single “Gestalt” perception during physical activity (p. 3). Borg defined PE as “... the feeling of how heavy and strenuous a physical task is” (p. 8). The following is also presented to subjects when instructed on how to rate the scale, “The perception of exertion depends mainly on the strain and fatigue in your muscles and on your feeling of breathlessness or aches in the chest” (p. 47). Other figures in the RPE field, Noble and Robertson [31, 32], followed Borg's general view of PE and defined it as “The subjective intensity of effort, strain, discomfort and/or fatigue that is felt during exercise”. To date, the work by Borg, Nobel and Robertson on PE has

been cited thousands of times. Other notable researchers have embraced the viewpoint that an integration of inputs from several bodily systems produces PE [1, 13, 33, 34].

A different perspective of PE was put forth by Marcora [11, 35]. According to Marcora, PE is solely a product of corollary discharge occurring in the brain. This perspective leans on the theory of corollary discharge, which postulates that when the motor cortex delivers signals to activate muscles, copies of these signals are simultaneously delivered to the sensory cortex. These copies are then interpreted as PE [36, 37]. In contrast to the perspective advocated by Borg and others, Marcora views PE to be independent of afferent feedback. For example, heart transplantation patients reported normal RPE values despite their denervated organ [38]. Marcora defines PE as the “Conscious sensation of how hard, heavy, and strenuous a physical task is” [35]. While this definition resembles Borg's, it is important to note that the terms fatigue and discomfort are absent from it, and from the subsequent instructions of the scale. This is because they represent different constructs that have distinct neurophysiological pathways [11]. Marcora and colleagues use Borg's RPE or CR-10 scale, and instruct participants to report how hard they are driving their working limbs [39], how heavily they are breathing, and the overall sensation of how strenuous the exercise is [40]. The precise instructions are modified to fit the exercise modality and research question.

Abbyss and colleagues [10] recently proposed a distinction between effort and exertion, two terms that are commonly used synonymously. They defined effort as the “The amount of mental or physical energy being given to a task”. This emphasizes effort as a process in which people invest a certain resource. Exertion was defined using Borg's definition “The feeling of how heavy and strenuous a physical task is”. This emphasizes exertion as a process associated with the physical and physiological stress induced by the exercise. In a sense, this distinction bridges the two viewpoints: effort is more aligned with Marcora's and the corollary discharge perspective, whereas exertion is more aligned with Borg's and the afferent feedback perspective. The requirement to narrow the definition of effort/exertion and to differentiate between perceptual experiences is generally agreed upon [12, 41]. However, it is debatable if effort and exertion should be distinguished precisely, because the terms are used synonymously. This fact could lead to the opposite intended effect and increase confusion in the RPE field. We offer possible solutions to these issues in Sect. 6.

3 Differentiating Perceptions

The perspective adopted by Borg and others on the underpinnings of PE likely led to the inclusion of the following terms within its definition: fatigue, discomfort, and heaviness. This

² For clarity purposes, we note that the constructs of effort and PE can be viewed as distinct. For example, Silvestrini and Gendola [25] distinguish between effort and perception of effort, stating that the two are often dissociated from another.

is problematic, as there is evidence that people can differentiate PE from such perceptions [12]. We now turn to discuss these terms in relation to PE. Importantly, we only discuss these three constructs, because they are included in Borg's or Marcora's definitions of PE. From a broader perspective, other constructs such as exercise induced pain, enjoyment, or even hunger can be of interest to the study of PE. However, such constructs require a thorough discussion in and of themselves and are beyond the scope of the present review.

3.1 Fatigue

Despite that some PE definitions include the term fatigue [2, 31], Borg [42] wrote that "the concept of fatigue should be distinguished from the concept of perceived exertion even though these two concepts have very much in common". Recently, in line with Borg's distinction, Micklewright et al. [43] developed a new single-item scale called rating of fatigue (ROF). The authors defined perceived fatigue as "... a feeling of diminishing capacity to cope with physical or mental stressors, either imagined or real". ROF and Borg's RPE were strongly correlated during an incremental physical task. Yet, people rated RPE as zero at the point of exercise termination, whereas ROF was still relatively high. This finding highlights the discriminant validity of fatigue during recovery. The strong relationship between the two scales ($r=0.99$) could partly stem from the fact that Borg's RPE instructions include the term fatigue. It is possible that when required to rate the two scales during exercise, participants rate them in a similar manner.

During incremental physical activities, fatigue and effort are expected to strongly correlate as both tend to increase simultaneously. This is because once muscular fatigue accumulates, it becomes more difficult to produce similar levels of force, and producing such force will require greater effort [44]. In contrast, short, maximal effort physical tasks completed when people are fresh, will likely lead to different perceptions. To illustrate, a single, three-second elbow flexion maximal isometric contraction should not lead to a meaningful perception of fatigue, but to maximal PE [45]. Also, in cases in which exercise difficulty decreases over time, different and even inverse relationships between perceptions may occur. For example, perception of fatigue may increase, while PE decreases. This proposal is supported by studies observing different RPE values in protocols, where exercise difficulty increased or decreased over time, although completing similar work in both conditions [46, 47]. Hence, exercise configuration can interact with the ratings of PE and fatigue.

3.2 Discomfort

With proper instructions, people can discriminate between PE and discomfort during exercise [48, 49]. To further

establish this division, Steele and colleagues recently developed modified CR-10 Borg scales that measure effort³ and discomfort [19]. Discomfort was defined as "the physiological and unpleasant sensations associated with exercise" [19]. The authors observed that participants can differentiate between effort and discomfort, and that exercise duration has a strong impact on perception of discomfort [50, 51]. A longer set, composed of more repetitions and/or time of muscles being under tension, led to higher perception of discomfort compared to effort, when the sets were taken to momentary failure [50, 51]. To illustrate, Stuart et al. [51] had participants complete dynamic back extensions to momentary failure while lifting either lighter (50%) or heavier (80%) loads that were calculated based on their maximal voluntary contractions. Whereas both conditions led to maximal RPE values, perception of discomfort did not reach maximal values in either condition, and was greater in the low force condition.

3.3 Heavy

Both Borg and Marcora include the term heavy in their definition of effort. This term is problematic, especially as it pertains to resistance training [52, 53]. Lifting a heavier load, absent of information on the number of repetitions completed relative to the repetition maximum (RM), is not necessarily indicative of effort [52, 53]. For example, it can be argued that completing one repetition of a 5RM load (1/5 RM = 20% of maximum) is easier compared to completing nine repetitions of a lighter 10RM load (9/10 RM = 90% of maximum). When completing an exercise to momentary failure, lifting lighter loads usually leads to higher RPE values compared to heavier loads [54, 55].⁴ Hence, the total number of repetitions completed relative to maximum is a more comprehensive way to estimate effort. The problem with including the term "heavy" as part the PE definition can be illustrated with the OMNI scales in resistance exercise [56]. Attached to the scale is a figure of a person holding a barbell above his head which represents PE. The increments from 0 to 10 are associated with heavier loads on the bar the person is holding. This figure could mislead people into rating the perceived heaviness of the object they are lifting, rather than their PE.

³ Steele and colleagues defined perceived effort as the demands of the activity or exercise relative to the current capacity to meet those demands. Note that the numerator and denominator were erroneously reversed in the paper (Steele 2019, personal communication, permission was granted to mention this information).

⁴ The higher RPE reported in such studies may actually represent people's perception of discomfort rather than effort.

4 Scales and Instructions

It is common for scientists to use altered versions of the Borg scales. In his book, Borg [30] highlights some errors and possible misuse of his scale: copying alterations, changes in format, verbal anchors, scale design, and shortening or modifying the instructions (p. 15). For example, the Borg CR 10 Scale includes a dot (●) below the digit 10. Whereas 10 is defined as “extremely strong” (explained as the strongest effort one has ever experienced), the dot is defined as the “absolute maximum”. The purpose of the dot is to avoid a ceiling effect by allowing people to rate higher values than 10 in case they should perceive such effort. However, in some studies, the dot is excluded from the scale [48, 57, 58]. The Borg CR 10 Scale also went through minor modifications over the years [59], yet the exact version used in studies is not always reported. It is unclear if—and to what extent—altering the scales or using different versions of the same scale affects ratings. Other effort scales are also widely used. For example, nine different OMNI scales were developed for distinct exercises and populations [2]. While many scales were validated against Borg’s RPE, the purpose of having multiple scales attempting to capture the same construct is unclear [1].

Distinct instructions and explanations are provided alongside RPE scales [2, 14, 48, 60]. Multiple or varying instructions and explanations can hinder measurement validity. Consider the importance of explaining what the upper limits of scales stand for (also known as anchoring, e.g., 10 is “maximal effort”). Without explicitly associating this number and term with a relevant event that can occur in the to-be-completed exercises, its actual meaning can be interpreted in various ways. For example, during a 10-km running time trial, when asked to rate RPE, some may interpret the upper limit in relation to the fastest they can run at a given point in time, whereas others as the fastest they can run at view of task completion. Unless instructed otherwise, both are legitimate ways of interpretation, and both could yield different ratings. In resistance training studies, the scales’ upper limit is normally associated with the terms “momentary failure” and “repetition maximum”. But inconsistencies in the way these two terms are defined and explained could lead people to rate their efforts differently [61, 62]. Some understand RM as the inability to complete another repetition, leading them to cease the set without trying to complete the next one. Others can understand RM as reaching the actual point during a repetition that they are unable to complete. These subtle differences—which are at least partly a function of the provided instructions—have the potential to influence subjects’ ratings.

5 Part vs. Whole-Body RPE

Scientists developed a division of RPE measurement between perceptions of effort occurring in (1) a specific muscle group (termed peripheral, local or differentiated RPE), (2) the cardiorespiratory system (termed central or respiratory-metabolic RPE), and (3) the body as a whole (termed undifferentiated RPE) [20–23, 63]. When asked to rate two or three of these RPE measurements, subjects rate specific muscle group effort the highest [22, 23]. While this approach to measuring effort is thought to enhance precision [23], it raises questions regarding the nature of the PE construct and its measurement.

If PE resembles “a social psychophysiological phenomenon” [64] and is presented as a “Gestalt” occurrence [65], then the meaning of its attribution to a specific body part is ambiguous. For example, one can rate how excited she feels, yet might be confused if also asked to rate how excited she feels in her quadriceps. Alternatively, if PE resembles, or is presented as a sensation that is attributed to a specific body part, then the meaning of its general “Gestalt” attribution is ambiguous. For example, when suffering from a toe injury, one can rate the amount of pain she feels in her toe, yet might be confused if also asked to rate the amount of pain she feels in general. Although people rate part and whole-body RPE, it is unclear if the reported values represent the same construct.

In view of the above, it is debatable if applying the same PE construct to measure both part and whole-body RPE is a sound administration strategy. The provided examples also emphasize the need to measure other perceptions in addition to, or instead of, RPE. For instance, it may be more suitable to collect whole-body RPE and perception of discomfort from specific body parts within a given study, rather than RPE from both. By doing so, we anticipate that it will be easier for people to grasp the notion of PE, which will thereby increase measurement validity. If one chooses to embrace part and whole-body viewpoints of PE, then it may be better to avoid simultaneously measuring RPE from both as commonly done. By asking people to rate both, there is a greater risk of receiving ratings to questions not posed (e.g., one rating addressing PE or affect and the other addressing pain or discomfort). However, such response inconsistencies are less likely to occur if one or the other administration strategies are exclusively used in the same study. Taken together, there is a need to evaluate if PE should be treated and administered as part or as a whole. Until this issue is resolved, we suggest considering the points put forth in this section when measuring and interpreting part and whole-body RPE values.

6 Solutions and Future Directions

Our proposed solutions are composed of two complementary approaches. The first is to narrow the scope of PE measurements by committing to fewer definitions, terms, and instructions. The second is to widen the scope of exercise and sports science by including measurements of other perceptions more frequently. We provide examples of useful single-item scales and when they can be used. We focus on single-item scales due to their similarities to RPE and their ease of administration during exercise. It is important to keep in mind that the scales to be introduced in the following sections may also suffer from methodological shortcomings, but these go beyond the scope of this article. Rather than following step-by-step solutions, our goals are to stimulate thought, explore new possibilities, and encourage new research directions.

6.1 Narrowing the Scope

A simple solution to the multiple definitions problem is to reduce the number of terms included within the PE definition. As illustrated above, people can differentiate between effort and other perceptions included in the popular PE definitions (fatigue, discomfort and heaviness). Of the existing prevalent definitions of PE, Marcora's is preferable, as it is simpler and includes fewer terms. But we propose narrowing this definition further. As mentioned previously, the term "heavy", which is part of Marcora's definition, can be misleading. Additionally, the added value of the following phrase "... the conscious sensation of ..." within the definition is debatable. Since people are asked to explicitly report their perceptions, a conscious state is mandatory.

The definition of effort offered by Abbiss and colleagues [10], i.e., "The amount of mental or physical energy being given to a task", is a step in the right direction. Psychologists Gendola and Wright define effort as "Mobilization of resources to carry out instrumental behavior" [66]. We embrace the characterization of effort as a process of investing certain resources, and view PE as the perception of their investment, irrespective of the actual relationship between the two. We thus propose that PE can be defined as: "The process of investing a given amount of one's perceived physical or mental resources out of the perceived maximum to perform a specific task". This can be simplified to perceived invested resources divided by perceived maximal resources. The addition of a reference point—the perceived maximum—can translate this definition into actionable measurement units. For example, the perceived maximum to complete a given task

corresponds to the upper range of effort scales (e.g., 10 in a 0–10 scale). However, the usefulness of this definition remains to be established.

Qualitative methods can assist coming to terms with the PE definitions. The Delphi technique is an iterative multi-stage process designed to transform experts' opinion into group consensus via questionnaires and interviews [67]. A Delphi study incorporating psychophysicists, exercise scientists and psychologists focused on the PE definitions can contribute to this goal. Due to the multiple views on PE, reaching a consensus among experts may be challenging. Alternatively, reaching an agreement about specific study designs that will resolve in favor of one definition of PE is a feasible task. Another qualitative avenue includes semi-structured interviews and focus groups of non-experts. In these designs, people are first provided with explanations and definitions of constructs and measurement instruments, and then asked to report if the explanations match their understanding of them. Such studies are successfully conducted in pain sciences, leading to interesting and actionable insights about the mismatches between pain definitions and measurement scales, and the meanings patients attribute to them [68–70]. Using a similar approach to investigate people's understanding of PE should lead to interesting insights. These qualitative designs coupled with the on-going quantitative applied and basic research will hopefully lead to a deeper understanding of the PE construct, and to an agreed upon PE definition.

6.2 Widening the Scope

It is our impression that exercise and sport scientists, at times, measure RPE even when attempting to answer different questions. The relative absence of scales that measure other perceptions is speculated to stem from the establishment of PE as the single most recognized subjective variable in exercise and sport science since the late 1960s [71]. We echo other scientists' recommendations [41, 71], and pose that as a field, we should measure other perceptual experiences more often. This will allow us to be more certain that we are receiving answers to the questions being asked. This recommendation is also in line with Abbiss and colleagues' [10] call to distinguish between effort and exertion, and treat them as separate perceptual experiences. Yet, as argued previously, it may be better to treat these two terms synonymously and generate other terms that are clearly distinct from effort/exertion to avoid misconceptions.

To illustrate the overreliance on PE, consider the following: some studies measured RPE post exercise as a tool to assess the abatement of PE [72, 73]. Part of the instructions in a study by Robertson and colleagues [72] included "... perception of effort is defined as the intensity

of effort, stress, discomfort, and/or fatigue that you feel during recovery from dynamic exercise". Three minutes post exercise, the reported RPE values were as high as 18 on the Borg's 6–20 scale [72]. While PE occurs during exercise, or can be reported as a memory of the exercise (e.g., session RPE [3]), it should not be experienced during rest. The high RPE scores recorded three minutes post exercise were likely perception of fatigue, stress or discomfort, but not effort. In contrast to RPE, the new ROF scale developed by Micklewright and colleagues [43] might be better suited to answer a recovery related question during rest.

Affective valence is a construct of potential interest. It refers to the positive and negative qualities of emotions and moods [74, 75]. During exercise, heterogeneous inputs, including fatigue, pain, reward and pride, are processed and acted upon [76, 77]. Since they reside on different continuums, they need to be converted to a common scale for them to be assessed and compared. This process, reflected in affective valence, allows for a decision to take place in real time (e.g., reduce running speed) [71, 76, 77]. Affective valence is thought to act as a 'common currency' in the tradeoffs between opposing motivations regulating exercise [71, 76, 77]. Affective valence is commonly measured by the Feeling Scale (FS): an 11-point bipolar scale ranging from +5 (very good) to –5 (very bad) [64]. Exercise psychologists have extensively studied the FS for the past 30 years, mostly among sedentary populations [74, 78]. A key finding of these studies is that people reporting lower FS ratings during exercise are less likely to continue exercising over time [74, 79, 80]. This finding is corroborated with the fact that people prefer to avoid unpleasant experiences [74, 76]. To date, however, the FS remains relatively unexplored by exercise and sport scientists. Measuring the perception of discomfort is another option. Discomfort is in many ways similar to negative affect measured by the FS. However, the Discomfort Scale [19], which ranges from 0 (no discomfort) to 10 (maximal discomfort), places greater emphasis on the negative affect compared to FS. When greater resolution is sought after, and the exercise is expected to cause considerable negative affect, the Discomfort Scale may be a better alternative.

From an applied perspective, lower FS scores reported by athletes during a physical task that is completed on a regular basis could indicate the beginning of a burnout or insufficient recovery. Another example is when two athletes are exercising at a similar RPE, but one rates the exercise as pleasurable while the other as not. This insight could lead coaches to adopt different training strategies. The observation that lifting heavier loads to task failure leads to lower perception of discomfort compared to lighter loads is also of practical value [51]. For example, lifting heavier loads may be a preferred option when a training program includes sets

completed to task failure. This is especially the case when the goal is to encourage exercise adherence, mostly among untrained populations.

Guiding participants to complete a number of repetitions within a particular distance from their estimated RM is a training strategy growing in popularity [4, 60, 81, 82]. This strategy allows prescribing and monitoring resistance training by achieving or avoiding certain intensity zones. For example, by following the estimated Repetitions to Failure Scale [60], subjects can be asked to terminate a set two repetitions away from RM. The similar Repetitions in Reserve (RIR) Scale translates the former into an RPE rating which ranges from 1 to 10 [4]. For example, subjects can be asked to reach an RPE score of 8, indicating that only two repetitions are left in reserve (i.e., 8RPE equals to 2RIR). While these scales can be viewed as versions of RPE scales, they are included in this section for two reasons. First, they offer a unique aspect that is absent from most scales: a perfect match between the measurement unit and the construct of interest: the number of repetitions left to task-failure. This makes rating these scales straightforward, which reduces the likelihood of misunderstandings, and increases measurement validity. Second, perception of repetitions left in reserve can be influenced by a wide array of perceptions, including, but not limited to, PE.

RIR scales are effective in monitoring training sessions, as they are correlated with bar velocity across exercises [83, 84], and are reasonably accurate at capturing RM [60, 81, 82]. They also allow for effective exercise regulation as sets, repetitions, and loads can be adjusted in view of one's perception of RIR. For example, two recent studies of 8 and 12 weeks' duration found that a RIR based method produced greater strength gains compared to percentage-based training among trained participants [6, 7]. Despite their clear benefits, RIR scores were found to be less accurate in untrained subjects [85] (although this finding is not consistent [81]), and with sets including high number of repetitions [81, 82]. Hence, RIR scales seem to be better suited for participants with resistance training background, performing relatively low number of repetitions.

7 Putting Scales into Practice

We conclude with practical recommendations on how single-item scales can be used in practice to reduce within and between laboratory variability, as well as confusion among participants and practitioners.

- It is vital to carefully consider which specific construct would be most suited to answer a given question, and subsequently, which scale is best suited for its measurement. These decisions should be justified before data collection, and reported in the subsequent manuscript.

- All laboratory members involved in data collection procedures should be aligned as to which scale, version of scale, instructions and explanations are to be used. Written scripts can assist this process. By making sure that this process is adhered to, within laboratory reliability and validity aspects concerning the scales are ensured.
- Authors of manuscripts should report which scale and version were used, coupled with the instructions and anchoring procedures. This information can be presented within the article, or as a supplementary document. This will allow for a richer interpretation of the results, replication attempts, meta-analytic procedures, and generally, increase between laboratories reliability and validity of scale measurement.
- Whenever possible, modifying validated scales and their instructions should be avoided as even small changes can impact scales' ratings. In cases in which scales development or modifications are justified, it is important to use construct, concurrent and discriminant validation procedures to ensure the scale's validity.
- An explicit, written question that subjects are expected to answer should be added to scales, e.g., "How effortful is the task?" for measuring PE. As single-item scales are used simultaneously (e.g., ROF and RPE), the addition of this question will remind the physically active subjects what exactly they are required to rate.

8 Conclusion

Despite the benefits of RPE scales, they suffer from limitations that warrant attention and action. These limitations include multiple PE definitions, scales, instructions, and the division of PE to body parts, all of which risk measurement validity. To overcome these problems, we proposed two conceptual solutions. The first is to narrow the number of PE definitions, the terms included within the PE definitions, the implemented scales, and instructions. The second is to incorporate other single-item scales that measure other perceptions more often. Following these recommendations will enhance measurement precision and expand measurement breadth.

Compliance with Ethical Standards

Funding No sources of funding were used for the preparation of this manuscript.

Conflict of interest Israel Halperin and Aviv Emanuel have no conflicts of interest that are directly relevant to the content of this manuscript.

References

1. Faulkner J, Eston RG. Perceived exertion research in the 21st century: developments, reflections and questions for the future. *J Exerc Sci Fit.* 2008;6(1):1–14.
2. Haile L, Gallagher M, Robertson RJ. Perceived exertion laboratory manual. New York: Springer; 2016.
3. Haddad M, Stylianides G, Djaoui L, Dellal A, Chamari K. Session-RPE method for training load monitoring: validity, ecological usefulness, and influencing factors. *Front Neurosci.* 2017;11:612.
4. Zourdos MC, Klemp A, Dolan C, Quiles JM, Schau KA, Jo E, et al. Novel resistance training-specific rating of perceived exertion scale measuring repetitions in reserve. *J Strength Cond Res.* 2016;30(1):267–75.
5. Chen MJ, Fan X, Moe ST. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *J Sports Sci.* 2002;20(11):873–99.
6. Helms ER, Byrnes RK, Cooke DM, Haischer MH, Carzoli JP, Johnson TK, et al. RPE vs. Percentage 1RM loading in periodized programs matched for sets and repetitions. *Front Physiol.* 2018;9:247.
7. Graham T, Cleather DJ. Autoregulation by "Repetitions in Reserve" leads to greater improvements in strength over a 12-week training program than fixed loading. *J Strength Cond Res.* 2019. <https://doi.org/10.1519/jsc.0000000000003164> (**publish ahead of print**).
8. Parfitt G, Evans H, Eston R. Perceptually regulated training at RPE13 is pleasant and improves physical health. *Med Sci Sports Exerc.* 2012;44(8):1613–8.
9. Buskard ANL, Jacobs KA, Eltoukhy MM, Strand KL, Villanueva L, Desai PP, et al. optimal approach to load progressions during strength training in older adults. *Med Sci Sports Exerc.* 2019. <https://doi.org/10.1249/mss.0000000000002038> (**publish ahead of print**).
10. Abbiss CR, Peiffer JJ, Meeusen R, Skorski S. Role of ratings of perceived exertion during self-paced exercise: what are we actually measuring? *Sports Med.* 2015;45(9):1235–43.
11. Marcora S. Perception of effort during exercise is independent of afferent feedback from skeletal muscles, heart, and lungs. *J Appl Physiol.* 2009;106(6):2060–2.
12. Pageaux B. Perception of effort in exercise science: definition, measurement and perspectives. *Eur J Sport Sci.* 2016;16(8):885–94.
13. Noakes TD. Fatigue is a brain-derived emotion that regulates the exercise behavior to ensure the protection of whole body homeostasis. *Front Physiol.* 2012;3:82.
14. Hutchinson JC, Tenenbaum G. Perceived effort—can it be considered gestalt? *J Sport Exerc Psychol.* 2006;7(5):463–76.
15. Venhorst A, Micklewright D, Noakes TD. Perceived fatigability: utility of a three-dimensional dynamical systems framework to better understand the psychophysiological regulation of goal-directed exercise behaviour. *Sports Med.* 2018;48(11):2479–95.
16. Pereira G, Souza DMD, Reichert FF, Smirmaul BPC. Evolution of perceived exertion concepts and mechanisms: a literature review. *Rev Bras Cineantropom Desempenho Hum.* 2014;16(5):579–87.
17. Lampropoulou S, Nowicky AV. Evaluation of the numeric rating scale for perception of effort during isometric elbow flexion exercise. *Eur J Appl Physiol.* 2012;112(3):1167–75.
18. Swart J, Lindsay TR, Lambert MI, Brown JC, Noakes TD. Perceptual cues in the regulation of exercise performance—physical sensations of exercise and awareness of effort interact as separate cues. *Br J Sports Med.* 2012;46(1):42–8.
19. Steele J, Fisher J, McKinnon S, McKinnon P. Differentiation between perceived effort and discomfort during resistance training

- in older adults: reliability of trainee ratings of effort and discomfort, and reliability and validity of trainer ratings of trainee effort. *J J Train.* 2016;6(1):1–8.
20. Pandolf KB, Billings DS, Drolet LL, Pimental NA, Sawka MN. Differentiated ratings of perceived exertion and various physiological responses during prolonged upper and lower body exercise. *Eur J Appl Physiol.* 1984;53(1):5–11.
 21. Shephard RJ, Vandewalle H, Gil V, Bouhlef E, Monod H. Respiratory, muscular, and overall perceptions of effort: the influence of hypoxia and muscle mass. *Med Sci Sports Exerc.* 1992;24(5):556–67.
 22. Faulkner J, Eston R. Overall and peripheral ratings of perceived exertion during a graded exercise test to volitional exhaustion in individuals of high and low fitness. *Eur J Appl Physiol.* 2007;101(5):613–20.
 23. Robertson RJ, Gillespie RL, McCarthy J, Rose KD. Differentiated perceptions of exertion: part I. Mode of integration of regional signals. *Percept Mot Skills.* 1979;49(3):683–9.
 24. Adcock R, Collier D. Measurement validity: a shared standard for qualitative and quantitative research. *Am Polit Sci Rev.* 2001;95(3):529–46.
 25. Silvestrini N, Gendolla GHE. Affect and cognitive control: Insights from research on effort mobilization. *Int J Psychophysiol.* 2019;143:116–25.
 26. de Morree HM, Marcora SM. Psychobiology of perceived effort during physical tasks. *Handbook of biobehavioral approaches to self-regulation.* New York: Springer; 2015. p. 255–70.
 27. Borg G. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med.* 1970;2:92–8.
 28. Borg G. A category scale with ratio properties for intermodal and interindividual comparisons. In: Geissler HG, Petzold P, editors. *Psychophysical judgement and the process of perception.* Berlin: VEB Deutscher Verlag der Wissenschaften; 1982. p. 25–34.
 29. Borg E, Borg G. A comparison of AME and CR100 for scaling perceived exertion. *Acta Physiol (Oxf).* 2002;109(2):157–75.
 30. Borg G. Borg's perceived exertion and pain scales. Champaign: Human kinetics; 1998.
 31. Noble BJ. Perceived exertion. Champaign: Human kinetics; 1996. p. 115–7.
 32. Robertson RJ, Noble BJ. Perception of physical exertion: methods, mediators, and applications. *Exerc Sport Sci Rev.* 1997;25(1):407–52.
 33. Hampson DB, Gibson ASC, Lambert MI, Noakes TD. The influence of sensory cues on the perception of exertion during exercise and central regulation of exercise performance. *Sports Med.* 2001;31(13):935–52.
 34. Amann M, Venturelli M, Ives SJ, McDaniel J, Layec G, Rossman MJ, et al. Peripheral fatigue limits endurance exercise via a sensory feedback-mediated reduction in spinal motoneuronal output. *J Appl Physiol.* 2013;115(3):355–64.
 35. Marcora SM. Effort: perception of. In: Goldstein EB, editor. *Encyclopedia of perception.* Thousand Oaks: Sage; 2008. p. 380–3.
 36. Cafarelli E. Peripheral contributions to the perception of effort. *Med Sci Sports Exerc.* 1982;14(5):382–9.
 37. McCloskey D. Corollary discharges: motor commands and perception. In: Brooks VB, editor. *Handbook of physiology: section 1: the nervous system; vol II: motor control, part 2.* Washington: American Physiological Society; 1981. p. 1415–47.
 38. Braith RW, Wood CE, Limacher MC, Pollock ML, Lowenthal DT, Phillips MI, et al. Abnormal neuroendocrine responses during exercise in heart transplant recipients. *Circulation.* 1992;86(5):1453–63.
 39. Pageaux B, Marcora S, Lepers R. Prolonged mental exertion does not alter neuromuscular function of the knee extensors. *Med Sci Sports Exerc.* 2013;45(12):2254–64.
 40. Pageaux B, Lepers R, Dietz KC, Marcora SM. Response inhibition impairs subsequent self-paced endurance performance. *Eur J Appl Physiol.* 2014;114(5):1095–105.
 41. Venhorst A, Micklewright D, Noakes TD. Towards a three-dimensional framework of centrally regulated and goal-directed exercise behaviour: a narrative review. *Br J Sports Med.* 2018;52(15):957–66.
 42. Borg G. Psychophysical studies of effort and exertion: some historical, theoretical and empirical aspects. In: Borg G, Ottoson D, editors. *The perception of exertion in physical work.* London: Macmillan; 1986. p. 3–12.
 43. Micklewright D, Gibson ASC, Gladwell V, Al Salman A. Development and validity of the rating-of-fatigue scale. *Sports Med.* 2017;47(11):2375–93.
 44. De Morree HM, Klein C, Marcora SM. Perception of effort reflects central motor command during movement execution. *Psychophysiology.* 2012;49(9):1242–53.
 45. Smirmaul BdPC. Sense of effort and other unpleasant sensations during exercise: clarifying concepts and mechanisms. *Br J Sports Med.* 2012;46(5):308–11.
 46. Meckel Y, Zach S, Eliakim A, Sindiani M. The interval-training paradox: physiological responses vs. subjective rate of perceived exertion. *Physiol Behav.* 2018;196:144–9.
 47. Greer BK, Young PR, Thompson B, Rickert BJ, Moran MF. Impact of direction of unloading influence on template rate of perceived exertion. *J Strength Cond Res.* 2018;32(12):3398–404.
 48. Christian RJ, Bishop D, Girado O, Billaut F. The role of sense of effort on self-selected cycling power output. *Front Physiol.* 2014;5:115.
 49. Loenneke J, Thiebaud R, Fahs C, Rossow L, Abe T, Bemben M. Blood flow restriction: effects of cuff type on fatigue and perceptual responses to resistance exercise. *Acta Physiol.* 2014;101(2):158–66.
 50. Fisher JP, Steele J. Heavier and lighter load resistance training to momentary failure produce similar increases in strength with differing degrees of discomfort. *Muscle Nerve.* 2017;56(4):797–803.
 51. Stuart C, Steele J, Gentil P, Giessing J, Fisher JP. Fatigue and perceptual responses of heavier-and lighter-load isolated lumbar extension resistance exercise in males and females. *PeerJ.* 2018;6:e4523.
 52. Steele J. Intensity; in-ten-si-ty; noun. 1. Often used ambiguously within resistance training. 2. Is it time to drop the term altogether? *Br J Sports Med.* 2014;48(22):1586–8.
 53. Fisher J, Smith D. Attempting to better define “intensity” for muscular performance: is it all wasted effort? *Eur J Appl Physiol.* 2012;112(2):4183–5.
 54. Pritchett RC, Green JM, Wickwire PJ, Kovacs M. Acute and session RPE responses during resistance training: Bouts to failure at 60% and 90% of 1RM. *S Afr J Sports Med.* 2009;21(1):23–6.
 55. Shimano T, Kraemer WJ, Spiering BA, Volek JS, Hatfield DL, Silvestre R, et al. Relationship between the number of repetitions and selected percentages of one repetition maximum in free weight exercises in trained and untrained men. *J Strength Cond Res.* 2006;20(4):819–23.
 56. Lagally KM, Robertson RJ. Construct validity of the OMNI resistance exercise scale. *J Strength Cond Res.* 2006;20(2):252.
 57. Faull OK, Dearlove DJ, Clarke K, Cox PJ. Beyond RPE: the perception of exercise under normal and ketotic conditions. *Front Physiol.* 2019;10:229.
 58. Fairman CM, LaFountain RL, Lucas AR, Focht BC. Monitoring resistance exercise intensity using ratings of perceived exertion in previously untrained patients with prostate cancer undergoing androgen deprivation therapy. *J Strength Cond Res.* 2018;32(5):1360–5.

59. Borg G. A general scale to rate symptoms and feelings related to problems of ergonomic and organizational importance. *G Ital Med Lav Ergon*. 2008;30(1):8–10.
60. Hackett DA, Johnson NA, Halaki M, Chow C-M. A novel scale to assess resistance-exercise effort. *Sports Sci*. 2012;30(13):1405–13.
61. Steele J, Fisher J, Giessing J, Gentil P. Clarity in reporting terminology and definitions of set endpoints in resistance training. *Muscle Nerve*. 2017;56(3):368–74.
62. Giebetasing J, Fisher J, Steele J, Rothe F, Raubold K, Eichmann B. The effects of low-volume resistance training with and without advanced techniques in trained subjects. *J Sports Med Phys Fit*. 2016;56(3):249–58.
63. Kinsman R, Weiser P, Stamper D. Multidimensional analysis of subjective symptomatology during prolonged strenuous exercise. *Ergonomics*. 1973;16(2):211–26.
64. Hardy CJ, Rejeski WJ. Not what, but how one feels: the measurement of affect during exercise. *J Sport Exerc Psychol*. 1989;11(3):304–17.
65. Borg G. Physical performance and perceived exertion [dissertation]. Lund: Gleerup; 1962.
66. Gendolla GH, Wright RA. Effort. In: Sander D, Scherer KR, editors. *Oxford companion to the affective sciences*. New York: Oxford University Press; 2009. p. 134–5.
67. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000;32(4):1008–15.
68. Robinson-Papp J, George MC, Dorfman D, Simpson DM. Barriers to chronic pain measurement: a qualitative study of patient perspectives. *Pain Med*. 2015;16(7):1256–64.
69. Williams ACDC, Davies HT, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. *Pain*. 2000;85(3):457–63.
70. Dannecker EA, Warne-Griggs MD, Royse LA, Hoffman KG. Listening to patients' voices: Workarounds patients use to construct pain intensity ratings. *Qual Health Res*. 2018;29(4):484–97.
71. Hartman ME, Ekkekakis P, Dicks ND, Pettitt RW. Dynamics of pleasure–displeasure at the limit of exercise tolerance: conceptualizing the sense of exertional physical fatigue as an affective response. *J Exp Biol*. 2019;222(3):jeb186585.
72. Robertson RJ, Nixon PA, Caspersen CJ, Metz KF, Abbott RA, Goss FL. Abatement of exertional perceptions following dynamic exercise: physiological mediators. *Med Sci Sports Exerc*. 1992;24(3):346–53.
73. Utter AC, Kang J, Nieman DC, Dumke CL, Mcanulty SR, Mcanulty LS. Carbohydrate attenuates perceived exertion during intermittent exercise and recovery. *Med Sci Sports Exerc*. 2007;39(5):880–5.
74. Ekkekakis P, Parfitt G, Petruzzello SJ. The pleasure and displeasure people feel when they exercise at different intensities. *Sports Med*. 2011;41(8):641–71.
75. Ekkekakis P. People have feelings! Exercise psychology in paradigmatic transition. *Curr Opin Psychol*. 2017;16:84–8.
76. Cabanac M. Pleasure: the common currency. *J Theor Biol*. 1992;155(2):173–200.
77. Cabanac M, Leblanc J. Physiological conflict in humans: fatigue vs. cold discomfort. *Am J Physiol*. 1983;244(5):621–8.
78. Ekkekakis P, Petruzzello SJ. Acute aerobic exercise and affect. *Sports Med*. 1999;28(5):337–47.
79. Williams DM, Dunsiger S, Ciccolo JT, Lewis BA, Albrecht AE, Marcus BH. Acute affective response to a moderate-intensity exercise stimulus predicts physical activity participation 6 and 12 months later. *Sport Exerc Psychol*. 2008;9(3):231–45.
80. Rhodes RE, Kates A. Can the affective response to exercise predict future motives and physical activity behavior? A systematic review of published evidence. *Ann Behav Med*. 2015;49(5):715–31.
81. Hackett DA, Cogley SP, Davies TB, Michael SW, Halaki M. Accuracy in estimating repetitions to failure during resistance exercise. *J Strength Cond Res*. 2017;31(8):2162–8.
82. Zourdos MC, Goldsmith JA, Helms ER, Trepeck C, Halle JL, Mendez KM, et al. Proximity to failure and total repetitions performed in a set influences accuracy of intraset repetitions in reserve-based rating of perceived exertion. *J Strength Cond Res*. 2019. <https://doi.org/10.1519/jsc.0000000000002995> (**publish ahead of print**).
83. Ormsbee MJ, Carzoli JP, Klemp A, Allman BR, Zourdos MC, Kim JS, et al. Efficacy of the repetitions in reserve-based rating of perceived exertion for the bench press in experienced and novice benchers. *J Strength Cond Res*. 2019;33(2):337–45.
84. Helms ER, Storey A, Cross MR, Brown SR, Lenetsky S, Ramsay H, et al. RPE and velocity relationships for the back squat, bench press, and deadlift in powerlifters. *J Strength Cond Res*. 2017;31(2):292–7.
85. Steele J, Endres A, Fisher J, Gentil P, Giessing J. Ability to predict repetitions to momentary failure is not perfectly accurate, though improves with resistance training experience. *PeerJ*. 2017;5:e4105.