Exercise Tolerance during VO_{2max} Testing is a Multifactorial Psychobiological Phenomenon

Word count: 5396

Midgley, A. W., Earle, K., McNaugton, L. R., Siegler, J. C., Clough, P., & Earle, F. (2017). Exercise tolerance during VO2 max testing is a multifactorial psychobiological phenomenon. Research in Sports Medicine, 25(4), (480 - 494). doi:10.1080/15438627.2017.1365294. ISSN 1543-8635

Abstract

Fifty-nine men completed a VO_{2max} test and a questionnaire to establish reasons for test termination, perceived exercise reserve (difference between actual test duration and the duration the individual perceived could have been achieved if continued until physical limitation), and perception of verbal encouragement. Participants gave between 1 and 11 factors as reasons for test termination, including leg fatigue, various perceptions of physical discomfort, safety concerns, and achievement of spontaneously set goals. The two most common main reasons were leg fatigue and breathing discomfort, which were predicted by pre-to-post test changes in pulmonary function ($p = 0.038$) and explosive leg strength ($p = 0.042$; $R^2 = 0.40$). Median (interquartile range) perceived exercise reserve, was 45 (50) s. Two-thirds of participants viewed verbal encouragement positively, whereas one-third had a neutral or negative perception. This study highlights the complexity of exercise tolerance during VO_{2max} testing and more research should explore these novel findings.

Key words: Effort; fatigue; goal setting; maximal oxygen uptake; verbal encouragement

INTRODUCTION

Hill, Long, and Lupton (1924) proposed that during the assessment of maximal oxygen uptake (VO_{2max}) the cardiovascular system's capacity to deliver oxygen to working muscles is reached, resulting in muscle hypoxia, a rapid increase in intramuscular lactate concentration, and muscle fatigue leading to task failure. This chain of events has since been termed the cardiovascular/anaerobic/catastrophic model of fatigue (Noakes & St Clair Gibson, 2004). Other physiological catastrophe models of fatigue have subsequently been proposed and debated (see McKenna & Hargreaves, 2008, for review), however, more recently there has been increased interest in psychological influences in the tolerance of exercise (Evans, Boggero, & Segerstrom, 2015; McCormick, Meijen, & Marcora, 2015) and the conception of a psychobiological model (Marcora, 2008).

The psychobiological model of exercise tolerance postulates that maximal exercise is terminated due to task disengagement caused by a perceived intolerable perception of effort, or perception that continuing exercise is not possible because a maximal effort has been given (Marcora, 2008). Within motivational frameworks, this reflects differences in the willingness of individuals to continue exercising towards their true physical limit (Kalsbeek, 1968) and the gap between motivational intensity and potential motivation (Brehm & Self, 1989). Empirical evidence supporting a voluntary control of effort and task disengagement during VO_{2max} testing comes from observations that individuals typically terminate the test because of perceived pain and discomfort rather than any physical limitation (Hamilton, Killian, Summers, & Jones, 1996; Myers et al., 1992; O'Donnell, Chau, & Webb, 1998). Although these studies provide important insight into exercise tolerance during VO_{2max} testing, only the main reason for task disengagement was reported, including either an option for combined leg fatigue and breathing discomfort (Hamilton et al., 1996; O'Donnell et al., 1998) or combined general and leg fatigue (Myers et al., 1992). Since various bodily sensations and motivational factors are associated with exercise (Brehm & Self, 1989; Tenenbaum et al., 1999), it is plausible that the decision to disengage a task is influenced by many factors, including personal goals (Weinberg, Bruya, & Jackson, 1985). Another issue is that leg fatigue and breathing discomfort are the most common main reasons for VO_{2max} test termination (Hamilton et al., 1996; Myers et al., 1992; O'Donnell et al., 1998), however, it is currently not known whether participants who terminate the test due to leg fatigue demonstrate greater pre-to-post changes in explosive leg strength than changes in pulmonary function, and vice-versa. Furthermore, the observation that a VO_{2max} test is typically continued until task disengagement rather than task failure supports the proposition of a perceived exercise reserve, defined as the time difference between the actual test duration and the test duration the individual perceives could have been achieved if exercise had been continued until task failure. No studies have attempted to quantify it, however.

Verbal encouragement is another important aspect of exercise tolerance during VO_{2max} testing, used to facilitate the achievement of a maximal effort by the participant (Andreacci et al., 2002). Accordingly, several studies observed that verbal encouragement increased exercise duration during VO_{2max} testing (Andreacci et al., 2002; Chitwood, Moffatt, Burke, Luchino, & Jordan, 1997; Moffatt, Chitwood, & Biggerstaff, 1994). Limited research has found this to be inconsistent across personality types (Chitwood et al., 1997), however, suggesting that some individuals might have a neutral or negative perception of verbal encouragement. Furthermore, despite verbal encouragement being an extremely common component of exercise testing (Halperin, Pyne, & Martin, 2015), preferences for the content and timing of delivery of the encouragement have not been investigated. From a practical perceptive, there is currently no empirical basis to inform evidence-based practice of what characterises and quantifies effective verbal encouragement.

The purpose of the present study was to investigate several related aspects of exercise tolerance during VO_{2max} testing. Specific aims were to 1) establish reasons for task disengagement; 2) establish any time difference between actual test duration and the test duration participants thought they could have achieved if exercise had been continued until task failure; and 3) determine participant preferences for verbal encouragement. This is the first study to investigate these specific aims, which are all important practical aspects of maximal exercise testing (American Thoracic Society & American College of Chest Physicians, 2003).

METHODS

Participants. Fifty-nine participants were recruited from the university population and local sports clubs over a period of 33 months. Inclusion criteria dictated that participants were male, aged between 18 and 40 years, physically active, non-smokers (> 5 years), had no chronic disease, acute illness, or injury, and not taking prescribed medications. The mean (SD) characteristics of the 59 participants were as follows: age, 28 (7) yr; height, 1.80 (0.07) m; body mass, 84.8 (12.7) kg; sum of seven skinfolds, 12.1 (4.7) %; training frequency, 4.9 (1.2) days per wk; and total training duration, 5.8 (2.4) hr per wk (plus one outlier of 19 h per wk). This prospective observational study, involving each participant visiting the laboratory twice, was approved by the university research ethics committee and all participants provided written consent. During all procedures only one participant and one investigator were present, and the same male investigator supervised all procedures.

Laboratory Visit 1. Upon arrival to the environmentally controlled laboratory, which was maintained at a temperature of around 20°C and a relative humidity of around 40%, participants were screened using a health questionnaire, automated blood pressure monitor (Omron M6, Omron Healthcare, Milton Keynes, UK) and spirometer (MicroLoop, Cardinal Health, Hoechberg, Germany). Height, body mass, and skinfold measurements were then taken. Tests of pulmonary function and explosive leg strength were then practiced until there was a difference of less than 5% between three consecutive measurements, or 8 trials of the test had been completed, whichever came first. This familiarization was undertaken in an effort to improve the accuracy and reliability of the tests during visit 2. Pulmonary function was determined using forced vital capacity (FVC) and forced expiratory volume in 1 s (FEV_1) using a MicroLoop spirometer (Cardinal Health, Hoechberg, Germany), in accordance with the joint guidelines of the American Thoracic Society and European Respiratory Society (Miller et al. 2005). Explosive leg strength was determined using a standing vertical jump test. The stretch-shortening cycle is an important aspect of muscle function during running and using an assessment of muscle function that incorporates the stretch-shortening cycle, such as the standing vertical jump, is regarded as an effective approach to assessing muscle fatigue resulting from such activities as running (Komi, 2000). Standardisation of the jumps was facilitated by instructing participants to stand on the mat with feet shoulder width apart and hands on the hips, and to quickly bend the knees and then jump as high as possible whilst keeping the hands on the hips at all times. The height of the jump was measured using a commercially available jump mat (Newtest System, Newtest Oy, Oulu, Finland). Each pulmonary function and jump test was separated by approximately 60 s of rest.

Laboratory Visit 2. Upon arrival participants signed a form to confirm compliance with pre-test procedures, consisting of no large or high fat meal within 4 hr before the visit, no food ingestion within 1 hr, and no intense or unaccustomed exercise within 24 hr. Pulmonary function and explosive leg strength were then determined as previously described for laboratory visit 1. Each participant performed three jumps and three pulmonary function tests and the highest values were used for data analysis. Each participant then completed a VO_{2max} test performed on a computer-controlled motorised treadmill set at a 1% incline (HP Cosmos, HP Cosmos Sports and Medical gmbh, Nussdorf-Traunstein, Germany). A 5-min warm-up of steady running was followed by an incremental phase characterised by a ramp rate of 0.1 $km \cdot h^{-1}$ every 6 s until the participant reached his limit of tolerance. Participants were asked to terminate the test by either pressing the emergency stop button, or straddling the treadmill belt, and to practice their preferred method three times during the warm-up. The treadmill console displaying running speed, time, and distance was visible to participants throughout the test. A safety harness was not worn by any participant during any part of the test, but a safety mat was positioned behind the treadmill at all times. Of the 59 participants, 4 had never been on a treadmill before, 10 had previously used a treadmill, and 45 were currently using a treadmill during training. No participants had any observed difficulty with treadmill running during the VO_{2max} test. Verbal encouragement given during the VO_{2max} test was standardised by providing it every 60 s between 5 and 8 minutes of the incremental phase of the test, and every 30 s thereafter. Before delivery each verbal encouragement statement was memorised from a prepared script hidden from each participant's view. The encouragement included statements that the participant was doing well, instructional statements such as "Keep pushing, keep it going as long as you can", and reminders that the test requires the participant to provide a maximum effort. Respired air was analysed throughout the test using an automated open-circuit gas analysis system (Oxycon Pro®, Jaeger, Hocksberg, Germany) connected to a silicone face mask. Peak pulmonary oxygen uptake and peak respiratory exchange ratio were regarded as the highest 30 s stationary time-averaged values. The gas analyzers were calibrated immediately before each VO_{2max} test using ambient air and certified standard gases containing 16.0 ± 0.02% oxygen and 5.0 ± 0.02% carbon dioxide (Cryoservice Ltd, Woercester, UK). Heart rate was continuously measured using a telemetric heart rate monitor (model T41, Polar Electro Oy, Kempele, Finland) and peak heart rate was regarded as the highest 5 s stationary time-averaged value. A fingertip capillary blood sample was collected 3 min after the VO_{2max} test and immediately analysed for the determination of blood lactate concentration (ABL800 FLEX analyser, Radiometer Ltd, Crawley, UK). Three and 10 min after the VO_{2max} test each participant repeated three trials of the pulmonary function and standing vertical jump tests, respectively. Fifteen minutes after termination of the VO_{2max} test participants completed an Exercise Tolerance Questionnaire to establish reasons for task disengagement, goal setting, perceived exercise reserve, and perceived usefulness of verbal encouragement. For each listed reason for task disengagement, there was a choice of the following responses: 'not a contributory factor', 'a minor contributory factor', 'a major contributory factor', and 'the only contributory factor'. A separate section allowed the addition of reasons not listed. Participants circled what they perceived as being the main reason for task disengagement. Perceived exercise reserve was established by asking: 'If you had kept running instead of terminating the test when you did, how long do you think it would have been before you either fell on the treadmill, or went off the back of the treadmill?' Treadmill running was used because test termination time is more clearly defined, compared to cycle ergometry where the test is typically allowed to continue for a short time close to the end while the test administrator verbally encourages the participant to increase pedal cadence or work rate because it has dropped below that which is required. Perceived usefulness of verbal encouragement in motivating participants to provide a maximal effort was determined by them marking a visual analogue scale (Aitken, 1969), consisting of a 100 mm horizontal line anchored at each end with the text '0 (not at all helpful)' and '100 (extremely helpful). Spaces were provided for participants to state whether they found specific encouragement statements or timings of encouragement particularly helpful or unhelpful. The Exercise Tolerance Questionnaire was developed from the findings of previous studies that investigated reasons for VO_{2max} test termination (Hamilton et al., 1996; Myers et al., 1992; O'Donnell et al., 1998) and the responses of 10 additional participants that each previously completed a semi-structured interview approximately 10 min after completing a VO_{2max} test. The latter findings are not reported here.

Data Analyses. Statistical analyses were completed using IBM SPSS statistics 22 (SPSS Inc., Chicago, IL). Reasons for test termination are reported as frequencies. Continuously distributed sample data are reported as the mean and standard deviation (SD), or the median and interquartile range (IQR), as appropriate. Multivariate stepwise logistic regression was used to investigate whether changes in explosive leg strength and pulmonary function pre-to-post VO_{2max} test could predict whether the main reason for task disengagement was leg fatigue rather than breathing discomfort. Two-tailed statistical significance was accepted as $p < 0.05$. The open-ended responses for participants' attitudes to verbal encouragement were content analysed to provide a condensed description. This inductive process involved a systematic examination of the responses in three phases. First, data were open coded by reading responses and generating headings which described the content. Categories were then freely generated from the headings and finally, categories were reduced by collapsing similar headings into higher-order categories.

RESULTS

Table 1 shows the summary statistics for the VO_{2max} test results.

Reasons for VO_{2max} Test Termination

Only one participant continued the VO_{2max} test until task failure, during which he fell off the back of the treadmill onto the safety mat. Figure 1A shows the distribution of the number of different reasons given for task disengagement for the other 58 participants, with between one and eleven reasons given, with a mode of five. Table 2 shows the specific reasons participants gave for terminating the VO_{2max} test and whether each reason was perceived as a minor or major contributory factor for task disengagement. In total there were 22 different reasons given by the 58 participants. These included a perceived inability to maintain the required running speed, various sensations of physical fatigue and discomfort, safety concerns, and whether spontaneously set goals had been achieved. Figure 1B shows the frequency distribution for the reason each participant perceived as being the main contributory factor for task disengagement. Forty-two participants set time, distance, or running speed performance goals before or during the VO_{2max} test, four participants set a physical perception goal, and two set a target heart rate goal. Of the 42 participants who set performance goals, 25 stated reaching their goal was a contributory factor for task disengagement, and three said it was the main contributory factor for task disengagement. Changes in standing vertical jump height ($p = 0.042$) and FEV_1 ($p = 0.038$) pre-to-post VO_{2max} test predicted whether the main reason for task disengagement was leg fatigue rather than breathing discomfort (overall logistic regression model: $p = 0.007$; Nagelkerke $R^2 = 0.40$).

Perceived Exercise Reserve

Figure 1C shows the frequency distribution for perceived exercise reserve formatted into 15 s time bins. The median (IQR) perceived exercise reserve was 45 (50) s, with a range of 4 to 120 s for 57 of the 58 participants who volitionally terminated the test. One participant was an outlier with a value of 180 s. Twenty-six participants reported a perceived exercise reserve of at least 1 min and nine participants at least 2 min. Only five participants reported a value of less than 15 s.

Perceptions of Verbal Encouragement

The median (IQR) value for how useful participants perceived the verbal encouragement in motivating them to provide a maximal effort was 73.0 (39), with a range of 0 to 100 on the 0 to 100 VAS scale. Figure 1D shows the frequency distribution of the data placed into deciles.

Attitudes to verbal encouragement varied across participants, ranging from highly positive to highly negative. Of the 44 open-ended responses, 29 referred to the verbal encouragement in a positive way, and five negatively. There were two broad categories of positive evaluation. Six responses reported verbal encouragement as generally facilitative, such as *"in general all encouragement was beneficial" and "any encouragement is helpful".* A second larger group of responses identified specific phrases that helped or supported participants in maintaining effort (n = 23). These responses can be described within four sub-categories, including general encouragement, *"just keep going",* use of power words, "*keep* *pumping those arms and legs"*, reference to maximum, *"keep going to max",* and positive reinforcement, *"you're doing really well'.* Of the five negative responses, one participant referred to the verbal encouragement as *"annoying"*, and four participants found the encouragement generally distracting, *"it was putting me off focusing"* and "*at near max I was keyed up and found the verbal encouragement a distraction"*. Two participants stated that the tone of the voice, rather than the words, was most important. The ten responses coded as neutral could be summarised in four categories, those unable to recall the verbal encouragement (n = 2), those who perceived no impact (n = 2), those who felt they did not need it to remain motivated (n = 3), and those who perceived some words/phrases positively and some negatively (n = 3).

Participants also were asked to comment on any preferences for timing of delivery of verbal encouragement. Of the 39 open-ended responses, five participants found the verbal encouragement to be helpful throughout the test, stating encouragement was *"helpful at all times"* and *"from the first words and thereafter"*. However, a much higher proportion of participants found verbal encouragement to be most helpful in the later stages *"towards the end"* (n = 16), or specifically referring to the point in the test when it was *"getting hard"* or they were *"flagging"* (n = 11). Of the remaining seven responses, one participant found only the first verbal encouragement helpful, *"first instance ok, rest unhelpful",* and one participant stated that the encouragement was particularly helpful when the "tester can see any decrease in effort or results begin to plateau". Four participants found the later encouragements to be distracting, for example, "*the calls towards the end were off-putting",* and *"near max a negative distraction",* with one participant perceiving the encouragement as annoying "most of the time".

DISCUSSION

A main finding of the present study was that each participant typically terminated the VO_{2max} test due to multiple reasons, including various perceptions of physical discomfort, safety concerns, and achievement of spontaneously set goals. Other notable findings were that some participants terminated the test a substantial amount of time before any perceived physical limitation, and around one-third of participants had a neutral or negative perception of the verbal encouragement that was used to try and motivate them during the test. These novel findings are discussed in each of the relevant sections below.

Reasons for VO_{2max} Test Termination

Participants in the present study collectively gave 14 different main reasons for test termination, which is more than the apparently healthy participants in previous studies that reported only two factors (leg fatigue and breathlessness; O'Donnell et al., 1998), three factors (leg discomfort, breathing discomfort, and one participant categorised as 'other'; Hamilton et al., 1996), and six factors (general/leg fatigue, breathlessness, ankle/knee/foot pain, back pain, dizziness, and nausea; Myers et al., 1992). A plausible explanation for the greater number of reasons observed in the present study is the more developed questionnaire that was used, which was based on questionnaires used in the previous studies and qualitative analysis of semi-structured interviews that investigated reasons for test termination. If traditional catastrophe models of fatigue are applied to treadmill-based VO_{2max} testing, the main reason for test termination would be a perceived inability to maintain the required leg turnover speed to keep up with the treadmill belt. Although this was the most common main reason for test termination in the present study, which is consistent with a previous study that used treadmill running (Myers et al., 1992), it was the main reason for only 28% of the participants. A plausible explanation for some of the inter-individual variability for why people perceive different sensations as main reasons for VO_{2max} test termination, is the relative functional capacity or resilience to fatigue of different bodily systems. This is supported by the observation that pre-to-post changes in explosive leg strength and pulmonary function were significant predictors for when the main reason for test termination was leg fatigue rather than breathing discomfort. In other words, those individuals with greater changes in explosive leg strength and less change in pulmonary function pre-to-post VO_{2max} test tended to terminate the test because of leg fatigue and vice versa. Another important consideration is that the perceived difficulty in maintaining leg speed was likely test protocol specific, given that treadmill speed increased over time while maintaining a constant grade of 1%. Protocols that result in relatively high treadmill grades, such as the Bruce protocol, have been shown to result in significantly lower VO_{2max} values than protocols with lower grades, which is thought to be partially a result of premature fatigue of the leg and trunk muscles due to participants' lack of familiarity with inclined running (Kang, Chaloupka, Mastrangelo, Biren, & Robertson, 2001). The present results should therefore be interpreted within the context of the characteristics of the specific treadmill protocol that was used.

A main finding of the present study is that the decision to terminate the VO_{2max} test was typically multifactorial, including factors not related to sensations of physical fatigue and discomfort, such as safety concerns and whether spontaneously set goals had been achieved. In fact, only one participant gave a single reason for test termination. These findings are novel, as previous studies reported only the main reason for task disengagement during VO_{2max} testing, which were all related to sensations of physical fatigue and discomfort (Hamilton et al., 1996; Myers et al., 1992; O'Donnell et al., 1998). Collectively, participants in the present study perceived 22 different reasons for task disengagement. Twelve of the contributory factors involved in test termination were directly related to sensations of physical discomfort, which might be expected given the 25 different sensations of physical discomfort previously reported by 10 runners during a demanding 9.1 km run (Tenenbaum et al., 1999). This ability to experience a complex array of uncomfortable physical sensations during exercise is via group III and IV afferents (O'Connor & Cook, 1999), although it is clear that exercise tolerance is not simply a matter of experienced afferent sensory sensations. Gate-Control Theory, for example, has been used to explain how physical discomfort during exercise influences exercise tolerance using an evaluation of the intensity and location of the discomfort, its negative consequences, and a decision about how to respond (Tenenbaum et al., 1999).

The present finding that task disengagement is typically because of sensations of physical discomfort and other contributory factors not related to physical discomfort is consistent with Hockey's model of motivational control, in which effort is managed by a compensatory control feedback loop (Hockey, 1997). This incorporates an effort budget, ongoing evaluation of effort investment, and evaluation of performance against goals. In the current context, effort would be determined largely by the experienced demands up to the point where the set effort budget is reached. Participants may opt to increase the effort budget at this point if the perceived value of the test remains high or a goal has not been achieved, or lower their investment if the costs of continuing exercise outweigh potential rewards or a goal has been achieved. Two important practical aspects are therefore goal setting and how highly the test is valued by the participant. The importance of personal goals was apparent in the present study, as 81% of participants spontaneously set one or more self-selected goals before or during the VO_{2max} test. This is similar to the 82% reported by Weinberg et al. (1985) for a maximal 3-min sit-up test. A notable finding of the present study, however, is that 25 of the 48 goal setters stated that achieving their goal influenced the decision to terminate the test, and three stated it was the main contributory factor. Future research should investigate the extent to which spontaneous goal setting is involved in the evaluative process leading to task disengagement during maximal exercise testing. For example, it would be interesting to establish the effect of obscuring the treadmill console display on spontaneous goal setting and test performance. With respect to how highly the test is valued by the participant, test administrators should be selective in what they say to participants before and during a VO_{2max} test to try and ensure the perceived value of the test remains high as perceptions of effort and discomfort increase.

Perceived Exercise Reserve

Evidence of an actual exercise reserve has been empirically investigated based on whether or not participants can exercise at a higher power output during a short bout of all-out exercise performed immediately after an exhaustive exercise bout (Ferguson, Wylde, Benson, Cannon, & Rossiter, 2016; Marales-Alamo et al., 2015; Marcora & Staiano, 2010). Findings from these studies, however, have been inconsistent and somewhat controversial due to some perceived methodological issues (Allen & Westerblad, 2010; Burnley, 2010; Morales-Alamo, Martin-Rincon, Perez-Valera, Marcora, & Calbet, 2016). The present study used a different methodological approach by investigating the extent of a perceived exercise reserve. Participants were informed before the VO_{2max} test, and during the test as part of verbal encouragement, of the importance of providing a maximal effort for obtaining accurate test results. Despite this, a novel finding of the present study was that many participants terminated the test at least 60 s before their perceived physical capacity, thereby supporting a psychobiological model of exercise tolerance during treadmill VO_{2max} testing (Marcora, 2008). The accuracy of perceived exercise reserve could be questioned, particularly considering the incremental nature of the VO_{2max} test. The most important consideration, however, is that many participants *believed* they could have exercised for considerably longer before task failure, but consciously decided to terminate the test. A certain amount of perceived exercise reserve would reasonably be expected during a treadmill-based VO_{2max} to allow the participant to stop the test safely; however, fear of falling was the main reason for stopping for only two participants.

The concept of a perceived exercise reserve is consistent with the observation that exercise tolerance can be increased by experimental interventions such as verbal encouragement (Moffatt et al., 1994) and exposure to a competitive environment (Wilmore, 1968). In other words, individuals typically perceive they could do more, but choose to do so only under certain conditions. It should be noted that a large inter-individual variation in perceived exercise reserve was observed in the present study, with values ranging from 4 to 180 s, reflecting considerable variability in the willingness of individuals to continue exercising towards their true perceived physical limit (Kalsbeek, 1968). This variability should be an important consideration in the interpretation of future studies investigating task failure during particular physical performance tests. Future research should further develop the concept of a perceived exercise reserve, including why people do not attempt to exercise longer during VO_{2max} testing if they believe they can. Such attributes as mental toughness, pain tolerance, and conscientiousness, as well as the value participants place on VO_{2max} test results and any personal goals, are all potential factors in explaining inter-individual variation in perceived exercise reserve.

Perceptions of Verbal Encouragement

The majority of the 44 participants in the present study who provided qualitative feedback perceived verbal encouragement as useful in helping motivate them to invest effort during the VO_{2max} test. This observation is consistent with experimental studies that reported increases in mean time to exhaustion of between 8% and 18% with the addition of verbal encouragement during VO_{2max} testing (Andreacci et al., 2002; Chitwood et al. 1997; Moffatt et al., 1994). A novel finding of the present study, however, is that approximately one-third of the 44 participants who gave qualitative feedback reported a neutral (n = 10) or negative (n = 5) perception of verbal encouragement. Inter-individual differences in personality traits might explain these varied perceptions. Chitwood et al. (1997) reported that verbal encouragement had no significant effect on the mean time to exhaustion of type A scorers during a VO_{2max} test, but increased the mean time to exhaustion of type B scorers by 15.7% (Chitwood et al., 1997). Type A scorers are characterised by extremes of competitiveness, time urgency, and aggressiveness, whereas Type B scorers are characterised by the relative absence of Type A characteristics (Carver, Coleman, & Glass, 1976). Similar personality effects were reported by Binboğa, Tok, Catikkas, Guven, & Dane (2013), where verbal encouragement improved maximal voluntary contraction force in low conscientious individuals, but not high conscientious individuals. It is notable that four of the five participants in the present study who perceived verbal encouragement negatively stated that it distracted them. A plausible explanation is that the encouragement interfered with their attentional focus strategy. If true, a reduction in exercise tolerance might be expected, since participants tend to demonstrate greater exercise tolerance when adopting their preferred attentional focus strategy (Baghurst, Thierry, & Holder, 2004). The large inter-individual differences in how useful verbal encouragement was perceived is further evidenced by a response range of between 0 and 100 on the 0-100 VAS scale. Given the qualitative responses evidenced that some people perceived verbal encouragement negatively, however, it would have been more informative if the VAS scale was centred around zero on a -50 to 50 scale, with the words ‘extremely demotivating’ and ‘extremely motivating’ anchoring the opposite ends of the scale.

Given the use of verbal encouragement during exercise testing is commonplace (Halperin et al., 2015), it is notable that previous research has investigated the efficacy of different frequencies of verbal encouragement, but not the content or timing of delivery. Andreacci et al. (2002) observed that providing verbal encouragement every 20 s during a VO_{2max} test increased mean exercise duration, but not when given every 60 or 180 s. The present study suggests a more effective strategy would be to progressively increase the frequency of encouragement during the test, as most participants perceived the encouragement as most useful near the end. Further support comes from the observation that spontaneous motivational self-talk during exercise is used more frequently during the most difficult part of exercise and when considering quitting (Gammage, Hardy, & Hall, 2001). Two of the five participants who perceived the encouragement negatively, however, stated that the encouragement was most distracting towards the end of the test, further highlighting the complex nature of the interaction between the characteristics of verbal encouragement and personal preference.

Responses for the preference for specific content of encouragement were more varied than timing of encouragement. Some participants found all encouragement useful, some preferred certain words or phrases, and some were not able to recall what was said, but stated that the tone of the voice was most important. These findings clearly highlight that further research employing more complex and focused methodology is needed, so that recommendations can be made regarding the characteristics of effective verbal encouragement with respect to content, tone, loudness, timing, and frequency. Studies that investigated the characteristics of self-talk used by individuals to motivate themselves during strenuous exercise should assist in this endeavour (Gammage et al., 2001; van Raalte, Morrey, Cornelius, & Brewer, 2015). Future research also should investigate whether verbal encouragement can have a negative effect on exercise tolerance in individuals with extreme personality types, or specific attentional focus strategies.

Conclusion

The purpose of the present study was to investigate several related aspects of exercise tolerance during VO_{2max} testing. The main findings were that apparently healthy men typically terminated a treadmill-based VO_{2max} test a considerable amount of time before any perceived physical limitation and due to multiple reasons. These reasons included a perceived inability to maintain the required running speed, various perceptions of physical discomfort, safety concerns, and whether spontaneously set goals had been achieved. Large inter-individual differences in the perceived usefulness of verbal encouragement for improving exercise tolerance, as well as the preference for timing and content of the verbal encouragement, also were observed. These findings reflect the complexity of exercise tolerance during treadmill-based VO_{2max} testing. The results also raise doubt about whether verbal encouragement should always be used during VO_{2max} testing, and when it is used, what should be said and when. The findings of the present study support a psychobiological model of exercise tolerance, but this should not be regarded as evidence to disregard other models. Different modalities and durations of exercise in different populations and environments are likely explained by different models of exercise tolerance. Consequently, generalising beyond the population, exercise test, and environmental conditions used in the present study should be undertaken with caution. Another consideration is that the participants were tested at different times of the day and it is possible that some of the observed inter-individual differences in responses were due to this. Furthermore, the present study explored several related aspects of exercise tolerance during VO_{2max} testing and more focused research is needed to further investigate the issues relating to the novel findings that have been presented and discussed.

Disclosure statement

The authors have no potential conflicts of interest relating to this research.

REFERENCES

Aitken, R.C. (1969). Measurement of feelings using visual analogue scales. *Proceedings of the Royal Society of Medicine*, *62*(10), 989-993.

Allen, D., & Westerblad, H. (2010). What limits exercise during high-intensity aerobic exercise? *European Journal of Applied Physiology*, *110*(3), 661-662.

American Thoracic Society & American College of Chest Physicians (2003). ATS/ACCP statement on cardiopulmonary exercise testing. *American Journal of Respiratory and Critical Care Medicine*, *167*(2), 211-277.

Andreacci, J.L., LeMura, L.M., Cohen, S.L., Urbansky, E.A., Chelland, S.A., & Von Duvillard, S.P. (2002). The effects of frequency of encouragement on performance during maximal exercise testing. *Journal of Sports Sc*iences, *20*(4), 345-352.

Baghurst, T., Thierry, G., & Holder, T. (2004). Evidence of a relationship between attentional styles and effective cognitive strategies during performance. *Athletic Insight*, *6*(1), 36-51.

Binboğa, E., Tok, S., Catikkas, F., Guven, S., & Dane, S. (2013). The effects of verbal encouragement and conscientiousness on maximal voluntary contraction of the triceps surae muscle in elite athletes. *Journal of Sports Sciences*, *31*(9), 982-988.

Brehm, J.W., & Self, E.A. (1989). The intensity of motivation. *Annual Review of Psychology*, *40*, 109-131.

Burnley, M. (2010). The limitation to exercise tolerance in humans: validity compromised by failing to account for the power-velocity relationship. *European Journal of Applied Physiology*, *109*(6), 1225-1226.

Carver, C.S., Coleman, A.E., & Glass, D.C. (1976). The coronary-prone behavior pattern and the suppression of fatigue on a treadmill test. *Journal of Personality and Social Psychology*, *33*(4), 460-466.

Chitwood, L.F., Moffatt, R.J., Burke, K., Luchino, P., & Jordan, J.C. (1997). Encouragement during maximal exercise testing of type A and type B scorers. *Perceptual and Motor Skills*, *84*(2), 507-512.

Evans, D.R., Boggero, I.A., & Segerstrom, S.C. (2015). The nature of self-regulatory fatigue and "ego-depletion": lessons from physical fatigue. *Personality and Social Psychology Review*, DOI: 10.1177/1088868315597841.

Ferguson, C., Wylde, L.A., Benson, A.P., Cannon, D.T., & Rossiter, H.B. (2016). No reserve in isokinetic cycling power at intolerance during ramp incremental exercise in endurance-trained men. *Journal of Applied Physiology*, *120*(1), 70-77.

Gammage, K.L., Hardy, J., & Hall, C.R. (2001). A description of self-talk in exercise. *Psychology of Sport and Exercise*, *2*, 233-247.

Halperin, I., Pyne, D.B., & Martin, D.T. (2015). Threats to internal validity in exercise science: a review of overlooked confounding variables. *International Journal of Sports Physiology and Performance*, *10*(7), 823-829.

Hamilton, A.L., Killian, K.J., Summers, E., & Jones, N.L. (1996). Symptom intensity and subjective limitation to exercise in patients with cardiorespiratory disorders. *Chest*, *110*(5), 1255-1263.

Hill, A.V., Long, C.N.H., & Lupton, H. (1924). Muscular exercise, lactic acid, and the supply and utilization of oxygen. *Proceedings of the Royal Society London Series B*, *97*(682), 155-176.

Hockey, G.R. (1997). Compensatory control in the regulation of human performance under stress and high workload; a cognitive-energetic framework. *Biological Psychology*, *45*(1-3), 73-93.

Kalsbeek, J.W. (1968). Measurement of mental workload and of acceptable load: possible applications in industry. *International Journal of Production Research*, *7*(1), 33-45.

Kang, J., Chaloupka, E.C., Mastrangelo, M.A., Biren, G.B., & Robertson, R.J. (2001). Physiological comparisons among three maximal treadmill exercise protocols in trained and untrained individuals. *European Journal of Applied Physiology*, *84*, 291-295.

Komi P.V. (2000). Strength-shortening cycle: a powerful model to study normal and fatigued muscle. *Journal of Biomechanics*, 33, 1197-1206.

Marcora, S.M. (2008). Do we really need a central governor to explain brain regulation of exercise performance? *European Journal of Applied Physiology*, *104*(5), 929-931.

Marcora, S.M., & Staiano, W. (2010). The limit to exercise tolerance in humans: mind over muscle? *European Journal of Applied Physiology*, *109*(4), 763-770.

McCormick, A., Meijen, C., & Marcora, S. (2015). Psychological determinants of whole-body endurance performance. *Sports Medicine*, *45*(7), 997-1015.

McKenna, M.J., & Hargreaves, M. (2008). Resolving fatigue mechanisms determining exercise performance: integrative physiology at its finest! *Journal of Applied Physiology*, *104*(1), 286-287.

Miller, M.R., Hankinson, J., Brusasco, V., Burgos, F., Casaburi, R., Coates, A., … Wanger, J. (2005). Standardisation of spirometry. *European Respiratory Journal*, *26*(2), 319-338.

Moffatt, R.J., Chitwood, L.F., & Biggerstaff, K.D. (1994). The influence of verbal encouragement during assessment of maximal oxygen uptake. *Journal of Sports Medicine and Physical Fitness*, *34*(1), 45-49.

Morales-Alamo, D., Losa-Reyna, J., Torres-Peralta, R., Martin-Rincon, M., Perez-Valera, M., Curtelin, D., … Calbet, J.A. (2015). What limits performance during whole-body incremental exercise to exhaustion in humans? *Journal of Physiology*, *593*(20), 4631-4648.

Morales-Alamo, D., Martin-Rincon, M., Perez-Valera, M., Marcora, S., & Calbet, J.A. (2016). No functional reserve at exhaustion in endurance-trained men? *Journal of Applied Physiology*, *120*(4), 476.

Myers, J., Buchanan, N., Smith, D., Neutel, J., Bowes, E., Walsh, D., & Froelicher, V.F. (1992). Individualized ramp treadmill: observations on a new protocol. *Chest*, *101*(5 Suppl.), 236S-241S.

Noakes, T.D., & St Clair Gibson. A. (2004). Logical limitations to the "catastrophe" models of fatigue during exercise in humans. *British Journal of Sports Medicine*, *38*(5), 648-649.

O'Connor, P.J., & Cook, D.B. (1999). Exercise and pain: the neurobiology, measurement, and laboratory study of pain in relation to exercise in humans. *Exercise and Sport Sciences Reviews*, *27*, 119-166.

O'Donnell, D.E., Chau, L.K., & Webb, K.A. (1998). Qualitative aspects of exertional dyspnea in patients with interstitial lung disease. *Journal of Applied Physiology*, *84*(96), 2000-2009.

Tenenbaum, G., Fogarty, G., Stewart, E., Calcagnini, N., Kirker, B., Thorne, G., & Christensen, S. (1999). Perceived discomfort in running: scale development and theoretical considerations. *Journal of Sports Sciences*, *17*(3), 183-196.

van Raalte, J.L., Morrey, R.B., Cornelius, A.E., Brewer, B.W. (2015). Self-talk of marathon runners. *The Sport Psychologist*, *29*, 258-260.

Weinberg, R., Bruya, L., & Jackson, A. (1985). The effects of goal proximity and goal specificity on endurance performance. *Journal of Sport Psychol*ogy, *7*, 296-305.

Wilmore, J.H. (1968). Influence of motivation on physical work capacity and performance. *Journal of Applied Physiol*ogy, *24*(4), 459-463.

Table 1. Summary statistics for test duration and peak physiological values for the VO_{2max} test (n = 59).

	Mean	SD	Minimum	Maximum
Test duration (min:s)	11:09	1:04	8:40	13:34
Pulmonary oxygen uptake ($ml \cdot kg^{-1} \cdot min^{-1}$)	50.5	6.8	36.0	65.2
Respiratory exchange ratio	1.21	0.07	1.08	1.38
Heart rate ($beats \cdot min^{-1}$)	190	11	153	218
Blood lactate concentration (mM)*	12.7	2.7	6.8	19.0

* Determined 3 min after test termination.

Table 2. Number of participants who stated that particular reasons for task disengagement during a VO_{2max} test were not a contributory factor, a minor contributory factor, or a major contributory factor (n = 58). One participant's results are not included since he experienced task failure.

Reason	Contributory Factor?		
	Major	Minor	No
Difficulty in maintain leg speed	28	20	10
Overall feeling of exhaustion	21	27	10
Breathlessness / breathing discomfort	24	22	12
Leg muscle ache / pain	16	26	16
Overall feeling of discomfort	8	30	20
Reached speed or time goal	7	18	33
Fear of falling on or off treadmill	7	10	41
Dizzy / light-headed	4	19	35
Lack of motivation to continue	3	8	47
Nausea	3	5	50
Concerned might faint / black out	1	6	51
Concerned might aggravate previous injury	1	5	52
Boredom	1	1	56
Concerned might experience cardiac event	1	0	57
Concerned might get injured	0	6	52
Sweat in the eyes	0	2	56

Reasons given under the open-ended 'other' category were as follows: unable to maintain arm speed (1 participant), arm ache (1 participant), back pain (1 participant), dry throat/mouth (2 participants), felt hot (1 participant), and needed to urinate (1 participant).

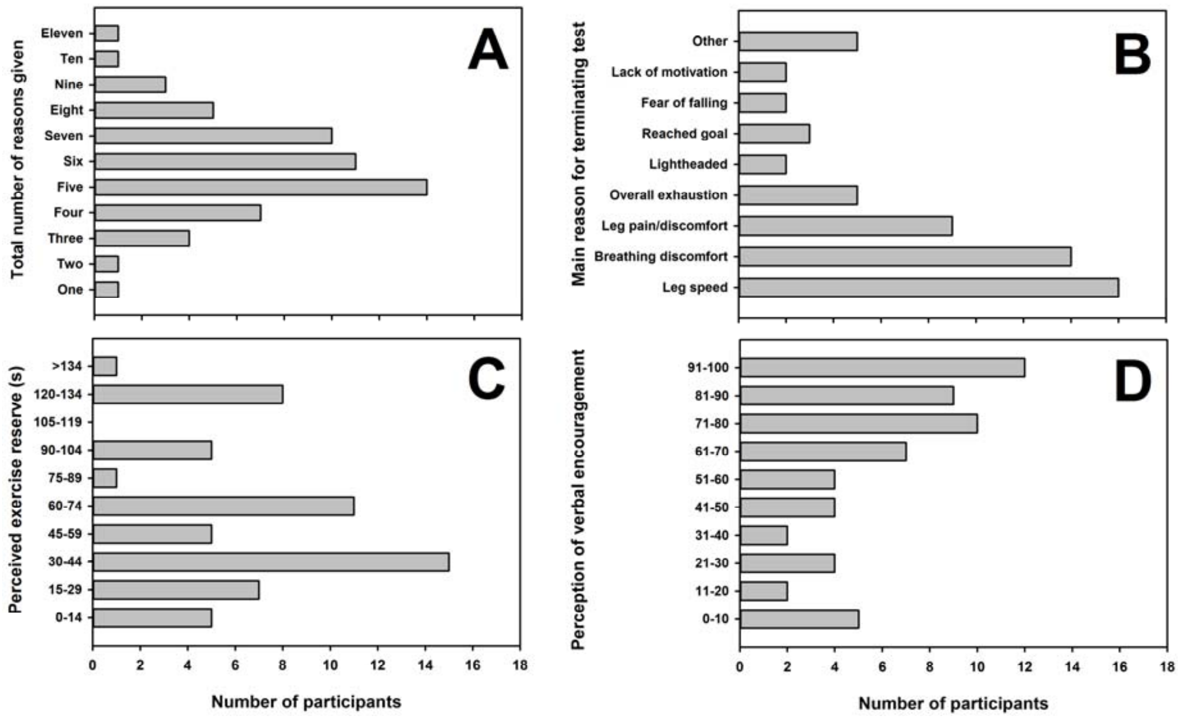

Figure 1. Frequency distributions for (**A**) the total number of reasons each participant gave for terminating the VO_{2max} test; (**B**) the number of participants who stated that a particular factor was the main reason for terminating the test; (**C**) the time participants estimated that they could have continued running before they either collapsed on the treadmill or fell off the back if they had not terminated the VO_{2max} test at the moment they did (i.e. perceived exercise reserve); and (**D**) how useful participants perceived the verbal encouragement for helping them invest effort during the VO_{2max} test, scored on a visual analogue scale ranging from 0 (not at all helpful) to 100 (extremely helpful). In panel B, other main reasons for terminating the test were back pain, overall feeling of discomfort, a feeling that might faint, concerned might experience a cardiac event, concerned might aggravate a previous injury, and a feeling of nausea.