Training to Fatigue: The Answer for Standardization When Assessing Muscle Hypertrophy?

Scott J. Dankel¹ · Matthew B. Jessee¹ · Kevin T. Mattocks¹ · J. Grant Mouser¹ · Brittany R. Counts¹ · Samuel L. Buckner¹ · Jeremy P. Loenneke¹

Abstract Studies examining resistance training are of importance given that increasing or maintaining muscle mass aids in the prevention or attenuation of chronic disease. Within the literature, it is common practice to administer a set number of target repetitions to be completed by all individuals (i.e. 3 sets of 10) while setting the load relative to each individual’s predetermined strength level (usually a one-repetition maximum). This is done under the assumption that all individuals are receiving a similar stimulus upon completing the protocol, but this does not take into account individual variability with regard to how fatiguing the protocol actually is. Another limitation that exists within the current literature is the reporting of exercise volume in absolute or relative terms that are not truly replicable as they are both load-dependent and will differ based on the number of repetitions individuals can complete at a given relative load. Given that the level of fatigue caused by an exercise protocol is a good indicator of its hypertrophic potential, the most appropriate way to ensure all individuals are given a common stimulus is to prescribe exercise to volitional fatigue. While some authors commonly employ this practice, others still prescribe an arbitrary number of repetitions, which may lead to unfair comparisons between exercise protocols. The purpose of this opinion piece is to provide evidence for the need to standardize studies examining muscle hypertrophy.

In our opinion, one way in which this can be accomplished is by prescribing all sets to volitional fatigue.

Key Points

The heterogeneity among resistance training protocols is problematic given the variability in repetitions that can be completed at a given relative intensity.

The hypertrophic potential of resistance training appears to be related to the fatigability of exercise and activation of motor units.

Performing exercise to volitional fatigue helps to ensure all individuals receive a similar stimulus and will allow for the comparison of other training variables such as sets, rest, intensity, etc.

1 Introduction

The American College of Sports Medicine (ACSM) and United States Department of Health and Human Services (USDHHS) recommend that individuals engage in resistance exercise at least twice per week as part of a comprehensive exercise program [1, 2]. Behavioral engagement in resistance exercise has been shown to reduce the odds of multimorbidity [3] and premature mortality [4], which may be attributed to increased muscle mass aiding in the prevention or attenuation of chronic diseases such as obesity and type II diabetes [5]. While the benefits of increasing muscle mass lend to the importance
of examining the efficacy of different resistance training protocols, the lack of standardization with respect to the number of repetitions prescribed makes it difficult to compare the importance of altering different training variables across studies (for review see Wernbom et al. [6]). Even studies administering the same protocol (e.g. 3 sets of 10 repetitions at 70 % one-repetition maximum [1RM]) across all individuals within a study population may provide a different stimulus as some individuals may be able to complete more repetitions at a given relative load [7, 8]. Given that muscle growth from resistance training appears to be dependent on increasing muscle activation through fatiguing protocols [9], a set number of repetitions may produce differential results within the same sample of individuals with respect to muscle growth, and may potentially contribute, in part, to the heterogeneous response to resistance training [10]. For this reason, no studies using an arbitrary number of repetitions can truly be replicated because the results will depend on the 'endurance' capacity of individuals included within the study population.

Studies attempting to provide equal stimuli between protocols attempt to do so by matching exercise volume expressed in total kilograms [11], but this can be problematic as the stimulus will differ based on the load being used. For example, exercising with a 30 or 90 % of 1RM load to volitional fatigue produced similar increases in skeletal muscle protein synthesis; however, matching the total volume produced differential results [12]. This lower elevation in skeletal muscle protein synthesis from the 30 % 1RM work-matched protocol can likely be attributed to the stimulus that was incapable of eliciting great enough fatigue and motor unit recruitment. Therefore, the reduced muscle activation from insufficient fatigue can explain the lack of elevation in skeletal muscle protein synthesis. This is evident in that only those muscle fibers sensing tension would upregulate the synthesis of new proteins [13], which demonstrates that exercise protocols can be more appropriately compared if they are performed to volitional fatigue as opposed to being volume-matched. Indeed, 10 weeks of training to volitional fatigue with 30 or 80 % 1RM loads produced similar increases in knee extensor muscle size [14] and, similarly, 6 weeks of bench press exercise performed to volitional fatigue with either 30 or 70 % 1RM elicited similar increases in muscle size of the pectoralis major and triceps brachii [15]. The idea that training to volitional fatigue allows for a more appropriate comparison of different training variables is supported by the recent meta-analytic studies assessing the importance of different training variables [16, 17]. The authors do not include studies that match total exercise volume, but rather only assess studies that performed exercise to volitional fatigue, likely because this allows for more appropriate comparisons. While this may have been a large reason the authors were limited to assessing only eight [16] or nine [17] studies, other meta-analyses do not exclude those not exercising to volitional fatigue [18], which may be problematic given the heterogeneity in the protocol itself.

The purpose of this current opinion was to propose a method to increase the homogeneity of resistance training protocols that are specifically examining muscle growth. Muscle strength will not be addressed in this current opinion because it appears to be primarily driven by the intensity of exercise [14], and this can easily be standardized by using a relative load across individuals (i.e. %1RM). Rather than administering an arbitrary number of repetitions to be completed, we encourage researchers to administer all sets performed to volitional fatigue to ensure a more relative stimulus is applied upon the completion of each set. Additionally, this allows for the reporting of exercise volume to be expressed in terms of ‘fatigu ing sets’ which can be easily replicated, interpreted and compared across studies when other training variables are controlled for.

2 Potential Problems with Reporting Volume in Kilograms

A common way to match the resistance exercise stimulus is to calculate the total volume of work completed in kilograms (repetitions × load lifted) [11]. Not only does this method of reporting volume have limitations as it is largely impacted by the absolute load being lifted, but it also does not take into account how fatiguing the training stimulus actually is. A recent study illustrated that bench press exercise performed at 30 or 75 % of an individual’s 1RM produced similar increases in muscle size [15], likely due to the similar levels of fatigue and muscle activation [9]. Simply representing the volume of these protocols in kilograms would suggest that the protocols differed drastically with regard to the muscle hypertrophic stimulus received. Given the average bench press 1RM in this study was approximately 61 kg for the 30 % 1RM condition, and approximately 51 kg for the 75 % 1RM condition, the 30 % 1RM condition completed approximately 141 repetitions to volitional fatigue using an 18.3 kg load for a total exercise volume of 2580.3 kg per training session, whereas the 75 % 1RM condition completed approximately 30 repetitions to volitional fatigue using a 38.25 kg load for a total exercise volume of 1147.5 kg per training session. Extrapolating this to the total exercise volume over the 18 training sessions performed, the 30 % 1RM condition completed 46,445 kg of exercise, compared with the 75 % 1RM condition, which completed 20,655 kg of exercise. This would suggest that the low load condition completed a
224% greater volume of exercise when, in reality, all individuals were given a similar relative stimulus. That is, all individuals performed exercise until they were fatigued to a point where they could not complete another repetition. Therefore, in such an instance we would encourage the author(s) to report the volume as ‘3 sets to volitional fatigue’ while also reporting the relative load used. This can then be easily replicated to ensure a common stimulus as opposed to simply reporting the total load lifted.

### 3 Potential Problems with Reporting Relative Volume

In addition to reporting resistance training volume in absolute kilograms lifted, training volume has also been expressed in relative terms by using the equation \((\text{repetitions} \times \%1\text{RM})\) [19]. Relative volume has also been used as a way to equate resistance training stimuli between groups [19] but this method has its own limitations, and this is again apparent when examining the results of the aforementioned study by Ogasawara et al. [15]. The 141 repetitions completed at 30% 1RM would result in a relative volume of 42.3 arbitrary units (au), which is nearly double that of the 22.5 au relative volume that would result from completing 30 repetitions at 75% 1RM. Therefore, even when expressed in relative terms, the total exercise volume after 18 sessions would still be significantly greater following training to volitional fatigue at 30% 1RM (761.4 au) when compared with 75% 1RM (405 au). Of note, this is just one example that was chosen to illustrate this point. This limitation with reporting relative volume is commonly observed when comparing low loads and high loads performed to volitional fatigue, and, as such, other examples can be found elsewhere [14, 20, 21]. These studies illustrate that even relative volume is not an appropriate way to express exercise volume as it is highly impacted by the relative load lifted (i.e. %1RM). By representing volume in sets to fatigue, we can eliminate a large amount of individual variability by ensuring that all individuals received a similar relative stimulus. That is, by reporting the number of times the muscle has been taken to volitional fatigue, we ensure all individuals complete a similar protocol while receiving an equally stressful stimulus provided by each set of exercise.

### 4 Relative Load is Not Always a Fair Stimulus

It has been observed that females are more resistant to fatigue than males [22], and endurance athletes are more fatigue-resistant than weight-trained individuals [8]. These observations can likely be attributed to physiologic differences, as well as the absolute loads being lifted at each relative intensity. For example, as the relative load is decreased, the sex [7] and sport-specific [8] differences in the time to reach fatigue becomes greater, with females and endurance athletes fatiguing at a much slower rate. For this reason, prescribing a certain number of repetitions at a specific percentage of an individual’s 1RM may produce differing results from one individual to the next depending on how fatiguing the exercise is. For example, prescribing 3 sets of 10 repetitions using a 70% 1RM load may cause some individuals (particularly those with higher absolute 1RMs or lower muscular endurance) to reach volitional fatigue, whereas other individuals (particularly those with lower 1RMs or greater muscular endurance) may complete all of the prescribed repetitions without much difficulty, given that they can perform upwards of 30 repetitions at that relative load [8]. Figure 1 illustrates data obtained from two studies conducted in our laboratory assessing the maximum number of repetitions completed by untrained individuals during elbow flexion exercise with either a 35% [23] or 70% [24] 1RM load. Individual variability clearly illustrates that individuals are not all receiving the

![Fig. 1](image)

Fig. 1 Individual variability in repetitions to volitional fatigue for the elbow flexion exercise at two different relative loads. Circles represent each individual, with solid lines representing the group median. Notably, repetitions completed at 35% 1RM ranged from 18 to 73, and at 70% ranged from 9 to 18. The repetitions at 35% [23] and 70% [24] were obtained from two previous studies in our laboratory. 1RM one-repetition maximum
same relative stimulus when an arbitrary number of repetitions are prescribed (e.g., 3 sets of 10 repetitions). Therefore, a certain training protocol may produce conflicting results with regard to muscle growth, but this may be largely related to the number of repetitions that can be completed at a given relative load. This brings about an important question as to why investigators take the time to standardize the load relative to an individual’s 1RM, but completely disregard differences in muscular endurance by simply having all individuals perform the same number of repetitions. While it is common for investigators to match groups in an attempt to equate for 1RM strength, it may be equally as important to balance groups based on the strength endurance of the individuals (i.e., repetitions that can be performed at a set load), particularly when an arbitrary number of repetitions are prescribed. Further adding to the need for a more individualized protocol is the idea that all individuals will recover from exercise at a different rate, making the repetitions of the subsequent sets likely to be highly variable as well. Of course, this could easily be eliminated if exercise is always performed to volitional fatigue whereby every individual receives a similar stimulus that brings the muscle to exhaustion ‘x’ amount of times. If sets are not taken to volitional fatigue, no study can truly be replicated because the stimuli provided will differ based on the physiology of the inclusive study population.

5 Increasing the Homogeneity

The benefits of performing all sets to volitional fatigue allows for the results to be more easily interpreted and compared across studies. It has been shown that relative loads between 20 and 90 % 1RM will likely provide a similar stimulus for muscle growth [12, 14, 25], with both producing similar increases in types I and II fiber size [14]. While exercise to volitional fatigue with loads under 20 % or over 90 % 1RM remain less explored, there is likely a relative loading range in which resistance exercise produces similar results with regard to muscle growth; however, the range in which this window extends beneath 20 % or beyond 90 % 1RM remains unknown. For example, too high of an exercise load may cause the muscle to fatigue too quickly, causing an insufficient stimulus for muscle growth [26] unless a larger number of sets are performed. Likewise, resistance exercise using very low intensities that do not fatigue the muscle within approximately 3 min may provide differential results in the time course of different protein synthetic subfractions [27]. However, since most resistance training studies use loads between 30 and 90 % 1RM, administering resistance exercise to volitional fatigue would allow for easier cross study comparisons by increasing the homogeneity of the exercise stimulus. Even if loads not falling between 30 and 90 % 1RM are used, this still allows for a more common stimulus to be applied within the same study population provided the load is high enough to induce volitional fatigue. Meta-analyses attempting to assess the importance of various training variables [6, 18] have been severely limited by the lack of homogeneity within studies. By taking all sets to volitional fatigue, we can truly assess the importance of different training variables as all individuals will receive a similar stimulus upon the completion of each set of exercise.

6 Physiological Rationale for Providing a More Homogenous Stimulus

The size principle of Henneman et al. [28] states that motor units are preferentially recruited so that lower threshold motor units containing smaller, weaker muscle fibers are recruited before higher threshold motor units containing larger, stronger fibers. Higher threshold motor units can be activated independent of the training load provided the exercise induces a sufficient level of fatigue to mandate their activation, and this will subsequently provide a level of tension capable of inducing muscle hypertrophy across a greater number of muscle fibers [13]. For this reason, training to volitional fatigue is not mandatory for muscle growth, but the magnitude of muscle hypertrophy will likely be dependent on the number of muscle fibers activated. While measurable muscle hypertrophy will occur provided the exercise is taken close enough to volitional fatigue to activate enough muscle fibers, training to volitional fatigue ensures that individuals activate all muscle fibers voluntarily possible with resistance exercise. That is, by prescribing a set number of repetitions, some individuals may not need to activate all muscle fibers because the exercise can be completed without their contribution, and these inactive muscle fibers are not exposed to the tension thought to initiate mechanotransduction-induced muscle hypertrophy. The similar muscle activation present during low (30 % 1RM) and high load (70 % 1RM) exercise to volitional fatigue (97 % vs. 82 %) illustrates that a similar stimulus is being applied [29]. While there is a slightly lower muscle activation present during low-load exercise (15 %), this can likely be attributed to motor unit cycling whereby the lower loads allow for some motor units to be cycled on and off to aid in recovery without the need for all motor units to be activated at once [30].

Interestingly, a recent paper [31] has proposed a similar idea with respect to creating a more homogenous response to endurance exercise. It was suggested that part of the interindividual variability in responses to aerobic exercise may be due to the prescription of exercise relative to an
individual’s whole-body aerobic capacity (VO₂ max), which does not account for differences in the peripheral musculature being exercised. The authors propose that individuals with greater skeletal muscle mass may need to exercise at a larger percentage of their whole-body VO₂ max to see similar local muscle adaptation. Within the resistance training literature, a similar idea has been proposed in that those with lower absolute strength levels may be able to complete more repetitions at a given relative load [32]. Thus, we may need to account for differences in absolute strength to make the training stimulus equal across individuals. Fortunately, our proposed method of prescribing all exercise to volitional fatigue provides an easy way to administer a more uniform stimulus across individuals at a given load as this would take into account these differences in strength endurance.

7 Potential Limitations with Using Sets to Volitional Fatigue

Previous studies have attempted to discourage the use of resistance exercise performed to volitional fatigue for several reasons, none of which appear to be supported by compelling evidence. A decrease in basal insulin-like growth factor-1 (IGF-1) has been observed following training to volitional fatigue, and this has been proposed to be a negative adaptation in response to the higher stress of training to volitional fatigue [33]. Additionally, it has been proposed that training to failure can result in a period of overtraining [34]. However, this hypothesis was made on the basis that strength plateaued after 5 weeks of training to volitional fatigue when compared with a periodized program that continued to increase strength. These results are more likely related to the principle of specificity, whereby the group training to volitional fatigue maintained a constant training load; however, the periodized group began training at a much higher intensity during the 5–7 week period when a strength divergence was present. Additionally, it is unlikely that performing 3 sets of an exercise to volitional fatigue would have drastic effects on overtraining, given that most protocols eliciting overtraining effects commonly involve extreme volumes of exercise [35]. Despite these previous observations, training to volitional fatigue has been repeatedly shown to increase muscle size [14, 15, 36, 37], refuting the argument that any adverse hormonal or neural adaptations are limiting muscle growth during this type of training.

Another argument against the use of training to volitional fatigue is that individuals may prematurely stop short of fatigue due to the lack of motivation to continue exercise, whereas prescribing a set number of repetitions gives individuals a goal number of repetitions to complete. Despite this potential limitation, most individuals will reach volitional fatigue within approximately 30 repetitions (approximately 90 s), assuming a load of ≥30 % 1RM is used [36]. Furthermore, individuals can be monitored and motivated to complete as many repetitions as possible, and this flexibility has benefits over prescribing protocols that do not allow deviation from a set number of repetitions. Using standard protocols that prescribe a set number of repetitions may be disadvantageous given that some individuals may complete all of the repetitions with ease, and the standardized protocol cannot be altered to ensure the individual is given a sufficient stimulus capable of inducing muscle growth. While the number of repetitions completed in the latter sets of multiset exercise protocols may differ based on differences in recovery, we do not feel as though this produces a differential stimulus as these subsequent sets will all be taken to volitional fatigue. Therefore, all individuals will likely be recruiting and loading a large portion of muscle fibers toward the end of each set, which would likely provide a similar stimulus for muscle hypertrophy [9, 13]. Data from our laboratory examining blood flow restricted exercise would support this hypothesis as differences in repetitions toward the latter sets did not result in differential muscle growth [38]. Furthermore, different rest intervals have been shown to modify the protein synthetic response even when all sets are performed to volitional fatigue [39], which may make it more difficult to compare studies employing different rest intervals. Despite this potential limitation, other studies have illustrated no differences in muscle hypertrophy with differing rest intervals [40, 41], and performing exercise to volitional fatigue would still serve to eliminate the unintentional variation in fatigue provided by each set of exercise. Finally, while this current opinion is focused on muscle hypertrophy, training to volitional fatigue is likely not necessary for standardization with respect to strength outcomes as strength appears to be primarily driven by the intensity and specificity of exercise (i.e. %1RM for a certain lift) as opposed to the volume completed [14].

8 Conclusions

To administer a more standardized stimulus across all individuals, and to properly answer questions regarding the importance of modifying different training variables, it would be of benefit to conform to a more uniform resistance training protocol. Even when resistance exercise protocols are prescribed relative to an individual’s 1RM, the stimulus will differ depending on the individual’s muscular endurance or absolute load they are capable of lifting. Therefore, rather than prescribing an arbitrary number of repetitions to be completed, allowing all
individuals to exercise to volitional fatigue will ensure a common stimulus that induces a similar level of fatigue across all individuals. Both the absolute (repetitions × load) and relative (repetitions × %1RM) loads are of little meaning as these will produce differential results among individuals and will differ based on the load used. Thus, we recommend that researchers report the total volume of exercise as the number of times the exercise was performed to volitional fatigue as this stimulus is more easily replicable. With respect to resistance exercise, there is a need for a more standardized protocol aimed at reducing individual variability by prescribing exercise to volitional fatigue as this will result in a common stimulus applied to all individuals.

Compliance with Ethical Standards

Funding The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this manuscript. This article was not supported by any funding.

Conflict of interest Scott Dankel, Matthew Jesse, Kevin Mattocks, J. Grant Mouser, Brittany Counts, Samuel Buckner and Jeremy Loenneke declare that they have no conflicts of interest relevant to the content of this article.

References


