Estimation of Repetitions to Failure for Monitoring Resistance Exercise Intensity: Building a Case for Application

Daniel A. Hackett, Stephen P. Cobley, and Mark Halaki

Discipline of Exercise and Sport Science, The University of Sydney, Sydney, New South Wales, Australia

Abstract

Hackett, DA, Cobley, SP, and Halaki, M. Estimation of repetitions to failure for monitoring resistance exercise intensity: Building a case for application. J Strength Cond Res 32 (5): 1352–1359, 2018—The purpose of this study was to (a) examine the accuracy of Estimated Repetitions to Failure (ERF) during resistance exercise between 2 sessions and (b) compare ERF to rating of perceived exertion (RPE) for determining proximity to momentary failure. Forty-eight adults with recreational resistance training experience performed 3 sets of 10 repetitions at 70% one-repetition maximum (1RM) and 80% 1RM for the chest press and leg press, respectively. At the completion of each set, participants reported their ERF and then continued repetitions to failure to determine actual repetitions to failure (ARF). Two sessions of the same experimental protocol were performed with 48 hours between bouts. For session 1, error in ERF was greater during the first sets compared with third sets for the chest press (2.0 vs. 0.6 repetitions and \(p < 0.001\)) and leg press (3.1 vs. 1.6 repetitions and \(p < 0.001\)). No differences for error in ERF were observed between sessions 1 and 2 for the chest press (\(p > 0.944\)); however, less error in ERF was found for the leg press during set 1 of session 2 (3.1 vs. 1.9 repetitions and \(p < 0.013\)). Strong to very strong relationships were found between ERF and ARF (\(r = 0.59–0.87\) and \(p < 0.01\)), whereas most relationships for RPE and ARF were small to moderate (\(r = 0.32\) to \(-0.42\) and \(p < 0.01\)). Improvement in the accuracy of ERF after a single training bout is minimal, whereas ERF compared with RPE seems to have greater sensitivity for discriminating momentary failure.

Key Words resistance training, RPE, training intensity, repetition maximum

Introduction

Resistance training intensity is generally prescribed from the percentage of a one-repetition maximum (%1RM) or the most load that can be lifted for a defined number of repetitions, known as a repetition maximum (RM). Although maximum effort is required during every set when resistance exercise is prescribed from RM, the effort required when performing sets at %1RM can vary greatly and ultimately depends on the number of repetitions that are performed. Furthermore, studies have shown interindividual variation in the number of repetitions performed to momentary failure at fixed %1RM (1,12,21). Therefore, it is unlikely that the same degree of effort or training stimulus would result for 2 individuals performing sets of a specific number of repetitions (e.g., 10 repetitions) at a fixed %1RM. Furthermore, the number of repetition maximum at a fixed %1RM can be different for the same individual when performing different exercises. This problem could be solved using RM so that effort among trainers is standardized. However, the high physiological and psychological demands of performing resistance training solely with RM may result in overtraining as well as injuries (24,25).

When resistance training is prescribed based on %1RM, the rating of perceived exertion (RPE) scale is often used to assist with standardizing training conditions between individuals (19). There are 2 types of RPE scales, which include the 6–20 category scale and 0–10 category ratio scale (CR-10), whereas the latter is considered better suited for resistance exercise (7,17,23). Resistance exercise intensity can be estimated from the RPE because the scale assesses subjective effort, strain, discomfort, and fatigue. However, several investigators have reported RPEs less than maximum during resistance exercise to volitional fatigue, indicating a mismatch between RPE and maximal effort (18,22).

An alternate approach to using the RPE scale to assist with assessing resistance exercise intensity is to have individuals report the number of repetitions possible after completion of a set. This can be performed with the use of the Estimated Repetitions to Failure (ERF) scale, which was previously validated in a cohort of experienced bodybuilders (9). The results of this previous study indicated that the margin of
error in ERF was approximately 1 repetition across 5 sets performed for both the bench press and squat (9). More recently, Hackett et al. (8) examined the accuracy of ERF in a large group of healthy adults with various levels of resistance training experience. It was shown that the error in ERF was ~1 repetition when 0–5 repetitions from momentary failure, there was less error for upper compared with lower-body exercises, and that men were more accurate than women for lower-body exercises. Furthermore, it was found that the accuracy of ERF was not influenced by resistance training experience. However, it remains largely unknown whether a repeated bout affects the accuracy of ERF and whether the ERF is just a surrogate to RPE.

The ability to accurately monitor resistance exercise intensity is essential for coaches and athletes. If ERF proves to be an accurate method to determine proximity to momentary failure, this may be a more effective tool compared with RPE to monitor resistance exercise intensity. Potential advantages of ERF include the ability to better equate exercise intensity between athletes or trainers compared with %1RM, as well as the ability to indirectly monitor the rate of recovery or adaptation between training sessions. As such, ERF could be used by coaches to modify training sessions and programs to optimize adaptations.

The purpose of this study was to (a) examine the accuracy of ERF during resistance training between 2 sessions and (b) compare ERF to RPE for determining proximity to momentary failure. It was hypothesized that the accuracy in ERF across sets would be similar between testing sessions. It was also speculated that ERF and not RPE would be strongly associated with actual repetitions to failure (ARF).

**METHODS**

**Experimental Approach to the Problem**

Each participant visited the laboratory on 2 occasions. The first visit involved 1RM testing and experimental session 1, whereas the experimental session 2 was conducted during the next visit. During the experimental sessions participants performed 3 sets of 10 repetitions for resistance exercises at a fixed percentage of 1RM (%1RM). Exercises were performed with a pin-loaded vertical chest press machine (Maxim, Kidman Park, South Australia) and a pin-loaded horizontal leg press machine (Kolossal, Sydney, New South Wales). Participants briefly paused when the prescribed number of repetitions for each set was reached while they reported their ERF and RPE and then continued to momentary failure. Two sessions of the same experimental protocol were performed with 48 hours between sessions to minimize confounding influences of previous exercise. Participants were instructed to maintain their normal diet during the days preceding visits, to consume their last meal at least 2 hours before exercise and to avoid using preworkout supplements (3). Participants were further instructed to refrain from resistance training or any other strenuous type of exercise 48 hours before visits. Absolute difference between ERF and ARF for each set and across sessions was used to determine changes in accuracy for ERF, whereas associations were used to examine the discriminative ability of ERF and RPE to determine momentary failure.

**Subjects**

Twenty-eight men (± SD age = 20-56 years, body mass = 80.4 ± 10.6 kg, and height = 175.6 ± 7.9 cm) and 20 women (age = 19-55 years, body mass = 61.4 ± 9.1 kg, and height = 163.6 ± 7.0 cm) participated in this study. Most participants (42 of 48) reported having ≥1 year resistance training experience at the recreational level. Participants were informed of the study purposes, procedures provided, and all potential risks before consent. Informed consent documents were signed by all participants before commencing the study, which was approved by the University of Sydney Human Research Ethics Committee.

**Procedures**

**One-Repetition Maximum.** Participants warmed up with 1–2 sets of 8–10 repetitions with light-moderate loads before 1RM testing for the chest press and leg press. After warm-up, loads were adjusted to enable participants to perform ≤10RM, so that 1RM for the exercises could be accurately estimated (20). If ≥10 repetitions could be performed or failure was not reached before 10 repetitions, the load was increased and 5-minute recovery was provided before the next RM attempt. The Brzycki 1RM prediction equation (5) was used to estimate the 1RM based on the load and repetitions performed. The equation is mathematically expressed as:

\[
1RM = \frac{load}{(1.0278 - 0.0278 \times \text{number of repetitions})}
\]

Standard error of estimate (SEE) of 1RM from the Brzycki equation for the chest press was previously found to be 1.67

---

**Table 1. Category ratio rating of perceived exertion scale.**

<table>
<thead>
<tr>
<th>Rating</th>
<th>Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Rest</td>
</tr>
<tr>
<td>1</td>
<td>Very, very easy</td>
</tr>
<tr>
<td>2</td>
<td>Easy</td>
</tr>
<tr>
<td>3</td>
<td>Moderate</td>
</tr>
<tr>
<td>4</td>
<td>Somewhat hard</td>
</tr>
<tr>
<td>5</td>
<td>Hard</td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Very hard</td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Maximal</td>
</tr>
</tbody>
</table>

*The verbal anchors have been changed slightly (e.g., light becomes easy, strong, or severe becomes hard). The participants were shown this scale at the conclusion of the exercise set and asked “how would you rate your effort for the set?”*
and 3.00 kg at 5RM and 10RM, respectively (20). For the leg press, the SEE was found to be 13.74 and 20.41 kg at 5RM and 10RM, respectively (20). The test-retest intraclass correlation coefficient for the above testing 1RM estimation protocol in our laboratory is $0.90$.

Familiarization of Estimated Repetitions to Failure and Rating of Perceived Exertion Scales. After 1RM testing, participants received information on how to use the ERF and RPE scales during the resistance exercises. Participants were instructed to use a memory-anchoring procedure to enable the linking of exercise intensities with their full ERF and RPE response range. This involved asking each participant to think of times during training when they reached levels of exertion that were equal to verbal cues at the bottom and top of the scales. Participants were also told that they would be asked to report their RPE and ERF at the completion of each set 10 of repetitions. Both scales were written on a board and placed directly in front of participants during the exercises. From the RPE scale, participants were asked “how would you rate your effort for the set?” A rating of “0” was associated with “no effort” (rest), and a rating of “10” was considered to be maximal exertion (Table 1). From the ERF scale, participants were asked “how many additional repetitions can you perform?” For example, a “0” indicated that the participant estimated that no additional repetitions could be completed (momentary failure reached) (Table 2).

**Experimental Sessions.** For sessions 1 and 2, participants performed 3 sets of 10 repetitions at 70 and 80% 1RM for the chest press and leg press, respectively, with 2–3 minutes recovery between sets. The rationale for different %1RM used for the chest and leg exercises was to have participants perform a similar number of repetitions to failure. Based on results from pilot testing, $\approx 20$ repetitions to failure were performed with 70% 1RM and 80% 1RM for the chest press and leg press, respectively. During lifts, participants were encouraged to complete each repetition through a full range of motion without deviating from the proper technique while keeping the lifting speed constant. On completion of 10 repetitions, participants paused briefly (i.e., for 5 seconds) at the end of the concentric phase and were then required to report their RPE and ERF (in that order). Participants then continued with the set performing repetitions until momentary failure, which was defined as the participant achieving volitional failure or the incapacity to perform the exercise with proper execution. The actual number performed to momentary failure was referred to as the ARF. Verbal encouragement (i.e., shouting positive words) was provided throughout sessions to ensure that “true” momentary failure was achieved.

**Statistical Analyses**

The error in estimation of repetitions to failure was calculated as the absolute
difference between ERF and ARF for each set. To assess the error, a 4-factor analysis of covariance (ANCOVA) was used with sets, exercise, and session serving as within-subject factors, sex as a between-subject factor, and experience as a covariate. Tukey post hoc tests were used when significant ANCOVA results were found. Actual repetitions to failure for each exercise between sessions (for corresponding sets) were analyzed using independent t-tests. Associations between ERF and ARF across sets for each participant and exercise were examined using Pearson’s correlations and linear least-products regression (15). These parametric tests were used to compare ERF and ARF because the data were interval, normally distributed (checked using the Kolmogorov-Smirnov test), and had similar variances. A Bland-Altman analysis between ARF and ERF for the chest press and leg press for sessions 1 and 2 was used to assess bias and the limits of agreement. Associations between RPE and ARF were evaluated using a Spearman’s rank correlation because of RPE being a non-parametric variable. Strength of the associations was qualitatively assessed using the following criteria: trivial ($r < 0.1$), small ($0.1 < r < 0.3$), moderate ($0.3 < r < 0.4$), strong ($0.4 < r < 0.7$), very strong ($0.7 < r < 0.9$), nearly perfect ($r > 0.90$–0.99), and perfect ($r = 1.0$) (13). All analyses were performed using Statistica version 10.0 (StatSoft Inc., Tulsa, AZ, USA). Data are presented as mean ± SD, and level of significance was set at $p < 0.05$.

**RESULTS**

After completing the 10 repetitions and ERF, participants

---

**Figure 1.** Bland-Altman plot for ERF and ARF for both sessions 1 and 2. ERF = estimated repetitions to failure; ARF = actual repetitions to failure. A. chest press session 1; B. chest press session 2; C. leg press session 1; D. leg press session 2.

**Figure 2.** Accuracy in estimation of repetitions to failure for men vs. women. *Significant differences between men and women ($p < 0.05$). †Significant difference to session 1 ($p < 0.05$). ‡Significant difference between exercises for the corresponding session ($p < 0.05$).
performed repetitions ranging from 0–11 repetitions. However, 0 repetitions were performed only in 3 instances (sets) for different participants, and during one set, a participant performed 11 repetitions. Therefore, most additional repetitions performed after participants ERF ranged from 1–10 repetitions. Table 3 shows the actual repetitions performed for all sets of exercises (chest press and leg press) during sessions 1 and 2. For set 3 of the chest press, the actual repetitions to failure were greater in session 2 (1.1 repetitions, \( p < 0.05 \)). There were no other differences for actual repetitions to failure between sessions for the corresponding sets of exercises. There was a systematic increase in the error between ARF and ERF with increasing number of repetitions as shown by a significant regression and significant positive slope for both the chest press (Figure 1A, B; \( p = 0.001 \)) and leg press (Figure 1C, D; \( p = 0.001 \)) during sessions 1 and 2.

Analysis of covariance results indicated that there was a significant main effect for sex (\( p < 0.001 \)) with men exhibiting less error in ERF than women, for set (\( p < 0.001 \)) with post hoc results indicating that the error in ERF during set 1 was greater than set 2 (\( p < 0.001 \)) and both sets 1 and 2 exhibited greater error than set 3 (\( p < 0.001, p = 0.020 \), respectively). No other main effects were significant (\( p > 0.076 \)). There was a significant interaction between session, exercise, and set (\( p = 0.032 \)) with post hoc results indicating as follows. All other interaction effects were not significant (\( p > 0.05 \)).

**Initial Session**

The error in ERF tended to decrease as sets progressed. For the chest press, there was greater error for set 1 (2.0 repetitions) compared with set 3 (0.6 repetitions) (\( p < 0.001 \)) with no difference between sets 1 and 2 (1.2 repetitions) (\( p = 0.143 \)) nor between sets 2 and 3 (\( p = 0.467 \)). For the leg press during the first session, the error in ERF was greater for set 1 (3.1 repetitions) compared with sets 2 (1.8 repetitions) (\( p < 0.001 \)) and 3 (1.6 repetitions) (\( p < 0.001 \))

![Figure 3. Accuracy in estimation of repetitions to failure between initial and second sessions. *Significant differences between sessions (\( p < 0.05 \)). † Significant difference to set 1 (\( p < 0.05 \)). ‡ Significant difference between exercises for the corresponding set (\( p < 0.05 \)).](image)

![Figure 4. Associations between actual repetitions to failure and estimated repetitions to failure in comparison with rating of RPE. ERF = estimated repetitions to failure; ARF = actual repetitions to failure; RPE = rating of perceived exertion. *Significant association (\( p < 0.05 \)).](image)
with no difference between sets 2 and 3 (\(p = 0.999\)). Between exercises, there was less error in ERF for chest press compared with the leg press during sets 1 (\(p = 0.005\)) and sets 3 (\(p = 0.013\)) with no difference during sets 2 (\(p = 0.575\)) (Figure 2).

**Second Session**

For the chest press, the error in ERF was greater for set 1 (2.4 repetitions) compared with sets 2 (1.2 repetitions) (\(p < 0.001\)) and 3 (0.9 repetitions) (\(p < 0.001\)) with no difference between sets 2 and 3 (\(p = 0.999\)) (Figure 2). However, for the leg press, there was no difference in the error in ERF across sets (\(p > 0.053\)) (Figure 2). Between exercises, there was no difference in the ERF for chest press compared with the leg press during all sets (\(p > 0.733\)) (Figure 3).

**Between the First and Second Sessions**

There were no differences between sessions in the error in ERF for chest press during any of the sets (\(p > 0.944\)). However, the error in ERF for leg press was higher during set 1 (3.1 vs. 1.9 repetitions, \(p = 0.013\)) with no difference during sets 2 and 3 (\(p = 0.922\)).

**Ratings of Perceived Exertion and Muscular Failure**

Strong to very strong relationships were found between ERF and ARF for sessions 1 and 2 (\(r = 0.59–0.87, p < 0.01\)) (Figure 4). By contrast, there were only 2 sets where strong relationships were found between RPE and ARF (\(r = 0.54\) and \(-0.55, p < 0.01\)), whereas the rest of the relationships were either small to moderate (\(r = -0.32\) to \(-0.42, p < 0.01\)) or trivial (\(r = -0.15\) and \(-0.18, p > 0.05\)).

**DISCUSSION**

This investigation found that improvements in the accuracy of ERF during resistance exercises are minimal after a single session. This was observed with improvement in accuracy of ERF only for the initial set of the leg press during the second session, which partially supports our original hypothesis. Also in agreement with our hypothesis, strong correlations were found between ERF and ARF across all sets of exercises, whereas weaker correlations were found between RPE and ARF. This suggests that ERF may be a more appropriate method to monitor resistance exercise performance in relation to proximity to momentary failure. The study also showed that the accuracy in ERF is greater for the chest press compared with leg press but only during the initial session.

Participants were shown to underestimate ERF for first sets by \(-2–3\) repetitions, with subsequent improvements in accuracy as sets progressed and an error in ERF of \(-1\) repetition in the final sets. Similar to the findings from previous studies, the proximity to momentary failure when reporting ERF was shown to influence the accuracy of ERF (8,9). Therefore, it seems that exertional sensations (e.g., muscle activation, afferent signals from Golgi tendon organs, muscle spindles, and mechanoreceptors) play an important role toward improvement in ERF accuracy (6,14,16). Furthermore, the general lack of improvement in accuracy of ERF during the second session may indicate that participants relied heavily on their first set of an exercise as a reference point to make their estimation. So it seems that current rather than previous exertional sensations are most influential for improving the accuracy in ERF. However, potentially 2 sessions is not enough time for participants to improve their accuracy when the error in ERF is not great (i.e., error of \(-1–2\) repetitions). This supported by the accuracy of ERF improving during the second session for the first set of leg press, in which the corresponding set during the initial session was the only instance where the error in ERF was \(-3\) repetitions.

Consistent with the findings from a previous study (8), the accuracy of ERF was greater when performing the chest press compared with leg press during the initial session. However, during the second session, no difference in accuracy of ERF was observed between exercises. An explanation for this finding is likely related to the improvement in accuracy of ERF for the first set of leg press during the second session. The greater error in ERF for the leg press compared with chest press during the initial session could be related to participants having lower self-efficacy for the leg press, perhaps associated with the heavier absolute loads used. Therefore, at least one session is required with utilizing the ERF for the leg press to ensure that the accuracy of ERF is similar to that of the chest press.

Men compared with women were more accurate with reporting of ERF, and this topic has been previously discussed (8). Briefly, a potential mechanism that may explain this finding could be the anatomical-physiological differences in male and female muscles, which may influence the sensory information to be input into the central nervous system to allow for the effort to be perceived (10). Future research is needed to confirm whether the accuracy of ERF during resistance training differs between sexes and, if confirmed, further exploration of possible mechanisms. Although it may seem that ERF may be a less-effective tool for women to monitor resistance exercise intensity, it is possible that women may improve their ability to ERF with further practice and this should also be explored in future studies.

Previous studies have demonstrated that active muscle RPE ratings increase during resistance exercise, when lifting heavier loads and when approaching muscular failure (6,14,16). In this study, relationships between RPE and repetitions to momentary failure were mainly small to moderate compared with the stronger relationships between ERF and momentary failure. Furthermore, it could be argued that, based on the findings from the present study, the momentary failure cannot be determined based solely on RPE. This is in agreement with previous studies where an RPE of less than 10 has been reported despite momentary failure being reached (18,22). Based on the evidence from this study, the case is being built toward ERF compared with RPE...
being a more appropriate method for monitoring of resistance exercise intensity.

Potentially the ability of RPE to discriminate momentary failure may be improved with the “repetitions in reserve” (RIR) scale that combines both RPE (CR-10) with estimating repetitions to failure (11). The RIR scale is gaining popularity among resistance trainers and coaches to quantify and practically use RPE for training purposes. Zourdos et al. (26) found that experienced compared to novice resistance trainers were more accurate with reporting RIR during squat. However, it should be noted that Zourdos et al. (26) did not directly assess accuracy of RIR, rather it was inferred based on the %1RM–RM continuum (2). Also, it is highly likely that the number of repetitions performed to failure (RM) at %1RM would differ between experienced and novice squatters (21), thus limiting conclusions made concerning the accuracy of RIR. The responses using the CR-10 also increase in a nonlinear and positively accelerating manner during exercise (4). Therefore, it seems unlikely when using the CR-10 that a one-point movement along the scale would equate with approximately 1 repetition. Because the CR-10 is being modified so that it relates to repetitions in reserve (e.g., RPE 9 = 1 repetition remaining), a better decision would be to not combine these 2 methods. Furthermore, the inverse relationship between RPE and estimation of repetitions to failure is not intuitive and aligned with the idea of capturing remaining capability under fatigue or when performing sets to momentary failure. Therefore, it seems more logically appropriate to use ERF instead of the RIR scale which resolves these issues.

A limitation of this study is that the conditions for the experimental sessions were different with 1RM testing occurring only before session 1. Therefore, it should be acknowledged that the ability to ERF during the first session may have been influenced by previous 1RM testing. The exertional sensations experienced from the IRM testing may have assisted participants with ERF during the first experimental session, at least for the initial sets of each exercise. Closer proximity to momentary failure during exercise sets may have resulted from previous IRM testing as a result of fatigue and assisted participants with accuracy of ERF during the first experimental session. However, a significant difference in the actual repetitions to failure was found for only set 3 of the chest press between sessions 1 and 2 (1.1 repetitions greater in second session). Therefore, it seems that differences in proximity to failure as a result of previous 1RM testing may not have influenced ERF. It is also important to note that the findings from this study can only be extrapolated to the loading intensities used and the specific exercises that were performed. Future research is needed to examine whether the accuracy of ERF found in this study can be achieved with more complex resistance exercises (e.g., squats and deadlifts). Previously, the accuracy of ERF for the bench press and squat has been shown to be good (error in ERF of approximately 1 repetition) in male bodybuilders (9). However, to date, no study has examined the accuracy of ERF with single-joint exercises (e.g., bicep curls and leg extensions).

In conclusion, the results suggest that there is little improvement in the accuracy of ERF after a single training bout. However, there seems to be a greater chance of improving accuracy of ERF after a single session if error in ERF for an exercise is >3 repetitions during the initial session. Rating of perceived exertion did not correlate as strongly as ERF with actual repetitions to failure, providing evidence that RPE is less sensitive for discriminating momentary failure. Therefore, ERF compared with RPE seems to be better suited for monitoring resistance exercise intensity.

**Practical Applications**

The ERF scale provides coaches, trainers, and athletes with a method to monitor proximity to momentary failure during resistance exercise with reasonable accuracy. By contrast, the RPE scale seems unable to discriminate momentary failure as well as being a subjective measure for which its accuracy cannot be quantified. The accuracy of ERF reported by a resistance trainer can be readily assessed periodically by coaches. Although, it needs to be emphasized that the accuracy of ERF is affected by the repetition range from momentary failure, with accuracy increasing as a lifter approaches failure. Also, caution is required when using the ERF scale with women because of their lower accuracy compared with men. However, with repeated application and user experience, the reliability and accuracy of ERF is likely to improve over time. As identified, repetitions performed to momentary failure at specific %1RM can vary considerably between individuals (1,12,21). This can lead to large differences in exertion and fatigue responses between individuals when prescription is based on a selected number of repetitions to be performed at a %1RM. Therefore, coaches could implement the use of the ERF scale within resistance training programs to better equate performances between athletes. Coaches could also use the ERF scale to help identify whether loads need to be adjusted and to help their athletes train at intensities that are more closely matched. For example, loads could be selected leading toward an ERF of 2–3 after sets of 10 repetitions. Another benefit of the ERF scale is that individual responses reported can assist with monitoring the rate of recovery or adaptation between training sessions. If ERF values are greater or less between training sessions where the similar exercises and loads were used, this could assist coaches with modifying the training sessions to optimize adaptations.

**Acknowledgments**

The authors have no conflicts of interest to disclose.

**References**